Bioanalysis
the JOURNAL, the ZONE

THE COLLECTABLES:
YEAR IN BIOANALYSIS – looking back to 2016

Contents

Large Molecules

Metabolites and Metabolomics

Quantification Studies

Foreword

We are pleased to present the 2016 edition of The Collectables: Year in Bioanalysis – a special supplement, highlighting content from the journal *Bioanalysis* and Bioanalysis Zone picked by the editors.

Journal Content Highlights

Bioanalysis publishes a variety of articles covering key advances in this ever-evolving field and 2016 saw a range of articles featuring in our most-read content. We saw several Research Articles presenting novel work and important advancements in the understanding of techniques, and Review articles highlighting recent significant advances in research, ongoing challenges, continue to be well received by our readers. Our free-to-view meeting reports and recommendation papers also proved to be popular articles in 2016.

With the aim of highlighting current and happening themes in the field of bioanalysis, we published several Special Focus Issues in 2016 on the topics 'Microscale Bioanalysis', 'Immunoaffinity MS', 'HRMS in DMPK' and a two-part series on 'Bioanalysis of Biomarkers'. We have a number of exciting themed issues planned for 2017, with the first issue of 2017 itself being a Special Focus Issue on 'Methods and Techniques for Metabolic Phenotyping'.

Bioanalysis Zone Highlights

Bioanalysis Zone is the online home of bioanalysis: we bring together the latest bioanaltyical news, views, research and products in one place, and provide a forum for the community to discuss the developments of this fast-moving field. In 2016 we focused on a variety of 'hot topics' with our special 3-month Spotlight features: 'Large molecules by LC–MS', 'Hybrid LBA-LC-MS' and 'Quantitative HMRS'. We saw exclusive interviews and commentaries from opinion leaders working within academia and the pharmaceutical industry in our 'Ask the Experts' feature on 'Microsampling' and 'Biomarkers'.

A highlight of 2016 was the independent Roundtable Discussions organized jointly by *Bioanalysis* and Bioanalysis Zone and held in Orlando, USA (April) and Barcelona, Spain (November) in which bioanalytical experts from Pharmaceutical Companies and Contract Research Organizations were brought together to discuss topical issues faced by the bioanalytical community and the changing world of bioanalysis, respectively.

In 2017 *Bionalaysis* and Bioanalysis Zone will continue to be at the forefront of bioanalysis with exciting new content and features, in line with key advances in the field, reported and analyzed by experts, providing an authoritative but accessible forum for the modern bioanalyst.

We hope you enjoy the supplement.

Sankeetha Nadarajah
Managing Commissioning Editor
s.nadarajah@future-science.com

Naamah Maundrell
Editor, Bioanalysis Zone
n.maundrell@future-science.com

Biomarkers

Bioanalysis

2016 White Paper on recent issues in bioanalysis: focus on biomarker assay validation (BAV): (Part 2 – Hybrid LBA/LCMS and input from regulatory agencies)

The 2016 10th Workshop on Recent Issues in Bioanalysis (10th WRIB) took place in Orlando, Florida with participation of close to 700 professionals from pharmaceutical/biopharmaceutical companies, biotechnology companies, contract research organizations, and regulatory agencies worldwide. WRIB was once again a 5-day, weeklong event – A Full Immersion Week of Bioanalysis including Biomarkers and Immunogenicity. As usual, it is specifically designed to facilitate sharing, reviewing, discussing and agreeing on approaches to address the most current issues of interest including both small and large molecules involving LCMS, hybrid LBA/LCMS, and LBA approaches, with the focus on biomarkers and immunogenicity. This 2016 White Paper encompasses recommendations emerging from the extensive discussions held during the workshop, and is aimed to provide the bioanalytical community with key information and practical solutions on topics and issues addressed, in an effort to enable advances in scientific excellence, improved quality and better regulatory compliance. This White Paper is published in 3 parts due to length. This part (Part 2) discusses the recommendations for Hybrid LBA/LCMS and regulatory inputs from major global health authorities. Parts 1 (small molecule bioanalysis using LCMS) and Part 3 (large molecule bioanalysis using LBA, biomarkers and immunogenicity) have been published in the *Bioanalysis* journal, issues 22 and 23, respectively.

First draft submitted: 30 September 2016; Accepted for publication: 3 October 2016; Published online: 18 November 2016

An Song[1], Anita Lee[2], Fabio Garofolo[*,3], Surinder Kaur[1], Jeff Duggan[4], Christopher Evans[5], Joe Palandra[6], Lorella Di Donato[7], Keyang Xu[1], Ronald Bauer[8], Mark Bustard[9], Linzhi Chen[4], Laurent Cocea[9], Stephanie Croft[10], Fabrizio Galliccia[11], Sam Haidar[12], Nicola Hughes[13], Akiko Ishii-Watabe[14], Rafiqul Islam[15], Barry Jones[16], John Kadavil[12], Carsten Krantz[17], Gustavo Mendes Lima Santos[18], Timothy Olah[19], João Pedras-Vasconcelos[12], Ludovicus Staelens[20], Yoshiro Saito[14], Natasha Savoie[21], Kara Scheibner[12], Susan Spitz[22], Nilufer Tampal[12], Eric Thomas[23], Stephen Vinter[24], Jason Wakelin-Smith[24], Jan Welink[25], Jianing Zeng[19] & Shaolian Zhou[26]

*Author for correspondence:
Tel.: +39 06 910 45 277
f.garofolo@angelini.it

Author affiliations can be found at the end of this article.

The views expressed in this article are those of the authors and do not reflect official policy of the US FDA, EU EMA, Health Canada, UK MHRA, Austria AGES, Italy AIFA, Brazil ANVISA, Japan MHLW and WHO. No official endorsement by the FDA, EMA, Health Canada, MHRA, AGES, AIFA, ANVISA, MHLW or WHO is intended or should be inferred.

Introduction

This year marked the 10th anniversary edition of the Workshop on Recent Issues in Bioanalysis (10th WRIB), which was held in Orlando, Florida from April 18–22, 2016. Nearly 700 professionals from pharmaceutical/biopharmaceutical companies, biotechnology companies, contract research organizations (CROs), and regulatory agencies worldwide were in attendance. In addition to the three focused and sequential workshop days, there were six advanced training sessions throughout the week providing attendees multiple choices to combine core workshop days and training, thereby maximizing their learning process in many fields of drug development expertise. As with prior WRIBs, this 10th edition was specifically designed to facilitate sharing, reviewing, discussing and agreeing upon scientific and regulatory approaches to address the most recent issues in both small and large molecule bioanalysis, biomarkers and immunogenicity. The areas of interest included LCMS, hybrid LBA/LCMS as well as LBA/cell-based approaches.

The chairs of the 2016 edition of the WRIB included Dr Eric Yang (GlaxoSmithKline), Dr Jan Welink (EMA), Dr An Song (Genentech), Dr Fabio Garofolo (Angelini Pharma), Dr Susan Richards (Sanofi), Dr Lakshmi Amaravadi (Sanofi) and Dr Renuka Pillutla (Bristol-Myers Squibb).

As usual, a number of regulatory agency representatives contributed actively to the 10th WRIB, including Dr Sam Haidar (US FDA), Dr Nilufer Tampal (US FDA), Dr John Kadavil (US FDA), Dr Kara Scheibner (US FDA), Dr João Pedras-Vasconcelos (US FDA), Dr Jan Welink (EU EMA), Dr Ronald Bauer (Austria AGES), Mr Jason Wake-

FUTURE SCIENCE part of fsg

Acronyms	
Abbreviation	**Definition**
ADA	Anti-drug antibody
ADC	Antibody–drug conjugate
AMED	Japan Agency for Medical Research and Development
ANDA	Abbreviated New Drug Application
API	Active Pharmaceutical Ingredient
BA	Bioavailability
BCS	Biopharmaceutics Classification System
BE	Bioequivalence
BLA	Biologics License Applications
BMV	Bioanalytical method validation
CDR	Complementarity Determining Regions
CRO	Contract research organization
CTA	Clinical Trial Application
ELISA	Enzyme-Linked Immunosorbent Assay
ELN	Electronic Laboratory Notebook
GCP	Good Clinical Practice
GCLP	Good Clinical Laboratory Practice
GLP	Good Laboratory Practices
HRMS	High Resolution Mass Spectrometry
IA	Immunoaffinity
IND	Investigational New Drug
LBA	Ligand Binding Assay
LCMS	Liquid chromatography mass spectrometry
LOD	Lower Limit of Detection
LLOQ	Lower limit of quantitation
LIMS	Laboratory Information Management System
mAb	Monoclonal antibody
MRD	Minimum Required Dilution
MRM	Multiple Reaction Monitoring
NDS	New Drug Submission
PBS	Phosphate-Buffered Saline
PD	Pharmacodynamic
PK	Pharmacokinetic
PrD	Proteolytic Digestion
PTM	Post-translational modifications
QA	Quality Assurance
QC	Quality control
SIL	Stable Isotope Labeled
SISCAPA	Stable Isotope Standard Capture with Anti-Peptide Antibodies
SOP	Standard Operating Procedure
WRIB	Workshop on Recent Issues in Bioanalysis

lin-Smith (UK MHRA), Mr Stephen Vinter (UK MHRA), Dr Fabrizio Galliccia (Italy AIFA), Mr Gustavo Mendes Lima Santos (Brazil ANVISA), Dr Mark Bustard (Health Canada), Dr Laurent Cocea (Health Canada), Dr Akiko Ishii-Watabe (Japan MHLW-NIHS), Dr Yoshiro Saito (Japan MHLW-NIHS) and Ms Stephanie Croft (WHO).

The entire workshop was designed to complete the drafting of this White Paper based on the daily working dinners and open panel discussions among the lecturers, regulators and attendees. Each core workshop day covered a wide-range of bioanalytical, biomarker and immunogenicity topics requested by members of the community, and included lectures from industry opinion leaders and regulatory representatives.

As with prior WRIB editions [1–12], a significant number of topics were addressed during the workshop and condensed into a series of relevant recommendations. In this current White Paper, the exchanges, consensus and resulting recommendations on 32 recent issues ('hot' topics) in bioanalysis, biomarkers and immunogenicity are presented. These 32 topics are distributed across the following areas:

Small molecules, peptides and small molecule biomarkers by LCMS:

- Method development challenges in bioanalysis (six topics);
- Bioanalytical regulatory challenges (five topics);
- Hybrid LBA/LCMS for biotherapeutics, biomarkers and immunogenicity:
  - Biomarker and immunogenicity assays (four topics);
  - PK assays (four topics);
- Large molecules by LBA and cell-based assays:
  - Immunogenicity (five topics);
  - Biomarkers (four topics);
  - PK assays (four topics).

In addition to the recommendations on the aforementioned topics, an additional section has been provided in the current White Paper that specifically focuses on key inputs from regulatory agencies.

Due to its length, the 2016 edition of this comprehensive White Paper has been divided into three parts for editorial reasons. This publication (Part 2) covers the recommendations for Hybrid LBA/LCMS and Regulatory Agencies' Inputs. Part 1 (Small Molecule, Peptides and Small Molecules Biomarkers using

LCMS) and Part 3 (Large Molecule Bioanalysis using LBA/Cell-based Assays, Biomarkers and Immunogenicity) are published in volume 8 of *Bioanalysis*, issues 22 and 23 (2016), respectively.

Discussion Topics

Biomarkers & immunogenicity assays

Biomarkers in solid tissues: benefits of hybrid LBA/LCMS & specific sample preparation techniques designed to minimize interferences

With hybrid LBA/LCMS or LCMS only methods, what are the best practices for sample processing, weighing of tissue, homogenization techniques, and pulverization versus sectioning of different tissue types? What are the best approaches for stability evaluations in prepared lysate, buffer selection, and storage conditions? What are the best practices for recovery and matrix effects, dilution linearity, sequential IA, and impact of lipids? What is the industry best practice in harvesting/freezing samples to preserve protein biomarker integrity, sample sourcing and normalization?

Successful validation of sensitive confirmatory protein biomarker assays by hybrid LBA/LCMS: practical recommendations to meet regulatory requirements

What is the importance of the use of SIL standards for hybrid LBA/LCMS biomarker assay validation versus other approaches? In cases when a surrogate tryptic peptide is used to quantitate a full-length protein analyte, what are the differences/caveats for using a full-length labeled protein as an internal standard compared to a winged peptide (with tryptic sites flanked by a few amino acids) compared to a fully tryptic labeled peptide? Have any comparisons been made to quantitate the same protein using all three approaches mentioned above? How well do full-length recombinant labeled proteins mimic/control for endogenous analytes? What are some criteria (in addition to correlation, Bland-Altman plots) that can be used to determine whether a hybrid LBA/LCMS assay and LBA assay can be used interchangeably to get to the same conclusions? Is multiplexing better? What are some of the benefits/caveats compared to multiplexed hybrid LBA/LCMS assays? Would a full validation be warranted if one wanted to go to a 'uniplex' hybrid LBA/LCMS assay with one of the analytes from the multiplex assay?

Latest advancements for improving drug tolerance & specificity by hybrid LBA/LCMS

Can LCMS be used as an alternate platform for ADA analysis? Can a LCMS-ADA assay be used for screening, titration, or both? How are LCMS-ADA assays validated? Are they similar to ADA immunoassays? What are the pros and cons of using a species-specific ADA positive control for validation? Is it important to determine cut-points for LCMS? Can LLOQ or LOD be used instead? Can LCMS-ADA assay improve drug tolerance of the assay?

Key Terms

Immunogenicity: The ability of a substance, including biotherapeutic substances, to stimulate an immune response in animals and humans.

Endogenous Substance: A molecule, protein or substance that originates from the biological matrix.

Catabolism: Formation of secondary molecules derived from complex molecules in living organisms.

Isotype: Antibody class and subclass of a given species.

Microsampling: Sampling of very small sample volumes (µl) from animals and humans to assess drug and chemical exposure in biological matrix.

Recent progress in multiplexing ADA by hybrid LBA/LCMS for antibody screening, titering & isotyping

How do you establish cut-points when background IgG, IgM, IgA or IgD levels are below the detection limits due to clean background? What does the assay sensitivity mean when it is established with a positive control, especially if the positive control is not a humanized antibody? Does determination of the absolute amount of ADA provide any additional value in comparison to a traditional ADA immunoassay approach? Can LCMS be used to determine ADAs in the presence of high circulating concentrations of drug to overcome the issue of drug tolerance? Can LCMS be used for immunoglobulin allotyping/isotyping by specific immunocapture: IgG, IgM and other immunoglobulin subclasses and/or by specific signature tryptic peptides as a surrogate for the measurement of intact ADA? More work has been published on the use of LCMS for ADA assays confirming the possibility of using LCMS for immunogenicity studies. Was this work submitted to Regulatory Agencies? If yes, did it receive any questions from Regulatory Agencies?

PK assays

Use of hybrid LBA/LCMS technology compared to LBA alone to overcome the PK, biomarker & immunogenicity assay challenges of novel biotherapeutic modalities

With tight timelines, it is difficult to wait until LBAs do not meet assay expectations to develop a hybrid

LBA/LCMS method. Do bioanalytical scientists conduct LBA and hybrid assay development in parallel or in series, if one approach is not working? What is the best/recommended approach for method development to save time? Do scientists have criteria (perhaps a decision tree) that determine when to develop a LBA and when to develop a hybrid assay? Should we start with a hybrid assay by default? Overall, can hybrid LBA/LCMS technology, compared to LBA alone, overcome the PK, PD, and ADA assay challenges of novel biotherapeutic modalities? What considerations and recommendations can we include to take advantage of adding LCMS to the traditional LBA approach as a starting strategy when developing assays for novel biotherapeutic modalities? Is it advantageous to perform a comparative evaluation of PK, PD and ADA to detect common interferences?

Strategies to avoid ADA interferences in PK assays by hybrid LBA/LCMS

Can a direct LCMS assay (reagent free i.e. total digest) be used in combination with ELISA to better understand the PK of biologics in the presence of ADA? Are other groups performing free/total/bound measurements for biologics? What platform(s) are used? What information is gained from this? Can the high specificity of LCMS eliminate interferences and discriminate among isobaric interferences, modifications, substitutions, unique signature peptides not present on ADA? Is it possible to use harsher conditions than in LBA for ADA dissociations like alkaline, acid, Guanidine combo treatments?

Progress and fine-tuning of validation criteria for hybrid LBA/LCMS PK methods: building on previous recommendations

What is the best approach to test clinical hybrid assays for ADA resistance, especially where humans have not yet been dosed? What is the best solid state method for hybrid assays? If using beads, is on-bead or off-bead digestion the best approach? Why? Should IA capture efficiency be routinely measured when Hybrid LBA/LCMS methods are validated? What is the best approach used by the industry to overcome problems with matrix effect determination for Hybrid LBA/LCMS methods? Is there any progress from previous recommendations on recovery, digestion efficiency, protein assay stability and critical reagent evaluations?

Evaluation & impact of mAb biotransformation in vivo on PK assay accuracy in non-clinical & clinical studies

How widely is biotherapeutic biotransformation characterized in the industry? What is the best approach for biotherapeutic biotransformation characterization: top down versus bottom up versus integrated approaches? Why is it important to understand biotherapeutic biotransformation? What is the impact of biotherapeutic biotransformation on PK assays, safety, and efficacy? Is the use of LCMS' or HRMS' high specificity attribute the best approach to identify, characterize and quantify biotherapeutic biotransformations? How is the quantification of minor modified biotherapeutics in the presence of a large amount of unmodified entities handled?

Discussions, consensus & conclusions

Biomarkers & immunogenicity assays

Biomarkers in solid tissues: benefits of hybrid LBA/LCMS & specific sample preparation techniques designed to minimize interferences

The expansion of new biotherapeutic modalities has increased the use of PK/PD modeling and system pharmacology. This has subsequently increased the demand from bioanalytical and biomarker groups with respect to quantitative measurements of biologics and biomarkers in both biological fluids and tissues. Tissue measurements come with increased scientific challenges compared to those in serum and/or plasma. The challenges come from many factors, such as procuring and handling the tissue sample, recovery of the analyte of interest and properly normalizing the target measurement of interest. However, quantitative tissue measurements, at the site of pharmacological action, can be critical as they may provide information to scientists about target validation, target coverage for dose projections, occupancy (total/bound), precision medicine for patient stratification, normal versus disease for target or biomarker measurements, membrane bound or intracellular proteins and, finally, drug disposition.

The use of Hybrid LBA/LCMS approaches offers an effective way to meet the requirements of sensitivity and cleanup associated with analyte measurement in solid tissue. However, best practice is strongly dependent on the analyte under investigation and intended use of the data. Hybrid LBA/LCMS methods are primarily used in discovery bioanalysis or for exploratory purposes since standardization and recovery tracking remains one of the major hurdles. In order to reduce potential interference from residual blood, in particular in PK analysis, it is recommended to wash the tissue sample using PBS upon tissue harvest and immediately flash freeze the sample in liquid nitrogen to prevent proteolysis. For exploratory, non-clinical studies, homogenization in acid at the collection site may be needed to better control the process, but obviously, this needs to be tested during assay development.

When the intent is to perform a single measurement representative of drug distribution in the entire tissue sample, heterogeneity of the target protein can be a challenge. Hence, pulverizing and homogenizing the tissue is recommended to get a representative overall measurement. It was agreed that in this situation consideration should be given to the stability of the analyte at elevated temperatures, since heat may be generated during the homogenization process. When possible, samples should be kept at low temperatures during the homogenization process. Typically, a stainless steel or similar bead-based approach performed directly in vials is appropriate. It should be noted that different beads are available for different tissues. To properly weigh the tissue, it is recommended that tubes be pre-weighed with lysis buffer containing protease inhibitor, and then weighed following sample addition. It is generally preferred by PK/PD scientists for protein concentrations to be normalized against tissue weight rather than protein content. Following pulverization/homogenization, the tissue is removed and can be reprocessed using additional buffer if there is any concern that all the analyte may not have been extracted. If less than 5–10% is left in the second extraction, there is no need to repeat the process.

It was agreed among the scientists that currently there are limited options to improve analyte recovery from tissues, especially when combined with IA sample processing. Lipids can often cause poor protein recovery especially for protein biomarkers. Therefore, assessing the effect of lipids on the recovery of the protein analyte is recommended; a lipid clearing solution may be used. The overall consensus is that dilution typically increases recovery, thus samples should be diluted as much as possible in the homogenate lysate to maximize recovery and minimize interference. However, dilution may affect the measurement of certain types of analyte forms: such as free and bound estimations.

Acceptable stability parameters of the analyte need to be defined *a priori,* depending on the sample collection process and the purpose of the data. Some attendees stated that it is difficult to test analyte stability in tissues directly, so it was recommended to bank homogenate lysate and test it over time. Different buffers and different storage conditions should be tested on the homogenate (e.g., freeze/thaw a lysate pool) to help assess if there are problems with stability. Commercial lysis buffers with protease inhibitors generally work well, and cryopreservatives (e.g., glycerol) can be considered. Furthermore, it was decided that it is not necessary to perform incurred sample stability. Instead, lysates should be banked in order to possibly look at additional protein targets. Finally, careful consideration should be put into tissue sourcing. Attendees felt that the quality of commercial tissue sources, while easy to obtain, can be uncertain due to the uncertain nature of collection and handling. Therefore it may be preferable to obtain tissues via collaboration with trusted academic sites and hospitals where possible.

Successful validation of sensitive confirmatory protein biomarker assays by hybrid LBA/LCMS: practical recommendations to meet regulatory requirements

LBA has historically been the method of choice for quantitating proteins and peptides. However, inconsistent specificity and lack of reproducibility among products are some of the challenges of this approach when using commercial kits [13,14]. Hybrid LBA/LCMS assays are increasingly used in addition to LBA because these are highly selective methods to quantify proteins and peptides in complex matrices therefore can provide confidence in measuring analytes of interest. Users of these hybrid LBA/LCMS assays indicate this new platform to be reproducible, robust, and specific for biomarker measurements. The addition of an IA step increases the sensitivity of the assay in order to measure very low levels of proteins in samples. These assay platforms can provide a 10-fold increase in sensitivity, which can improve detection significantly [15]. Nevertheless, factors such as cost, sample volume requirements, assay throughput, and ease of implementation continue to be limiting factors to consider when deciding on the assay platform.

One of the cornerstones of a robust LCMS method is the use of an appropriate internal standard. Industry best practice is to use a SIL internal standard if available. This is equally important for Hybrid LBA/LCMS assays, and in cases where a surrogate tryptic peptide is used to quantify a full-length protein analyte, scientists have the option of using a full-length SIL protein, a winged SIL peptide where the tryptic peptide of interest is flanked by extra amino acids, or a labeled tryptic peptide of the exact length. If a labeled protein is available, then it is the best choice and if possible to use the same cell line as was used to produce the recombinant protein to maintain constant PTMs. However, industry experience indicates there is no determinable advantage in using a full-length protein versus a winged peptide; any differences encountered were likely due to other issues within the assay even if, in theory, having a full-length protein internal standard can be very useful when IA recovery is in question. Differences were observed when using a labeled protein versus a peptide only when severe solubility issues were present (i.e., membrane or PEGylated proteins). SISCAPA was considered an option. On-column capture should be used in this case. In conclusion, the

overall recommendation is that as long as validation tests demonstrate that the assay is reproducible and robust, there was no preference for the type of internal standard used. Moreover, when assaying endogenous analytes, using a full-length recombinant labeled protein may be impacted by differential expressed protein isoforms and PTMs. Lot-to-lot variability of recombinant proteins is a significant challenge, and should be mitigated by minimizing the number of lots used, performing bridging assays and ensuring robust reagent characterization.

To determine whether a Hybrid LBA/LCMS assay and LBA assay can be used interchangeably to get to the same study conclusions, correlation and Bland-Altman plots are still the recommended criteria. It is difficult to absolutely match historical data using different qualitative assessments. When moving to a uniplex hybrid LBA/LCMS assay from an already validated multiplex assay, it was agreed that some cross-validation experiments were minimally recommended to assess comparability of results.

Latest advancements for improving drug tolerance & specificity by hybrid LBA/LCMS

Therapeutic proteins and peptides have the potential to elicit immune responses resulting in ADAs that can pose problems for both patient safety and product efficacy. During drug development, immunogenicity is usually examined using a risk-based approach along with specific strategies for developing 'fit-for-purpose' bioanalytical approaches. LBAs are the most widely used platforms for ADA detection and characterization because of their high sensitivity and throughput. During the past decade, LCMS and Hybrid LBA/LCMS methods have been demonstrated to be successful techniques for the quantitation of biotherapeutics for PK assays and protein biomarkers in biological matrices, mainly owing to their high specificity, selectivity, multiplexing capability and wide dynamic range. More recently, Hybrid LBA/LCMS methodologies have been developed for simultaneous semi-quantitation and isotyping of the ADA in human and animal plasma or serum. ADAs are captured by biotinylated drug or the drug–ADA complex is captured by an anti-drug antibody. ADAs are then eluted and subjected to trypsin digestion followed by LCMS detection of specific peptides for each antibody isotype. The relative concentration of each ADA isotype in the extraction buffer is quantified using a calibration curve which is generated from reference materials of each ADA isotype prepared in the extraction buffer.

The proof-of-concept of this hybrid methodology is demonstrated by detecting pre-existing ADA in human plasma [16]. Either biotinylated drug or biotinylated ADA could be used as the immunocapture reagent, each with its own merits and shortcomings. Biotinylated drug can readily capture ADA, but drug interference could be an issue if drug levels in the samples are high. On the other hand, immunocapture using an ADA eliminates the drug tolerance issue which is often associated with immunoassays, providing that the ADA is able to capture the drug-ADA complex in addition to free drug. With this method, unique peptides from each ADA isotype/subclass can be identified and monitored by LCMS. ADA isotyping can be performed by the detection of isotype-unique peptides. An absolute ADA amount can be obtained however the result is only semi-quantitative since the polyclonal ADA response cannot be mimicked with calibrants and commercially available Ig isotypes are spiked only post-elution to create the calibration curve. Similar to traditional ADA LBA assays, cut-points at 95% confidence interval can be established using undosed controls. Endogenous Ig interferences need to be reduced in order to improve assay sensitivity. Human positive ADA controls or the controls containing human Fc are recommended for more accurate estimation of the LLOQ/LOD. Protein A/G beads can also be used to capture total immunoglobulins and then the drug-specific surrogate peptides can be measured to reflect the amount of the immunoglobulins bound with the drug. This application is only suitable to the drug which does not have binding affinity to Protein A/G, such as Fab domain antibody or other non-Fc proteins.

It was agreed that, with further improvements and industry experience, Hybrid LBA/LCMS methodology could become a useful tool in immunogenicity assessments. At present, this methodology is used for experimental purposes and not used in regulatory data acquisition. It can be easily implemented in bioanalytical lab settings for routine ADA isotyping and semi-quantitation. The key issue is the reliability of immunocapture step. As ADA levels measured by Hybrid LBA/LCMS represent absolute amounts relative to the positive control used, one can compare ADA isotype levels between samples, between studies, as well as between different biotherapeutics, providing that consistency in positive controls is achieved to determine recovery. A database of such information could be gradually built and provide valuable insight to better understand immunogenicity of biotherapeutics.

It was also agreed that Hybrid LBA/LCMS ADA assays could become an alternative for screening, isotyping, and titration without a need for confirmatory assays due to better specificity (lower false positives). In cases where there is interference from the drug itself, there could be benefits for a confirmatory assay, but this is assay and positive control dependent

and more information is needed prior to widespread adoption. Each IgG response is different from patient to patient, so it is critical to compare post-dose samples to the same patient's pre-dose sample. In the case where Hybrid LBA/LCMS could be used for screening assays, a separate titration may not be required because the screening assay provides the relative concentration of ADA. However, the assay is only semi-quantitative because reference standards are not spiked into the matrix. Commercial IgG or IgM, not the actual ADA, is used to run the calibration curve. The selected surrogate peptide may be significantly different from the actual ADAs, even if the same isotype but difference species.

Neither the 2009 draft regulatory guidance for ADA assays [17] nor its 2016 revised version [18] discuss the use of Hybrid LBA/LCMS for immunogenicity assessment. Therefore, it is recommended that the requirements for the validation of Hybrid LBA/LCMS ADA assays be leveraged from the best practices and lessons learned from LBA ADA assays. In addition, because of the absence of regulatory guidelines regarding validation of Hybrid LBA/LCMS ADA assays it is recommended that sponsors contact the Regulatory Agencies and discuss assay validation plans with them prior to conduction. Currently, using a species-specific ADA positive control for validations can help establish detection limit, and is essential for surrogate peptide selection in Hybrid LBA/LCMS assays. Usually due to availability reasons, positive controls are often monoclonal whereas an ADA response is polyclonal. However, positive controls cannot represent all endogenous ADAs, introducing limits to the method validation. This is the same as for the LBA ADA assays. Cut-points are important statistically when a broad population is evaluated, so it is likely both cut-points and LLOQ/LOD are needed.

Recent progress in multiplexing ADA by hybrid LBA/LCMS for antibody screening, titering & isotyping

Currently, LBAs are widely used platforms for evaluating the immunogenicity of protein therapeutics, and a tiered approach is used to screen, confirm and titer the ADA present in patient samples. However, drug interference in the assay can prevent low level and low affinity ADAs from being detected. When it comes to isotyping of ADAs, although new technologies are being developed with LBAs, these technologies have not been broadly utilized in the pharmaceutical industry due to their complexity, as well as the infrequent and unclear need for isotyping. It was agreed that Hybrid LBA/LCMS assays can be used for semi-quantitative measurement of ADAs through either direct measurement of IgGs, IgMs, IgAs, IgEs and IgDs and their respective subclasses bound to the drug, or indirect measurement of the drug bound to ADAs. However, the use of Hybrid LBA/LCMS assays to determine the relative amount of ADA does not show additional value in comparison to a traditional ADA LBA approach. There are still limitations to using LCMS for very low levels or low affinities of ADA due to lower sensitivity than LBAs. Additionally, it is not clear why isytoping would add value in most cases. Sensitivity for a Hybrid LBA/LCMS assay will be established using cut-points when the positive control is not a same species antibody because different surrogate peptides are measured.

There has been recent progress in multiplexing ADA measurement and isotyping immunoglobulin subtypes by Hybrid LBA/LCMS [16]. This technology provides a way of overcoming the drug tolerance issue as long as adequate sensitivity can be established. However, background issues are still an overall challenge. More work has been published on the use of Hybrid LBA/LCMS for ADA assays confirming the possibility of using LCMS for immunogenicity studies and providing an alternative to the traditional tiered approach that uses three separate assays. However, regulatory agencies report that they have not yet reviewed any submitted ADA data using Hybrid LBA/LCMS at this time. Despite being a young application for Hybrid LBA/LCMS, multiplexing isoyping of ADAs is already used as a proof-of-concept for ADA determination but as industry and regulatory experience increase more of these assays are expected.

PK assays

Use of hybrid LBA/LCMS technology compared to LBA alone to overcome the PK, biomarker & immunogenicity assay challenges of novel biotherapeutic modalities

Novel biotherapeutic modalities include fusion proteins, bispecific antibodies, ADCs, and other emerging new entities. Developing assays for these modalities involves overcoming additional challenges to those already evident with monoclonal antibody therapeutics. The challenges, including assay interference from endogenous protein drug targets (usually a soluble target) or ADAs in PK assays, drug or target in ADA assays and target or binding partners for PD assays, must be overcome to develop an accurate and robust method. It was discussed whether a comparative evaluation of PK, PD and ADA data could help detect common interferences. It is recommended that the biology would first have to be examined to decide if common interferences are likely. Further, the interferences would be different based on the platform, limiting the practicality of this

approach. The procurement of the necessary reagents may also be a challenge for the novel therapeutics. There were many advantages acknowledged for combining LCMS with traditional LBA to create a Hybrid LBA/LCMS approach. LCMS can better characterize exactly what is being measured, improve the understanding of the analyte as well as the biology involved and also detect *in vivo* modifications (e.g., clipping of biotherapeutics with an endogenous counterpart). Method development is generally faster compared to traditional LBA, and comes with improved selectivity and less stringent reagent requirements. It is worthwhile to mention here that Hybrid LBA/LCMS assays can use reagents that do not work in LBA assays due to better selectivity of the mass spectrometers. It is also generally more resistant to interference due primarily to the specificity offered by the detection step and the use of only one antibody for capture. However, it should be noted that the Hybrid LBA/LCMS and LBA may not measure the same species, depending on the assay formats. Therefore, in some cases it is desirable to use both assays for bioanalysis in order to get a better understanding of PK/PD. On the other hand, caution should be exercised when comparing results from the two assays.

Overall, it was agreed that presently LBAs are often still used as the default approach largely due to the confidence resulting from a successful history with this technology, regulatory acceptance and existing in-house and CRO capabilities. Hybrid LBA/LCMS assays can also be used as the main PK assay or when issues like those discussed above arise. For a more proactive approach, it is recommended that scientists should begin developing LBA and hybrid methods in parallel with a clear path as to when to continue using one methodology based on initial result generation. This approach will work if a clear strategy is in place, allowing scientists to make an informed decision based on the pros and cons of both methodologies, as well as taking into account the phase and the biology of the drug that it would be used for. The current reality is that most scientists work in series due to financial and resource limitations. Furthermore, using a hybrid method limits the ability to transfer the method to another lab such as a CRO, who may not have that platform capability. It was agreed that Hybrid LBA/LCMS is a possible orthogonal method to LBA for PK and PD assays [19,20].

Strategies to avoid ADA interferences in PK assays by hybrid LBA/LCMS

The ability to accurately quantify total concentrations of biopharmaceutical therapeutics in the presence of ADAs is sometimes required to understand the safety and/or efficacy of these molecules in both preclinical and clinical studies. These studies are traditionally supported using LBAs which require binding of the therapeutic to a capture reagent (e.g., antigen capture or anti-idiotypic capture) and a detection reagent (e.g., Anti-Human IgG antibody or anti-idiotypic antibody) for quantitation. However, the presence of an ADA may compete with the binding of the therapeutic to the capture and/or detection reagent, thus preventing measurement and masking the presence of the circulating molecule. This interference can thereby have a damaging impact on safety and/or efficacy interpretations [21]. Recommendations have previously been made for assessing the impact of immunogenicity on PK assays [9,12,22]. Since then, an orthogonal approach like LCMS has emerged as an alternative tool to better understand the PK of biologics in the presence of ADA. It was agreed that the high selectivity of LCMS can minimize interferences and discriminate among isobaric interferences, modifications, substitutions, and unique signature peptides not present on an ADA. However, bioanalytical method characterization of ADA tolerance is imperative. Some traditional ways for solving ADA tolerance issues include acid dissociation, acid capture elution (ACE), biotinylated or passive adsorption, and increasing minimum required dilutions. When traditional LBA options do not solve these method issues, it is recommended that LCMS or Hybrid LBA/LCMS methods with harsher conditions than those typically used for LBA (e.g., alkaline, acid or guanidine combo treatments) could be used to achieve ADA dissociations. The appropriate platform for PK/PD endpoint assessments should be based on a biology driven strategy with consideration given towards the collective endpoints for interpretation.

Progress & fine-tuning of validation criteria for hybrid LBA/LCMS PK methods: building on previous recommendations

Much attention has been paid in the last several years on the validation of protein bioanalytical methods that quantify surrogate peptides via LCMS after proteolytic digestions [6,8,11]. An enhancement of these LCMS methods increasing in prevalence is the Hybrid LBA/LCMS approach. This approach utilizes affinity capture reagents to enrich the protein drug from matrix prior to digestion. These reagents can range from less specific affinity reagents such as protein A to highly specific immune-affinity capture reagents such as anti-idiotypic monoclonal antibodies or drug target. Techniques for coupling these affinity capture methods to the LCMS normally employ solid state methods using micro particles such as magnetic or sepharose beads

coated with streptavidin that binds to biotinylated reagents. The result is an increase in method sensitivity of up to 100 times allowing such assays to be validated and applied to a number of clinical and non-clinical studies with good success [11].

It was agreed that beads were the best solid state method for Hybrid LBA/LCMS assays, although the choice is assay dependent. On-bead digestion is recommended because of its ease and lack of extra steps when compared to off-bead digestion, although it was acknowledged that the off-bead approach is ideal because it gives the most complete digestion and is subject to less interference from non-specifically bound materials found on the beads. For many applications on-bead digestion is often sufficient for the purpose of the assay.

Along with the enhanced performance of Hybrid LBA/LCMS methods, some of the potential complications associated with LBA may become apparent. These include possible effects due to ADAs, circulating ligands, or specific interferences in patient populations and/or dosed subjects. Since these factors can directly impact quantitation and recovery, they must be evaluated carefully during method validation and monitored during sample analysis. Hence, it is recommended for Hybrid LBA/LCMS methods to a) study selectivity in both pooled control and pre-dose patient matrices; b) examine the effects of specific interfering ligands, such as soluble targets, at expected physiological concentrations upon drug quantitation in spiked QC samples; and c) examine the impact of added ADAs of known high titer, if available upon quantitation, in spiked QC samples. The best approach to test clinical Hybrid LBA/LCMS assays for ADA resistance, especially in drug naïve patients, is to use pre-existing ADAs, animal ADAs as positive controls.

It is recommended that recovery and digestion efficiency in validation are examined at low, mid and high QC concentrations. Special validation experiments designed to specifically measure affinity capture recovery should be tested with both spiked control and patient pre-dose samples to determine whether there is any effect of the patient matrix on immunocapture efficiency. The recovery results provide information on selectivity and matrix effect, which is a known difficulty when using mass spectrometry as a detection tool. Testing patient pre-dose samples as naive and spiked at the LLOQ concentration is the recommended method for testing matrix effect and selectivity.

It is suggested that parallelism experiments be performed during validation on incurred samples by diluting with disease patient, pre-dose, and control matrices, followed by quantification. The results are then compared to those diluted in buffer.

Finally, it was agreed that the issue of reagent management requires considerations similar to LBA reagents. Hence, reagent stability, suitability, and lot-to-lot comparability can often be evaluated on a fit-for-purpose basis. Many of these reagent management concerns can be handled by working in close collaboration with the reagent group of the organization due to their experience and extensive use of methodology to verify, characterize and maintain these critical reagents.

Evaluation & impact of mAb biotransformations in vivo on PK assay accuracy in non-clinical & clinical studies

Drug development is a slow process with a high failure rate, especially when new drug modalities are being introduced. At the same time, biotherapeutics are becoming even more structurally complex, with formats including bispecifics, fusion proteins, antibody fragments and cyclic peptides. Analytical methods to understand the fate of biotherapeutics *in vivo*, including structural stability or catabolism, have the potential to provide valuable information during drug development and may help speed up the process in getting drugs to patients. Historically, such a depth of structural information has not been available for biotherapeutic development. However more recently, mass spectrometry-based approaches have been used to explore large molecule *in vivo* catabolism [23,24] and is considered the best approach to identify, characterize and potentially quantify biotherapeutic biotransformations such as antibody-drug conjugates. Hybrid LBA/LCMS approaches developed for ADCs can be extended to mAbs and other modalities to identify biotransformations including deamidation, clipping and isomerization [11]. Such biotransformations in the proximity of the CDR target binding or Fc effector function binding regions of a mAb biotherapeutic has the potential to alter the *in vivo* pharmacology, affecting efficacy, safety, and PD evaluations. Biotransformations in the CDR region for mAbs also has the potential to alter binding to assay reagents in LBA or Hybrid LBA/LCMS PK assays and yield anomalous PK measurements. Understanding biotherapeutic biotransformations, therefore, will help scientists determine the impact on efficacy, safety, PD, and PK, as well as help to correlate in vitro monitoring. Traditionally, characterization of biotherapeutic *in vivo* catabolism is very limited and only done when needed. Top-down or middle-down approaches, which analyse the intact biotherapeutic or its sub-units (e.g., light chains and heavy chains of the mAb) via immunoaffinity capture and mass analysis. Another option is a bottom-up approach, where MRM or HRMS full scan analyses are used to target specific peptides or an open

proteomics discovery approach to identify biotransformations. However, it was agreed that an integrated approach would be the preferred option. In the case where it is necessary to quantify minor biotransformed biotherapeutics in the presence of a large amount of unmodified biotherapeutics, it is recommended to follow the MIST guidance for small molecules [25] present in quantities of more than 10%. The dynamic range of the method should allow for proper analysis. For quantities below 10%, the minor biotherapeutic is unlikely to be relevant, depending on the importance of biotransformation.

Key input from regulatory agencies

The 10th WRIB continued with its long standing contributions from several international regulatory agencies, including the US FDA, EU EMA, Health Canada, UK MHRA, Austria AGES, Italy AIFA, Brazil ANVISA, Japan MHLW and WHO. These agencies shared their views on multiple topics of interest for the global bioanalytical community attending this event, to provide clarification on unresolved issues or expectations. Regulators provided input on a variety of topics across all three core days, including:

- Microsampling ICH
- Extract stability
- Batch acceptance criteria
- Electronic data integrity
- Endogenous compounds
- Validation of biologics and biosimilars
- Immunogenicity issues
- Inspection programs, observations and new guidelines

An update on the recently released draft S3A Q&A document focusing on microsampling for toxicokinetic studies [26] was provided by a representative from Japan MHLW, who is the rapporteur of the ICH implementation working group that prepared the document. The document contains questions addressing microsampling covering scope, application, effect on safety and issues regarding bioanalytical methods. It was started in early 2015 and took approximately a year to complete and was released for comments. Consultation was open until September 2016 on the ICH web site, and it is anticipated that the final version will be adopted by spring or summer of 2017.

The subject of extract stability was brought back to the table by the US FDA, following up on the original discussion in the 2009 White Paper in Bioanalysis [1]. The Agency stated that there are still differences in industry practices with regards to the evaluation of extract stability. The US FDA representative reminded those present that the objective of the extract stability evaluation is to determine stability of the analyte(s) and internal standard(s) post-extraction. This is referred to as the post-preparative stability in the US FDA Guidance on Bioanalytical Method Validation [27]. Post-extraction storage of samples may be at room temperature or colder, and extract stability can address stability of the analyte following the resident time of the samples on the autosampler for the duration of the run. This is particularly relevant if the run is interrupted for some reason, and injection is continued; the analyte and IS stability should be preserved from the beginning of the run to the end of the run. Two procedures were presented. The first was to inject high and low QC stability samples along with a freshly prepared set of calibrators and analyze the QCs. The stability QC samples are then subjected to storage conditions for a determined period of time and reinjected using freshly spiked calibrators. Calculated concentrations are then compared to nominal values for determination of stability. Some advantages of this approach are that the use of ratios addresses possible differences in ionization or other sources of variability, and freshly prepared calibrators are used in case of similar degradation between analyte and IS. Some disadvantages to consider are that the fresh calibrators are not processed with the stability samples, introducing additional variability (e.g., extraction variability), and once the sealing mat is punctured for the stability QCs, there is a potential for evaporation. Another method was proposed, where the stability samples are extracted with a fresh set of calibrators and batch QCs. The curve and batch QCs are injected immediately, but the stability QCs are not injected with these. Instead, stability samples are immediately stored and injected on a future day (after a period of aging), but read against the originally injected curve. In this case, all samples are processed (extracted) together and there are no issues with evaporation, since the sealing mat of the stability samples remain unpierced. This method, however, assumes stable instrument response and spray conditions over the course of the experiment. Neither method was favored as the preferred strategy. However, the pros and cons of each approach should be considered when determining the most appropriate design of the experiment based on the method.

The US FDA also introduced the topic of batch acceptance. This is particularly pertinent when multiple 96-well plates are used, and each plate is processed individually but injected in sequence in one run.

Regulators have come across data where batch criteria were applied across the multiple plates, as opposed to each plate individually. However, in cases such as this, where each plate was processed individually, the expectation is that there is at least a complete set of QC samples on each plate along with study samples. This way, once the instrumental run criteria pass, each batch can then be evaluated independently. This method provides the most information about the quality of the data generated and is consistent with good science, US FDA recommendations and the EU EMA guideline on BMV [28]. It is important that SOPs clearly define what constitutes a batch.

Austria AGES and US FDA addressed the increased focus by regulatory agencies on electronic data. Data integrity is of primary concern, thus inspectors are very vigilant in identifying signs of manipulation or falsification. Data integrity issues are not limited to e-records, electronic storage media and software, but also involve all information management systems used by a company. Appropriate data management, as well as maintenance of data integrity, are areas prone to risk and should have an adequate level of risk analysis. Poor data management can open the door for unpredictable malpractices. Electronic laboratory notebooks have benefitted many companies by improving efficiency and productivity, prolonging record longevity, allowing for efficient and global information sharing and retrieval, and by facilitating real time data review. Industry has been discussing the impact of their use for many years [8,11,29]. Regulators indicate that they are reviewing data from a variety of different ELN systems, as one size does not appear to fit all, resulting in many challenges occurring during site inspections. ELN formats, structure and granularity differ from site to site, forcing regulators to take time in learning the navigation of the system. Hyperlinked information in worksheets may be read-only or links may be broken, thereby causing delays during the inspection. This may result in extended waiting times for accessing all required information while waiting for site staff knowledgeable in the system. Furthermore, analytical data audits become challenging when different aspects of the analytical data are documented in different systems (e.g., old versus new, paper versus ELN). Therefore, the inspection scope has been expanding in order to comprehend the IT aspects of ELNs for assuring data integrity. Documentation of study data within an ELN should still follow all predicate rules. The individuals who carried out each aspect of study conduct should be clearly captured, as well as the date and time when the study events actually occurred, and not just when they were recorded. This is usually done via e-signatures and timestamps. Access must be available to the audit trail to determine what changes occurred and the individuals who made them. It should also be noted that efforts are made to inspect the ELN source files and not just PDF export files. It was recommended that any access granted to the inspectors should not affect the system performance. However, inspectors should still be able to completely reconstruct the study within the system. To quickly orient themselves with the ELN structure and possible links to other LIMS at the firm, regulators may request a system map. Finally, ELN SOPs should be in place to elaborate on topics such as the overall workflow and archiving, version control of forms and templates used, audit trails, the ELN review and approval process, and the roles and responsibilities within the ELN. Apart from ELN systems, all electronic data are within the scope of a regulatory inspection. Regulators still consistently note cases where there are no personnel with expertise in computer validation at the firm. Gaps in the proper validation of a computer system increase the risk of data integrity issues. The evaluation of risk must include assessments on the possibility of deleting or modifying data from outside the system, network storage of analytical and PK data without transparent access control, intermediate study results that flow prematurely via unauthorized access to the software, and ineffective QC and QA verification practices. Access to the software must be controlled and justified by the job description. The data flow must be clearly described in the processes such that the verification of batch results, data approval, release for PK and handoffs to pharmacokinetic scientists and QA are clearly tracked and documented. Any weaknesses of the quality system, data management and the use of computerized systems make it difficult for regulators to trust the position of management and results in significant suspicion about the reliability of trial data, thereby requiring further investigations during the inspection.

US FDA provides some guidance on the analysis of endogenous compounds [30–32], and some white papers have also provided recommendations on this topic [33,34]. Endogenous compounds can originate from the endocrine glands (e.g., glucagon, estrogens) or from dietary uptake (e.g., iron, omega 3 fatty acids). When designing a study for endogenous compounds, several points need consideration. First, the selection of the matrix should be appropriate. The PK characteristics of the drug substance should be considered in the selection of the matrix (e.g., selecting urine to assay potassium because potassium concentrations in the body are controlled largely by the kidneys through urinary excretion). The stability of the analyte *in vivo* is also important. Consideration for the use of low background level populations may reduce the need

for baseline level corrections (e.g., post-menopausal women for estrogen studies). If corrections are needed, an evaluation of the characteristics of the baseline levels is beneficial (i.e. expected variations and magnitude of variations), as well as the methods that will be used to perform the correction (i.e. point-to-point, average of pre-dose values). Method validation of endogenous compounds includes some extra challenges. It is necessary to reliably determine the baseline level of the compound in the matrix of choice. If a surrogate matrix is used, additional validation data are needed to show equivalent performance of the analytical method between the surrogate and the subject matrix. QC samples prepared in authentic matrix should always be included during accuracy and precision. Stability in authentic matrix is recommended, even for rare matrices. Recovery between the analyte-free matrix and authentic matrix should be comparable and complete. Assay performance in multiple lots of surrogate and authentic matrix should be cross-validated to prove that there are no matrix effects. It was made clear that while the fit-for-purpose concept for the validation of biomarker assays is appropriate for different stages of drug development, for ANDAs, submission of a fully validated analytical method is expected.

US FDA shared their experience on the validation of immunogenicity assays for biologics and biosimilars. Indeed, these submissions have increased over the last several years as more pharmaceutical products of these types are being developed and marketed. As a result, regulatory experience with these types of submissions is more readily available. Cases exist where a low positive control concentration in an ADA assay was selected with an assay value much higher than the assay cut point and/or study sample values. This is concerning because there is no indication of precision at the low end of the assay range, resulting in uncertainty that lower ADA concentrations can be detected. Guidance [16,35] and literature [36] recommend selecting a low positive control concentration resulting in assay failure at a rate of ~1%. Non-reproducibility between Tier 1 and Tier 2 assays also raises concerns for the inspectors, because assay reliability is then uncertain and the potential for false negatives in the Tier 1 screening assay increases. There is a potential for misidentification of ADA positive samples and inaccurate assessment of the ADA response rate when samples are defined as negative during the neutralizing assay because they were less than the screening cut point even when they were confirmed positive and no further analysis is performed. Confirmatory cut-points should not be set too high, nor be variable; otherwise there is no assurance of assay specificity. Titer assay cut points are of equal concern when dilutions of patient samples for the Tier 3 assays are performed with normal matrix instead of patient matrix, due to e.g. unavailability. In these cases, reported titer is not indicative of the true titer, and there is a potential failure to err on the conservative side of titer reporting. Finally, using a LIMS system for immunogenicity assays brings risks for run identification errors, failure to monitor raw assay values of positive controls and for the deactivation of valid assay results. Immunogenicity data generated using LIMS systems should ensure that no acceptable data is deactivated and that there are pre-established criteria for reporting of final titer values and for governing sample repetition specific to the titer assay.

US FDA, Health Canada and Japan MHLW addressed some immunogenicity issues. Health Canada presented some typical issues that are seen during the review of immunogenicity data as part of marketing approval submissions. Currently, it is very difficult to determine the impact of drug immunogenicity on the benefit/risk profile of a biologic. Safety of the patients should be evaluated in the context of the particular disease being treated as well as by comparison to similar products on the market or in development, which may shed light on alleviating any safety issues. For example, adverse events related to immune reactions should be monitored within a meaningful timeframe, which can range from hours to several months after the last dose of treatment. Determining the impact of ADAs on efficacy is also a challenge. Some examples were provided where data was unable to show either the neutralizing or the non-neutralizing antibodies' impact on efficacy. Another case was shown where the long-term persistence of ADA was widely variable across the patients tested, and no conclusions could be drawn on the impact. Finally, the quality of the methods for detecting ADAs could be questionable. Methods were sometimes not adequately validated, especially with regards to specificity and sensitivity, and so data drawn from these methods were not considered reliable. In light of these challenges, Health Canada clarified that, at the CTA stage (equivalent to the IND step in the US), the apparent immunogenicity of the test drug alone is not considered a 'show stopper' as long as the risk/benefit profile of the drug is acceptable. Furthermore, at the NDS stage (equivalent to the BLA stage in the US), the agency's clinical decisions are based on the efficacy response and the benefit/risk profile of the drug, since, typically, study results conclude that no correlation between immunogenicity and efficacy/safety was found. Health Canada evaluates the role of immunogenicity in each submission on a case-by-case basis and collaborates with other regulatory agencies to develop a consistent approach. Japan MHLW indicated that they do not have specific guidelines for immunogenicity

assessments. However, EU EMA and US FDA guidelines/guidance addressing the subject are good references to use for Japanese submissions, as the basic concepts are similar to the approach taken in Japan. The strategy for immunogenicity assessments, including the sampling schedule, depends largely on the clinical target product profile. In order to fill the gap of lack of guidance in Japan, an AMED research group is in the process of creating a document that includes some points to consider for ensuring the reliability of ADA assays. This document will be the basis for a future guideline for immunogenicity assessments of therapeutic proteins in Japan. Some considerations include cut point determinations, drug tolerance and biosimilars. Some additional questions regarding immunogenicity were posed to the regulatory panel. First, the situation was presented where some screening positives came in under the cut point in the confirmatory assay. This was considered acceptable because some false positives do occur and are expected, and the confirmatory assay is used to confirm this. Regulators agreed with this, however, their concern is raised when a significant portion of the samples demonstrate screening false positives. In this case, an investigation would be warranted. Regulators were also asked their opinion on using a raised titer cut point instead of a screening cut point. It was acknowledged that some scientists follow this practice, and it is supported in the literature, however, one should err on the side of caution when deciding to accept this approach.

Several agencies presented summaries of their inspection programs as well as some examples of their observations and new guidelines:

UK MHRA highlighted that although bioanalysis is often discussed under the guise of GLP regulations [37,38], it is important to underline that clinical study samples must be analyzed under GCP [39]. Often, bioanalysts claim that since they only deal with bioanalytical samples and not clinical patients, GCP is less relevant. However, subjects have certain rights and these rights extend to the way that clinical samples are taken, processed, analyzed and reported. Guidance exists for laboratories that perform the analysis of clinical trial samples [40,41] in order to capture the relevant portions of each set of practices. However, harmonization between agencies on the applicability of GCP (UK MHRA, US FDA via 21CFR320, Health Canada, Japan MHLW), GLP (ANVISA) or GCLP (WHO) to clinical bioanalysis is not consistent. That being said, regulators expect that the roles and responsibilities outlined in the GCP guidelines be followed. Training of staff working on clinical studies must be appropriate in education, experience and technical trainings received on the job.

UK MHRA continued its regulatory input by stating that staff involved in the analysis of samples from clinical trials must have an appropriate educational background, relevant experience and have received appropriate technical training relating to the work they undertake. GCP training commensurate with the employees' roles and responsibilities is required, and should be refreshed at appropriate intervals. Finally, unblinding of samples was discussed. The bioanalytical principal investigator first needs to consider whether there is an actual need to unblind the samples and what the consequences of unblinding for the trial might be. Any request to unblind must be approved by the sponsor. The process for unblinding should be outlined in an SOP and well documented including the rationale for the decision. Once the decision to unblind has been taken, control of the unblinded data is essential and consideration should be given to who needs to know and how the information will be communicated. Storage, transfer and retention of the unblind data also need to be considered. The clinical study report should include details of the need to unblind study samples and whether there is any perceived impact on the conduct of the trial.

Italy AIFA underlined that although a sponsor may transfer any or all of their trial-related duties and functions to a CRO, the ultimate responsibility for the quality and integrity of the study resides with the sponsor. Therefore, the sponsor as well as the applicant, if needed, should demonstrate the controls in place in order to guarantee the quality and integrity of the study data. For the clinical portion of the study, this is required via the monitoring function. If monitoring is extended to the bioanalytical portion of the study, qualification of the monitor in the specific field should be demonstrated and documented. Independent audits, performed by qualified people, are also an acceptable way to verify the bioanalytical phase of the study. In any case, laboratories should be included by the sponsors in their qualification programs of vendors/third parties.

WHO presented a list of case studies on clinical trial issues, violations and impact on bioanalytical data reliability. Most products submitted for WHO prequalification are generics and are therefore supported by clinical bioequivalence studies or biowaivers. WHO inspections have shown that data integrity is a worldwide issue and both the sponsors and the CRO are responsible. An increasing trend on the number of falsified study data occurrences have been seen and could be due to a number of reasons, including an increased rate of detection due to recent focus and training of inspectors in the area of data integrity. It was hypothesized that one of the motives

for data manipulation could be that the sponsors believe that it is less onerous to falsify bioanalytical data than to reformulate a failing product. On the other hand, the CROs want to provide sponsors passing studies to avoid the sponsors seeking out other competing CROs. WHO, in collaboration with other agencies, are actively working to assure data integrity by increasing their focus on data appearing suspicious, by including inspectors with IT skills as part of inspection teams, as well as increasing awareness on the topic by the publication of guidance to industry [42,43].

Brazil ANVISA presented some summaries of new technical notes that have been released since last year's WRIB conference. The first covers bioequivalence studies for generic products which contain vitamins above the daily intake and associated with APIs [44]. Bioequivalence to a Brazil ANVISA reference drug must be shown. A BCS biowaiver is not applicable. PK parameters of the unaltered vitamin in the appropriate matrix are required. Baseline values at specific intervals are required to establish a mean concentration per subject, which is then subtracted from the calculated concentrations. Vitamin intake from food must be controlled; a detailed menu must be submitted. The technical note regarding PK interactions for fixed dose combinations [45] is mandatory to follow when the efficacy and safety of the combination has not been documented in the literature or when one of the molecules is a new molecule in Brazil. Two BA studies would be needed comparing the combination dose to each individual Brazil ANVISA reference product. Regarding regulations for PK studies for biosimilars [46], the applicable LBA or LCMS validation criteria [47] is expected. Reporting must follow resolution RE 895/2003. It is recommended to conduct these types of studies in Brazil ANVISA certified CROs.

Health Canada advised attendees of three new documents that they had recently released. The first is a draft guidance document sent for consultation, which discusses the disclosure of confidential business information [48]. The final version of their Proposed Policy on Bioequivalence Standards for Highly Variable Drug Products [49] had also just been recently released as well as a guidance on Post-Notice of Compliance (NOC) Changes [50]. Following this update, the latest perspective on their notice on the Clarification of Bioanalytical Method Validation Procedures [51] was discussed. This notice was released by Health Canada in October 2015 with regards to the number of tubes used to perform matrix-based stability evaluations. Immediately following the notice, the agency began issuing deficiency letters asking for clarification, or in some cases, additional stability information based on the specifics of each case. The interpretation of the notice by the industry was not uniform, and deficiency letters were asking for different stability evaluations to be conducted depending on the details of the case, causing generalized confusion over the scope of the requests. The impact was wide-reaching and costly for applicants in both resources and time. In order to clarify the agency's position on this topic, an addendum to the notice was issued in March 2016 [52] and discussed at the conference. The notice applies to bioanalytical methods (primarily employing LCMS) for the measurement of small molecules in BE studies that are pivotal to the regulatory assessment of human therapeutic drug products in generic and innovator submissions. Health Canada's interpretation and expectations for all matrix-based stability experiments (i.e. long-term, freeze-thaw and bench top) is that a minimum of 3 tubes be stored and assessed at each QC concentration. Typically in the past, matrix-based stability experiments were reported using six values (n = 6), and it had been assumed that each value represented an individual tube. However, the agency came to understand that some companies were storing one single bulk tube, but then processing and misreporting the data as six replicates. Health Canada considers this a misrepresentation of the data, since sound scientific practice indicates that this modified approach should be reported as a single (n = 1) stability tube. The US FDA and EU EMA BMV guidance/guidelines also specify that a minimum of triplicate values must be reported. During the discussion the other regulators were in line with Health Canada's thinking and expectations. It should be noted, however, that there was considerable disagreement from attendees, who claim that their results show no impact from storing a single bulk tube or replicate tubes. In response, the agency noted that they were open to industry compiling data to support their position, but strongly cautioned that such data is specific to an ingredient and to conditions of the method, so that extrapolating across ingredients and methods would likely not be acceptable. Going forward, data submitted prior to the October 2015 notice will be evaluated on a case-by-case basis to determine if additional stability data will be requested. However, it is expected that all submission after October 2015 contain matrix stability performed using a minimum of three tubes per QC concentration.

Recommendations

Below is a summary of the recommendations made during the 10th WRIB.

Biomarkers & immunogenicity assays

1. It is recommended to wash solid tissue samples with PBS during sample collection to remove possible blood contamination and flash freeze as soon as possible. When heterogeneity of the analyte distribution is a concern, use of pulverization and homogenization techniques is recommended for the entire tissue sample to ensure a representative overall measurement. Lipids can often cause a poor recovery, therefore performing a correlation with lipid concentration and recovery of protein analyte is recommended using a commercially available lipid clearing solution. It was agreed that it is very challenging to test analyte stability in tissues, so it is suggested to test stability of homogenate lysate over time instead.

2. For protein biomarker assays by Hybrid LBA/LCMS, as long as validation tests demonstrate that the assay is reproducible, there is no preference which type of labeled internal standard, whole protein, winged peptide or exact peptide, should be used. To determine whether a Hybrid LBA/LCMS assay and LBA assay can be used interchangeably, it is recommended that correlation and Bland-Altman plots be used. When changing to a uniplex Hybrid LBA/LCMS assay from a validated multiplexed LBA assay, some cross-validation tests are recommended to establish assay data comparability.

3. Overall, Hybrid LBA/LCMS ADA assays can be an alternative for screening, isotyping, and titration assays without needing a confirmatory assay. In cases where there is interference from drug or from endogenous substances, there could be some benefits of using Hybrid LBA/LCMS ADA assays for mitigating drug tolerance issues and decreasing the false positive rate. The requirements for the validation of Hybrid LBA/LCMS ADA assays should be leveraged from the best practices and lessons learned from LBA ADA assays, since currently there are no regulatory guidelines specific for Hybrid LBA/LCMS approaches. Using a species-specific ADA positive control for validations can help establish detection limits during method validation, and is essential for Hybrid LBA/LCMS assays for surrogate peptide selection.

4. Hybrid LBA/LCMS can be successfully applied to semi-quantitative measurement of ADAs through direct measurement of IgGs, IgMs, IgAs, IgEs and IgDs and their respective subclasses bound to the drug, or indirect measurement of the drug bound to ADAs. However, more industry experience is still needed before this technology is more widely adopted. The way of determining ADA positives and negatives and interpretation of ADA data also needs to be reconsidered and reevaluated due to the different techniques in measuring ADA.

PK assays

1. Hybrid LBA/LCMS assays provide structural characterization information during PK analysis and improve the understanding of the analyte and biology involved in addition to quantification. Method development is generally faster compared with traditional LBA, and comes with improved selectivity and less stringent reagent requirements. It is also generally less prone to matrix interference due to high specificity offered by the detection step. Scientists ideally should consider developing LBA and Hybrid methods in parallel with a clear stopping point to deciding which one to proceed with based on available results. .

2. The high selectivity of LCMS assays can provide greater discrimination between interferences, modifications, substitutions, and unique signature peptides than that afforded by traditional LBA approaches. When traditional LBA strategies to avoid ADA interferences in PK assays do not solve the issues, it is recommended to use LCMS or Hybrid LBA/LCMS assays with harsher conditions (e.g., alkaline, acid or guanidine combo treatments) for ADA dissociations than those typically used for LBA. A clear understanding of the analytical interferences and limitations is important.

3. Beads have been the most commonly used solid state method of choice for Hybrid LBA/LCMS assays. On-bead digestion is recommended because of its ease and lack of extra step when compared to off-bead digestion. The best approach to test clinical Hybrid LBA/LCMS assays for ADA measurement, especially where humans have not yet been dosed, is to use pre-existing ADAs, animal ADAs as positive controls. Selectivity should be tested in both pooled control and pre-dose patient matrices; the effects of specific ligands, such as soluble targets, at expected physiological concentrations upon drug quantitation in spiked QCs should be examined; as well as the impact of added ADAs of known high titers. Recovery and digestion efficiency should be examined during method validation. Recovery results, when patient pre-dose samples are tested naive and spiked at the LLOQ concentrations, provide information on selectivity

and matrix effects. Parallelism experiments should be performed during validation and potentially also with (pooled) incurred samples. The issue of reagent management for Hybrid LBA/LCMS assays requires considerations similar to LBA reagents including reagent stability, suitability, and lot-to-lot comparability.

4. Hybrid LBA/LCMS methods are increasingly used to identify, characterize and quantify biotherapeutic biotransformations using an integrated approach. Understanding biotransformations will help scientists determine the impact on efficacy, safety, PD, and PK, as well as help to correlate with *in vitro* monitoring. If biotransformations of biotherapeutics are present in quantities of more than 10% of the starting material, it is recommended to follow the MIST guidance regarding quantification recommendations. For quantities below 10%, the minor modified biotherapeutic is likely to have less impact on the overall activity, depending on how important the biotransformation is to the biotherapeutic.

Acknowledgements

The authors would like to acknowledge the US FDA, Europe EMA, UK MHRA, Austria AGES, Italy AIFA, Brazil ANVISA, Health Canada, Japan MHWL and WHO for supporting this workshop. S Richards (Sanofi), L Amaravadi (Sanofi/Genzyme), R Pillutla (Bristol-Myers Squibb), H Birnboeck (F. Hoffmann-La Roche Ltd.), F Garofolo (Angelini Pharma) for chairing the workshop and/or the white paper discussions. All the workshop attendees and members of the bioanalytical community who have sent comments and suggestions to complete this White Paper. W Garofolo, L Lu, X Wang, M Losauro, N Savoie, A Hernandez, K Kalaydjian and J. Conception for the assistance in the organization of the event. Future Science Group as a trusted partner.

Financial & competing interests disclosure

The authors have no relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript. This includes employment, consultancies, honoraria, stock ownership or options, expert testimony, grants or patents received or pending, or royalties.

No writing assistance was utilized in the production of this manuscript.

[1]Genentech, South San Francisco, CA, USA
[2]Merck & Co., Inc., Rahway, NJ, USA
[3]Angelini Pharma, Pomezia, RM, Italy
[4]Boehringer-Ingelheim, Ridgefield, CT, USA
[5]GlaxoSmithKline, King of Prussia, PA, USA
[6]Pfizer, Andover, MA, USA
[7]Caprion Biosciences, Montreal, QC, Canada
[8]Austria AGES, Wien, Austria
[9]Health Canada, Ottawa, ON, Canada
[10]WHO, Geneva, Switzerland
[11]Italy AIFA, Rome, Italy
[12]US FDA, Silver Spring, MD, USA
[13]Bioanalytical Laboratory Services a Division of LifeLabs LP, Toronto, ON, Canada
[14]Japan MHLW-NIHS, Tokyo, Japan
[15]Celerion, Lincoln, NE, USA
[16]Q2 Solutions, Ithaca, NY, USA
[17]Novartis, Basel, Switzerland
[18]Brazil ANVISA, Brasilia, Brazil
[19]Bristol-Myers Squibb, Princeton, NJ, USA
[20]UCB Biopharma, Braine L'Alleud, Belgium
[21]CFABS, Montreal, QC, Canada
[22]Formerly MedImmune, Gaithersburg, MD, USA; currently Incyte Corporation, Wilmington, DE, USA
[23]Covance, Indianapolis, IN, USA
[24]UK MHRA, London, UK
[25]Dutch MEB, Utrecht, The Netherlands
[26]Roche Pharma Research and Early Development, Roche Innovation Center, Basel, Switzerland

References

1 Savoie N, Booth BP, Bradley T *et al.* 2008 White Paper: the 2nd Calibration and Validation Group Workshop on recent issues in good laboratory practice bioanalysis. *Bioanalysis* 1(1), 19–30 (2009).

2 Savoie N, Garofolo F, van Amsterdam P *et al.* 2009 White Paper on recent issues in regulated bioanalysis from the 3rd Calibration and Validation Group Workshop. *Bioanalysis* 2(1), 53–68 (2010).

3 Savoie N, Garofolo F, van Amsterdam P *et al.* 2010 White Paper on recent issues in regulated bioanalysis and global harmonization of bioanalytical guidance. *Bioanalysis* 2(12), 1945–1960 (2010).

4 Garofolo F, Rocci M, Dumont I *et al.* 2011 White Paper on recent issues in bioanalysis and regulatory findings from audits and inspections. *Bioanalysis* 3(18), 2081–2096 (2011).

5 DeSilva B, Garofolo F, Rocci M *et al.* 2012 White Paper on recent issues in bioanalysis and alignment of multiple guidelines. *Bioanalysis* 4(18), 2213–2226 (2012).

6 Stevenson L, Rocci M, Garofolo F *et al.* 2013 White Paper on recent issues in bioanalysis: "hybrid" – the best of LBA & LC/MS. *Bioanalysis* 5(23), 2903–2918 (2013).

7 Fluhler E, Hayes R, Garofolo F *et al.* 2014 White Paper on recent issues in bioanalysis: a full immersion in bioanalysis (Part 1 – small molecules by LCMS). *Bioanalysis* 6(22), 3039–3049 (2014).

8 Dufield D, Neubert H, Garofolo F *et al.* 2014 White Paper on recent issues in bioanalysis: a full immersion in bioanalysis (Part 2 – hybrid LBA/LCMS, ELN & regulatory agencies' input). *Bioanalysis* 6(23), 3237–3249 (2014).

9 Stevenson L, Amaravadi L, Myler H *et al.* 2014 White Paper on recent issues in bioanalysis: a full immersion in bioanalysis (Part 3 – LBA and immunogenicity). *Bioanalysis* 6(24), 3355–3368 (2014).

10 Welink J, Fluhler E, Hughes N *et al.* 2015 White Paper on recent issues in bioanalysis: focus on new technologies and biomarkers (Part 1 – small molecules by LCMS). *Bioanalysis* 7(22), 2913–2925 (2015).

11 Ackermann B, Neubert H, Hughes N *et al.* 2015 White Paper on recent issues in bioanalysis: focus on new technologies and biomarkers (Part 2 – hybrid LBA/LCMS and input from regulatory agencies). *Bioanalysis* 7(23), 3019–3034 (2015).

12 Amaravadi L, Song A, Myler H *et al.* 2015 White Paper on recent issues in bioanalysis: focus on new technologies and biomarkers (Part 3 – LBA, biomarkers and immunogenicity). *Bioanalysis* 7(24), 3107–3124 (2015).

13 Bak MJ, Albrechtsen NW, Pedersen J, Holst JJ. Specificity and sensitivity of commercially available assays for glucagon-like peptide-1 (GLP-1): implications for GLP-1 measurements in clinical studies. *Diabetes Obes. Metab.* 16(11), 1155–1164 (2014).

14 Bak MJ, Albrechtsen NW, Pedersen J *et al.* Specificity and sensitivity of commercially available assays for glucagon and oxyntomodulin measurement in humans. *Eur. J. Endocrinol.* 170(4), 529–538 (2014).

15 Lee AY, Chappell DL, Bak MJ *et al.* Multiplexed quantification of proglucagon-derived peptides by immunoaffinity enrichment and tandem mass spectrometry after a meal tolerance test. *Clin. Chem.* 62(1), 227–235 (2016).

16 Chen LZ, Roos D, Philip E. Development of an immunocapture-LC/MS assay for simultaneous ADA isotyping and semi-quantitation. *J. Immunol. Res.* 2016, (2016).

17 US Department of Health and Human Services, US FDA, Center for Drug Evaluation and Research, Biologics Evaluation and Research. *Draft Guidance for Industry, Assay Development for Immunogenicity Testing of Therapeutic Proteins*. Rockville, MD, USA (2009).

18 US Department of Health and Human Services, US FDA, Center for Drug Evaluation and Research, Biologics Evaluation and Research. *Draft Guidance for Industry, Assay Development and Validation for Immunogenicity Testing of Therapeutic Proteins*. Rockville, MD, USA (2016).

19 Onami I, Ayabe M, Murao N, Ishigai M. A versatile method for protein-based antigen bioanalysis in non-clinical pharmacokinetics studies of a human monoclonal antibody drug by an immunoaffinity liquid chromatography–tandem mass spectrometry. *J. Chromatogr. A.* 1334, 64–71 (2014).

20 Wang SJ, Wu ST, Gokemeijer J *et al.* Attribution of the discrepancy between ELISA and LC–MS/MS assay results of a PEGylated scaffold protein in post-dose monkey plasma samples due to the presence of anti-drug antibodies. *Anal. Bioanal. Chem.* 402(3), 1229–1239 (2012).

21 Chirmule N, Jawa V, Meibohm B. Immunogenicity to therapeutic proteins: impact on PK/PD and efficacy. *AAPS J.* 14(2), 296–302 (2012).

22 Sailstad JM, Amaravadi L, Clements-Egan A *et al.* A White Paper – consensus and recommendations of a global harmonization team on assessing the impact of immunogenicity on pharmacokinetic measurements. *AAPS J.* 16(3), 488–98 (2014).

23 Xu K, Liu L, Saad OM *et al.* Characterization of intact antibody-drug conjugates from plasma/serum in vivo by affinity capture capillary liquid chromatography–mass spectrometry. *Anal. Biochem.* 412(1), 56–66 (2011).

24 Kaur S, Xu K, Saad O. Analysis of antibody drug conjugates by bead-based affinity capture and mass spectrometry: US8541178 (2013).

25 US FDA. *Guidance for Industry: Safety Testing of Drug Metabolites* (2008).

26 *ICH S3A Guideline: Note for Guidance on Toxicokinetics: the Assessment of Systemic Exposure in Toxicity Studies Questions and Answers* (2016).

27 US Department of Health and Human Services, US FDA, Center for Drug Evaluation and Research, Center for Veterinary Medicine. *Guidance for Industry, Bioanalytical Method Validation*. Rockville, MD, USA (2001).

28 European Medicines Agency. Committee for Medicinal Products for Human Use (CHMP). *Guideline on Bioanalytical Method Validation*. EMEA/CHMP/EWP/192217/2009. London, UK (2011).

29 Rocci M, Lowes S, Shoup R *et al.* 7th GCC Insights: incurred samples use; fit-for-purpose validation, solution stability, electronic laboratory notebook and hyperlipidemic matrix testing. *Bioanalysis* 6(20), 2713–2720 (2014).

30 US Department of Health and Human Services, US FDA, Center for Drug Evaluation and Research, Center for Veterinary Medicine. *Draft Guidance for Industry, Bioanalytical Method Validation*. Rockville, MD, USA (2013).

31 US Department of Health and Human Services, US FDA, Center for Drug Evaluation and Research. *Draft Guidance for Industry, Bioequivalence Studies with Pharmacokinetic Endpoints for Drugs Submitted Under an ANDA*. Rockville, MD, USA (2013).

32 US Department of Health and Human Services, US FDA, Center for Drug Evaluation and Research. Product-specific recommendations for generic drug development. http://www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm075207.htm

33 Nicholson R, Lowes S, Caturla MC *et al.* Conference report: 6th GCC focus on LBA: critical reagents, positive controls and reference standards; specificity for endogenous compounds; biomarkers; biosimilars. *Bioanalysis* 4(19), 2335–2342 (2012).

34 Hayes R, LeLacheur R, Dumont I *et al.* 9th GCC closed forum: CAPA in regulated bioanalysis; method robustness, biosimilars, preclinical method validation, endogenous biomarkers, whole blood stability, regulatory audit experiences and electronic laboratory notebooks. *Bioanalysis* 8(6), 487–495 (2016).

35 *General Chapter <1106> Immunogenicity Assays – Design and Validation of Immunoassays to Detect Anti-Drug Antibodies,*

US Pharmacopeia and the National Formulary (USP-NF) (2014).

36 Shankar G, Devanarayan V, Amaravadi L *et al.* Recommendations for the validation of immunoassays used for detection of host antibodies against biotechnology products. *J. Pharm. Biomed. Anal.* 48, 1267–1281 (2008).

37 *Title 21 – Food and Drugs, Chapter I – Food and Drug Administration Department of Health and Human Services Subchapter A – General: Part 58 – Good Laboratory Practice for Nonclinical Laboratory Studies*.

38 *OECD Series on Principles of Good Laboratory Practice and Compliance Monitoring, Number 1: Principles on Good Laboratory Practice* (1997).

39 *ICH Guideline for Good Clinical Practice E6(R1)* (1996).

40 European Medicines Agency. GCP Inspectors Working Group. *Draft Reflection Paper on Guidance for Laboratories that Perform the Analysis or Evaluation of Clinical Trial Samples*. EMA/INS/GCP/532137/2010 (2010).

41 World Health Organization. *Good Clinical Laboratory Practice (GCLP)* (2009).

42 World Health Organization. *WHO Expert Committee on Specifications for Pharmaceutical Preparations. Fifties Report Geneva, Annex 9 Guidance for Organizations Performing In Vivo Bioequivalence Studies* (2016). http://www.who.int/medicines/publications/pharmprep/WHO_TRS_996_annex09.pdf

43 World Health Organization. *WHO Expert Committee on Specifications for Pharmaceutical Preparations. Fifties Report Geneva, Annex 5 WHO Guidance on Good Data and Record Management Practices* (2016). http://www.who.int/medicines/publications/pharmprep/WHO_TRS_996_annex05.pdf

44 ANVISA. Technical Note Nº 07/2015 – Bioequivalence Studies for Vitamins, Brazilian Health Surveillance Agency, Brazil (2015). http://portal.anvisa.gov.br/documents/506392/0/CETER000.pdf/ff22cf14-c67a-4d37-99d0-88a74700e2a6

45 ANVISA. *Technical Note Nº 09/2015 – Pharmacokinetic Interaction for Fixed Dose Combinations and Clinical Trial Approval, Brazilian Health Surveillance Agency, Brazil* (2015). http://portal.anvisa.gov.br/documents/506392/0/NOTA±009±2015.pdf/7ef000d6-a39a-4b06-bbab-4c5ed21f4db4

46 ANVISA. *Technical Note Nº 02/2015 – PK Studies for Biological Products, Brazilian Health Surveillance Agency, Brazil* (2015). http://portal.anvisa.gov.br/documents/33836/2593824/Nota±de±esclarecimento±n%C2%BA±02±de±2015/c9423232-fdf0-4e68-8435-a7c663fda26a

47 ANVISA. *Resolution RDC no. 27, May 17, 2012. Validation of Bioanalytical Methods, Brazilian Health Surveillance Agency, Brazil* (2012). http://portal.anvisa.gov.br/legislacao#/visualizar/28854

48 Health Canada, Resource Management and Operations Directorate. *Draft Guidance Document – Disclosure of Confidential Business Information Under Paragraph 21.1(3)(c) of the Food and Drugs Act*. Ottawa, Ontario, Canada (2016).

49 Health Canada. *Notice: Proposed Policy on Bioequivalence Standards for Highly Variable Drug Products. (TBD)*.

50 Health Canada, Bureau of Policy, Science and International Programs Enquiries. *Post-notice of Compliance (NOC) Changes: Quality Document*. Ottawa, Ontario, Canada (2016).

51 Health Canada, Bureau of Pharmaceutical Sciences Therapeutic Products Directorate. *Notice: Clarification of Bioanalytical Method Validation Procedures* (2015).

52 Health Canada, Bureau of Pharmaceutical Sciences Therapeutic Products Directorate. *Addendum to Notice: Clarification of Bioanalytical Method Validation Procedures* (2016).

Ask the Experts: Biomarkers

Biomarkers are now increasingly used in the pharmaceutical industry for early proof-of-concept studies and, furthermore, much research is focusing on the development of diagnostic and predictive biomarkers for point-of-care biosensor systems in the clinical environment and beyond. With this increase in demand and research, how can the field be standardized and regulated, what are the key compliance and logistical challenges, and how can they be overcome?

Check out what some of the key leaders thought about biomarkers in the excerpt below:

What remains the greatest challenge in identifying biomarkers?

"The shear variety and number of biomarkers is the greatest challenge. Identifying important biomarkers for a specific disease still requires understanding of many biological mechanisms and pathways. These include small molecules that can be determined following metabolomics investigational models, proteins and peptides that can be determined following proteomic investigational models, lipids following lipidomics, etc.

In addition, having the appropriate instrumentation, data management and reference standards is critical. These can be routine though esoteric.

Finally, another great challenge that remains is the complexity of data analysis – e.g., historical data for patients, confidence in the correlation between the biomarker and disease/outcome over time as both analytical technology and treatments change over time, and other complexities around data analysis."

Patrick Bennett (PPD)

"One of most significant challenges we face at Biogen in the biomarker field is around identifying biomarkers related to central nervous system (CNS) indications and therapeutics, particularly biomarkers that measure target engagement or a pharmacodynamic response within the CNS.

Given that the CNS is an anatomical location not feasibly accessible in clinical studies, we rely on sampling cerebral spinal fluid (CSF) and/or neuroimaging to evaluate neural-specific biomarkers. These approaches are particularly challenging due to complex technical considerations and the evolving science behind the biomarkers of interest.

With respect to CSF, the levels of proteins are incredibly low and not all CNS-specific pathways/mechanisms are mirrored in the CSF. Neuroimaging techniques have great promise in understanding CNS disease biology, but the lengthy development time for identifying and optimizing novel imaging agents is not always aligned with key decision points of the drug development program."

Devagi Mehta, Chris Stebbins, Danielle Graham and Lauren Stevenson (Biogen)

"I think the challenge is not necessarily in identifying biomarkers but rather in identifying the RIGHT biomarkers. Quite often we see new biomarkers come along but without sufficient physiological information as to their behaviour or variation – both in normal and disease states.

Furthermore, with the advent of the 'ultra-sensitive' platforms, I think that new biomarkers are sometimes being put forward for investigation prematurely as often there are technological (method) issues still to be answered before we can reliably investigate the changes and levels of the biomarker. Often the question is how can we prove that we're measuring what we think we are? And this is sometimes not substantially addressed."

John Allinson (LGC Group)

Special Focus Issue: Bioanalysis of Biomarkers

Bioanalysis published two Special Focus issues on Bioanalysis of Biomarkers in 2016. Biomarkers come in many forms with their complementary technologies, but are central to current drug development. The themed issues contain a collection of articles on biomarkers in drug development, technology assessments and scientific applications.

The two Special Focus issues have been Guest Edited by Mark E. Arnold (Covance Laboratories, Inc., USA), Hendrik Neubert (Pfizer, Inc., USA), Lauren F Stevenson (Biogen, USA) and Fabio Garofolo (Angelini Pharma, Italy). The Guest Editors provided a scene-setting foreword in the first issue, which explains,

"Our intention going in was to highlight the variety of challenges that measurement of biomarkers poses and the breadth of technological solutions being applied. In reaching out to peers and the scientific community, we were delighted to see the extensive response from not only those working with technology typically seen in the pages of this journal … but also from those using an array of other technologies…".

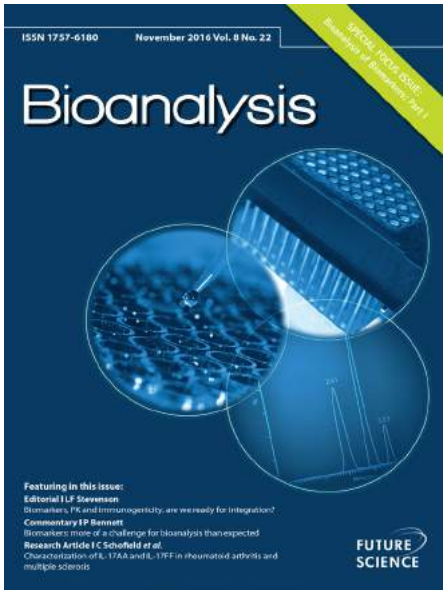
The Special Focus issues contain a variety of content including Editorial, Review and Research Articles looking at the challenges in measurement of biomarkers, and the range of technological solutions (both tested and new) available to overcome these challenges. The potential use of alternative analytes as biomarkers, and associated bioanalytical advancements are also touched upon. The '2016 White Paper in Bioanalysis' recommendations from the 10th WRIB, focusing on Biomarker Assay Validation (BAV), are also featured in these issues.

"I am delighted to have worked with experts from the industry to produce this timely and comprehensive two-part Special Focus Issue discussing the challenges, advancements and next steps in biomarker bioanalysis" stated Sankeetha Nadarajah, Editor of *Bioanalysis*.

Both issues can be viewed here:

Part I of the Special Focus Issue on Bioanalysis of Biomarkers

Part II of the Special Focus Issue on Bioanalysis of Biomarkers



HRMS

Special Focus Issue: HRMS in DPMK

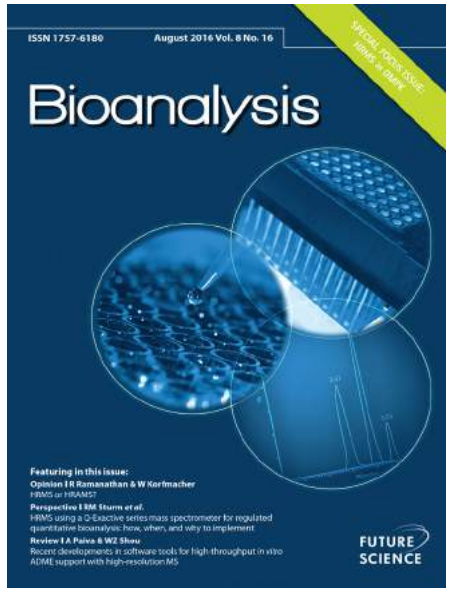
Bioanalysis, published a Special Focus issue on HRMS in DMPK, which is a follow-up to the two Special Focus issues of *Bioanalysis* on HRMS that were published in 2012 and 2013. The issue contains a collection of articles that provide an update on how HRMS is being used by scientists engaged in new drug discovery and drug development.

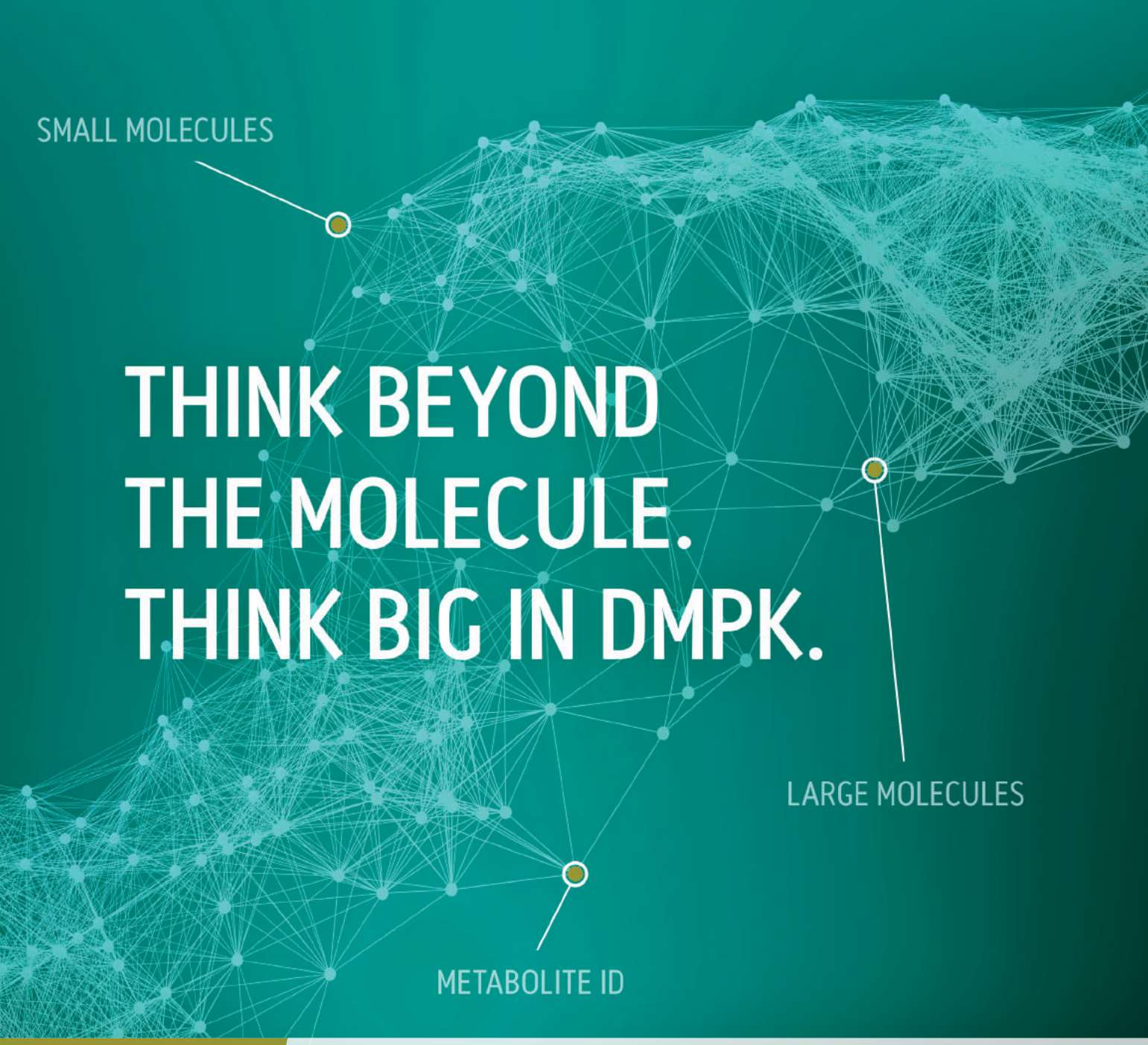
Walter Korfmacher, one of the Guest Editors from Sanofi commented, *"This special issue on HRMS will be of interest to those bioanalytical scientists currently using HRMS as well as to those who are planning to use HRMS systems in the near future."*

The special issue includes a range of articles including an opinion article discussing the rationales between the terminologies HRMS and HRAMS, and a Commentary on the outsource-ability of HRMS-based DMPK assays. It also includes several Research Articles on expanding the applications of HRMS, a Perspective article looking at the use of a Q-Exactive series mass spectrometer for regulated quantitative bioanalysis, and a Review on recent developments in software and automation tools for high-throughput discovery *in vitro* ADME assays based on HRMS.

"Today's HRMS systems are starting to reach new heights and are capable of providing triple quadrupole like quantitative sensitivity, while providing options to gain qualitative information. However, are HRMS systems ready to replace triple quadrupoles? Can CRO partners accommodate HRMS-based DMPK assays? To learn more about HRMS in DMPK applications and to understand how the technology has evolved, pharmaceutical, CRO, regulatory agency, and chemical industry bioanalytical scientist should take advantage of the special issue on HRMS" said Ragu Ramanathan, Guest Editor from Pfizer.

View the full issue – www.future-science.com/toc/bio/8/16





THINK BEYOND THE MOLECULE. THINK BIG IN DMPK.

INNOVATIVE DMPK SOLUTIONS

Today's bioanalysts don't have it easy. Whether optimizing methods for small molecules, making the transition from small to large bioanalysis or ensuring a metabolite isn't missed, the demands are greater than ever. Drug candidates are becoming increasingly complex, and you need holistic solutions that address each step in the workflow. To think beyond the size of the molecule, visit DMPK.waters.com.

Waters

THE SCIENCE OF WHAT'S POSSIBLE.®

PHARMACEUTICAL ▪ HEALTH SCIENCES ▪ FOOD ▪ ENVIRONMENTAL ▪ CHEMICAL MATERIALS

Perspective
Special Focus Issue – HRMS in DMPK

Bioanalysis

HRMS using a Q-Exactive series mass spectrometer for regulated quantitative bioanalysis: how, when, and why to implement

High-resolution MS (HRMS) has seen an uptake in use for discovery qual/quan workflows, however, its utilization in late discovery/development has been slow. Past reports comparing HRMS to triple quadrupole (QQQ) instrumentation to date have indicated that HRMS instruments are capable of producing data acceptable for regulated bioanalysis, however lack the sensitivity required for sub ng/ml LLOQ assays. Recent advances in HRMS instrumentation have closed the sensitivity gap with QQQ and have even provided improved selectivity and sensitivity over QQQ SRM assays. Herein, the authors will describe how, when, and why HRMS (specifically Q-Exactive series mass spectrometers) should be considered for implementation in regulated quantitative bioanalysis assays.

Robert M Sturm*[,1],
Barry R Jones[1],
Daniel E Mulvana[1]
& Steve Lowes[1]
[1]Q[2] Solutions, 19 Brown Road, Ithaca, NY 14850, USA
*Author for correspondence:
Tel.: +1 607 216 6523
Fax: +1 607 266 0749
sturmrob@gmail.com

First draft submitted: 30 March 2016; Accepted for publication: 21 June 2016; Published online: 27 July 2016

Keywords: development • high-resolution MS • HRAM • HRMS • mass extraction window • orbitrap • Q-Exactive

In the early 2010s there was great interest regarding the potential use of high-resolution MS (HRMS, HRAM, HR-MS, HR/MS) in the drug discovery and development workspaces. In 2011 Ramanathan *et al.* [1] and Korfmacher [2] posed the following question. Is it time for a paradigm shift in drug discovery bioanalysis from SRM using triple quadrupole instruments (QQQ) to HRMS analysis? The key term here is drug discovery. During drug discovery HRMS instruments are commonly operated in full-scan MS1 mode (Qual/Quan approach) where both quantitative (drug concentration) and qualitative (drug metabolites, post-translational modifications, degradants, biomarkers, and formulation materials) information can be collected in a single data acquisition. At the time of data acquisition the analyst may have certain targets for which they are seeking information. However, data from full-scan information may be mined at a later time for additional supporting information to help determine the drug's performance. Because data were collected in full-scan mode on an HRMS instrument, the existing raw data files may be processed to extract the information rather than requiring a new sample preparation and LC–MS analysis. From this description, it is clear that HRMS instruments operated in Qual/Quan mode have a definite efficiency advantage over triple quadrupole SRM workflows. There are several descriptions in the literature describing the use of HRMS quadrupole time-of-flight (Q-TOF) and orbitrap-based instruments for this type of application [3–8].

In 2013, Fung *et al.* [9] and Huang *et al.* [10] both published perspective articles in *Bioanalysis* describing the current use of HRMS in bioanalysis. At that time it was well accepted that HRMS instruments could produce accurate and precise quantitative data. There are examples of HRMS instruments providing excellent quantitative performance for liquid biofluid [11–25] and dried matrix

FUTURE SCIENCE part of fsg

spot samples [26–29]. However, there was some question regarding if/when this approach would become more mainstay in drug development. The authors prudently detailed challenges related to mainstream use of HRMS in the development arena. The challenges are as follows:

- Qualitative data are less important at the development stage so justifying large data files is more difficult [9].
- HRMS instruments lack the sensitivity required for low pg/ml LLOQ and triple quadrupole instruments (QQQ) remain the most sensitive platform for demanding assays [9,10].
- Data storage and retrieval systems need to be well thought out regarding the handling of the large data files generated by HRMS. Can the current information technology infrastructure within an organization adapt the data management capacity generated by HRMS instruments? [9,10].
- Data processing between QQQ and HRMS instruments are fundamentally different [9,10]. For QQQ instruments the extracted ion chromatograms (XIC) are defined by the data acquisition parameters. The analyst simply uses the data processing software to extract the acquired SRM transition information from the dataset. For HRMS instruments, the instrument data acquisition parameters do not precisely define what ions are used for quantification. Rather, defined accurate masses and mass extraction windows (MEW) of the analytes and internal standards are used postacquisition for software processing to define the construction of the integrated XICs. Can the same MEW generate uniform selectivity throughout an entire drug program or will there be a point where a change in MEW will be necessary? If a change in MEW is deemed necessary, such changes would require partial validation of the method [9].
- HRMS full scan data are rich in information. This is why there has been great success for the platform in discovery. During GLP and clinical studies it may be tempting to mine the data for other metabolites/biomarkers that were not included in the validated method. However, there could be regulatory compliance issues since this is beyond the scope of the validated method and patient consent may be lacking [9]. In addition, PK/TK-validated methods for GLP and clinical studies are designed specifically to optimize sample extraction procedures for selectivity/sensitivity of the target analyte. The qualitative data in these sample sets may not be representative of metabolite/biomarker distributions and therefore unreliable for secondary investigations [10].

There are several HRMS platforms (Q-TOF, TOF, Fourier transform ion cyclotron resonance and orbitrap-based mass analyzers) available to bioanalysts. The general challenges and best practices for HRMS implementation discussed herein can likely be applied to the other common HRMS platforms listed above, however specific method development strategies outlined below will focus on data acquisition modes specific to Q-Exactive series mass spectrometers. The aim of this article is to provide a foundation for successful implementation of Q-Exactive series-based HRMS platforms into regulated quantitative bioanalysis laboratories. Specifically, instrument validation considerations, method development/validation strategies and examples where HRMS address QQQ limitations will be referenced and presented. Despite this article focusing on HRMS implementation in a drug development setting, the general strategies can be applied to other regulated industries including environmental analysis, food safety/quality, forensic science, and sports-doping testing.

How is HRMS implemented for quantitative bioanalysis in a regulated environment?

As discussed above, HRMS has been well accepted in the discovery arena for qual/quan workflows. In 2013 the outlook for HRMS in late discovery/development was cloudy [9,10]. The selectivity gained through HRMS was impressive but it was still uncertain if HRMS instruments had the wherewithal to be a serious contender in regulated quantitative bioanalysis, an area dominated by QQQ instrumentation. This article will cover how HRMS instrumentation, specifically Thermo Q-Exactive series mass spectrometers, can be implemented in a regulated bioanalysis setting. The authors will share their experiences with the technology and where they foresee HRMS can greatly improve the quality of assays that can be developed to support a promising drug candidate's journey toward approval.

Specific consideration should be given to the operational differences HRMS presents over QQQ quantitation. Regulatory auditors and inspectors have become comfortable with LC–MS data but this confidence is derived from decades of QQQ-generated chromatograms and integrated peak areas. As described above, and further elaborated below, the process of generating a chromatogram from an HRMS instrument is a postacquisition activity and thus redefines what raw data are. This is a very important concept that requires

education, scientific defense, and control measures. To date, regulatory authorities have expressed support for HRMS use in bioanalytical application but it is incumbent on analysts to be able to explain, justify, and reconstruct experimental parameters that fundamentally influence the data being reported.

Considerations for HRMS software validation

Instrumentation that uses software to create, modify, maintain, archive, retrieve or transmit nonclinical and clinical data that may be used in a regulatory filing should be thoroughly vetted prior to implementation in a bioanalytical laboratory (instrument validation). One of the prerequisites for an instrument to be considered for regulated bioanalysis is that it possess software capable of meeting the recommendations set forth by regulators to ensure the integrity and traceability of data acquired [30]. Therefore, it is advantageous to work with an instrument company that has an established history of manufacturing mass spectrometers and software for regulated bioanalysis. It should be noted that these software packages may have been created with QQQ instrumentation in mind. Therefore, it is the responsibility of the bioanalytical laboratory to ensure that the software acquires data under the secure data system as intended, tracks users' actions accurately, and maintains data integrity during storage.

The main difference between QQQ and HRMS software rests in how the data are acquired and processed. For QQQ instruments, the acquisition parameters define the scan filters (SRM transitions) that are used to generate extracted ion chromatograms for an analytical run. For HRMS, the acquisition parameters do not define the processing parameters. For acquisition, analysts define the mass range and experimental mode (full-scan MS1 or MS/MS) in which the data will be collected but do not de-facto limit how the data can be processed. Quantitative information for HRMS instruments is extracted postacquisition using defined analyte and internal standard accurate masses and MEW. Other than scheduled calibration and resolution settings, the user does not have control over what mass is recorded by the instrument, but does have input regarding how narrow of a mass range will be used to generate XICs for quantitative analysis. The HRMS instrument generates an in-run accurate mass that will not always be the exact same as previously observed or predicted for the target compound. The MEW is used to factor in the uncertainty in the target compound's accurate mass across each scan and enables the generation of representative XICs. For example, if the target ion has a predicted mass-to-charge ratio (*m/z*) of 762.394 the observed apex of the ion peak may fluctuate from *m/z* 762.390 to *m/z* 762.399. The MEW for this compound should factor in this mass uncertainty so the XIC produced for this ion in a quantitative assay should contain peak information from approximately *m/z* 762.388 to *m/z* 762.400 (MEW = ±0.006 *m/z* units). It is important to ensure that the MEW input into the processing software is accurate as the MEW affects accuracy and selectivity of an assay significantly [19]. The software's audit trail can help confirm that the same MEW is used from run to run for a given study.

Considerations for instrument calibration frequency

In our experience, the Q-Exactive series mass spectrometer has excellent day to day mass accuracy stability. Weekly mass accuracy calibration has been recommended by Thermo [31]. When we first started using a Q-Exactive series mass spectrometer for targeted quantification we conservatively calibrated mass accuracy prior to each run (e.g., daily calibration). As time went on, the HRMS scientific team began to monitor the mass accuracy over extended periods after mass calibration. It was determined that a minimum calibration of every 3 days could be implemented operationally without any negative effects on data integrity. This time frame is also convenient as it enables several runs to be queued up over the weekend without interruption. It is up to the organization's scientific team to evaluate the mass accuracy performance of any HRMS instrument prior to assigning a calibration schedule. Once defined, the calibration schedule should be captured in the standard operation procedure for the instrument platform.

Considerations for determining the optimum MEW for a method

As mentioned above, the observed accurate mass of an ion can vary slightly from scan to scan. In the authors' experience, on a Q-Exactive series mass spectrometer the observed *m/z* of an ion that has a charge state of 1 and a low molecular weight (<500 Da) displays a very static accurate mass and rarely varies more than ±0.001 *m/z* unit. For multiply charged ions with *m/z* values exceeding *m/z* 500 the authors have observed greater mass accuracy fluctuations ranging from ±0.004 *m/z* unit to ±0.008 *m/z* unit. In general, the mass accuracy for ions used for quantification do not vary greater than 5 ppm across an assay's concentration dynamic range (typically three orders of magnitude).

Selection of an MEW that is both selective and quantitative across an assay's dynamic range is critical. The MEW must be small enough to provide selectivity from background ions and large enough to provide coverage of the target ion's mass accuracy fluctuation

over the concentration range and data collection duration [19,32]. Conceptually, there are two ways to define the accurate mass for a given quantitative method. In both approaches, the MEW tolerance should remain static. In the first approach, the analytical method document would contain the exact mass for the analyte and internal standard calculated *in silico* using each compound's chemical formula and apply that to all runs. Since the observed accurate mass will vary from the exact mass from run-to-run the defined MEW will be asymmetric around the analyte and internal standard accurate masses. This approach will necessitate the analyst to specify that a larger MEW be used for validation and sample analysis so that the target ions can be quantified accurately and reproducibly. A drawback to this approach is that the necessity of a larger MEW may introduce background noise affecting the ruggedness of the method's selectivity [33]. In a 2012 Bioanalysis article, Ramagiri *et al.* eloquently presented the processing challenges associated with HRMS data [19]. Using a Q-TOF, the authors presented the negative impact that a wider MEW can have on signal-to-noise (S/N) at different analyte concentration levels.

The second approach specifies that the defined MEW will be applied to the in-run accurate mass determined by averaging the scans present across the full-width at half-maximum of a mid-level standard XIC. This approach results in a symmetric MEW used for quantification and enables the analyst to develop a method with a smaller MEW resulting in maximum selectivity. In practice, the process of extracting the accurate masses for the quantitative ions is very straightforward. The analytical method document would specify the analyte and internal standard accurate masses observed in method development/validation and the mass tolerance (e.g., MEW) within which the target ions should be observed. The analyst identifies the ions present in the averaged mass spectrum within the defined mass tolerance and then integrates the chromatographic peaks within the run using those values. Figure 1 displays the raw data output for a mid-concentration standard curve point using HRMS MS/MS (PRM) acquisition method on a Q-Exactive series mass spectrometer. The analyst can clearly identify the accurate masses for the predefined analyte (y_5, y_6, y_7) and internal standard (y_5 and y_7) ions that will be used to generate XICs since they are very abundant at this concentration level. The second approach maximizes selectivity for the HRMS approach.

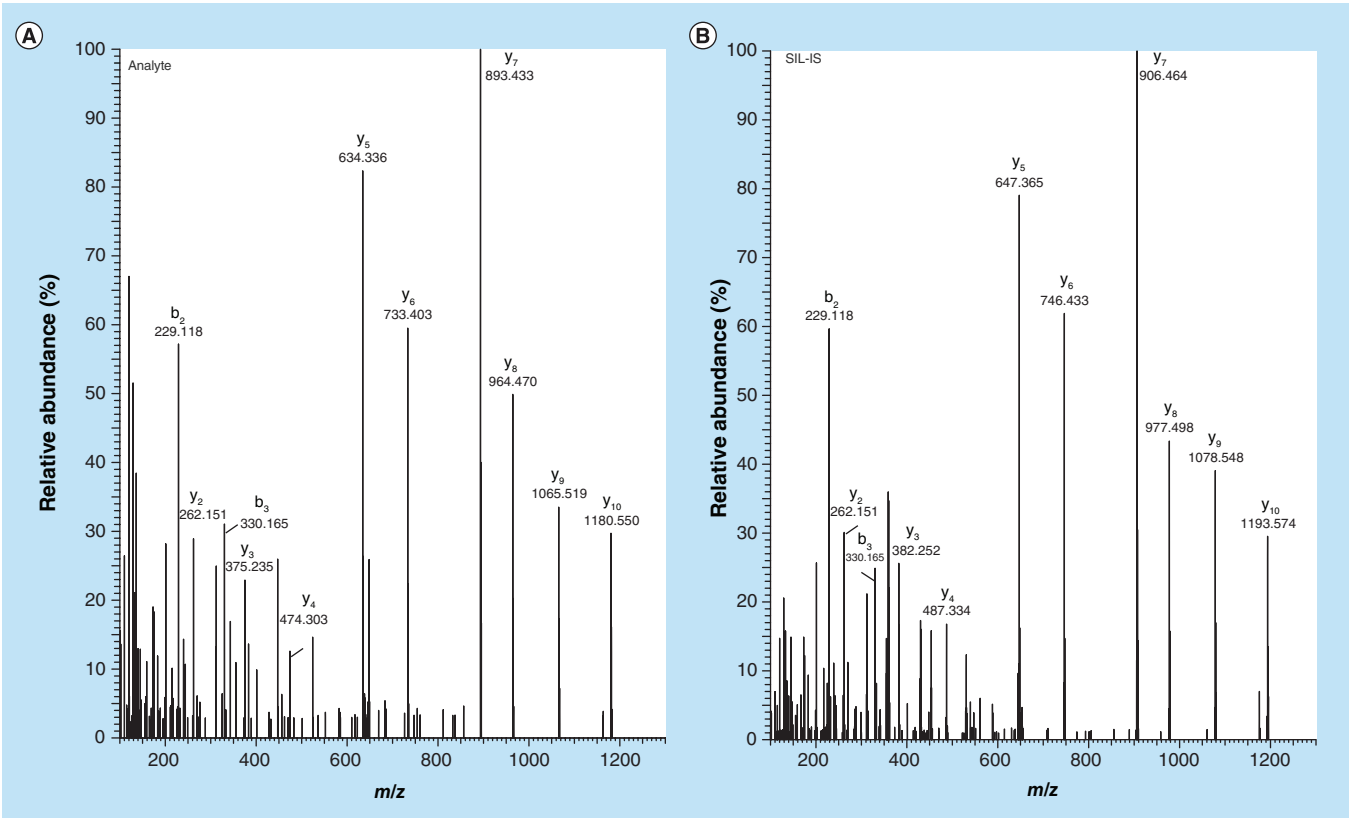


Figure 1. Example of raw data output for a mid-concentration standard curve sample using high-resolution MS/MS acquisition on a Q-Exactive series mass spectrometer. The analyst can clearly pick out the accurate masses for the predefined analyte (y_5, y_6, y_7) and internal standard (y_5 and y_7) ions that will be used to generate XICs.

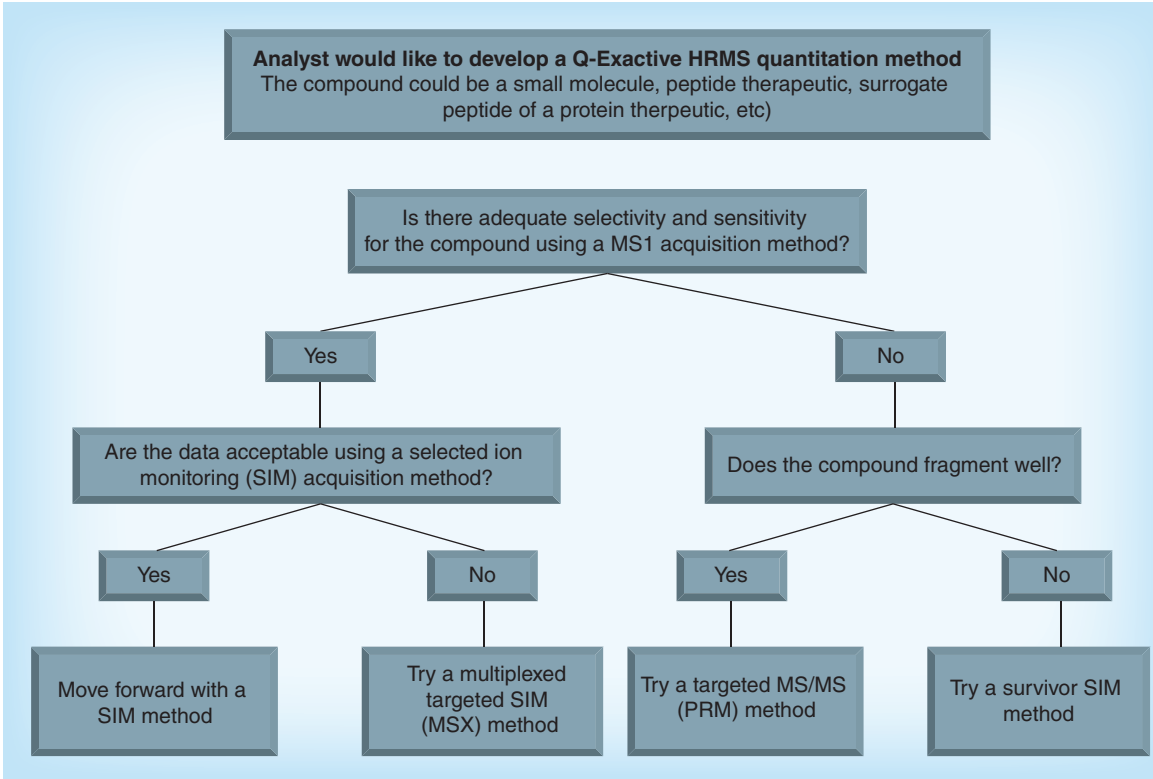


Figure 2. Method development decision tree for developing high-resolution MS acquisition methods for a Q-Exactive series mass spectrometer.

Acquisition mode decision tree when creating a method for a Q-Exactive series mass spectrometer (method development)

Gallien *et al.* detailed well the capabilities and performance of a Q-Exactive series mass spectrometer for quantitative applications [34]. The authors' suggest that readers review this article if unfamiliar with the quantitative capabilities of a Q-Exactive series mass spectrometer. The most common HRMS quantification mode is full scan MS1. This is the mode often utilized in qual/quan workflows and is likely the mode traditionally used when comparing overall sensitivity performance between HRMS and QQQ instruments [9,14,18,20,27]. In late stage discovery/development, the main focus is to quantitate drugs and metabolites in biological matrices with excellent selectivity and sensitivity. If HRMS using a Q-Exactive series mass spectrometer is required, the best sensitivity for an MS1 acquisition will be achieved by minimizing the mass range of ions that are able to enter the C-trap [24]. This observation is attributed to the fixed capacity of ions that a C-trap can hold for a given automatic gain control setting. Once the C-trap fills to the specified automatic gain control, no additional ions (background or analyte) are allowed to enter the C-trap until the following scan acquisition. If the quadrupole mass filter is not operated to limit population of ions that enter the C-trap, the authors have observed that C-trap fill times will commonly be less than 10 ms and the fraction of ions derived from the analyte and internal standard are suboptimal. For maximum sensitivity, the C-trap should be filled for as long as possible (required resolution for the orbitrap acquisition should limit C-trap fill time, e.g., maximum injection time) with the purest population of analyte and internal standard ions. This is why during the authors rarely operate in full scan mode during method development for quantitative assays on an HRMS instrument.

Figure 2 displays the decision tree matrix that method development scientists employ when developing an HRMS method for a Q-Exactive series mass spectrometer. For Q-TOF users, please see the Ramagiri *et al.* publication where a method development decision tree for Q-TOF targeted protein/peptide quantification is presented [19]. If there is adequate selectivity in an MS1 level acquisition, the simplest acquisition mode would be to measure the analyte and internal standard using a selected ion monitoring (SIM) acquisition method. SIM acquisition utilizes quadrupole selectivity to bracket the analyte and internal standard precursor ions so they both travel through the quadrupole and enter into the C-trap together. After the ions are collected in the C-trap, they are injected into the orbitrap mass analyzer for

analysis. If there are background ions present between the analyte and internal standard ion envelopes which dramatically affect the purity of the ion population within the C-trap resulting in non-optimal C-trap fill times and poor SIM data, a multiplexed targeted SIM acquistion may be applied to exclude the background ions from the C-trap. During a multiplexed targeted SIM acquisition, the quadrupole rasters between analyte and the internal standard ion envelopes such that analyte ions and internal standard ions enter the C-trap separately. By operating the quadrupole in this manner, background ions in the mass region between the analyte and internal standard ion envelopes are excluded from the C-trap fill resulting in a purer population of ions of interest trapped in the C-trap. After the analyte and internal standard ions are trapped together in the C-trap they are injected together into the orbitrap mass analyzer for mass analysis (same as SIM mode). If selectivity for the analyte and internal standard are poor in MS1 acquired data even at a resolving power of 140,000 at *m/z* 200, MS/MS (PRM [35]) data collection should be considered. This acquisition mode normally results in superior sensitivities when mass analysis is limited by selectivity issues in either MS1 HRMS analysis or QQQ SRM (an example will be discussed later). Lastly, a pseudo-SRM acquisition can be used to increase selectivity for ions that have poor MS1 selectivity and do not fragment well by collision induced dissociation [36]. On a Q-Exactive series mass spectrometer, this pseudo-SRM approach has been termed 'Survivor-SIM' [37]. In this approach, collision energy is applied to break up the background ion precursors that crowd the MS1 precursor ion envelope of the analyte. The predefined accurate mass ions of the precursor analyte are used for quantification in this MS/MS approach. An example of the Survivor-SIM scan mode will be described later in this article.

Should all the ions within the ion envelope be summed together or should only select ions be used?

The common practice in the authors' laboratory for HRMS quantification is to use the quadrupole on a Q-Exactive series mass spectrometer to select one charge state envelope of the analyte to monitor for quantification. Operation of the HRMS instrument in this manner does not allow a user to sum all the ions from the multiple charge state forms of the analyte as is sometimes done in the qual/quan approach [9]. If only one charge state of the analyte and internal standard are monitored, should all of the ions present within each species' ion envelope be summed? In the authors' experience and others' [19], the answer is no. During method development each ion within the analyte and the internal standard ion envelope should be interrogated over several runs to determine the best ions to generate an XIC from. As described previously [19], although the HRMS approach 'promises' to show an analyst every ion present in a mass spectrum, matrix background interferences can affect the purity of a given HRMS ion peak. This was true for the z = +4 charge state of hepcidin-25 when analyzed using a Q-TOF mass spectrometer [19] and can even be present on a mass analyzer capable of higher resolving power like a Q-Exactive series mass spectrometer. In addition to matrix interferences, the authors have also observed varying mass accuracy fluctuations from scan-to-scan and across the dynamic range of an assay from ions within the same charge ion envelope, particularly for multiply charged species. Inclusion of ions that have poor mass accuracy performance can severely hinder the ruggedness of an assay. Currently, there are no automated solutions to the evaluation of different ion combinations and MEWs for quantitative assays. It would be a great boon to the method development process if vendors developed software that could iteratively test combinations of different ions for quantitative performance. However, once the analyst has thoroughly evaluated the ions in an envelope, the selectivity and performance of an HRMS method is outstanding.

Defining system suitability requirements: if there is no noise how useful is S/N criteria?

Figure 3 displays XICs for a matrix zero sample blank and LLOQ for a small molecule HRMS MS/MS (PRM) assay developed at the authors' laboratory [17]. The MEW for the XICs displayed was ±0.003 *m/z* units. For this assay no observable gain in signal was observed by increasing the MEW. The reader's attention should be drawn to the absence of ion current observed in the matrix zero sample blank. The selectivity is outstanding, especially for an assay that was experiencing significant trouble meeting LLOQ requirements on a QQQ instrument operated in SRM mode due to background noise (Figure 4). The lack of ion current observed in the matrix zero sample blank is not unique to this assay, and is typical of all the HRMS MS/MS (PRM) methods developed to date at the authors' laboratory. This presents a challenge for the analyst, as the lack of background response complicates evaluation of S/N as an estimation of assay performance. In common analytical methods where background noise is present, the limits of detection and quantitation are estimated by evaluation of S/N ratios. However, it is difficult to apply this concept to HRMS data when no noise is observed. Rather than

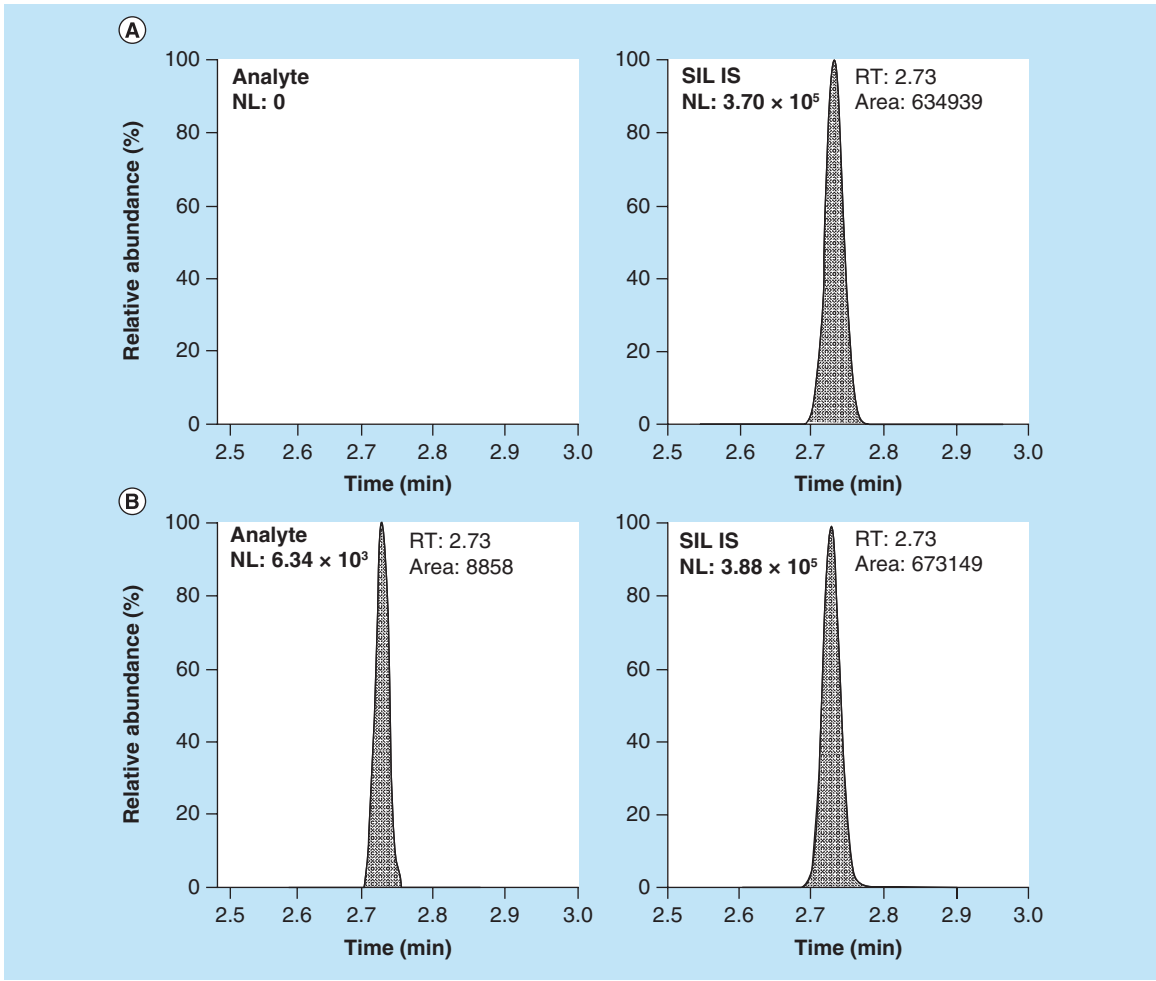


Figure 3. Example XICs for a (A) matrix zero sample blank and (B) LLOQ for a small molecule HRMS MS/MS (PRM) assay. Note that there is no observable noise in the analyte channel. How can S/N criteria be used to determine LOD and LLOQ?
HRMS: High-resolution MS; S/N: Signal-to-noise.

defining the LLOQ based on S/N, we believe the LLOQ should first be defined by the needs of the assay and then demonstrated quantitatively to be measurable within acceptable accuracy and precision criteria.

Prior to any validation experiment or sample analysis batch, a system suitability consisting minimally of a low concentration standard is acquired to assess the condition of the experimental extraction and LC–MS system. Common system suitability criteria are based around S/N at a specified concentration. As S/N is often not a useful metric for HRMS data, it is critical to track the area and spectral intensity of system suitability samples during method development and validation so appropriate criteria can be assigned. In practice, the authors have found that using either a minimal peak area or spectral intensity threshold for the system suitability sample can confirm the appropriateness of the extraction and LC–MS equipment for the subsequent analytical run.

Data file size minimization

The data file size of HRMS raw files was an important challenge topic in previous HRMS perspective articles [10,14]. It is true that SRM data files are smaller than HRMS data files. It is also true that SRM data are much simpler than HRMS data. The data output for SRM is the number of ions hitting a detector during a specific duration for the defined precursor to product transition. The data output for an HRMS experiment is a high-quality HRMS mass spectrum for each scan. An analyst can see exactly what they are measuring rather than assuming that the QQQ mass spectrometer is transferring only the specified precursor and product ions through Q1 and Q3. With quality comes data size. The size of a typical 5 min SRM raw file is approximately 0.100 Mb (internal data). The size of a 5 min HRMS raw file can vary depending on acquisition mode and time. For a 5 min acquisition in full scan mode (*m/z* 200 to *m/z* 1500) the file size can be approximately 60 Mb

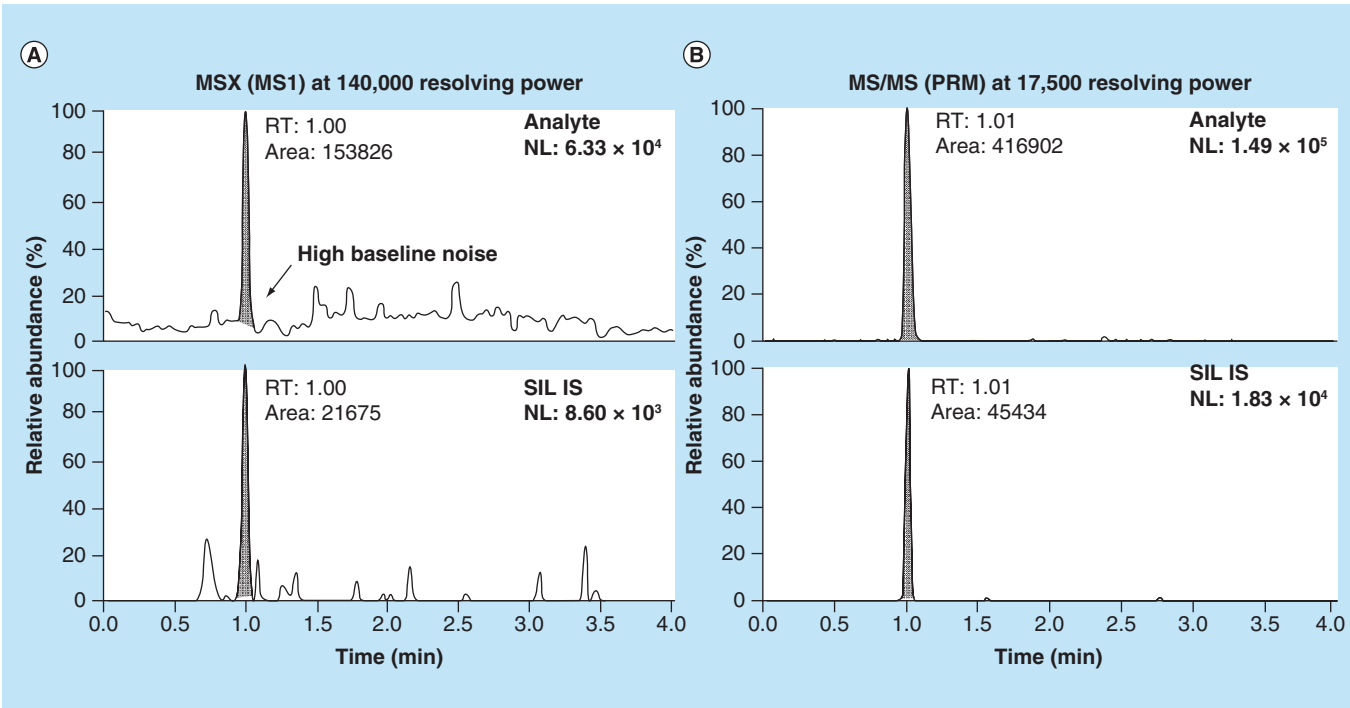


Figure 4. Example that selectivity is not always achievable in MS1 mode on a high-resolution MS instrument. (A) MS1 XIC for a small molecule (MEW = 0.003 *m/z* units) using 140,000 resolving power (*m/z* 200) at the ULOQ. **(B)** MS/MS acquisition at 17,500 resolving power (*m/z* 200) at the ULOQ. Note the SRM transition used for **(B)** exhibited high background noise on a QQQ supporting the use of HRMS MS/MS acquisition for this analyte.

and in SIM mode (6 *m/z* unit window) approximately 20 Mb (internal data). When the SIM mode acquisition time is reduced to 0.8 min (±0.4 min around the XIC apex for the analyte/internal standard determined for system suitability samples) the file size can be reduced to approximately 3 Mb (internal data). For a 96-well block run, the SRM will produce approximately 10 Mb of data and the HRMS instrument operated for 0.8 min in SIM mode will produce approximately 300 Mb of data (30-times greater than SRM). Data file size can be minimized by limiting mass range and the amount of time that data are collected during a sample acquisition. The total data size can be manageable with proper IT infrastructure in place. However, vendors should continue to strive to minimize the data footprint that HRMS data can command. In the case of Q-TOF technology, it has been reported that analysts and vendors have been able to minimize quantitative QTOF data file size by operating in TOF-MRM mode (0.100 Mb per raw file) which is a major feat for HRMS technology [38].

Analyst education

Considering the advantages HRMS affords to early stage drug discovery with combined quantitative and qualitative aspects, and the sensitivity improvements to targeted quantification achieved via improved selectivity, why has not adoption of HRMS in regulated bioanalysis been more widespread? One point of resistance to this paradigm shift is the experience gap with skilled users of the technology. As QQQ mass spectrometers have dominated the instrument fleets of small molecule bioanalytical labs, the nuances specific to SRM measurements have dominated our rhetoric, our infrastructure, and our operational and scientific mindsets. Key to bridging this gap is the effective training of our analysts in the HRMS technologies and techniques. The nucleus of any effort to introduce advanced and disruptive approaches in a contract research laboratory rests within the method development/validation team. It is here where investments in skilled HRMS scientist(s) should be made along with the associated instrumentation acquisitions. Collaborative efforts with sponsors to apply HRMS to bioanalytical challenges allows for the hands-on experience necessary to standardize internal practices and set operating procedures. Once utility is demonstrated and methods are developed, transfer of knowledge to production staff is accomplished through established standard operation procedures and hands-on training via the method development function.

When & why is HRMS implemented for quantitative bioanalysis in a regulated environment?

Selectivity challenged assays

High-resolution accurate MS instruments will not replace a fleet of QQQ instruments for targeted quan-

tification. Quantification of small molecules by QQQ instruments operated in SRM mode has been the gold standard for LC–MS quantitative approaches for the last 20 years [39–42]. Over this time period advances in drug design have increased the specificity of therapeutics to their targets resulting in decreased dosing amounts and subsequently the necessity for more sensitive modes of detection. For QQQ systems, an assay's sensitivity is often limited by the selectivity of the SRM transition compared with the background ion current. Solutions developed over the years to address selectivity-challenged assays include the use of ion mobility to separate ions based on charge, *m/z*, and/or collisional cross-section [36,43,44]; the employment of a second fragmentation stage for product ion fragmentation (SRM cubed; SRM^3 [45]) and use of HRMS to differentiate background interference ions from target analyte ions [17,19,23,37,38]. The prior two techniques have been used in niche circumstances without wide adoption. HRMS, on the other hand, has been gaining traction in the bioanalysis community and has risen to the forefront of alternative selectivity technologies for LC–MS quantitative bioanalysis [8,13,17,19,21,23,24,34,35,37,38,46,47].

The 2015 WRIB conference in Miami, FL, USA was led off by Chris Evans who posed the following question. When you need a sensitive and selective method, the answer could be HR-MS [38]?! Evans presented three examples of assays that required increased sensitivity but experienced selectivity issues on a QQQ. In each of these examples HRMS selectivity proved to be a reliable tool able to mitigate the QQQ background interferences to push the LLOQ into the low pg/ml range. An example of HRMS power to push sensitivity limits by overcoming background interferences is displayed in Figure 4. Here, the HRMS approach was used to develop an assay for a small molecule compound, where selectivity interferences prevented quantitation below 10 pg/ml by SRM. Using the decision tree model displayed in Figure 2, the authors first attempted to develop the assay in MS1 mode. Figure 4A shows that there was constant background noise for the analyte and internal standard precursor ions even when the Q-Exactive series mass spectrometer was operated using maximum resolving power of 140,000 at *m/z* 200 and an MEW of ±0.003 *m/z* units. Following the decision tree, MS/MS (PRM) mode at varying resolving powers was employed. Figure 4B displays the excellent selectivity that was achieved in MS/MS (PRM) mode even under 'low' resolution conditions for an orbitrap (resolving power of 17,500 at *m/z* 200). The excellent selectivity at the 'low' resolution value enabled the authors' to operate the orbitrap at 12 Hz scan speed enabling excellent sampling of a 6 s wide (baseline) chromatographic peak. Under these conditions the authors were able to validate the assay with a 2 pg/ml LLOQ, a five-fold improvement over that achieved using a QQQ platform. Additional examples of this type of HRMS performance over SRM can be found in the literature [21,32].

It is important to evaluate new technologies for bioanalysis, especially if they simplify analytical workflows and increase the ruggedness of a method. Figure 5 is an example where a simpler, more user-friendly nanoelectrospray (nanoESI) source was tested during development of a large molecule quantitative assay requiring a demanding 7 pg/ml LLOQ [21]. Figure 5A shows that the 7 pg/ml LLOQ could not be achieved using the newer nanoESI source on a QQQ instrument due to sensitivity limitations rising from a high background ion current. Figure 5B shows the 7 pg/ml LLOQ on a Q-Exactive series mass spectrometer operated in MS/MS (PRM) mode. The background interference observed on the QQQ instrumentation was mitigated by the accurate mass selectivity. This assay has been successfully validated on the HRMS instrument platform.

Analytes that do not fragment well

It can be near impossible to develop a selective and sensitive SRM method for a compound that does not fragment well. For these compound types, HRMS is an excellent alternative to QQQ SRM or QQQ pseudo-SRM strategies. In 2013, two papers detailed the successful application of HRMS to bile acid quantification for discovery [13] and regulated bioanalysis [23] programs. In both publications, the authors alluded to the challenges of measuring bile acid concentrations by SRM. These challenges include the wide variety of naturally occurring bile acids which are difficult to separate chromatographically, the lack of fragmentation of unconjugated bile acids, and the use of either pseudo-SRM approaches or SRM transitions based on adducts which result in high background noise and poor selectivity [23]. Using an HRMS instrument operated in MS1 full scan (*m/z* 350 to *m/z* 550) or targeted SIM modes, the authors' were able to qualify or validate panel assays for 19 and 6 bile acids, respectively. In both cases, the background interference hindering QQQ quantitation method development was reduced by use of an HRMS MS1 approach.

Another class of compounds that do not fragment well using collision-induced dissociation are disulfide-rich cyclic peptides [37,48]. The low intensity product ions observed in MS/MS spectra of these compounds are very complex and difficult to interpret [48]. HRMS operated in MS1 mode is an attractive platform for quantitative analysis of these compounds. However, as noted above, occasionally even at its highest resolving

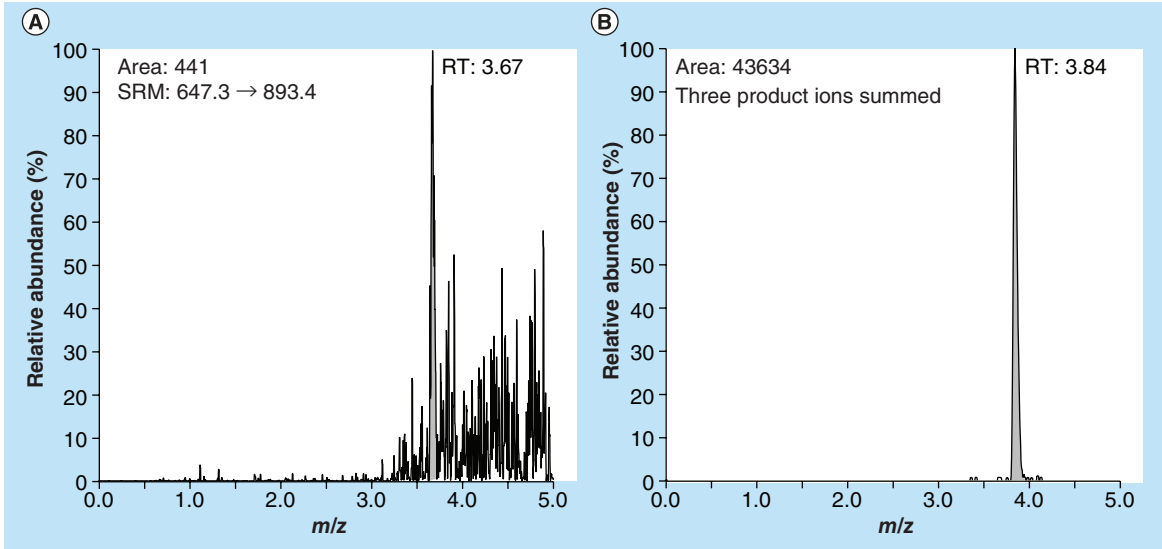


Figure 5. Example where high-resolution MS selectivity can improve the performance for a demanding LLOQ assay facilitating higher throughput. (A) LLOQ on a Thermo TSQ Vantage. **(B)** LLOQ on a Thermo Q-Exactive series mass spectrometer.

power an HRMS instrument cannot totally eliminate chemical noise which interferes with accurate detection of the analyte. For situations where chemical interferences continue to confound the generation of a selective/sensitive analytical method for a fragmentation resistant compound even at maximum resolving power, Ciccimaro *et al.* presented a novel scan mode (Survivor-SIM) to address this niche area [37]. Survivor-SIM is akin to pseudo-SRM on a QQQ. The precursor ion is selected for fragmentation in the second stage of the mass spectrometer and the fragmentation-resistant 'survivor' ion is subsequently used for quantitation. Although feasible on a QQQ mass spectrometer, the lack of resolving power and mass accuracy normally results in a high chemical background [36]. Performing the same acquisition type (Survivor-SIM) on an HRMS instrument is a different case. Figure 6 shows an example of Survivor-SIM in operation. Figure 6A displays the mass range *m/z* 616 to *m/z* 618. Within this mass range is the z = +5 charge ion envelope for the peptide of interest in human plasma. As displayed, the software was unable to determine the charge state using any of the ions observed in this window. Figure 6B displays the same mass range when a higher normalized collision energy of 24 was applied for fragmentation in the higher energy collision dissociation cell on a Q-Exactive series mass spectrometer. As displayed, the background interference ions have been fragmented into smaller product ions leaving the z = +5 charge ion envelope for the peptide of interest readily observable. This acquisition mode enables users to accurately quantify difficult to fragment analytes at a lower LLOQ by simplifying the population of ions within the *m/z* region of the analyte of interest [37].

Large molecule bioanalysis

Due to their size, charge state heterogeneity and fragmentation challenges the quantitation of peptide and protein therapeutics and biomarkers may be best performed on HRMS instrument platforms rather than QQQ platforms. There are numerous publications supporting this position [12,18,19,21,37,38,47,49]. Measurement of peptide/protein therapeutics in MS1 mode has the potential of simplifying analytical workflows by enabling whole molecule analysis, rather than bottom-up, surrogate peptide quantification of a proteolytic or chemical fragment. HRMS instruments can also allow analysts to easily determine the charge state of different ion envelopes (as displayed for the z = +5 ion envelope in Figure 6). Using a QQQ operated in unit mass resolution, it is nearly impossible to determine charge state of an ion above a z of +3.

When the measurement of the peptide/protein therapeutic (either intact or surrogate peptide) is conducted in MS/MS (PRM) mode, the entire product ion distribution can be obtained. Although an analyst may only perform quantification using single or multiple ions from the product ion spectra, the remaining ions provide confidence to the analyst that they are monitoring the correct analyte species (especially important in bottom-up approaches) [50]. In addition, bottom-up approaches make the chemical background much more complex because single peptides and proteins have now been broken down into dozens of smaller fragments that can complicate the selective detection of a target biotherapeutic surrogate peptide [51]. HRMS detection makes the complex sample more understandable adding greater selectivity to the analysis.

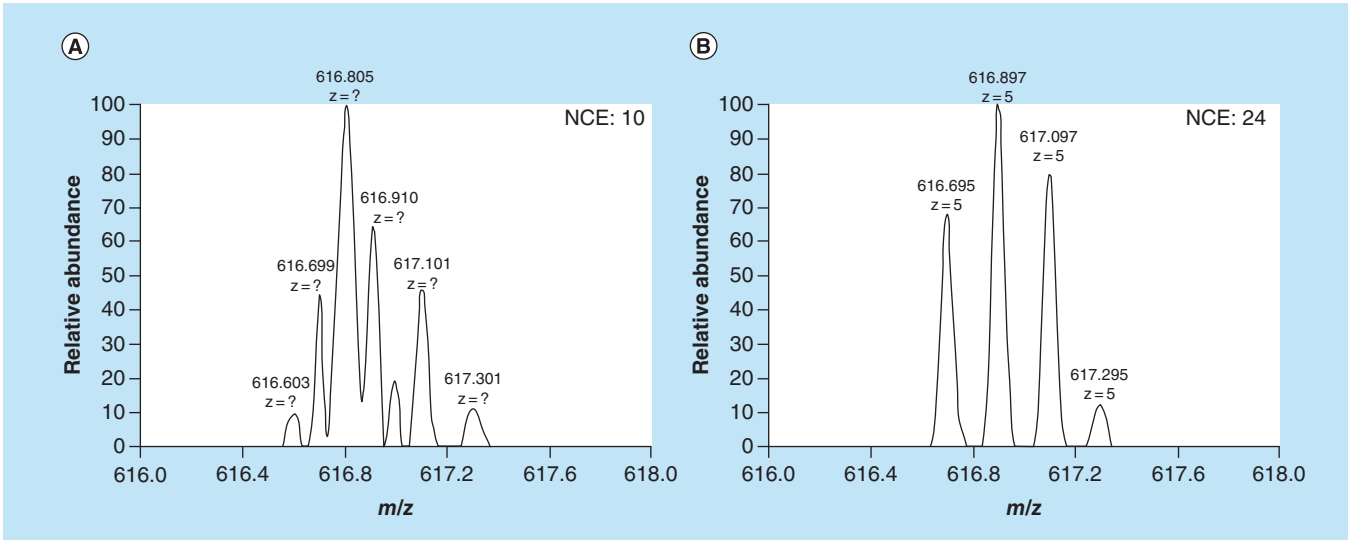


Figure 6. Example of how Survivor-SIM can improve selectivity and therefore sensitivity for a difficult to fragment disulfide-rich cyclic peptide. (A) MS/MS spectrum of a precursor ion with low MS/MS normalized collision energy (NCE = 10). **(B)** MS/MS spectrum of precursor ion with moderate MS/MS normalized collision energy (NCE = 24). Note that the ion envelope for the precursor ion is much cleaner after removing easier to fragment background ions from the target *m/z* region.

Conclusion

Advances in HRMS hardware, software, data compression algorithms, and users' general experience with the technology over the past few years have enabled the bioanalysis community to become much more comfortable with using HRMS instrumentation to support new chemical entity programs [38,49]. HRMS orthogonal selectivity to QQQ SRM approaches make the technology an excellent complement to a bioanalytical labs fleet of QQQ mass spectrometers.

Future perspective

The debate regarding which technique is more sensitive, QQQ or HRMS, will continue. There will continue to be presentations and publications stating that HRMS is more sensitive for a specific analyte and there will be publications to the contrary. The more important concept to remember is that HRMS may offer an analyst the opportunity to lower the LLOQ for a QQQ assay suffering from selectivity challenges. HRMS may be the method of choice in the future for panel assays as less upfront work is required prior to moving into validation and sample analysis. In panel assays, the extraction and chromatography need to be more general than single analyte methods. Therefore, there is a greater opportunity for selectivity challenges to arise for any of the analytes or their internal standards using a QQQ SRM approach. In the near future, the authors' believe they will see additional HRMS instruments entering pharmaceutical and CRO laboratories and sponsor demand for HRMS assays will increase especially as additional biotherapeutics enter the product pipeline.

Acknowledgements

The authors thank Thermo Scientific for their advice and support in implementing HRMS technology in a regulated bioanalytical lab and Ms Susan Swift-Spencer of Q^2 Solutions for her assistance with figure creation.

Financial & competing interests disclosure

The authors have no relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript. This includes employment, consultancies, honoraria, stock ownership or options, expert testimony, grants or patents received or pending, or royalties.

No writing assistance was utilized in the production of this manuscript.

Executive summary

- High-resolution MS (HRMS) instrumentation can be successfully implemented into a regulated bioanalysis laboratory. Laboratory processes need to be thoroughly assessed to determine appropriate operating procedures prior to use of HRMS instrumentation for regulated bioanalysis.
- HRMS instruments will not replace QQQ instrumentation. HRMS instrumentation will complement existing technology and offer a platform for selectivity-challenged assays.
- HRMS will have a more significant impact to large molecule biotherapeutic and biomarker applications than to conventional small molecule applications.

References

Papers of special note have been highlighted as:
• of interest; •• of considerable interest

1 Ramanathan R, Jemal M, Ramagiri S *et al.* It is time for a paradigm shift in drug discovery bioanalysis: from SRM to HRMS. *J. Mass. Spectrom.* 46(6), 595–601 (2011).

2 Korfmacher W. High-resolution mass spectrometry will dramatically change our drug-discovery bioanalysis procedures. *Bioanalysis* 3(11), 1169–1171 (2011).

3 Bateman KP, Kellmann M, Muenster H, Papp R, Taylor L. Quantitative-qualitative data acquisition using a benchtop orbitrap mass spectrometer. *J. Am. Soc. Mass. Spectrom.* 20(8), 1441–1450 (2009).

4 Zhang H, Zhang D, Ray K, Zhu M. Mass defect filter technique and its applications to drug metabolite identification by high-resolution mass spectrometry. *J. Mass. Spectrom.* 44(7), 999–1016 (2009).

5 Hopfgartner G, Tonoli D, Varesio E. High-resolution mass spectrometry for integrated qualitative and quantitative analysis of pharmaceuticals in biological matrices. *Anal. Bioanal. Chem.* 402(8), 2587–2596 (2012).

6 Meyer M, Maurer H. Current applications of high-resolution mass spectrometry in drug metabolism studies. *Anal. Bioanal. Chem.* 403(5), 1221–1231 (2012).

7 Zhang J, Maloney J, Drexler D *et al.* Cassette incubation followed by bioanalysis using high-resolution MS for *in vitro* ADME screening assays. *Bioanalysis* 4(5), 581–593 (2012).

8 Murphy K, Bennett P, Duczak N Jr. High-throughput quantitation of large molecules using multiplexed chromatography and high-resolution/accurate mass LC-MS. *Bioanalysis* 4(9), 1013–1024 (2012).

9 Fung E, Jemal M, Aubry A. High-resolution MS in regulated bioanalysis: where are we now and where do we go from here? *Bioanalysis* 5(10), 1277–1284 (2013).

• **Excellent, thoughtful review of the state of HRMS in regulated bioanalysis in 2013.**

10 Huang M, Lin Z, Weng N. Applications of high-resolution ms in bioanalysis. *Bioanalysis* 5(10), 1269–1276 (2013).

• **Excellent, thoughtful review of the state of HRMS in regulated bioanalysis in 2013.**

11 Bruce SJ, Rochat B, Beguin A *et al.* Analysis and quantification of vitamin D metabolites in serum by ultra-performance liquid chromatography coupled to tandem mass spectrometry and high-resolution mass spectrometry – a method comparison and validation. *Rapid Commun. Mass Spectrom.* 27(1), 200–206 (2013).

12 Dillen L, Cools W, Vereyken L *et al.* Comparison of triple quadrupole and high-resolution TOF-MS for quantification of peptides. *Bioanalysis* 4(5), 565–579 (2012).

13 Ding J, Lund ET, Zulkoski J, Lindsay JP, Mckenzie DL. High-throughput bioanalysis of bile acids and their conjugates using UHPLC coupled to HRMS. *Bioanalysis* 5(20), 2481–2494 (2013).

14 Fung E, Xia YQ, Aubry A, Zeng J, Olah T, Jemal M. Full-scan high resolution accurate mass spectrometry (HRMS) in regulated bioanalysis: LC-HRMS for the quantitation of prednisone and prednisolone in human plasma. *J. Chromatogr. B Anal. Technol. Biomed. Life Sci.* 879(27), 2919–2927 (2011).

15 Li Y, Hackman M, Wang C. Quantitation of polypeptides (glucagon and salmon calcitonin) in plasma samples by 'high resolution' on a triple quadrupole mass spectrometer. *Bioanalysis* 4(6), 685–691 (2012).

16 Mekhssian K, Mess J, Garofolo F. Application of high-resolution MS in the quantification of a therapeutic monoclonal antibody in human plasma. *Bioanalysis* 6(13), 1767–1779 (2014).

17 Mulvana DE, Sturm R, Bowen CL *et al.* Overcoming triple quadrupole selectivity challenges by use of high-resolution accurate mass spectrometry to enable sensitive quantitation of a small molecule therapeutic in rabbit plasma. Presented at: *21st International Reid Bioanalytical Forum*. Guildford, UK, 7–10 September 2015.

18 Plumb RS, Fujimoto G, Mather J *et al.* Comparison of the quantification of a therapeutic protein using nominal and accurate mass MS/MS. *Bioanalysis* 4(5), 605–615 (2012).

19 Ramagiri S, Garofolo F. Large molecule bioanalysis using Q-TOF without predigestion and its data processing challenges. *Bioanalysis* 4(5), 529–540 (2012).

•• **Excellent discussion of a mass extractions window's effect on selectivity and sensitivity of a quantitative HRMS assay.**

20 Rousu T, Tolonen A. Comparison of unit resolution SRM and TOF-MS at 12,000 mass resolution for quantitative bioanalysis of 11 steroids from human plasma. *Bioanalysis* 4(5), 555–563 (2012).

21 Sturm RM, Jones BR, Shan L, Buckholz JE, Biondolillo RF, Mulvana DE. PD biomarker measurement by LC/MS/MS – resolving sensitivity and throughput challenges using high resolution mass spectrometry. Presented at: *16th Annual Land O'Lakes Bioanalytical Conference*. Madison, WI, USA, 13–16 July 2015.

22 Tonoli D, Varesio E, Hopfgartner G. Quantification of acetaminophen and two of its metabolites in human plasma by ultra-high performance liquid chromatography-low and high resolution tandem mass spectrometry. *J. Chromatogr. B Anal. Technol. Biomed. Life Sci.* 904, 42–50 (2012).

23 Voelker T, Wang H, Irish M *et al.* Method development and validation of six bile acids for regulated bioanalysis: improving selectivity and sensitivity. *Bioanalysis* 5(10), 1229–1248 (2013).

24 Wang H, Bennett P. Performance assessment of microflow LC combined with high-resolution MS in bioanalysis. *Bioanalysis* 5(10), 1249–1267 (2013).

25 Wei C, Grace JE Jr., Zvyaga TA, Drexler DM. Utility of high-resolution accurate MS to eliminate interferences in the bioanalysis of ribavirin and its phosphate metabolites. *Bioanalysis* 4(15), 1895–1905 (2012).

26 L Lawson G, Cocks E, Tanna S. Quantitative determination of atenolol in dried blood spot samples by LC-HRMS: a potential method for assessing medication adherence. *J. Chromatogr. B: Anal. Technol. Biomed. Life Sci.* 897, 72–79 (2012).

27 Li Y, Henion J, Abbott R, Wang P. Dried blood spots as a sampling technique for the quantitative determination of

guanfacine in clinical studies. *Bioanalysis* 3(22), 2501–2514 (2011).

28 Oliveira RV, Henion J, Wickremsinhe ER. Automated direct extraction and analysis of dried blood spots employing on-line SPE high-resolution accurate mass bioanalysis. *Bioanalysis* 6(15), 2027–2041 (2014).

29 Sturm R, Henion J, Abbott R, Wang P. Novel membrane devices and their potential utility in blood sample collection prior to analysis of dried plasma spots. *Bioanalysis* 7(16), 1987–2002 (2015).

30 FDA Guidance for Industry Computerized Systems Used in Clinical Investigations (2007). www.fda.gov

31 Thermo Scientific. Best practices in Q exactive maintenance and troubleshooting. Presented at: *iOrbi 5*. Ithaca, NY, USA, 6 December 2014.

32 Bowen CL, Kehler JR, Boram S, Abberley L, Szapacs M, Evans C. Modify on the fly: triple quad to high resolution in support of a dermal clinical study requiring an ultra low LLOQ. *Bioanalysis* 8(3), 205–214 (2016).

•• **Excellent example of how HRMS was utilized to further the development of a therapeutic drug when the SRM sensitivity suffered from selectivity issues.**

33 Glauser G, Grund B, Gassner AL *et al.* Validation of the mass-extraction-window for quantitative methods using liquid chromatography high resolution mass spectrometry. *Anal. Chem.* 88(6), 3264–3271 (2016).

34 Gallien S, Domon B. Quantitative proteomics using the high resolution accurate mass capabilities of the quadrupole-orbitrap mass spectrometer. *Bioanalysis* 6(16), 2159–2170 (2014).

35 Peterson A, Russell J, Bailey D, Westphall M, Coon J. Parallel reaction monitoring for high resolution and high mass accuracy quantitative, targeted proteomics. *Mol. Cell. Proteomics* 11(11), 1475–1488 (2012).

36 Fu Y, Xia Y, Flarakos J *et al.* Differential mobility spectrometry coupled with multiple ion monitoring in regulated LC-MS/MS bioanalysis of a therapeutic cyclic peptide in human plasma. *Anal. Chem.* 88(7), 3655–3661 (2016).

37 Ciccimaro E, Ranasinghe A, D'arienzo C *et al.* Strategy to improve the quantitative LC-MS analysis of molecular ions resistant to gas-phase collision induced dissociation: application to disulfide-rich cyclic peptides. *Anal. Chem.* 86(23), 11523–11527 (2014).

• **Describes application of Survivor-SIM to a peptide therapeutic.**

38 Evans C. When you need a sensitive and selective method, the answer could be HR-MS?! Presented at: *9th Workshop on Recent Issues in Bioanalysis*. Miami, FL, USA, 13–17 April 2015.

39 Henion J, Brewer E, Rule G. Sample preparation for LC/MS/MS: analyzing biological and environmental samples. *Anal. Chem.* 70(19), A650–A656 (1998).

40 Jemal M. High-throughput quantitative bioanalysis by LC/MS/MS. *Biomed. Chromatogr.* 14(6), 422–429 (2000).

41 Korfmacher W. Foundation review: principles and applications of LC-MS in new drug discovery. *Drug Discov. Today* 10(20), 1357–1367 (2005).

42 Zhou S, Song Q, Tang Y, Naidong W. Critical review of development, validation, and transfer for high throughput bioanalytical LC-MS/MS methods. *Curr. Pharm. Anal.* 1(1), 3–14 (2005).

43 Xia YQ, Wu ST, Jemal M. LC-FAIMS-MS/MS for quantification of a peptide in plasma and evaluation of FAIMS global selectivity from plasma components. *Anal. Chem.* 80(18), 7137–7143 (2008).

44 Zerr A, Heunier L, Wood S, Struwe P. Application of differential ion mobility mass spectrometry to peptide quantitation. Presented at: *European Bioanalysis Forum*. Barcelona, Spain, 14–16 November 2012.

45 Sidibe J, Varesio E, Hopfgartner G. Quantification of ghrelin and des-acyl ghrelin in human plasma by using cubic-selected reaction-monitoring LCMS. *Bioanalysis* 6(10), 1373–1383 (2014).

46 Campbell J, Ramagiri S. Recognizing the potential benefits and pitfalls of high-resolution MS. *Bioanalysis* 5(10), 1157–1160 (2013).

47 Dillen L, Cuyckens F. High-resolution MS: first choice for peptide quantification? *Bioanalysis* 5(10), 1145–1148 (2013).

48 Liu W, Ng J, Meluzzi D, Bandeira N *et al.* Interpretation of tandem mass spectra obtained from cyclic nonribosomal peptides. *Anal. Chem.* 81(11), 4200–4209 (2009).

49 Kellie J, Kehler J, Szapacs M. Application of high-resolution MS for development of peptide and large-molecule drug candidates. *Bioanalysis* 8(3), 169–177 (2016).

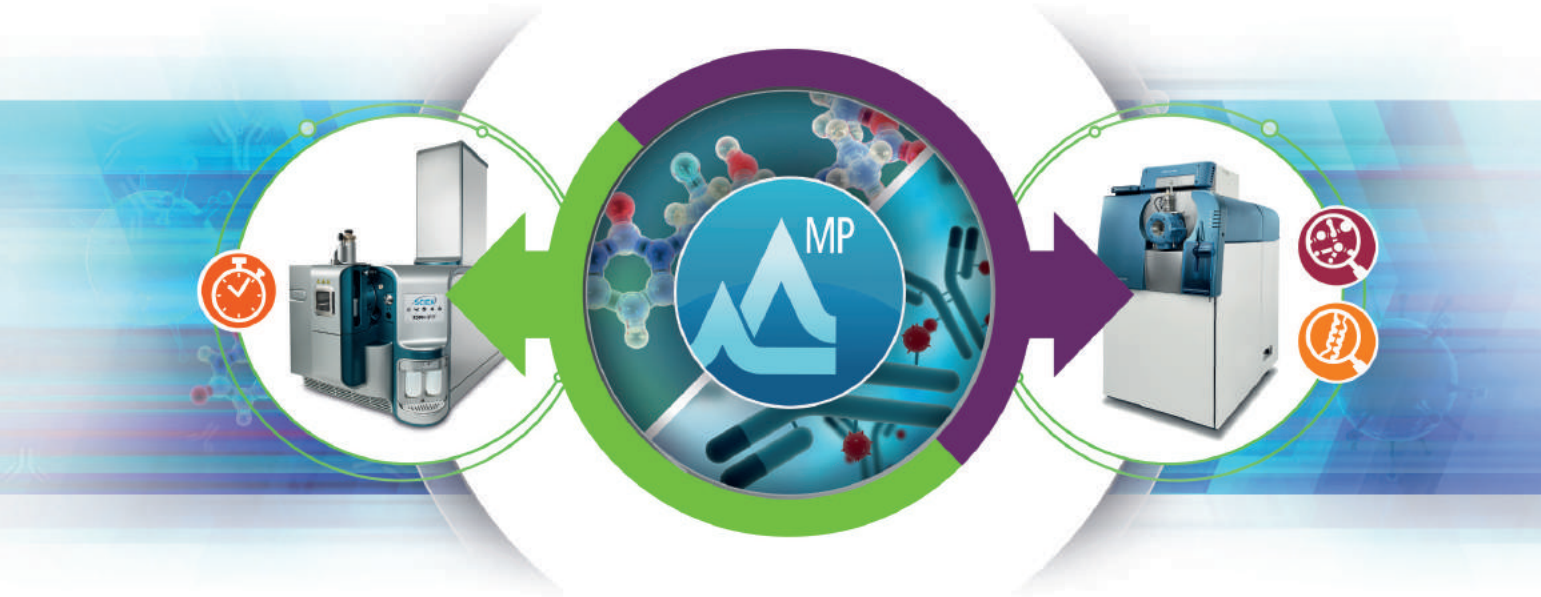
• **Provides evidence that HRMS and QQQ platforms provide identical PK data from an in-life sample set.**

50 Delatour T, Mottier P, Gremaud E. Limits of suspicion, recognition and confirmation as concepts that account for the confirmation transitions at the detection limit for quantification by liquid chromatography-tandem mass spectrometry. *J. Chromatogr. A* 1169(1), 103–110 (2007).

51 Zheng N, Zeng J, Manney A *et al.* Quantitation of a PEGylated protein in monkey serum by UHPLC-HRMS using a surrogate disulfide-containing peptide: a new approach to bioanalysis and *in vivo* stability evaluation of disulfide-rich protein therapeutics. *Anal. Chim. Acta* 916, 42–51 (2016).

High Productivity, Total Confidence

Routine and Advanced Biotransform Solutions



Whether you are doing high-throughput quick-look MetID and soft spot analysis, or comprehensive identification of drug metabolites and peptide catabolites, you need fast and accurate results, every time. SCIEX has developed the most advanced and integrated Biotransform Solutions available, giving you a complete workflow with the ultimate confidence in results.

Do More: Get high-throughput, high-quality metabolite ID with the streamlined and easy-to-use X500 QTOF System, so you can consistently meet challenging MetID timelines.

See More: Gain unbeatable detection and quantitation of virtually every detectable metabolite in a single injection, using data independent SWATH® Acquisition.

Achieve More: Interpret your metabolism data faster with powerful, automated processing from MetabolitePilot™ Software 2.0 that interprets and correlates structure and profiles for both small and large molecules.

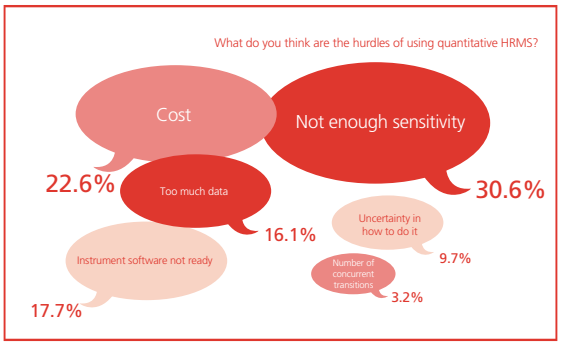
Explore the Biotransform Solutions from SCIEX, and learn how to transform your metabolism studies.

SCIEX.com/biotransform

SCIEX

Answers for Science.
Knowledge for Life.™

Spotlight: Quantitative HRMS



Recent significant improvements in high resolution mass spectrometry (HRMS) technology have challenged traditional methods of quantitative and qualitative workflows. In this Spotlight event hosted by Bioanalysis Zone, we explored how recent developments in this cutting-edge technology have been received by the bioanalytical community, and look to the future applications of HRMS in the field.

In a series of engaging webinars and feature articles from *Bioanalysis*, we looked at the broad range of applications of the technique, its associated benefits and challenges, and future developments of HRMS.

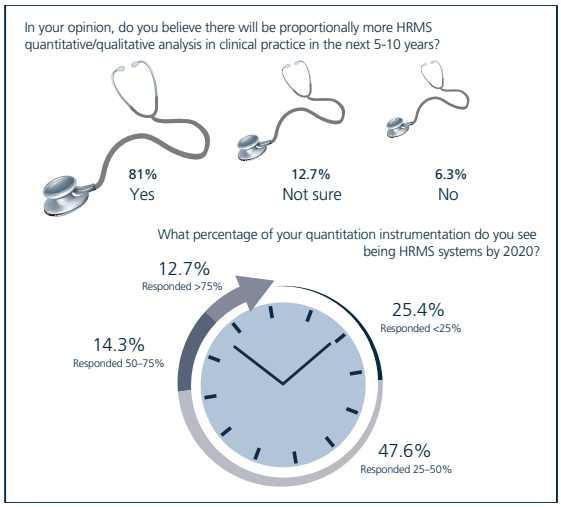
Why has HRMS technology increasingly gained attention from many within the bioanalytical community? What are the key benefits offered by this instrument? What key challenges need to be addressed for the successful application of HRMS? This Spotlight engaged key experts within the field to offer opinion and discourse on these questions.

The Spotlight series began with a survey from which we aimed to gain insight on the use of HRMS within the bioanalytical industry. The survey produced some very exciting insights into what the bioanalytical community thinks of HRMS; opinions that were further discussed by our panel of leading experts in a live panel discussion event.

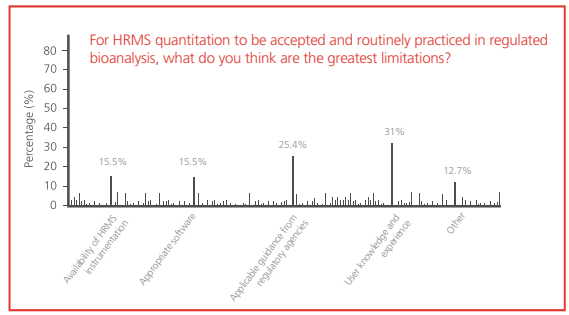
One of the recurring messages throughout this Spotlight focus was that the virtues of HRMS were indisputable. Our results reflected what the bioanalytical community have previously voiced: HRMS has many virtues and has the potential to impact on a very large proportion of chemical space.

The ability of HRMS to simultaneously collect both quantitative and qualitative data for targeted and non-targeted analysis is one of the foremost reasons that attract many to the technology. Other equally important factors include HRMS' multiplexing capabilities, adaptability, and its ability to obtain rich data sets.

Another exciting feature of this technology is its broad range of applications. HRMS' multiplexing capabilities is needed in both small and large molecules. The technology has been around for a long time and while all the scientific groundwork has already been laid out, according to Russell Mortishire-Smith (Waters Corporation), one of the experts who participated in the panel discussion: "… the real revolution in HRMS has come in the last 10 years. Over the next 5–10 years many of the applications of HRMS will be formally validated and put into everyday use."



It is clear these are exciting times for the application of HRMS, and the future is thus very bright and promising for the technique. It could have implications in the clinical context, in fields such as personalized medicine, where it will allow enhanced studies in proteomics and metabolomics, in addition to reducing our reliance on LC for separation, especially when LC is hard to maintain in the laboratory.



If the bioanalytical community is unequivocal in the benefits of HRMS – data from our survey did point out what may have seen as reluctance by some in switching to using HRMS. To explain what this result implied Diego Cortes (R&D Manager, PPD® Laboratories) responded: "More than reluctance there are regulatory concerns from labs and pharma companies to adopt HRMS for quantitation in biomed. Dialogues between regulated bioanalytical community and regulatory agencies need to take place to address potential regulatory concerns the bioanalytical community have."

Thus the reluctance is not based on the inadequacy of the technology but rather on the question raised: Can simultaneous quantitative and qualitative analysis fit in regulated bioanalysis?

In addition to the panel discussion event, we also held three mini webinars as part of this Spotlight. In these webinars by leading experts, Diego Cortes (PPD® Laboratories), Michal Kliman (Mass Spec) and Yves le Blanc (SCIEX) presented some of the advantages of HRMS based on their experiences.

Key Points

- HRMS has been available for a long time; originally it was applied in qualitative applications, however the introduction of spectrometers with improved mass resolution over wider mass ranges has allowed for the application of HRMS in quantitative workflows.
- In the clinical setting, HRMS is increasingly being adopted due to its ease of operation, reasonable cost and maintenance. HRMS can also be combined with qualitative exploration of data in the small molecule field.

Research Article

Bioanalysis

Hybridizing techniques

Antibody–drug conjugate bioanalysis using LB-LC–MS/MS hybrid assays: strategies, methodology and correlation to ligand-binding assays

Background: Antibody–drug conjugates (ADCs) are complex drug constructs with multiple species in the heterogeneous mixture that contribute to their efficacy and toxicity. The bioanalysis of ADCs involves multiple assays and analytical platforms. **Methods:** A series of ligand binding and LC–MS/MS (LB-LC–MS/MS) hybrid assays, through different combinations of anti-idiotype (anti-Id), anti-payload, or generic capture reagents, and cathepsin-B or trypsin enzyme digestion, were developed and evaluated for the analysis of conjugated-payload as well as for species traditionally measured by ligand-binding assays, total-antibody and conjugated-antibody. **Results & conclusion:** Hybrid assays are complementary or viable alternatives to ligand-binding assay for ADC bioanalysis and PK/PD modeling. The fit-for-purpose choice of analytes, assays and platforms and an integrated strategy from Discovery to Development for ADC PK and bioanalysis are recommended.

First draft submitted: 26 January 2016; Accepted for publication: 16 May 2016; Published online: 9 June 2016

Keywords: antibody–drug conjugate • antibody conjugated-payload • DAR-bias • DAR-sensitive ADC assay • drug antibody ratio • fit-for-purpose bioanalysis • generic reagents • immunocapture LC–MS/MS hybrid assay • ligand binding-LC–MS/MS hybrid assay • soluble target

Jian Wang*[,1], Huidong Gu[1], Ang Liu[1], Alexander Kozhich[1], Vangipuram Rangan[2], Heather Myler[1], Linlin Luo[1], Richard Wong[3], Huadong Sun[3], Bonnie Wang[4], Heather E Vezina[5], Shrikant Deshpande[2], Yan Zhang[1], Zheng Yang[3], Timothy V Olah[3], Anne-Francoise Aubry[1], Mark E Arnold[1], Renuka Pillutla[1] & Binodh DeSilva[1]

[1]Analytical & Bioanalytical Operations, Bristol-Myers Squibb, Princeton, NJ 08543, USA

[2]Biologics Discovery California, Bristol-Myers Squibb, Redwood City, CA 95035, USA

[3]Preclinical Candidate Optimization, Bristol-Myers Squibb, Princeton, NJ 08543, USA

[4]Drug Safety Evaluation, Bristol-Myers Squibb, New Brunswick, NJ 08903, USA

[5]Clinical Pharmacology & Pharmacometrics, Bristol-Myers Squibb, Princeton, NJ 08543, USA

*Author for correspondence:
Tel.: +1 609 252 3856
jian.wang@bms.com

This decade has seen the US FDA approvals of two antibody–drug conjugates (ADCs), Adcetris® (brentuximab vedotin) and Kadcyla® (ado-trastuzumab emtansine) for the treatment of cancer [1,2] with many more ADCs in clinical development. ADCs are drug constructs consisting of a monoclonal antibody (mAb), attached through a chemical linker to a small molecule, often cytotoxic, drug or payload. ADCs combine the target binding specificity of mAbs and the potency of cytotoxic drugs [3,4]. Many ADC drug candidates are currently at different preclinical and clinical development stages in this rapidly evolving field [5,6].

Because of the heterogeneous nature of ADC molecules, particularly for random-conjugated ADCs, multiple species may contribute to the efficacy and toxicity of ADCs. ADC PK evaluation, human dose projection, PK/PD modeling and the associated bioanalytical PK assays are complicated. ADC bioanalytical testing has, traditionally, used both ligand-binding assays (LBAs) and LC–MS/MS approaches. Currently, industry-wide standard practices, strategies, and regulatory guidelines are still developing. ADC PK and bioanalytical strategies have been one of the most hotly discussed topics industry-wide during the last 5–10 years. Two recent special focus issues have been dedicated to ADC PK and bioanalysis recently [7,8]. As a highlight, the AAPS Drug Conjugate Working Group issued a position paper on bioanalysis of ADCs in which definitions of ADC analytes and bioanalytical assays were discussed [9]. There were comprehensive reviews on ADC bioanalytical assay

FUTURE SCIENCE part of fsg

strategies and challenges [10–13], and on *in vivo* biotransformation [14] and characterization of ADC catabolism [15]. An overview of analytes and ADC bioanalytical strategies was also outlined in a recent industry white paper on ADME characterization of ADCs from the 'ADC ADME working group' of the International Consortium for Innovation and Quality in Pharmaceutical Development [16] and in reviews of ADME properties of therapeutic proteins [17], and of characterization of ADCs by MS [18]. In depth coverage of LBA in ADC bioanalysis from early to late preclinical development was the subject of one recent review [19], while the validation of an integrated series of LBAs for clinical studies was described in another article [20]. An integrated, multiplatform strategy for the bioanalysis of a microtubule polymerization inhibitor ADC [21] and details of PK assays for T-DM1 [22] have been reported. While not directly related to bioanalysis, but influential in consideration of the required assays, PK considerations and PK/PD modeling also have been the subject of review and research articles [23–27].

Based on literature and presentations from industry and regulatory authorities, four major analytes are recommended for ADC PK: total-antibody, which includes antibody conjugated to the payload and unconjugated (naked) antibody; conjugated-antibody, defined as the antibody conjugated to at least one payload; conjugated-payload, defined as payload conjugated to the antibody; unconjugated-payload, defined as payload molecules deconjugated from the antibody [9,10]. Both conjugated-antibody and conjugated-payload are considered measures of the ADC active species, describing ADCs from the perspective of either the antibody or the payload [10]. It is unclear, at this time, whether the conjugated-antibody or the conjugated-payload correlates more closely with efficacy or toxicity of ADCs. The ratio of conjugated-payload and total-antibody is defined as the average drug to antibody ratio (DAR) and its change *in vivo* is considered a good indication of the degree of ADC deconjugation and other biotransformation processes.

Historically, the primary assays for total-antibody and conjugated-antibody have been LBAs and the unconjugated-payload assay has been by LC–MS/MS, while the conjugated-payload is predominately measured by affinity capture or immunocapture LC–MS/MS hybrid assays [9–11]. Since these hybrid assays are truly the combination of ligand binding platforms and LC–MS technology, we chose to name them as LB-LC–MS/MS hybrid assays, instead, in this article. It was also reported that DAR-sensitive conjugated-antibody assays can be used to monitor *in vivo* payload change especially for ADCs with noncleavable linkers [11,19,28]. Affinity-capture followed by capillary LC–MS or hydrophobic interaction chromatography (HIC) to characterize intact ADC DAR distribution in biological matrices have been reported [10,29–32]. Immunocapture combined with cathepsin-B cleavage followed by LC–MS/MS for the quantitation of conjugated-payload has been discussed [9,10] and a case study of assay development, validation, and application has been published [33]. A recent publication reports a protein A capture LC–MS/MS hybrid assay with papain (a cysteine protease) cleavage for the quantitation of valine-citrullin-linked MMAE ADCs [34]. Method development and assay validations for the analysis of unconjugated-payload of ADCs by LC–MS/MS assays have been the topic in recent publications [35–37].

Immunocapture LC–MS/MS assays have been increasingly used in bioanalysis of protein therapeutics in recent years [38–41]. To our best knowledge, though discussed in various review articles [9–11], there have been few publications detailing LB-LC–MS/MS hybrid assay methodology, procedures and assay characteristics for ADC PK bioanalysis, particularly for total-antibody and conjugated-antibody. Meanwhile, in ADC PK and bioanalysis, questions such as "which assays and analytes are preferred, conjugated-antibody or conjugated-payload?" and "what assay format to use, LBAs or hybrid assays in preclinical and clinical PK studies?" have remained hot topics in the literature, workshops and conferences. This article describes a series of LB-LC–MS/MS hybrid assays for ADC PK analysis, which have been used in preclinical and/or clinical studies and attempts to provide some answers to these questions: Can hybrid assays play additional roles beyond measuring conjugated-payload? Are hybrid assays viable alternatives to LBA for the analysis of total-antibody? Can hybrid assays be used as an alternative or complement to LBAs for the analysis of conjugated-antibody? Can ADC bioanalysis be supported in a LC–MS based laboratory without full LBA capabilities? What is the ideal ADC bioanalytical strategy from early Discovery to late clinical stage? All hybrid assays were developed, and validated or qualified, using protocols that include tests normally required for both LBA and LC–MS/MS assays [42–46]. Method development and validation or qualification of hybrid assays and results from PK/toxicokinetic (TK) studies are discussed. This article aims at providing technical guidance for the scientists conducting various hybrid assays for ADC bioanalysis. The article will also address some of the highly discussed topics in ADC bioanalysis as mentioned above. The comparison or correlation between hybrid assays and LBAs will not only be limited to assay platform but also incorporate ADC analytes and assay DAR characteristic which brings the discussion to a wider scope and to better

clarity. Discovery and Development integrated ADC bioanalysis strategies are proposed herein as well. The ADC described in this manuscript is in early development and consists of a microtubule polymerization inhibitor as the payload attached to a specific undisclosed mAb through random lysine conjugation [21,47]. General statements made in this article are applicable to this ADC and other random-conjugated ADCs, while they may not apply to ADCs of other constructs.

Experimental

LBAs for total-antibody & conjugated-antibody

Methods utilizing a sandwich immunoassay format and the standard multiarray microplate (Meso Scale Discovery, MD, USA) platform were used in preclinical PK and TK studies. The method details have been described in Myler *et al.* [21]. The methods used in the clinical PK evaluation were modified to a Gyrolab xP Workstation (Gyros Inc, Uppsala, Sweden) format [20]. To be noted in the published methods, conjugated-antibody in which the antibody conjugates with the active payload was defined as 'Active-ADC,' and the conjugated-antibody (payload + metabolite) in which the antibody conjugates with either payload (active) or payload metabolite (inactive) was defined as 'Total ADC.' This manuscript discusses ADC bioanalytical methodology in general, so the widely used nomenclatures for ADC analytes are chosen for the convenience of communication.

LB-LC–MS/MS hybrid assays for conjugated-payload

The biotinylation and the immunocapture procedure using streptavidin cartridges with specific biotinylated anti-Id or anti-payload capture mAb for the analysis of conjugated-payload, and the subsequent cathepsin-B cleavage and LC–MS/MS detection have been described in details elsewhere [33].

Briefly, a specific anti-Id mAb was first labeled with biotin at a molar challenge ratio of 12:1, and then immobilized on high-capacity streptavidin-coated cartridges using an Agilent Bravo liquid handler (Agilent Technologies, MA, USA). Plasma samples were loaded on the cartridges at a slow flow rate of 2 µl/min, which was considered to be critical to ensure quantitative capture between immobilized capture mAb and ADC in the biological sample. After multiple washes of the cartridges and the syringes, the captured analytes were eluted into a collection plate by reducing the pH to 2. The eluted samples were neutralized immediately to a slightly acidic condition which was favorable for the following enzyme cleavage step. The ADC described in this manuscript contains a cleavable dipeptide linker that can be cleaved specifically by lysosomal enzymes. A specific lysosomal enzyme, cathepsin B, was used and the *in vivo* condition was mimicked to release the payload from the ADC. The cleavage was conducted in a buffer containing 1.34 units of the enzyme at an optimal pH of 5–6 in the presence of reducing reagent DTT. The mixture was incubated at 25°C for 3 h and the reaction was stopped by adding cold acetonitrile. LC–MS/MS was carried out using a Shimadzu (Tokyo, Japan) Nexera HPLC system interfaced to a SCIEX (CA, USA) Triple Quad 5500 mass spectrometer. A Waters (MA, USA) Acquity UPLC BEH C18 column (1.7 µm, 2.1 × 50 mm) was used for chromatographic separation and the mobile phases consisted of mobile phase A (5 mM ammonium bicarbonate containing 0.005% ammonium hydroxide in 95:5 water: acetonitrile) and mobile phase B (5:95 water:acetonitrile). Positive ion electrospray MS in multiple reaction monitoring mode was used for the detection. The transitions of *m/z* 771→98 and *m/z* 781→98 were monitored for the payload and the internal standard (IS) ($^{13}C_9$, ^{15}N-payload), respectively.

Protein A or protein G cartridges were also used for the analysis of conjugated-payload. The cartridges were equilibrated with 50 µl of PBST (Phosphate buffered saline containing 0.05% Tween 20, Sigma-Aldrich, Sweden) or DPBS (Dulbecco's phosphate buffered saline, Lonza, MD, USA) at a flow rate of 25 µl/min. Plasma samples (25 µl) were loaded on the cartridges at a flow rate of 2 µl/min. The cartridges were washed with 50 µl of washing solution, which consisted of 50 mM ammonium bicarbonate and 0.1% BSA in DPBS. The elution and neutralization steps were the same as described in the previous section.

LB-LC–MS/MS hybrid assays for total-antibody & conjugated-antibody

Immunocapture

Cartridge format

The biotinylation and the immunocapture procedure, using streptavidin cartridges with specific biotinylated anti-Id or anti-payload capture mAbs for the analysis of total-antibody or conjugated-antibody, and protein A cartridge for the analysis of total-antibody, are the same as described above for the analysis of conjugated-payload analysis.

Magnetic-beads format

Immunocapture was also conducted with Dynabeads Myone streptavidin T1 magnetic beads (Life Technologies, CA, USA) using biotinylated mouse anti-Id or mouse anti-payload mAb. Mouse anti-payload mAb bound to protein G beads were also explored.

Beads (1 ml) were washed three-times with 1 ml of PBST solution, and then resuspended with 1 ml of

PBST. Based on the bead capacities (200 µg of mAb per 1 ml of streptavidin T1 beads, and 240 µg of mAb per 1 ml protein G beads), the beads were saturated by adding about 400 µg of mAb for streptavidin T1 beads and 480 µg of mAb for protein G beads, respectively. The beads and mAb mixture was incubated at room temperature for 1 h on a rotator. After incubation, the beads were washed again three-times with 1 ml of PBST, and then resuspended with 1 ml of PBST containing 1% BSA (bovine serum albumin). The immobilized beads were stored at 4°C for future use.

PBST (10 µl) was added to 10 µl of each serum sample in a 96-well plate, and then 50 µl of capture beads was added to each sample. The plate was incubated at room temperature for 30 min in a Thermomixer R model 5355 incubator (Eppendorf, Hamburg, Germany) at a vortexing speed of 900 rpm (15 s on/off) for the immunocapture. The plate was first washed three-times with 200 µl of PBST, and then one time with 200 µl of 50 mM ammonium bicarbonate using Janus Mini automated liquid handler from Perkin Elmer (MA, USA).

Trypsin digestion for total-antibody & conjugated-antibody

To each of the samples obtained from immunocapture, using either beads format or cartridge format, 100 µl of 50 mM ammonium bicarbonate containing 10% methanol was added, the samples were then incubated at 90°C for 30 min for thermal denaturation. After the samples cooled to room temperature, 25 µl of Promega trypsin (~50 µg/ml) was added into each sample for digestion at 60°C for 2 h. The digestion was stopped by adding 10 µl of 10% formic acid in water, and 20 µl of 200 ng/ml of stable isotopic labeled signature peptide IS working solution in 20% acetonitrile and 80% water was added before LC–MS/MS analysis of total Ab or conjugated-antibody.

LC–MS/MS method

LC–MS/MS was carried out using a Shimadzu (Tokyo, Japan) Nexera HPLC system interfaced to a SCIEX (CA, USA) Triple Quad 6500 mass spectrometer equipped with a TurboIonspray™ source. Forty (40) µl of digested solution was injected into a Waters Acuity UPLC HSS T3 column (1.7 µm, 2.1 × 50 mm) with a gradient elution using mobile phases of 0.1% formic acid in water (A) and 0.1% formic acid in acetonitrile (B) at a flow rate of 0.8 ml/min. The column temperature and the autosampler temperature were 60°C and 5°C, respectively. Positive ion electrospray in multiple reaction monitoring mode was used for the detection. The optimized source temperature, turbo ionspray voltage, declustering potential and collision energy were set to 600°C, 5500 V, 50 V and 28 eV, respectively. The transitions of *m/z* 531.3 -> 785.4 and 534.3 -> 785.4 were monitored for the signature peptide [48] from the CDR region of the antibody and the IS (stable labeled signature peptide).

LB-LC–MS/MS hybrid assays developed using different immunocapture reagents on different immunocapture platforms for the analysis of conjugated-payload, total-antibody and conjugated-antibody in different biological matrices are summarized in Supplementary Table 1.

Results & discussion

Nomenclature of ADC analytes & complexity of ADC bioanalysis

The four recommended ADC analytes (total-antibody, conjugated-antibody, conjugated-payload and unconjugated-payload) [9,10] illustrate the relationship between antibody and payload in random-conjugated ADCs. The terms 'total' and 'free,' which are often used in bioanalysis of biologics [49,50], take a slightly more complex meaning in ADC bioanalysis. In protein therapeutics, 'total' refers to the sum of bound and unbound to soluble targets. In the case of ADC, 'total' can also mean conjugated plus unconjugated antibody. Then, there are two meanings of 'total' and 'free' in ADC bioanalysis. 'Free' could mean 'unconjugated' to the antibody or 'unbound' to the soluble target. To avoid confusion in communications in ADC bioanalysis, we would prefer not to use 'free,' but rather 'unconjugated' or 'unbound.' If the term 'total' has to be used for simplicity in naming or labeling ADC species for PK analysis, it has to be defined clearly.

In addition to the four assays, there are other assays that may be used to answer specific questions related to safety, efficacy or disposition. The two dimensional plots presented in Figure 1 facilitate the understanding of the complexity and requirements. The horizontal axis describes conjugation/deconjugation, and the vertical describes bound/unbound to the soluble target. For both 'antibody' and 'payload,' the analytes can exist in four different forms *in vivo*: from conjugated and bound in the lower left quarter to unconjugated and unbound in the upper right. The conjugated-antibody (assay 2 as commonly referred [10]) exists in the upper left quarter, representing antibody conjugated to the payload but unbound to the soluble target. In clinical studies, we are also often asked to measure the total conjugated-antibody concentration, unbound and bound to soluble target in patient samples, which could correlate better to total efficacy. This total conjugated-antibody (could be referred as assay 2-II), a new assay, can be measured by either LBA or hybrid assay using appropriate capture and detection approaches

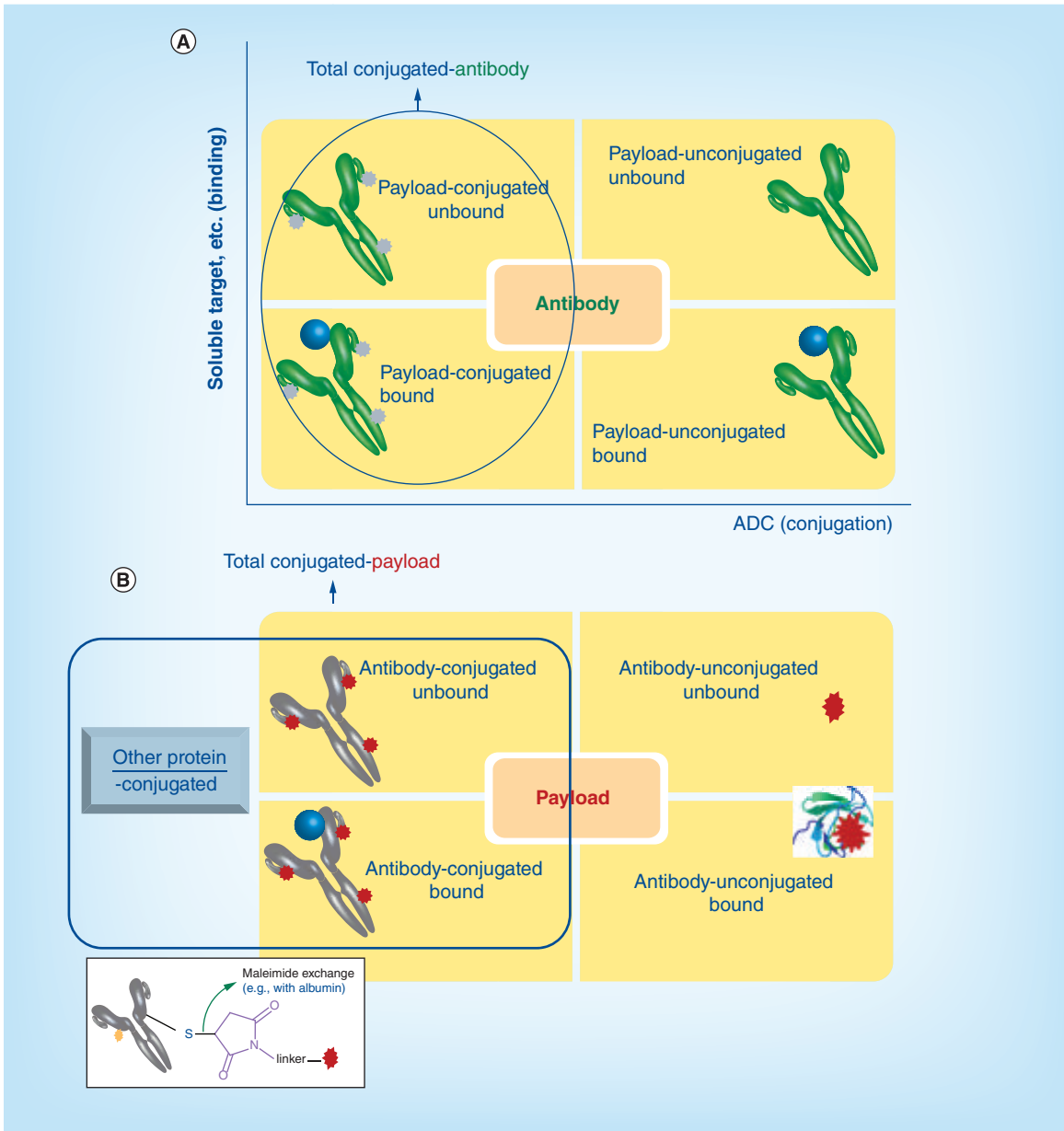


Figure 1. Complexity of antibody–drug conjugate analytes with soluble targets considered. (A) Total conjugated-antibody: conjugated-antibody unbound and bound to soluble targets. **(B)** Total conjugated-payload: payload conjugated to ADC mAb which are unbound or bound to soluble targets as well as payload conjugated to other plasma proteins. Gray dots in **(A)** and red dots in **(B)** represent payload; blue dots represent soluble target.

(reagents or MS-MS technique). Similarly, the analyte in the 'antibody conjugated-payload' (assay 3 as commonly referred [10]) is antibody linked to at least one payload and unbound to the soluble target. In some ADC chemistry, the payload can be conjugated to other plasma proteins through linker exchange [14,51–53]. Quantifying the payload conjugated to the drug antibody and all other plasma proteins could provide useful information for payload-dependent toxicity. This new analyte 'total conjugated-payload' can be measured by a new hybrid assay (could be referred as assay 3-II) in which an anti-payload capture reagent is used that can capture conjugated and unconjugated payload.

Development of LB-LC–MS/MS hybrid assays for ADC bioanalysis as complementary or alternative assays to LBAs

An immunocapture LC–MS/MS hybrid assay can be technically considered as an LBA with LC–MS/MS as the detector (or the immunocapture step can be considered as a specific extraction procedure for the highly specific mass spectrometer). No matter the perspec-

tive used, hybrid assays utilize the strengths of both approaches. Hybrid assays include three essential steps (Figure 2). The first step is immunocapture which is equivalent to the first step of a traditional LBA, the second step is enzyme cleavage or digestion and the last step is LC separation and MS/MS detection. Like in LBAs, a variety of specific capture reagents can be used including target antigens, anti-Id antibodies or anti-payload antibodies. Protein A and G can also be used as generic capture reagents. Biotinylated-antibodies can be immobilized on streptavidin coated magnetic beads or cartridges. In the case of the random-conjugated ADC of interest, enzyme digestion can be done with cathepsin-B to release the payload or by trypsin digestion to generate a signature surrogate peptide [48], which are then measured by LC–MS/MS. Various combinations of capture reagent and enzyme allows tailoring the assay for a specific analyte as described below.

Three automated immunocapture platforms in cartridge-format have been used in our routine operation: AssayMAP Bravo™ from Agilent, MSIA™ tips from Thermo and PhyTips™ from PhyNexus. Each hybrid method was optimized for immunocapture and enzyme cleavage or digestion steps. The cathepsin-B cleavage step was optimized for enzyme content, reaction time and temperature under vendor recommended pH and solution conditions. The trypsin digestion optimization and assay development are similar to that for mAbs or protein therapeutics. The difference is that a signature peptide without lysine is preferred or required for ADCs randomly conjugated at lysine residues. Validation of hybrid assays for ADCs need to fulfill the requirements for both small molecule LC–MS/MS and protein therapeutic LBAs and DAR sensitivity evaluation.

Currently hybrid assays have been developed in our laboratory for conjugated-payload, total-antibody and conjugated-antibody with different combinations of capture reagents and enzymes: anti-Id, anti-payload, cathepsin-B and trypsin (Figure 3 & Supplementary Table 1). The hybrid assays are labeled as H-1 through H-4 for ease of referencing in internal projects support. The correlations of these assays with commonly discussed assays 1–4 in literature [10] are also provided in Figure 3 in conjunction with the nomenclature of ADC analytes/assays discussed above. The anti-Id/cathepsin-B conjugated-payload assay (H-1) is the essential and primary application of hybrid assay complementary to (DAR-insensitive) conjugated-antibody measured by LBA [9,10]. The assay has been validated and used to support ADC preclinical and clinical studies [20,21,33]. Combining anti-Id capture and trypsin digestion measures total-antibody (H-3) as an alternative to the commonly used LBA.

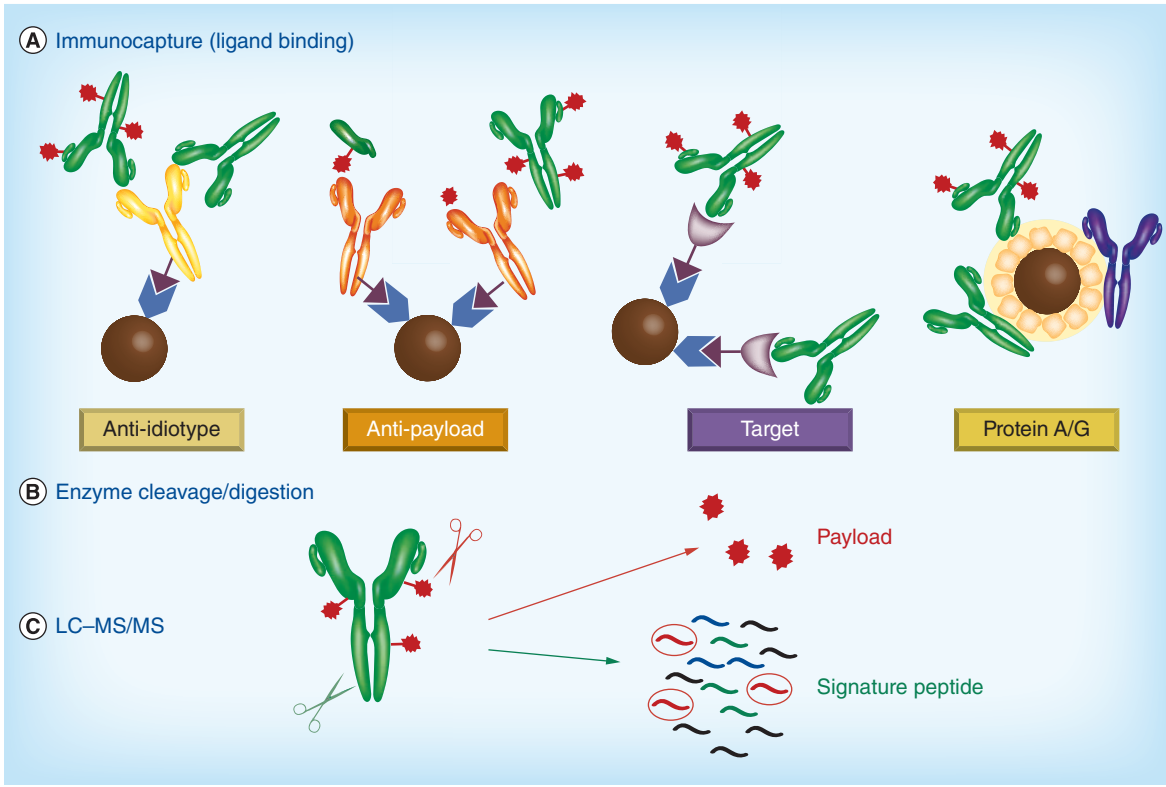


Figure 2. Procedure of LB-LC–MS/MS hybrid assays. (A) Immunocapture (ligand binding), **(B)** enzyme cleavage/digestion, **(C)** LC–MS/MS quantitation.

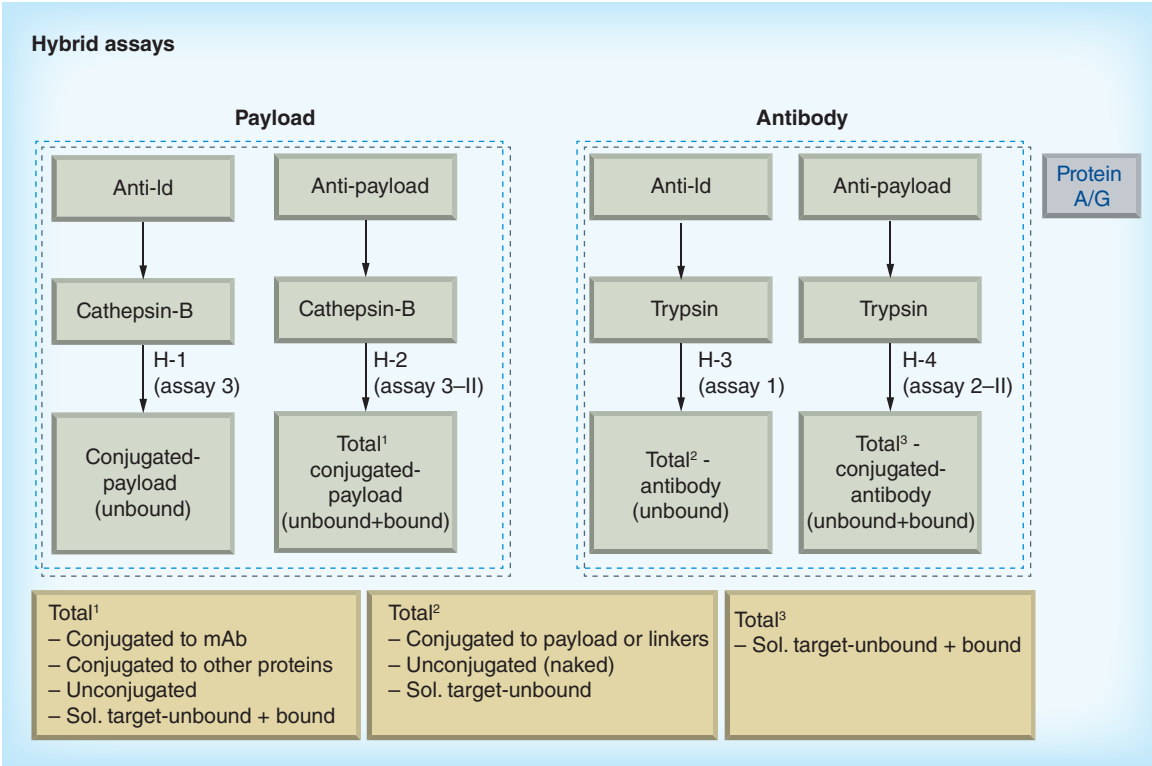


Figure 3. Hybrid assays developed with combinations of different capture reagents and enzymes. (Anti-Id, anti-payload, protein A/G, cathepsin-B and trypsin) for the analysis of conjugated-payload (H-1, assay 3), total conjugated-payload (H-2, assay 3-II), total-antibody (H-3, assay 1) and total conjugated-antibody (H-4, assay 2-II).

Anti-payload capture followed by cathepsin-B cleavage (H-2) or trypsin digestion (H-4) are advanced hybrid assays that quantify total conjugated-payload and total conjugated-antibody (bound and unbound to soluble targets), respectively. As discussed above, the total conjugated-payload assay (H-2), when needed, could provide insight to payload-dependent toxicity. Meanwhile, in an example discussed in the following section, the total conjugated-antibody hybrid assay (H-4) was used to analyze samples from a rat PK study for the ADC compound studied. Generic capture reagents such as protein A or G were also evaluated for conjugated-payload hybrid assays and results were compared with that from assays using specific reagents.

Performance & characteristics of the LB-LC–MS/MS hybrid assay for conjugated-payload

The conjugated payload was fully validated in cyno plasma, as previously reported [33], according to current FDA Guidance and EMA Guidelines on Bioanalytical Method Validation [54–57] and qualified for sample analysis in rat serum and human plasma. The experiments included the determination of the lower limit of quantification, accuracy and precision, selectivity, carryover, recovery, assay robustness and analyte stability. In addition, an ADC-specific validation experiment, the ability to accurately measure different DAR species was conducted. Here is a brief summary highlighting the key results.

- The overall recovery of the assay was 57%, with recoveries of immunocapture step and the enzyme cleavage efficiency of 78 and 73%, respectively;
- The calibration standard curve range was 0.303 to 151.442 ng/ml for conjugated payload, which was equivalent to the range of 20.0 to 10,000.0 ng/ml for the intact ADC. The assay demonstrated good linearity, accuracy and precision;
- Acceptable specificity and sensitivity were determined using ten individual lots of blank plasma unspiked and spiked at LLOQ and high QC levels;
- Conjugated payload was stable in monkey plasma for at least 24 h at room temperature, 38 days at -70°C and following five freeze–thaw cycles.
- The assay allowed accurate quantification of conjugated payload for the ADC having varying DAR. Supplementary Table 2 shows the DAR characteristics of conjugated-payload assay in cyno, human plasma and rat serum using anti-Id capture.

When anti-Id mAb capture is used, the conjugated payload assay measures ADC species that contain at

least one unbound mAb arm and is available to bind the target on tumor cells. Soluble target in circulation may block the binding sites on the ADC. In preclinical species, the soluble target may not be cross-reactive to the ADC, or the level of soluble target may be too low in relation to the ADC level to have a significant impact on the measurement, which was the case in rat in this example. Therefore, the assay tolerance to the presence of soluble target was not tested in animals. However, the levels of soluble target in the patient populations are projected to be high and variable, and, thus, the human assay was thoroughly characterized for the potential interference of soluble target across the estimated range of 10 to 1000 ng/ml. The results are shown in Figure 4. An acceptance criterion of 80% recovery was used to determine the significance of the impact. In the 'no impact zone,' soluble target levels had no impact on accurate quantitation of antibody-conjugated payload. However, relatively high levels of soluble target were found to interfere with the recovery of the analyte at low concentrations. In this 'interference zone,' a portion of ADC cannot be captured and quantified, presumably because both arms of the antibody are occupied by soluble target. Furthermore, due to this concentration-dependent impact of soluble target, a stable-isotope labeled ADC IS (SIL IS) may not be a good candidate for the measurement of conjugated payload in human samples as the IS binds to the soluble target as well, and its response vary in incurred samples and are different from that in calibration curves.

Specific versus generic immunocapture in conjugated-payload assays

Generic capture reagents such as protein A and G have been used for ADC conjugated-payload assays [10,34]. While a comprehensive comparison will be reported in a separate publication, Figure 5 compares the quantitation of conjugated-payload of the ADC of interest at ADC equivalent LLOQ concentration (20 ng/ml) in human plasma (containing lower amount of the soluble target) using different capture reagents anti-Id, protein A, Protein G in AssayMAP cartridges of 100 µg mAb capture capacity. The responses in the assays using Protein A and G are generally higher than that from the assay using anti-Id capture. This result indicates that, while in preclinical species without soluble target or human mAb cross-reactive soluble target, generic capture reagents (protein A and protein G) are comparable to the specific anti-Id capture reagent for LB-LC–MS/MS hybrid conjugated-payload assays in their ability to capture the ADC (data not shown), the higher recovery from a protein A or G in human plasma containing soluble target could be due to the fact that the conjugated-payload concentrations measured include the conjugated-payload, in which the antibody bound and unbound to soluble target (total) as discussed earlier, while anti-Id capture assay measures payload attached to the antibody unbound to the soluble target. The difference in concentrations from assays using generic and anti-Id captures provides information regarding 'free' and 'total' ADC, with regards to its binding to the soluble target or anti-drug

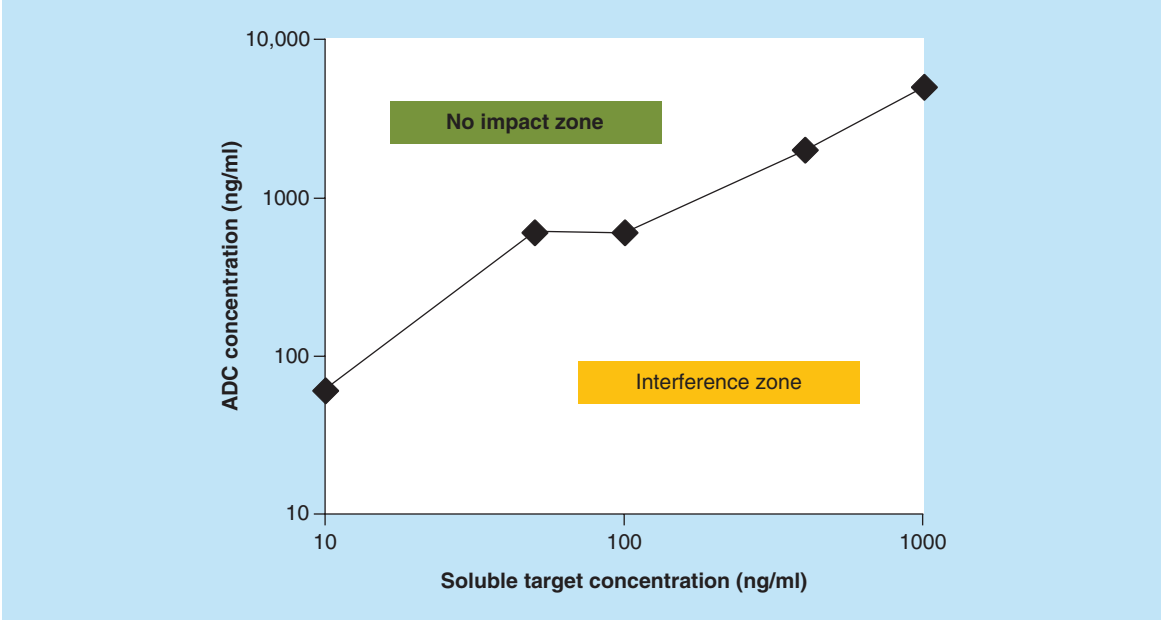


Figure 4. Tolerability to shed/soluble targets in conjugated-payload assay in human plasma.
ADC: Antibody–drug conjugate.

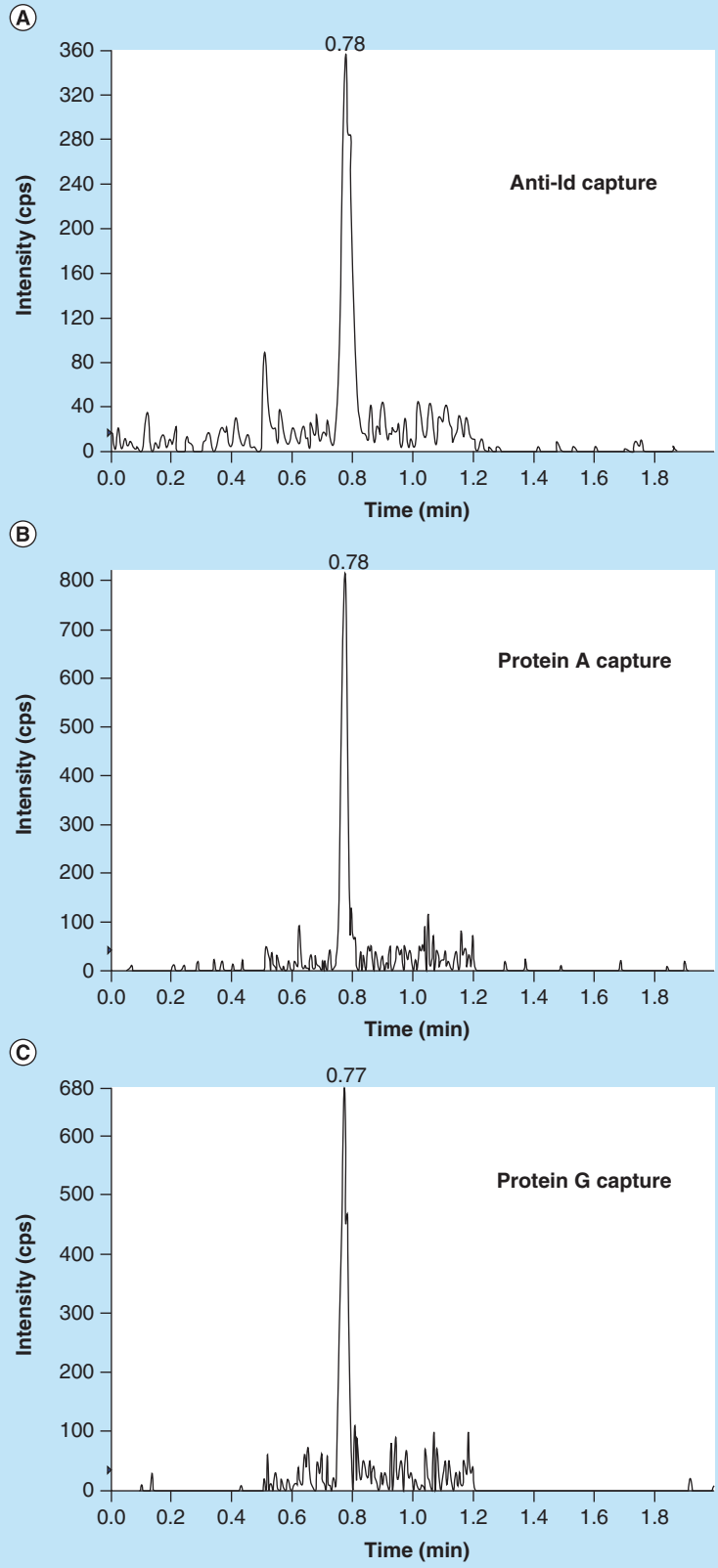


Figure 5. Chromatograms of conjugated-payload at antibody–drug conjugate equivalent LLOQ concentration (20 ng/ml) in human plasma (containing lower amount of soluble target) using different capture reagents.
(A) Anti-Id, **(B)** protein A, **(C)** protein G.

antibodies (ADAs) in clinical samples [49,50]. The recommendation is that generic capture reagents could be used in preclinical studies, especially in early Discovery, to screen antibodies and payloads when specific reagents (anti-Id and anti-payload) are unavailable; anti-Id capture reagent should be used in clinical assays if measuring soluble target unbound 'bio-active' ADC is desirable.

Performance & characteristics of LB-LC–MS/MS hybrid assays for total-antibody and conjugated-antibody

Hybrid assays with immunocapture performed on streptavidin and protein A cartridges, on an Agilent AssayMAP Bravo system, followed by trypsin digestion and LC–MS/MS analysis were developed and qualified for the quantitation of total-antibody and conjugated-antibody. During synthetic conjugation of the drug product, payloads are randomly conjugated to more than 70 lysine residues of the antibody for the ADC being studied. As lysine is one of the tryptic cleavage sties, to avoid interference of the linker to the trypsin digestion and LC–MS/MS analysis, a signature peptide without lysine, in the heavy chain CDR region, was identified and used for quantitation. Confirmatory peptides are often used in LC–MS quantitation of proteins [38–41]. They were not used in this work and will be considered in future hybrid assays for total-antibody and conjugated-antibody.

The assay performance was evaluated for both total-antibody and conjugated-antibody assays in different biological matrices, in other words, rat serum, monkey plasma and human plasma, primarily using cartridges on AssayMAP Bravo. Calibration curves ranged from 0.5 to 150 µg/ml in both assays. Supplementary Table 3 shows that the quality control samples met acceptance criteria in a conjugated-antibody assay in cyno plasma. Similar standard curves and quality control samples performance were observed for all the assays developed in other matrices. All assays achieved 0.5 µg/ml LLOQ, using 12.5 µl or less of biological sample. Assay LLOQ could be improved from 0.5 to 0.1 µg/ml by using 50 µl of biological sample. Assay specificity was tested for all assays using six different lots of matrix, and the results from the conjugated-antibody cyno plasma assay using streptavidin cartridges are shown in Supplementary Table 4. For the quantitation of total-antibody using anti-Id capture, in preclinical studies where soluble target interference is negligible, ideally, a stable-isotope labeled mAb with the labeling positions within the signature peptide would be used as the assay IS to track ADC analyte during the entire process of sample preparation, including immunocapture, trypsin digestion and LC–MS/MS analysis. As a labeled mAb was not available for the ADC of interest, a stable-isotope labeled signature peptide was used as the IS and added into the samples after trypsin digestion to track the analyte during LC separation, and MS ionization and detection. During the immunocapture and trypsin digestion process, where no IS was used, robust standard curves and QC performance was consistently demonstrated. For the total-antibody assay in human plasma using anti-Id capture, due to the presence of soluble target as discussed in the conjugated-payload assay, SIL IS of signature peptide is the choice for the assay. Meanwhile, the hybrid conjugated-antibody assay, using anti-payload capture theoretically measures conjugated-antibody unbound and bound to the soluble target which is a total conjugated-antibody assay and unaffected by the presence of soluble target.

Similarly to the approach used in the conjugated-payload assay validation, evaluation of DAR sensitivity or bias of the total-antibody and conjugated-antibody assays was accomplished by measuring enriched DAR 2 and DAR 4 standards at two concentrations against standard curves prepared with the DAR 3 reference material. No DAR bias was observed for the conjugated-antibody (Table 1) and total-antibody (Supplementary Table 5) hybrid assays using anti-payload and anti-Id capture reagents, respectively, when conducted as described above using cartridges on the AssayMAP Bravo platform.

Comparison of immunocapture using cartridges & magnetic-beads

In our experience, DAR sensitivity for LB-LC–MS/MS conjugated-antibody assays depends on the platform used for the immunocapture. For the ADC of interest, the hybrid conjugated-antibody assay using anti-payload capture was DAR-insensitive in the AssayMAP Bravo cartridge immunocapture platform but was DAR-sensitive in the magnetic-beads platform, which showed a bias for DAR 2 QC samples. While the exact mechanism is unknown and warrants further investigation, it was hypothesized that the improved recovery of different DAR species in a chromatography extraction column equivalent process during immunocapture in cartridges (vs mainly surface interaction in beads format immunocapture and plate format LBAs) could partially account for the absence of DAR bias. It is plausible that different DAR species have different thermodynamic equilibrium constant but have similar rate constant for forward binding reaction (Kon). In addition to this observed different DAR characterizes for anti-payload capture conjugated-antibody assays for the random-conjugated ADC of interest, other major differences between cartridges and magnetic-beads immunocapture for various analytes are capacity

Table 1. Drug to antibody ratio (DAR) characterization of conjugated-antibody assay in rat serum with anti-payload capture using cartridges on a AssayMAP Bravo: No DAR bias when DAR 2 and DAR 4 materials were measured against DAR 3 standard.

Conjugated-antibody (anti-payload capture)	Measured mean (μg/ml)	Adjusted mean† (μg/ml)	Adjusted nominal† (μg/ml)	CV (n = 3)	DEV%
DAR2 at 5 μg/ml	4.44	4.22	4.25	9.1	-0.8
DAR4 at 5 μg/ml	5.20	4.94	4.95	3.6	-0.3
DAR2 at 75 μg/ml	70.3	66.8	63.8	4.5	4.7
DAR4 at 75 μg/ml	68.6	65.2	74.3	8.1	-12.3

†Concentration was adjusted as DAR distribution of the standards follows Poisson distribution:
For ADC with DAR 2: 15% of naked antibody.
For ADC with DAR 3: 5% of naked antibody.
For ADC with DAR 4: 1% of naked antibody.

and the different degree of automation. A comparison of the characteristics between these two platforms of immunocapture is shown in Supplementary Table 6.

Applications of hybrid assays in preclinical PK studies

The random-conjugated ADC of interest was dosed in a discovery rat PK study designed to compare various ADC-related analytes measured by LBAs and hybrid assays. The PK profiles of five analytes measured by hybrid assays are presented in Figure 6A. These are total-antibody (red square, H3 assay), conjugated-antibody (blue diamond, H-4 assay) and conjugated-payload (blue cross, H-1 assay) which are three of the four commonly measured ADC analytes [10] (the unconjugated-payload was not measured here). As reported in earlier publications, the payload is inactivated by metabolism and there are different anti-payload capture reagents available that are specific to the payload or to both payload and the metabolite, respectively [20,21,58]. Payload and metabolite conjugated-antibody (green diamond, also an H-4 assay including the metabolite) was also measured which is labeled as conjugated-antibody (payload + metabolite) as well as conjugated-(payload + metabolite) (green cross, also an H-1 assay including the metabolite). It should be noted that, as discussed in the earlier section, that the H-4 hybrid assay using anti-payload capture can capture conjugated-antibody both unbound and bound to soluble target. However, the soluble target level in rats is negligible, and it does not cross-react with human mAb of ADC. Therefore, it is likely that the two H-4 assays only measure unbound conjugated-antibody or conjugated-antibody (payload + metabolite). The three antibody assays, total-antibody, conjugated-antibody and conjugated-antibody (payload + metabolite), were proven to be DAR-insensitive while the two payload assays were DAR-proportional. The ratio of conjugated-payload to total-antibody (average DAR, payload) and the ratio of conjugated-(payload + metabolite) to total-antibody (average DAR, payload and metabolite) across the PK time course were plotted in Figure 6B. This example demonstrated that hybrid assays are capable and sufficient to support ADC preclinical PK studies, highlighting the versatility of the LC–MS/MS platform for ADC support. The comparison of the average DAR from conjugated-payload and from conjugated-(payload + metabolite) (Figure 6B) shows that at late time points (>200 h), on average, about one third to half of the payload is deconjugated and the rest of the conjugated (payload related) species are combination of payload and metabolite in which the metabolite is about a third to half of the payload.

DAR-insensitive versus DAR-sensitive assays

One complicating factor and a center of debate in ADC bioanalysis, especially in the random-conjugated ADCs, is related to the *in vivo* DAR change. The drug substance is a mixture of ADC species of different DARs, typically 0 through 8 [9–11]. Unfortunately, most often, the individual DAR1–8 ADC reference standards are unavailable and only the drug substance with average DAR of 3 to 4 is used as the reference standard. As broadly reported in the literature, the DAR distribution *in vivo* is expected to change over time, following the administration of randomly conjugated ADCs, due to deconjugation, biotransformation and/or different clearance rates of the ADC species with different DARs [9–11]. The consequence for the quantitative bioanalysis is that the distribution of analyte species in the incurred samples is different from that in the calibrants. This is analogous to quantifying multiple metabolites of a small molecule drug using LC–MS/MS assays when the reference standards of the metabolites are unavailable and the reference standard of the parent compound is used to quantify all species. The mass spectrometric responses of the metabolites and the parent drug may not be the same leading to inaccurate measurements.

While the conjugated-payload assay is DAR-sensitive, DAR-proportional in fact, by definition, the total-

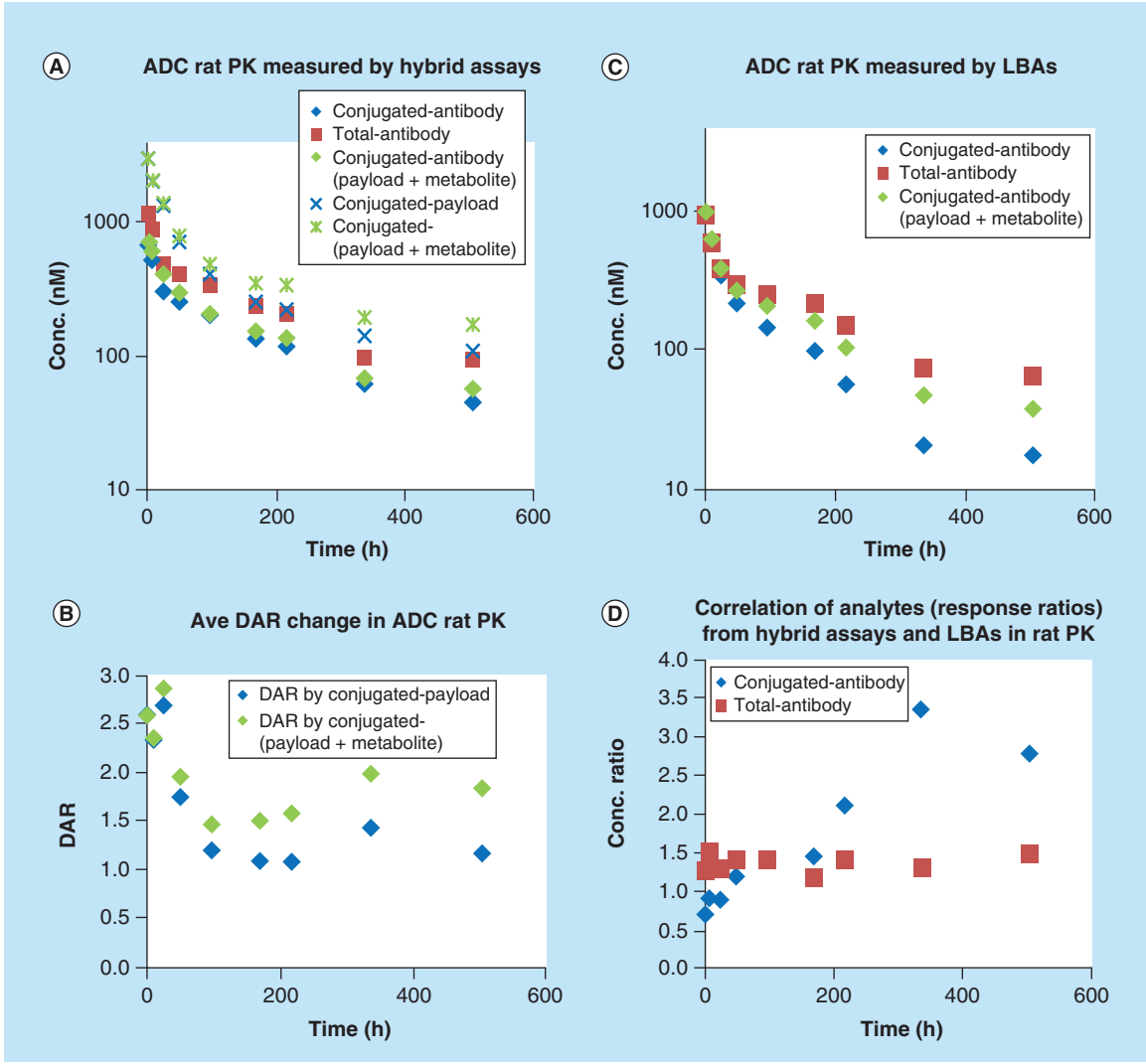


Figure 6. Hybrid assays and ligand-binding assay in antibody–drug conjugate rat PK. (A) Hybrid assays measured conjugated-antibody (blue diamond), conjugated-antibody (payload + metabolite) (green diamond), total-antibody (red square), conjugated-payload (blue cross) and conjugated-(payload + metabolite) (green cross). **(B)** *In vivo* average DAR change measured by hybrid assays, DAR by conjugated-payload (blue diamond) and DAR by conjugated-(payload + metabolite) (green diamond). **(C)** LBAs measured conjugated-antibody (blue diamond), conjugated-antibody (payload + metabolite) (green diamond) and total-antibody (red square). **(D)** Correlations of analytes (response ratios) from hybrid assays and LBAs. Conjugated-antibody (blue diamond) and total-antibody (red square).
Conjugated-antibody (payload + metabolite): payload and metabolite conjugated-antibody; conjugated-(payload + metabolite): antibody conjugated-payload conjugated to payload and the metabolite.
ADC:Antibody–drug conjugate; DAR: Drug to antibody ratio; LBA: Ligand-binding assay.

antibody and conjugated-antibody assays measure the antibody portion of ADCs and should ideally be DAR independent [9–11]. It has been widely reported and discussed that assay format, the capture and detection reagents used, and assay conditions affect the DAR characteristics of conjugated-antibody LBAs [9–11,19,28]. It is generally believed that the conjugated-antibody LBAs using anti-payload capture are more likely to be DAR-insensitive, but this may not always be the case and can be ADC molecules and reagents dependent. As an example, for the random-conjugated ADC of interest, the conjugated-antibody LBA developed using anti-payload capture reagents was DAR-sensitive. In comparison, the hybrid conjugated-antibody assays in Figure 6A using anti-payload capture in cartridge format immunocapture were proven to be DAR-insensitive as shown in Table 1.

In addition to demonstrating the capabilities of various hybrid assays in analyzing ADCs analytes (Figure 6A), the rat PK study was also aimed at evalu-

ating the correlation between LBAs and hybrid assays for ADC analytes, such as total-antibody and conjugated-antibody with assay DAR characteristics in the context. As shown in Figure 6C, the samples from the rat PK study were also analyzed for total-antibody, conjugated-antibody and payload and metabolite conjugated-antibody in the established traditional LBA methodology [20,21,58]. Several observations can be made when comparing Figure 6A & 6C and as shown in Figure 6D. One of the concerns of a DAR-sensitive conjugated-antibody assay, in comparison to a DAR-insensitive one, is the underestimation of the antibody concentrations at late time points (up to theefold when average DAR decreases from the nominal of 3 at early time points to about 1 at late time points as commonly seen for random-conjugated ADCs) [10,28]. This could affect the reported trough concentrations and exposure (AUC). Figure 6D shows that the concentration ratios of two conjugated-antibody (blue diamond) measured by DAR-insensitive hybrid in Figure 6A and DAR-sensitive LBA in Figure 6C increased from 1 at early time points to around 3 at later time points in the PK time course. The difference in AUC of the conjugated-antibody concentration versus time is about twofold (data not shown). This observation matched as reported by Stephan [28] and Kumar [19] in their LBA assay format evaluation and comparison of PK parameters. Figure 6D also shows the constant concentration ratios of total-antibody measured by the hybrid and LBA throughout the PK time course since both assays had the same DAR characteristics. It should be noted that this constant ratio ideally should be close to 1 showing the complete equivalency of the two assays measured by different assay platforms. It could be due to experimental errors or other unknown reasons that the observed constant ratio is approximately 1.3. As a result, the C_{max} from the hybrid total-antibody is higher than those from the hybrid conjugated-antibody. If the difference of concentrations at late time points of PK profiles between conjugated-antibody (payload + metabolite) (green diamond) and conjugated-antibody (blue diamond) is due to payload metabolism, the DAR-sensitive LBAs (Figure 6C) could have overestimated the degree of metabolism compared with the DAR-insensitive hybrid assays (Figure 6A).

Which conjugated-antibody assay is preferred, DAR-insensitive or sensitive? The question remains whether this threefold difference in trough concentration and twofold difference in AUC matter for the correlation between exposure and toxicity/efficacy if both DAR-insensitive and sensitive conjugated-antibody concentrations are used at different stages of ADC development [16]. Despite of the diversity of structures, heterogeneity of components present *in vivo* and their continuous change are common for ADCs [8]. In most publications on ADC bioanalysis including well cited reviews [10] and the AAPS position paper [11], DAR-insensitive conjugated-antibody assay is preferred. In a recent review paper on ADC bioanalysis LBAs, Kumar *et al.* [19] suggested to apply DAR-insensitive assays for IND toxicology studies with the rationale that this would generate better correlation between toxicology and exposure of all DAR species. We, then, may conclude that, in certain situations such as in this example, an LB-LC–MS/MS hybrid assay could be complementary to LBA, providing desired DAR-insensitive conjugated-antibody data.

Hybrid conjugated-payload assay versus conjugated-payload (DAR-sensitive conjugated-antibody) LBA

Three analytes were measured in a cyno PK study of the ADC of interest and the PK profiles are presented in Figure 7A, with conjugated-antibody (LBA) in blue, total-antibody (LBA) in red and conjugated-payload (hybrid) in green. The average DAR *in vivo* decreased from around 3, as in the dosed ADC material, to below 1 after about 1 week as is commonly seen for random-conjugated ADCs (Figure 7B). This indicates that, at later time points, on average, there is one payload attached to each antibody. As previously reported, conjugated-payload (green) and conjugated-antibody (blue) profiles were parallel throughout the PK time course with a ratio approximately equal to the starting average DAR of 3.

The results shows a good correlation between a DAR-sensitive conjugated-antibody LBA and a conjugated-payload hybrid assay for the quantitation of the active component of ADCs. The DAR-proportional conjugated-payload data validates the DAR-sensitive conjugated-antibody data [21]. Stephan stated that DAR-sensitive conjugated-antibody assays can be used to monitor *in vivo* DAR change when conjugated-payload assays are unavailable or for ADCs with non-cleavable linkers [28]. Kumar *et al.* [19] expressed that DAR-sensitive conjugated-antibody LBAs or LC–MS/MS based assays in early Discovery may be useful to better describe the changes in conjugated-payload over time and associated PK parameters. The argument is based on the hypothesis that conjugated payload is the main driver of efficacy at the site of action [19,25,59]. They further stated that ideally, a DAR-sensitive LBA would be equivalent to a conjugated-payload assay, as we observed in this example.

This suggests that the DAR-sensitive conjugated-antibody assay is not measuring the conjugated-antibody but rather is measuring the conjugated-payload. DAR-sensitive conjugated-antibody is expressed as

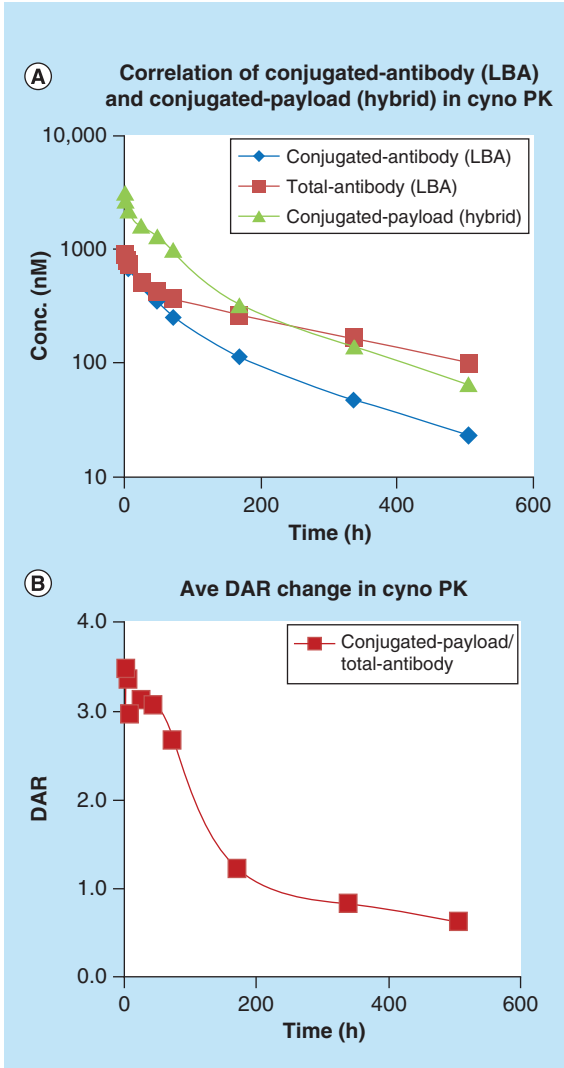


Figure 7. Correlation of conjugated-antibody (ligand-binding assay) and conjugated-payload (hybrid) in antibody–drug conjugate cyno monkey PK. (A) PK profiles of DAR-sensitive conjugated-antibody in ligand-binding assay (LBA) (blue diamond) which is equivalent to conjugated-payload in LBA, total-antibody in LBA (red square); conjugated-payload in hybrid assay (green triangle). **(B)** *In vivo* change of average drug to antibody ratio.
DAR: Drug to antibody ratio; LBA: Ligand-binding assay.

the equivalent concentration of ADC standard. Conjugated-payload measures the payload attached to the drug antibody and by definition is DAR-proportional and has clear PK and PD meaning. If a DAR-proportional assay is desired to monitor the payload change *in vivo*, then the conjugated-payload is the more direct assay. Changing the concentration unit of DAR-proportional conjugated-antibody (e.g., multiplying the nominal DAR value in the ADC dosing material) will convert the result to conjugated-payload. In that regard a perfect DAR-sensitive conjugated-antibody assay should be called a DAR-proportional assay, which is indeed a conjugated-payload assay; while DAR-sensitive assays refer to the assays which may not be able to stoichiometrically measure small molecule drug for all DAR species [19].

Since both assays are essentially the same, there is no need to conduct both assays in PK studies. Which conjugated-payload assay to use, DAR-sensitive LBA or hybrid assay? Hybrid conjugated-payload assay has less dependence on reagent availability especially in early Discovery. Generic reagents such as protein A and G can be readily available to be used to screen antibody, linkers, and payload to select the lead candidate. LBA conjugated-payload assay does require the anti-payload antibody. For the ADCs with noncleavable linkers LBA conjugated-payload assay is the assay of choice to monitor *in vivo* payload change. Post Discovery, i.e., IND enabling preclinical studies and clinical studies, if LBA conjugated-payload (aka DAR-proportional conjugated-antibody) is still desired, the assay can be processed in parallel with total-antibody and other assays in a harmonized and integrated fashion when same anti-Id capture reagent is used [20]. Alternatively, the fully established hybrid conjugated-payload assay, used since early Discovery can also be efficiently and effectively used to support ADC projects at later development stages with either generic or specific capture reagents.

Discovery & development integrated ADC PK/BA strategies: choice of ADC analytes & assays in discovery PK, development TK & clinical PK studies

Currently, there is no industry-wide consensus for ADC bioanalysis except for the fact that total-antibody, conjugated-antibody and/or conjugated-payload, and unconjugated-payload are recommended as the major analytes for ADC PK [9–11,19]. It is understood that each unique ADC molecule may require a specific combination of analytes using a fit-for-purpose approach. The key elements of Genentech's strategy [10,22] include investigating DAR distribution *in vivo* and testing all quantitative assays with individual DARs to ensure accurate quantitation of all ADC analytes in early discovery. Key PK analytes in nonclinical and clinical studies are total-antibody (LBA), conjugated-antibody or antibody conjugated-payload (hybrid assay) and unconjugated-payload (LC–MS). The conjugated-payload represents the active ADC drug component and is the preferred analyte when possible in a hybrid assay using protein A as the generic capture reagent. Pfizer proposed a stage-specific BA strategy [9,19] where a DAR-sensitive conjugated-antibody LBA was applied in early discovery and a DAR-insensitive con-

jugated-antibody LBA was preferred in development (IND-enabling and clinical studies). Pfizer scientists did acknowledge that evolution of assays could generate different PK profile and parameters at different stages of ADC development [19] as discussed above and reported by Genentech [28].

With limited experience on ADC performance in a clinical setting, it is still uncertain whether conjugated-antibody or conjugated-payload correlates more closely with efficacy and safety though it has been suggested that conjugated-payload is the key analyte for efficacy while (DAR-insensitive) conjugated-antibody is the preferred analyte for toxicology correlations [19]. To maintain the continuity of bioanalytical analytes and assays at different stages of ADC development, one strategy to be considered is that in early discovery, because of the efficiency, robustness and simplicity, a generic protein A or G capture conjugated-payload assay is ideal for screening and candidate selection. Generic capture total-antibody in LBA or hybrid format can be applied. In late discovery and early development, DAR-insensitive conjugated-antibody in LBA or hybrid format should be established if possible and both conjugated-antibody and conjugated-payload are measured. Assay performance in regard to DAR sensitivity should be evaluated at this stage using either enriched average DAR mixtures or purified individual DAR standards. It should be kept in mind that isolating individual DAR species for randomly conjugated ADCs is technically challenging and a fit-for-purpose assay strategy should be considered. Continuation with only one of the antibody-conjugate assays in late development is recommended, either a DAR-insensitive conjugated-antibody or a DAR-proportional conjugated-payload in either LBA or hybrid format depending on the technical strength of a particular bioanalytical lab and PK/PD correlations in early clinical studies. While currently there is a general belief that site-specific ADCs are the future of ADC development, and that ADC bioanalysis will be much simplified with reduced *in vivo* heterogeneity, we would like to caution that determining optimal conjugation sites for site-specific ADCs could be a challenging task. Either site-specific ADC shifts the bioanalytical focus and difficulties to different stages or randomly conjugated ADCs will continue to be one of the major directions of ADC development.

Conclusion

The complex nature of ADCs presents unique bioanalytical challenges and requires the measurement of multiple analytes. Both LBA and LC–MS/MS assays and their combination, LB-LC–MS/MS hybrid assays, are used in PK bioanalysis of ADCs. The use of 'LB' rather than 'affinity capture,' 'immunocapture' in the terminology of hybrid assays could facilitate the understanding of the nature of these assays and promote their flexible and fit-for-purpose applications. A series of hybrid assays, complementary or as substitutions for LBAs, were developed for a random-conjugated ADC, by combining anti-Id or anti-payload capture, and cathepsin-B or trypsin enzymes for the analysis of conjugated-payload, total-antibody and conjugated-antibody. Hybrid assays are capable of exclusively supporting ADC bioanalysis for PK studies. Generic reagents such as protein A and protein G have also been used in both the conjugated-payload and total-antibody hybrid assays. In general, hybrid assays using generic capture reagents are equivalent to those using specific anti-Id capture reagents in preclinical species when soluble target interference is negligible. On the contrary, in human plasms/serum samples, hybrid assays using protein A and G capture will measure analytes (conjugated-payload or total-antibody) in which the mAb unbound and bound to the soluble target while assays using anti-Id capture measures soluble target unbound ADC analytes. When considering which technology or platform to use in ADC bioanalysis and PK modeling, the comparison of LBAs and hybrid assays should include factors related to analytes being measured, assay DAR characteristic and platform availability. LBAs and hybrid assays can be complementary or alternative to each other in ADC bioanalysis. To be specific: DAR-proportional hybrid conjugated-payload assay is complementary to DAR-insensitive conjugated-antibody LBA and an alternative to DAR-sensitive conjugated-antibody LBA. DAR-insensitive hybrid conjugated-antibody assay is complementary to DAR-sensitive conjugated-antibody LBA. DAR-insensitive total-antibody hybrid assay and LBAs are alternatives to each other. The results of these assay investigations and comparisons demonstrate the possibilities of flexible and fit-for-purpose ADC bioanalytical assay strategies at different stages of ADC development and also tailored to the capabilities and preference of a particular bioanalytical laboratory with a focus on LBAs or LB-LC–MS/MS hybrid assays, and to the particular ADC constructs worked on. The strategy applied to ADC bioanalysis continues to evolve and industry-wide harmonization is desirable.

Future perspective

Diversified bioanalytical assays will continue to be used to support ADC PKs. More clinical data on a

variety of ADCs will help to decide which ADC analyte correlates more closely with safety and efficacy. We will see more applications of hybrid LB-LC–MS/MS conjugated-payload assays in Discovery to measure preclinical PK studies in addition to evaluating *in vitro* and *in vivo* linker and payload stabilities and payload metabolism using generic reagents. Either hybrid or LBA conjugated-payload (DAR-sensitive conjugated-antibody) assay can continue into Development to analyze regulated preclinical and clinical studies with a focus on efficacy correlation. As ADC toxicity can be both payload- and antigen-dependent, DAR-insensitive conjugated-antibody plays a significant role in establishing safety margins, and correlating preclinical and clinical ADC exposures. While LBAs have been the primary assays for total-antibody and (DAR-insensitive) conjugated-antibody, with more pharmaceutical companies and CROs working on ADCs and more experience industry-wide, and particularly because of the integration and close collaboration between LBA- and LC–MS/MS-based laboratories, there will be more applications of LB-LC–MS/MS hybrid assays for the analysis of these two ADC-related analytes. Immunocapture at the ADC and peptide-level coupled with micro- or nano-LCs will increase the sensitivity of these LB-LC–MS/MS hybrid assays.

Supplementary data

To view the supplementary data that accompany this paper please visit the journal website at: www.future-science.com/doi/full/10.4155/bio-2016-0017

Acknowledgements

The authors would like to thank the following for contributions to the assay strategy development, execution of the studies and experiments, and discussions: J Brailsford, J Zeng, R Iyer, X Gu, W Li, M Zhu, S Piccoli, D Passmore, D Desai, J Haulenbeek, F Zambito, S Gangwar, K Kelly, P Lemaire, A Davulcu, S, Brueggemeier, M Hay, C Yeung, A Tam, J Jackson.

Financial & competing interests disclosure

The authors of this article are current or past employees of Bristol-Myers Squibb Company (BMS). All financial support for the studies reported herein was provided by BMS.

Executive summary

- Antibody–drug conjugate (ADC) molecules are heterogeneous mixtures and *in vivo* have a continuous dynamic change of composition. Randomly conjugated ADCs have higher degrees of complexity. Although each unique ADC may require a specific group of analytes, in general, four analytes are commonly measured in ADC PK studies including total-antibody, conjugated-antibody, conjugated-payload and unconjugated-payload using LBA, LC–MS/MS and ligand binding (LB)-LC/MS/MS hybrid assays.
- LB-LC–MS/MS hybrid assays were developed to measure total-antibody, conjugated-antibody and conjugated-payload in ADC PK studies. These assays are capable of fully supporting ADC bioanalytical needs and may be considered as alternative or complementary to traditional ligand-binding assays (LBAs).
- The drug to antibody ratio (DAR) sensitive hybrid conjugated-payload assay measured different DAR species accurately against an ADC reference standard with average DAR of 3. The assay results correlated well with the LBA conjugated-payload (DAR-sensitive conjugate-antibody) assay.
- Hybrid total-antibody and conjugated-antibody assays were DAR-insensitive using anti-Id and anti-payload as capture reagents, respectively.
- The immunocapture in the cartridge and beads formats generated different DAR characteristics for hybrid conjugated-antibody assays using the same anti-payload capture reagent for a random-conjugated ADC studied. More investigations and comparisons are warranted.
- Currently there is no industry-wide standard strategy or procedures for ADC bioanalysis.
- While stage specific assay strategies are currently being discussed throughout the industry, a Discovery and Development integrated ADC bioanalysis strategy with the focus of assay continuity, flexibility and fit-for-purpose is proposed;
- Due to their independency on specific reagents, DAR-proportional hybrid conjugated-payload assays have an edge in early Discovery for compound screening and candidate selection.
- DAR-sensitive conjugated-payload in hybrid or LBA format can be continued into late Discovery and early Development stages while DAR-insensitive conjugated-antibody assays in either LBA or hybrid formats should be developed.
- Either DAR-sensitive conjugated-payload or DAR-insensitive conjugated-antibody assay could be carried into later Development once efficacy/toxicity and exposure relationship (of a specific analyte) has been established.
- ADC bioanalytical strategies continue to evolve. The choice of LBA or hybrid platforms for the quantitation of the major ADC PK analytes can be made on a case-by-case, fit-for-purpose fashion and is also dependent on the strength of a particular bioanalytical lab and its ADC bioanalytical philosophy.

The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.

Ethical conduct of research

The authors state that they have obtained appropriate institutional review board approval or have followed the principles outlined in the Declaration of Helsinki for all human or animal experimental investigations. In addition, for investigations involving human subjects, informed consent has been obtained from the participants involved.

References

Papers of special note have been highlighted as:
• of interest; •• of considerable interest

1 US FDA. FDA approves Adcetris to treat two types of lymphoma (2011). www.fda.gov

2 US FDA. FDA approves new treatment for late-stage breast cancer (2013). www.fda.gov

3 Goldmacher V, Kovtun Y. Antibody–drug conjugates: using monoclonal antibodies for delivery of cytotoxic payloads to cancer cells. *Ther. Deliv.* 2(3), 397–416 (2011).

4 Perez HL, Cardarelli PM, Deshpande S *et al.* Antibody–drug conjugates: current status and future directions. *Drug Discov. Today* 19(7), 869–881 (2013).

5 Beck A, Reichert JM. Antibody–drug conjugates: present and future. *MAbs* 6, 15–17 (2014).

6 Mullard A. Maturing antibody–drug conjugate pipeline hits 30. *Nat. Rev. Drug Discov.* 12, 329–332 (2013).

7 Kaur S. Bioanalysis special focus issue on antibody–drug conjugates. *Bioanalysis* 5(9), 981–983 (2013).

8 Gorovits B. Bioanlaysis of antibody–drug conjugates. *Bioanalysis* 7(13), 1559–1560 (2015).

9 Gorovits B, Alley SC, Bilic S *et al.* 2011 Bioanalysis of antibody–drug conjugates: American Association of Pharmaceutical Scientists Antibody-Drug Conjugate Working Group position paper. *Bioanalysis* 5(9), 997–1006 (2013).

• **Collaborative manuscript whereby many key opinion leaders across industry have addressed bioanalytical challenges for antibody–drug conjugate (ADC).**

10 Kaur S, Xu K, Saad OM *et al.* Bioanalytical assay strategies for the development of antibody–drug conjugate biotherapeutics. *Bioanalysis* 5(2), 201–226 (2013).

•• **High impact manuscript describing the potential ADC species and associated bioanalytical strategies employed to understand drug disposition.**

11 Stephan JP, Kozak KR, Wong WL. Challenges in developing bioanalytical assays for characterization of antibody–drug conjugates. *Bioanalysis* 3(6), 677–700 (2011).

12 Sauerborn M, Dongen W. Practical considerations for the pharmacokinetic and immunogenic assessment of antibody–drug conjugates. *Biodrugs* 28, 383–391 (2014).

13 Clark T, Han X, King L *et al.* Insights into antibody–drug conjugates: bioanalysis and biomeasures in discovery. *Bioanalysis* 5(9), 985–987 (2013).

14 Tumey LN, Rago B, Han X. *In vivo* biotransformation of antibody–drug conjugates. *Bioanalysis* 7(13), 1649–1664 (2015).

15 Saad OM, Shen BQ, Xu K *et al.* Bioanalytical approaches for characterizing catabolism of antibody–drug conjugates. *Bioanalysis* 7(13), 1583–1604 (2015).

16 Kraynov E, Kamath AV, Walles M *et al.* Current approaches for ADME characterization of antibody–drug conjugates: an industry White Paper. *Drug Metab. Dispos.* 44(5), 617–623 (2015).

17 Tibbitts J, Canter D, Graff R, Smith A, Khawli LA. Key factors influencing ADME properties of therapeutic proteins: a need for ADME characterization in drug discovery and development. *MAbs* 8(2), 229–45 (2016).

18 Beck A, Terral G, Debaene F *et al.* Cutting-edge mass spectrometry methods for the multi-level structural characterization of antibody–drug conjugates. *Expert Rev. Proteomics* 13(2), 157–183 (2016).

19 Kumar S, King LE, Clark TH. Antibody–drug conjugates nonclinical support: from early to late nonclinical bioanalysis using ligand-binding assays. *Bioanalysis* 7(13), 1605–1617 (2015).

20 Myler H, Rangan VS, Kozhich A *et al.* Validaiton of an integrated series of ligand binding assays for the quantitative determination of antibody drug conjugate in biological matrices. *Bioanalysis* 8(6), 519–531 (2016).

21 Myler H, Rangan VS, Wang J *et al.* An integrated multiplatform bioanalytical strategy for antibody–drug conjugates: a novel case study. *Bioanalysis.* 7(13), 1569–1582 (2015).

22 Dere R, Yi JH, Lei C *et al.* PK assays for antibody–drug conjugates: case study with ado-trastuzumab emtansine. *Bioanalysis* 5(9), 1025–1040 (2013).

23 Lin K, Tibbitts J. Pharmacokinetic consideration for antibody drug conjugates. *Pharm. Res.* 29, 2354–2366 (2012).

24 Shah D, Barletta F, Betts A *et al.* Key bioanalytical measurements for antibody–drug conjugate development: PK/PD modelers' perspective. *Bioanalysis* 5(9), 989–992 (2013).

25 Bender B, Leipold DD, Xu K *et al.* A mechanistic pharmacokinetic model elucidating the disposition of trasutuzumab emtansine (T-DM1), an antibody–drug conjugate (ADC) for treatment of metastatic breast cancer. *AAPS J.* 16(5), 994–1008 (2014).

26 Sukumaran S, Gadkar K, Zhang C *et al.* Mechanism-based pharmacokinetic/ pharmacodynamic model for THIOMAB™ drug conjugates. *Pharm Res.* 32(6), 1884–1893 (2014).

27 Khot A, Sharma S, Shah D. Integration of bioanalytical measurements using PK-PD modeling and simulation: implications for antibody–drug conjugate development. *Bioanalysis* 7(13), 1633–1648 (2015).

28 Stephan JP, Chan P, Lee C *et al.* Anti-CD22-MCC-DM1 and MC-MMAF conjuagtes: impact of assay formation on pharmacokinetic parameters determination. *Bioconj. Chem.* 19, 1673–1683 (2008).

29 Xu K, Liu L, Saad OM *et al.* Characterization of intact antibody–drug conjugates from plasma/serum *in vivo* by affinity capture capillary liquid chromatography-mass spectrometry. *Anal. Biochem.* 412, 56–66 (2011).

30 Xu K, Liu L, Dere R *et al.* Characterization of the drug-to-antibody ratio distribution for antibody–drug conjugates in plasma/serum. *Bioanalysis* 5(9), 1057–1071 (2013).

31 Luo Q, Chung HH, Borths C *et al.* Structural characterization of a monoclonal antibody-maytansinoid immunoconjugate. *Anal. Chem.* 88(1), 695–702 (2016).

32 Redman EA, Mellors JS, Starkey JA, Ramsey JM. Characterization of intact antibody drug conjugate variants using microfluidic CE-MS. *Anal. Chem.* 88(4), 2220–2226 (2016).

33 Liu A, Kozhich A, Passmore D *et al.* Quantitative bioanalysis of antibody-conjugated payload in monkey plasma using a hybrid immuno-capture LC–MS/MS approach: assay development, validation, and a case study. *J. Chromatogr. B* 1002, 54–62 (2015).

34 Sanderson RJ, Nicholas ND, Baker Lee C *et al.* Antibody-conjugated drug assay for protease-cleavable antibody–drug conjugates. *Bioanalysis* 8(1), 55–63 (2016).

35 Birdsall RE, McCarthy SM, Janin-Bussat MC *et al.* A sensitive multidimensional method for the detection, characterization, and quantification of trace free drug species in antibody–drug conjugate samples using mass spectral detection. *MAbs* 8(2), 306–317 (2016).

36 Li Y, Gu C, Gruenhagen J, Yehl P, Chetwyn NP, Medley CD. An enzymatic deconjugation method for the analysis of small molecule active drugs on antibody–drug conjugates. *MAbs* 8(4), 698–705 (2016).

37 Heudi O, Barteau S, Picard F, Kretz O. Quantitative analysis of maytansinoid (DM1) in human serum by on-line solid phase extraction coupled with liquid chromatography tandem mass spectrometry -Method validation and its application to clinical samples. *J. Pharm. Biomed. Anal.* 120, 322–332 (2016).

38 Neubert H, Muirhead D, Kabir M *et al.* Sequential protein and peptide immunoaffinity capture for mass spectrometry-based quantification of total human β-nerve growth factor. *Anal. Chem.* 85, 1719–1726 (2013).

39 Lund H, Lovsletten K, Paus E *et al.* Immuno-MS based targeted proteomics: highly specific sensitive and reproducible human chorionic gonadotropin determination for clinical diagnostics and doping analysis. *Anal. Chem.* 84, 7926–7932 (2012).

40 Whiteaker JR, Zhao L, Zhang HY *et al.* Antibody-based enrichment of peptides on magnetic beads for mass-spectrometry-based quantification of serum biomarkers. *Anal. Biochem.* 362, 44–54 (2007).

41 Schlosser G, Vekey K, Malorni A *et al.* Combination of solid-phase affinity capture on paramagnetic beads and mass spectrometry to study non-covalent interactions: example of minor groove binding drugs. *Rapid Commun. Mass Spectrom.* 19, 3307–3314 (2005).

42 Dufield D, Neubert H, Garofolo F *et al.* 2014 White Paper on recent issues in bioanalysis: a full immersion in bioanalysis (Part 2 – hybrid LBA/LCMS, ELN & regulatory agencies' input). *Bioanalysis* 6(23), 3237–3249 (2014).

43 Knutsson M, Schmidt R. LC–MS/MS of large molecules in a regulated bioanalytical enviroment-which acceptance criteria to apply. *Bioanalysis* 5(18), 2211–2214(2013).

44 DeSilva B, Smith W, Weiner R *et al.* Recommendations for the bioanalytical method validation of ligand-binding assays to support pharmacokinetic assessments of macromolecules. *Pharm. Res.* 20, 1885–1900 (2003).

45 Duggan JX, Vazvaei F, Jenkins R. Bioanalytical method validation considerations for LC–MS/MS assays of therapeutic proteins. *Bioanalysis* 7(11), 1389–1395 (2015).

46 Jenkins R, Duggan JX, Aubry AF *et al.* Recommendations for validation of LC–MS/MS bioanalytical methods for protein biotherapeutics. *AAPS J.* 17(1), 1–16 (2015).

47 Rao C, Rangan VS, Deshpande S. Challenges in antibody–drug conjugate discovery: a bioconjugation and analytical perspective, *Bioanalysis* 7(13), 1561–1564 (2015).

48 Lee JW. Generic method approaches for monoclonal antibody therapeutics analysis using both ligand binding and LC–MS/MS techniques. *Bioanalysis* 8(1), 19–27 (2016).

49 Lee JM, Kelley M, King LE *et al.* Bioanalytical approaches to quantify "total" and "free" therapeutic antibodies and their targets: technical challenges and PK/PD applications over the course of drug development. *AAPS J.* 13, 99–110 (2011).

50 Mayer AP, Hottenstein CS. Ligand-binding assay development: what do you *want* to measure versus what you *are* measuring? *AAPS J.* 18(2), 287–289 (2015).

51 Shen BQ, Xu K, Liu L *et al.* Conjugation site modulates the *in vivo* stability and therapeutic activity of antibody–drug conjugates. *Nat. Biotechnol.* 30(2), 184–189 (2012).

52 Alley SC, Benjamin DR, Jeffrey SC *et al.* Contribution of linker stability to the activities of anticancer immunoconjugates. *Bioconjug. Chem.* 19(3), 759–765 (2008).

53 Tumey LN, Charati M, He T *et al.* Mild method for succinimide hydrolysis on ADCs: impact on ADC potency, stability, exposure, and efficacy. *Bioconjug. Chem.* 25(10), 1871–1880 (2014).

54 US Department of Health and Human Services, Food Drug Administration, Center for Drug Evaluation and Research. Guidance for Industry: Bioanalytical Method Validation. Rockville, MD, USA (2013). www.fda.gov

55 European Medicines Agency. Guideline on Bioanalytical Method Validation. London, UK (2011). www.ema.europa.eu

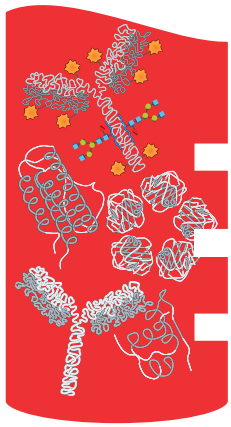
56 Japan Ministry of Health and Labor Welfare. Guideline on Bioanalytical Method (Ligand Binding Assay) Validation in Pharmaceutical Development (2014). www.nihs.go.jp

57 China Pharmacopoeia (ChP). Draft Guidance on Bioanalytical Method Validation (2014). http://knex2014.europeanbioanalysisforum.eu

58 Vangipuram SR, Myler H, Kozhich A *et al.* Biotransformation and stability of antibody–drug conjugates: payload metabolism and linker cleavage delineation. *Bioanalysis* 7(11), 1319–1323 (2015).

59 Wada R, Erickson HK, Lewis Phillips GD *et al.* Mechanistic pharmacokinetic/ pharmacodynamic modeling of *in vivo* tumor uptake, catabolism, and tumor response of trastuzumab maytansinoid conjugates. *Cancer Chemother. Pharmacol.* 74(5), 969–980 (2014).

thermoscientific



Revitalize Your Quantitation Workflows

The world of Pharma and Biopharma has evolved significantly in the recent past, and the domain of therapeutics is wider than ever before. The demand to achieve quality results faster while reducing cost continues to grow.

Today's world of quantitation warrants a complete workflow that provides increased confidence in analysis and quantitation for all molecule types, every time, and better than ever before. Thermo Fisher offers you the quantitation capabilities you've always aspired to have.

Discover more: **thermofisher.com/revitalize-quan**

ThermoFisher SCIENTIFIC

Special Focus Issue: Immunoaffinity MS

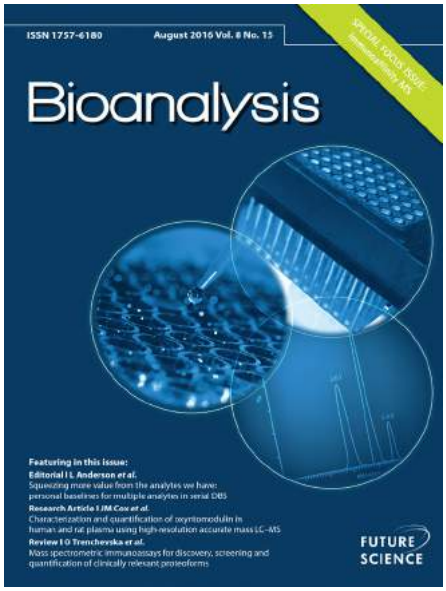
Bioanalysis published a Special Focus issue spotlighting the new developments in immunoaffinity-mass spectrometry (IA-MS) methods. The combination of IA sample preparation with MS detection allows analytical scientists to keep on track with advances in biological sciences and creates new opportunities for research.

"IA-MS has become a recognized analytical genre with demonstrated impact to both biomarkers and biotherapeutics. However, to achieve broader adoption it is essential that we remove barriers to implementation and focus on applications where hybrid methods outperform their component parts." said Guest Editor Bradley L Ackerman, from Eli Lilly and Company.

The issue includes a range of Editorial and Research Articles looking at utilizing the inherent value of IA-MS, and also the wider application of the technique, which may pose new challenges.

"We are excited to be publishing this Special Focus issue of Bioanalysis, which includes a number of research articles on novel applications of IA-MS technology" said Sankeetha Nadarajah, Editor of *Bioanalysis*. *"The issue aims to highlight the ongoing developments in these areas, challenges in adopting the hybrid technique and also provide a future outlook on how the field will evolve."*

To read the full issue, please click here – www.future-science.com/toc/bio/8/15



Bioanalysis

Hybridizing LBA with LC–MS/MS: the new norm for biologics quantification

"By combining traditional LBA immunoaffinity assays with the latest LC–MS/MS technologies, bioanalysts can achieve efficient immunocapture and sample extraction of large molecules, followed by highly sensitive and selective data analysis..."

First draft submitted: 5 January 2016; **Accepted for publication:** 12 January 2016; **Published online:** 26 February 2016

Keywords: biologics bioanalysis • hybrid LBA–LC–MS/MS • immunocapture-enrichment • LC–MS/MS • ligand-binding assay • protein therapeutics

Suma Ramagiri
Author for correspondence:
SCIEX, 500 Old Connecticut Path, Framingham, MA 01701, USA
suma.ramagiri@sciex.com

Ian Moore
SCIEX, 71 Four Valley Drive, Concord, ON L4K 4V8, Canada

Over the past 20 years, protein and peptide drug development has advanced significantly, becoming a rapidly growing segment of the biotherapeutics market. It has generated treatments for a wide variety of diseases ranging from inflammatory disorders to cancer [1,2]. This expansion of biologics-based medicines has pushed drug manufacturers to adopt new bioanalytical and analytical approaches to quantify and characterize multiple modalities of protein-based therapeutics efficiently with particular focus on the high-throughput needs of early-stage PK studies. Ligand-binding assays (LBAs) are widely used for bioanalysis of small and large molecules, but these assays can be lengthy to develop and have limited selectivity and sensitivity. LC–MS/MS-based bioanalysis can offer important advantages over immunoassay-based techniques, but LC–MS/MS with large molecules introduces several challenges affecting throughput, operating expertise and ease of use.

Bioanalytical techniques for large-molecule quantification

Traditional approaches for quantification of biologic drugs typically involve either LBAs, such as ELISA, or UV detection of individual peptides using HPLC separation [3,4]. One of the challenging parts of biologic drug quantification is developing a selective and sensitive assay, which can accurately measure and not over- or underestimate the drug concentration levels. This is very important during preclinical or clinical safety/efficacy testing. The gold standard has been LBAs but due to lack of specificity, this technique has been recently complemented with LC–MS. More recently, LC–MS/MS-based methods [5–7] have come to the forefront as a feasible approach for the quantification of various therapeutic peptides, proteins and monoclonal antibodies (mAbs) in biological matrices with many of these methods relying on proteolytic digestion of the target mAb and quantification of multiple unique signature peptides, which are equivalent to levels of the whole protein. The LC–MS/MS method complements the traditional LBA approach and provides the specificity and sensitivity necessary for overcoming LBA limitations, such as matrix interferences and time-consuming reagent development. Further incorporation of immunoaffinity-enrichment steps, which concentrate the target biologic drug prior to proteolysis, not only lowers the background complexity but also captures the active and circulating isoform of the drug, which makes the assay functionalized [8,9].

Pros & cons of ligand-binding & LC–MS/MS techniques for large-molecule bioanalysis

The main driver behind exploring alternate techniques is to increase the selectivity of an assay. Each assay platform, either LBA or LC–MS/MS, has demonstrated its strengths

FUTURE SCIENCE part of fsg

and limitations [10]. The ultimate end goal for any technique is to provide most reliable and accurate results. Traditionally, large-molecule PK assays have employed LBAs that rely on immunoaffinity detection of a unique epitope on the protein or peptide of interest. This approach can detect both physiologically active and 'free' forms of circulating large-molecule drugs, which makes the assay functionalized. Even though, the high specificity of the antibody-based interactions can track an analyte with high sensitivity but are typically expensive and time-consuming to develop. Specific antibody reagents to be developed for each mAb variant, a process that is often not compatible with the compressed time frames encountered during the initial stages of drug development. On the other hand, LC–MS/MS is widely used for highly selective and sensitive bioanalysis of small molecules. However, large-molecule bioanalysis using LC–MS/MS also presents challenges such as the need for extensive and tedious sample preparation and the expertise to operate complex LC–MS instrumentation including the software to process the data limits wider adoption of the technique. LBAs have limited dynamics and are subjected to limited selectivity and antibody cross-reactivity. This results in lack of specificity from interference and high-background levels that are not suitable for meeting the biopharmaceutical industry's requirements to detect diverse peptide, and protein therapeutics with ever-increasing sensitivity and reproducibility. LC–MS/MS assays can bring all the benefits of further selectivity with either multiple reaction monitoring (MRM) quantification or high-resolution quantification, in addition to the ability to measure multiple components and even biotransformation products such as catabolites of the drug [11]. Each technique shares opportunities and limitations to keep up with the ever-increasing demand for selective and sensitive biologics bioanalysis [12]. By combining traditional LBA immunoaffinity assays with the latest LC–MS/MS technologies, bioanalysts can achieve efficient immunocapture and sample extraction of large molecules, followed by highly sensitive and selective data analysis. Hybrid LBA–LC–MS/MS opens up new prospects to accelerate the drug discovery and development [13].

"The shift from small molecule to large molecule is happening exponentially, so are the techniques to characterize and quantify these molecules in the biological matrices."

Hyphenating LBA & LC–MS/MS opportunities

With the increasing complexity of multidomain therapeutic drug candidates (e.g., antibody–drug conjugates, pegylated proteins and fusion proteins), the specificity of LC–MS/MS detection combined with the selective enrichment advantages of ligand binding will help to fully characterize the PK/PD properties of these new drug entities. LC–MS/MS methods have the advantage of simultaneous detection of multiple peptides or molecules derived from a protein construct. This offers the opportunity to characterize the metabolism, catabolism and protein stability of both bioactive and inactive regions of the construct and can provide complimentary information to LBAs**.** Additional information complimentary to LBAs can be obtained by the ability of LC–MS/MS to quantitate multiple proteins in the same sample such as target and therapeutic or therapeutic and antidrug antibodies.

Hybrid LBA–LC–MS/MS assays for biologics bioanalysis

A recent example of the utility of combining ligand binding with the specificity of LC–MS detection was provided by Liu *et al.*, who described a hybrid immune-capture LC–MS/MS assay for the quantification of the payload molecules of an ADC [14]. The hybrid antibody-conjugated payload assay has three steps: immunoaffinity-enrichment with an anti-idiotype capture mAb, payload cleavage using a specific lysosomal enzyme and LC–MS/MS analysis of the cleaved payload. This combined approach provides direct measurement of the bioactive species, payload molecule linked to the mAb and not the other possible analytes, which include unconjugated payload and its metabolites and unconjugated mAb. A unique characteristic of multidomain peptide therapeutics is their heterogeneity, this is reflected in the drug–antibody ratio (DAR) of ADCs. The DAR *in vivo* is dynamic and changes over the time course of PK measurements, so that the calibration curve ADC reference material might not behave the same as the ADC with varying DAR. The authors showed their hybrid assay with direct measurement of the payload molecule could quantify ADCs in QC samples of DAR 2.0 and 4.0 using calibration samples prepared with a DAR of 3.0. An example of the complementary nature of a hybrid LBA–LC–MS/MS assay and LBA was provided by their characterization of the change in DAR over the time course of the PK measurement. The data were generated by using the ratio of the antibody-conjugated payload concentration from the hybrid assay to the total-Ab concentration from an LBA and showed that the average DAR decreased from 3 to 1 postdosing. This data, generated from multiple assays can give unique insights into the metabolism and disposition of the ADC that a singular assay could not. Finally, the multiplexing ability of MS detection was demonstrated by monitoring an inactive form of the ADC

that consisted of a metabolite of the payload still conjugated to the mAb. This study is an excellent example of the power of hybrid LBA–LC–MS/MS techniques to deliver a comprehensive bioanalytical strategy for ADC PK characterization [14].

The complementarity of integrating the data from orthogonal techniques was illustrated in a recent presentation by Krantz at EBF 2015 [15]. In a preclinical PK study of a PEGylated protein–drug candidate known to be proteolytically unstable, both an LBA and a multiplexing LC–MS/MS assay were employed to characterize its' PK behavior. The multiplexed LC–MS/MS assay simultaneously measured three peptides from distinct regions of the protein. The PK profile of the drug candidate as measured by LBA only correlated with the PK profile as measured by one of the three signature peptides. It was confirmed that the LBA results and the LC–MS/MS results of the correlating peptide represented only the bioactive portion of the drug candidate. The major species present in serum were truncated PEGylated forms of the protein which were represented by the assay of the other two peptides. The LBA on its own was not able to detect these metabolites. In a follow-up clinical study of the unconjugated therapeutic protein, a second LBA employing a different capture pAb was employed along with the three peptide multiplexed LC–MS/MS assay [15]. When examining the PK data correlation between the assays from the unconjugated protein study, it was discovered that the converse was true. The LBA PK profile correlated with the LC–MS/MS assay of the peptides representing the inactive region of the protein. In this example, LC–MS/MS was used to characterize the LBA and helped to determine what the LBA was measuring.

Conclusion

The shift from small molecule to large molecule is happening exponentially, so are the techniques to characterize and quantify these molecules in the biological matrices. There are many challenges in dealing with large-molecule drugs and one of the most important and a common challenge is a selective and specific bioanalytical assay. A Gold standard assay such as LBA still has its place given throughput and ease of use. But the expectation around patient safety and efficacy demands more selective assay platforms such as LC–MS/MS system, perhaps hybrid LBA–LC–MS/MS assay is the strategy to go. Comparable cost analysis [16] between different LBAs and LC–MS/MS assays makes it easier for scientists to put more emphasis in choosing scientifically sound and selective techniques for the study. Hybrid LBA–LC–MS/MS assay can bring all the benefits of further selectivity with either MRM quantification or high-resolution quantification in addition to the ability to measure multiple components and even metabolite's/catabolite's down the road.

Future perspective

Although LC–MS/MS technologies have certainly advanced to be more suitable for biologics bioanalysis, the array of MS technologies and techniques, tedious sample preparation methods and reagents could be daunting for nonexperts needing to develop a new biologics bioanalysis workflow. An integrated package for biologics analysis that includes sample preparation kits, automated immunocapture-enrichment process, MS technologies and standardized software with ready-to-use protocols makes the process easier to implement without compromising data quality. Significantly improved selectivity can also be achieved in biological samples that have numerous highly abundant protein interferences by using MRM [3] and differential ion mobility separation technology. Further advances in software and automation are being made to increase throughput, and enhance method development and optimization. The latest instrumentation and software developments will deliver significant improvements in the quality and reliability of bioanalysis studies, delivering more robust and compliant data that have important implications for patient safety. These developments are making LC–MS/MS more accessible for biologics bioanalysis, helping scientists to accelerate their large-molecule drug development.

Financial & competing interests disclosure

S Ramagiri is employed by SCIEX (USA) and I Moore is employed by SCIEX (CA). The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.

References

1 Ecker DM, Jones SD, Levine HL. The therapeutic monoclonal antibody market. *MAbs* 7(1), 9–14 (2015).

2 Fosgerau K, Hoffmann, T. Peptide therapeutics: current status and future directions. *Drug Discov. Today* 20(1), 122–128 (2015).

3 Damen CW, Schellens JH, Beijnen JH. Bioanalytical methods for the quantification of therapeutic monoclonal antibodies and their application in clinical pharmacokinetic studies. *Hum. Antibodies* 18(3), 47–73 (2009).

4 Roman J, Qiu J, Dornadula G, Hamuro L, Bakhtiar R, Verch T. Application of miniaturized immunoassays to

discovery pharmacokinetic bioanalysis. *J. Pharmacol. Toxicol. Methods* 63(3), 227–235 (2011).

5 Heudi O, Barteau S, Zimmer D *et al*. Towards absolute quantification of therapeutic monoclonal antibody in serum by LC–MS/MS using isotope-labeled antibody standard and protein cleavage isotope dilution mass spectrometry. *Anal. Chem.* 80(11), 4200–4207 (2008).

6 Li H, Ortiz R, Tran LT *et al*. Simultaneous analysis of multiple monoclonal antibody biotherapeutics by LC–MS/MS method in rat plasma following cassette-dosing. *AAPS J.* 15(2), 337–346 (2013).

7 Li H, Ortiz R, Tran L *et al*. General LC-MS/MS method approach to quantify therapeutic monoclonal antibodies using a common whole antibody internal standard with application to preclinical studies. *Anal. Chem.* 84(3), 1267–1273 (2012).

8 Ezan E, Dubois M, Becher F. Bioanalysis of recombinant proteins and antibodies by mass spectrometry. *Analyst* 134(5), 825–834 (2009).

9 Xu K, Liu L, Maia M *et al*. A multiplexed hybrid LC-MS/MS pharmacokinetic assay to measure two co-administered monoclonal antibodies in a clinical study. *Bioanalysis* 6(13), 1781–1794 (2014).

10 Zheng J, Mehl J, Zhu Y, Xin B, Olah T. Application and challenges in using LC-MS assays for absolute quantitative analysis of therapeutic proteins in drug discovery. *Bioanalysis* 6(6), 859–879 (2014).

11 Campbell JL, Le Blanc JY. Peptide and protein drug analysis by MS: challenges and opportunities for the discovery environment. *Bioanalysis* 3(6), 645–657 (2011).

12 Ezan E, Bitsch F. Critical comparison of MS and immunoassays for the bioanalysis of therapeutic antibodies. *Bioanalysis* 1(8), 1375–1388 (2009).

13 Zhang Q, Spellman DS, Song Y *et al.* Generic automated method for liquid chromatography–multiple reaction monitoring mass spectrometry based monoclonal antibody quantitation for preclinical pharmacokinetic studies. *Anal. Chem.* 86(17), 8776–8784 (2014).

14 Liu A, Kozhich A, Passmore D *et al.* Quantitative bioanalysis of antibody-conjugated payload in monkey plasma using a hybrid immuno-capture LC–MS/MS approach: assay development, validation, and a case study. *J. Chromatogr. B* 1002, 54–62 (2015).

15 Krantz C. LC-MS for large molecules vs. ligand binding assays: orthogonal readout or contradictory methods? Presented at: *EBF Open Symposium*, Barcelona, Spain, 18 November 2015.

16 de Dios K, Manibusan A, Marsden R, Pinkstaff J. Comparison of bioanalytical methods for the quantitation of PEGylated human insulin. *J. Immunol. Methods* 396(1), 1–7 (2013).

SHIMADZU
Excellence in Science

Switch + Go

Nexera UC/s allows measurements by liquid chromatography (LC) as well as supercritical fluid chromatography (SFC) on a single system. An increased range of compounds can be analyzed as LC and SFC offer very different selectivities for analytes of interest. Switching between LC and SFC permits rapid screening for optimum separation conditions, resulting in improved analytical efficiency.

Improved analytical results and efficiency
using two different separation techniques

Smaller footprint, reduced cost of acquisition
while benefiting from a full SFC/UHPLC setup

Automated workflow
to create LC/SFC screening sequence

Upgrade of existing LC to SFC functionality
without the need to buy an additional instrument

Nexera UC/s

www.shimadzu.eu/nexera-ucs

Spotlight: Hybrid LBA-LC–MS

Spotlight summary written by Carsten Krantz (Mannheim University, Germany)

Bioanalysis Zone focused on LBA/LC–MS, covering the development of LBA/LC–MS technologies, benefits of the techniques and more. In conjunction with this Spotlight focus we conducted a survey to gain insight into how you use this technology and the regulatory environment you work in. The survey addressed methodological practices followed in LBA/LC–MS, including acceptance criteria used for drug analysis and much more. In this Commentary, Carsten Krantz provides a summary of this recent survey.

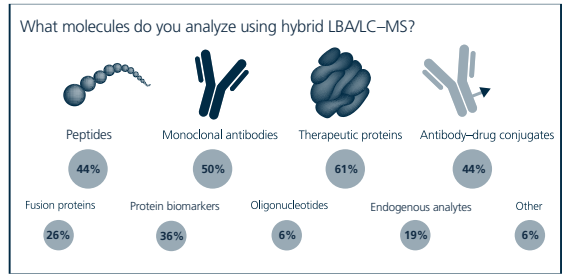
Ligand-binding assay (LBA) and LC–MS/MS technologies are widely used in the bioanalytical field for the analysis of small or large molecules. While LBA technology has been used for many years for quantification of proteins or biotherapeutics in biological matrices, LC–MS/MS has recently come more into focus as a valuable tool for the bioanalytical scientist to answer specific questions that were previously unaddressed.

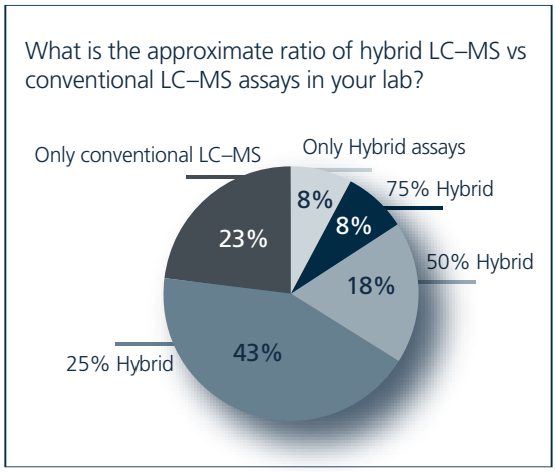
LC–MS/MS brings several advantages including short assay development times, high selectivity and specificity, less stringent need for critical reagents – although may suffer from a lack of sensitivity due to a highly complex matrix that interferes with sensitive detection of the surrogate peptide. In order to minimize this problem, scientists have combined the specific detection of a mass spectrometer with the highly selective functionality of a LBA.

In these immunoaffinity LC–MS assays (also known as hybrid LBA/LC–MS assays), a specific binding reagent coated to a solid matrix is used as a bait to specifically pull down the analyte of interest. Unbound small and large contaminants can be removed by washing and the analyte of interest is selectively captured and subsequently quantified in a mass spectrometer. Useful capture reagents for hybrid LBA/LC–MS assays include generic reagents that selectively capture species–specific immunoglobulin, target protein, anti-idiotype antibodies, binding proteins and receptors. Alternatively, the capture can be performed with an antipeptide antibody reagent when samples are predigested.

In order to gain further insights into the use of hybrid LBA/LC–MS assays among bioanalytical laboratories, Bioanalysis Zone recently conducted a survey. Of all the individuals who took part in the survey, the majority (72%) were from North America, 24% from Europe and 4% from Asia. Half of the participants work in pharma (36% in large companies and 14% in medium to small sized pharma companies), 23% come from a CRO or CMO, 13% work in the Biotechnology sector and 5% represent either an equipment vendor or are working in a different field (e.g., large biopharma, consumables and equipment vendors or agrochemical and seeds). Strikingly, only 4% of all participants belong to academia, highlighting that essentially all research and development in hybrid LBA/LC–MS is done with an industrial perspective.

Among responders who use hybrid LBA/LC–MS assays, 44% describe themselves as a manager or director and 20% have a leading position within their company (Director, CEO, VP). On the laboratory level, 32% of Scientists/Chemists but only 2% of technicians use these assays. When asked where hybrid LBA/LC–MS assays are applied the majority answered they are analyzing peptides (44%), mAbs (50%) or ADCs (44%) or therapeutic proteins (61%). Fusion proteins (26%), protein biomarkers (36%) endogenous analytes (19%) and oligonucletides (6%) were mentioned less

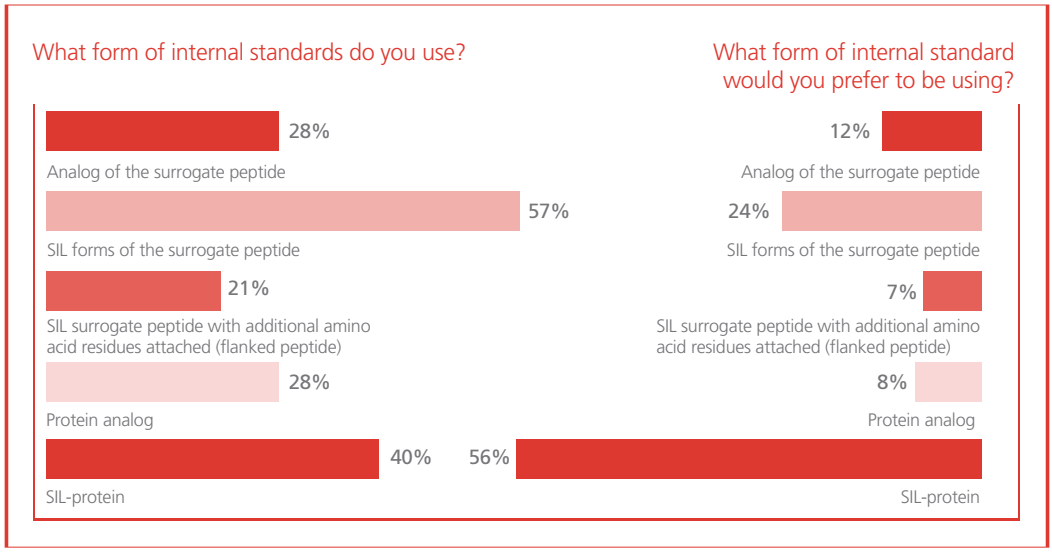


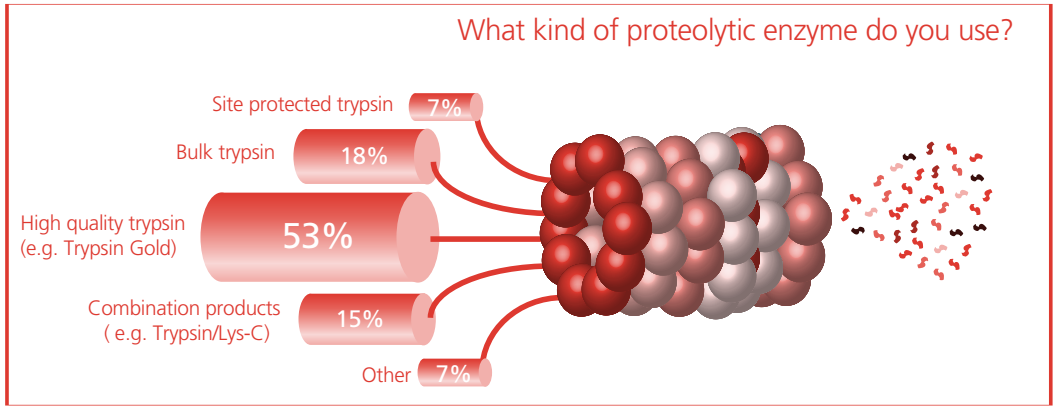


frequently. Interestingly, these analytes are measured predominantly in a discovery or preclinical setting (22% early discovery, 17% late discovery, 24 % preclinical IND enabling Tox studies). Approximately 30% of hybrid LBA/LC–MS are used in early clinical development and only 7% in late Phase III studies, indicating that hybrid LBA/LC–MS is still a relatively new assay format. It will be interesting to see if we will see more clinical applications as projects move through the pipeline.

The notion of a rather young technology that still needs to find its niche in the bioanalytical laboratory is reinforced by the survey result indicating that for many of the labs (43%) hybrid LBA/LC–MS represents only one-quarter of all LC–MS assays (hybrid LBA/LC–MS and conventional direct assays). Only 18% of the labs have half of their LC–MS assays in the hybrid format and only 8% use more than three-quarters or all their assays with hybrid LBA/LC–MS. Nevertheless, just 23% of labs use exclusively conventional LC–MS assays, indicating that hybrid LBA/LC–MS have been introduced in the majority of laboratories. The predominant application for hybrid assays is PK/PD studies to determine the drug (76%) or the drug target/biomarker (46%). Many survey responders (42%) indicated that they apply these assays for other drug-related studies in the DMPK environment, 14% use hybrid LBA/LC–MS in precision medicine and 21% to answer safety-related questions. Despite having many similarities and being applied for similar purposes, interestingly in only 56% of all cases are LBA and LC–MS capabilities under the same management.

Regarding the procedure for the hybrid LBA/LC–MS, a key issue to address is the internal standard. Forty percent of responders apply a SIL protein or, if available, a protein analog (28%; e.g., from another species) that can be spiked into the sample and is thus exposed to all assay steps. However, a protein internal standard is not always available and 57% of responders use a stable isotope-labeled form of the surrogate peptide or alternatively an analog of the surrogate peptide. Interestingly, only 21% use a stable isotope-labeled peptide with additional amino acids (i.e., flanking or winged peptide) that are removed during digestion. Apparently, it is believed that these extra amino acids do not have a major advantage compared with conventional SIL peptides.





The preference for a full SIL protein is reflected also in the answer that 56% of responders would select a SIL protein if available. Here, just 24% would prefer SIL peptides, only 12% would go with a peptide analog and only 8% or 7% of all responders would use a protein analog or a flanking SIL peptide. Depending on the binding epitope of the capturing reagent, the immunoaffinity capturing applied for hybrid LBA/LC–MS can be at the protein or the peptide level (SISCAPA approach). The majority of responders (62%) prefer the immunocapture step on the intact level prior to enzymatic digestion, only 5% choose an immunocapture approach after digestion. A so-called double hybrid LBA/LC–MS – a capture at the intact analyte level followed by digestion and subsequently a second capture on the peptide level – is preferred by 16%. Seventeen percent of responders specified that they usually use none of these steps, which could indicate the preference for a direct or conventional LC–MS assay.

Besides the capturing reagent, an important step for a hybrid LBA/LC–MS is the question of whether the captured proteins are digested directly on the beads together with the capturing reagent or if the bound proteins are first eluted and then digested. The latter usually gives higher sensitivity but introduces an elution step and a possible need for pH neutralization prior to the digestion, all steps where losses may be incurred. Nevertheless, 52% of all responders prefer to elute and digest, whereas only 26% prefer to digest directly on the beads. A small minority (7%) of responders go a step further and dry-down the eluate prior to digest. Fifteen percent of responders use none of the above mentioned steps.

Since a large molecule LC–MS assay is completely dependent on efficient release of the surrogate peptide, the enzyme used is clearly very important. The enzyme of choice is trypsin, but there are many different preparations of trypsin on the market that differ in their purity, residual activity of chemotrypsin, or auto-proteolytic activity. A majority (53%) of all responders use a high-quality trypsin e.g., Trypsin Gold. The unprotected lower grade (but more affordable) bulk trypsin is preferred by 18%. Fifteen percent prefer a combination product e.g., Trypsin/LysC and 7% either use site protected trypsin to prevent auto-proteolysis or some other enzyme source.

Similarly to small molecule LC–MS, triple quadrupole mass spectrometers combined with a conventional HPLC or UPLC are the most common used systems for hybrid LBA/LC–MS. The majority of all responders (64%) use these quantitation workhorses. High-resolution MS like Orbitrap or Q-TOF combined with these LCs are used only by 12% of survey responders. Fourteen percent attach a micro or nano LC to their QqQ and 10% use these LCs in combination with a high-resolution instrument. When it comes to acceptance criteria, 33% use the LBA 4-6-20 (25% at LLOQ) criteria. Around a quarter of all responders apply the small-molecule 4-6-15 acceptance criteria (20% at LLOQ). Thirty eight percent are not biased by either small molecule LC–MS or LBA criteria but rather use a case-by-case base tailored to the application of the assay. A crucial experiment to evaluate robustness of any bioanalytical assay is the ability to reliably remeasure study samples. For these so-called incurred sample reanalyzes (ISR), 34% of all survey responders apply the small molecule acceptance criteria (2/3 of all samples within 20% of the mean) and 31% use the LBA acceptance criteria (≤30%), and 35% respondents indicated that currently no ISR is performed.

When asked where participants see the main limitation of LBAs, 68% responded that the requirement of specific antibodies constitutes a major drawback. These tool antibodies usually require a substantial amount

of time to generate, which is a problem for 67% of survey responders. Additionally, 42% of responders highlight the limited selectivity and multiplexing functionality of most LBAs. Thirty percent believe that matrix interferences – which are often encountered in LBAs – constitute a major disadvantage.

Interestingly, only 18% indicated cost as a drawback of a LBA. When asked a similar question about disadvantages of conventional LC–MS assays for large molecules, the majority (70%) responded that an inherent problem of these assays is insensitivity when used without further enrichment tools. Forty two percent believe that the complex sample preparation of conventional LC–MS/MS assays is a main limitation. In the absence of an immunocapture step, the analyte measured is the 'total' concentration, which may or may not equate to the pharmacologically relevant concentration, and 37% of survey responders see this as a main limitation of conventional LC–MS/MS assays.

About a quarter of survey responders see the lack of regulatory guidelines (26%), or the required expertise to run such an assay (25%) as a limitation. Similarly 25% of survey participants believe that conventional LC–MS/MS assays are still a developing science. Lack of experience at CRO (14%) or limited throughput (18%) were also considerations. Only 4% stated uncertainty regarding setting acceptance criteria constitutes a problem.

Conversely, when benefits of hybrid LBA/LC–MS were discussed, 39% of responders indicate that improvement of sensitivity is a main driving force. Twelve percent mentioned that selectivity, or in some cases the more meaningful picture of drug levels when hybrid LBA/LC–MS are used, is important. Four percent emphasize the possibility to measure free drug and only 2% highlight the multiplexing possibility of hybrid LBA/LC–MS. Thirty percent believe that all these benefits are a main advantage to select hybrid LBA/LC–MS assays.

To improve the advantages of these assays even further, 77% of responders believe that kits or automated processes would help in the work. Finally, 91% of survey participants indicated that they believe the proportion of hybrid LBA/LC–MS will increase in the next 3–5 years and that quantification and quantitative analysis of target engagement for biologics will become a significant growth area for hybrid LBA/LC–MS (70%) further highlighting the strong potential of this technology in the upcoming years.

Key Points

- LBA is historically the gold standard for biologics and is employed extensively in their analysis. LC–MS, in contrast, is considered the gold standard for small molecule analysis. Until recently, LC–MS was reserved for niche applications in the area of large molecule bioanalysis.
- Both techniques have differing benefits and drawbacks: moreover, when run in parallel, the results obtained can vary. This has led to extensive discussion based on which result is 'correct'. However, the discussion recently has shifted to, 'how can we use the techniques orthogonally and, in some cases, simultaneously'.
- LBA/LC–MS is an example of the simultaneous use of both techniques and the collaboration between scientists with a LC–MS background and those with a LBA background.

Industry

Ask the Experts: Microsampling

In this 'Ask the Experts' feature we explored some of the recent developments in microsampling with leading experts in the field. For the past couple of years, there have been a number of exciting developments in microsampling. The development of techniques such as VAMS and solid phase microextraction (BioSPME) have transformed sample preparations and methodological practices.

In this feature we discussed some of the various microsampling techniques, including advantages and challenges faced when using these techniques. Check out what some of the key leaders thought about microsampling in the excerpt below:

How do you see microsampling techniques evolving over the next 5–10 years? What challenges do you think will remain or appear?

"Manipulation of microsamples remains a significant challenge for many labs. High throughput automated analysis of microsampling devices seems to be within reach. There is a significant trend towards automation, and microsampling workflows will be strongly affected by it. Completely automated workflows, from sample receipt to answer, are conceivable at this point and I think within 5–10 years they will be a reality."

Stuart Kushon (Neoteryx)

"The main challenge is full acceptance so that microsampling becomes the default approach for preclinical safety testing. The complete removal of satellite groups is the second challenge, as microsampling is currently primarily focused on exposure measurements…I envisage that over the next 5–10 years we will see a lot of innovation around the sampling technique with new technology being put forward, for example the MSW2 device from Shimadzu (Kyoto, Japan)."

Tim Sangster (Charles River)

"There will be continued advancements in the specificity of microsampling techniques. Microsampling devices containing antibodies for highly specific assays are within the foreseeable future. Evolution of microsampling with biosensors has the potential to allow rapid patient assessment in under developed countries that do not have access to advanced laboratory techniques."

Craig Aurand (Supelco/Sigma-Aldrich)

"From a technical perspective, I think there is scope for the further development of novel sample collection devices. From a study design perspective, I hope we will see the logistic simplicity of microsampling increasingly applied in late-stage clinical studies to improve PK data generation by 'decoupling' collection of blood samples from clinical site visits. This would allow for samples to be taken at home, possibly by the patients themselves. This could enable wider participation in PK aspects of late-stage studies, and allow for collection of samples at time points that would be impractical in a clinic setting."

Chris James (Amgen Inc.)

NEW INVESTIGATOR AWARD 2017

SPONSORED BY Waters

THE SCIENCE OF WHAT'S POSSIBLE.™

Congratulations to Matthew Lockett

Winner of the Bioanalysis Zone 2017 New Investigator Award, receiving a US$1000 prize and a year's subscription to *Bioanalysis!*

Meet Matthew and find out about his research by attending his presentation at 17.40pm, Wednesday 15th November

Each year, *Bioanalysis* and Bioanalysis Zone run the New Investigator Award (formerly known as the Young Investigator Award) to identify and reward promising early career researchers in our community.

Do you know an outstanding early career researcher?

Download our 2018 nomination form at
http://www.bioanalysis-zone.com/awards/new-investigator-award/

Bioanalysis ZONE

fsg CONNECT

www.bioanalysis-zone.com
info@bioanalysis-zone.com
+44 (0)20 8371 6090

Bioanalysis

9th GCC Closed Forum: CAPA in regulated bioanalysis; method robustness, biosimilars, preclinical method validation, endogenous biomarkers, whole blood stability, regulatory audit experiences and electronic laboratory notebooks

Roger Hayes[1], Richard LeLacheur[2], Isabelle Dumont[3], Philippe Couerbe[4], Afshin Safavi[5], Rafiq Islam[6], Colin Pattison[7], Stephanie Cape[8], Mario Rocci[9], Chad Briscoe[10], Laura Cojocaru[11], Elizabeth Groeber[12], Luigi Silvestro[13], Jennifer Bravo[2], Ron Shoup[14], Manon Verville[3], Jennifer Zimmer[15], Maria Cruz Caturla[16], Ardeshir Khadang[17], James Bourdage[18], Nicola Hughes[19], Saadya Fatmi[20], Lorella Di Donato[21], Curtis Sheldon[6], Anahita Keyhani[22], Christina Satterwhite[23], Mathilde Yu[24], Michele Fiscella[25], James Hulse[26], Zhongping (John) Lin[27], Wei Garofolo*[,28], Natasha Savoie[28], Yi Qun Xiao[29], Kai Kurylak[29], Sarah Harris[30], Manju Saxena[30], Mike Buonarati[30], Ann Lévesque[31], Nadine Boudreau[31], Jenny Lin[32], Masood U Khan[33], Gene Ray[34], Yansheng Liu[34], Allan Xu[35], Gunjan Soni[36], Ian Ward[37], Clare Kingsley[37], Hanna Ritzén[38], Edward Tabler[39], Bob Nicholson[39], Patrick Bennett[39], Nico van de Merbel[40], Shane Karnik[41], Mohammed Bouhajib[42], Jaap Wieling[43], Daniel Mulvana[44], Benno Ingelse[45], Mike Allen[46], Michele Malone[47] & Xinping Fang[48]

Note: Due to the equality principles of Global CRO Council for Bioanalysis (GCC), the authors are presented in alphabetical order of company name, with the exception of the first twelve authors who provided major contributions to topics discussed as the session chairs of the meeting (presented in alphabetical order).

*Author for correspondence: wei@global-cro-council.org
Author affiliations can be found at the end of this article.

9th GCC Closed Forum, Miami, FL, USA, 13 April 2015

The 9th GCC Closed Forum was held just prior to the 2015 Workshop on Recent Issues in Bioanalysis (WRIB) in Miami, FL, USA on 13 April 2015. In attendance were 58 senior-level participants, from eight countries, representing 38 CRO companies offering bioanalytical services. The objective of this meeting was for CRO bioanalytical representatives to meet and discuss scientific and regulatory issues specific to bioanalysis. The issues selected at this year's closed forum include CAPA, biosimilars, preclinical method validation, endogenous biomarkers, whole blood stability, and ELNs. A summary of the industry's best practices and the conclusions from the discussion of these topics is included in this meeting report.

First draft submitted: 8 January 2016; Accepted for publication: 26 January 2016; Published online: 26 February 2016

Keywords: biomarker • biosimilar • blood partitioning

The 9th Global CRO Council in Bioanalysis (GCC) Closed Forum took place on 13 April 2015, preceding the 9th Workshop on Recent Issues in Bioanalysis in Miami, FL, USA. The GCC was founded in 2010 to allow senior management level CRO rep-

resentatives to share experiences and opinions related to scientific and regulatory issues in bioanalysis [1]. Eight prior closed forums have periodically occurred since the founding of the GCC, typically associated with other major bioanalytical conferences, to discuss a variety of topics and issues of particular interest to bioanalytical CROs. Conference reports describe the discussions, allowing a glimpse into the opinions and challenges faced by this subsection of the industry [2–6]. Additionally, several White Papers have been published by the GCC on topics of particular widespread interest [7–12]. These White Papers provide recommendations that are meant to be applicable to the global bioanalytical community.

Roger Hayes, chair of the 9th GCC, began the meeting in traditional fashion, with the official admonition statement [1], which instructed participants on the guidelines to follow during the discussions. Following this, all participants introduced themselves and then this year's discussion topics were presented. The eight topics for the 9th GCC Closed Forum were as follows:

- The role of CAPA in regulated bioanalysis.
- Method robustness in regulated bioanalysis.
- Best practices for supporting biosimilar studies using ligand binding methods.
- Status of method validation requirements in the preclinical space.
- Best practices for measuring endogenous biomarkers beyond prior GCC recommendations.
- Whole blood stability and blood partitioning.
- Regulatory audit experiences.
- Electronic laboratory notebooks and LIMS systems.

Discussion topics

The role of CAPA in regulated bioanalysis

A robust Corrective Action and Preventive Action (CAPA) program is required in a Chemistry, Manufacturing and Controls (CMC)/Good Manufacturing Practices (GMP) environment in accordance with the Q10 GMP guidance document [13]. The core value of such a program is to determine the root cause of any quality issue resulting from an investigation into a nonconformance, complaint or other undesirable situation. Once determined, the root cause is addressed by a two-pronged approach. The first is the corrective action, which is implemented to address the cause of a detected issue. The second is the preventive action, which is implemented to prevent the detected issue from recurring.

In the context of manufacturing drug products, CMC/GMP principles outlined in the Q10 document are applied to ensure that drugs are consistently produced and controlled in such a way as to meet the quality standards appropriate for their intended use, as required by the marketing authority. The CAPA program, when implemented properly, should result in a greater understanding of the product and processes and, therefore, their improvement. Part of the regulatory requirements of the program includes complete and thorough documentation of the entire process involved in a CAPA, which typically includes seven basic steps:

- The **identification** of the problem/potential problem, nonconformity or incident;
- An **evaluation** of the magnitude of the problem and potential impact on the company;
- The development of an **investigation** procedure with assignments of responsibility;
- Performing a thorough (root cause) **analysis** of the problem with appropriate documentation;
- Creating an **action plan** listing all the tasks that must be completed to correct and prevent the problem from recurring;
- The **implementation** of the plan;
- A thorough **follow-up** with verification of the completion of all tasks, and an assessment of the appropriateness and effectiveness of the actions taken.

Based on feedback from sponsors to CROs, there appears to be an expectation that GLP/regulated bioanalytical laboratories meet these GMP requirements despite the fact that they are out of scope for GMP regulations. There is a fundamental difference between GMP and GLP, however, which gives pause to the CROs when trying to decide how best to fulfill the expectations of their clients. At a high level, GMP is concerned with the integrity and quality of a manufacturing product intended for human use, whereas GLP protects the integrity and quality of laboratory data used to support a product application.

There are undisputed benefits to implementing a CAPA program in bioanalysis, including better control of processes and methods, better understanding of issues, a decrease in the risks of failures and an increase in data accuracy and reliability. Conversely, there are potential drawbacks if a CAPA program is implemented inappropriately. These include additional work and costs as well as inadequate investigations leading to a waste of resources (human, time, finan-

cial) with no significant added value to the generated bioanalytical data.

In an effort to understand how and if CROs are electing to implement CAPA programs, a survey was sent to CRO member companies asking to share their approaches and philosophies. A majority of the respondents indicated that a CAPA program is pertinent to regulated bioanalysis, but should be judiciously applied and is not intended for all quality issues. The results indicate that most respondents believe that the quality issues that should be addressed by a CAPA program are those that are recurrent or systemic in nature, not isolated, or those that have significant impact on the bioanalytical data being generated.

The discussion with the attendees seemed to support these survey results. There was agreement that while the idea of a CAPA program is a good one, the extent of a GMP-like program is not required, and the investigation of quality issues should be outlined in a written procedure or Standard Operating Procedure (SOP). There is an assumption in GMP that methods and processes are designed to never fail; however, in GLP, tolerance limits are built into the methods, and failures do not have the same impact. For example, the instability of an analyte during a sample extraction procedure requires a method adjustment, not necessarily a full investigation into the cause of the stability issue. It was noted that most attendees perform trending at their companies in order to decide whether a CAPA approach to the remediation of quality issues is appropriate. Some examples of issues that may require a CAPA approach can include significant documentation lapses, lost samples or chronic contamination issues.

In conclusion, in a regulated bioanalytical CRO environment, CAPA elements should be thoughtfully applied by considering a balance between expected benefits, resources needed and potential risks for the organization. The CAPA approach should be tailored and adapted to fit the area of regulated bioanalysis. Each quality issue should be evaluated on a case-by-case basis prior to applying CAPA; the impact, frequency and significance of the issue should drive the decision to implement CAPA (i.e., risk-based CAPA).

Method robustness in regulated bioanalysis

Another topic discussed was that of 'GMP creep' into regulated bioanalysis and the expectation of how method robustness should be evaluated. The ICH Q2(R1) guideline for validation of GMP analytical methods [14] refers to small but deliberate variations in the method that could impact the reliability of the results. This is further referenced in the 2015 US FDA guidance document [15], which calls for a systematic approach for method robustness. However, the notion of evaluating the variations in the bioanalytical method is not excluded from the international Bioanalytical Method Validation guidelines [16–19]. All four documents speak of performing precision and accuracy evaluations on different days. The FDA guidance further recommends performing the tests with different analysts, equipment, reagents and laboratories. The EMA recommends testing at least one run in a size equivalent to the anticipated study sample runs. However, GMP validation parameters suggest showing method reliability with many more variables, including pH of the mobile phase, temperature, flow rate, etc.

CRO member companies were surveyed to determine their approach to method robustness. The majority of respondents indicated that method robustness in regulated bioanalysis should be evaluated and so there was further inquiry into the variables that were tested. Most believed that validation runs should be performed on different days. Different analysts would only be required when the size of the study and the intended number of analysts used during the application of the method warrants it (i.e., such an assessment may not be necessary or applicable for small studies or when only a single analyst is involved). The use of different columns should be evaluated only if the need for multiple columns is anticipated (e.g., analysis of large numbers of study samples and use of multiple chromatographic systems). If the use of different lots of critical reagents is expected in ligand binding assays (LBA) or hybrid LBA/LC–MS methods, then this variable may need to be tested. There was no consensus on the evaluation of run size. Some suggested that only the largest anticipated run size should be tested, while others believed that because variable run sizes were inevitable during method validation, any additional tests were not required. Still others believed it depended on how many study samples were anticipated during the use of the method. Finally, respondents were asked if any other variables should be evaluated. Many mentioned that there might be other method specific parameters that may need to be investigated, but that this was method dependent.

In conclusion, method robustness testing in bioanalysis confirms that method performance will not be affected by variations in some parameters. The parameters to include in the testing should take into account the assay and the conditions expected during sample analysis and should be tiered based on the purpose of the method.

Best practices for supporting biosimilar studies using ligand binding methods

Regulatory agencies have provided guidance documents relevant to biosimilars, since these medicinal

products have been submitted with increasing frequency [20,21]. However, despite these guidance documents, the relative infancy of these types of submissions has resulted in numerous, inconsistent approaches to meeting the recommendations requested by the agencies for PK assays, immunogenicity assays and critical reagent characterization. In order to determine how the industry is applying these recommendations and guidances, survey questions were sent to the GCC member companies. The first set of questions inquired after PK assays used to support biosimilar studies. The results for this type of assay indicate that the majority of respondents use one assay to measure both the innovator and the biosimilar drugs. This is based on the assumption that both drugs are identical, which is often not the case. Therefore, care should be exercised to ensure that this is the best approach. Should two assays be needed, samples should be run with both assays and bridging assays are needed to be able to appropriately evaluate the results. Calibrators are typically prepared with the biosimilar; however, QC samples during both method validation and sample analysis are prepared with both the biosimilar and the innovator test articles. Method comparability is demonstrated by using back-calculated concentrations (interpolated from one curve) of one set of QCs against the other with acceptance criteria of ±20%. Differences in concentrations between biosimilars and reference (innovator) products need to be investigated. It was noted in the discussions that correction factors may not be considered acceptable by the regulatory agencies.

Multiple lots of both reference and biosimilar drugs are typically evaluated during method development. Key reagent characterization, including purity, purification method and formulation, is performed for both the biosimilar and the innovator; however, the binding characteristics of the capture/detection reagents to the biosimilar and reference product are not typically evaluated. Relative potency testing is performed only if there are differences between the protein determination methods.

Regarding immunogenicity assays, respondents were evenly divided on the use of one or two assays. For those who use one assay, respondents use the biosimilar as the drug and use only one positive control. In the one assay format, evaluating the confirmatory cutpoint by using only the biosimilar is common practice for the majority of respondents. For those that use two assays, the acceptance criterion for proving biosimilarity is results within 2–3-fold for sensitivity, cutpoint and drug tolerance. For neutralization assays for biosimilars, cell-based neutralizing antibody assays are typically selected although ligand binding assays can be used.

Status of method validation requirements in the preclinical space

The GCC membership was surveyed regarding the different approaches to bioanalytical method validation used under differing circumstances, and the confusion surrounding the use of the popular term 'GLP validation.' GLP regulations detail quality systems, controls and documentation and do not specifically address bioanalytical method validation requirements. These are defined in guidance documents and regulations [16–20].

Therefore, initial survey questions attempted to determine how industry labeled GLP and GCP bioanalytical work. The majority of respondents use the term 'Regulated Bioanalysis,' and less than a third distinguish between clinical and preclinical bioanalytical validation. It should be noted that approximately 60% of respondents stated that they performed more than 75% clinical work. Of this group, less than 25% of the clinical work were bioequivalence (BE) studies. The attendees discussed at length whether GLP regulations should be used to support clinical studies, even though they fall out of scope. There was no consensus, although the use of Good Clinical Laboratory Practices [22] was suggested as an alternative approach. This may be limited by individual regulatory authorities' use of GCP inspections for laboratories supporting clinical trials, however.

For those that perform a variety of study types (clinical, preclinical, BE, non-BE), the majority stated that validation testing is not tiered based on the type of study. Further, regulatory audit experience seems to demonstrate that auditors tend to audit against full GLP compliance, not using a tiered framework, although they do also accept this latter approach.

The use of a tiered approach to method validation was evaluated via the survey. The responses indicate that there are two situations that favor its use: 1) analytical challenges preclude it; and 2) the end use of the data does not require full validation. Attendees noted that sponsors are generally hesitant to use methods that have not undergone full validation. When a tiered validation is used, there is no consistent language used to describe it, and a variety of terms are employed; for example, partial validation, qualification, exploratory method. Although no consensus was reached at the closed forum, attendees did desire that harmonized language for these types of methods should be agreed upon. The use of tiered validation for preclinical studies seemed to be favorably accepted, but was much less acceptable when used for clinical studies.

Several scenarios were proposed to identify when respondents believe that using a tiered validation would be appropriate. A tiered approach was supported by the majority of respondents for at least the secondary

study endpoint, if the situation warrants and regardless of human or animal species. One example would be if rare matrices were being assayed (e.g., tissues or cerebrospinal fluid). For common matrices, respondents were more comfortable using a fully validated method. Other potential situations for using a tiered approach would be for biomarkers or minor, inactive metabolites.

Best practices for measuring endogenous biomarkers beyond prior GCC recommendations

Down-regulated biomarkers are endogenous substrates that, upon treatment, result in lower concentrations. Therefore, potentially very low concentrations are expected during the quantification of these analytes, causing a particular challenge when selecting a suitable assay matrix and subsequently trying to meet selectivity, recovery and bias requirements during method validation. In order to determine industry best practices for the quantification of endogenous biomarkers beyond those already recommended by the GCC [23], a survey was sent to both LBA and LC–MS bioanalysts asking what types of matrices are used in their assays.

LBA bioanalysts were evenly divided over the use of authentic matrix or buffer matrix to prepare standards when the endogenous level of the biomarker is below the LLOQ. Conversely, when the endogenous level is very high, buffer matrix was most commonly selected, with a still significant number of respondents electing to use a surrogate matrix (i.e., from another species). It was specified that when anything other than authentic matrix is used, dilutional linearity and parallelism must be verified. For the preparation of validation QC samples when endogenous concentrations are low, bioanalysts overwhelmingly select authentic matrix, pooled matrix is preferred and then fortified with a commercially obtained recombinant biomarker or the high calibrator from the assay kit. When endogenous biomarker concentrations are high, the use of diluted authentic matrix was recommended, although there were differences on how the diluted matrix was prepared. These different methods included mixing various individual lots of authentic matrix to obtain the desired concentrations, diluting authentic matrix with buffer or diluting authentic matrix with depleted matrix. It was noted that respondents typically screen between 10 and 30 individual lots of matrix to find appropriate endogenous levels for their assays.

Different approaches are used by LC–MS bioanalysts. When endogenous biomarker concentrations are low, typically multiple lots of authentic matrix are screened and those with the lowest levels are selected and used for preparing standards. A significant number of respondents also use a surrogate matrix. Authentic matrix is then used to prepare validation QC samples. When very high endogenous biomarker concentrations are expected, standards are prepared with surrogate matrix. Validation QC samples are prepared with authentic matrix for mid and high QCs, and diluted or surrogate matrix is used for low and LLOQ QCs. Some options for surrogate matrix include phosphate-buffered saline; phosphate-buffered saline with bovine serum albumin or human serum albumin in order to take into account the protein content of authentic biological matrix and the increased solubility of hydrophobic compounds; synthetic matrices; or stripped matrix (note that carbon stripping may be variable from lot to lot, and affinity extraction is costly and time consuming). However, the use of a surrogate matrix has some specific issues that must be investigated, such as analyte solubility differences in synthetic aqueous matrix, partial insolubility or partial precipitation or adsorption of nonpolar analytes and analyte recovery differences in surrogate matrix. If authentic matrix is preferred, then it may be possible to use a surrogate analyte in the form of a stable isotope-labeled (SIL) analyte, assuming that the properties of the authentic analyte and SIL analyte are the same and the mass difference between the two is sufficient such that the quantification will not be affected. The internal standard would then be a differently labeled compound or an analog of the analyte. If an SIL analyte is selected, then a correction factor must be applied by comparing the response factors of both the authentic and SIL analyte. The correction factor can be determined by preparing two curves each using surrogate and authentic matrices, determining the response factors and then taking the ratio. Of the two approaches presented above, that is, surrogate analyte/authentic matrix versus authentic analyte/surrogate matrix, the comparison data using the former gave the best results. The effect of matrix variability should be investigated and stability must be done in the authentic matrix.

Whole blood stability & blood partitioning

The Bioanalytical Method Validation guidelines and regulations [16–20] recommend that analyte stability should be considered from the time of sample collection through all handling until analysis. As a response to this recommendation, to address stability during the sample collection procedure, the industry began performing whole blood stability during method development or validation. The European Bioanalysis Forum provided recommendations in 2011 [24], reporting that many bioanalysts agree that plasma stability is an acceptable surrogate for blood stability in approximately 95% of the cases, with the exception of mol-

ecules containing certain functional groups such as *N*-oxides or hydroxamic acids. They recommended analyzing whole blood using a qualified assay. That same year, the GCC recommended assessing whole blood stability using a validated plasma method [7], wherein time 0 samples were drawn and immediately centrifuged to harvest the plasma and compared with stability samples which were left at testing conditions prior to harvesting the plasma. However, any partitioning of the analyte into red blood cells, cell surface binding or platelet binding is not taken into account. One attendee recounted their experience of an FDA-directed audit for a product submission. The goal of the audit appeared to be to understand how the compound behaved in whole blood and hemolyzed plasma. However, when the question of whether there had been any regulatory feedback on the conduct of whole blood stability testing during routine audits was posed during a survey, the majority of respondents indicated that there had been none. The limited feedback that had been received during audits seemed to indicate that the agencies were satisfied with the generally accepted testing of harvested plasma after the blood had remained at ambient temperature for at least 2 h.

The survey attempted to determine the performance of the whole blood stability evaluation by bioanalysts following the earlier GCC recommendations [7]. The majority of respondents routinely conduct sample collection stability testing, typically in support of regulated studies (i.e., GCP or GLP studies). A significant number only perform this test for BE studies, and only a small number perform this test in support of discovery studies. In general, respondents believe that conducting stability testing in whole blood adds scientific value beyond what is learned through routine plasma testing. The evaluation is typically performed by processing the blood sample to plasma prior to analysis and determining the relative concentration versus the control sample. Survey results were less conclusive when the respondents were asked if their test procedure allowed the analyte to reach partitioning equilibrium prior to conducting the experiment. The majority did allow for equilibration, and time 0 and test samples were kept at the same temperature. However, comments indicated that in some cases, partitioning is only evaluated as needed if an issue is detected. It was proposed to use a radiolabeled analyte to definitively determine the partition coefficient in order to properly demonstrate whole blood stability, but this option was largely deemed unnecessary.

If testing results indicated that a significant percentage of the analyte partitions into the red blood cells, most respondents would consider redeveloping the method using whole blood as the matrix, although a significant percentage would not necessarily do so; it would be evaluated on a case-by-case basis, typically by the sponsor. Finally, if instability in whole blood is demonstrated, there were a variety of opinions regarding the validation of the accompanying method. Many thought that the method could not be used in its current form. Others believed that the method could still be considered valid and used to support regulated studies while still others believed that it could only support nonregulated studies. It was clear, however, that the impact of the instability needs to be assessed and documented, and the resulting actions to mitigate the impact must be clearly indicated. Some potential actions include adjusting sample handling procedures to ensure stability by changing the temperature, shortening the processing time or using a stabilizer in the collection tubes.

In conclusion, the original recommendations provided by the GCC [7] were still supported and no new updates appear necessary.

Regulatory audit experiences

A discussion surrounding the regulatory experiences of attendees occurred, with a variety of findings being presented along with suggestions for compliance, when available.

- Test article characterization: the quality assurance (QA) unit failed to review the raw characterization data of the test article, and only assured that the certificate of analysis (provided by the sponsor) was available. Nor did the QA unit determine if the sponsor was compliant with the GLP regulations under 21 US Code of Federal Regulations 58. It is now unclear if there is an expectation that the CRO would need to perform an external audit to assure compliance. Attendees suggested considering including a compliance exception in the QA statement.

- Freeze–thaw stability: two findings were related to freeze–thaw stability. First, freeze-thaw stability of incurred sample reanalysis (ISR) samples exceeded the validated number of freeze–thaw cycles. Attendees agreed that freeze–thaw of these samples is inherently supported by the acceptability of the ISR results and no additional validation testing is required. Second, the thaw time during the evaluation was not representative of the time that study samples were thawed. It is suggested to thaw validation samples at least 2 h between cycles.

- Training SOPs: proficiency training on SOPs was not included in the training program. A process for proficiency training on SOPs is suggested.

- Run interruptions: regulatory auditors cited that runs should not be interrupted for interim analysis of run acceptance, nor can assumptions be made regarding run acceptance prior to completion of the run. It is recommended to ensure that SOPs do not allow for interim analysis of run acceptance.

- Sponsor communication: regulatory auditors cited that correspondence and decisions for all events need to be documented and maintained within the study. It is suggested that SOPs outline the requirements for capturing, retaining and filing significant communications (including verbal ones) and correspondence. Specifically, these should include significant decisions that impact or affect study conduct.

- Sample storage: the SOP was lacking in the description for quarantine procedures and allowed for the use of open access storage units for temporary storage. It is suggested that SOPs outline the procedure for quarantining samples when they are received compromised. Additionally, the use of open access storage units is not recommended.

- Instrumentation SOPs: some instrumentation SOPs are not harmonized with user manual specifications and recommendations. It is recommended to include references to user manuals in applicable SOPs.

- Freezer mapping: regulatory auditors cited that temperature mapping of all storage units must be conducted because the results of this mapping must then be used to determine proper temperature probe placement.

- Time stamping documentation: regulatory auditors cited that time stamping is required on critical documentation to demonstrate that approvals of activities are conducted prior to their execution. It is suggested that critical forms be reformatted to include the requirement for recording the time of each dated signature.

Electronic laboratory notebooks & LIMS systems

The topic of electronic laboratory notebooks (ELNs) was revisited again at this closed forum after being discussed during the 7th GCC meeting [5]. At that time, the CRO community was divided on the use of ELNs. Most member companies were still using paper, and although users of these systems saw benefits to quality and documentation, cost and a variety of types of studies causing adaptability issues were seen as deterrents to implementation. In order to determine if there had been a change to this position after two more years of experience, a survey was sent to member companies. Results indicate that only a third of respondents use an ELN at this time.

Queries into additional electronic systems were also asked. Most GCC members do not use electronic archiving software, but elect to archive electronic data manually. Thermo Watson LIMS™ is used by a majority of respondents, or a custom/in-house developed system. If Watson is used, it is principally for routine analysis or method validation, automated reassay selection and sample tracking.

Future perspective

The GCC will continue to provide recommendations on hot topics of global interest in small and large molecule bioanalysis, biomarkers and immunogenicity, and expand its membership by coordinating its activities with the regional and international meetings held by the pharmaceutical industry. Please contact the GCC [25] for the exact date and time of future meetings, and for all membership information.

Acknowledgements

The GCC would like to thank the following: Roger Hayes (MPI Research) for chairing the 9th GCC Closed Forum; Mario Rocci (ICON Laboratories) and Isabelle Dumont (Algorithme Pharma) for designing the survey, collecting answers, preparing survey results and chairing the sessions on 'The place of CAPA in regulated bioanalysis' and 'Method robustness in regulated bioanalysis'; Rafiq Islam (Celerion), Dominique Gouty (BioAgilytix) and Chris Beaver (inVentiv Health Clinical) for designing the survey, collecting answers, preparing survey results and chairing the session on 'Best practices for supporting biosimilar studies'; Richard LeLacheur (Agilux Labs), Elizabeth Groeber (WIL Research) and Colin Pattison (Charles River Labs) for designing the survey, collecting answers, preparing survey results and chairing the session on 'Status of method validation requirements in the preclinical space'; Afshin Safavi (BioAgilytix), John Kamerud (Eurofins Pharma Bioanalytics Services) and Laura Cojocaru (Tandem Labs-now Covance) for designing the survey, collecting answers, preparing survey results and chairing the session on 'Best practices for measuring endogenous biomarkers beyond prior GCC recommendations'; Philippe Couerbe (Atlanbio) and Stephanie Cape (Covance) for chairing the session on 'Whole blood stability and blood partitioning'; Roger Hayes (MPI Research) and Chad Briscoe (PRA Health Sciences) for chairing the sessions on 'Regulatory audit experiences' and 'Electronic laboratory notebooks and LIMS systems'; All the GCC member company representatives who filled in the numerous surveys used to prepare the discussion of the 9th GCC Closed Forum; All the member company representatives who have sent comments and suggestions to complete this commentary; Natasha Savoie (GCC) for taking the minutes of the 9th GCC Closed

Forum and drafting the first version of this commentary; Wei Garofolo (GCC) for organizing the logistics of the meeting and coordinating the review of this commentary.

Financial & competing interests disclosure

The authors have no relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript. This includes employment, consultancies, honoraria, stock ownership or options, expert testimony, grants or patents received or pending, or royalties.

No writing assistance was utilized in the production of this manuscript.

Author affiliations

[1]MPI Research, Mattawan, MI, USA
[2]Agilux Laboratories, Worcester, MA, USA
[3]Algorithme Pharma, Laval, QC, Canada
[4]Atlanbio, Saint-Nazaire, France
[5]BioAgilytix, Durham, NC, USA
[6]Celerion, Lincoln, NB, USA
[7]Charles River, Edinburgh, UK
[8]Covance, Madison, WI, USA
[9]ICON, Whitesboro, NY, USA
[10]PRA Health Sciences, Lenexa, KS, USA
[11]Tandem Labs, West Trenton, NJ, USA, currently at Covance, WestTrenton, NJ, USA
[12]WIL Research, Ashland, OH, USA
[13]3S-Pharmacological Cons. & Res. GmbH, Harpstedt, Germany
[14]AIT Bioscience Indianapolis, IN, USA
[15]Alturas Analytics, Moscow, ID, USA
[16]Anapharm Europe, Barcelona, Spain
[17]Axis Clinicals, Dilworth, MN, USA
[18]BASi, West Lafayette, IN, USA
[19]Bioanalytical Laboratory Services, Toronto, ON, Canada
[20]Biotrial Bioanalytical Services Inc., Laval, QC, Canada
[21]CAPRION, Montreal, QC, Canada
[22]Charles River, Senneville, QC, CAN, currently at Algorithme Pharma, Laval, QC, Canada
[23]Charles River, Reno, NV, USA
[24]CIRION BioPharma Research, Laval, QC, Canada
[25]Covance, Chantilly, VA, USA
[26]Eurofins Pharma Bioanalytics Services, St Charles, MO, USA
[27]Frontage Laboratories, Inc., Exton, PA, USA
[28]Global CRO Council for Bioanalysis (GCC)
[29]Global CRO Council for Bioanalysis, Envigo, East Millstone, NJ, USA
[30]Intertek USA, San Diego, CA, USA
[31]Previously at InVentiv Health Clinical, Quebec City, QC, CAN
[32]CMIC, Inc., Hoffman Estates, IL, USA
[33]KCAS Bioanalytical Services, Shawnee, KS, USA, currently at Alliance Pharma, Malvern, PA, USA
[34]KCAS Bioanalytical Services, Shawnee, KS, USA
[35]Keystone Bioanalytical, Inc., North Wales, PA, USA
[36]Lambda Therapeutic Research, Ahmedabad, India
[37]LGC, Fordham, UK
[38]Mercodia AB, Uppsala, Sweden
[39]PPD, Richmond, VA, USA
[40]PRA Health Sciences, Assen, The Netherlands
[41]Pyxant Labs, Colorado Springs, CO, USA
[42]Pharma Medica, Toronto, ON, Canada
[43]QPS, Groningen, The Netherlands, currently at Antaeus Biopharma, Groningen, Netherlands
[44]Quintiles, Ithaca, NY, USA
[45]Quintiles, Oss, The Netherlands
[46]Tandem Labs, Durham, NC, USA, currently at Covance, Durham,NC, USA
[47]Worldwide Clinical Trials, Austin, TX, USA
[48]WuXi / XBL, 107 Morgan Lane, Plainsboro, NJ, USA

References

1 Premkumar N, Lowes S, Jersey J *et al.* Formation of a global contract research organization council for bioanalysis. *Bioanalysis* 2(11), 1797–1800 (2010).

2 Breda N, Garofolo F, Cruz Caturla M *et al.* The 3rd Global CRO council for bioanalysis at the international reid bioanalytical forum. *Bioanalysis* 3(24), 2721–2727 (2011).

3 Lowes S, Jersey J, Shoup R *et al.* 4th Global CRO Council for Bioanalysis: coadministered drugs stability, EMA/US FDA Guidelines, 483s and Carryover. *Bioanalysis* 4(7), 763–768 (2012).

4 Nicholson R, Lowes S, Cruz Caturla M *et al.* 6th GCC focus on LBA: critical reagents, positive controls and reference standards; specificity for endogenous compounds; biomarkers; biosimilars. *Bioanalysis* 4(19), 2335–2342 (2012).

5 Rocci M, Lowes S, Shoup R *et al.* 7th GCC: use of incurred samples; fit-for-purpose validation; solution stability; electronic lab notebooks; hyperlipidemic matrix testing. *Bioanalysis* 6(20), 2713–2720 (2014).

6 Bower J, Fast D, Garofolo F *et al.* 8th GCC: consolidated feedback to US FDA on the 2013 Draft FDA Guidance on Bioanalytical Method Validation. *Bioanalysis* 6(22), 2957–2963 (2014).

7 Lowes S, Jersey J, Shoup R *et al.* Recommendations on: Internal Standard Criteria, Stability, Incurred Sample Reanalysis and Recent 483s by the Global CRO Council for Bioanalysis. *Bioanalysis* 3(12), 1323–1332 (2011).

8 Boterman M, Doig M, Breda M *et al.* Recommendations on the interpretation of the new EMA Guideline on Bioanalytical Method Validation by Global CRO Council for Bioanalysis (GCC). *Bioanalysis* 4(6), 651–660 (2012).

9 Sangster T, Maltas J, Struwe P *et al.* Recommendations on ISR in multi-analyte assays, QA/bioanalytical consultants and GCP by Global CRO Council for Bioanalysis (GCC). *Bioanalysis* 4(14), 1723–1730 (2012).

10 Lowes S, Boterman M, Doig M *et al.* Recommendations on bioanalytical method stability implications of co-administered and co-formulated drugs by Global CRO Council for Bioanalysis (GCC). *Bioanalysis* 4(17), 2117–2126 (2012).

11 Hougton R, Gouty D, Allinson J *et al.* Recommendations on biomarker bioanalytical method validation by GCC. *Bioanalysis* 4(20), 2439–2446 (2012).

12 Lowes S, LeLacheur R, Shoup R *et al.* Recommendations on Incurred Sample Stability (ISS) by GCC. *Bioanalysis* 6(18), 2385–2390 (2014).

13 US Department of Health and Human Services, US FDA, Center for drug evaluation and research, center for biologics evaluation and research. guidance for industry: Q10 pharmaceutical quality system (2009). www.fda.gov/downloads/Drugs/.../Guidances/ucm073517.pdf

14 ICH Harmonised tripartite guideline: Q2(R1) validation of analytical procedures: text and methodology (1996). www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Quality/Q2_R1/Step4/Q2_R1__Guideline.pdf

15 US Department of Health and Human Services, US FDA, Center for drug evaluation and research, center for biologics evaluation and research. guidance for industry: analytical procedures and methods validation for drugs and biologics (2015). www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm386366.pdf

16 US Department of Health and Human Services, US FDA, Center for Drug Evaluation and Research, Center for Veterinary Medicine. Draft Guidance for Industry, Bioanalytical Method Validation, Rockville, MD USA (2013). www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm368107.pdf

17 US Department of Health and Human Services, US FDA, Center for Drug Evaluation and Research, Center for Veterinary Medicine. Guidance for Industry, Bioanalytical Method Validation, Rockville, MD USA (2001). www.fda.gov/downloads/Drugs/Guidance/ucm070107.pdf

18 European Medicines Agency. Committee for Medicinal Products for Human Use (CHMP). Guideline on Bioanalytical Method Validation. London, UK (2011). EMEA/CHMP/EWP/192217/2009. www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2011/08/WC500109686.pdf

19 ANVISA. Resolution RDC no 27, May 17, 2012. Brazilian Health Surveillance Agency, Brazil (2012). http://bcn2012.europeanbioanalysisforum.eu/slides/day%203/7%20updates%20from%20the%20globe/3_tavares.pdf

20 European Medicines Agency. Committee for Medicinal Products for Human Use (CHMP). Guideline on similar biological medicinal products. London, UK (2014). CHMP/437/04 Rev 1. www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500003517.pdf

21 US Department of Health and Human Services, US FDA, Center for Drug Evaluation and Research, Center for Biologics Evaluation and Research. Scientific Considerations in Demonstrating Biosimilarity to a Reference Product (2015). www.fda.gov/downloads/DrugsGuidanceComplianceRegulatoryInformation/Guidances/UCM291128.pdf

22 European Medicines Agency. Committee for Medicinal Products for Human Use (CHMP). Draft Reflection paper on guidance for laboratories that perform the analysis or evaluation of clinical trial samples. London, UK (2010). EMA/INS/GCP/532137/2010 www.ema.europa.eu/docs/en_GB/document_library/Regulatory_and_procedural_guideline/2010/09/WC500096987.pdf

23 Houghton R, Gouty D Allinson J *et al.* Recommendations on biomarker bioanalytical method validation by GCC. *Bioanalysis*, 4(20), 2439–2446 (2012).

24 Freisleben A, Brudny-Kloppel M, Mulder H *et al.* Blood stability testing: European Bioanalysis Forum view on current challenges for regulated bioanalysis. *Bioanalysis* 3(12), 1333–1336 (2011).

25 Global CRO Council for Bioanalysis. www.global-cro-council.org

Bioanalysis

ISSN: 1757-6180
Frequency per year: 24

IMPACT FACTOR:
2.673 (2016)

In association with
Bioanalysis ZONE
www.bioanalysis-zone.com

BIOANALYSIS COVERS KEY AREAS SUCH AS:

- Analyte extraction and sample preparation, including microsampling
- Chromatography and separation sciences
- MS, NMR and other key detection methods, including ligand binding assays
- Drug and metabolite assays
- Biomarker assays, diagnostic assays and test kits
- Method development and validation approaches
- Innovative bioanalytical methods and new instrumentation and equipment
- Laboratory automation and efficiency
- Regulatory and compliance issues
- Data processing and management

INDEXING

MEDLINE/Index Medicus, EMBASE/Excerpta Medica, Chemical Abstracts, BIOSIS Previews, BIOSIS Reviews Reports and Meetings, Journal Citation Reports/Science Edition, Science Citation Index Expanded™ (SciSearch®), Scopus®

"*Bioanalysis* is an outstanding journal, offering the latest in analysis and filling a serious gap in our knowledge of analysis of biological samples."

– **Sanjay Garg**, The University of South Australia, Australia

www.future-science.com

To claim your free trial, contact us at: trials@futurescience.com

newlands press part of fsg

Bioanalysis

Bioanalytical inspections: organizational changes and regulatory perspectives

"The Division of Bioequivalence and GLP Compliance, responsible for CDER's Bioavailablity/Bioequivalence Inspection Program and the Good Laboratory Practice Inspection Program, became the Office of Study Integrity and Surveillance."

Keywords: batch acceptance • biosimilars • extract stability • immunogenicity testing • regulatory • US FDA

First draft submitted: 3 March 2016; **Accepted for publication:** 17 March 2016; **Published online:** 15 April 2016

Sam H Haidar
Author for correspondence:
US Food & Drug Administration, 10903 New Hampshire Ave., Silver Spring, MD 20993, USA
sam.haidar@fda.hhs.gov

Kara A Scheibner
US Food & Drug Administration, 10903 New Hampshire Ave., Silver Spring, MD 20993, USA

During January 2015, several offices within the Center for Drug Evaluation and Research (CDER) were re-organized to improve consistency of decision-making, and to streamline policy, procedures and communications. Some Divisions were elevated to Office status because of the increased number of personnel resulting from the passage of the Generic Drug User Fee Amendments of 2012 (GDUFA). The resources provided by the new user fees supported hiring more personnel to address the significantly increased workload and newly imposed timelines. The Division of Bioequivalence and GLP Compliance (DBGLPC), responsible for CDER's Bioavailablity/Bioequivalence (BA/BE) Inspection Program and the Good Laboratory Practice (GLP) Inspection Program, became the Office of Study Integrity and Surveillance (OSIS). Additionally, OSIS became the newest suboffice under CDER's Office of Translational Sciences. Previously, this group was in the Office of Scientific Investigations under CDER's Office of Compliance. OSIS currently includes the Division of Generic Drug Bioequivalence Evaluation conducting and reviewing work under the BA/BE Inspection Program, and the Division of New Drug Bioequivalence Evaluation, which, in addition to the BA/BE program, also conducts and reviews inspections under the GLP Inspection Program. Concurrent with the reorganization, OSIS began to implement a surveillance approach to inspections in conjunction with the application-based inspections, conducted previously.

This commentary provides an overview of changes in procedures within OSIS; it will also briefly discuss topics of interest that came up during the past year, including assays related to therapeutic biologics and immunogenicity testing.

Procedural changes

Prior to January 2015, scientific reviewers in CDER's Office of Generic Drugs or in the Office of Clinical Pharmacology usually determined the need for inspections during their reviews of studies supporting drug applications submitted to the Center. If an inspection was deemed necessary, a written consult requesting an inspection was generated by the review division and submitted to DBGLPC. DBGLPC prepared assignment memoranda for inspections of the clinical components of BA/BE studies and sent them to the Office of Regulatory Affairs (ORA) to conduct the inspections. The memoranda provided the background information for the inspection. If the inspection request was for the analytical component of a BE study, the same procedure was followed; however, analytical inspections included participation by DBGLPC staff to provide any necessary scientific expertise.

Beginning in January 2015, the new office (OSIS) began implementing a surveillance

FUTURE SCIENCE part of fsg

approach to inspections. Surveillance is conducted contemporaneously with study-directed or application-based inspections, requested by the review divisions. As a result, OSIS is now responsible for selecting sites for routine inspections of clinical and bioanalytical sites that conduct studies for submitted applications. Currently, project management and members of the Collaboration, Risk-Evaluation and Surveillance Team (CREST) within OSIS evaluate sites in incoming applications to determine the need for inspections based on site history and other factors. The objective is to optimize resources and streamline procedures commensurate with CDER's significant increase of obligations under GDUFA.

> "Beginning in January 2015, the new office (OSIS) began implementing a surveillance approach to inspections."

Under surveillance inspections, greater emphasis is placed on collecting information about overall site performance. Inspectors document the knowledge and experience of personnel, site procedures, commitment to good scientific principles and adherence to regulations. Inspection procedures include evaluating relevant studies, in addition to data audits and thorough review of individual studies. The information collected about the site plays a significant role in determining the timing and frequency of future inspections.

In addition to the organizational changes during 2015, OSIS conducted several inspections in support of therapeutic biologics applications to CDER, including studies supporting biosimilars. Inspection and evaluation of immunogenicity studies presented a new set of challenges. This was reflected by the difficulties some sponsors experienced while conducting regulated studies. As a result, CDER identified some studies requiring extensive review and the potential need for significant modification. This topic is discussed in greater detail further in this article. Topics related to analytical methods of small molecules were also examined within the office, when inspections encountered multiple approaches to similar types of testing. Two examples, extract stability and the inclusion of quality control samples in extraction batches, are described further in the following sections.

Processed sample (extract) stability

Testing conducted on samples postextraction (other than concentrations) can serve different objectives or address different concerns. Examples include: when a run is interrupted due to instrument failure or other reasons; when a run is stored in the autosampler for a period of time prior to injection; when processed samples are stored at room temperature or in the refrigerator; stability of samples at the beginning of a run compared with those at the end of the run; and finally, reinjection reproducibility where a whole run may be injected twice and variability in the results is determined. Different terminology has been ascribed to the testing listed above [1,2]. For example, the US FDA's Guidance for Industry: Bioanalytical Method Validation, published in 2001 refers to testing conducted postextraction as postpreparative stability [3]. Others have used terms such as processed sample stability, extract stability, autosampler stability and processed sample integrity. This has led to some confusion in industry, especially because some testing may require the use of freshly prepared standard curves or calibrators (e.g., storing extracted samples for a period of time), and some do not (e.g., re-injection reproducibility). In addition to the experience we have gained through inspections, OSIS staff has participated in efforts to identify the various approaches for testing processed samples in an effort to determine best practices, and provide better recommendations. Efforts in this regard are ongoing.

Several factors should be considered when conducting stability testing on extracted samples. As described in the 2009 White Paper on Recent Issues in Regulated Bioanalysis from the 3rd Calibration and Validation Group Workshop, there is always the possibility of similar degradation between the internal standard and the analyte of interest [1]. Simply reinjecting samples after storage could provide misleading results. For example, the ratio between the internal standard and the analyte may be the same, although significant degradation may have taken place. To address this concern, freshly prepared calibrators may be used. Another approach, which may be used when samples are left in the autosampler for a period of time, involves injection of a run the first time, waiting for a given number of hours, injecting the samples a second time, then determining concentration based on the results from the calibrators in the first run. As such, stability concerns would be detected without the need for using freshly prepared calibrators. Other factors to consider include evaporation and possible chemical reactions resulting from prolonged exposure to air (oxygen). This may occur after sample vial membranes are punctured by the sampling needle following the first injection of the run. Excessive evaporation can be expected when the reconstitution solution has a large component of organic solvent(s).

Acceptance of processed batches

Next we consider the topic of run acceptance in cases of runs comprising two or more individually (sepa-

rately) processed (extracted) batches. The European Medicines Agency (EMA) Guideline on bioanalytical method validation recommends the following: "In case a whole run comprises more batches, acceptance criteria should be applied to the whole run and to the individual batches. The run can be acceptable, although a batch might have to be rejected, as criteria were not met" [4]. As such, QCs at three different concentrations (low, medium and high) are to be included in each batch. This is consistent with FDA's position. Practical or technical limitations notwithstanding, FDA currently recommends the inclusion of low, medium and high QCs in duplicates in individually processed batches. If singlet QCs are utilized at the three concentrations, then failure of any of the QCs within the batch could cast doubt on the accuracy of the unknowns within that batch, even if the run as a whole passes. Then, the batch may need to be repeated. This will be determined on a case-by-case basis.

Biologics & biosimilars

In 2015, review and evaluation of a large number of biologics applications, including those supporting biosimilars, revealed specific areas of scientific concern within immunogenicity studies. The following sections highlight some of the critical observations OSIS has made over the past year while inspecting antidrug antibody (ADA) and neutralizing antibody (NAb) assays in support of immunogenicity assessments for therapeutic proteins.

Selection of a relevant low positive control antibody

Selection of a positive control antibody for use in ADA/immunogenicity testing can be challenging. Available affinity-purified antibodies are not necessarily representative of antibodies produced in response to a therapeutic protein, and results obtained are potentially the 'best case scenario.' Thus, it is important that assay values of selected positive control concentrations are representative of antibody levels present in study samples. This is especially critical for the low positive control (LPC); the LPC should adequately monitor performance of the ADA screening assay at response levels surrounding the screening assay cut point. The language in the FDA draft guidance, "Assay Development for Immunogenicity Testing of Therapeutic Proteins" [5] is that the selected LPC concentration should result in an assay rejection rate of 1% (i.e., the LPC assay signal should fall below the cut point 1% of the time).

During recent inspections of biologics applications with immunogenicity testing, we have observed instances of selected LPC concentrations that yield an assay response significantly higher than the screening cut point; in some studies the LPC assay response has been up to tenfold greater. Conversely, study sample assay response values have been very close to the screening assay cut point, and therefore significantly less than the LPC response. Using an LPC concentration with a high signal response compared with the cut point prevents adequate monitoring of assay performance and precision of assay values around the cut point. Consequently, the precision of study sample assay values that fall in the low range of the assay cannot be confirmed, and there is little confidence in the ability of the assay to distinguish potentially ADA-positive versus potentially ADA-negative samples in the screening assay. Relevance of the LPC is also an important consideration in the confirmatory assay. If the assay signal of the LPC is significantly higher than the screening cut point, consistent signal inhibition in the confirmatory assay does not accurately reflect potential study samples with a signal response in the low range of the assay. Thus, demonstration of consistent assay signal inhibition (above the confirmatory cut point) at an LPC concentration yielding assay values close to the screening cut point is critical to show the ability of the assay to positively detect ADAs in all study samples.

Reproducibility between Tier 1 & Tier 2 assays

During inspection of studies, we have observed study samples that screen positive in the Tier 1 assay, but do not 'screen positive' in the Tier 2 assay (i.e., the unspiked study samples within a confirmatory assay have assay response values below the screening cut point). The result of this scenario is a degree of uncertainty in the reliability of the assay: if study samples screen positive in Tier 1 and negative in Tier 2, there is a distinct possibility that the converse scenario could occur (i.e., study samples that screened negative in Tier 1 could test positive in Tier 2). Thus, in these studies, the assay appears to have significant reproducibility concerns, and potentially positive samples may be missed in the initial screening (Tier 1) assay.

In addition, this scenario poses a question of how to handle samples that screen positive in Tier 1, 'screen' negative in Tier 2 and confirm positive in Tier 2. We have seen examples of such samples reported as positive, placing confidence in the highly specific confirmatory assay. However, there are also examples of such samples reported as negative; justification for reporting the sample as negative was based on the fact that a sample which screened negative in the confirmatory assay cannot be confirmed positive, thus placing greater confidence in the screening assay.

These inconsistencies present several questions that should be discussed and addressed within the immu-

nogenicity testing community: should there be an expectation for reproducibility of unspiked study samples in ADA assays as an indicator of a reliable assay? Should there be a consensus among firms in how these data are reported? If yes, does scientific rationale favor the confirmatory assay or the screening assay? Should such study results trigger repeat analysis? And, is there an underlying cause of assay nonreproducibility that should be considered?

Interference in cell-based neutralizing antibody assays

Using a cell-based assay format to determine presence of neutralizing antibodies in study samples is highly desirable because of the functional relevance of the assay. However, cell-based assays also add additional layers of complexity and variability that must be addressed during method validation. One such additional layer is interference from endogenous molecules (e.g., cytokines) present in control serum samples and serum from study samples. Assessment of potential interference from these molecules is critical to ensuring reliability of the assay, and prevention of false positive or negative results.

In assessing the potential for interference in a NAb assay, choice of cell type is highly important. The ability of a cell to respond to both the therapeutic protein of choice (e.g., erythropoietin), and to other potential stimulating molecules, should factor into the decision of which cells are best suited for a NAb assay. If the cell line chosen for the assay responds to multiple stimuli, the specific concentration level at which interference in the assay will be present should be determined for each molecule. This knowledge should subsequently be applied when testing the study samples. Assessing the level of potentially interfering endogenous molecules is critical in all study samples, for example, via an ELISA assay. This is particularly true in patient populations where levels of these molecules can be elevated due to a specific disease. If levels of interfering molecules cannot be assessed, an alternative method should be used to determine whether study samples contain molecules that could potentially stimulate cells in the absence of study drug. Finally, if endogenous molecules are present in specific study samples at concentration levels that will interfere with the assay, a method to negate this interference should be assessed and validated. For example, high levels of endogenous cytokines can often be removed using monoclonal antibody depletion protocols prior to study sample analysis in the NAb assay.

Conclusion

The topics of discussion presented herein represent a small subset of the processes, methodology and scientific rationale that are critical to traditional BE studies, and to bioanalysis of biologics and biosimilars. Further information on these and other topics is available in the manuscripts and guidances referenced in this commentary.

Disclaimer

The views expressed in this article do not necessarily represent the official views or policies of the US FDA.

Acknowledgements

The authors wish to thank William (Bill) Taylor and Sean Kassim for their insightful comments and edits of this Commentary.

Financial & competing interests disclosure

The authors have no relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript. This includes employment, consultancies, honoraria, stock ownership or options, expert testimony, grants or patents received or pending, or royalties.

No writing assistance was utilized in the production of this manuscript.

References

1 Savoie N, Garofolo F, van Amsterdam P *et al.* 2009 White Paper on Recent Issues in Regulated Bioanalysis from The 3rd Calibration and Validation Group Workshop. *Bioanalysis* 2(1), 53–68 (2010).

2 Nowatzke W, Woolf E. Best practices during bioanalytical method validation for the characterization of assay reagents and the evaluation of analyte stability in assay standards, quality controls, and study samples. *AAPS J.* 9(2), E117–E122 (2007).

3 US Food and Drug Administration. Guidance for Industry: Bioanalytical Method Validation. US Department of Health and Human Services, US FDA, Center for Drug Evaluation and Research, MD, USA (2001). www.fda.gov/downloads/Drugs/Guidances/ucm070107.pdf

4 European Medicines Agency. Guideline on bioanalytical method validation. EMEA, Committee for Proprietary Medicinal Products (CPMP), London, UK (2011). www.ema.europa.eu

5 US Food and Drug Administration. Guidance for Industry: Assay Development for Immunogenicity Testing of Therapeutic Proteins (Draft). US Department of Health and Human Services, US FDA, Center for Drug Evaluation and Research, MD, USA (2009). www.fda.gov

Bioanalytical sciences

LGC offers a world-class bioanalytical capability with extensive experience in high-integrity analytical science.

- **Small molecule bioanalysis** (LC-MS/MS, BioMS™, Peptides)
- **Biopharmaceuticals** (PK, ADA, NAbs)
- **Biomarkers**
- **Microbiology**
- **Molecular biology** (qPCR, DNA sequencing)

www.lgcgroup.com

Science for a safer world

@LGC_dds

Roundtable discussions: Issues facing the Bioanalytical Community & The Changing World of Bioanalysis

In April 2016, Bioanalysis Zone and *Bioanalysis* organized an independent Roundtable Discussion at Hilton Orlando Lake Buena Vista, Orlando, Florida, USA, in which bioanalytical experts from Pharma and CROs were brought together to discuss topical issues faced by the quantitative bioanalytical community.

Chaired by Neil Spooner (Senior Editor, *Bioanalysis*), the discussion focused on three key areas.

In part 1 of this panel discussion, a panel of leading CRO and Pharma Directors assess the current situation of outsourcing in bioanalysis, and discuss the direction of outsourcing of regulated bioanalysis in the future.

Discussions included:

- Outsourcing strategies
- Drivers and business models in outsourcing
- Scientific exchange and interactions between CROs and Pharmas
- Sharing and discussing data between CROs and Pharmas
- Cost–effectiveness of outsourcing

In part 2 of this panel discussion, the panelists discuss the divide between investments being made by Pharmas and CROs in new bioanalytical techniques and technologies. Discussions included:

- Investments made by Pharmas and CROs for new constructs
- Is there a real divide between CRO and Pharma?
- If so, why is this occurring?
- What are the gaps?
- What are the solutions?

In part 3 of the panel discussion, the panelists discuss the growing skills gap in bioanalytical laboratories. Discussions included:

- Reality of the skills gap
- In which technical areas is the skills gap most apparent?
- Where does the skills gap occur – Pharmas and/or CROs?
- How do we bridge the gap?

You can view the full Round Table Discussion Report published in *Bioanalysis* online (Spooner N, Cape S, Hayes R *et al.* Issues Facing the Bioanalytical Community – Summary of Roundtable Discussions. *Bioanalysis* 8[21], 2189–2193 [2016]).

You can also view highlights from the Bioanalysis Zone Roundtable Discussion in the accompanying supplement online.

In November 2016 Bioanalysis Zone and *Bioanalysis* organized an independent Roundtable Discussion at the Hesperia Tower Hotel and Conference Center, Barcelona, Spain, in which bioanalytical experts from Pharmas and CROs were brought together to discuss topical issues in the Changing World of Bioanalysis.

Chaired by Neil Spooner (Senior Editor, *Bioanalysis*), the discussion focused on three key areas.

Part 1: How do we access the required skills?

- Decreasing the current skills gap evident in most bioanalytical labs, particularly in the younger generation of bioanalysts
- Collaborations between industry and academia to produce a newer generation of bioanalysts that are better able to deal with more challenging bioanalytical problems
- Training both new and existing staff to deal with more challenging bioanalytical problems (more challenging molecules and newer constructs) and cross-training staff within organisations
- The changing nature of the CRO–Pharma relationship

Part 2: What will the future bioanalytical laboratory look like and what skills will this require?

- The need for bioanalysts to acquire more biological skills in order to understand a newer generation of complex molecules
- Convergence of different units of companies and merging of skills
- The role of automation in the bioanalytical lab of the future

Part 3: How do we embed science in the Pharma–CRO partnership?

- There is a need for more scientific interactions between CRO and Pharma, the model of which will vary based on client needs and the nature of molecules being studied.
- Scientific dialogues should be increased between scientists from both CRO and Pharma, and outsourcing managers, to provide a more comprehensive idea of requirements from both sides.
- The transfer of knowledge on science and technology between Pharma and CRO (and Biotech) should continue, potentially through dedicated roles within the organizations and appropriate training.
- Working across teams, such as CRO scientists working with internal project teams at Pharma, could also aid the process and which has already begun in the industry.

The full Roundtable Discussion Report and videos of the discussions that took place are available on *Bioanalysis* and Bioanalysis Zone.

ENVIGO

At Envigo, the positives are in more than just our name

+ Leading non-clinical CRO with research model portfolio
+ An extensive global footprint
+ Proven scientific excellence and regulatory knowledge
+ Expert data interpretation and insights
+ A truly customer-centric culture
+ Renowned animal welfare and enrichment

E-book download

'The true value of a non-clinical CRO'

envigo.com/truevalue

New Investigator Award

Over the past 7 years, Bioanalysis Zone is proud to promote and recognize the work of highly talented early-career researchers. With the support of our sponsor Waters Corporation we are honored to offer nominees and winners of our annual New Investigator Award a springboard to help establish themselves in the exciting world of bioanalysis.

Panagiotis A Vorkas was named the winner of the 2016 New Investigator Award. He was nominated by Prof Elaine Holmes, his supervisor at Imperial College London (UK) in honor for his work in developing a metabolic profiling workflow utilizing two untargeted UPLC–MS methods which, when combined, expanded metabolome coverage. The rigorous optimization of these methods and implementation of new and much needed concepts in quality control, suggestions in post-acquisition artefact exclusion, structural assignment of a wide range of metabolites and development of an experimental set-up to achieve this resulted in several key publications. Application of the developed pipeline in different manifestations of cardiovascular disease, demonstrated its superior ability to expand metabolite coverage and most importantly their efficiency – as untargeted methods – to detect previously unreported but biologically important molecules and elucidate dysregulated pathways in disease.

"I am deeply honored to receive this year's Bioanalysis Zone New Investigator Award. I see this award as a huge encouragement in my career to continue contributing to advancements in bioanalysis. I am confident that the award will enhance my career prospects and hope to enjoy the professional success of past recipients."

Panagiotis A Vorkas

"It's been delightful running this year's award. As ever, the NIA has gone from strength to strength each year attracting a number of hugely talented bioanalysts working on a variety of projects spanning a wide range of present day bioanalytical challenges…I would like to congratulate Panagiotis Vorkas and to all this year's nominees for their hard work and contribution to the bioanalytical field."

Ayan Ali, Digital Editor for Bioanalysis Zone (2016)

"I am a great believer in encouraging the development of the new and exciting talent we need to face the bioanalytical challenges of the future and for them to bring forwards the novel approaches and ideas that will be the lifeblood of our discipline in the years to come."

Dr. Neil Spooner (Spooner Bioanalytical Solutions)

"We are a company that, above all, values the quality of science. To be true to our reputation, we should be supporting others who share similar ideals. We have excellent scientists who work at Waters…It's our responsibility to support others and their research, whose work we care about and believe in. Waters plays an important role in advancing science. We primarily do this through technological innovation, but we also champion the scientists – new and experienced – who are moving scientific research forward."

Dr. Erin Chambers, Director in the Scientific Operations group (Waters Corporation)

Bioanalysis

Large Molecule

A current perspective of supercharging reagents and peptide bioanalysis

First draft submitted: 17 November 2015; **Accepted for publication:** 23 November 2015; **Published online:** 25 January 2016

Keywords: charge state envelope • glucagon • liraglutide • *m*-NBA • peptide LC–MS/MS analysis • supercharging

Richard G Kay
Author for correspondence:
LGC Limited, Newmarket Road, Fordham, Cambridgeshire, CB7 5WW, UK
richard.kay@lgcgroup.com

James Howard
LGC Limited, Newmarket Road, Fordham, Cambridgeshire, CB7 5WW, UK

Sara Stensson
Ferring Pharmaceuticals A/S, Kay Fiskers Plads 11, Copenhagen, DK-2300, Denmark

The bioanalysis community has been using LC–MS-based approaches for quantifying small molecule drugs and biomarkers extracted from plasma for at least a couple of decades. During this time LC–MS/MS technology has improved significantly, especially the sensitivity of triple quadrupole instruments. Over the past decade, these instruments have become sufficiently sensitive to detect and quantify larger analytes like peptides and small proteins once they have been extracted from complex matrices. However, these analytes are significantly more challenging to work with than small molecules, and the achievable sensitivity using LC–MS/MS instrumentation is hindered by the same mechanism that enables these higher molecular weight analytes to be detected by bench top instruments in the first place. This hindrance is the acquisition of multiple number of charges in the electrospray source – for example, a hypothetical 3500 kDa peptide with a single charge (m/z 3501) will not be visible using any of the commercially available triple quadrupoles. However, the addition of two more positive charges to the peptide generates an m/z of 1167.7, which is within the m/z range of all commercial instruments. The acquisition of charges in the electrospray source is not a well-controlled process, and the hypothetical peptide would likely be present with m/z values of 1751, 1167.7, 876 and 701 corresponding to the 2, 3, 4 and 5+ charged molecules, respectively. The ultimate number of charges any given peptide can obtain is usually driven by the number of chargeable groups, which is determined by the peptides primary amino acid structure, and in some cases by its secondary structure. The acquisition of multiple charge states splits the available ion current for the peptide into parts, and although these data are useful for characterization purposes, splitting ion current is severely detrimental for quantitative analysis. Most SRM-based analyses select only a single precursor-product transition due to the need to perform the same analysis for an internal standard, and also to have a short duty cycle so that the peptide peak has sufficient data points for integration purposes. The process of selecting a single precursor out of the multitude means that the analyst is immediately disposing of a significant amount of signal from the peptide. It is possible to monitor multiple transitions from the same peptide and sum the data, however, the increase in signal is often accompanied by a larger increase in background noise, negating any benefits. This sacrifice of sensitivity is the main downside of using LC–MS/MS platforms for peptide bioanalysis.

Discussions have taken place at many bioanalytical conferences on approaches to modify the precursor spectrum of large peptides to reduce the number of charge states, however no overall strategy has been presented. However, there were in fact articles in the

FUTURE SCIENCE part of fsg

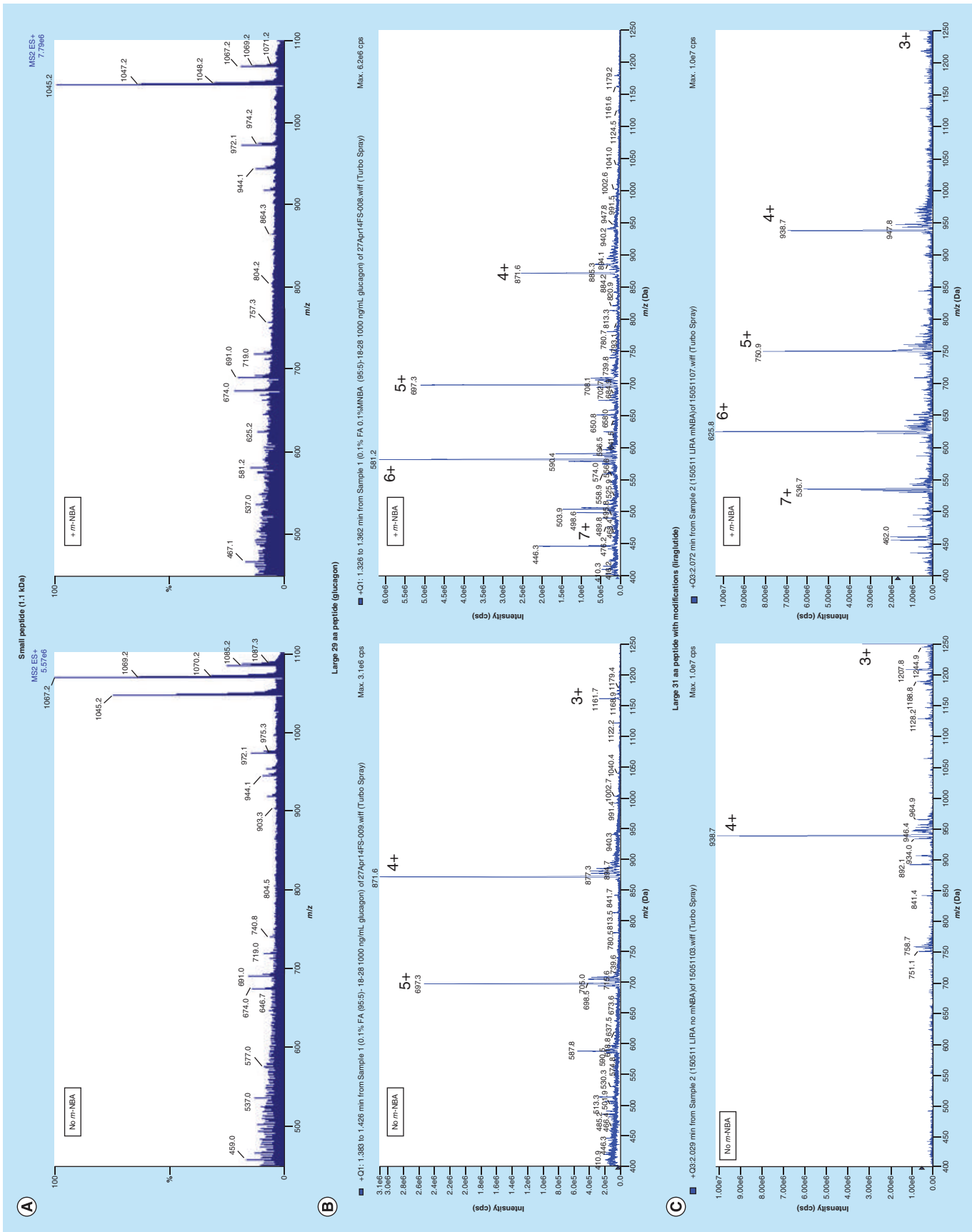

A
Small peptide (1.1 kDa)
MS2 ES+
5.57e6
No m-NBA
+ m-NBA
MS2 ES+
7.79e6
m/z
%
B
Large 29 aa peptide (glucagon)
+Q1: 1.383 to 1.426 min from Sample 1 (0.1% FA (95:5)- 18-28 1000 ng/mL glucagon) of 27Apr14FS-009.wiff (Turbo Spray)
Max. 3.1e6 cps
+Q1: 1.326 to 1.362 min from Sample 1 (0.1% FA 0.1%MNBA (95:5)-18-28 1000 ng/mL glucagon) of 27Apr14FS-008.wiff (Turbo Spray)
Max. 6.2e6 cps
Intensity (cps)
m/z (Da)
4+
5+
6+
7+
3+
C
Large 31 aa peptide with modifications (liraglutide)
+Q3:2.029 min from Sample 2 (150511 LIRA no mNBA)of 150511103.wiff (Turbo Spray)
Max. 1.0e7 cps
+Q3:2.072 min from Sample 2 (150511 LIRA mNBA)of 150511107.wiff (Turbo Spray)
Max. 1.0e7 cps

Figure 1. Charge state envelopes of three peptides with and without the presence of *m*-NBA (see facing page).
(A) A small peptide. **(B)** Glucagon, a 29 aa peptide. **(C)** Liraglutide, a 31 aa palmitolyated peptide.

literature discussing a group of mobile phase modifiers which could be added to increase the charge states of compounds used in normal LC–MS conditions. The earliest manuscript discussing the use of these compounds was in fact back in 2000, where Lavarone *et al.* discussed the effect of ethylene glycol and 2-methoxyethanol on cytochrome C and myoglobin [1]. The same group published further work in 2001, where they demonstrated the supercharging of cytochrome C using glycerol and *m*-nitrobenzyl alcohol (*m*-NBA), and showed that the later was a more effective 'supercharger' at lower concentrations [2]. Another group [3] looked at the potential use of *m*-NBA, dimethyl sulphoxide (DMSO), and sulpholane and how they effected the charge state of cytochrome C and other small proteins. These groups discussed the use of superchargers mainly for increasing charge states on proteins, however did not mention the potential for quantitation. There is currently a single manuscript in the literature discussing the use of superchargers (*m*-NBA) to improve sensitivity for peptide LC–MS/MS analyses, where Halquist *et al.* used *m*-NBA to influence the charge states of the gut peptide Oxytomodulin [4] and its effect on sensitivity. The application of these superchargers has somehow escaped the majority of the bioanalytical community, and therefore their potential benefits for large and small molecules have gone amiss. This commentary discusses the application of superchargers, in-particular *m*-NBA, and how they might affect peptide LC–MS/MS analyses for both large and small peptides.

What are superchargers & how do they work?

Lavarone and Williams hypothesized that superchargers increase the surface tension of the droplets in the electrospray source, which in turn increases the Rayleigh limit, decreasing the radius of the droplet prior to coulombic explosion [5]. This decreased droplet volume would result in a concomitant increase in charge density, potentially ensuring that readily chargeable groups are in fact charged. This would therefore increase the intensity of the more highly charged species in the MS, and reduce the relative intensity of lower charge states. The initial work by Lavarone *et al.* on superchargers described the use of ethylene glycol and 2-methoxyethanol [1]. They later assessed glycerol as a supercharger however they found that *m*-NBA was more effective at lower concentrations and was also an MS friendly reagent [2]. Other supercharging solvent modifiers have been investigated, where Lomeli *et al.* tested a number of compounds structurally similar to *m*-NBA, some of which gave similar or improved supercharging characteristics for myoglobin at a lower overall concentration [6].

The effect of superchargers on intact peptide spectra

The effect of superchargers on the precursor spectrum of peptides has been assessed on a number of peptides to date, and in every case, significant changes to the peptide spectra have been detected, especially with larger peptides where higher numbers of charges have been observed when supercharges have been used. Example spectra from a small peptide (1.1 kDa), glucagon (a 29 amino acid, 3.5 kDa peptide) and liraglutide (a 31 amino acid, 3.8 kDa peptide with a palmitoyl modification) are included in Figure 1.

The data presented in Figure 1 show that *m*-NBA had significant effects on the charge state envelope of the two larger peptides. For glucagon, the addition of *m*-NBA dramatically reduced the level of the 4+ charge state, and completely abolished evidence of the 3+, however it results in the previously absent 6+ charge state being the most intense peak. For liraglutide, the change in the precursor spectrum is the most significant shift we have seen yet, where the 4+ charge state was heavily dominant in normal conditions, but with *m*-NBA present, 5, 6 and 7+ charge states appear at significant levels. This could suggest that liraglutide (a GLP-I analog with a palmitoyl side chain) has poor ionization efficiency in normal LC–MS/MS conditions. The peptide elutes at high organic solvent conditions (for a peptide) in the region of 45% ACN, therefore the low efficiency is not due to a low concentration of organic solvent. This suggests that the palmitoylation is reducing the ability of the peptide to acquire significant numbers of charges, limiting it to three or four charges. The addition of *m*-NBA enables the peptide to acquire significantly higher numbers of charge states, and therefore suggests that its addition changes the ionization process sufficiently to overcome the dampening effect of the palmitoylation.

The effect of *m*-NBA on the small peptide is slightly different than for the other larger peptides, as no obvious change in charge state distribution can be seen in the presence of *m*-NBA. However, the addition of *m*-NBA gave a twofold increase in intensity as the abundance of the sodium adduct ion (*m*/*z* 1067.2) and the potassium ion adduct (*m*/*z* 1083.2) was significantly reduced, and the production of the proton adduct ion was favored.

Superchargers & their application to peptide quantitation

Small peptide bioanalysis

Small endogenous peptides, such as vasopressin and oxytocin, are very potent and therefore need to be quantified at low pg/ml levels, thus there is always an interest to find ways to obtain more sensitivity for such peptides. Even though the benefits of using superchargers to increase sensitivity on smaller peptides is less obvious, their utility was assessed, as it would be a very simple way to obtain more sensitivity.

As is displayed in Figure 1, the change in charge state distribution after addition of *m*-NBA was completely different for the small peptide compared with the larger peptides. And when quantifying the singly charged peptide in the presence of 0.05% *m*-NBA, it increased the peptide peak area response by approximately 50%. However, the background noise was also elevated with the addition of *m*-NBA, so the benefit in signal-to-noise was less pronounced. Despite achieving similar S/N ratios, the intra-assay precision and accuracy results (at four concentration levels) with and without supercharger indicated a gain in precision after the addition of the *m*-NBA, especially at the LLQ level (5.00 pg/ml). The value of the precision was 6 and 13%, respectively. This suggests that the *m*-NBA improved the robustness of the ionization of the peptide at the LLOQ, despite the similar S/N value.

Large peptide bioanalysis

The potential benefit of using superchargers for large peptide quantitation is clear, as the reduction in the number of charge states for a large peptide could increase the proportion of the total peptide ion current that is being monitored by the SRM method. This has been demonstrated in buffer standards, however this does not always translate into sensitivity gains when analyzing extracted plasma samples. This is believed to be due to the superchargers increasing the signal of background noise as well as peptide. This was observed in the analysis of endogenous glucagon levels using LC–MS/MS, where an existing formic acid mobile phase approach [7] was equivalent in terms of sensitivity to the *m*-NBA approach, however, the background noise associated with the *m*-NBA analysis was increased, resulting in a decrease in the overall S/N ratio. For this reason, the use of *m*-NBA for glucagon analysis was not progressed.

To date, LGC have taken one large peptide (~4 kDa in size) through a regulated validation and sample analysis using *m*-NBA as a mobile phase modifier. The *m*-NBA was required to improve the sensitivity of the method so that a challenging LLOQ could be reached. The SRM method for the supercharged peptide used a higher charge state than the normal analysis approach, which gave a boost in sensitivity. The validation of the method was successful, and generated very good inter-batch precision (4.2 and 5.1% at LLOQ) and accuracy values (2.4 and 3.1% at LLOQ). The method was then used successfully for a clinical sample analysis project.

How simple is testing *m*-NBA for peptide analysis?

The process of testing whether a supercharger makes a positive impact is extremely simple and can be performed quickly. A bottle of the existing organic mobile phase with *m*-NBA present can be set up on a secondary line on the LC system, and if looking at singly charged molecules, the analyst can perform a comparison of the effect of solvents with and without superchargers within minutes. If a peptide is being analyzed that has previously shown to become multiply charged, then a full scan analysis would need to be performed to assess the presence of additional multiply charged species. The fragmentation pattern of these higher charged peptides would need to be investigated and suitable SRM transitions developed to obtain the best sensitivity for comparison purposes. However, since no shift in retention time has been demonstrated when *m*-NBA has been present in the organic phase, the analyst will know where their analyte should elute, aiding the decision as to whether *m*-NBA has had a positive impact on the signal to noise.

Overall conclusion

Having assessed the use of superchargers on a number of peptides (both large and small), it suggests that *m*-NBA increases the ionization of everything in the source. Therefore the cleanliness of the extract can play a major part in whether an increase in analyte signal to noise is likely. If an increase in signal to noise has been observed, it is likely that the use of *m*-NBA could be carried through to validation and sample analysis as using superchargers is robust and reproducible. Furthermore, no detrimental effect on mass spectrometers has so far been detected, as solvent diversion has always been used such that the solvent flow only enters the MS system when it is required, in order to keep the instrument clean. To date, the only supercharger that has been tested is *m*-NBA and it is possible that other modifiers could result in larger sensitivity gains, but this remains to be investigated.

Financial & competing interests disclosure

The authors have no relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript. This includes employment, consultancies, honoraria, stock ownership or options, expert testimony, grants or patents received or pending, or royalties.

No writing assistance was utilized in the production of this manuscript.

References

1 Lavarone AT, Jurchen JC, Williams ER. Effects of solvent on the maximum charge state and charge state distribution of protein ions produced by electrospray ionization. *J. Am. Soc. Mass Spectrom.* 11(11), 976–985 (2000).

2 Lavarone AT, Jurchen JC, Williams ER. Supercharged protein and peptide ions formed by electrospray ionization. *Anal. Chem.* 73(7), 1455–1460 (2001).

3 Miladinovic SM, Fornelli L, Lu Y, Piech KM, Girault HH, Tsybin YO. In-spray supercharging of peptides and proteins in electrospray ionization mass spectrometry. *Anal. Chem.* 84, 4647–4651 (2012).

4 Halquist MS, Sakagami M, Karnes HT. Determination of oxyntomodulin, an anorectic polypeptide, in rat plasma using 2D-LC–MS/MS coupled with ion pair chromatography. *J. Chrom. B* 903, 102–111 (2012).

5 Lavarone AT, Williams ER. Mechanism of charging and supercharging molecules in electrospray ionization. *J. Am. Chem. Soc.* 125(8), 2319–2327 (2003).

6 Lomeli SH, Peng IX, Yin S *et al.* New reagents for increasing ESI multiple charging of proteins and protein complexes. *J. Am. Soc. Mass Spectrom.* 21, 127–131 (2010).

7 Howard JW, Kay RG, Tan T *et al*. Development of a high throughput UHPLC–MS/MS (SRM) method for the quantitation of endogenous glucagon from human plasma. *Bioanalysis* 6(24), 3295–3309 (2014).

Research Article

Bioanalysis

Development of an ultra-sensitive Simoa assay to enable GDF11 detection: a comparison across bioanalytical platforms

Background: Four bioanalytical platforms were evaluated to optimize sensitivity and enable detection of recombinant human GDF11 in biological matrices; ELISA, Meso Scale Discovery, Gyrolab xP Workstation and Simoa HD-1. **Results & methodology:** After completion of custom assay development, the single-molecule ELISA (Simoa) achieved the greatest sensitivity with a lower limit of quantitation of 0.1 ng/ml, an improvement of 100-fold over the next sensitive platform (MSD). **Discussion & conclusion:** This improvement was essential to enable detection of GDF11 in biological samples, and without the technology the sensitivity achieved on the other platforms would not have been sufficient. Other factors such as ease of use, cost, assay time and automation capability can also be considered when developing custom immunoassays, based on the requirements of the bioanalyst.

Maria Myzithras*[,1], Hua Li[1], Tammy Bigwarfe[1], Erica Waltz[1], Priyanka Gupta[1], Sarah Low[1], David B Hayes[1], Scott MacDonnell[1], Jennifer Ahlberg[1], Michael Franti[1] & Simon Roberts[1]

[1]Boehringer Ingelheim Pharmaceuticals, 900 Ridgebury Road, Ridgefield, CT 06877, USA

*Author for correspondence: maria.myzithras@boehringer-ingelheim.com

First draft submitted: 9 October 2015; Accepted for publication: 29 January 2016; Published online: 26 February 2016

Keywords: ELISA • GDF11 • Gyrolab • immunoassay • ligand-binding assay (LBA) • MSD • Simoa

In the past several years there have been many advances in ligand-binding assays that have been replacing traditional plate-based ELISAs with increased automation and sensitivity for detecting proteins [1]. Some of these new methodologies and technologies include the Meso Scale Discovery (MSD) platform, Gyrolab xP Workstation and the single-molecule ELISA (Simoa) HD-1 Analyzer. Each platform has unique features as well as advantages and disadvantages depending on the user's desired bioanalytical data output.

Our aim was to develop a sensitive and specific bioanalytical assay to detect GDF11 in biological matrices. GDF11 is a homodimer in the TGFB2 superfamily which is thought to circulate at subnanogram concentrations in various biological species [2,3] and developing a custom immunoassay with such low sensitivity requirements can be difficult. Previous publications have attempted to quantitate GDF11 with Western blots [2], which are only qualitative, or with nonspecific SOMAmers [4], which also detect closely-related GDF8 protein (myostatin), highlighting the need for a custom immunoassay, to be developed in-house. A monoclonal anti-GDF11 antibody from R&D Systems (MN, USA) was purchased, qualified and used as both a capture and detection reagent for all assays across multiple platforms to determine their respective sensitivities and dynamic ranges. Human recombinant GDF11 was also purchased from R&D systems and qualified in house for use in assay development. The four platforms evaluated were: ELISA, MSD, Gyrolab xP Workstation and Simoa HD-1.

Although there are several commercially available kits designated to quantitate human GDF11, they either lacked specificity for active (i.e., homodimeric) human recombinant GDF11 and/or selectivity against human recombinant GDF8 (myostatin) [3].

FUTURE SCIENCE part of fsg

Materials & methods

Antibody & reagent evaluation

Several anti-GDF11 antibodies are commercially available from vendors however, after in-house evaluation many of these antibodies were found to cross-react with GDF8 [5] as others have determined as well [4]. R&D Systems anti-GDF11 monoclonal mouse IgG1 antibody (Cat# MAB1958, clone 743833) was specific in binding to human recombinant GDF11 (R&D Systems cat 1958-GD/CF) and was not cross reactive to human recombinant GDF8 [5]. Due to the limited number of anti-GDF11 reagents available the R&D systems antibody was used as both a capture and detection reagent in all assay formats discussed, which likely precludes the measurement of GDF11 when bound to any of its numerous binding partners in biological matrices. In addition to human GDF11 the antibody can also recognize mouse and rat GDF11 since the active form of the protein is 100% homologous across these three species. Because of this homology, custom in-house antibody generation would be difficult and time consuming; therefore, efforts were focused on utilizing the one commercially available tool for assay development, even though it may not be the most ideal situation.

Reagent qualification

Recombinant human/mouse/rat GDF11 protein was purchased from R&D Systems, catalog # 1958-GD (lots MMW1314061 and MMW1314081). The recombinant protein was *E. coli*-derived Asn299-Ser407, with an N-terminal methionine. Sequence was confirmed in-house by mass spectrometry. The two lots were prepared from the same bulk batch and purchased lyophilized in separate months. The lyophilized powder was reconstituted to approximately 0.3 to 0.5 mg/ml with 30 mM Acetate, pH 4.5 buffer. Because GDF11 is not soluble in standard SEC running buffers, sedimentation velocity analytical ultracentrifugation (SV-AUC) was used to assess the aggregation state of the reagent. The SV-AUC experiment was performed on a Beckman Coulter Proteome Lab XLI. Sample was loaded into a charcoal filled epon cell with matching buffer in the reference channel, and 100 scans of concentration versus radial position were collected at 60,000 rpm. The data were analyzed by the continuous c(s) model using SEDFIT and GDF11 was found to be >95% disulfide linked dimer in the pH 4.5 by AUC, thus confirming the quality of the protein.

The anti-human GDF11 antibody was purchased from R&D Systems, catalog number MAB19581 (lot CGIM02). It is a monoclonal mouse IgG1 Clone # 743833 which was raised against *E. coli* derived recombinant human GDF11 (Asn299Ser407). Analytical size exclusion chromatography was performed to assess antibody purity using a Waters Acquity UPLC system. The system was run at 0.5 ml/min for 5 min in 50 mM NaPhosphate 0.2 M L-arginine 0.05% NaAzide pH 6.8. Ten micrograms of material was injected into an Acquity UPLC BEH c18 column (1.7 µM Product# 186002350) and results were read at 280 nm. The antibody was found to be 93% monomeric and 6% aggregate, thus confirming the quality of the antibody.

Reagent labeling

After qualification, the antibody was labeled for various assays as follows**.**

Biotinylation for the ELISA, MSD and Gyrolab platforms was performed at a concentration of 1 mg/ml using a no-weigh EZ-Link Sulfo-NHS-LC-Biotin kit (Pierce Thermo Scientific) using a 1:10 antibody to biotin challenge ratio and 1 h incubation with shaking (600 rpm) at room temperature. The resulting biotinylated antibody was desalted using a 0.5 ml Zeba 7K MWCO spin column (Thermo Scientific), prewashed three-times with 300 µl of 1X PBS.

The anti-GDF11 antibody was also labeled at 1 mg/ml with an MSD Sulfo-Tag NHS-Ester (Meso Scale Discovery) using a 1:10 antibody to Sulfo ratio, and also an Alexa Fluor 647 label (Molecular Probes Labeling Kit, Life Technologies), both according to manufacturer's instructions. Both labeled antibodies were desalted using 0.5 ml Zeba 7 K MWCO spin columns (Thermo Scientific), prewashed three-times with 300 µl of 1X PBS.

For the Simoa instrument the antibody was biotinylated at a 1.0 mg/ml concentration using a 50X NHS-PEG Biotin (Thermo Fisher) ratio and desalted into Quanterix proprietary buffers following the manufacturer's recommendations. Antibody was then conjugated to proprietary paramagnetic beads (Quanterix) at 0.5 mg/ml using proprietary wash and resuspension buffers following the specific protocol provided by Quanterix. Bead count and aggregation were assessed using a 'bead aggregation protocol' on the Simoa instrument, and verified at above 80% monomeric for use in the assay.

Other common reagents used in various assay formats include 1X PBS, pH 7.4 (Gibco), Bovine Serum Albumin (BSA, SeraCare), Tween 20 (Sigma), Low Cross Buffer (Candor Biosciences) and pooled mouse serum or plasma (Innovative Research).

ELISA

A 96-well polystyrene microtiter plate (Nunc Maxisorp) for ELISA was coated with 50 µl per well of 2 µg/ml unlabeled anti-GDF11 antibody, overnight at 2–8°C. Unbound capture reagent was washed away with 300 µl per well wash buffer (1X PBS with 0.05%

Tween 20) three-times and the wells were blocked with 5% BSA with 0.05% Tween 20 and incubated at room temperature on a plate shaker (300 rpm, 1 h). The plate was washed three-times with wash buffer and 50 µl per well of standards and QCs in binding buffer (5% BSA with 0.05% Tween 20) were added and incubated at room temperature on a plate shaker (300 rpm, 1 h). The plate was washed three-times with wash buffer and 50 µl per well of biotinylated anti-GDF11 at a concentration of 2 µg/ml was added and incubated at room temperature on a plate shaker (300 rpm, 1 h). The plate was washed three-times with wash buffer and 100 µl per well horseradish peroxidase (HRP)-conjugated to Streptavidin (Jackson ImmunoResearch) was added at 1 mg/ml in binding buffer and incubated at room temperature on a plate shaker (300 rpm, 1 h). Following three washes, 50 µl per well of tetramethyl benzidine substrate was added causing an enzymatic reaction with any remaining HRP. The reaction was stopped with 50 µl per well 1 M H2SO4 and the optical density (OD) was measured using a SpectraMax M3 microplate reader at 450 and 650 nm dual wavelength. Softmax Pro software (v5.4) was used for calibration standard curve fitting using a 4-parameter logistic model and back calculation of all QC concentrations.

MSD

An MSD assay to detect human recombinant GDF11 was developed using a homogenous assay format. A master mix was prepared, combining 4.0 µg/ml biotinylated anti-GDF11 antibody and 2.0 µg/ml sulfo-labeled anti-GDF11 antibody in binding buffer (5% BSA in 1X PBS with 0.05% Tween 20). This master mix was added to a 96-well, nonbinding, light-blocking plate (Fisher Scientific) at 50 µl per well. Twenty-five microliter of standards and QCs (diluted in binding buffer) were added per well in duplicate to the nonbinding plate containing the master mix. The nonbinding plate was incubated at room temperature on a plate shaker (500 rpm, 1.5 h). In parallel, an MSD streptavidin (SA) gold 96-well plate (Meso Scale Discovery) was blocked using 150 µl blocking buffer (5% BSA with 1X PBS with 0.05% Tween 20) and incubated at room temperature on a plate shaker (500 rpm, 1.5 h). After incubation, the MSD plate was washed three-times with 300 µl per well wash buffer (0.05% Tween 20 in 1X PBS). Fifty microliter aliquots of each sample mix from nonbinding plates were added to MSD plates and incubated at room temperature on a plate shaker (600 rpm, 1.5 h). After incubation, the plate was washed three-times and 150 µl of 2x Read Buffer T (MSD, stock diluted 2X in H20) was added to each well and read immediately on the MSD Sector Imager 2400 using an electrochemiluminescent signal. Quality control concentrations were back-calculated using standard curves fitted using a four-parameter logistics equation using MSD Discovery Workbench software.

Gyrolab

A custom run was developed on the Gyrolab by first preparing biotinylated anti-GDF11 antibody at 100 µg/ml in Gyros wash buffer (1X PBS with 0.01% Tween 20) for use as capture reagent. Alexa Fluor 647 labeled anti-GDF11 antibody was spun at 14,000 g for 4 min before being diluted at a concentration of 20 nM in Low Cross Buffer for use as detection reagent. Ten microliters of standard curve and QC dilutions were prepared in Low Cross Buffer, and all reagents (including several wells of wash buffer) were transferred to 96-well PCR plates (Fisher Scientific). The PCR plates and Gyrolab BioAffy 200 nl CDs were loaded into the instrument carousels and the default three-step 'Wizard' assay protocol was run, which adds capture, analyte and detection in separate steps and washes between each step. The system runs the fully automated immunoassay at nanoliter scale (only 200 nl sample used but requires 4 µl dead volume) using a microfluidic system contained within custom 'CDs.' Each CD has microstructures that flow sample, reagents and wash through chambers and a streptavidin coated affinity column using centrifugal and capillary forces [6]. Each microstructure represents one sample and microfluidic interaction with its own inlet for reagents, with the detection step being laser-induced fluorescence detection. Quality control concentrations were back-calculated using standard curves fitted to a four-parameter logistics equation (at 5% PMT) using Gyrolab Evaluator software.

Simoa

An anti-GDF11 antibody was conjugated at a concentration of 0.5 mg/ml to paramagnetic beads for use as capture antibody and also biotinylated at a 60X biotin ratio for use as detection antibody, following manufacturer's instructions, for use on the Single Molecule Array system (Simoa HD-1 Analyzer by Quanterix). Capture beads were diluted in Bead Diluent (Quanterix) at a concentration of 5.0 E+6 beads and detection antibody diluted in Detection Diluent (Quanterix) at a concentration of 1.8 µg/ml. All reagents were loaded into the instrument reagent bay, including Streptavidin B-galactosidase reporter enzyme (SBG, Quanterix) at 100 pM and one vial of resorufin-B-D-galactopyranoside substrate (RGP, Quanterix). Standards and quality controls and were diluted in 5% BSA with 0.05% Tween 20 and were then loaded (100 µl per replicate)

into the instrument in glass vials (VWR) or 96-well nonbinding plates (Nunc, Fisher Scientific). All sample incubations, washes (using a proprietary buffer system by Quanterix) and detection steps are performed using a 'two step homebrew assay' default protocol, all of which are fully automated. Capture beads are incubated with sample and detection antibody, are washed and re-suspended in RGP, and transferred to femtoliter volume micro-arrays on a custom Simoa 'CD.' The wells are sealed with mineral oil and the imaging process is initiated. Detection was accomplished using the SBG reporter enzyme and RGP substrate and subsequently read using time-lapsed fluorescence detection at multiple wavelengths to calculate the Average Enzymes per Bead (AEB), which also enables 'digital ELISA' readout [7]. Quality control concentrations were derived by plotting standard curve concentrations versus AEB in a five-parameter logistics equation with $1/y^2$ weighting using Simoa HD-1 Analyzer data reduction software.

Results & discussion

A summary of each bioanalytical assay format and platform investigated are shown in Figure 1.

Assay qualification & performance

The ELISA, MSD and Gyrolab assays were developed following nonregulated exploratory guidelines which included specificity, matrix interference, establishment of minimum required dilution (MRD) and quality control performance in pooled serum or plasma. Standard curves on these platforms were first assessed in buffer, and then in serum or plasma, and the MRD was determined using the highest percentage that showed the same curve overlay and performance compared with buffer. All standards and QC's for assay development for the Simoa were run in buffer, in absence of any serum due to interference from endogenous GDF11 because of the increased sensitivity of the assay compared with the others. It is common to use matrix alternatives when developing nonregulatory exploratory biomarker assays since validation guidelines are flexible, as it is a challenging area and many assays are fit-for-purpose [8]. The updated US FDA bioanalytical method validation guidance document includes only a short section on biomarker assay development guidelines, which discusses that scientists should attempt to validate in a method similar to PK assays. However, meeting the same qualifications may not be possible in all biomarker assays, so validation that is 'deemed appropriate' by the scientist will suffice [9]. A recent WRIB white paper discusses the challenges that are faced when developing biomarker assays and what validation steps can be acceptable during assay development. Since our reagents and reference protein were

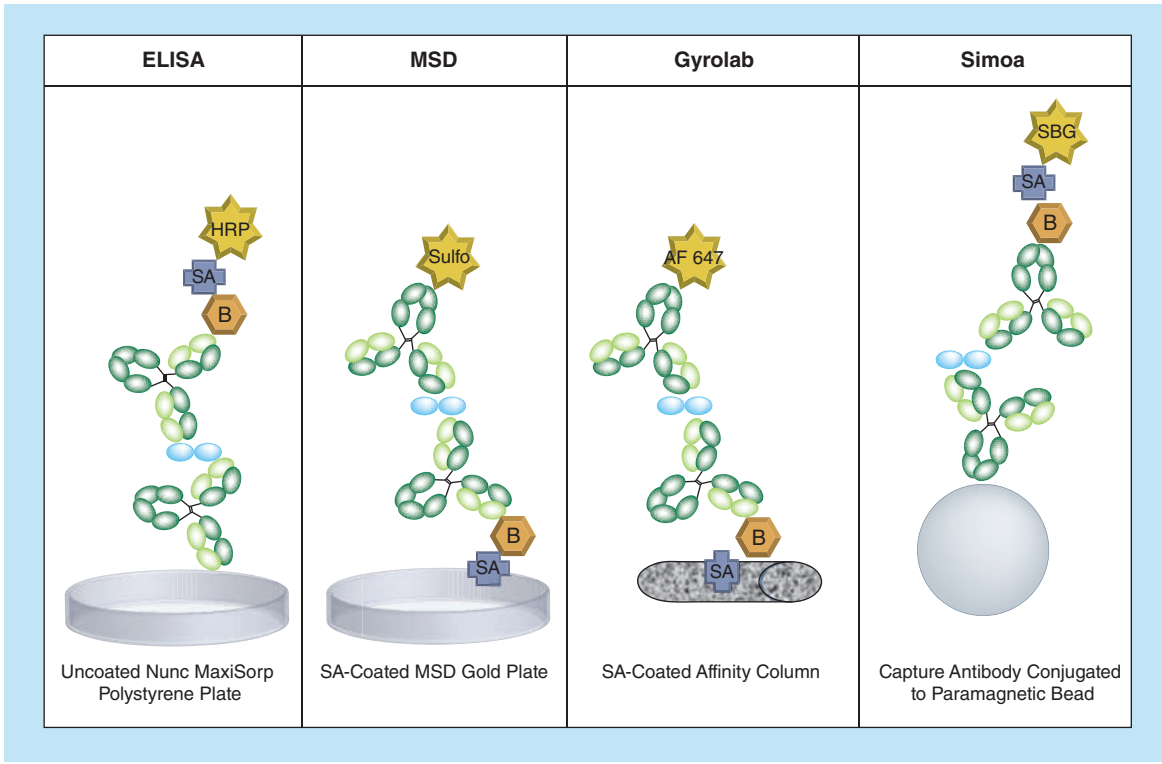


Figure 1. Illustration of each bioanalytical assay.

Table 1. Comparison of assay sensitivities, dynamic range and characteristics.

Platform	Dynamic range (ng/ml)	MRD	LLOQ (ng/ml)	ULOQ (ng/ml)	Minimum sample volume (µl)	QC Bias range (%, abs.)	Format and readout
ELISA	6–171	10	60	1710	50	10.2–26.4	Plate based, colorimetric
MSD	1.1–1093	10	11	10,930	25	0.3–27.3	Plate based, electro-chemiluminescence
Gyrolab	7.8–8000	5	39	40,000	4	0.7–21.6	Microfluidics affinity based, fluorescence
Simoa	0.05–30	2	0.1	60	100	0.0–24.0	Bead based, fluorescence and digital readout

MRD: Minimum required dilution.

well-characterized with rigorous QC testing, our Simoa assay method with appropriate controls meets sufficient exploratory validation criteria in a non-clinical setting for a quantitative assay, and buffer use for matrix was justified [8]. The MRD of Simoa was determined by sample dilution linearity testing. The estimated quantitation range of each assay was established using the LLOQ and ULOQ, which were determined by spiking in known values of recombinant GDF11 protein as quality controls and assessing recovery within ±30% accuracy of the nominal value and multiplying by MRD for each assay.

Sensitivity

The LLOQ and ULOQ for each assay format are reported in Table 1 and were determined using the accuracy (within ±30% nominal value) of spiked GDF11 and multiplying by MRD. Dilution linearity and MRD were determined on ELISA, MSD and Gyrolab by testing different mouse and human serum percentages with standard curve and QC performance. These results are shown in Table 1. The comparison of full standard curves representing each assay type is shown in Figure 2. The most sensitive assay format was found to be the Simoa with an LLOQ of 0.1 ng/ml in buffer, followed by the MSD with 11 ng/ml, Gyrolab at 39 ng/ml, and lastly the ELISA with an LLOQ of 60 ng/ml. Compared to the other platforms, the Simoa increases quantifiable range by approximately 100-fold, with the capability to deliver 3-logs worth of data. It is apparent in this figure that ELISA, MSD and Gyrolab all are approaching a signal over background of 1 at roughly 1–2 ng/ml (assay noise), and therefore values cannot be quantified within a linear range with confidence, which is evident by the established dynamic ranges in Table 1.

Sample volume requirement

Although the Simoa assay was most sensitive it required the highest sample volume (100 µl) per replicate, which can be limiting when analyzing rodent, preclinical samples and valuable clinical samples. The Gyrolab required the least amount of sample as it only uses 0.2 µl per replicate (dead volume 4 µl) in the microfluidics system, which would be optimal for microsampling and rodent studies, however the sensitivity was compromised at 39 ng/ml. Gyros also offer 1000 nl CDs, which load more sample volume into the affinity column and chambers for increased sensitivity, but no improvement was seen when developing our custom assay (data not shown). The LLOQ of MSD was 11 ng/ml, with a sample requirement of 25 µl, which could be a good compromise if reagent volume is a limiting factor. The ELISA has a sample requirement of 50 µl, which can still be limiting with rodent studies, and coupled with reduced sensitivity; it is not a good alternative in this case.

Cost, assay development & run time

There are factors in addition to sensitivity that need to be considered when selecting an immunoassay which are described in Table 2. Each custom assay requires reagent modifications and proprietary components which can increase cost and various assay development and run times.

The GDF11 methods developed on all platforms had similar custom assay development times of approximately 1–3 weeks depending on the desired assay; however each requires multiple reagent modifications and multiple proprietary components. A traditional sandwich ELISA may not require any additional reagent modification or labeling, but in our case biotinylated detection was used because of the lack of available reagents. Generic antispecies antibodies already coupled to HRP can be purchased in many cases, which can improve on reagent cost and assay development time. The other platforms all require biotinylation and an additional fluorescent labeling (MSD, Gyrolab) or bead conjugation (Simoa). Bead conjugation is the most time consuming reagent modification

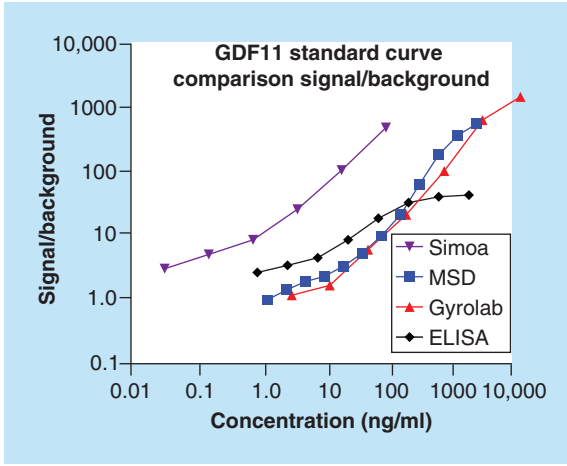


Figure 2. Comparison of GDF11 standard curves with each bioanalytical assay. Plotted as a function of mean signal/background due to the differences in data readout (Gyrolab = Response, MSD = RLU, Simoa = AEB, ELISA = OD).

step, with the conjugation protocol and bead aggregation verification run requiring almost an entire day of development time.

ELISA and MSD are both plate based and require more hands-on bench time, with washing and liquid handling done by hand or with incorporating some automated steps. Automation can be incorporated to aid in blocking, washing or sample addition steps, but the full time required to complete the entire assay still remains the same, at 3–6 h, due to fixed incubation times, which are longer at 1–2 h each. The Gyrolab and Simoa are fully automated with shorter run times (reflected in the instrument technology which requires shorter incubation times compared with plate-based methods), but these conveniences are offset by the higher cost of operating the instruments. The Gyrolab requires proprietary CDs that are approximately three-times more expensive than MSD plates. The Gyrolab also has optional sample and detection buffers available for purchase which can drive up cost; however they are not necessary components for methods development. The Simoa CDs are roughly the same cost as Gyrolab, however many more proprietary items such as reagent buffers, wash buffers and consumables (tips, cuvettes, sealing oil) add to the cost of operating the instrument. The Simoa buffers cannot be substituted due to the specific salt, pH and detergent requirements of each component, which are custom designed for the instrument automation and sample processing/imaging.

Focus on Simoa technology

The Simoa HD-1 by Quanterix is a newer platform, with mainstream implementation in several large pharmaceutical and biotech companies in the past several years. The increased sensitivity that is being reported in their marketed assay kits and 'homebrew' custom assays, compared with other platforms, can be achieved due to their unique bead-based technology. The bead-based antibody-analyte capture can serve as sample enrichment, and the resulting immunocomplexes (antibody conjugated to bead, analyte, biotinylated detection antibody, and SGB enzyme) are trapped in individual femtoliter-sized wells in single arrays (single-molecule array or SiMoa) on custom CDs. There are 216,000 wells per array (24 arrays per CD) available to accept each bead, with or without an immunocomplex present, therefore detection is theoretically possible down to a single molecule per well. A time-lapsed fluorescence signal is produced which is imaged by microscope optics (creating a 'digital' readout), and because each immunocomplex is confined to one 50-fl sized well, the signal produced is much more concentrated compared with other LBA plate-based assays, which can lead to better assay sensitivity. Plate based assays such as ELISA or MSD have a much higher surface area and volume per well, therefore the signal produced by an enzymatic/substrate reaction is diffused and higher amounts of target analyte is required before producing adequate signal over background to accurately quantitate [7].

Conclusion

We were successful at developing a custom assay to detect human recombinant GDF11 that met our sensitivity requirements after evaluating four methodologies; ELISA, MSD, Gyrolab and Simoa. The Simoa increased our sensitivity compared with the other three platforms by approximately 100-fold and was used for sample analysis. There are many other aspects that need to be considered when developing custom immunoassays to detect various analytes in biological matrices such as assay development and run time, cost, reagent modification and automation capability, that should be considered when developing custom immunoassays across multiple platforms . The Simoa outperformed other immunoassay formats with an LLOQ of 0.1 ng/ml, but comes with an added cost and more labor intensive reagent modifications. However, even with these limitations, the Simoa enabled GDF11 detection in human and mouse serum samples. Without this technology, the sensitivity achieved on the other three platforms would not have been sufficient.

Future perspective

The Simoa accomplished the desired sensitivity however the bead based nature of the assay also has downsides, and one of the largest limitations already

Table 2. Comparison of platforms to guide immunoassay selection.

Platform	Approx. assay development time	Reagent modifications	Proprietary components	Approx. assay run time	Approx. cost	Fully automated	Custom multiplexing capability
ELISA	1–2 weeks	Labeled detection	None	3–5 h	+	No	No
MSD	2–3 weeks	Labeled capture and detection	Microplates, read buffer	3–6 h	++	No	Yes
Gyrolab	1–2 weeks	Labeled capture and detection	CDs, optional sample buffers	1–2 h	+++	Yes	No
Simoa	2–3 weeks	Conjugated capture, labeled detection	Reagent buffers, wash buffers, CDs, paramagnetic beads, consumables	2–3 h	++++	Yes	Yes

Run time and approximate cost are calculated based on one 96-well plate of samples.

mentioned is the sample volume requirement of 100 µl per replicate. One future advantage of the Simoa is the ability to multiplex [10], which could offset the limitations concerning volume, which other platforms currently do not offer with such increased sensitivity. Homebrew multiplex kits are available to develop custom assays to detect up to 3–6 analytes in one 100 µl sample volume, with the same increased sensitivity compared with traditional immunoassays. Since the technology is newer, it is anticipated that it will only continue to improve in the future, with the instrument gaining more features and less volume and reagent requirements. Coupled with the sensitivity increase and ability to multiplex, we think this instrument will be crucial for enabling critical biomarker detection as well as custom low-level protein detection in the near future.

Financial & competing interests disclosure

The authors have no relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript. This includes employment, consultancies, honoraria, stock ownership or options, expert testimony, grants or patents received or pending, or royalties.

No writing assistance was utilized in the production of this manuscript.

Ethical conduct of research

The authors state that they have obtained appropriate institutional review board approval or have followed the principles outlined in the Declaration of Helsinki for all human or animal experimental investigations. In addition, for investigations involving human subjects, informed consent has been obtained from the participants involved.

Executive summary

- Four bioanalytical ligand-binding assay-based platforms were evaluated to optimize sensitivity and enable detection of recombinant human GDF11 in biological matrices; ELISA, MSD, Gyrolab xP Workstation and Simoa HD-1.
- After completion of custom assay development on each platform, it was found that the single-molecule ELISA (Simoa) achieved the greatest sensitivity.
- The Simoa had an LLOQ of 0.1 ng/ml, an improvement of approximately 100-fold over the next sensitive assay (MSD).
- Coupled with the increased sensitivity, automation capability and ability to multiplex, the Simoa will be an important bioanalytical tool in the future to enable critical biomarker detection

References

Papers of special note have been highlighted as:
• of interest; •• of considerable interest

1 Mora J, Given Chunyk A, Dysinger M *et al.* Next generation ligand binding assays-review of emerging technologies' capabilities to enhance throughput and multiplexing. *AAPS J.* 16(6), 1175–1184 (2014).

• **Highlights multiple platforms for ultrasensitive detection of proteins. This review helped guide our decision to invest in the Simoa HD-1 as opposed to other platforms.**

2 Loffredo FS, Steinhauser ML, Jay SM *et al.* Growth differentiation factor 11 is a circulating factor that reverses age-related cardiac hypertrophy. *Cell* 153(4), 828–839 (2013).

•• **Highlights the original thought that recombinant GDF11 therapy can have beneficial cardiac effects in aged mice; however, GDF11 protein is 'quantitated' using a western blot, which is only a qualitative assessment. This highlighted the need for a custom assay to accurately measure protein concentration, which did not exist yet.**

3 Rodgers BD, Eldridge JA. Reduced circulating GDF11 Is unlikely responsible for age-dependent changes in mouse heart, muscle, and brain. *Endocrinology* 156(11), 3885–3888 (2015).

4 Egerman MA, Cadena SM, Gilbert JA *et al.* GDF11 increases with age and inhibits skeletal muscle regeneration. *Cell Metab.* 22(1), 164–174 (2015).

•• **Highlights the fact that reagents used in several other papers (SOMAmers, various antibodies) were nonspecific and also detected GDF8. An MSD assay was established with the same reagents (R&D Systems) but they report higher sensitivity; however no standard curve, QC or assay qualification was shown, so there was some doubt with reported results and LLOQ.**

5 Smith SC, Zhang X, Zhang X *et al.* GDF11 does not rescue aging-related pathological hypertrophy. *Circ. Res.* 117(11), 926–932 (2015).

•• **Puts the original Loffredo *et al.* GDF11 publication into doubt, and also shows our Simoa assay enabled the detection and quantitation of GDF11 dosed in mouse at 0.1 mg/kg IP daily, which the previous paper could not show.**

6 Mora JR, Obenauer-Kutner L, Vimal Patel V. Application of the Gyrolab platform to ligand-binding assays: a user's perspective. *Bioanalysis* 2(10), 1711–1715 (2010).

7 Rissin DM, Kan CW, Campbell TG *et al.* Single-molecule enzyme-linked immunosorbent assay detects serum proteins at subfemtomolar concentrations. *Nat. Biotechnol.* 28(6), 595–599 (2010).

8 Amaravadi L, Song A, Myler H *et al.* 2015 White Paper on recent issues in bioanalysis: focus on new technologies and biomarkers (Part 3 – LBA, biomarkers and immunogenicity). *Bioanalysis* 7(24), 3107–3124 (2015).

9 US Department of Health and Human Services, US FDA, Center for Drug Evaluation and Research, Center for Veterinary Medicine. Guidance for Industry, Bioanalytical Method Validation. (2013). www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm368107.pdf

10 Rivnak AJ, Rissin DM, Kan CW *et al.* A fully-automated, six-plex single molecule immunoassay for measuring cytokines in blood. *J. Immunol. Methods* 424, 20–27 (2015).

• **Highlights the Simoa technology and the theory behind how the assay works, and how increased sensitivity is achieved.**

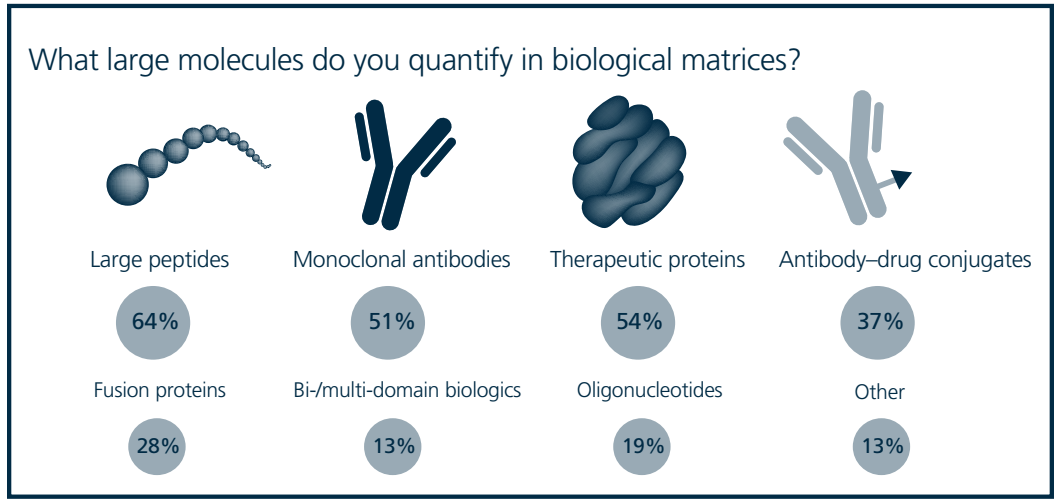
Spotlight: Large molecule analysis by LC–MS

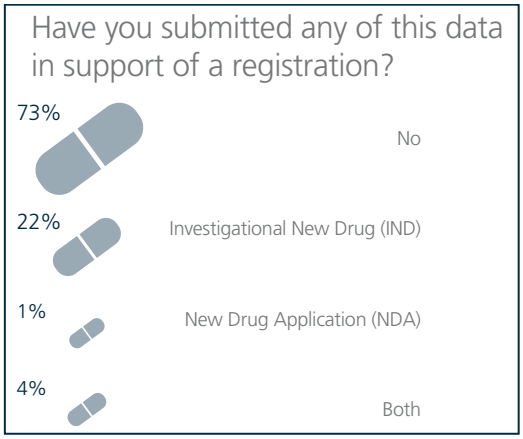
Spotlight summary written by Faye Vazvaei (Roche Innovation Center, NY, USA)

As the complexity and diversity of biological therapies increases, the application of MS in quantifying large molecules in biological fluids has also grown. While the use of LC-MS/MS for the analysis of small molecules and small peptides and the use of ligand binding assays (LBAs) for the analysis of protein and large peptides have matured into the leading quantitative technologies over the last couple of decades, the application of MS-based methodologies to quantify proteins and large peptides is slowly but surely finding its place in various bioanalytical laboratories.

Bioanalytical scientists and practitioners are continually looking for new ways and developing new methods to answer the questions raised by project teams regarding the kinetics of the large molecule *in vivo*. LC-MS has offered scientists an economical and efficient way of developing methods at an early discovery stage without the need for highly specific affinity reagents. It has also evolved as an orthogonal approach for the investigation of biotherapeutics when there are cross-reactivity issues with the LBAs – and in the area of biotherapeutics disposition and biotransformation *in vivo*. As the use of this technology is growing within the pharmaceutical industry, Bioanalysis Zone conducted a survey on Large Molecules Analysis by LC-MS. This survey provides a snapshot of the current industry practices in the area of large molecules bioanalysis by LC-MS.

Reflecting on the survey results, it quickly becomes apparent that the 103 respondents work with multiple types of molecules. The majority of respondents work on large peptides (64%) followed by monoclonal antibodies (51%) and antibody–drug conjugates (37%). Approximately 13% of the respondents indicated that they quantify oligonucleotides, and another 13% indicated that they measure a variety of other molecules ranging from protein biomarkers, glycerides and pesticides to antioxidants and plant protein. It is remarkable to see the use of MS in quantitative analysis within the diversity of these modalities.



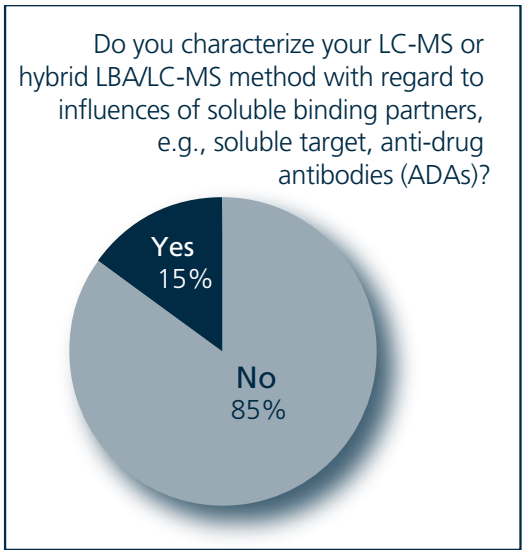


It was also interesting to see that this survey reached many countries across the globe. The survey respondents were from 18 countries, with the USA comprising the plurality of those responding to the question (41%). Other respondents indicated their country of residence as India (11%), The Netherlands (8%), Canada (7%), Japan and France (6% each), the UK, China, Germany and Iceland (3% each), and Andorra, Belgium, Brazil, Denmark, New Zealand, Republic of Korea, Spain and Sweden (1.4% each). The industry demographics were varied with pharma (including generic companies) constituting 28% of all respondents, followed by CROs (CROs) at 22% and academia at 13%. Other responding demographics included biotechnology, equipment vendor, and government research and food supplements.

The majority of the respondents stated that they work in a regulatory environment, with only 25% working exclusively in the non-regulated space. Approximately three-quarters (74%) of the respondents use their bioanalytical assays during the preclinical stage. This is not surprising, given the rapid speed of method development needed during the discovery phase. As expected, this number declines to 45% for early clinical development and 22% at late stage (phase III), where specific antibodies are more likely to be available for the LBA platform. The survey reveals the growing need for biosimilar assays: approximately 23% of the respondents utilize the assay for biosimilars.

Of those working in the regulated space, only three respondents (4.3% of all respondents) indicated that they have submitted LC-MS data in support of both IND and NDA applications, whereas approximately 22% indicated that they have submitted data only in support of IND applications. This indicates that the use of MS as a quantitative application for large molecules is still developing and is in its expansion phase. Surprisingly, only one respondent submitted data exclusively in support of an NDA application. From the question and the responses given in the survey, it is not clear what type of data was submitted in support of either IND and/or NDA application.

One of the questions on the minds of many advanced and beginner users of the technology is whether a specific strategy should be followed in selecting the bioanalytical platform at different stages of drug development. Approximately 55% of the respondents indicated that they do follow a strategy in their decision making.



These considerations may include the lack of availability of specific antibodies for the protein of interest, or the requirement for obtaining information to understand better the compound's pharmacokinetics. At first glance, it might sound surprising that the other 45% indicated that they do not have a strategy in choosing their assay platform. A closer look at the survey may indicate that the respective respondents may be supporting 'generic'-type assays in the preclinical stage, CROs supporting a sponsor's methodology for a particular application or involved in collaborative activities (academia and instrument vendors).

Like any technology, there are advantages as well as disadvantages in the use of LC-MS in large molecule quantitative bioanalysis. The advantages listed by the survey respondents seem

to be enough of an incentive for approximately 73% of them to believe that the use of MS technique with or without immunoaffinity capture will increase in the future. The perceived advantages of this technique by the respondents are: faster method development, improved specificity, wider dynamic range, ability to multiplex, the possibility to differentiate between different variants and addressing LBA cross-reactivity issues, among others. When asked, 85% of survey participants said that they do not characterize their assay with regard to influence of soluble binding partners. The perceived disadvantages are also numerous, with the availability of a stable isotopically labeled molecule ranking as one of the main challenges, together with the sample preparation issues and a lack of desired sensitivity. Perhaps the most important challenge for the quantitation via proteolytic digestion is to show that the measured surrogate peptide(s) is representative of the molecule. In addition, there is no regulatory guidance for the latter. Quite a few of those who responded (85%) said there is a need for a harmonized descriptive nomenclature in the large molecule quantitative space.

In conclusion, the survey shows that this technique is beyond its infancy, and is developing towards greater maturity in the coming years. As the industry moves forward in utilizing these technologies, it is hoped that the experience, both in the science and the regulatory filings, will be shared and the industry will collectively benefit from the opportunities that any proven technology provides: understanding the molecule and its fate *in vivo*.

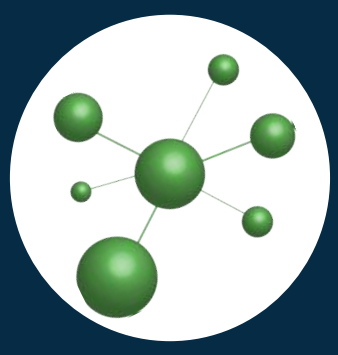
Key Points

- In recent years there has been an increase in large molecule therapeutics due to their target specificity, lower toxicity and higher potency, with many pharmaceutical companies now offering a vast portfolio of large molecule drugs.
- This increase in large molecule therapeutics has been accompanied by an increase in the use of LC–MS. Even though LBA's are still the gold standard for large molecule quantification, LC–MS has many benefits, such as a broader linear dynamic range allowing for multiplexing.
- There are still a range of challenges associated with the use of LC–MS for these compounds.

Bioanalysis ZONE

Join our free bioanalysis knowledge hub!

Explore – Stay up-to-date with the latest bioanalytical news, MEDLINE-indexed articles and videos



Connect – Share knowledge with a network of international experts from the bioanalysis community

Enhance – Learn through multimedia content, including webinars, spotlights and panel discussions

www.bioanalysis-zone.com

fsg CONNECT

Bioanalysis

Large-scale implementation of sequential protein and peptide immunoaffinity enrichment LC/nanoLC–MS/MS for human β-nerve growth factor

Background: A previously described immunoaffinity (IA)-LC-MS/MS assay for human β-nerve growth factor (β-NGF) was implemented to support large-scale sample testing for multiple clinical trials. **Methodology & results**: The procedure was modified to increase throughput by simultaneous preparation of two 96-well plates and LC duty-cycle reduction. Robustness of the LC method and nano-ESI was ensured during large-scale assay execution by closely monitoring and, if needed, replacing system components prior to failure. Following validation, the assay was used to analyze approximately 19,000 samples from multiple clinical studies over several years. **Conclusion**: Routine implementation of the β-NGF IA-LC–MS/MS assay supported drug development programs. This optimized assay format now serves as a template for other clinical protein biomarker assays.

Gary A Schultz*[,1], Kathlyn McCardle[1] & Hendrik Neubert[2]

[1]Q Squared Solutions BioSciences LLC, Ithaca, NY 14850, USA

[2]Pharmacokinetics Dynamics & Metabolism, Pfizer, Inc., Andover, MA 01810, USA

*Author for correspondence:
Tel.: +1 607 273 6538
gary.a.schultz@gmail.com

First draft submitted: 9 December 2015; Accepted for publication: 22 February 2016; Published online: 23 March 2016

Keywords: antipeptide antibody column • biomarkers • extended SIL peptide • immunoaffinity • LC–MS • method development

Protein and peptide immunoaffinity (IA) techniques linked to LC–MS/MS have surfaced as powerful bioanalytical tools for the quantification of protein biomarkers [1–10]. Assays based on IA-LC–MS/MS offer good specificity, sensitivity and reproducibility, providing a high degree of data confidence, which ultimately enables robust decision-making with biomarker data [11]. With the advent of innovative IA workflows and more sensitive MS instrumentation, it has become possible to robustly quantify low-abundance biomarkers in the sub-ng/ml to low pg/ml range in serum and plasma [12–14]. Many of the reported assays have been demonstrated in preclinical or exploratory clinical investigations, oftentimes with sample numbers not exceeding a few dozen, or perhaps a few hundred, per study. However, there is an emerging need to utilize this technology in support of larger interventional clinical trials and other larger patient studies in order to take advantage of the discriminatory power of quantitative protein MS for clinical decision-making [15]. Depending on the study design and measured endpoints, such clinical studies can generate sample numbers in excess of a 1000 per study. This represents an opportunity to adapt the performance characteristics of such IA-LC–MS/MS assays to achieve higher throughput and to meet the robustness requirements for larger-scale clinical implementation.

Recently, an assay has been described for the quantification of total human β-nerve growth factor (β-NGF) concentrations in human serum using sequential protein and tryptic peptide IA enrichment followed by nanoLC–MS/MS [13]. β-NGF plays a role in pain modulation through interaction with p75 and TrkA neurotrophin receptors and has become a target for monoclonal antibody analgesics. NGF is comprised of two identical 13.5 kDa subunits and is found within

FUTURE SCIENCE part of fsg

a range of 15–50 pg/ml in serum of healthy subjects. Despite the complexity of the β-NGF assay format, the utility in translational research and early clinical studies has been demonstrated [13]. The described bioanalytical performance characteristics of this assay paved the way to enable its adaptation in larger clinical trials. To this end, further optimization of throughput and robustness was required, for example, via simultaneous preparation of more than a single 96-well plate and reduction of LC duty cycle.

Herein, we describe method considerations for the large-scale implementation of a multiday IA extraction and digestion method combined with multiday LC/nanoLC–MS/MS batch analysis. The analytical method must be designed to ensure robustness of an LC method with nanoESI for high-throughput analysis. Furthermore, the integrity of critical reagents from validation through completion of studies must be ensured. Validation of an optimized quantitative IA-LC–MS/MS assay for β-NGF quantification in human serum is described, which ultimately led to the implementation of the method to support clinical trials. The validated method was used to analyze approximately 19,000 clinical samples for multiple studies over several years. This example paves the way for broader implementation of this technology in routine clinical sample analysis of protein biomarkers in general.

Materials & methods

Chemicals & reagents

Recombinant human β-NGF protein was purchased as a lyophilized powder from R&D Systems (MN, USA) and prepared at a concentration of 200 µg/ml. The extended stable isotope-labeled (SIL) peptide standard AWRFIRIDTACVC(V*) (L*)SRKAVRRA, (V* [$^{13}C_5$ $^{15}N_1$], L* $^{13}C_6$ $^{15}N_1$]) was custom synthesized by Thermo Fisher Scientific (Ulm, Germany), quantified by amino acid analysis and provided in solution at 1.5 nmol/ml. Antihuman β-NGF (goat polyclonal antibody), used as the capture antibody, was purchased from R&D Systems and biotinylated in-house. The biotinylated stock solution was stored in single-use aliquots at -70°C and diluted to 37.5 µg/ml in 10 mM phosphate-buffered saline (PBS) before use. The ligand affinity purified rabbit polyclonal anti-β-NGF peptide antibody for IA-LC was obtained from Lampire Biological Laboratories (PA, USA) in 1 mg/ml aliquots. Anti-β-NGF peptide antibody columns were prepared as previously described [6,13,16]. Human serum and dried ultrapurity bovine serum albumin (BSA) were purchased from Sigma-Aldrich, Inc. (MO, USA). Streptavidin-coated magnetic beads (Dynabeads® MyOne™ Streptavidin C1) were purchased from Life Technologies (NY, USA). Ammonium acetate, 3-[(3-cholamidopropyl)-dimethylammonio]-1-propane sulfonate (CHAPS), 0.1 M hydrochloric acid (HCl), iodoacetamide, PBS tablets, sodium chloride, 2 M Tris HCl buffer, ammonium formate, ethanol and sodium azide were purchased from Sigma-Aldrich, Inc. Tris(2-carboxyethyl)phosphine HCl (TCEP) and formic acid (98%) were purchased from EMD Chemicals (PA, USA). The MS-grade trypsin was purchased from Promega Corporation (WI, USA). Isopropanol was purchased from JT Baker (PA, USA). HPLC-grade acetonitrile and water were purchased from Honeywell Burdick and Jackson (NJ, USA).

Stock solution preparation, balance, storage & use

Human β-NGF stock solution was prepared by the addition of 0.485 ml of 5% BSA in 10 mM PBS (surrogate matrix) to 97 µg of lyophilized human β-NGF protein (corrected for purity) to a concentration of 200 µg/ml. To demonstrate the preparation accuracy of each stock solution, two separate stock solutions were prepared and an aliquot from each vial was diluted in buffer (n = 6). Internal standard was added to the neat-diluted stock solution aliquots followed by reduction, alkylation and digestion with trypsin, according to the described method, to assess the response from each preparation. Stock solutions were considered balanced if the mean response from each stock (n = 6) was <10% different. Newly prepared stock solution was aliquoted into one-time use protein LoBind® tubes (Eppendorf via Fisher Scientific, NY, USA) at a 25 µl volume and stored at -70°C until use. One frozen-stock aliquot was used upon thawing for preparation of the standard and quality control (QC) samples provided that the stock had demonstrated balance and was within established stability. When a new preparation of stock solution was needed, new stocks were diluted and balanced, and compared with the previously frozen-stock aliquots as a bridge to demonstrate that the new vials provided an equivalent response. This proved particularly important when a new lot of recombinant human β-NGF reference material was used.

Calibration & QC sample preparation

Calibration and QC samples were prepared in a wet ice bath from 7 to 450 pg/ml in surrogate matrix consisting of 5% BSA in 10 mM PBS on the day of analysis. For run acceptance, at least 75% of individual standard curve replicates must be within 25% of theoretical concentration. Human β-NGF is present in control human serum, and human serum void of β-NGF was unavailable. The concentrations of human β-NGF were screened in multiple individual serum lots prior to validation to determine appropriate spiking concen-

trations for QC samples across the analytical range. A human serum pool was prepared in sufficient volume and aliquoted into single use vials that were stored at -70°C. The endogenous (END) concentration of this control serum pool was determined within-run by analysis of replicates (n = 6) over three precision and accuracy runs. A grand mean (n = 18) END concentration was determined and used as the theoretical value of the END QC in each validation run. The QC samples were prepared in the qualified control matrix pool by spiking with human β-NGF at 20, 400 and 650 pg/ml. Spiked QC samples were subsequently diluted 1:1 with surrogate matrix in the extraction plate to the target concentration of 10, 200 and 325 pg/ml.

Extraction procedure

A minimum twofold dilution was applied to study samples by diluting 300 µl of serum 1:1 with surrogate matrix. QC samples were prepared at the target concentration in 1:1 human control matrix to surrogate matrix. Sample aliquots of 300 µl were made into 1-ml 96-well sample plates placed in an ice bath followed by dilution with 300 µl of surrogate matrix. Each sample was buffered with 100 µl of 0.1 M ammonium acetate, 0.5 M sodium chloride, 0.1% w/v CHAPS (CHAPS buffer), prior to the addition of 750 ng of biotinylated capture antibody. The sample plate was incubated overnight at approximately 4°C while gently shaking (400–500 rpm). After approximately 16 h, the human β-NGF/antibody complex was captured by incubation with 20 µl C1 streptavidin magnetic beads for 60 min at room temperature. Plates were loaded to a Hamilton STAR (NV, USA) for automated liquid handling using a 96-channel pipetting head and magnet stands for magnetic bead collection. Using the robot, the beads were collected on the magnetic stand and washed three-times with 300 µl of CHAPS buffer, followed by one wash with 300 µl of 10 mM PBS, pH 7.4. Biotinylated anti-β-NGF complex was captured on the magnetic beads. Human β-NGF was eluted from the beads with 140 µl of 30 mM HCl, transferred to a 0.5-ml 96-well elution plate and then neutralized with 15 µl of 2 M Tris HCl, pH 8.

Post extraction, 9 fmol of the SIL internal standard peptide was added to each sample before reduction, alkylation and trypsin digestion. Reduction and alkylation were performed by incubating with 15 µl of 75 mM TCEP for 45 min at 56°C, and then with 15 µl of 150 mM iodoacetamide for 30 min at room temperature. Finally, 1.6 µg of trypsin was added to each sample for overnight digestion at 37°C. β-NGF was quantified using the tryptic peptide IDTACVCVLSR; the extended SIL peptide generates IDTACVC(V)(L)SR upon trypsin digestion. Finally on day 3, enzyme activity was deactivated by heating sample plates to 100°C for 30 min.

Liquid chromatography

A Dionex UltiMate® 3000 system (CA, USA) containing the following modules was used for this assay: WPS-3000 TPL autosampler, SRD-3600 and SR-3000 solvent racks, LPG-3600 and LPG-3400M pumps and FLM-3100 and FLM-3300 flow managers. Each flow manager contains a 10-port valve in a thermostat-controlled oven. The LPG-3600 contains two ternary pumping systems (designated as 'Loading Pump' and 'Micropump1') and the LPG-3400 is a quaternary pump (designated as 'Nano LC Pump'). The LPG-3400 was connected to the FLM-3300 flow manager equipped with a 1000:1 flow splitter to deliver a flow of 600 nl/min to the nanoLC column held in the oven at 75°C. The Loading Pump and Micropump1 were configured to deliver flow rates ranging from 0.3 ml/min up to 1 ml/min to the antipeptide and trapping columns during the flow program. The antipeptide antibody column was held at 30°C in the FLM-3100 flow manager oven. See Figure 1 for details on how the pumps and columns were connected.

The LC pump and valve configuration were optimized for throughput through the use of multiple pumps to deliver the mobile phases necessary for the antipeptide antibody column and trap column. This LC configuration eliminated all of the gradient delay times normally associated with using a single pump for a column. In this way, when each valve rotates, the desired mobile phase is delivered to each column without delay. This configuration reduced the analysis time from 18 min/sample to 10.2 min/sample when using a single pump per column plumbing configuration [13].

The loading pump delivered 130 µl sample injections to the antipeptide antibody column at 0.3 ml/min for 3 min using 25 mM ammonium formate, pH 7, with 0.01% sodium azide. Simultaneously, Micropump1 delivered 0.1% formic acid in water at 0.3 ml/min to the trap column (0.3 × 5 mm Acclaim® C18; Dionex), and Nano LC Pump delivered an initial mobile phase of 88% water/12% acetonitrile/0.1% formic acid to the nanoLC column (75 µm i.d. × 15 cm Acclaim C18; Dionex). At 3 min, the antipeptide antibody column was eluted to the trap column and washed for 2 min. Simultaneously, the Loading Pump composition was changed to 1% formic acid and 1% ethanol in water to wash the antipeptide antibody column at 1 ml/min. At 5 min after the injection, valve 2 was switched, connecting the Nano LC Pump to the trap column for gradient elution of the β-NGF surrogate peptide to the nanoLC column. The nanoLC gradient begins at 4.7 min from 12% mobile phase B (MPB) to 30%

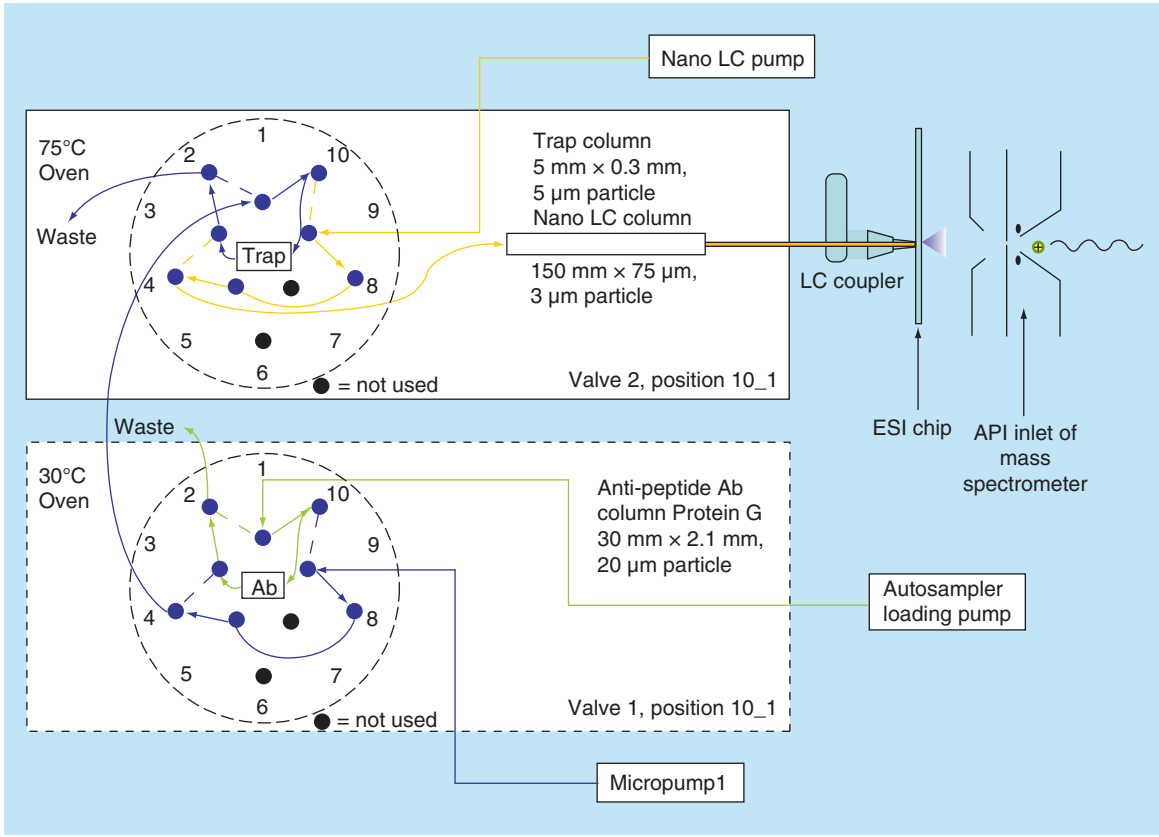


Figure 1. LC flow path and connections including the online peptide immunoaffinity capture module.

MPB at 6.5 min; then the gradient ramps to 90% MPB at 7.4 min. At 8 min, the Loading Pump re-equilibrated the antipeptide antibody column in preparation for the next injection; at 9 min, valve 2 rotates to wash the trap column with 50% acetonitrile/25% isopropanol/0.1% formic acid at 0.3 ml/min. The injection to injection cycle time of the method was 10.2 min.

Mass spectrometry

A Thermo Vantage™ triple quadrupole mass spectrometer (Thermo Scientific) fitted with an Advion TriVersa NanoMate® (TVNM) with ESI Chip™ nanoelectrospray ion source (Advion, Inc., NY, USA) was used. The TVNM was chosen for this work due to its reliable delivery of a stable nanoelectrospray across many days using a single nozzle. The TriVersa NanoMate ChipSoft (Advion, Inc.) software monitors the spray current, which varies with solvent conductivity that is dependent on the composition of effluent. Higher spray currents result with higher water and acid composition of the mobile phase whereas lower spray currents result at higher acetonitrile compositions. If the spray current deviates from user-defined high and low current values (an indication of a problem with the spray), the TVNM was programmed to switch the LC effluent to a new nozzle so that data acquisition continued uninterrupted. The TVNM was operated at a spray voltage of 1.7 kV with spray sensing thresholds of 30 and 300 nA. The β-NGF surrogate peptide was quantified by SRM of the $[M+2H]^{2+}$ molecular ion at *m/z* 647.3 fragmenting to *m/z* 893.4 (y_7^{1+}). Two qualifier ions (y_6^{1+} and y_8^{1+}) were also monitored for the NGF peptide. The SIL-analyte SRM was *m/z* 653.9 to *m/z* 906.4 (y_7^{1+}). A 50 ms dwell time was used. Considering a 5 ms inter-dwell time delay between the four SRM transitions and a peak width at baseline of typically 6 s, approximately 24 SRM data points were recorded per peak.

Method validation

Three precision and accuracy batches were completed at five serum QC concentrations assayed n = 6 for each validation batch. A large pool of control human serum was used for preparation of the END QC and validation QCs at low, mid and high concentrations by spiking 20, 400 and 650 pg/ml of human β-NGF in serum, respectively. The 25- and 100-fold dilutions were validated by spiking 10,000 pg/ml human β-NGF in human serum and subsequently diluting in surrogate matrix. Dilution accuracy near the LLOQ was assessed by twofold dilution of the END QC with surrogate matrix. Spiked validation QCs were pre-

pared fresh for each batch, except for low- and high-concentration QC aliquots stored at -20°C or -70°C for stability assessments.

Results & discussion

Method optimization

This assay presented several method-development challenges. An SIL human β-NGF protein was not available, so the assay variability prior to internal standard addition was minimized by using automated sample preparation with a liquid-handling robotic system. Evaluation of analyte recovery during the β-NGF IA step was performed to ensure that the recovery was maximized. In practice, this included the selection of an amount of anti-NGF capture antibody (750 ng per sample), which was in large excess compared with the amount of endogenous NGF in serum (e.g., baseline NGF is ~20–40 pg/ml). This also involved evaluation of recovery at different incubation times with the capture antibody at several concentrations. Time intervals evaluated were limited to times that enabled a single analyst to perform the sample preparation in one workday (10 h or less) and overnight incubation (16 h). The recovery reached steady-state using an overnight incubation (data not shown). Trypsin digestion was evaluated after approximately 3 h to determine if adequate sensitivity could be achieved without overnight digestion. Adequate sensitivity, as measured by the signal/noise ratio at the target LLOQ was only achieved with overnight digestion.

The flow configuration for this assay utilized separate pumps for loading and elution of the antipeptide antibody column, thus eliminating LC-gradient mixing delays and greatly reducing the injection-to-injection cycle time. This configuration allowed immediate delivery of the elution and wash solvents to the columns upon valve rotation. Furthermore, this approach was used to wash the trap column at 0.3 ml/min, greatly speeding the wash and equilibration of the trap and nanoLC columns. In this nanoLC assay, the primary source of carryover was found to be the trap column. Hundreds of column volumes were delivered to the trap column during a short period of time, minimizing carryover and LC–MS method time.

This assay required the use of an antipeptide antibody column to achieve the target LLOQ. The sensitivity gain using the antipeptide antibody column was determined by injecting the same extracts with and without use of the antipeptide antibody column. The data (not shown) demonstrated an 18-fold improvement in peak area response for the internal standard. It is likely that most of the sensitivity loss with the two-column configuration is due to ion suppression during the 5-min gradient and that the lost sensitivity could be recovered with the use of a longer, shallow gradient. However, the longer cycle time for a two-column configuration would make the method impractical for high-throughput applications. To maximize sample throughput, use of an antipeptide antibody column not only contributes to assay speed but also assay robustness by selective enrichment of the target peptides. Therefore, for methods intended for analysis of hundreds or thousands of samples, investment in the preparation of antipeptide antibodies provides improved assay performance.

Method validation

Following method optimization, the quantitative IA-LC–MS/MS assay for human β-NGF was validated using a 7.0–450 pg/ml calibration range. As this assay incorporates two IA steps akin to traditional ligand binding assays, certain practices commonly used in ligand binding assays analysis of protein biomarkers were adapted to this new assay format, including the use of duplicate standards, a nonlinear curve fit and wider acceptance criteria for accuracy and precision. The data fit through the duplicate standards was quadratic applying a $1/x^2$ weighting factor (Figure 2). Acceptance criteria of 25% precision and 25% accuracy were applied for all QC samples at all concentrations (END β-NGF, END β-NGF diluted in surrogate matrix, β-NGF spikes into serum and dilutions of high β-NGF spikes into serum) and samples for stability evaluation. In the context of this assay, accuracy refers to the relative agreement of a measured β-NGF concen-

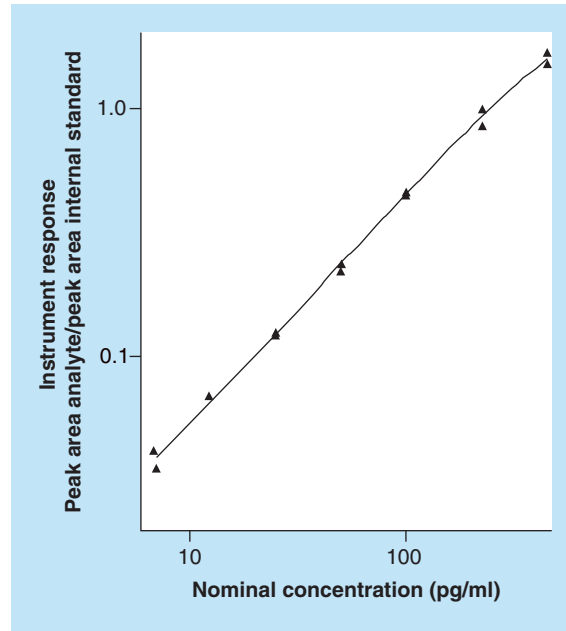


Figure 2. Representative calibration curve from assay validation ranging from 7 to 450 pg/ml β-nerve growth factor.

Table 1. Summary statistics of intra- and inter-assay precision and accuracy of the IA-LC–MS/MS assay for β-NGF.

Human serum	β-NGF target concentration (pg/ml)																					
	2× endogenous Dil 23.7			Endogenous 23.7			VS low 33.7			VS mid 224			VS high 349			100× VS Dil 10,000			25× VS Dil 10,000		
Batch	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	
n	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	
Intra batch																						
Mean (pg/ml)	26.0	25.0	25.5	25.2	23.3	22.6	34.4	33.1	30.6	215	203	190	337	301	307	11,700	11,300	11,100	10,800	9,140	9,440	
CV (%)	8.9	5.6	9.3	3.4	7.1	7.1	9.2	9.5	5.7	3.8	3.7	2.8	4.2	6.0	7.9	4.6	6.0	8.5	3.3	4.4	6.6	
Inaccuracy (%)	9.8	5.6	7.7	6.4	-1.6	-4.6	2.1	-1.7	-9.1	-3.9	-9.3	-15.1	-3.4	-13.7	-12.0	16.8	10.8	13.8	7.8	-5.8	-2.2	
Inter batch																						
Mean (pg/ml)		25.5			23.7			32.7			203			315			11,400			9,800	
CV (%)		7.8			7.4			9.3			6.1			7.8			6.5			6.6	
Inaccuracy (%)		7.7			N/A			-2.9			-9.3			-9.7			13.8			-2.2	

CV: Coefficient of variation; Dil: Dilution; IA: Immunoaffinity; N/A: Not available; NGF: Nerve growth factor; VS: Validation sample.

tration (END and recombinant β-NGF spikes) with the excepted concentration based on calibration with recombinant β-NGF. Intrabatch precision (% CV) on three testing occasions using six replicates at each of the tested NGF concentrations in human serum was no higher than 9.5% (Table 1). Intrabatch accuracies percent relative error (%RE) ranged from -15.1 to 16.8%. Interbatch precision was between 6.1 and 9.3% CV and interbatch accuracy ranged from -9.7% to 13.8% at all concentrations. Dilution linearity was also confirmed with serum samples fortified with 10,000 pg/ml NGF that were diluted 100- or 25-fold with surrogate matrix. Example extracted ion chromatograms for the calibration standards at the LLOQ at 7 pg/ml and ULOQ at 450 pg/ml are shown in Figure 3. Carryover was considered acceptable if the analyte response in the second carryover blank after a ULOQ injection was <25% of the mean response of the LLOQ calibration standards; carryover was generally <20%.

Stability of the surrogate and corresponding SIL peptide in the final digested extract was demonstrated for 64 h at 37°C, 53 h at 4°C and 55 days at -20°C. The stability of spiked human β-NGF in biological matrix was demonstrated for up to four cycles of freezing and thawing at -20°C and -70°C and for 4 h at 4°C. Long-term stability of spiked human β-NGF in biological matrix frozen at -20°C and -70°C was tested up to 706 days. Results confirmed analyte stability under these conditions.

Large-scale method implementation

Large-scale method implementation required approximately 55 steps completed over 45 h, including two overnight procedures, which required careful planning and execution of the sample extraction method. Sample preparation began on the afternoon of day 1, which included fresh preparations of standards in surrogate matrix, QCs in biological matrix and aliquoting of these and study samples into 2-ml protein LoBind plates according to a Watson™ (Thermo Scientific) laboratory information management system (LIMS)-generated worklist. Biotinylated antibody was added and the 96-well plates were incubated at 4°C while shaking overnight. Day 2 activities were mostly automated using a robotic liquid-handling system, including antibody-bead incubation, bead washing, acid elution and transfer of the captured protein into clean 96-well plates. Internal standard was added followed by reduction, alkylation and trypsin digestion overnight at 37°C. Day 3 activities included 30 min of heating at 100°C to deactivate enzyme activity. The majority of day 2 activities were automated and occurred mostly in the morning, which allowed day 1 preparations for a second batch to begin in the afternoon. This overlap-

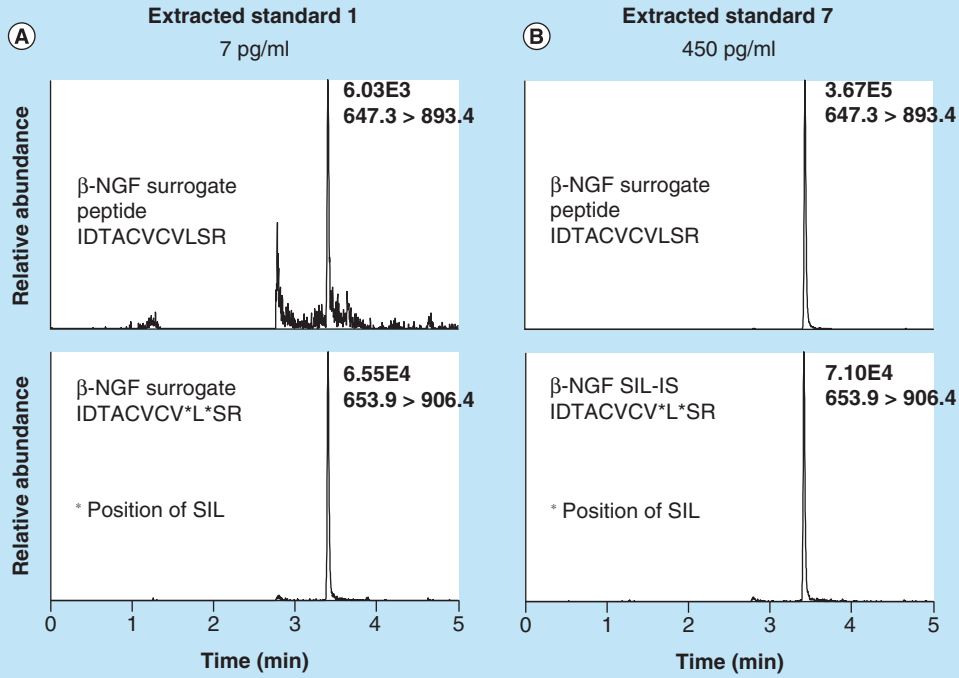


Figure 3. Extracted ion chromatogram of the surrogate peptide IDTACVCVLSR and the corresponding SIL-IS peptide for (A) an extracted LLOQ standard 1 at 7.0 pg/ml and (B) an extracted ULOQ standard 7 at 450 pg/ml.
NGF: Nerve growth factor; SIL-IS: Stable isotope-labeled internal standard.

ping batch workflow enabled the preparation of >500 clinical study samples during a 5-day workweek using a single analyst and robot. The average preparation and analysis time per batch was therefore reduced to approximately 30 h.

A single 96-well plate consumed approximately 17 h of LC–MS time and a two 96-well plate batch consumed approximately 35 h. A schematic of this workflow is shown in Figure 4. During analysis of larger studies, two-plate batches were extracted and analyzed as a routine practice, creating an offset in the time needed to prepare sample batches and the time needed to be analyzed by LC–MS. This led to a backup in the LC–MS operation using a single

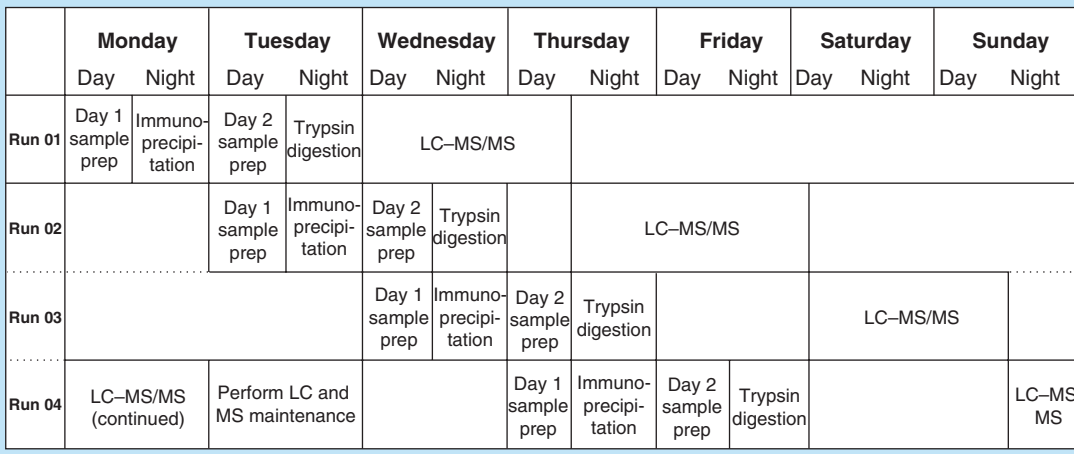


Figure 4. Scheduling of two-plate immunoaffinity extraction and LC–MS/MS batches for a typical week that enabled the analysis of four two-plate batches containing 135 clinical samples per batch per instrument, a total of 540 clinical samples per week.
prep: Preparation.

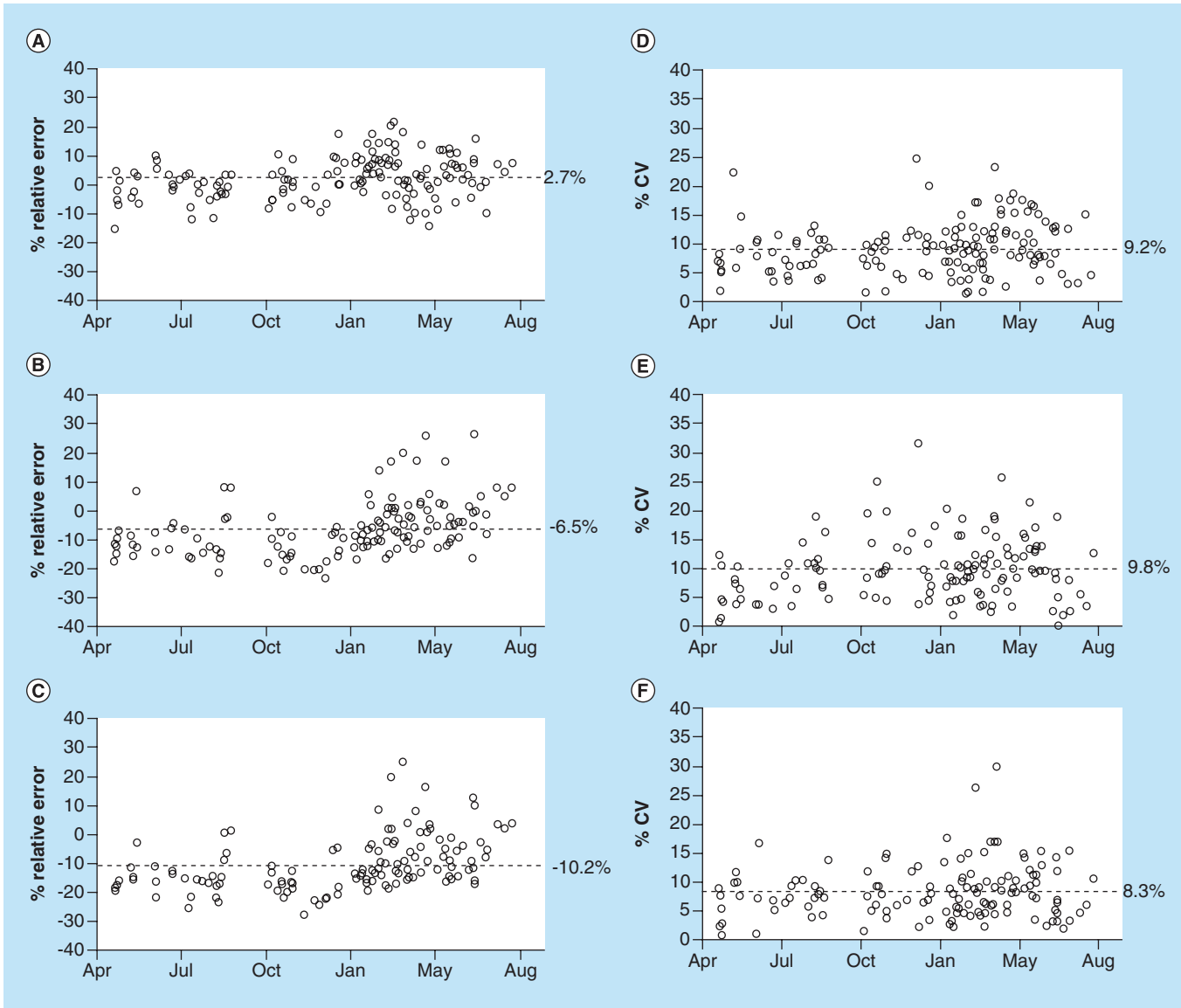


Figure 5. Longitudinal assessment of mean intrabatch percentage relative error for (A) QC1 (B), QC2 and (C) QC3 and percentage CV for (D)QC1, (E) QC2 and (F) QC3 analyzed during study support. Each data point is the average of four quality control (QC) replicates. The dotted lines indicate the grand mean of percentage relative error or percentage CV across all acceptable batches for each QC. QCs: endogenous β-NGF in human serum spiked with 10 pg/ml β-NGF (QC1), spiked with 200 pg/ml β-NGF (QC2) and spiked with 325 pg/ml β-NGF (QC3).
CV: Coefficient of variation.

LC–MS system and required additional extract stability to be established beyond the autosampler extract stability conducted as part of the method validation. Additional validation experiments were performed to evaluate the stability of extracts following digestion stored at -20°C for 55 days before LC–MS analysis. This data met the predefined acceptance criteria for stability, which enabled the decoupling of sample preparation from sample analysis. This allowed for improvements in laboratory efficiency to extract samples when resources were available and to do the same with the LC–MS.

The LC–MS system was set up and configured on the morning of day 2 and the first batch was submitted to the instrument. This batch was typically one that was prepared the previous week and stored at -20°C. Subsequent two-plate batches were analyzed sequentially throughout the week, with the last batch finishing on day 7 of the week. The next week began with performing any LC or MS maintenance and calibration requirements prior to restarting LC–MS analysis. Generally, LC maintenance involved changing the trap column once weekly or sooner, as needed, and replacing and purging new solvents into the system. Mass spectrome-

ter maintenance included replacing the ion transfer tube with a clean one and mass calibration and verification.

The workflow schedule (Figure 4) is described for a single MS supported by wet laboratory extractions requiring a single analyst. As needed, this workflow may also be applied to three mass spectrometers supported by two wet laboratory analysts, or other combinations as needed. In this case, the offset in extraction vs analysis time would be minimized to further improve efficiency and throughput.

Assay performance during sample analysis

To maximize the number of clinical samples analyzed during routine sample testing, the majority of the batch runs consisted of two 96-well plates. A standard curve was placed at the beginning and end of the batch to bracket all study samples with an equal number of QC sample replicates distributed on each plate (n = 4, two per plate). For run acceptance, at least 50% of QC replicates at each low, mid and high concentration level and two-thirds of QC replicates overall were within 25% of theoretical concentration. A two-plate batch enabled analysis of approximately 135 clinical samples during a single week, a single analyst could prepare four batches for analysis on a single LC–MS system for a total of 540 samples. Over the course of 18 months, a total of 18,744 samples (excluding standards and QCs) were analyzed in 173 runs, with an overall run success rate of 95%. The performance of the QCs that were analyzed alongside the study samples was tracked longitudinally (Figure 5). The graphs show mean percent RE and mean percent CV (each from four replicates per batch) for QCs 1–3 for each accepted batch over the duration of the study. The grand mean of percent REs across the entire study was 2.7% for QC1, -6.5% for QC2 and -10.2% for QC3. The grand mean of percent CVs across the entire study was 9.2% for QC1, 9.8% for QC2 and 8.3% for QC3. No significant performance drift was discernible.

The pump-back pressures were monitored across each batch (see Figure 6 as an example) and throughout the week to monitor system performance. Columns and valves were replaced between batches according to the operational pressures known for each column. The antipeptide antibody columns were robust and demonstrated consistent performance across several thousand injections. In fact, throughout sample analysis only four antipeptide antibody columns were used.

The trap column was replaced most frequently at a rate of one or two weekly, depending on the increase

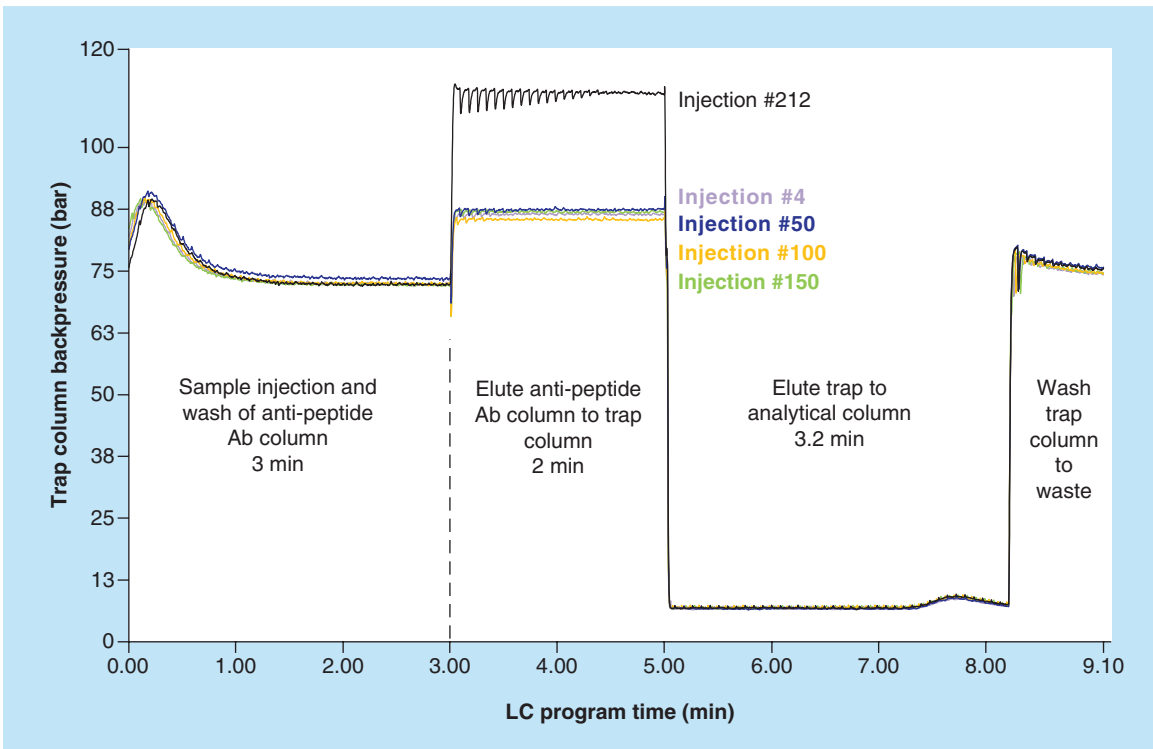


Figure 6. Overlay of the trap column back pressure for five samples from a two-plate batch. The backpressure for sample injection 212 indicated a pressure increase from previous samples in the batch during the eluting of the sample from the peptide antibody column to the trap column. As this was the last sample for this batch, the trap column was replaced prior to initiating the next batch of samples.
Ab: Antibody.

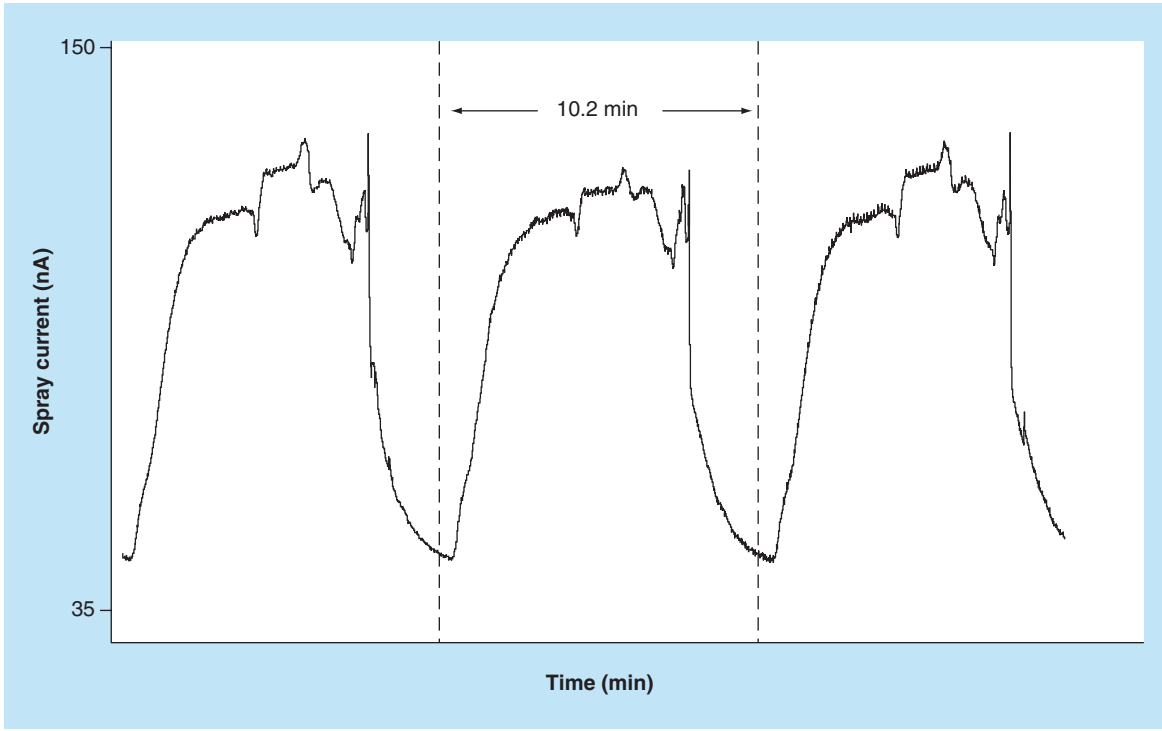


Figure 7. Nanoelectrospray current over three consecutive gradient cycles. Under normal conditions, the spray current repeated in a predictable pattern. When the spray current deviated outside of this pattern, the sample flow was switched to a new nozzle in less than 3 s.

in column back pressure. The pressure increased either slowly across many samples or abruptly when the valve rotated. It was observed that columns failed for two reasons: due to the slow accumulation of sample extract material; and/or, rotor material accumulation from the valve rotor wear. The pumps used herein enabled tracking of the pump hours of operation and the valve rotations. This information was used to preemptively replace parts and columns prior to reaching a point of failure. This approach enabled a relatively high run acceptance of 95% for this complex workflow.

The analytical column lasted for approximately 2 weeks (1700 injections). It was replaced primarily due to poor chromatography (increased peak width and tailing). Occasionally, an abrupt increase in column pressure occurred. In such a case, inspection of the column inlets using a stereomicroscope indicated that small black particles from the valve rotor were accumulating on the column inlet.

The TVNM software enabled the monitoring of the spray current across the batch runs. Under normal operation, there was a predictable high and low spray current observed (Figure 7). It was determined that when the spray current deviated outside of this normal range, the spray would cease. The TVNM software was programmed to switch to a new nozzle within 3 s upon deviating outside of a predefined range, without requiring user intervention. Some nozzles would last for an entire week whereas others would last for only a single sample. There was no predictability to spray failures, so the spray sensing feature was critical to the success of this analysis.

Conclusion

This report highlights the performance and successful implementation of a validated bioanalytical IA-LC–MS/MS assay for human β-NGF biomarker in high-throughput, routine analysis. This IA LC/nanoLC–MS/MS assay has proven to be robust with a low run-failure rate and reproducibly sensitive. To the best of our knowledge, this assay is also the first demonstration of large-scale clinical implementation of a protein assay that includes a nanoliter/min analytical flow rate in the final chromatographic stage and nanoelectrospray MS. The robust assay performance enabled the support of clinical development programs of a therapeutic antibody with nearly 19,000 samples.

Future perspective

Routine implementation of IA-LC–MS/MS workflows in clinical studies has become reality. As the IA-LC–MS/MS technology matures further and more precedent is set for protein biomarker quantification using the methodology of antibody to capture protein biomarkers prior to trypsin digestion combined with an online antipeptide antibody enrichment of a surrogate

peptide, this multiday assay format may be used as a template for other assays. This can be accomplished by substituting commercially available reagent antibodies for IA and custom antipeptide antibodies for online IA capture. However, the basic components of this assay workflow and their scheduling for larger-scale analysis are translatable to other biomarker assays. Modification of buffers and digestion conditions, as well as optimization of extraction recovery for each new protein analyte, may be needed.

The current assay format also presents additional opportunities for method optimization for any potential continued implementation of this assay and other assays of this kind. This includes exploring whether a reduced sample volume can be accommodated to support studies with sample limitations, evaluating additional surrogate matrix buffers that are more generalizable to other biological matrices and other biomarkers, accelerating trypsin digestion to reduce the overall sample preparation time, and assessing if extract stability can be maintained in the absence of heat deactivation of trypsin. It is expected that adding the SIL peptide to the bead elution or neutralization buffer, instead of a separate low-volume addition, may further improve overall assay precision and remove one sample preparation step.

Furthermore, the inclusion of a prequalified END QC within each batch would provide the assay tracking data necessary to ensure consistency throughout reagent lot changes for long-term studies. It is anticipated that improved instrumentation and future investigations by us and others, will advance this assay format further to enable a more widespread implementation of the IA-LC–MS/MS technology for clinical protein biomarker applications.

Authors' contributions

All authors were involved in the data analysis, design and performance of the research, and the writing of the manuscript.

Financial & competing interests disclosure

The laboratory work was performed at Q2 Solutions (formerly Quintiles and Advion) and paid for by Pfizer. At the time of this work, G Schultz and K McCardle were employees of Advion and Quintiles. K McCardle is an employee of Q2 Solutions. H Neubert is an employee of, and holds stock or stock options in Pfizer. The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

Assistance in preparing the manuscript for submission (editing and formatting of tables and text) was provided by J Oleynek, at Engage Scientific Solutions, and was funded by Eli Lilly and Co. and Pfizer, Inc.

Executive summary

Background

- An existing β-nerve growth factor immunoaffinity-LC–MS/MS assay was optimized to allow increased throughput and enhanced analytical performance to support drug-development programs.

Experimental

- The procedure was modified to increase throughput by simultaneous preparation of two 96-well plates and LC duty-cycle reduction.
- This assay was validated and implemented to measure the β-nerve growth factor concentrations in approximately 19,000 clinical samples.

Results & discussion

- This immunoaffinity LC/nanoLC–MS/MS assay is robust, reproducible and sensitive, and has a low run-failure rate.
- This assay is the first demonstration of large-scale clinical implementation of a protein assay that includes a nanoliter/min analytical flow rate in the final chromatographic stage and nanoelectrospray MS.

Conclusion

- This optimized assay format is a template for other clinical protein biomarker assays.

References

Papers of special note have been highlighted as:
• of interest; •• of considerable interest

1 Ackermann BL. Hybrid immunoaffinity–mass spectrometric methods for efficient protein biomarker verification in pharmaceutical development. *Bioanalysis* 1(2), 265–268 (2009).

2 Ackermann BL, Berna MJ. Coupling immunoaffinity techniques with MS for quantitative analysis of low-abundance protein biomarkers. *Expert Rev. Proteomics* 4(2), 175–186 (2007).

• **A pivotal early review of the immunoaffinity applications of MS of protein biomarkers.**

3 Ocana MF, Neubert H. An immunoaffinity liquid chromatography–tandem mass spectrometry assay for the quantitation of matrix metalloproteinase 9 in mouse serum. *Anal. Biochem.* 399(2), 202–210 (2010).

4 Kushnir MM, Rockwood AL, Roberts WL, Abraham D, Hoofnagle AN, Meikle AW. Measurement of thyroglobulin by liquid chromatography–tandem mass spectrometry in serum and plasma in the presence of antithyroglobulin autoantibodies. *Clin. Chem.* 59(6), 982–990 (2013).

5 McAvoy T, Lassman ME, Spellman DS *et al.* Quantification of tau in cerebrospinal fluid by immunoaffinity enrichment and tandem mass spectrometry. *Clin. Chem.* 60(4), 683–689 (2014).

6 Neubert H, Gale J, Muirhead D. Online high-flow peptide immunoaffinity enrichment and nanoflow LC–MS/MS: assay development for total salivary pepsin/pepsinogen. *Clin. Chem.* 56(9), 1413–1423 (2010).

7 Niederkofler EE, Phillips DA, Krastins B *et al.* Targeted selected reaction monitoring mass spectrometric immunoassay for insulin-like growth factor 1. *PLoS ONE* 8(11), e81125 (2013).

8 Wang W, Choi BK, Li W *et al.* Quantification of intact and truncated stromal cell-derived factor-1alpha in circulation by immunoaffinity enrichment and tandem mass spectrometry. *J. Am. Soc. Mass Spectrom.* 25(4), 614–625 (2014).

9 Wang W, Walker ND, Zhu L-J *et al.* Quantification of circulating D-dimer by peptide immunoaffinity enrichment and tandem mass spectrometry. *Anal. Chem.* 84(15), 6891–6898 (2012).

10 Zhang H, Xiao Q, Xin B *et al.* Development of a highly sensitive liquid chromatography/tandem mass spectrometry method to quantify total and free levels of a target protein, interferon-gamma-inducible protein-10, at picomolar levels in human serum. *Rapid Commun. Mass. Spectrom.* 28(13), 1535–1543 (2014).

11 Neubert H, Palandra J, Fernandez Ocana M. Quantification of biotherapeutic targets: new opportunities with immunoaffinity LC–MS/MS. *Bioanalysis* 6(13), 1731–1733 (2014).

12 Palandra J, Finelli A, Zhu M, Masferrer J, Neubert H. Highly specific and sensitive measurements of human and monkey interleukin 21 using sequential protein and tryptic peptide immunoaffinity LC–MS/MS. *Anal. Chem.* 85(11), 5522–5529 (2013).

13 Neubert H, Muirhead D, Kabir M, Grace C, Cleton A, Arends R. Sequential protein and peptide immunoaffinity capture for mass spectrometry-based quantification of total human beta-nerve growth factor. *Anal. Chem.* 85(3), 1719–1726 (2013).

•• **Original β-nerve growth factor immunoaffinity-LC–MS/MS assay highlighting robust assay performance and the utility in translational research and early clinical studies.**

14 Lee AYH, Chappell DL, Bak MJ *et al.* Multiplexed quantification of proglucagon-derived peptides by immunoaffinity enrichment and tandem mass spectrometry after a meal tolerance test. *Clin. Chem.* 62(1), 227–235 (2016).

15 Ackermann BL. Understanding the role of immunoaffinity-based mass spectrometry methods for clinical applications. *Clin. Chem.* 58(12), 1620–1622 (2012).

16 Dufield DR, Radabaugh MR. Online immunoaffinity LC/MS/MS. A general method to increase sensitivity and specificity: How do you do it and what do you need? *Methods* 56(2), 236–245 (2012).

• **A practical guide to configuring and optimally using antipeptide antibody columns.**

Bioanalysis

Overview on biotherapeutic proteins: impact on bioanalysis

This article provides an overview of the information and factors relevant to designing bioanalytical strategies in support of *in vivo* nonclinical and clinical studies of protein therapeutics. The summarized information includes representative types of the therapeutic proteins, their key structural characteristics, the relationship between post-translational modifications and function, issues during purification and formulation, PK of therapeutic proteins and immunogenicity. The effect of each of those on bioanalysis strategy has been pointed out. The impacts of structural variant and 'free'/'bound' forms on PK assessment have been discussed.

Keywords: bioanalysis • biotherapeutic proteins • immunogenicity • PK • protein therapeutics

Yan J Zhang*[,1], Linlin Luo[1] & Dharmesh D Desai[1]

[1]Analytical & Bioanalytical Development, Bioanalytical Sciences-Biologics, Bristol-Myers Squibb, NJ, USA

*Author for correspondence: yan.j.zhang@bms.com

Since the discovery of protective properties of passive immunization in blood transferred from pathogen-infected animals in the late 19th century, protein therapy has become an increasingly important mode of medical treatment. The early successes include passive administration of animal sera for treatment of infectious diseases such as diphtheria, scarlet fever, pneumococcal pneumonia and meningococcal meningitis [1–3] and use of pancreatic extracts in the treatment of diabetes mellitus by the 1920s [4]. Improvement in protein purification has enabled use of purified immunoglobulin (e.g., hyperimmune IgG preparations) and other therapeutic proteins (e.g., insulin, growth factors) derived from animal and subsequently from human sources to provide treatment or protective benefit to patients [5–7]. Recombinant DNA technology revolutionized protein therapy by providing means of mass production of homogenous proteins of interest from living cells and by using well-defined bioprocesses. In addition, the hybridoma technology introduced in 1975 by Kohler and Milstein [8] enabled production of monoclonal antibodies as targeted therapeutics for a wide variety of diseases. Recombinant or hybridoma technology derived therapeutic products are referred to as 'biopharmaceuticals,' 'biotherapeutics,' 'biologicals' or 'biologics.' These terms normally encompass recombinant therapeutic polypeptides and proteins (including antibodies). The therapeutic peptides which are generated through chemical synthesis are important types of therapeutics but are not considered as biotherapeutics, and thus are not the focus of this chapter.

Since human insulin [9], the first recombinant protein therapeutic that received approval from USA FDA in 1982, the numbers of protein therapeutics either in development or available on market after approval have increased dramatically. As of 2014, over 246 biopharmaceutical products have been approved for clinical use in the treatment of a large variety of diseases including cancer, autoimmune disorders, metabolic diseases and infectious diseases [10].

With the expansion of protein engineering capabilities for drug design and an increase in the number of drug targets that are being pursued, there is an increasing demand placed on bioanalysis to support PK and PD studies. In order to apply the best bioanalytical strategies for sample analysis to meet

FUTURE SCIENCE part of fsg

Key terms

PK: Described as 'what the body does to a drug' and defined as the study of the time course of drug absorption, distribution, metabolism and excretion.

PD: Described as 'what a drug does to the body' and defined as the study of the biochemical and physiological effects of drugs and the mechanisms of their actions, including the correlation of their actions and effects with their protein structure.

Biotransformation: The conversion of drugs and other xenobiotics from one form to another within the body, often associated with change in pharmacologic activity.

specific study needs, it is important to understand the characteristics of the therapeutic proteins, which 'forms' of the analyte to measure, and the factors that may impact measurement of the protein therapeutic.

This chapter provides an overview of information relevant to bioanalysis of protein therapeutics including consideration of therapeutic protein structure; stability and purity, and the impact of these components on the exposure–response relationship.

Therapeutic protein types

The advances in protein engineering have resulted in an expanding diversity in the types of protein therapeutics. Recombinant peptides and proteins have been widely used to treat different diseases (Table 1). However, peptides and proteins less than 60 kDa usually have a short circulation half-life. Strategies for increasing the half-life of these relatively small biotherapeutics include the attachment of chemical polymers (e.g., polyethylene glycol [PEG]) or making a fusion protein by recombinantly adding a half-life extension protein (e.g., albumin, transferrin, antibody Fc region). These strategies showed longer time in circulation and less C_{max} to trough difference which can improve efficacy/safety properties. Other advantages of engineering a fusion protein include enabling the protein to target a specific cell surface component enhancing biological activity at the target site (e.g., fusing a cytokine to an antibody) [11], increasing the avidity for the target (e.g., bivalent Fc fusion protein) and conferring bi-specific functions (e.g., target binding and Fc-dependent cytotoxicity).

A large number of biotherapeutics are monoclonal antibodies (mAb) or mAb-derived molecules. They show specific pharmacological effects with long circulatory half-lives and are less likely to undergo significant biotransformation *in vivo*. Advances in antibody engineering and mAb production techniques have not only progressed mAb from murine origin to humanized to fully human, but have also optimized the properties including stability and effector function. In addition, more mAb-based modalities, such as antibody–drug conjugates (ADC), bispecific antibodies and domain antibodies, are being developed to address more complex therapeutic needs. Between 2010 and April 2014, mAb represent almost 27% of all biologics approvals [10]. By mid 2014, a total of 58 monoclonal antibody products have been approved in Europe and/or the USA, and approximately 300 more candidates are currently being developed [12].

Nonantibody-based scaffold proteins are an emerging category of biotherapeutics [13]. The advancement in techniques such as phage, cell and ribosomal display enabled *in vitro* generation of large scale of nonimmunoglobulin molecule libraries. Proteins with high affinity and specificity for a given target can be isolated from these libraries. Additional desired properties of engineered non-Ab scaffold proteins can include lower cost of goods (e.g., ease of manufacturing and consistent product profiles), increased on-target effects (e.g., smaller size for potentially efficient tissue penetration and adjustable half-life for less frequent administration), decreased off-target effects (e.g., Fc-mediated Ab- or complement-dependent cytotoxicity) and intellectual property (e.g., avoid patent issues with mAb targets).

Table 1 summarizes the representative types of the therapeutic proteins and key structural characteristics. A number of GLP-1 receptor (GLP-1R) agonists are listed to show good examples of multiple approaches (i.e., recombinant peptides and fusion proteins) to address the same therapeutic need (treatment of Type 2 diabetes) [14]. Each of the therapeutic protein modalities requires different bioanalytical strategies and methods that are suited for their unique physical/chemical properties.

Post-translational modification

The vast majority of therapeutic proteins are produced from prokaryotic- and eukaryotic-based expression systems, such as recombinant *E. coli*, yeast and mammalian cell lines. A small minority of them are produced from insect, plant cell lines and transgenic animals. Many factors including cost of goods, expertise availability, patenting status and scientific considerations influence the selection of an appropriate expression system for a given therapeutic protein. The post-translational modification (PTM) requirement of the protein is one of the key scientific factors for selecting an appropriate expression system. For example, an *E. coli*-based system which generates aglycosylated proteins is not suitable for producing proteins that require glycosylation for optimal activities. Mammalian cell lines are currently the predominant therapeutic protein production system owing to their ability to perform

Table 1. Selective examples of therapeutic protein types.

Type		MW kDa	Example(s)	Ref.
Peptides and recombinant proteins		0.2–60	Exenatide: a synthetic peptide with 53% homology to human GLP-1; approved in USA (2005) and EU (2007)	[14,15]
			Liragutide: a recombinant acylated peptide with 97% homology to human GLP-1; approved in EU and Japan (2009), as well as in USA (2010)	
Fuson proteins	PEGylated	1–60 (excluding PEG)	LY2428757: a pegylated human GLP-1; in Phase II trials	[14,16,17]
	Fc fusion	>55	Dulagutide: human GLP-1 fused to human IgG4 Fc; in Phase III trials	[14,18]
	Albumin fusion	>67	Albiglutide (Tanzeum: two repeats of human GLP-1 fused to human albumin; approval in USA	[14,19]
mAb based	mAb	~150	Muromonab-CD3 (Orthoclone OKT3), the first monoclonal antibody approved	[20]
	ADC	>150	Brentuximab vedotin (Adcetris®): antihuman CD40 linked with Auristatin; approved in USA (2011) and EU (2012) for treatment of Hodgkin's lymphoma and anaplastic large cell lymphoma	[21]
	Domain Ab	~15	In late 1980, a repertoire of isolated murine VH domains was screened for binding to lysozyme in Cambridge, UK	[22]
	ScFv	~55	Blinatumomab: a fusion protein composed of two single-chain antibodies providing specific binding to CD19 and CD3; received FDA priority review designation in acute lymphoblastic leukemia in October 2014	[23]
	Bi-specific IgG	~150	Catumaxomab: the first bsAb on the market (US, 2009) for treatment of malignant ascites in patients with EpCAM-positive carcinomas; ongoing Phase II for treatment of ovarian and gastric cancer	[24]
Non-Ab scaffolds	Adnectin	~10	CT322: the first Adnectin binding specifically to VEGFR2, was evaluated in Phase II clinical trials	[25]
	DARPin (designed ankyrin repeat proteins)	~15	MP0112: the first therapeutic DARP in candidate, is a VEGF inhibitor and has entered clinical trials for the treatment of wet macular degeneration and diabetic macular edema in early 2010	[26]
	Affibody molecules	~6	first clinical data on imaging using an HER2-binding affibody molecule in 2005	[27]
	Kunitz domain	~7	Ecallantide: a kallikrein inhibitor derived from Kunitz domain to treat hereditary angioedema; approved by the FDA (2009)	[28]

complex PTMs that are often required for optimal secretion, stability and drug activities.

The majority of proteins derived from eukaryotes undergo PTMs. Several hundred PTMs have been reported [29], which occur not only to the nascent ribosome-translated polypeptide but also at anytime during the life of the protein. Nonsequence-derived PTMs can alter protein function by changing either the chemical nature of an amino acid (e.g., deamidation, deimination, oxidation and carboxylation) or protein structures (e.g., formation of disulfide bridges, and proteolytic cleavage). While proper PTMs are necessary for biotherapeutic production and function, undesirable PTMs can cause issues leading to potentially inconsistent drug efficacy and unwanted toxicity. In the context of biotherapeutic proteins, the primary concern for PTMs is during purification, formulation and storage processes as well as *in vivo* after administration. As an example, oxidation of methionine residues in a1-antitrypsin leads to a loss of

Key terms

Protein aggregation: A phenomenon or process (chemical or physical) by which proteins adopt a conformation that causes their polymerization into polymer aggregates.

Immunogenicity: The ability of a substance to elicit immune responses, which can be humoral and/or cell-mediated immune responses.

Neonatal Fc receptor (FcRn): A cell surface receptor that binds to the Fc portion of IgG and also binds to albumin. Endocytic recycling of both IgG and albumin by FcRn result in increased half life of both of these proteins.

Antidrug antibody: Either pre-existing or treatment-induced antibody specific for drug; may also refer to the positive control (PC) monoclonal or polyclonal ADA that is used as a quality control during immunogenicity method development, validation and production.

its critical anti-elastase activity required for the treatment of emphysema [30]. Deamidation of asparagine residues to form aspartic acid and iso-aspartic acid can lead to degradation and alteration of protein activity, particularly during long-term storage [31].

Understanding the relationship between PTM structure and function has enabled the rational engineering of PTMs to enhance therapeutic properties. Implementation of suitable analytical measurements is crucial to monitoring undesirable PTMs. Understanding the relationship between PTMs and the protein function is important to employ appropriate bioanalytical strategies to measure functionally relevant forms of the drug for appropriate exposure and response assessments.

Structural variants & impact on bioanalysis

After administration into animals or human subjects, therapeutic proteins may undergo additional PTMs, also referred as biotransformation, resulting in further structural modifications and different structural variants co-existing *in vivo*. Proteolytically cleaved fragments from both endopeptidase and exopeptidase activities have been observed for peptides conjugated to Fc, such as in the cases of romiplostim and Fc–FGF21 conjugates [32–35]. An increase in aspartate isomerization *in vivo* at the CDR region of a model mAb was reported with significant impact on the function of the molecule [31].

From bioanalytical perspective, biotransformation may confound quantification of active or other desired forms. The assay reagents for a ligand-binding assay may not be able to distinguish between different variants. As in the case of FGF21-derived protein therapeutics, a ligand-binding assay that measures the level of total protein including both proteolytic fragments and the intact protein will overestimate the active drug exposure. On the other hand, an LC-MS assay following an unmodified surrogate peptide may miss the quantification of the modified form, resulting in underestimated exposure. Therefore, it is crucial to understand which structural variants to measure and what bioanalytical technology to measure them. Due to the greater structural complexity of proteins compared with small molecule drugs, it is more challenging to quantitatively determine the structure–activity relationships (SARs) for therapeutic protein candidates. Risk-based assessment based on amino acid sequences and *in vitro* data may be used to determine whether structural variants have safety and/or efficacy liabilities, and thus would require specific monitoring of that entity.

Purification & formulation

Unlike small molecule drugs which are chemically synthesized, therapeutic proteins are produced using different biologic expression/production systems involving a genetically modified host cell. Therefore, recovery and purification of these proteins, often by multiple chromatographic techniques, is an integral part of the manufacturing process. The goals of the purification are to ensure the purity, activity and safety of the finished product, while optimizing the manufacturability and exchanging the buffer into the final formulation. The processing has to be economic, effective and reliable for separation of impurities including product related variants (e.g., modified forms and aggregates), process-related materials (e.g., cell culture media, leachates from the separation media), host cell components and contaminant species that should not be included in the drug product (e.g., nucleic acids, viruses, pyrogens, residual host cell proteins). After purification, the purified bulk drug substance is subjected to a series of formulation, fill and finish steps and then converted into the final drug product (at the desired concentration and in the desired formulation and container or delivery device). These processing steps have the potential to further impact therapeutic protein stability [36].

One of the prominent issues during the purification and storage is product aggregation. Such aggregation can range considerably in size from dimers through subvisible to visible particles, and can be covalent or noncovalent, native or non-native, soluble or insoluble and reversible or irreversible. Recombinant protein aggregation can happen in high-producing cell lines where the protein-folding facilitating chaperones become overloaded with the recombinant protein [31]. The mechanism of protein aggregation during purification and storage still remains to be fully understood.

Studies suggest that aggregates can be heterogeneous and the mechanism is dependent on the nature of the stress applied [37,38]. External stress factors can lead to loss of native structure of biotherapeutic proteins. The examples of those factors include air–liquid and liquid–solid interfaces, temperature changes, pH values, high ionic solution conditions and mechanical shear. In addition, proteins tend to be attracted to hydrophobic surfaces where they can more likely unfold. Structurally altered proteins and partially folded intermediates which expose hydrophobic surface and then allow intermolecular interactions are thought to play a role in forming aggregates [39,40]. The native proteins can also form aggregates through absorption to microparticles or at high concentrations. Aggregation not only can reduce yield thus impacting cost, but also has been correlated to increased incidences in immunogenicity and immune-mediated adverse events in patients.

Appropriate selection of excipients, physical state and storage conditions is essential for formulation of biotherapeutics. In addition, the intended route of administration and the 'marketability' of the final drug product are important factors to be considered. Aggregation is also a concern during formulation. There have been developments in the use of excipients to reduce aggregation in formulated products [41]. Those excipients include surfactants (e.g., polysorbate 20 and 80) and amino acids (e.g., arginine, lysine and glutamic acid) to prevent protein adsorption at the air–liquid interface, and carbohydrates (e.g., sucrose, dextrose) to cover the surface of hydrophobic bonding sites and thus to preserve the native structures.

PK & immunogenicity

PK of therapeutic proteins

PK of therapeutic drugs is usually described in terms of ADME – the drug's ability to be Absorbed, Distributed, Metabolized and Excreted by the body. Small molecule therapeutics are typically administered orally. Degradation by liver enzymes and renal clearance are the primary components of metabolism and excretion for small molecules. By contrast, the parenteral route of administration (e.g., intravenous, intramuscular or subcutaneous [sc.]) is the standard route of administration (ROA) for therapeutic protein drugs due to poor oral delivery. With the development of more convenient delivery devices, outpatient sc. auto-injection systems have increasingly become a preferred route of administration.

After an sc. injection, therapeutic proteins pass through the interstitium and then enter the blood or lymphatic systems. Proteins >20 kDa are generally believed to be favorably absorbed into the lymphatic system before being transported into the vascular circulation and the tissue space [42]. Systemic bioavailability of peptide and protein drugs following sc. is quite variable, ranging from 20 to 100% [43]. For large protein therapeutics such as mAbs (~150 kDa), distribution is generally restricted to the blood stream and extracellular space with low tissue/blood ratios. If the target is expressed on the surface of cells, in order to exert their pharmacological activity, therapeutic proteins have to extravasate across the endothelial walls of blood capillaries, pass through the interstitium and reach the target on the cell surface. There are many factors that can affect extravasation of protein drugs, including biological factors (i.e., capillary structure differences, the disease state of tissue and the flow rate of blood),and physicochemical properties of therapeutic proteins (i.e., molecular size, shape, charge and polarity). The abundance of binding factors (i.e., targets and neonatal Fc receptor (FcRn) for some mAbs) in tissues and the binding kinetics can also influence tissue distribution of therapeutic proteins. Therapeutic proteins are cleared from the circulation by a number of different mechanisms including extracellular proteolysis, endocytosis (nonspecific or receptor-mediated) followed by intracellular degradation, antidrug antibody (ADA) mediated clearance and renal excretion. The biophysical properties of the therapeutic protein will dictate the predominant mechanism of clearance. Binding to the cell membrane associated target results in the internalization of therapeutic proteins into lysosome compartment and thus degradation and clearance. This type of clearance is referred to as target-mediated drug disposition (TMDD). Therapeutic proteins that induce a specific antibody response (also see 'Immunogenicity of therapeutic proteins' section below) may result in the formation of immune complexes which are cleared by either complement or Fc receptors.

Other factors can also affect the PK of therapeutic proteins. IgG mAbs with isotypes of 1, 2 and 4 can bind to the FcRn in the acidic (<pH 6.0) environment of the endosomes after pinocytosis. The bound proteins are transported back to cell surface and released into the extracellular space at near neutral pH. This mechanism of FcRn-mediated recycling is used both by IgG antibodies and serum albumin to extend their half-life in circulation. Since the number of most membrane targets is limited, target binding can be saturated at high doses of therapeutic protein; whereas FcRn is of high abundance and the binding is not saturable at most dosages. Consequently, some therapeutic proteins with cell membrane targets have nonlinear PK at low doses and linear PK at high doses.

For PK evaluation, the meaningful exposure data include the assessment of all active components of the

biological product. Measuring every variant (active or inactive) can be done but does not provide value-added data. Because the active product is not a single chemical but a mixture of closely related complex biological substances, it is important to select a bioanalytical method which measures the most pharmacological relevant form (see sections on 'Structural variants and impact on bioanalysis' and 'Free and bound forms and impact on PK assessment').

Free & bound forms & impact on PK assessment

The majority of therapeutic proteins are designed to interact with specific circulating or cell membrane bound target proteins via noncovalent, high affinity, reversible binding that elicit a pharmacological effect. Once administered to animals or to humans, therapeutic proteins can bind to various components such as circulating target proteins, antidrug antibodies and other endogenous entities in circulation. As a result, the therapeutic proteins can co-exist in free, partially bound and bound forms [44]. In most cases, the forms which have at least one target site free of binding partners (often referred as free or partially bound forms) are considered bioactive forms due to their target-binding potential, whereas the 'fully bound' forms which have no target binding site free are considered inactive forms. The levels of these circulating binding partners may change with time in relation to the drug PD, immunogenicity or disease state.

For the assessment of therapeutic proteins, whether to measure free or bound levels of therapeutics continues to be an ongoing topic of discussion in the bioanalytical community. Measurement of free drug concentration was considered to provide a better estimate of efficacious concentrations and safety margin [45]. In the draft US FDA guidance on 'Clinical Pharmacology Data to Support a Demonstration of Biosimilarity to a Reference Product', it states that analytical assays should be able to detect the active and/or free product instead of the total product, particularly if binding to a soluble ligand is a necessary step for activity and clinical effect [46]. The AAPS Ligand-Binding Assay Bioanalytical Focus Group published a White Paper on challenges and issues of measuring free and total drugs and targets, and how data should be used to support drug discovery and development [44]. The design of a bioanalytical method should consider which variant/form of the biotherapeutic protein should be measured to achieve the intended clinical or nonclinical data. This design is frequently referred to as 'fit for purpose' [47].

Box 1. Example of product-specific immunogenicity risk factors.

Product-specific immunogenicity risk factors
- Amino acid sequence
- Post-translational modifications
- Protein conformation
- Aggregates
- Expression system
- Formulation
- Product impurities
- Mechanism of action
- Route of administration
- Dose, frequency and duration

Immunogenicity of therapeutic protein

All biologic therapeutics have the potential to elicit an immune response against the drug, which is typically assessed by detecting for the presence of antidrug antibody. The risk of immunogenicity on patient safety is the primary concern. For example, it was reported that neutralizing anti-erythropoietin antibodies could develop in patients with anemia of chronic renal failure during treatment with epoetin, a recombinant human erythropoietin, leading to potential pure red-cell aplasia [48]. A death of a multiple sclerosis (MS) patient who had developed neutralizing antibodies (Nab) during treatment with natalizumab was noted in another case report. It was speculated that the formation of ADA rendered the patient functionally untreated leading to a severe rebound inflammation [49]. Even though the cases with severe consequences are rare, regulatory agencies have indicated that an immunogenicity risk assessment should be conducted during all stages of therapeutic product development [50]. Furthermore, a strategy to mitigate immunogenicity-related adverse events should also be developed as the therapeutic product advances through clinical trials. Immunogenicity risk can be subdivided into patient-specific and product-specific factors. Patient-specific immunogenicity risk factors include dosing amount/frequency, route of administration and patient immune status. An example of product-specific factors that may be considered for immunogenicity risk assessment is presented in Box 1.

In addition to the safety concerns, ADA may alter the PK and subsequently PD of a protein. Antibody formation can cause either increased or decreased clearance of the therapeutic protein. Furthermore, ADA may interfere with bioanalytical assay measurements.

Immunogenicity has been typically monitored by determining the presence of binding ADAs using immunoassays and the presence of Nabs using a biological assay. Closely monitoring the formation of ADA has become an integral part of biotherapeutic development throughout a therapeutic's life cycle, from preapproval to postmarket commitment.

Conclusion & future perspective

With an increase in the number of protein therapeutics that are being pursued in biopharma industry, there is an increasing demand placed on bioanalysis to support PK/PD studies. In order to apply the best bioanalytical strategies to meet specific study needs, it is important to understand the characteristics of the therapeutic proteins, which 'forms' of the analyte to measure and what factors may impact measurement of the protein therapeutics.

In comparison to small molecule drugs, therapeutic proteins are not only larger in molecule size but also with higher complexity in structures. They are produced by cells and exist as a heterogeneous mixture with variants from different post-translational modifications and aggregation, and are susceptible to changes. After being dosed into an animal or a human body, therapeutic proteins can undergo further structural changes, or biotransformation. In addition, due to biological interactions, a therapeutic molecule can exist as bound, partially bound, free, or high order complex forms. Consequently, therapeutic proteins are present as a mixture of closely related complex biological substances with both active and inactive variants and forms inside a body. Further, all biologic therapeutics have the potential to elicit an immune response against the drug. The drug bound with antidrug antibodies may or may not still retain the desirable activities. It is important to understand during assay development what forms of the drug are most relevant to pharmacological assessment and which technology to use to measure the intended forms or variants.

Financial & competing interests disclosure

The authors received salaries from BMS, but BMS has no interest or financial conflict with the subject matter or materials discussed in the manuscript. The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.

Executive summary

- Recombinant or hybridoma technology derived therapeutic products are usually referred to as 'biopharmaceuticals,' 'biotherapeutics,' 'biologicals' or 'biologics.'
- The advances in protein engineering have resulted in an expanding diversity in the types of protein therapeutics. Each of the therapeutic protein modalities requires different bioanalytical strategies and methods that are suited for their unique physical/chemical properties.
- Understanding the relationship between post-translational modifications and the protein function is important to employ appropriate bioanalytical strategies to measure functionally relevant forms of the drug for appropriate exposure and response assessments.
- Biotransformation may confound quantification of active or other desired forms. It is crucial to understand which structural variants to measure and what bioanalytical technology to measure them.
- Aggregation not only can reduce yield thus impacting cost, but also has been correlated with increased incidences in immunogenicity and immune-mediated adverse events in patients.
- For PK evaluation, the meaningful exposure data include the assessment of all active components of the biological product.
- The therapeutic proteins can co-exist in free, partially bound and bound forms. The design of a bioanalytical method should consider which form of the biotherapeutic protein to be measured to fit for the intended purpose.
- Immunogenicity may alter the PK and subsequently PD of a therapeutic protein. Furthermore, antidrug antibody may interfere with bioanalytical assay measurements.

References

Papers of special note have been highlighted as:
• of interest; •• of considerable interest

1 Casadevall A. Passive antibody therapies: progress and continuing challenges. *Clin. Immunol.* 93(1), 5–15 (1999).

2 Keller MA, Stiehm ER. Passive immunity in prevention and treatment of infectious diseases. *Clin. Microbiol. Rev.* 13(4), 602–614 (2000).

3 Reynolds T, De Zafra C, Kim A, Gelzleichter TR. Chapter 1 - Overview of biopharmaceuticals and comparison with small-molecule drug development. In: *Nonclinical Development of Novel Biologics, Biosimilars, Vaccines and Specialty Biologics*. Plitnick LM, Herzyk DJ (Eds). Academic Press, San Diego, CA, USA, 3–33 (2013).

4 Banting FG, Best CH, Collip JB, Campbell WR, Fletcher AA. Pancreatic extracts in the treatment of diabetes mellitus. *Can. Med. Assoc. J.* 12(3), 141–146 (1922).

5 Berger M. A history of immune globulin therapy, from the Harvard crash program to monoclonal antibodies. *Curr. Allergy Asthma Rep.* 2(5), 368–378 (2002).

6 Snydman DR. Historical overview of the use of cytomegalovirus hyperimmune globulin in organ transplantation. *Transpl. Infect. Dis.* 3(Suppl. 2), 6–13 (2001).

7 Frasier SD. The not-so-good old days: working with pituitary growth hormone in North America, 1956 to 1985. *J. Pediatr.* 131(Suppl. 1 II), S1–S4 (1997).

8 Kohler G, Milstein C. Continuous cultures of fused cells secreting antibody of predefined specificity. *Nature* 256(5517), 495–497 (1975).

9 Crea R, Kraszewski A, Hirose T, Itakura K. Chemical synthesis of genes for human insulin. *Proc. Natl Acad. Sci. USA* 75(12), 5765–5769 (1978).

• **This made mass production of insulin possible.**

10 Walsh G. Biopharmaceutical Benchmarks 2014. *Nat. Biotechnol.* 32(10), 992–1000 (2014).

• **Indicates an unprecedented approval rate for protein therapeutics.**

11 Chang CH, Gupta P, Goldenberg DM. Advances and challenges in developing cytokine fusion proteins as improved therapeutics. *Expert Opin. Drug Discov.* 4(2), 181–194 (2009).

12 Ecker DM, Jones SD, Levine HL. The therapeutic monoclonal antibody market. *mAbs* 7(1), 9–14 (2015).

• **Shows the increasing numbers of mAbs on market and under development.**

13 Löfblom J, Frejd FY, Ståhl S. Non-immunoglobulin based protein scaffolds. *Curr. Opin. Biotechnol.* 22(6), 843–848 (2011).

14 Sivertsen J, Rosenmeier J, Holst JJ, Vilsboll T. The effect of glucagon-like peptide 1 on cardiovascular risk. *Nat. Rev. Cardiol.* 9(4), 209–222 (2012).

15 Kaspar AA, Reichert JM. Future directions for peptide therapeutics development. *Drug Discov. Today* 18(17–18), 807–817 (2013).

16 Jung SK, Deluca PP, Kang CL. Emerging PEGylated drugs. *Expert Opin. Emerg. Drugs* 14(2), 363–380 (2009).

17 Hamidi M, Rafiei P, Azadi A. Designing PEGylated therapeutic molecules: advantages in ADMET properties. *Expert Opin. Drug Discov.* 3(11), 1293–1307 (2008).

18 Shimamoto G, Gegg C, Boone T, Quéva C. Peptibodies: a flexible alternative format to antibodies. *mAbs* 4(5), 586–591 (2012).

19 Sleep D, Cameron J, Evans LR. Albumin as a versatile platform for drug half-life extension. *Biochim. Biophys. Acta* 1830(12), 5526–5534 (2013).

20 Buss NaPS, Henderson SJ, Mcfarlane M, Shenton JM, De Haan L. Monoclonal antibody therapeutics: history and future. *Curr. Opin. Pharmacol.* 12(5), 615–622 (2012).

21 Sassoon I, Blanc V. Antibody-drug conjugate (ADC) clinical pipeline: a review. *Methods Mol. Biol.* 1045, 1–27 (2013).

22 Holt LJ, Herring C, Jespers LS, Woolven BP, Tomlinson IM. Domain antibodies: proteins for therapy. *Trends Biotechnol.* 21(11), 484–490 (2003).

23 Nagorsen D, Kufer P, Baeuerle PA, Bargou R. Blinatumomab: a historical perspective. *Pharmacol. Ther.* 136(3), 334–342 (2012).

24 Byrne H, Conroy PJ, Whisstock JC, O'kennedy RJ. A tale of two specificities: bispecific antibodies for therapeutic and diagnostic applications. *Trends Biotechnol.* 31(11), 621–632 (2013).

25 Lipovsek D. Adnectins: engineered target-binding protein therapeutics. *Protein Eng. Des. Sel.* 24(1–2), 3–9 (2011).

26 Stumpp MT, Amstutz P. DARPins: a true alternative to antibodies. *Curr. Opin. Drug Discov. Dev.* 10(2), 153–159 (2007).

27 Löfblom J, Feldwisch J, Tolmachev V, Carlsson J, Ståhl S, Frejd FY. Affibody molecules: engineered proteins for therapeutic, diagnostic and biotechnological applications. *FEBS Lett.* 584(12), 2670–2680 (2010).

28 Lehmann A. Ecallantide (DX-88), a plasma kallikrein inhibitor for the treatment of hereditary angioedema and the prevention of blood loss in on-pump cardiothoracic surgery. *Expert Opin. Biol. Ther.* 8(8), 1187–1199 (2008).

29 Khoury GA, Baliban RC, Floudas CA. Proteome-wide post-translational modification statistics: frequency analysis and curation of the swiss-prot database. *Sci. Rep.* 1, pii: srep00090 (2011).

30 Taggart C, Cervantes-Laurean D, Kim G *et al.* Oxidation of either methionine 351 or methionine 358 in alpha 1-antitrypsin causes loss of anti-neutrophil elastase activity. *J. Biol. Chem.* 275(35), 27258–27265 (2000).

31 Prueksaritanont T, Tang C. ADME of biologics-what have we learned from small molecules? *AAPS J.* 14(3), 410–419 (2012).

32 Hall MP, Gegg C, Walker K *et al.* Ligand-binding mass spectrometry to study biotransformation of fusion protein drugs and guide immunoassay development: strategic approach and application to peptibodies targeting the thrombopoietin receptor. *AAPS J.* 12(4), 576–585 (2010).

33 Hecht R, Li YS, Sun J *et al.* Rationale-based engineering of a potent long-acting FGF21 analog for the treatment of type 2 diabetes. *PLoS ONE* 7(11), e49345 (2012).

34 Yie J, Wang W, Deng L *et al.* Understanding the physical interactions in the FGF21/FGFR/beta-Klotho complex: structural requirements and implications in FGF21 signaling. *Chem. Biol. Drug Des.* 79(4), 398–410 (2012).

35 Hager T, Spahr C, Xu J, Salimi-Moosavi H, Hall M. Differential enzyme-linked immunosorbent assay and ligand-binding mass spectrometry for analysis of biotransformation of protein therapeutics: application to various FGF21 modalities. *Anal. Chem.* 85(5), 2731–2738 (2013).

36 Rathore N, Rajan RS. Current perspectives on stability of protein drug products during formulation, fill and finish operations. *Biotechnol. Prog.* 24(3), 504–514 (2008).

37 Joubert MK, Luo Q, Nashed-Samuel Y, Wypych J, Narhi LO. Classification and characterization of therapeutic antibody aggregates. *J. Biol. Chem.* 286(28), 25118–25133 (2011).

• **Suggests how aggregates are formed for therapeutic antibodies.**

38 Krishnan S, Raibekas AA. Multistep aggregation pathway of human interleukin-1 receptor antagonist: kinetic, structural, and morphological characterization. *Biophys. J.* 96(1), 199–208 (2009).

39 Acosta-Sampson L, King J. Partially folded aggregation intermediates of human gammaD-, gammaC-, and gammaS-

crystallin are recognized and bound by human alphaB-crystallin chaperone. *J. Mol. Biol.* 401(1), 134–152 (2010).

40 Kim D, Yu MH. Folding pathway of human alpha 1-antitrypsin: characterization of an intermediate that is active but prone to aggregation. *Biochem. Biophys. Res. Commun.* 226(2), 378–384 (1996).

41 Ratanji KD, Derrick JP, Dearman RJ, Kimber I. Immunogenicity of therapeutic proteins: influence of aggregation. *J. Immunotoxicol.* 11(2), 99–109 (2014).

42 Lee JW. ADME of monoclonal antibody biotherapeutics: knowledge gaps and emerging tools. *Bioanalysis* 5(16), 2003–2014 (2013).

43 Lin JH. Pharmacokinetics of biotech drugs: peptides, proteins and monoclonal antibodies. *Curr. Drug Metabol.* 10(7), 661–691 (2009).

44 Lee JW, Kelley M, King LE *et al.* Bioanalytical approaches to quantify "total" and "free" therapeutic antibodies and their targets: technical challenges and PK/PD applications over the course of drug development. *AAPS J.* 13(1), 99–110 (2011).

• **Explains what to measure, total versus free, in large molecule PK/PD bioanalysis.**

45 Kuang B, King L, Wang HF. Therapeutic monoclonal antibody concentration monitoring: free or total? *Bioanalysis* 2(6), 1125–1140 (2010).

46 FDA. Clinical pharmacology data to support a demonstration of biosimilarity to a reference product (2014). www.fda.gov

47 Desilva B, Smith W, Weiner R *et al.* Recommendations for the bioanalytical method validation of ligand-binding assays to support pharmacokinetic assessments of macromolecules. *Pharm. Res.* 20(11), 1885–1900 (2003).

•• **Sets the recommendation for PK validation of large molecules.**

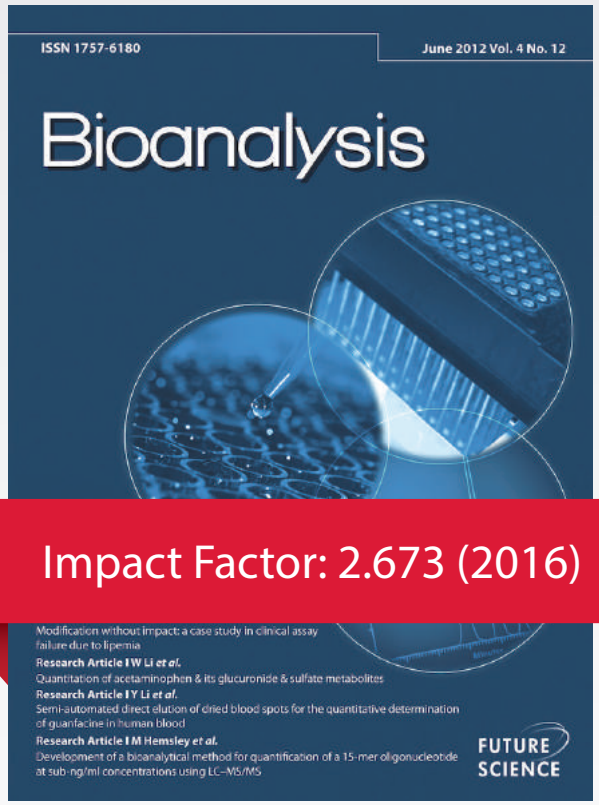
48 Casadevall N, Nataf J, Viron B *et al.* Pure red-cell aplasia and antierythropoietin antibodies in patients treated with recombinant erythropoietin. *N. Engl. J. Med.* 346(7), 469–475 (2002).

49 Svenningsson A, Dring AM, Fogdell-Hahn A *et al.* Fatal neuroinflammation in a case of multiple sclerosis with anti-natalizumab antibodies. *Neurology* 80(10), 965–967 (2013).

50 US Dep. of Health and Human Services, FDA, CDER, CBER. Guidance for industry immunogenicity assessment for therapeutic protein products. (2014). www.fda.gov

•• **The most updated US FDA guidance for immunogenicity testing of protein therapeutics.**

CALL FOR PAPERS



Impact Factor: 2.673 (2016)

Bioanalysis provides the busy bioanalyst with a forum for the rapid publication of original research and critical reviews of all the latest relevant and significant developments. Focusing on techniques and methods used for the detection or quantitative study of analytes in human or animal biological samples, Bioanalysis encourages the submission of articles describing forward-looking applications.

Issues per Year: 24
Average Monthly Full Text Views: 13,554

- Medline indexed
- Open Access and Accelerated Publication options available

Submit your paper

s.nadarajah@future-science.com

newlands press

Research Article

Bioanalysis

Validation of an integrated series of ligand-binding assays for the quantitative determination of antibody–drug conjugates in biological matrices

Background: The bioanalytical strategy for antibody–drug conjugates (ADC) includes multiple integrated measurements of pharmacologically relevant ADC. **Methods & results:** Three ligand-binding assays were validated for the measurement of total antibody, active ADC and total ADC. Accuracy and precision demonstrate%bias from -8 to 14%, %CV from 3 to 11% and total error from 3 to 21%, with >98% samples meeting incurred sample reanalysis criteria. Each assay met stability, selectivity, dilutional integrity, carry over and specificity criteria with no interference from associated metabolite/impurity. Given the active ADC assay sensitivity to payload, active ADC was used to assess drug to antibody ratio. **Discussion & conclusion:** Implementation of a microfluidic automated platform enabled high throughput sample analysis of multiple analytes with minimal sample processing.

First draft submitted: 16 November 2015; Accepted for publication: 19 January 2015; Published online: 26 February 2016

Keywords: antibody–drug conjugate • drug antibody ratio • ligand-binding assays • selectivity • specificity • total antibody • payload

Heather Myler*[,1], Vangipuram S Rangan[2], Alex Kozhich[1], Brian Hoffpauir[3], Donna Dail[1], Jennifer Cummings[1], Mark Saewert[1], Amy Manney[1], Ang Liu[4], Chetana Rao[2], Jian Wang[4], Renuka Pillutla[1] & Binodh DeSilva[1]

[1]Bioanalytical Sciences-Biologics, Bristol-Myers Squibb, Lawrenceville, NJ, USA
[2]Biologics Discovery California, Bristol-Myers Squibb, Redwood City, CA, USA
[3]Immunochemistry, PPD Laboratories®, Richmond, VA, USA
[4]Bioanalytical Sciences-LCMS, Bristol-Myers Squibb, Lawrenceville, NJ 08648, USA
*Author for correspondence:
Tel.: +1 609 252 3600
heather.myler@bms.com

Antibody–drug conjugates (ADC) are a re-emerging class of cancer therapeutics that have shown promise in treating cancer patients with the goal of maximizing therapeutic exposure at the tumor site while minimizing systemic exposure and the severe side effects of conventional chemotherapeutics. The growing database for ADC-based targeted therapies, including monoclonal antibodies conjugated to DNA alkylators, DNA strand breakers, tubulin inhibitors and tubulin stabilizers [1–3], continues to broaden our knowledge base and provides advantages to the development of additional targeted therapy options including peptide drug conjugates, nanocarriers and viral drug delivery systems.

Randomly conjugated ADC, resulting in highly heterogeneous mixtures of drug product with varying numbers of payload conjugated to each antibody molecule at variable conjugation sites, have provided significant challenges for the analytical community. One of the challenges to the bioanalytical community is whether to develop assays that are sensitive or insensitive to the drug antibody ratio (DAR). There are a wide variety of opinions currently in debate, but there is yet to be a consensus on which is more correlative with safety and efficacy parameters. Furthermore, this may be different for different ADC. There are many valuable case studies in the literature that we can use to guide our evolving bioanalytical strategies for quantifying ADC [4–8], however, the complex nature of ADC along with the potential for activity modulating biotransformation events currently necessitates a multianalyte bioanalytical strategy comprised of measuring the intact ADC (Ab conjugated to payload), total Ab and unconjugated payload.

This manuscript describes the validation of the three ligand binding bioanalytical assays utilizing highly specific reagents piv-

FUTURE SCIENCE part of fsg

otal to our current bioanalytical strategy to measure total antibody, active ADC and total ADC. The total antibody (Ab) assay measures total circulating therapeutic antibody irrespective of the presence or absence of conjugated payload or payload metabolites.

[Total Ab] = [Active ADC] + [Inactive ADC] + [unconjugated Ab]

The active ADC assay measures free or partially free antibody (one or both arms of the antibody are not bound to soluble target [9]) conjugated to active payload. When developed for quantitative payload sensitivity, the active ADC ligand-binding assay (LBA) can be used to determine DAR and should be highly analogous to the antibody conjugated payload assay measured via immunocapture LC–MS/MS (IC-LC–MS/MS). The total ADC assay measures free or partially free antibody [9] conjugated to active payload or conjugated to inactive payload metabolites.

[Total ADC] = [Active ADC] + [Inactive ADC].

It is likely that site specific technologies will further impact bioanalytical strategies. As ADC and other multiple domain therapeutics become more homogeneous through the implementation of site specific conjugations [10,11], analytical methods for characterizing those therapeutics become less complex. Limited information is available in the literature for ADC that have payload that undergo modification exclusive from deconjugation as is the case discussed in this manuscript, preceding manuscripts [12,13] and manuscripts in preparation [Wang *et al.* Antibody–drug conjugate bioanalysis using lb-lc–ms/ms hybrid assays: strategies, methodology and complementarity to lba (2016), Submitted]. These biotransformation events can have significant impact on the potency of the molecule, making it important to differentiate between the resulting active and inactive analytes. Antipayload antibodies that specifically recognize the parental active version of the payload and the metabolized/inactive version of the payload greatly facilitate the measurement of pharmacologically relevant ADC.

Experimental

LBAs: total Ab, active ADC & total ADC

All of the LBAs were technically and operationally harmonized to improve efficiency and data integration as described in Myler *et al.* [12] and in the methods sections below. In brief, the same standards, quality controls (QC), patient samples, capture Ab and buffers were used for all analytes measured in the LBA assays. Distinct detection Abs were used to provide analyte specificity for the active ADC, total Ab, or total ADC analytes. A single, 20-µl aliquot was used to generate results for all three analytes. Additionally, the antibody conjugated payload immunocapture-LC–MS/MS assay described in Liu *et al.* [13] uses the same capture antibody as the three LBA to provide further continuity.

Nonclinical toxicokinetic methods

The nonclinical methods used to support the GLP toxicology studies utilized a sandwich immunoassay format and the standard MULTI-ARRAY® microplate (Meso Scale Discovery [MSD], Gaithersburg, MD) platform. Nonclinical method details are described in Myler *et al.* [12].

Clinical pharmacokinetic methods

The following instruments, chemicals, noncritical reagents and matrices were used: EZ-Link™ NHS-LC–LC-Biotin (Thermo Scientific, MA, USA), Alexa Fluor647 NHS ester (Succinimidyl Ester; Thermo Scientific, MA, USA), Human Serum, CD-1 mouse serum (Bioreclamation), 10% Tween-20 (Thermo Scientific, MA, USA), Glycine (JT Baker, PA, USA), 10% sodium dodecyl sulfate (SDS; Life Technologies, NY, USA), Immunoassay Diluent (ANP Tech., DE, USA), Milli-Q 'Synthesis' Water Treatment/Purification System (EMD Millipore, MA, USA), Stericup-GP, 0.22 µm, polyethersulfone, 1000/1000 ml, radiosterilized filters (EMD Millipore).

A mouse anti-idiotype (anti-Id) monoclonal antibody was biotinylated [12] and used as the capture reagent for all ADC analytes. The same anti-id Ab was also conjugated to Alexa-647 and used as a detection reagent for the total Ab assay. An antipayload mAb that binds specifically to only the active form of the payload (described above) was conjugated to Alexa-647 (Thermo Scientific, MA, USA) and used to detect the active ADC analyte. An antipayload mAb that binds to both the active and inactive form of the payload was conjugated to Alexa-647 NHS ester (Thermo Scientific, MA, USA) and used to detect the total ADC analyte.

The clinical method for quantitation of active ADC, total Ab and total ADC utilized Bioaffy 1000 microfluidic CDs containing streptavidin-coated columns with a Gyrolab xP Workstation (Gyros Inc, Uppsala, Sweden). Each automated Gyros run was executed using the Gyrolab control software (version 5.4). Three Gyrolab runs were utilized to measure the unique ADC analytes. Two wash solutions were used for all Gyrolab runs. Wash solution 1 consisted of 0.05% Tween 20 in PBS and was connected to station 1 and also used as the hydraulic solution for sample and reagent needles. Wash solution 2 consisted of 50 mM glycine, 0.5% SDS, pH 9.5 and was connected to wash station 2. All wash solutions were filtered prior to use. The minimum required dilutions (MRD) for all analytes were

prepared with immunoassay diluent containing 2% CD-1 mouse serum. The MRD for the active ADC assay was 1:20, total Ab assay was 1:50 and total ADC assay was 1:100. Total Ab and total ADC MRD were prepared from the active ADC MRD.

All Gyros runs were executed as laboratory information management system (LIMS) runs, and work lists were generated by PPD Laboratories® proprietary LIMS (Assist LIMS, Version 5.4). Gyros runs for the active ADC and total ADC analytes utilized a three-step Gyrolab method (Capture-Analyte-Detection) with a PMT gain of 5%. The capture Ab for the active ADC and total ADC analytes was diluted to 100 µg/ml in Immunoassay Diluent. Rexxip F Buffer (Gyros Inc.) was used to dilute the active ADC detection Ab (4 µg/ml) and the total ADC detection Ab (0.5 µg/ml). For the total Ab, a master mix containing 50 µg/ml biotin- and Alexa-647-conjugated anti-Id Ab was mixed with an equivalent volume of the total Ab MRD and then incubated for 30–120 min prior to loading sample plates and initiating the run on the Gyrolab xP workstation. The Gyros run for the total Ab analyte used a one-step Gyrolab method with PMT gain of 5% to deliver samples containing the bridged complexes to the CDs. A four parameter logistic regression curve with $1/Y^2$ weighting was used to calculate the concentrations of active ADC, total Ab, total ADC in the samples.

Method validation experimental design

Bioanalytical methods employed to support regulated studies such as IND-enabling toxicology studies and clinical studies are validated according to specified criteria as defined in various Bioanalytical Method Validation Guidance's and Guidelines [4,14–17]. Standard parameters to be assessed include: intra- and interassay accuracy and precision, selectivity, specificity, dilutional integrity and/or carryover, and stability. Acceptance criteria for these parameters can be found in the cited guidance's but generally consist of precision requirements of ≤20% coefficient of variation (%CV) and accuracy within ±20% bias (or difference from theoretical). Experimental design and criteria details described below pertain specifically to the clinical assay validation but are highly comparable to that used for the nonclinical assay validation.

Accuracy & precision

Intra- and inter-assay accuracy and precision are evaluated as described in DeSilva *et al.* [4]. All validation samples were prepared in 100% matrix prior to performing the MRD. Standard levels included: 100 (low anchor point), 200 (LLOQ), 400, 800, 1600, 3200, 6400, 12,800 (ULOQ) and 25,600 ng/ml (high anchor point). QC levels included: LLOQ at 200 ng/ml; low QC (LQC) at 600 ng/ml; mid QC (MQC) at 2000 ng/ml; high QC (HQC) at 10,000 ng/ml and ULOQ at 12,800 ng/ml. QCs at each level were analyzed ($n \geq 3$) in six separate runs each assay. The dilution QC at 128,000 ng/ml, was diluted 1:50 in 100% human serum prior to performing the MRD. The intra-assay and inter-assay precision (%CV) of the LLOQ was required to be ≤25.0%; precision at all other levels were required to be ≤20.0%. The intra-assay and interassay accuracy as assessed by % difference from theoretical (DFT) or bias at the LLOQ was required to be within ±25.0% and within ±20.0% at all other levels. Total error (precision + |accuracy|, or TE) was required to be ≤ 40.0% at the LLOQ and ≤30.0% at all other levels.

Selectivity

Matrix samples from at least ten individual human donors from each of the applicable patient populations including healthy human donors, ovarian cancer patients, mesothelioma patients, pancreatic cancer patients, non-small-cell lung cancer patients and gastric cancer patients were analyzed at blank, LLOQ and HQC levels to evaluate assay selectivity.

Eighty percent of the blank samples from each population were required to quantitate less than the LLOQ. For each population indicated above, at least 80.0% of the samples at the LLOQ level were required to quantitate within ±25.0% of the theoretical value, and at least 80.0% of the samples at the HQC level were required to quantitate within ± 20.0% of the theoretical value.

Runs containing selectivity samples must also contain pooled matrix controls prepared at equivalent concentrations (LLOQ and HQC). For each run containing selectivity samples at the LLOQ level, at least two of the three pooled matrix control replicates at the LLOQ were required to quantitate within ±25.0% of the theoretical value and have a %CV ≤25.0%. For each run containing selectivity samples at the HCQ level, at least two of the three pooled matrix control replicates were required to quantitate within ±20.0% of the theoretical and have a %CV ≤20.0%. Selectivity samples analyzed in runs containing pooled matrix control samples (at the same level) that failed to meet the above acceptance criteria were re-evaluated in a run with acceptable controls.

Selectivity testing was also done in hemolytic and lipemic samples. To evaluate the effect of sample hemolysis on study sample quantitation, low and high QCs were prepared in matrix hemolyzed to approximately 1100 mg/dl free hemoglobin and analyzed (n = 6 each level). To evaluate the effect of lipemia on study

sample quantitation, low and high QCs were prepared in lipemic matrix with >300 mg/dl triglycerides and analyzed (n = 6 each level). The %CV and mean accuracy was required to be ≤20.0% and within ±20.0%, respectively. The blank hemolyzed matrix was required to quantitate below the LLOQ.

Specificity

Anticipated supraphysiological levels of soluble target (0, 10.0, 30.0, 90.0 270, 810, 2430, 7290 ng/ml), inactive ADC (0.00, 400, 1200, 1800 ng/ml), unconjugated Ab (0.00, 400, 1200, 1800 ng/ml) and anti-drug antibody (ADA; 0.00, 16.0, 80.0, 400, 2000 and 10,000 ng/ml) were evaluated minimally at the blank, LQC and HQC levels to assess interference in the total antibody, total ADC and active ADC assays. Each interference sample (drug plus potential interfering factor), prepared in pooled serum, was incubated at 37°C for at 1–2 h and then frozen at -80°C ± 10°C prior to analysis. Noninterference of the compounds was demonstrated if the interference sample was within ±20.0% of the theoretical value. The ADA used in this experiment was an affinity purified polyclonal Ab derived from hyperimmunized rabbit serum. Rabbits were immunized with drug substance.

The inactive ADC reagent was a discovery grade reagent used to provide qualitative results and facilitate a better understanding of the specificity of each of the total Ab, total ADC and active ADC assays. Although quantitative acceptance criteria were not applied to the inactive ADC interference evaluations, inactive ADC concentrations were expected to be additive for the total Ab and total ADC assays. To demonstrate specificity of the active ADC assay for active analyte only, inactive ADC was required to not be measurable in blank control samples or to positively bias LQC and HQC samples. In order to determine a lack of interference of the inactive ADC (metabolite or impurity) in the active ADC assay, inactive ADC was required to not interfere in the quantitation of the active ADC high and low QCs.

The results for unconjugated Ab were expected to be additive for the total Ab assay only. In order to determine a lack of interference of the unconjugated Ab (metabolite or impurity) in the total ADC and active ADC assays, unconjugated Ab was required to not interfere in the quantitation of the total ADC and active ADC high and low QCs.

Stability

The following stability conditions were evaluated: long-term stability in frozen matrix, freeze–thaw stability, analyte stability in thawed matrix and processed sample stability. Stability samples were spiked at the HQC, LQC and Dil QC levels in pooled normal human serum. For all stability tests five values were required to calculate stability statistics for each QC level tested and the %CV and the mean accuracy of the replicate determinations were required to be ≤20.0% and within ± 20.0%, respectively.

The effect of freezing the analyte in biological matrix and storing samples in this condition over an extended period of time was evaluated by analyzing low, high, and Dil QCs (n = 6) stored in a cryofreezer to support cryofreezer storage and stored in a freezer to support freezer storage. Long-term stability ≥12 months is ongoing.

The influence of the physical processes of freezing and thawing on analyte stability in biological matrix was evaluated by subjecting low, high, and Dil QCs (n = 6) to eight freeze–thaw cycles (8FT). For the first freeze–thaw cycle, samples were maintained at -80°C for a minimum of 24 h and for subsequent cycles, the samples are held in the freezer for at least 12 h. Following each freezer cycle, the samples are removed from the freezer until completely thawed.

To evaluate whether holding samples in a thawed state at room temperature for at least 6 h or on wet ice for at least 24 h adversely affects analyte stability in biological matrix, low, high and Dil QCs (n = 6) were maintained frozen for a minimum of 24 h and then thawed and maintained at room temperature for at least 6 h or on wet ice for at least 24 h.

To evaluate whether holding samples that have been diluted in the minimum required dilution (MRD) buffer adversely affects analyte stability, low, high and Dil QCs (n = 6) were diluted to the MRD with MRD buffer and then maintained at least 16 h at room temperature. Additional sets of processed stability samples (n = 6) were diluted to the MRD in MRD buffer and maintained for 24 and 48 h at 2–8°C.

Dilutional integrity & hook effect

The ability to dilute study samples >ULOQ was evaluated by preparing a Dil QC at or above the estimated maximum concentration (C_{max}) of ADC in study samples and diluting it into assay range. The Dil QC was prepared by spiking reference material into pooled normal human serum to achieve a concentration of 128,000 ng/ml drug in at least 98% serum. Serial dilutions were prepared by diluting the Dil QC in 100% pooled normal human serum prior to diluting to the MRD. The dilution series extended above and into the range of quantitation and were analyzed at 128,000 ng/ml (Dil 1), 32,000 ng/ml (Dil 4), 8000 ng/ml (Dil 16), 2000 ng/ml (Dil 64) and 500 ng/ml (Dil 256).

The %CV and the mean accuracy for each dilution QC replicate was required to be ≤20.0% for those dilu-

tions within the range of quantitation. For dilutions where the expected response after dilution is above the highest calibration curve point, the result was required to read above the highest calibration standard or greater than the ULOQ. Samples with concentrations of analyte greater than the ULOQ that quantitate within the range of quantitation are indicative of a prozone or hook effect.

Carryover

To evaluate the potential carryover of transfer needles, a QC pool prepared at the LLOQ and diluted to the MRD in MRD buffer was transferred from the sample plate to the CD by all eight sample needles (one replicate; 2 CD structures) before and immediately after each needle had transferred one replicate (two CD structures) of a Dil QC sample (128,000 ng/ml) from the sample plate to the Gyros CD. Thee-fourths of the LLOQ samples transferred before and after the Dil QC were required to have %CV and mean accuracy values ≤25.0% and within ± 25.0%, respectively.

DAR sensitivity

Due to the discovery grade quality of the DAR reagents, DAR sensitivity testing was done outside of the formal validation as part of method development to further characterize the active ADC assay which is the assay thought to best correlate with safety and efficacy. The standards and QCs were prepared from reference material (mean DAR 3.0) and drug substance with mean DAR of 3.9, 2.0, 1.7 and 0.8 were evaluated at the low, middle and high ends of the standard curve. Purified DAR species testing is in progress and will be included in a subsequent manuscript.

Results & discussion

The bioanalytical strategy consisted of several independent measures of the ADC including: total antibody, active ADC, total ADC, Ab-conjugated payload, unconjugated payload, soluble/shed target, a panel of biomarkers indicative of programmed cell death and a series of immunogenicity assays designed to elucidate antidrug antibody specificity and is described in Myler *et al.* [12]. The three LBA used to assess PK were technically and operationally harmonized to maximize functionality and data integration. To further increase efficiency in the clinical setting, the platform was changed from a plate-based electrochemiluminescense platform to a CD-based microfluidics platform that required only a single, 20 µl aliquot to generate results for the three analytes. The same standards, QCs, patient samples, capture Ab and buffers were used for all analytes. Distinct detection Abs were used to provide specificity for the active ADC, total Ab, or total ADC analytes. Each sample was diluted to the analyte-specific MRD and was run in accordance with Figure 1. The diluted samples were loaded onto the Gyrolab and analyzed for active ADC, total Ab and total ADC in series. The bridging incubation for the total Ab occurred during the active ADC run time so as to eliminate down time due to incubation steps. Each CD has a run time of 1 h, enabling the generation of approximately 200 results in 7 h by a single analyst (6-h run time and 1-h sample preparation).

Validation results

The assays described herein were validated in accordance with current global regulatory guidance's [14–17] and are consistent with applicable white papers [7].

Accuracy & precision

The assay range for the three clinical assays (LLOQ at 200 ng/ml to ULOQ at 12,800 ng/ml) was established and validated so that clinical sample dilution was minimized, thereby maximizing operational efficiency.

Table 1 provides a summary of the clinical validation accuracy and precision data for the active ADC, total Ab and total ADC assays. For all 3 LBA, total error ranged from 4.4 to 21.4, %Bias ranged from -2.6 to 13.8 and %CV (interassay) ranged from 3.8 to 10.7. For the active ADC alone, total error ranged from 6.6 to 16.8, %Bias ranged from 2.0 to 8.9 and %CV (interassay) ranged from 3.8 to 7.9, well within the acceptable limits per regulatory guidelines. Supplementary Table 1 provides a validation summary for the nonclinical validation of the total Ab and active ADC assays in rat and monkey serum. For all LBA in both nonclinical species matrix, total error ranged from 3 to 15, %Bias ranged from -8 to 10 and %CV (interassay) ranged from 3 to 9. All parameters tested including accuracy and precision, robustness, hook effect, dilutional integrity, selectivity, specificity and stability met acceptance criteria. Additionally, 54/54 (100%) of the total Ab and 52/54 (96%) of the active ADC samples met incurred sample reanalysis (ISR) criteria, were within 30% of the original value.

Selectivity

Over 80% of samples prepared at LLOQ & HQC levels in healthy human donor, ovarian cancer, mesothelioma, pancreatic cancer, non-small-cell lung cancer and gastric cancer patient serum met acceptance criteria, within ±25% bias at LLOQ and ±20% bias at HQC; 100% of blank samples < LLOQ (Figure 2). There was no effect from hemolysis, ≥1100 mg/dl of hemoglobin, or lipemia, >300 mg/dl triglycerides, on the quantitation of the active ADC, total Ab, or total ADC (data not shown).

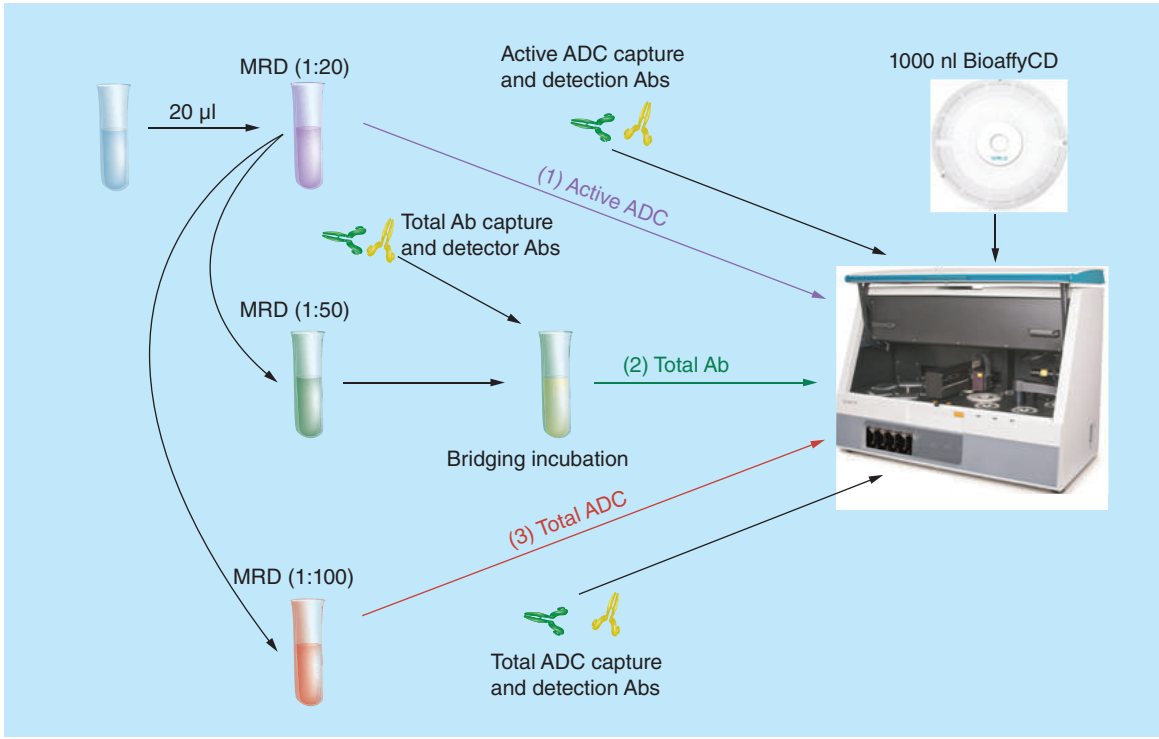


Figure 1. Automated Gyrolab™ xP workstation enabled development of single, multianalyte PK Method: 192 results in 7 h by one analyst.

Specificity

Target sensitivity

The relevance of soluble and/or shed target levels in the ADC PK assays are described in greater detail in Myler *et al.* [12]. The clinical assay validation results for this parameter are described herein. Shed target has no impact on samples that do not contain ADC. In the active ADC and total Ab assays, no interference was observed for the LQC with soluble target ≤90 ng/ml (10, 30 and 90 ng/ml); negative bias was observed for LQC spiked with soluble target ≥270 ng/ml (270, 810, 2430 and 7290 ng/ml). No interference was observed for HQC spiked with soluble target ≤810 ng/ml (10, 30, 90, 270 and 810 ng/ml); negative bias observed for HQC spiked with soluble target at 2430 and 7290 ng/ml. In the total ADC assay no interference was observed for the LQC spiked with soluble target ≤270 ng/ml (10, 30, 90 and 270 ng/ml); negative bias was observed for LQC spiked with soluble target ≥810 ng/ml (810, 2430 and 7290 ng/ml). No interference was observed for HQC spiked with soluble target ≤2430 ng/ml (10, 30, 90, 270, 810 and 2430 ng/ml); negative bias was observed for HQC spiked with soluble target at 7290 ng/ml. As discussed previously, the measurement of free (at least one Ab arm not bound to target) active (nonmetabolized payload) ADC is the desired measurement as this allows for a more accurate determination of bioactive ADC.

Anti-drug antibody interference

ADA impact on PK is assessed during method validation by spiking an ADA positive control into high and low quality controls. The impact of ADA on exposure response analysis is described in the literature and was recently addressed by the global bioanalytical consortium [18]. No interference was observed at the blank or HQC spiked with ADA up to 10,000 ng/ml (16, 80, 400, 2000 and 10,000 ng/ml). In the active ADC assay, no interference was observed for LQC spiked with ADA ≤400 ng/ml (16, 80 and 400 ng/ml); negative bias was observed for LQC spiked with ADA at 2000 and 10,000 ng/ml. In the total Ab and total ADC assays, no interference observed for LQC spiked with ADA ≤2000 ng/ml (16, 80, 400 and 2000 ng/ml); negative bias was observed for LQC spiked with ADA at 10,000 ng/ml.

Analyte specificity & impurity/metabolite tolerance

To test the specificity of the active ADC, total Ab and total ADC assays, drug product with a mean DAR of 3.0 was spiked with increasing levels of inactive ADC that had chemically inactivated payload or unconjugated Ab that had no payload. Inactive ADC should be detectable in the total Ab and total ADC assays only. Unconjugated Ab should only be detectable in the total Ab assay. Impurity and or metabolite levels were

Table 1. Clinical assay accuracy and precision.

Analyte	Sample	Nominal conc. (ng/ml)	Precision		Accuracy	Total error
			Intra-assay %CV	Interassay %CV	Mean %Bias	(Mean %Bias) + Interbatch %CV
Active ADC	LLOQ	200	3.8	3.8	3.4	7.2
	LQC	600	5.9	6.2	2.6	8.7
	MQC	2000	4.5	4.5	2.1	6.6
	HQC	10,000	6.4	6.8	5.3	12.0
	ULOQ	12,800	4.5	5.2	2.0	7.2
	Dil QC	128,000	2.8	7.9	8.9	16.8
Total Ab	LLOQ	200	6.4	9.4	-2.6	12.0
	LQC	600	4.8	6.6	2.6	9.2
	MQC	2000	5.1	5.1	2.9	8.0
	HQC	10,000	3.2	4.7	3.1	7.9
	ULOQ	12,800	1.9	3.9	0.5	4.4
	Dil QC	128,000	2.1	6.8	5.8	12.5
Total ADC	LLOQ	200	3.1	4.3	3.2	7.5
	LQC	600	3.4	4.9	3.6	8.5
	MQC	2000	4.9	5.7	1.7	7.4
	HQC	10,000	4.2	7.6	13.8	21.4
	ULOQ	12,800	5.1	10.7	10.0	20.7
	Dil QC	128,000	3.5	9.8	10.5	20.3

n = 18, 6 runs, three replicates each run.
%CV = (standard deviation/mean) ×100;
%Bias = (mean calculated concentration-nominal concentration)/nominal concentration × 100 or;
%Bias = (mean calculated concentration/ nominal concentration)-1] × 100 (calculation used in above data table).

selected based upon the nonclinical primate model where active ADC was shown to comprise 20–25% of the heterogeneous mixture 504 h after dosing [12].

The presence of inactive ADC or unconjugated Ab was not expected to interfere with or positively bias the quantitation of the active ADC analyte, because the detection Ab for active ADC assay is specific for the active form of payload. The results presented in Table 2 show that all blank samples spiked with increasing levels of inactive ADC or unconjugated Ab impurity or metabolite quantitated below the LLOQ, and that all LQC and HQC samples spiked with increasing levels of impurity or metabolite quantitated within ± 20% of the nominal ADC concentration, demonstrating specificity for the active payload and insignificant interference from unconjugated Ab and inactive ADC when spiked at up to 75% of the heterogeneous mixture.

Inactive ADC and unconjugated Ab were both expected to be measurable in the total Ab assay. The expected result for the total Ab assay spiked with unconjugated Ab was expected to be equal to the sum of the nominal concentration of ADC and unconjugated Ab. The results presented in Table 3 show that all samples spiked with increasing levels of unconjugated Ab quantitated within ± 20% of the expected result [ADC + unconjugated Ab], demonstrating equal quantitation of unconjugated Ab and Ab conjugated to payload (mean DAR of 3.0) in the total Ab assay at up to 75% unconjugated Ab impurity or metabolite. The concentration of unconjugated Ab was not accounted for in the analytical specifications for the inactive ADC reagent. Thus, over-recovery was anticipated in the total Ab assay and only qualitative data was reported for this impurity. Results presented in Table 3 indicate that increasing concentrations of the inactive ADC in samples prepared at blank, low and high QC levels resulted in cumulative increases in measured concentrations of total Ab. It is understood that higher levels of unconjugated Ab impurity exist in this formulation than were accounted for in the specifications and/or calculations, thus resulting in over-recovery in the total Ab assay.

The presence of unconjugated Ab was not expected to interfere with or positively bias the quantitation of the total ADC analyte, because the detection Ab for total ADC assay is specific for the payload. The results

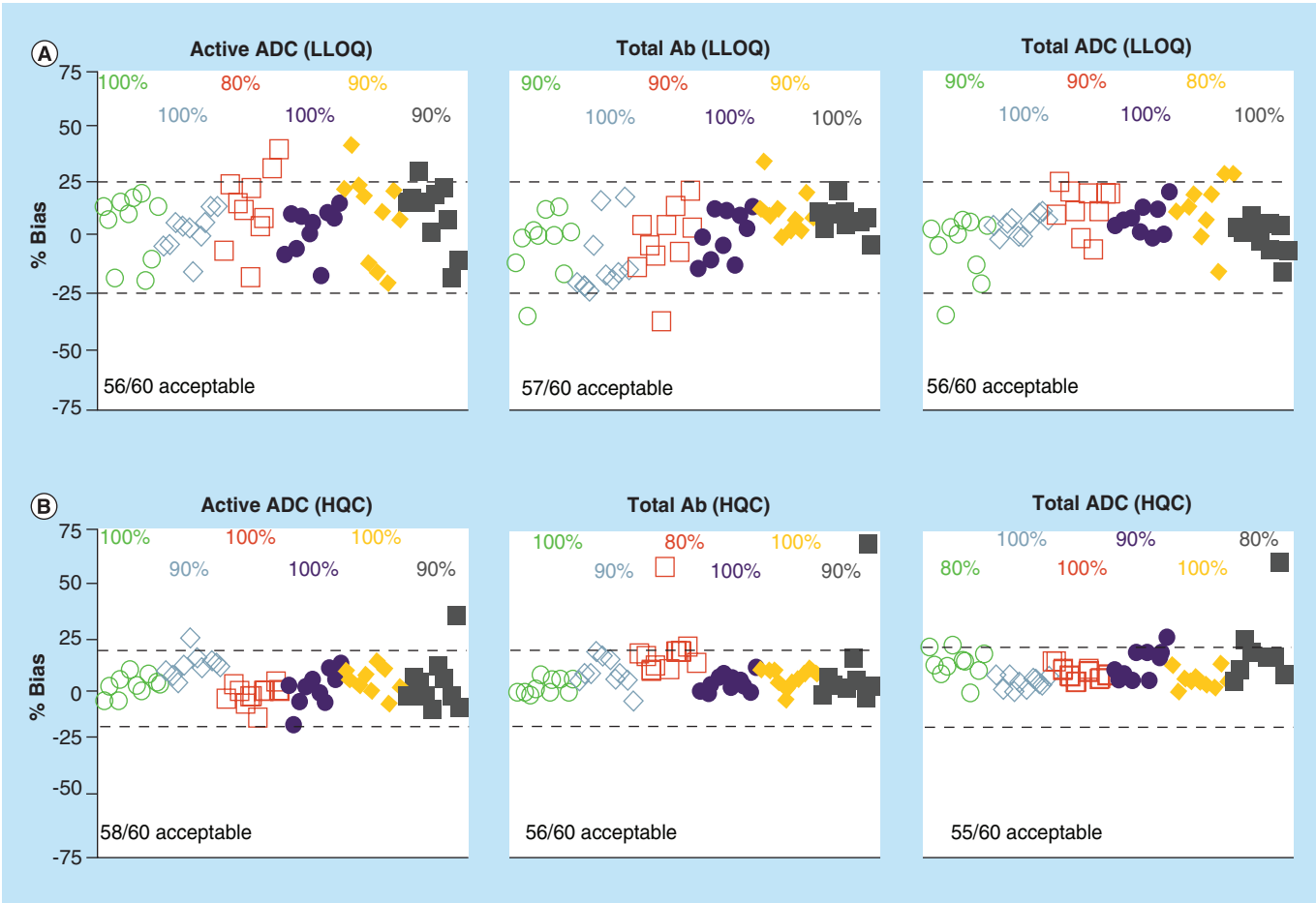


Figure 2. Clinical assay selectivity. Open circles indicate ovarian cancer; open diamonds indicates mesothelioma; open squares indicate non-small-cell lung cancer; closed circles indicate pancreatic cancer; closed diamonds indicate gastric cancer; closed squares indicate healthy human serum. In each panel, the color-matched percentages listed above each population denote the percent of individuals with acceptable %Bias. Overall results for all individuals are listed at the bottom of each graph.

presented in Table 4 show that all blank samples spiked with increasing levels of unconjugated Ab quantitated below the LLOQ, and that all LQC and HQC samples spiked with increasing levels of this impurity/metabolite quantitated within ±20% of the nominal ADC concentration, demonstrating specificity for Ab conjugated to payload and insignificant interference from unconjugated Ab impurity/metabolite at up to 75% of the heterogeneous mixture. Although quantitative acceptance criteria were not applied to inactive ADC interference evaluations due to the discovery grade nature of this reagent, as described above, the result of increasing inactive ADC concentrations was expected to be additive in the total ADC assay. Results presented in Table 4 indicate that increasing concentrations of the inactive ADC in samples prepared at blank, low, and high QC levels resulted in cumulative increases in measured concentrations of total ADC. The DAR normalized data are shown and indicate acceptable % bias; however, these data are only included for informational purposes as quantitative criteria were not applied in the validation. DAR normalization is discussed in greater detail in the DAR sensitivity section below.

In summary, the total Ab assay, Table 3, is DAR independent and thus does not require normalization based upon DAR. The active ADC and total ADC assays are DAR dependent or DAR sensitive, and thus do require normalization based upon DAR (Table 4). The active ADC and total ADC assays are insensitive to unconjugated Ab and thus the presumed additional unconjugated Ab in the inactive ADC reagent has no impact on the these assays. The total Ab assay, however, is sensitive to unconjugated Ab thus the over-recovery due to presumed levels of unconjugated Ab impurity.

Analyte stability

A set of low, high and Dil QCs (n = 3 vials each) were tested in each assay, active ADC, total Ab and total ADC, under a variety of conditions to establish analyte stability in the applicable matrix including: frozen serum stability (ongoing at -70°C and -20°C),

Table 2. Active ADC assay specificity: impact of impurities and metabolites.

Spiked conc. (ng/ml)			Expected result (ng/ml)	Observed result (ng/ml)	%Bias from expected
ADC (DAR 3.0)	Unconjugated Ab	Inactive ADC			
0	400	0	0	<LLOQ	N/A
0	1800	0	0	<LLOQ	N/A
0	0	400	0	<LLOQ	N/A
0	0	1800	0	<LLOQ	N/A
600	400	0	600	573	-4.5
600	1800	0	600	597	-0.5
600	0	400	600	647	7.9
600	0	1800	600	581	-3.1
10,000	400	0	10,000	9044	-9.6
10,000	1800	0	10,000	8686	-13.1
10,000	0	400	10,000	9349	-6.5
10,000	0	1800	10,000	8339	-16.6

The active ADC assay measures Ab conjugated to active payload only with insignificant interference from up to 75% unconjugated Ab and inactive ADC metabolites and/or impurity.

thawed serum stability (6 h at RT and 24 h on wet ice), processed sample stability (post MRD; 16 h at RT and 48 h at 2–8°C), and freeze–thaw stability (eight cycles). All sample data met the criteria, within ± 20% bias, for demonstrating stability under these conditions. The %Bias data is summarized in Table 5. Stability in the active ADC assay shows that the payload is not hydrolyzed nor is the linker cleaved under these conditions. Similarly, stability in the total ADC assay supports that the linker is not cleaved under these conditions. Stability in the total Ab assay indicates that the target binding region is still intact. These stability data have been cumulatively graphed in Supplementary Figure 1 to show the comparability of stability in each of the assays for each of the analytes.

Table 3. Total antibody assay specificity: impact of impurities and metabolites.

Spiked conc. (ng/ml)			Expected result (ng/ml)	Observed result (ng/ml)	%Bias from expected
ADC (DAR 3.0)	Unconjugated Ab	Inactive ADC			
0	400	0	400	475	18.9
0	1800	0	1800	2051	13.9
0	0	400	400	574	Detectable
0	0	1800	1800	2258	Detectable
600	400	0	1000	1025	2.5
600	1800	0	2400	2557	6.6
600	0	400	1000	1250	Detectable
600	0	1800	2400	2907	Detectable
10,000	400	0	10,400	11,788	13.3
10,000	1800	0	11,800	12,970	9.9
10,000	0	400	10,400	11,762	Detectable
10,000	0	1800	11,800	13,534	Detectable

The total antibody assay measures all Ab related species, Ab conjugated to active payload (active ADC), Ab conjugated to inactive payload (inactive ADC) and Ab w/ no payload (unconjugated Ab).
Over-recovery of the inactive ADC is attributed to levels of unconjugated Ab impurity that were not accounted for in the specifications. Quantitative criteria were not applied to this analyte during validation.

Table 4. Total ADC assay specificity: impact of impurities and metabolites.

Spiked conc. (ng/ml)			Expected result (ng/ml)	Observed result (ng/ml)	%Bias from expected
ADC (DAR 3.0)	Unconjugated Ab	Inactive ADC (DAR 2.2)			
0	400	0	<LLOQ	<LLOQ	N/A
0	1800	0	<LLOQ	<LLOQ	N/A
0	0	400	293	287	-2.0
0	0	1800	1320	1066	-19.2
600	400	0	600	576	-4.0
600	1800	0	600	645	7.5
600	0	400	893	911	2.0
600	0	1800	1920	1630	-15.1
10,000	400	0	10,000	11,255	12.6
10,000	1800	0	10,000	11,216	12.2
10,000	0	400	10,293	11,445	11.2
10,000	0	1800	11,320	12,116	7.0

The total ADC assay measures Ab conjugated to active and inactive payload with insignificant interference from up to 75% unconjugated Ab metabolites and/or impurity.
Quantitative criteria were not applied to the inactive ADC analyte during validation due to the discovery grade nature of this reagent. The expected results are normalized for DAR 2.2 calibrated to the DAR 3.0 calibration curve. The impact of DAR on ADC quantitation is discussed in greater detail below.

Dilutional integrity & Hook effect

The Dil QC containing 128,000 ng/ml was assessed undiluted and at 4-, 16-, 64- and 256-fold dilutions. All results within the assay range, that is diluted at least 16-fold in 100% human serum prior to the MRD, were within ± 20% bias in all three assays. Dilutions where the concentration after dilution were above the highest calibration curve point, were >ULOQ, thus there was no 'hook effect' observed at concentrations up to 128,000 ng/ml (data not shown).

Carryover

The potential carryover of transfer needles was evaluated by transferring LLOQ QC samples from the sample plate to the CD by all eight sample needles before and immediately after each needle transferred one replicate of Dil QC (128,000 ng/ml) from the sample plate to the CD. As shown in online available Supplementary Table 2, there was no evidence of carryover in any of the three assays (active ADC data shown).

DAR sensitivity

DAR sensitivity was tested in the active ADC assay by spiking nominal low or high concentrations of various DAR species into the assay including mean DAR 0.8, 1.7, 2.0 and 3.9, normalizing the results for DAR and assessing the %Bias. As shown in Table 6, the active ADC assay is shown to be quantitatively proportional for each DAR species. For example, DAR 0.8 is 27% of DAR 3.0, the calibrator DAR, thus the normalized expected result for a DAR 0.8 species would be 27% of the nominal concentration. At a nominal concentration of 1570 ng/ml DAR 3.0, DAR 0.8 would be expected to yield 27% of that or 419 ng/ml. The observed result for these conditions was 432 ng/ml resulting in a 3% bias. DAR species, 0.8 through 3.9, were tested at nominal concentrations of 1570 ng/ml and 6820 ng/ml, translating to 419 through 8866 ng/ml once DAR normalized. All except one of the conditions tested had %bias within ±20%. Due to the discovery grade quality of these DAR species, these evaluations were done in method development and were not performed during method validation. The characterization of DAR species, including the distribution around the mean DAR value is likely to impact the quality of associated results, so caution should be taken when carrying out these experiments and analyzing associated data.

In addition to the *in vitro* testing of individual DAR species discussed above, *in vivo* testing also shows the active ADC assay to be quantitatively DAR proportional as demonstrated by the parallelism to the Ab-conjugated payload immunocapture LC–MS/MS (Supplementary Table 3). More detail on this can be found in Myler *et al.*

Conclusion

The inherent complexity of randomly conjugated ADC, exacerbated by *in vivo* processing, necessitates

Table 5. Active ADC, total Ab and total ADC stability.

	%Bias								
	Active ADC			Total Ab			Total ADC		
	LQC	HQC	Dil QC	LQC	HQC	Dil QC	LQC	HQC	Dil QC
Freeze–thaw, eight cycles	1.0	9.1	19.9	14.5	3.0	12.2	5.0	3.1	12.8
Thawed serum, wet ice, 24 h	1.4	10.5	13.6	6.5	17.9	17.9	2.1	7.1	3.5
Thawed serum, RT, 6 h	-1.3	7.6	12.6	12.1	0.8	11.4	2.4	2.8	4.3
Processed sample, ≥16 h, RT	4.8	2.0	6.8	12.2	8.1	12.1	5.1	7.2	10.1
Processed sample, ≥48 h, 2–8°C	-1.6	-0.9	8.7	5.5	12.0	8.5	3.5	0.1	8.8
Frozen serum, 39 days, -25°C ± 5 C	6.2	-7.0	7.4	0.8	7.2	12.3	3.6	8.6	12.7
Frozen serum, 40 days, -80°C ± 10°C	3.8	-5.0	7.6	-1.8	5.7	11.7	2.0	1.8	8.7

the implementation of a series of bioanalytical methods to adequately describe the pharmacological and PK properties. Each assay described herein provides information about a particular ADC species present in circulation at any given time. It is important to understand the quantitative relationships among the different ADC species which, in combination with the bioanalytical data, can be accomplished through PK modeling and simulation approaches [19,20].

The automated microfluidic Gyrolab platform allowed for the development and validation of a single method that could be run by a single analyst to quantitate all three ADC analytes, active ADC, total Ab and total ADC, in series in under 4 h using a single 20 µl sample (Figure 1). This semiautomated, single method, multianalyte approach offered efficiency gains over other more laborious methods by reducing the documentation and perhaps more significantly reducing the sample processing burden while substantially increasing analyst throughput. The assays were determined to meet all validation criteria as described in the health authority guidance's on bioanalytical method validation [14–17].

Future perspective

As the number of ADC in the clinic is increasing, bioanalytical strategies to identify and quantitate ADC analytes for PK analysis continues to evolve. At different stages of ADC drug development, the need for essential bioanalytical methods varies. We are still identifying what is needed when but in general, the focus of bioanalytical assays in the early drug discovery stage is to provide exposure information, metabolism information, toxicokinetic (TK) characterization, and to predict toxicity and efficacy in humans. Understanding these parameters can assist in the selection of a lead ADC molecule whereas in the later nonclinical and clinical drug development stages, bioanalytical focus shifts to providing regulatory compliant PK/TK data for a comprehensive understanding of the pharmacology and toxicology of the therapeutic.

Due to the highly heterogeneous nature of randomly conjugated ADCs, a combination of LBA and LC–MS/MS-based bioanalytical assays are used to support ADC programs. If random conjugation transitions to site-specific conjugation for ADC, some of the bioanalytical considerations related to DAR that we have addressed herein, such as DAR dynamics, may become unnecessary. LBA assays due to their simplicity, throughput, accuracy and reproducibility will continue to play a major role in ADC bio-

Table 6. DAR sensitivity

DAR	0.8	1.7	2.0	3.9
% of DAR 3.0	27%	57%	67%	130%
Nominal conc. (ng/ml)			1570	
DAR normalized conc. (ng/ml)	419	890	1047	2041
Result (ng/ml)	432	928	1165	2385
%Bias of normalized result	3	4	11	17
Nominal conc. (ng/ml)			6820	
DAR normalized conc. (ng/ml)	1819	3865	4547	8866
Result (ng/ml)	1680	4210	5095	11200
%Bias of normalized result	-8	9	12	26

DAR normalized concentration indicates that the concentration was normalized to the DAR 3.0 calibrator. For example, DAR 0.8 is 27% of DAR 3.0 (calibrator DAR) and 419 ng/ml (DAR 0.8) is 27% of 1570 ng/ml (DAR 3.0).

analysis. Compared to the conventional colorimetric (ELISA) platforms, second and third generation platforms including chemiluminescence, ECL and fluorescent microfluidics offer better sensitivity, a broader dynamic range, and decreased sample volume; some come with automation and/or multiplexing options. As new classes of payload are being evaluated, we will continue to generate highly specific reagents, such as that described herein, to support the PK and immunogenicity assessment of these new classes. These reagents are used across the pipeline where a specific payload is employed. This reuse potential makes these reagents quite cost effective.

Tracing the biotransformation of payload associated with ADC in circulation continues to be of utmost importance. Immuno capture based LC–MS conjugated payload assays are going to continue to be critical to evaluate the *in vitro* and *in vivo* serum stability of ADC and to help identify pertinent metabolites. As more fully discussed in Myler *et al.*, an integrated set of assays can be used to quantify the most prevalent and relevant analytes [12].

Supplementary data

To view the supplementary data that accompany this paper please visit the journal website at: www.future-science.com/doi/full/10.4155/bio.16.13

Financial & competing interests disclosure

The authors have no relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript. This includes employment, consultancies, honoraria, stock ownership or options, expert testimony, grants or patents received or pending, or royalties.

No writing assistance was utilized in the production of this manuscript.

Ethical conduct of research

The authors state that they have obtained appropriate institutional review board approval or have followed the principles outlined in the Declaration of Helsinki for all human or animal experimental investigations. In addition, for investigations involving human subjects, informed consent has been obtained from the participants involved.

Executive summary

- Unique bioanalytical strategies are needed to identify, characterize and quantify the ADC species relevant to safety and efficacy.
- Reagents that allow for the derivation of active and inactive ADC are a crucial part of our bioanalytical PK strategy which includes the measurement of total antibody, active ADC and total ADC by LBA, conjugated payload by IC-LC–MS/MS and unconjugated payload by LC–MS.
- All LBA were validated according to current regulatory guidelines for bioanalytical method development to support nonclinical IND-enabling toxicology studies and clinical studies.
- All validation criteria were met and additional characterization data were generated to enhance assay and PK understanding.
- Given the accuracy of the active ADC assay in proportionally measuring payload conjugated to antibody, active ADC/total Ab can be used to assess DAR.
- The number of assays used to characterize ADC PK may be reduced to only include the most value added endpoints following the evaluation of sufficient clinical data.
- Clinical ligand-binding assays were automated using a miniaturized platform to optimize robustness and throughput.
- Minimal sample processing, reproducible results.
- One analyst can generate 192 results (64/assay) each 7-h shift, equivalent to 1 result every 2 min.

References

Papers of special note have been highlighted as:
• of interest; •• of considerable interest

1 Diamantis N, Banerji U. Antibody–drug conjugates – an emerging class of cancer treatment. *Br. J. Cancer* doi: 10.1038/bjc.2015.435 (2016) (Epub ahead of print).

2 Peters C, Brown S. Antibody–drug conjugates as novel anti-cancer chemotherapeutics. *Biosci. Rep.* 35(4), 1–20 (2015).

3 Perez HL, Cardarelli PM, Deshpande S *et al.* Antibody–drug conjugates: current status and future directions. *Drug Discov. Today* 19(7), 869–881 (2014).

4 DeSilva B, Smith W, Weiner R *et al.* Recommendations for the bioanalytical method validation of ligand-binding assays to support pharmacokinetic assessments of macromolecules. *Pharm. Res.* 20, 1885–1900 (2003).

•• **High-impact manuscript describing requirements for biotherapeutic bioanalytical method validation.**

5 Myler HA, Given A, Kolz K *et al.* Biotherapeutic bioanalysis: a multi-indication case study review. *Bioanalysis* 3(6), 623–643 (2011).

6 Kaur S, Xu K, Saad OM *et al.* Bioanalytical assay strategies for the development of antibody–drug conjugate biotherapeutics. *Bioanalysis* 5(2), 201–226 (2013).

•• **High impact manuscript describing the potential ADC species and associated bioanalytical strategies employed to understand drug disposition.**

7 Gorovits B, Alley SC, Bilic S *et al.* 2011 Bioanalysis of antibody–drug conjugates: American Association of Pharmaceutical Scientists Antibody–Drug Conjugate Working Group position paper. *Bioanalysis* 5(9), 997–1006 (2013).

• **Collaborative manuscript whereby many key opinion leaders across industry have addressed bioanalytical challenges for ADC.**

8 Hamblett KJ, Senter PD, Chace DF *et al.* Effects of drug loading on the antitumor activity of a monoclonal antibody drug conjugate. *Clin. Cancer Res.* 10(20), 7063–7070 (2004).

9 Lee JW, Kelley M, King LE *et al.* Bioanalytical approaches to quantify "total" and "free" therapeutic antibodies and their targets: technical challenges and PK/PD applications over the course of drug development. *AAPS J.* 13(1), 99–110 (2011).

• **Describes the importance of fit-for-purpose assay design, specifically with respect to target-therapeutic binding status.**

10 Cho H, Daniel T, Buechler YJ *et al.* Optimized clinical performance of growth hormone with an expanded genetic code. *Proc. Natl Acad. Sci. USA* 108(22), 9060–9065 (2011).

11 Sukumaran S, Gadkar K, Zhang C *et al.* Mechanism-based pharmacokinetic/pharmacodynamic model for THIOMAB™ drug conjugates. *Pharm. Res.* 32(6), 1884–1893 (2014).

12 Myler H, Rangan V, Wang J *et al.* An integrated multiplatform bioanalytical strategy for antibody–drug conjugates: a novel case study. *Bioanalysis* 7(13), 1569–1582 (2015).

•• **Preceding manuscript describing the potential ADC species and associated bioanalytical strategies employed to understand drug disposition.**

13 Liu A, Kozhich A, Passmore D *et al.* Quantitative bioanalysis of antibody-conjugated payload in monkeyplasma using a hybrid immuno-capture LC–MS/MS approach: assay development, validation, and a case study. *J. Chromatogr. B* 1002, 54–62 (2015).

14 US Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research. Guidance for Industry: Bioanalytical Method Validation. Rockville, MD (2013). www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm368107.pdf

15 European Medicines Agency. Guideline on Bioanalytical Method Validation. London, England (2011). www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2011/08/WC500109686.pdf

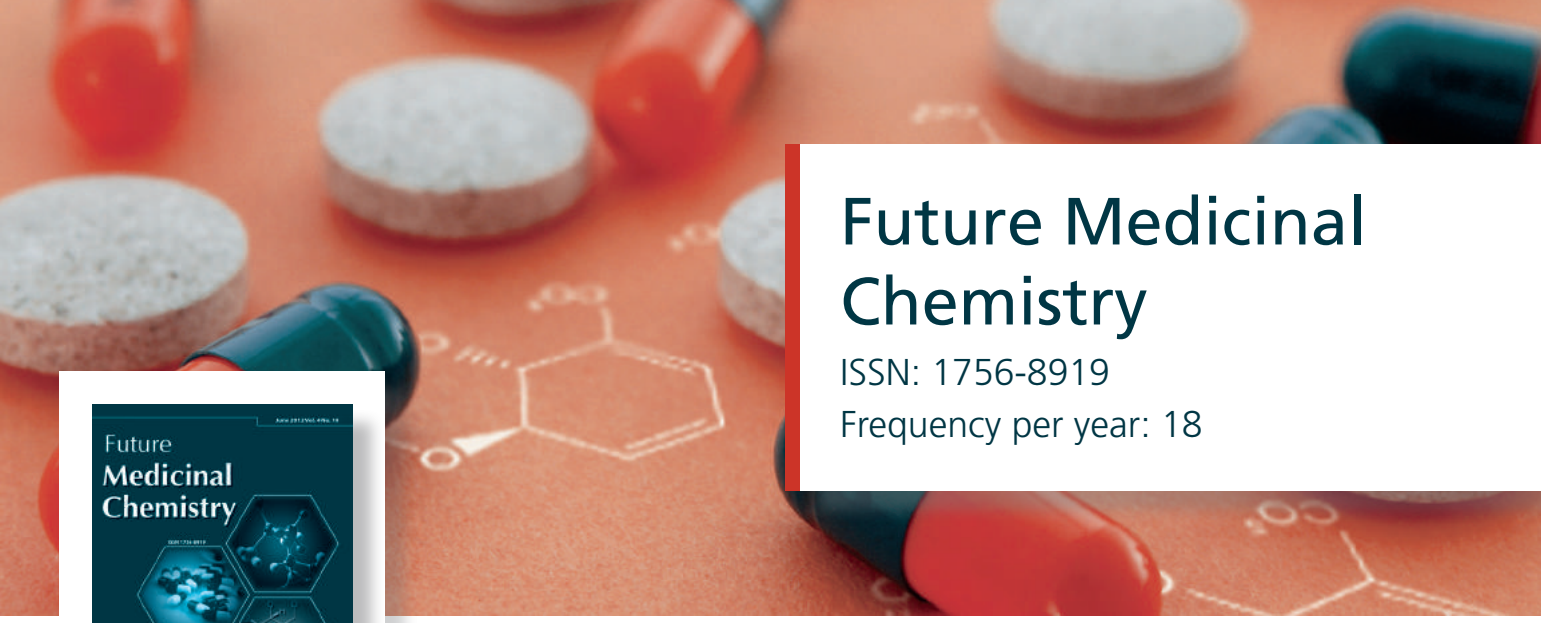
16 Japan Ministry of Health and Labor Welfare. Guideline on Bioanalytical Method (Ligand Binding Assay) Validation in Pharmaceutical Development (2014). www.nihs.go.jp/drug/BMV/BMV-LBA_draft_140124_E_rev.pdf

17 China Pharmacopoeia (ChP). Draft Guidance on Bioanalytical Method Validation (2014). http://knex2014.europeanbioanalysisforum.eu/site/ebf_knex2014/assets/slides/Dafang-Zhong--Guidelines-on-BMV-in-China.pdf

18 Sailstad JM, Amaravadi L, Clements-Egan A *et al.* A white paper–consensus and recommendations of a global harmonization team on assessing the impact of immunogenicity on pharmacokinetic measurements. *AAPS J.* 16(3), 488–498 (2014).

19 Gibiansky L, Gibiansky E. Target-mediated drug disposition model and its approximations for antibody–drug conjugates. *J. Pharmacokinet. Pharmacodyn.* 41, 35–47 (2014).

20 Lu D, Gibiansky L, Jin JY *et al.* Integrated pharmacokinetic model for antibody–drug-conjugate (ADC) in patients with B-cell malignancy: implications for optimal sampling. *J. Pharmacokinet. Pharmacodyn.* 41, S44 (2014).



Future Medicinal Chemistry

ISSN: 1756-8919
Frequency per year: 18

IMPACT FACTOR:
3.556 (2016)

In association with
MedChemNet
www.medchemnet.com

FUTURE MEDICINAL CHEMISTRY COVERS KEY AREAS SUCH AS:

- Drug discovery technologies
- Design and synthesis of novel lead compounds
- Computational approaches to synthesis and modeling
- Biophysical techniques for structural characterization
- Overviews of novel therapeutic strategies and emerging targets
- Drug targeting and delivery
- ADME/Tox investigations
- Commentary on science policy, economic and IP issues

"All medicinal chemists and further drug discovery experts – whether highly experienced or just starting out, whether in pharma or academia, should look to *Future Medicinal Chemistry* for the latest in emerging research in addition to timely and well-produced reviews of the globes most recent advances in not just medicinal chemistry, but drug discovery in general."

– **Jonathan Baell**, Senior Editor, Monash University, Australia

INDEXING

MEDLINE®, Chemical Abstracts EMBASE/ Excerpta Medica, BIOSIS Reviews, BIOSIS Reviews Reports and Meetings, Science Citation Index ExpandedTM (Scisearch®), Chemistry Citation Index®, Journal Citation Reports/Science Edition®, Scopus

www.future-science.com

To claim your free trial, contact us at: trials@futurescience.com

newlands press part of fsg

Review

Bioanalysis

Metabolites and Metabolomics

Analytical pitfalls and challenges in clinical metabolomics

Metabolomics-based strategies have become an integral part of modern clinical research, allowing for a better understanding of pathophysiological conditions and disease mechanisms, as well as providing innovative tools for more adequate diagnostic and prognosis approaches. Metabolomics is considered an essential tool in precision medicine, which aims for personalized prevention and tailor-made treatments. Nevertheless, multiple pitfalls may be encountered in clinical metabolomics during the entire workflow, hampering the quality of the data and, thus, the biological interpretation. This review describes the challenges underlying metabolomics-based experiments, discussing step by step the potential pitfalls of the analytical process, including study design, sample collection, storage, as well as preparation, chromatographic and electrophoretic separation, detection and data analysis. Moreover, it offers practical solutions and strategies to tackle these challenges, ensuring the generation of high-quality data.

Isabelle Kohler[1], Aswin Verhoeven[1], Rico JE Derks[1] & Martin Giera*[,1]

[1]Center for Proteomics & Metabolomics, Leiden University Medical Center, Albinusdreef 2, 2300RC2300 RC Leiden, The Netherlands

*Author for correspondence:
Tel.: +31 71 5269526
m.a.giera@lumc.nl

First draft submitted: 7 April 2016; Accepted for publication: 20 May 2016;
Published online: 21 June 2016

Keywords: capillary electrophoresis • clinical metabolomics • data analysis • GC • LC • lipidomics • MS • NMR

Background

Already in the 1940s, the concept that every individual has a metabolic pattern reflected by its bodily fluids has been investigated by Roger Williams and his associates, who at that time had to rely on paper chromatography [1]. Technical achievements, particularly in the field of analytical chemistry, allowed Dalgliesh *et al.* to later carry out the first GC–MS-based analysis of urine, marking the beginning of what is known today as MS-based metabolomics [2]. However, it took more than 30 years before the terms metabolome and metabolomics were coined [3]. Nowadays, metabolomics is integrated into the 'omics' cascade from genome, transcriptome and proteome to metabolome. Strictly speaking, metabolomics has been distinguished from metabonomics. Metabolomics is defined as the *'global unbiased analysis of the metabolites present within a biological system in an identified and quantified manner'* [4], while metabonomics refers to the *'quantitative measurement of the multiparametric time-related metabolic response of a complex (multicellular) system to a pathophysiological stimulus or genetic modification'* [5]. Within the field of metabolomics, it is generally accepted to further discriminate between targeted and untargeted metabolomics. Targeted approaches encompass the quantitative measurements of a selection of known metabolites involved in given biochemical pathway(s), while untargeted metabolomics is defined as the global unbiased analysis of all small molecules that constitute the metabolome [4,6–9]. Nontargeted approaches may be further divided into metabolic profiling, which focuses on specific classes of compounds, and metabolic fingerprinting,

FUTURE SCIENCE part of fsg

which represents a global approach based on the measurement of metabolic patterns ('fingerprints') [10]. Metabolomics forms an important tool for the system-wide understanding of biological processes and the discovery of novel molecular disease markers [11,12]. From a clinical perspective, metabolomics and clinical chemistry form two complementary fields, where metabolomics is ideally applied for the generation of new hypotheses and disease mechanisms, while clinical chemistry subsequently follows up such processes, enabling the accurate determination and validation of diagnostic markers. In clinical chemistry, numerous guidelines addressing good laboratory practice and accurate validation procedures for markers of interest are available [13–18]. Likewise, an increasing number of protocols and guidelines have been published over the last decade in clinical metabolomics [8,19–26]. However, the more hypothesis-driven nature of clinical metabolomics does not allow for the simple adaptation of clinical chemistry guidelines. Although there is certain overlap between the analytical challenges, the complex field of clinical metabolomics presents its own specific pitfalls and challenges. A proper knowledge of such challenges is therefore essential to generate adequate and reliable results and allows for appropriate biological interpretation. This review describes the multiple potential pitfalls encountered in untargeted and targeted clinical metabolomics linked to sample collection and preparation, compounds separation, MS and NMR detection, as well as data analysis; and discusses strategies to tackle such challenges.

Sample collection & preparation

The results expected from a metabolomics study are greatly impacted by the experimental design, since an inappropriate study design may lead to misleading statistical outcomes and, subsequently, inaccurate biological interpretations [27]. Appropriate collection, handling, storage and preparation of the samples are crucial for the generation of reliable data since sample contamination, degradation or low analyte recovery will have a significant effect on the analytical results.

Experimental design

Study design

Association (case–control) studies, which aim at finding significant associations between metabolome, clinical factors and a specific disease or health status, represent by far the majority of metabolomics experiments. In order to highlight these subtle associations with high confidence, the potential confounding factors between cases and controls including gender, age, ethnicity, BMI, diet, lifestyle (e.g., smoking habits, physical activity, etc.), health status or medications/xenobiotics should be as similar as possible [24,27]. The lipidome is particularly affected by the gender (higher plasma concentrations for triglycerides, ceramides, gangliosides and phospholipids in male), BMI, medications (e.g., statins) and the diet [28]. Some of these confounders, such as age and gender, are easily controlled but a full control and knowledge of all factors is difficult [29]. Ideally, the selected subjects should therefore be admitted to a clinic, so that environmental factors are restricted to their minimal, which remain difficult and costly in large-scale studies.

The number of subjects required for a specific study to produce meaningful outcomes with a high degree of confidence is difficult to estimate, notably in untargeted metabolomics cohorts. In principle, a power calculation can be used to estimate the number of samples required for a study [30]. However, in the majority of untargeted studies, the number of features of interest expected is *a priori* unknown [24,28]. Moreover, the biological and analytical variability of the samples within a group is usually undefined at the time of recruitment. In such cases, the number of subjects can be estimated based on previous studies, or ideally on a pilot study [28]. The inclusion of biological replicates is also recommended and preferred over analytical replicates since biological variance usually exceeds the analytical one [26].

Sampling conditions

Various biological materials have been already analyzed in metabolomics studies. The large majority of human metabolomics studies consists of the analysis of urine and/or blood-derived samples, which will be solely discussed hereafter.

Plasma and serum metabolomics involve some important considerations. First, a consistent sampling site throughout the whole study is essential since venous and artery blood have shown a different metabolic composition [31]. Furthermore, circadian oscillations can significantly influence the metabolome, particularly for lipids that can present large circadian fluctuations [28]. Finally, serum and plasma also differ in their metabolic composition, for instance in amino acids and carbohydrate levels [7,32]. The metabolic profile of plasma samples is also influenced by the anticoagulant used during sample collection, namely, heparin, citrate or EDTA. For instance, heparin may lead to ion suppression, while citrate plasma presents lower levels for some lipid classes. EDTA shows the advantage to chelate the potential divalent metal cations present in the sample that may accelerate the hydrolysis of important energy metabolites, but may also lead to ion suppression [28,31,33–35].

Urine represents an attractive matrix for metabolomics studies as its composition can be vastly affected

during diseases. Compared to blood-derived matrices, the composition of urine is influenced to a greater extent by environmental factors [9]. Moreover, urine samples present substantial concentration differences, requiring additional normalization steps during data analysis. An alternative practical solution relies on the 24-h collection of urine, which is nonetheless hardly applicable in clinical practice [9,31]. The time point of urine collection during micturition may significantly affect the urinary composition, mainly due to bacterial contamination. A mid-stream urine sample is usually favored, except when studying urinary tract infections where first-void volume is preferred [36,37]. Finally, in large-scale human metabolomics studies, samples are frequently collected at different sites. It is therefore important to provide written procedures or guidelines to all researchers, clinicians and laboratory staff involved in the project to ensure the highest reproducibility throughout sample collection and handling [24,37]. Multiple errors can indeed already occur at this stage, for example, use of improper containers or wrong storage of collected samples, particularly when home sampling is allowed in the study protocol [24,37].

Analytical workflow

In clinical metabolomics, the outcome of a biological study relies heavily on the analytical quality of the acquired data. In case of unreliable experiments, the analytical variability will contribute significantly to the total variability, hampering valid comparisons between populations. This is particularly true for untargeted approaches where semiquantitative data of largely unknown analytes are gathered.

Numerous strategies can be implemented to prevent batch effects and obtain high-quality data throughout the entire study. Since chromatographic and MS performance may decrease over time, it is recommended to divide larger studies into smaller batches and plan routine maintenance in-between, such as cleaning of ionization source components, tuning/recalibrating the mass analyzer, change of GC injector liners, etc. [7,24]. In order to ensure repeatability and reproducibility among all batches, the analytical workflow should include prior randomization of samples, repeated injections of blank samples and multiple injection of quality controls (QCs) throughout the analytical batch [7,8,24,38]. Randomization lowers the risk of bias and ensures that no correlation between metabolite levels and analysis order is introduced. The distribution of the samples in each batch should be carefully controlled to ensure that no bias has been introduced during the randomization [24]. The injection of blank samples (i.e., prepared blank solvents) prior to samples is recommended to stabilize the system and monitor the potential contaminants present. In lipids analysis using LC–MS, blank injections are also recommended within the batches to avoid possible carryover since some lipid compounds such as fatty acids tend to stick to the stationary phase [27].

QC samples typically consist of a pooled mixture obtained by mixing one aliquot of all samples, subject to the same preanalytical and analytical conditions. If the sample volume is too low, the QC can also be a commercial alternative with a representative composition [7]. QCs have multiple roles, including equilibrating the analytical platform, controlling the signal intensity prior to starting a sequence, estimating the intermediate precision between batches and allowing for postanalysis signal correction and normalization within or between batches [7,8]. For the latter, a filtering step is typically introduced during the data analysis where metabolic features with an excessive signal drift (relative standard deviation for peak areas >30% [13]) are removed from the data matrix [9,38]. QCs are typically injected at the beginning of a sequence prior to the first sample, intermittently through the run (every 5–12 injections), and at the end of the sequence [24]. Figure 1 illustrates how regular QC injections can be used to evaluate the analytical consistency of the data, using principal component analysis (PCA). In Figure 1A, the QC samples show a relevant time-trend, rendering the obtained data highly unreliable. In Figure 1B, the QCs are tightly clustered together, showing that the biological variability is greater than the analytical variation. In addition to multiple injections of QCs, the instrumental reproducibility can also be monitored by employing internal standards (IS), which should be selected according to the expected compound classes [8].

Sample storage

Sample storage conditions strongly impact the qualitative and quantitative metabolome due to chemical and enzymatic degradation. Numerous classes of compounds may be affected by inadequate handling and storage of collected samples, notably the energy metabolites [33], fatty acids undergoing autoxidative processes [39], amino acids [40] or biomarkers such as malondialdehyde [41]. Since a relatively long time may elapse between the collection of the samples and their analysis, multiple parameters must be monitored to lower the risks for degradation and interconversion, including temperature, light, humidity, time span, quenching and number of freeze–thaw cycles.

The quenching step, in other words, the rapid inhibition of metabolic activity, is essential to avoid any residual enzymatic activity. This is particularly relevant in blood-derived samples that, compared

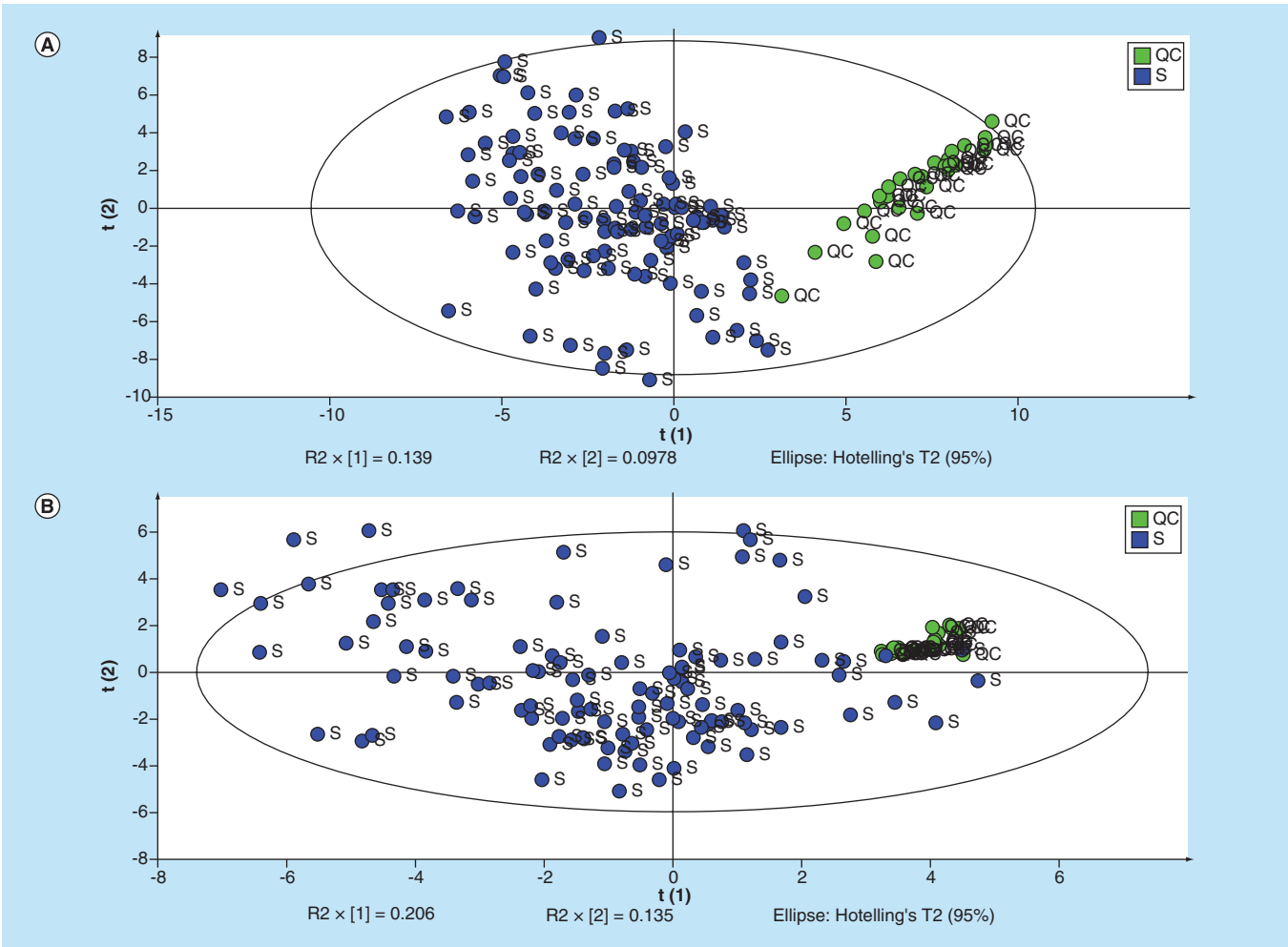


Figure 1. Scores plots obtained with principal component analysis on a set of samples (labeled S, blue circles) and quality controls (labeled QC, green circles). A tight clustering of all QCs in the principal component analysis model shows a good repeatability and analytical consistency of the instrument. In **(A)**, the QCs are not clustered and show a significant time-trend, leading to low-quality data. In **(B)**, all QCs are tightly clustered, illustrating a good analytical consistency. The scores plot in **(B)** shows that the possible separations observed between the samples are not due to instrumental variations but due to the samples themselves. Data obtained on the analysis of human urine samples with LC–MS (a Dionex UltiMate 3000 RSLC system hyphenated to a Bruker MaXis Impact UHR-qTOF mass analyzer).

with urine, contain relatively high concentrations of enzymes active even after the collection step [24]. Quenching should be fast enough since some energy metabolites such as adenosine triphosphate and glucose-6-phosphate show a metabolic activity of <1 s [35], while catecholamines have a half-life of few minutes in plasma [27]. The most popular method is the addition of ice-cold 60–100% methanol (MeOH) with or without addition of pH modifiers prior to snap-freezing in liquid nitrogen [35].

After collection, samples should be stored as fast as possible at -80°C or in liquid nitrogen. Many studies have shown that sample storage at -80°C for <6 months does not affect the large majority of metabolites, including lipids. However, this is not the case for instable compounds such as succinate, acetate, urea, lactate and glutamine that showed modified levels in urine samples stored for 1 week at -80°C [42]. In addition, it has been shown that for a large majority of metabolites in urine samples, the long-term storage at either -20°C or -80°C led to comparable metabolic profiles [43,44]. It is worth mentioning that not only the analyte concentrations might change but also modifications in matrix components might affect the analytical procedures [45].

Compared to blood-derived samples, urine is prone to bacterial contamination during or after sample collection that can lead to modification of the metabolome upon storage, especially nonfrozen. Preservatives such as 0.1% sodium azide can therefore be added to the samples, preventing bacterial growth and compound degradation, notably if samples have to be temporarily stored at 4°C [37,44,46,47]. Although numerous studies

aimed at evaluating the effect of freeze–thaw cycles on the metabolome, a consensus is still lacking. Nevertheless, multiple freeze–thaw cycles should be avoided as much as possible, for example, via preparation of sample aliquots after collection [28,48].

Sample preparation

The selection of the optimal sample preparation procedure relies on multiple factors, including the matrix analyzed, the physicochemical properties of the compounds of interest, the volume of sample and the platform(s) used for analysis. It is a critical step due to the potential analytical errors that can be introduced [49]. Most of these errors are operator-dependent, for example, low experimental repeatability or contamination of samples, and can therefore be significantly reduced with automation. Large-scale studies represent an additional challenge due to the number of samples that need to be pretreated prior to analysis. In this context, two strategies can be considered, namely, prepare all the samples in 1 day, divide them in batches and freeze them, and analyze each batch every day; or divide the samples in several batches and prepare them every day prior to analysis [8].

Untargeted metabolomics

In untargeted approaches, the sample preparation step is kept as simple as possible to obtain the widest metabolite coverage. Nonselective sample pretreatments such as protein precipitation (PP) for plasma/serum and simple dilution for urine samples (referred to as dilute-and-shoot) are generally used [7,50].

Organic solvent-based PP allows for the extraction of both polar and hydrophobic compounds. For human plasma, acetonitrile (MeCN) and acetone appear more efficient in terms of protein removal while MeOH, ethanol and methanolic mixtures result in a wider metabolite coverage [51]. On the other hand, MeOH and acetone give a better protein removal in serum samples compared with pure MeCN, while MeOH leads to better isolation of polar metabolites [52]. Compared to blood-derived samples, which show a high protein content (60–80 g/l), urine has a significantly lower protein concentration (0.5–1 g/l). Therefore, urine samples are usually simply filtrated and/or centrifuged after thawing, and diluted (typically five- to tenfold) with water or buffer. The filtration of urine samples on 0.20-µm cellulose filters removes materials in suspension and prevents bacterial growth during storage [53,54]. Dilution with buffer should be preferred over pure water since the pH of urine may vary extensively (i.e., pH 3–7.5). This is particularly important in capillary electrophoresis (CE) where the separation performance and repeatability is significantly affected by the salt composition and ionic strength of the matrix [55]. Hydrophilic interaction chromatography (HILIC) requires the dilution of urine with an aqueous organic solution to ensure acceptable separation performance since the presence of large amounts of water in the injected sample will lead to severe peak distortion [54]. Diluting the samples four- to five-fold with a mixture of MeCN/water 3:1 (*v/v*) has shown to give sufficient dilution of proteins and salts while keeping acceptable peak shapes [54]. In case of proteinuria, which is a rather common clinical abnormality, an additional centrifugation step should be carried out prior to injection of the supernatant, especially if the samples have not been previously filtered.

In lipidomics, nonselective plasma/serum pretreatments include liquid extraction with a mixture of chloroform/MeOH 2:1 (*v/v*) or chloroform/MeOH 1:2 (*v/v*) with subsequent addition of chloroform and water [7]. A methodology employing methyl-*tert*-butylether as the extraction phase was developed by Matyash *et al.*, allowing for a straightforward collection of the upper organic phase [7,28,56]. The use of isopropanol as precipitating agent may also be considered, enabling a broad coverage of plasma lipids with a higher recovery compared with other PP procedures [57].

Targeted metabolomics

In targeted approaches, the physicochemical properties of the compounds of interest are *a priori* known, which allows for a more selective sample preparation enabling a sufficient sample cleanup as well as preconcentration, typically using SPE or liquid–liquid extraction (LLE).

For blood-derived samples, a PP step is generally first carried out to remove proteins and avoid the formation of a foaming emulsion in LLE or clogging of the SPE support. Urine samples are diluted with buffer to normalize and reach adequate pH prior to extraction. LLE shows the advantage of separating both polar and hydrophobic analytes into two fractions that can subsequently be analyzed independently. However, very polar and/or ionized compounds show a poor extraction efficiency and LLE remains predominantly used in lipid analysis. Furthermore, the LLE procedure is difficult to automate, suffers from a low sample throughput and requires high amounts of solvents, which are clear limitations in large-scale metabolomics studies. Supported liquid extraction, where samples are adsorbed on a porous solid support material (diatomaceous earth support), allows for partial automation of the extraction process with a lower solvent consumption. Commercial 96-well plates are nowadays available that may be promising in metabolomics analysis [50].

Due to its ease of automation and the possibility of online combination, SPE is more widely used. The

large variability of stationary phase chemistries commercially available in many different formats (e.g., SPE cartridges, 96-well plates) enables for a tailored sample preparation. For instance, hydrophobic sorbents are particularly well suited for sample cleanup (desalting and delipidation) and enrichment of analytes with low polarity, while mixed-mode sorbents containing cationic or anionic moieties lead to excellent recoveries and preconcentration for polar and/or ionized compounds.

Besides analytical errors, the sample preparation step also increases the potential risk for compounds degradation. During the entire procedure, analytes are subject to contact with air and light, and possibly with elevated temperature [53]. In order to compensate for all analytical errors and ensure accurate data, targeted approaches require the addition of IS. Isotopically labeled IS, especially ^{13}C-labeled, are recommended [58]. Nevertheless, in most of the studies, labeled IS cannot be used for every compound due to high costs or because they are not commercially available. In such cases, the labeled standards should be carefully selected to correct not only for analytical errors during the sample preparation step but also for potential MS ion suppression/enhancement [53].

Chromatographic & electrophoretic techniques

A separation step is generally performed prior to MS detection using chromatographic (i.e., GC and LC) or electrophoretic techniques (i.e., CE). All techniques present specific benefits and limitations discussed hereafter.

GC

GC has been the first separation technique used in metabolomics and remains very popular. Separation in GC is achieved by fractionally vaporizing the analytes using temperature gradients while guiding an inert carrier gas (i.e., helium, nitrogen or hydrogen) over the stationary phase. Hence, separation is mainly driven by differences in the evaporation temperature. This process is obviously linked to an intrinsic evaporability of the analytes, which may lead to degradation of thermolabile components. Nevertheless, GC shows advances particularly in terms of separation efficiency. This superior separation efficiency is explained by the absence of eddy-diffusion during the chromatographic separation compared with LC [59].

However, GC presents three main pitfalls: first, the possible loss of thermolabile analytes, as previously mentioned; second, a rather cumbersome sample preparation step, particularly due to the need of analyte derivatization and third, a generally higher variability compared with LC-based metabolomics [19].

Thermal degradation

Fang *et al.* recently published a controversial study where the authors argued that elevated temperatures such as the ones during GC analysis severely altered the analyzed metabolome [60]. In this study, LC–MS was used to compare heated and nonheated plasma samples, thereby trying to mimic GC conditions. According to the authors, up to 40% of the heated compounds were destroyed, modified or degraded. This study caused much discussion, as it questioned the long-standing and prominent role of GC–MS, particularly in exploratory metabolomics [61]. Although the claims of the study might be rather far-fetched, it underlined the crucial point of analyte degradation during GC–MS analysis. However, not only thermal exposure of analytes can jeopardize the successful implementation of GC–MS analysis in clinical metabolomics but also other practical factors such as the use of different carrier gases, differences in the applied inlet liners (i.e., differential liner design), different injection techniques (i.e., hot needle versus cold needle injection) or the presence of impurities in the liner or on the column head. Some of the aforementioned points have been previously investigated [62,63]. Practical recommendations for ensuring reliable GC–MS-based metabolomics data include the use of a consistent liner design (i.e., tapered, goose neck or double goose neck), the constant use of deactivated inlet liners as well as glass wool, a regular maintenance schedule, injection of system suitability tests and/or QCs, regular change of GC columns and the use of the highest quality derivatization reagents and solvents [35,63,64].

Sample preparation

GC-based metabolomics analysis usually involves a two-step sample preparation procedure where keto functionalities are first derivatized using pyridine methoxylamine followed by silylation with either *N*-methyl-*N*-trimethylsilyl-trifluoroacetamide (MSTFA) or *N*-methyl-*N*-dimethyl-*tert*-butyl-bistrifluoroacetamide (MtBSTFA). It is well known that MtBSTFA forming *tert*-butyldimethylsilyl derivatives give rise to more intense and information-rich fragment ions especially when using electron ionization (EI) [65,66]. Unfortunately, the derivatization efficiency of MtBSTFA is rather poor, particularly for secondary and tertiary alcohols, likely due to sterical hindrance potentially leading to partial derivatization [67,68]. On the other hand, the stability of trimethylsilylesters obtained by the derivatization of carboxylic acids with MSTFA is several orders of magnitude lower than for their *tert*-butyldimethylsilyl counter parts, fostering the use of MtBSTFA.

A tight control and, if possible, automation of the derivatization procedures is recommended as some

derivatives (e.g., trimethylsilyl esters) are well known to rather rapidly undergo degradation over time, necessitating equal post-derivatization periods [40]. A possible alternative to silylation reagents may be the use of chloroformate or methylation reagents [59,67]. Some of the aforementioned pitfalls were addressed by Villas Boas *et al.*, who investigated compound methylation using methylchloroformate as an alternative derivatization strategy [69]. The authors concluded that methylation is superior to trimethylsilylation; however, further detailed studies are required to evaluate the use of alternative derivatization protocols.

LC

Compared to GC, LC allows for the direct analysis of rather polar and thermolabile compounds without the need for derivatization and adds additional selectivity through interactions between stationary and mobile phases. Among all the chromatographic modes that can be used in metabolomics and lipidomics, reversed-phase LC (RPLC) remains the gold standard while an increasing amount of metabolomics studies are performed using HILIC [7,35,38,54,70]. Over the last decade, RPLC–MS has become very competitive in multiple fields along with the development of innovative technologies such as the use of stationary phases equipped with superficially porous particles (core–shell technology) or sub 2-µm fully porous particles (ultrahigh pressure LC, UHPLC) [71–73]. Both approaches provide fast and high-resolution separations (increased peak capacity), which allows for high-throughput analysis of closely related metabolites.

Reversed phase LC

The large majority of RPLC metabolomics studies involve the combination of an octadecyl-modified stationary phase (C18) with an aqueous-organic mobile phase composed of either MeOH or MeCN containing 0.1% formic acid. This simple and versatile combination leads to acceptable peak shapes for most metabolites, good ionization efficiency in the positive ESI mode and excellent retention time repeatability. In metabolomics, RPLC shows two major drawbacks, namely, a possible ion suppression particularly caused by phospholipids in plasma/serum analysis (matrix effects) and a poor retention of highly polar components, especially in urine analysis [74]. Moreover, the use of formic acid in the mobile phase may lead to a relatively poor sensitivity in negative ESI mode. Finally, compared with GC–MS, RPLC–MS suffers from a lack of commercially available spectral libraries that makes compounds identification rather complex in untargeted approaches. In-house-built libraries containing both retention and spectral information thus offer a valuable tool for metabolite identification.

Hydrophilic interaction chromatography

In HILIC chromatographic mode, a polar stationary phase (i.e., bare silica or silica modified by functional groups such as diol, amine, amide or zwitterionic moieties) is used with a relatively hydrophobic mobile phase composed of an aqueous-organic mixture (usually MeCN with volatile buffers), forming a water-enriched layer of partially immobilized eluent on the stationary phase. The retention mechanisms are rather complex and involve hydrophilic partitioning, dipole–dipole interaction, hydrogen bonds and electrostatic interaction [75–77]. HILIC shows several advantages that render this technique particularly interesting in metabolomics, such as an enhanced MS signal due to a better eluent desolvation, the possibility of directly injecting organic extracts usually obtained by PP or extraction, and an orthogonal selectivity compared with RPLC [78]. However, HILIC remains rarely used in this field. This reluctance is likely explained by the complex mechanisms that render HILIC less flexible and not as straightforward as RPLC. The major drawbacks of this technique rely on the higher variability in the retention times, the longer equilibration times and the poorer peak shapes observed [54]. However, high-quality data can be ensured provided that good practice is followed during the analytical workflow, including a careful selection of the sample diluent and volume of injection, the use of volatile salt buffers at a fixed and repeatable pH value and ionic strength, the investigation of the matrix effects and the adequate data preprocessing parameters in untargeted approaches. These aspects have been recently discussed by Kohler *et al.* in a practical review [54].

Capillary electrophoresis

CE is a miniaturized technique where charged compounds are separated according to their charge-to-size ratio in an electrically conductive liquid phase (background electrolyte) under the influence of an electric field [48]. The hyphenation of CE with MS provides interesting advantages for metabolomics applications including a high selectivity and separation efficiency, especially for polar and/or charged compounds which show a poor retention in RPLC; an orthogonal mechanism of separation compared with chromatographic techniques, leading to complementary metabolic coverage; the separation of chiral compounds; its applicability to volume-limited samples and the possibility for single-cell metabolomics [79–83].

However, compared with other MS-based techniques, CE–MS is little used in metabolomics; it suf-

fers from a lower overall sensitivity and a poor reproducibility in migration times due to adsorption of matrix material on the capillary surface. Moreover, the commercial systems are not equipped with autosamplers adapted for high-throughput analysis and feature a lower injection repeatability compared with LC systems. The latter drawbacks foster the systematic use of IS to correct the injection variability. Finally, CE–MS coupling is not as straightforward as LC–MS or GC–MS and requires additional practical expertise to obtain a stable MS signal and reliable data.

The limitations differ depending on the interface used for CE–MS hyphenation. There are currently two different interfaces commercially available, namely, the sheath-liquid and the porous tip sheathless interface [84]. The sheath-liquid interface is typically considered more stable, repeatable and cost effective but suffers from a lower sensitivity due to dilution of the CE effluent with the sheath liquid. On the other hand, the commercial sheathless interface provides a higher sensitivity and separation efficiency but at higher costs, and with less robustness and flexibility [85,86]. Today, the optimal CE–MS setup for metabolomics approaches relies on a combination of sheath-liquid interface, online preconcentration techniques, systematic use of IS, regular injection of QCs, use of electrophoretic mobility instead of migration times, adequate postanalysis alignment of the data and sufficient practical expertise. The limitations of the technique certainly explain the reluctance in using it in metabolomics. It is only with further developments in CE instrumentation, notably with a stable and sensitive CE–MS interface, that this technique might be fully exploited as a complementary tool in addition to chromatographic techniques.

MS

MS has been largely used in metabolomics and remains the gold standard detection technique after chromatographic or electrophoretic separation. It is only due to the high-information content of MS data that identification and quantification of hundreds to thousands of analytes in a single analytical run can be achieved. The major ionization techniques used in MS-based metabolomics are EI, chemical ionization (CI), ESI and atmospheric pressure chemical ionization (APCI). While EI and CI are typically used in combination with GC, APCI can be used in combination with GC, LC and CE; and ESI with LC and CE.

Ionization techniques

Historically, EI was the first ionization technique employed for MS experiments. In EI, electrons emitted by a heated tungsten filament are accelerated. These electrons get into close proximity with the neutral molecules eluting from typically a GC column, leading to the formation of radical ions and charged fragments. EI is generally described as a 'hard' ionization technique; this can be very useful analyzing neutral molecules, although certain analytes such as esters or ketocarboxylic acids can undergo severe fragmentation. This may render the obtained spectral data very difficult to interpret when only low molecular-weight fragments are observed. Nevertheless, the major advantage of EI is its highly standardized and repeatable performance. Over the years, this has led to extensive spectral libraries containing more than 250,000 compounds, which make GC–EI–MS a very attractive technology for identifying unknown metabolites of interest [63]. Traditionally, a kinetic energy of 70 eV is used for EI, allowing for high spectral comparison. However, many modern GC–EI–MS instruments offer the possibility to modify this value, thereby changing the energy content of the emitted electrons. Although there have been some reports showing that reducing the electron energy causes less fragmentation, this usually also severely decreases the overall sensitivity.

CI is considered a 'soft' ionization technique where a reactant gas (frequently methane) is used in combination with an EI source, causing formation of primary ions stemming from the reagent gas. The formed reagent gas ions subsequently ionize the molecules of interest, mainly by collisional processes [87]. Compared to EI, CI therefore frequently allows the determination of the molecular ion. CI in negative mode has, for example, been widely used for the analysis of fatty acids [88], while its use in metabolomics studies has remained rather limited. This may be explained by the lack of spectral libraries since most of them are built with GC–EI–MS data.

ESI is certainly the most widely applied ionization technique for LC–MS-based metabolomics [89]. Although the detailed mechanisms of the ESI process are not fully understood yet, it is commonly accepted that initially an electrospray dispersion of a liquid in a nozzle generates charged droplets when forming the so-called Taylor cone. When charged droplets are generated, the solvent starts to evaporate, leading to an increase in charge density of the droplet. At one point, the so-called Rayleigh limit will be exceeded leading to formation of secondary droplets in the form of a Taylor cone from the primary droplets. This process is repeated leading to the formation of very small droplets, being the primary source of ions detected in an MS. Advantages of ESI are a straightforward hyphenation to LC and CE, its high ionization efficiency, virtually no restriction in mass range and the direct ionization of liquids. However, ESI may be prone to significant matrix effects, particularly

jeopardizing quantitative bioanalysis. Multiple studies have already addressed the strategies to ensure accurate quantitation, making ESI the most important ionization technique in modern bioanalysis despite its limitations [58,90]. In clinical metabolomics, matrix effects are not only jeopardizing the study outcome but they can also lead to misinterpretations, as shown in Figure 2. In their study, Giera and colleagues investigated the lipid profile of osteo- and rheumatoid arthritis patients [91]. Based on PCA and partial least-squares discriminant analysis (PLS-DA) of the LC–ESI–MS data acquired in negative ESI mode (Figure 2A), the clustering between synovial fluid samples was initially hypothesized to possibly be linked to the sphingomyleoid SM(d18:2/16:0), which was ranked as the third most important variable responsible for samples classification. However, after selective hydrolysis of phosphatidylcholines, no difference could statistically be observed between the samples as emphasized by the PCA scores plot shown in Figure 2B. This unexpected outcome was explained by the presence of oxidized phospholipids which overlaid SM(d18:2/16:0). Hence, the coelution of oxidized phospholipids with SM(d18:2/16:0) led to ion suppression, particularly in the negative ESI mode. The strong ion suppression of SM(d18:2/16:0) severely influenced the statistical model, causing the model to be based on the ion suppression effect instead of the compound itself.

APCI is a 'soft' ionization technique usually coupled to LC for the ionization of less polar and neutral compounds but it has also been hyphenated to both CE and GC [92,93]. In APCI, the liquid or gaseous eluent flows through a heated nebulizer; the obtained mixture of hot liquid and vapors expands into the atmospheric pressure interface where it is ionized by a corona discharge, leading to a proton transfer from the solvent molecules to the analytes. In contrast with ESI, which relies on a liquid-phase ionization, APCI ionization occurs in the gas phase. GC–APCI–MS has been scarcely reported in metabolomics studies even though it shows less fragmentation than EI and, in some cases, is more sensitive than EI and CI [94]. The main limitation of GC–APCI–MS in metabolomics applications, shared by LC–APCI–MS as well as GC–CI–MS, is the lack of spectral libraries for compound identification. A web-based library for GC–APCI–MS containing retention indices and MS(/MS) spectra for a range of endogenous and exogenous compounds (>150 compounds) has been presented by Pacchiarotta *et al.* [92]. However, as ionization using APCI is more condition-dependent (e.g., humidity, temperature, flow conditions, source geometry, cone voltages, etc.) than EI, comparison of spectra between different instruments might be limited [94].

Mass analyzers

Depending on the purpose of the study, a wide range of mass analyzers are used in clinical metabolomics, notably high-resolution MS such as quadrupole time-of-flight (qToF) or Orbitrap analyzers, as well as quadrupoles, ion traps and other hybrid instruments. Each mass analyzer shows advantages and limitations; therefore, many studies involve the use of more than one system to gather complementary information [35].

In targeted studies, mainly triple quadrupole or quadrupole linear ion trap mass analyzers are used due to their high sensitivity and selectivity, particularly when used in selective reaction monitoring (SRM) mode. Although the combination of LC with highly selective SRM transitions delivers a high degree of selectivity, it remains insufficient for discriminating geometrical isomers that show identical tandem mass spectra (see the 'Stereoisomers & chirality' section).

In untargeted metabolomics, qToF systems are largely employed as they provide high mass resolution and accuracy, as well as high scan speed. Notably, the obtained high-resolution spectral data are beneficial for compound identification, facilitating database searches and generation of elemental composition. However, based on recent studies, the unambiguous identification of compounds of interest should not solely rely on high-resolution MS, but should always include tandem mass spectra measurement, and ideally comparison of retention times and MS(/MS) spectra with authentic synthetic standard material [26].

Stereoisomers & chirality

Stereoisomers (including geometrical isomers), diastereomers, as well as enantiomers play crucial biological roles, rendering their separation and correct assignment highly important in clinical metabolomics. As an example, the diastereomers and geometrical isomers 5S,12S-diHETE and LTB_4 do not only arise out from different pathways but also show a large difference in their respective biological activity [95,96]. As such substances show identical tandem mass spectra and, in many cases, similar retention behavior, additional dimensions of separation are needed, for instance differential mobility spectrometry (DMS), as emphasized in Figure 3 [97]. DMS has recently shown to facilitate the separation of lipids as well as isomers from the leukotriene and protectin classes [96,98–100]. Another relevant example concerns D- and L-lactic acids, of which only D-lactic acid appears to be linked to the short bowel syndrome [101]. Therefore, it is of high importance for understanding the biological relevance of stereochemical isomers to ensure their separation and correct assignment. Both points have recently been underlined by Struys in the context of

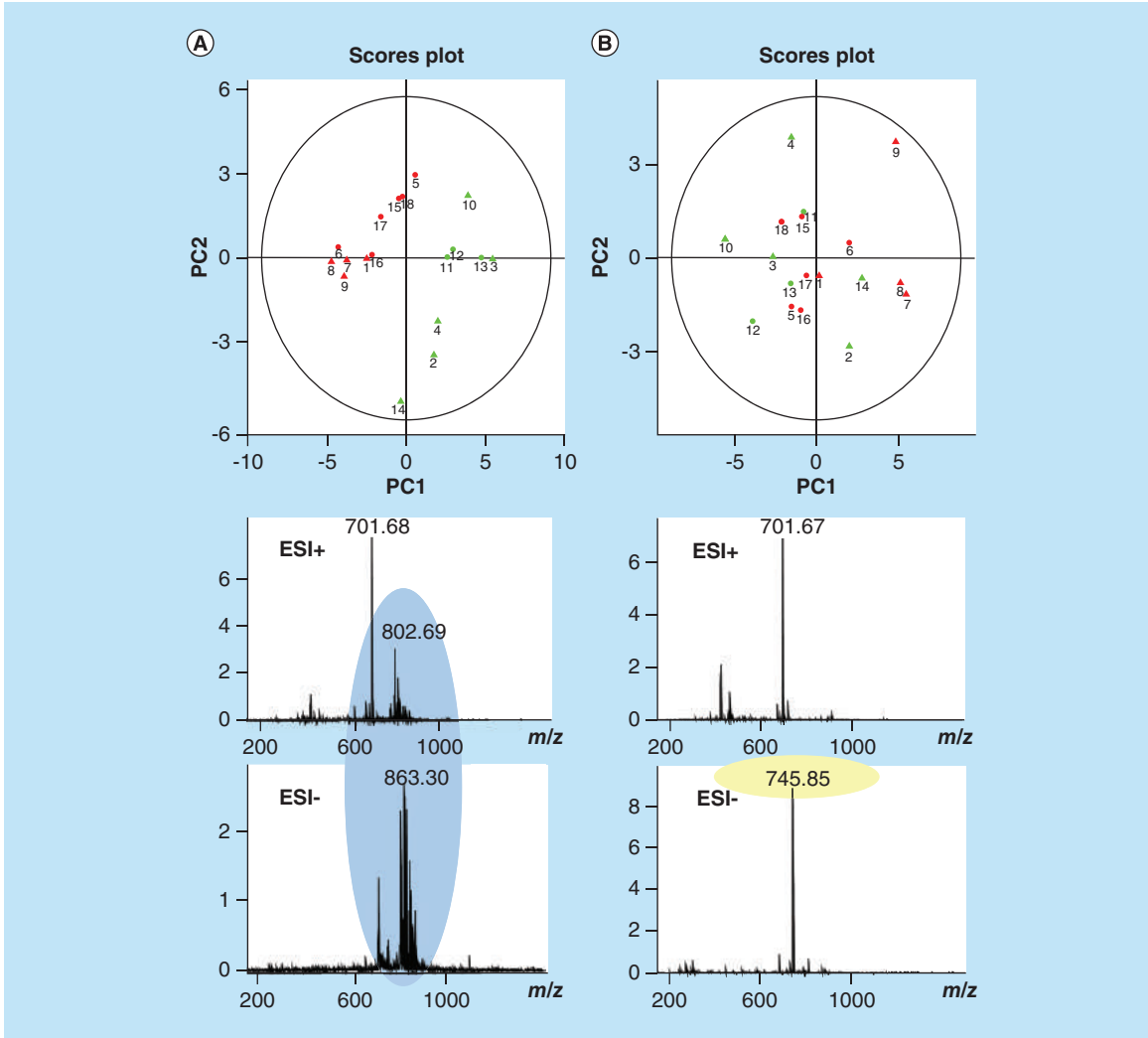


Figure 2. Illustration of the matrix effects in a clinical metabolomics study. The scores plots were built via principal component analysis of synovial fluid samples on LC–ESI–MS data in negative ESI mode. **(A)** Before hydrolysis of phospholipids. **(B)** After selective hydrolysis of phospholipids. The mass spectra correspond to the retention time of SM(d18:2/16:0) and were recorded in positive ESI mode (ESI+) and negative ESI mode (ESI-). For experimental conditions, see [91].

the analysis of D- and L-2-hydroxyglutarate where only D-2-hydroxyglutarate plays a crucial role in cancer [102]. While diastereomers and geometrical isomers usually present different physicochemical properties, which allow for their separation by chromatography, this is not the case for enantiomers. Enantiomer separation requires the use of chiral stationary phases or the addition of a chiral selector in the CE background electrolyte [103]. Although modern chiral phases allow the separation of various enantiomers, the choice of the optimal column chemistry or chiral selector remains an empirical process and the retention/migration behavior of enantiomers cannot be predicted [104]. Even if some reports have been recently published about chiral metabolomics (including NMR-based analysis), much work still needs to be done. However, chiral gas-phase separations, NMR shift reagents, chiral CE and derivatization reagents are certainly a way forward [105].

NMR spectroscopy

NMR is complementary to MS in clinical metabolomics. In NMR, the interaction of the magnetic dipole of atomic nuclei with an applied magnetic field is used to probe the chemical environment of those nuclei. This results in an isotope-specific NMR spectrum, where the position of the peaks in the spectrum is characteristic of the chemical group containing the nuclei. Notably, the peak area is linearly related to the number of nuclei in the sample, and thus the metabolite concentration. NMR is a nondestructive technique and requires a relatively simple sample

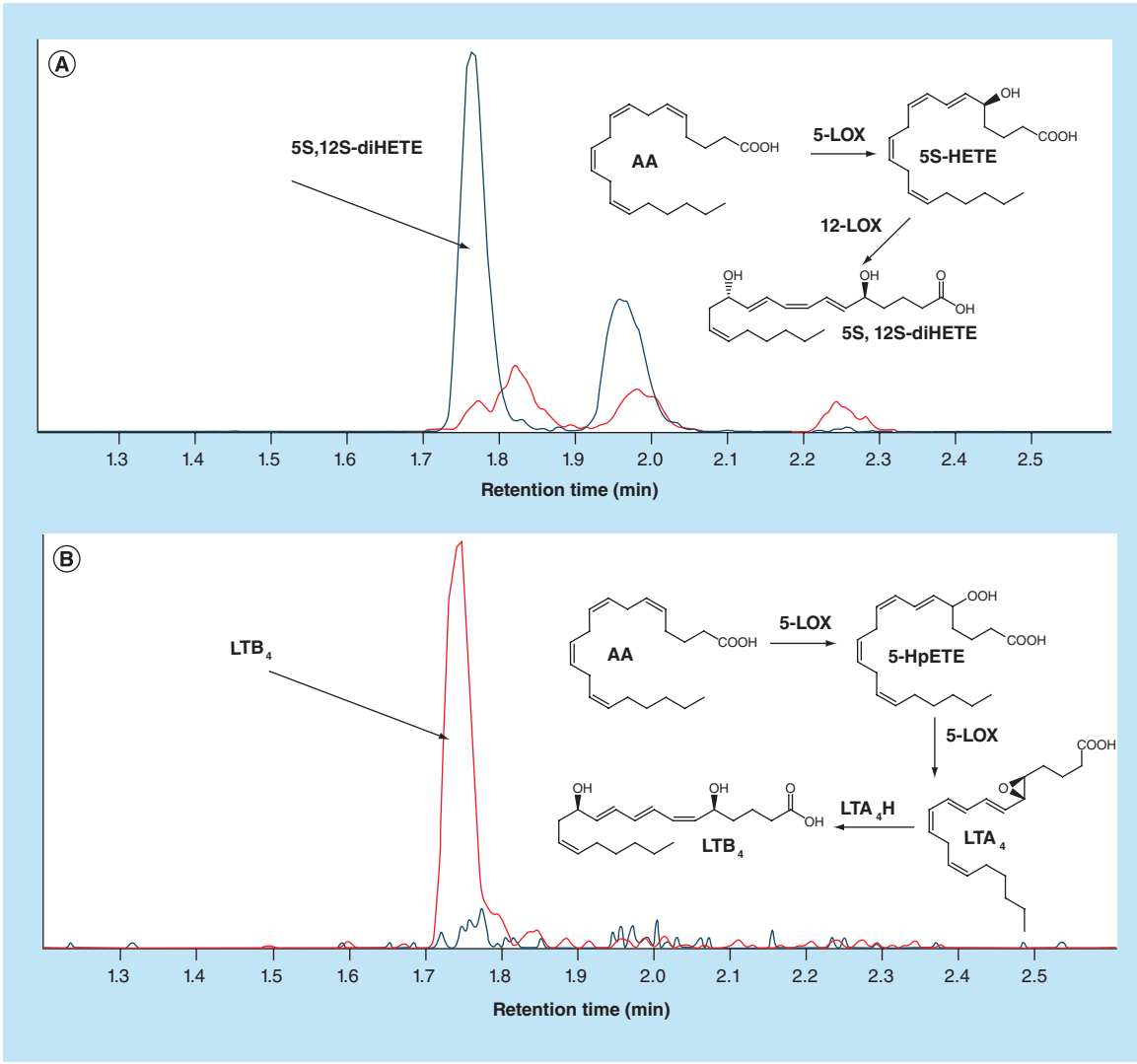


Figure 3. Differential mobility spectrometry-based separation of 5S,12S-diHETE (A) and LTB_4 (B). For experimental conditions, see [96].
Reproduced with permission from [96].

preparation. Although NMR has been successfully hyphenated to LC [106,107], clinical metabolomics studies usually involve direct NMR measurements. Considering the low dispersion of chemical shifts as compared with the peak width, this approach leads to relatively crowded but very reproducible spectra.

A common limitation of NMR is its relatively low sensitivity compared with MS-based techniques. The magnetic polarization of the nuclei is given by the population difference of the nuclear energy levels that are determined by the orientation of the nuclear magnetic dipole. The energy difference of these levels is small, leading to a tiny population difference of these energy levels at room temperature. This results in a relatively low sensitivity even for the isotopes with the best properties, in other words, the spin-½ nuclei with large magnetogyric ratios (^{1}H and ^{19}F). The sensitivity can be improved by using very strong superconducting magnets and cryoprobes that operate the receiver circuitry at low temperatures to reduce electronic noise. Typically, multiple scans are accumulated to further improve the S/N ratio. After relaxation, allowing the spin system for returning to thermal equilibrium (ca. few seconds), scans can be repeated to a certain extent (S/N ratio ~ $n^{½}$) to increase the sensitivity. However, metabolites with concentrations lower than 1 µM will usually be below the detection limit.

The number of metabolites visible in the proton (^{1}H)-NMR spectrum ranges from about 50 in serum/plasma samples [108] to roughly 200 in urine [109]. Figure 4 shows an example of a plasma, serum and urine ^{1}H-NMR spectrum. The number of peaks that a metabolite generates is determined by the number of

nonequivalent protons in the molecule. Additionally, the magnetic moment of the nuclei is transferred by the bonding electrons to neighboring nuclei by the J coupling. This coupling splits each peak into doublets, triplets, quadruplets or other multiplets of varying complexity. In a commonly used NMR system with a 14 Tesla magnet, the proton spectrum is 6 kHz wide. With an average peak width of 1 Hz, J couplings of 15 Hz and multiple peaks for each metabolite, peak overlap is common.

Whatever the matrix, water is by far the molecule with the highest concentration. In turn, the protons of water dominate the NMR spectrum unless special care is taken to suppress the water signals, such as using a combination of water presaturation and pulsed field gradients. A low irradiation field strength is used to generate the presaturation for the water signal; however, the metabolite peaks near the water signal are also affected and reduced in intensity. Moreover, the peaks from protons that exchange with water can be lower than expected as a result of the water presaturation.

In case of urine analysis, the sample pH can affect the chemical shift of certain peaks, particularly those close to nitrogen with the exception of quaternary ammonium salts. The addition of phosphate buffer during the sample preparation stabilizes the pH but because the final mixture is a compromise between metabolite concentration and pH stability, some peaks will still show significant shifts. Alignment algorithms exist to compensate the peak shifts due to these effects [110–113] but compared with chromatogram/electropherogram alignment, other nearby peaks do not necessarily shift to the same extent or even the same direction. The salt concentration affects the inductance of the radio frequency (RF) coil, influencing the RF strengths of the pulses and the signal-to-noise of the recorded signal. Automatic tuning and matching of the RF circuitry and automatic pulse calibration before every measurement are essential, but the excitation bandwidth and free induction decay (FID) reception performance of the probe will still be different from sample to sample depending on the salt concentration.

Because of the difficulty of deconvoluting the peaks to quantify individual metabolites, NMR metabolomics studies are often performed in a nontargeted and nonquantitative fashion. Therefore, even if some signals cannot be reliably deconvoluted or assigned to a metabolite, they can still contribute to the diagnostic models. An NMR spectrum routinely consists of several ten thousand points to accurately describe the shapes of the various peaks. In order to limit the number of variables, a so-called binning procedure is applied, which divides the whole spectrum in specific intervals prior to integra-

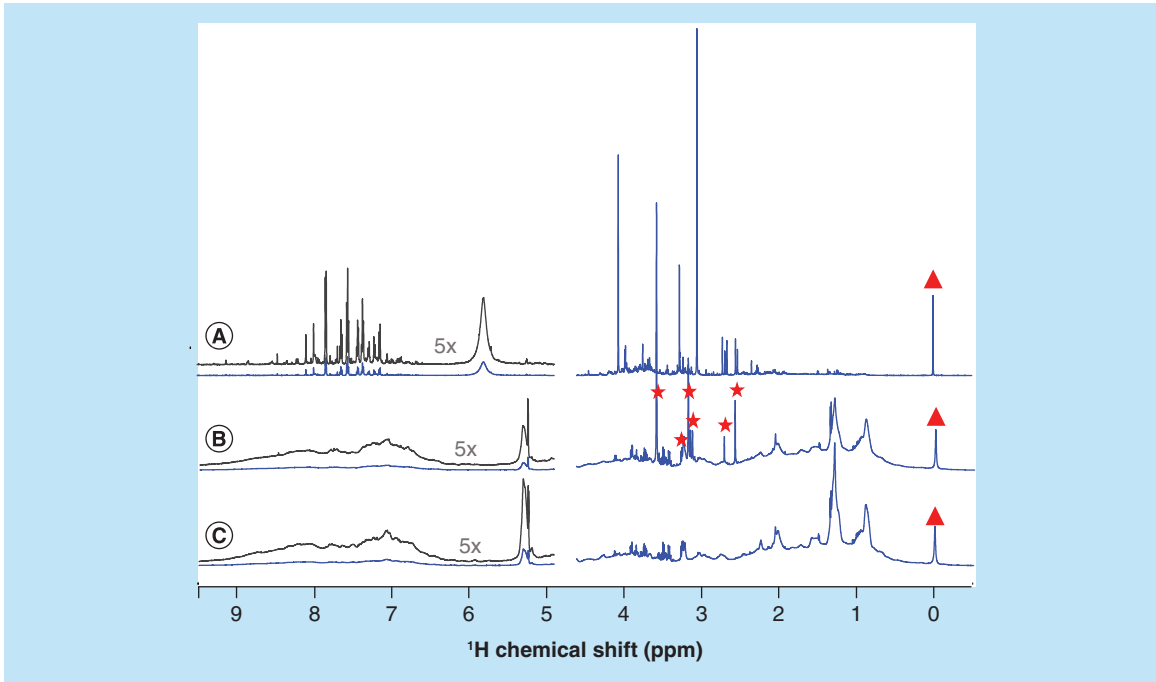


Figure 4. NOESY 1D NMR spectra of urine (A), EDTA plasma (B) and human serum (C), recorded on a spectrometer equipped with a 14.1 T magnet. The area around the water peak is not displayed. The urine spectrum was referenced with respect to the resonance of the TSP reference material at 0 ppm, while the serum and plasma spectra were referenced to the anomeric glucose peak at 5.23 ppm. The non-metabolite resonances of TSP and EDTA, both free and as complex, are indicated with triangles and stars, respectively. A 5x-zoom is shown for the low-field regions (light gray).

tion. There are different strategies for setting the location of the interval edges. The simplest approach is to use equidistant binning, where consecutive patches of spectrum of equal size define the bins, but shows the risk that a bin edge falls on top of a peak. A better approach is adaptive binning, in which an algorithm tries to find the optimal positions of the bin edges, ensuring that most peaks entirely fall in one bin [114–116].

Binning methods include information from unidentified and unresolved metabolites, but this does not ensure the generation of better models in metabolomics. Indeed, in many cohorts, participants may be on different medication regimes. In the case of indiscriminate inclusion of bins, all the exogenous metabolites will be also included in the model. In these situations, the quantification of only those metabolites that can be reliably assigned and deconvoluted should be performed. Several software tools have been developed to aid with the fitting and deconvolution of NMR peaks, both freely available (e.g., BATMAN [117]) or commercial (AMIX by Bruker Corporation or Chenomx by Chenomx Inc.).

Peak identification requires some ingenuity in NMR metabolomics. The functional group to which the proton belongs roughly determines the area of the spectrum where the peak appears. Databases exist that list the chemical shifts of the peaks of most metabolites in pure solutions, notably the Human Metabolome Database [118], the Biological Magnetic Resonance Data Bank [119], the Madison Metabolomics Consortium Database [120] and databases that are supplied with the commercial software packages Chenomx and AMIX. However, since the solvent, the pH and the salt concentration affect the exact chemical shift, the position of a peak in a biofluid can be different from those in standard solutions, rendering the identification challenging. The multiplicity of the peak may give additional information but increases the likelihood of peak overlap. Hence, in addition to a standard 1D spectrum, a 2D J-resolved spectrum is usually recorded for every sample that separates the effects of the chemical shift interaction and the J-coupling into two different spectral dimensions, significantly reducing peak overlap [121]. Nevertheless, several peaks with the same multiplicity often appear near each other. For example, the branched chain amino acids all produce doublets around 1 ppm. NMR offers a wide range of different pulse sequences that can help identifying metabolites by showing the connectivity of a proton to both nearby and remote protons, or to neighboring carbon atoms. These more complex pulse sequences typically yield multidimensional spectra from which a connectivity network can be extracted that is often unique for a specific metabolite. Databases include information on 2D spectra, and software tools exist to aid with the assignment of the 2D crosspeaks [122,123]. These 2D spectra also take longer to record and are, therefore, usually only applied on specific samples even if some metabolomics studies have been performed using 2D experiments routinely [124]. Since different peaks from the same metabolite respond in identical ways to concentration variations, finding high correlations between the peak areas of different resonances over cohorts can also be used to assign peaks to metabolites, a method named statistical total correlation spectroscopy (STOCSY) [125]. Compounds identification may be eventually confirmed by spiking standards in the studied matrix.

NMR also represents a powerful tool in quantitative metabolomics. Indeed, in most cases, the peak area is linearly related to the number of equivalent protons that generate the peak. However, inaccurate quantitation may result when the metabolite has different properties than the reference compound, for example, different relaxation behavior, attenuation by the presaturation field or lower excitation efficiency for the nuclear spins of the reference peak. Some metabolites, or even the reference compound itself, can associate with macromolecules or supramolecular assemblies present in the sample, for example, the commonly used reference compound trimethylsilyl propionate (TSP) with albumin and lipoproteins in serum. Other methods use an electronically generated reference signal or a combination of a reference sample and the length of the excitation pulse, which is inversely proportional to the sensitivity of the receiver circuitry.

Data analysis

Data analysis relies on several steps to get from raw data to a list of relevant metabolites, including data preprocessing, data pretreatment and multivariate data analysis and/or statistical analysis.

Data preprocessing

In data preprocessing, two steps are usually carried out, namely, feature alignment and peak picking. In liquid-phase separation techniques (LC–MS and CE–MS), nonlinear shifts occur in the retention/migration time due to pressure fluctuations, changes in column temperature or modifications of the mobile phase/background electrolyte composition. Shifts in the *m/z* dimension are not that big, but still important. Setting the *m/z* parameters wrong in the alignment can lead to erroneous results. Multiple algorithms are nowadays available for alignment, making the selection of the appropriate tool rather challenging [126,127]. These algorithms often combine peak picking and alignment since the latter is typically carried out on the data matrix containing the picked features, and

not on the raw data. This lowers the number of computational steps but may lead to cumulated errors in both peak picking and alignment processes. Moreover, the available alignment packages or methods preclude from performing each step with different packages. Besides the selection of the optimal algorithm, the alignment and peak picking parameters have to be carefully chosen, which represents a cumbersome task since multiple parameters influence the number of features detected, including the analytical platform, the experimental conditions and the matrix. Some guidelines are available in literature to help in the selection of the adequate algorithm and parameters [126,127].

Data preprocessing also includes a normalization step for matrices that show relevant differences in their composition, such as urine [128,129]. In this context, numerous normalization approaches can be considered based on urine volume, osmolality, creatinine concentration or total useful MS signal (i.e., components that are common to all samples) [128]. The optimal strategy for correction of analyte dilution is still matter of discussion. Creatinine concentration may be used to correct for the variability observed in individual sample volumes; however, in many diseases, the renal function and glomerular filtration are affected, which impacts urinary creatinine concentration and leads to inaccurate correction. Total area normalization and total useful MS signals are two alternative approaches for NMR and MS data, respectively, which show the drawback of being strongly influenced by very large signals, resulting in possibly incorrectly normalized data [128]. For both NMR and MS, an alternative is the so-called probabilistic quotient normalization that addresses this problem by calculating the most probable dilution factor based on a reference sample (e.g., single sample or median of QC samples), therefore providing adequate normalization for most matrices commonly encountered in clinical metabolomics [129]. The combination of different approaches may also significantly improve the normalization [130].

Data pretreatment

Data pretreatment includes scaling, centering and transformation to convert the feature list into a different scale suitable for the data analysis *per se* [131]. Transformation of the data is often required for multivariate analysis (MVA) methods, such as PCA, PLS-DA and orthogonal PLS-DA (OPLS-DA), which are very popular in metabolomics. Log or power transformations are generally used to obtain a more symmetric distribution and to correct for heteroscedasticity. Both methods show the same advantages but log transformation is less adapted to values close to zero. Routine scaling methods are unit variance (UV) and pareto scaling that both use a measure of the data dispersion as a scaling factor. In UV scaling, all metabolites become equally important, which increase the influence of small-fold changes. However, this also increases the influence of measurements errors, which are relatively large for low intensities. With pareto scaling, the square root of the standard deviation is used as a scaling factor, leading to a higher decrease in large-fold changes compared with small-fold changes. Therefore, compared with the original peak list, the large-fold changes become less dominant. Without a priori information, UV scaling can be used in a first instance since it is the most objective way to scale the data [131].

Multivariate data analysis

Getting biological information from MVA results can be quite challenging and lead to misinterpretation. A significant difference between data obtained from diseased individuals versus controls in the PCA or PLS-DA scores plot may imply the discovery of a potential biomarker. It is crucial to verify that this difference is not an artifact of the data processing and truly reflects a biological phenomenon. PCA class separation only occurs if the within-class variation is smaller than the between-class variation. Misleading class separations are not caused by biological variation, but derive from sample preparation problems, experimental bias or inadequate data preprocessing [131,132]. Supervised methods such as PLS-DA and OPLS-DA tend to overfit the data, especially when the number of variables is much higher than the samples, which is often the case in metabolomics. This overfitting phenomenon is illustrated in Figure 5. Figure 5A shows a PLS-DA scores plot built on a random data matrix consisting out of 40 samples and 100 variables. The data matrix includes two classes with 20 samples per class. A perfect separation is observed on the scores plot of this PLS-DA model with two clusters; however, based on the method validation step, the model is not valid [104,133]. Indeed, as highlighted in Figure 5B, the cross-validated PLS-DA score plot of the same random data matrix clearly shows that there is no class separation. An optimal model validation would rely on separate training and test datasets, where the test dataset is not used to create the training model. However, this is often not possible due to the low number of samples. Alternative approaches are the internal cross-validation methods, especially the so-called leave-n-out method, where the dataset is partitioned into subsets, and each of this subset is then used as a validation set [134]. The internal cross-validation is assessed by reporting the Q2 value, which may still be relatively high even in case of a poor model. Permutation testing can be used to overcome this limitation, where the class labels are randomly permutated and a new PLS-DA model is calculated.

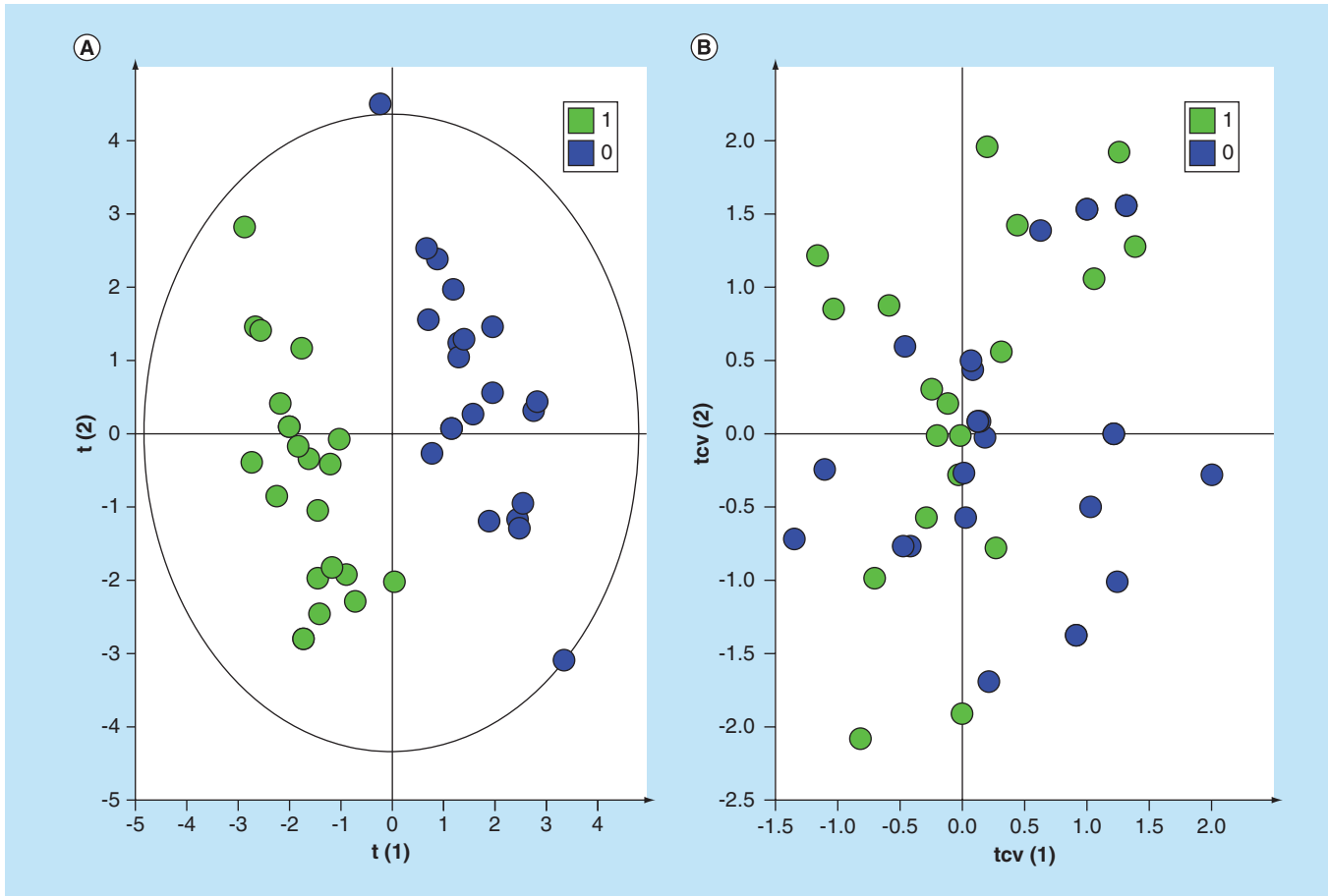


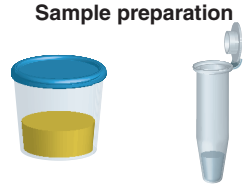
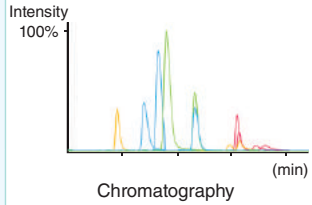
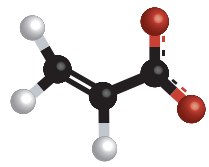
Figure 5. Scores plot obtained with PLS-DA on a random data matrix containing 40 samples and 100 variables. The data matrix includes two classes (green and blue circles, labeled '1' and '0,' respectively) containing 20 samples each. **(A)** Initial scores plot observed with PLS-DA and **(B)** cross-validated PLS-DA scores plot.

The new PLS-DA model is expected to lead to a poor prediction, and, in comparison to the real model, produce a much lower Q2 value. If this test is done many times, a distribution of Q2 value is generated, which can be thus used for hypothesis testing [104].

Metabolite identification & biological interpretation

The information gathered using MVA usually leads to a set of variables that show a relative importance in the discrimination between the study groups. These discriminating features are ranked according to their importance in sample classification. In untargeted metabolomics, the first challenge arises with the selection of variables, namely, the biomarkers candidates that will be subsequently identified prior to data interpretation. The accurate selection of biomarkers candidates is crucial since it significantly impacts the outcome of subsequent validation studies [135]. Even though early metabolomics studies focused on the discovery of single metabolites to explain pathological conditions, it is now widely accepted that pathological processes rely on more complex metabolic changes, fostering the identification of as many metabolites as possible [54]. However, identification of unknown metabolites is a very time-consuming and cost-effective process, which involves rigorous procedures to increase the confidence in metabolites identification. A consensus on minimum requirements for metabolite identification has been proposed [26] to help for standardization between laboratories and ensure that the same terminology is used within the community. Still, the basis for what constitutes a valid metabolite identification remains a matter of debate [26].

Once a set of specific metabolites have been successfully identified with the highest confidence, the biological interpretation of the data represents a major challenge. Indeed, in the large majority of clinical metabolomics-based studies, the candidate biomarkers do not show any obvious functional or pathway associations at first glance. It is only by combining multiple omics data, including genomics, proteomics and transcriptomics data, merged with systems biology and mechanistic approaches (e.g., genome-scale metabolic

Pitfalls and challenges		Solutions
• Confounding factors (e.g., gender, age, diet, lifestyle) • Number of samples per group • Analytic variability > biological variability	**Experimental design**	• Control and limit cofounding factors; admission to clinic to restrict environmental influences • Appropriate number of samples based on pilot study; biological replicates • Regular injection of QCs, blanks, labelled IS; sample randomization
• Metabolome changes – Arterial vs venous blood – Circadian oscillations – Pre- vs post-prandial status – Anticoagulants – Time between sampling and storage	 Study design Sampling Workflow	• Use SOPs describing sampling procedures (matrix, sampling site, containers, storage conditions) • Reproducible sampling conditions
• Chemical and enzymatic degradation and/or interconversion, impacting qualitative and quantitative information	**Sample storage** 	• Immediate quenching (60–100% MeOH) and snap-freezing in liquid nitrogen • Sample storage in adequate containers at -80°C, preferably ≤6 months; ≤3 freeze–thaw cycles • Addition of preservatives if needed (urine)
• Analytical errors (low repeatability, contamination, degradation) • Between-batch variability (large-scale studies)	**Sample preparation** 	• Automation of the process; use of labelled IS; work in cold rooms • Single batch preparation of samples or division of samples into several batches
• Metabolite coverage		• Keep the sample preparation as simple as possible; fit-for-purpose
• GC – Degradation of thermolabile compounds – Cumbersome sample preparation – Significant impact of experimental conditions on data quality	**Separation**  Chromatography Electrophoresis	• GC – Liquid-based techniques for labile analytes – Automation (equal post-derivatization time) – Detailed SOPs
• RPLC: poor retention of polar compounds		• RPLC: use HILIC or CE for polar compounds
• HILIC: poor repeatability; large peak shape		• HILIC: adequate equilibration; reproducible buffers; high proportion of MeCN in sample
• CE: poor migration times reproducibility		• CE: rinsing procedures; coated capillaries; injection of IS; report electrophoretic mobility
• EI: difficult interpretation of fragment spectra in low *m/z* range	**Ionization** 	• EI: lower kinetic energy (≤70 eV) or use CI
• CI: lack of spectral libraries		• CI: in-house libraries or EI for identification
• ESI: matrix effects (suppression/enhancement)		• ESI: systematic evaluation of matrix effects during method development; use of IS
• APCI: lack of spectral libraries; polarity range		• APCI: in-house libraries; ESI for polar compounds

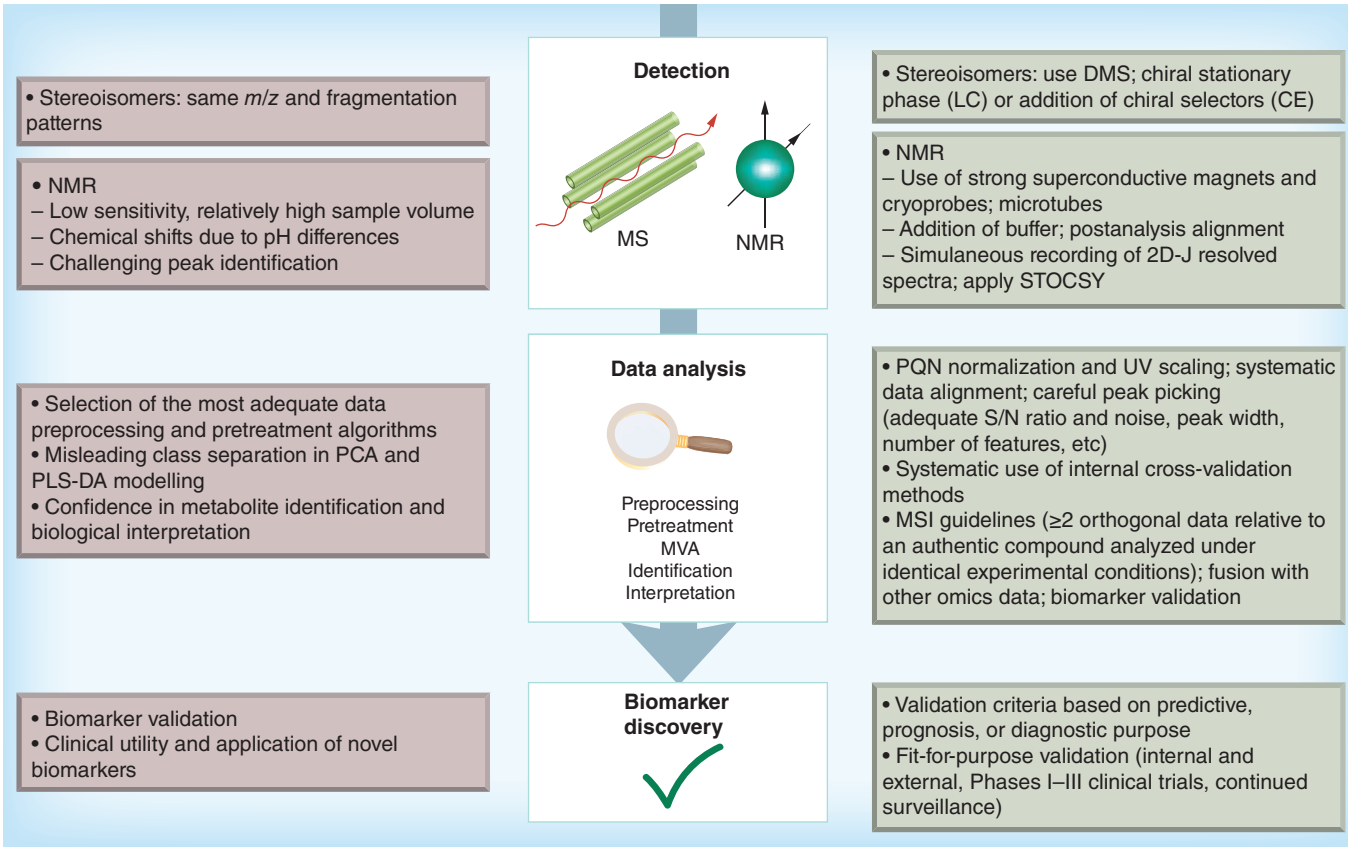


Figure 6. Graphical summary of the analytical pitfalls, challenges and possible solutions that may be encountered in clinical metabolomics (cont. from facing page).

APCI: Atmospheric pressure ionization; CE: Capillary electrophoresis; CI: Chemical ionization; DMS: Differential mobility spectrometry; EI: Electron ionization; ESI: Electrospray ionization; HILIC: Hydrophilic interaction chromatography; IS: Internal standard; MeCN: Acetonitrile; MeOH: Methanol; MSI: Metabolomics Standards Initiative; MVA: Multivariate analysis; PCA: Principal component analysis; PLS-DA: Partial least-squares discriminant analysis; PQN: Probabilistic quotient normalization; QC: Quality control; RPLC: Reversed-phase LC; SOP: Standard operating procedure; STOCSY: Statistical total correlation spectroscopy; UV: Unit variance.

modeling) that relevant pathways and biomolecular mechanisms may be brought to light.

Validation of the candidate biomarkers represents the last and essential step of the process, confirming the clinical utility. The purpose of the biomarkers (predictive, prognosis or diagnostic markers) will determine the sensitivity and specificity required for clinical applications [136]. The potential pitfalls that may come up during the validation stage are similar to the ones encountered in the biomarker discovery stage. Moreover, additional challenges arise during validation studies due to, for instance, larger sets of samples and less significant differences between cases and controls compared with discovery experiments. These challenges are numerous and beyond the scope of the current review [137,138].

Current status of metabolomics in clinical applications

The clinical utility of metabolomics has been discussed and emphasized for multiple diseases, including cardiovascular diseases [139], obesity and Type 2 diabetes [140,141], rheumatoid arthritis and osteoarthritis [142], multiple sclerosis [143], infectious diseases [144], neurological and psychiatric diseases [145], kidney diseases [146] and cancer [147,148]. Nonetheless, among these pathologies, a very small spectrum of validated biomarkers is today used in clinical practice. Serum creatinine is probably one of the most known metabolites used in clinical routine to estimate the glomerular filtration rate or monitor renal malfunctions despite a poor specificity. Another example is the assessment of equine estriols in maternal urine during pregnancy, which can be used to screen for Smith-Lemli-Optiz-syndrome, a rare hereditary disorder of cholesterol metabolism leading to mental retardation and multiple malformations. These metabolites, which play an essential role for early prenatal screening of the disease, have aroused via the analysis of the urinary sterolome [149].

Such progress in disease diagnosis regrettably remains more an exception rather than the rule, as

illustrated with the currently unsuccessful contribution of clinical metabolomics in cancer research. Cancer remains one of the most thrilling challenges of this century with a high rate of therapeutic failure leading to dramatic consequences. The current therapeutic approaches are commonly largely empirical, resulting in inappropriate therapies and high risks of toxicity. However, recent developments, especially in omics technologies, fostered the implementation of more personalized strategies allowing for enhanced clinical benefits. Few biomarkers have been adopted successfully in routine clinical care, for instance, prostate-specific antigen (screening of prostate cancer), α-fetoprotein (screening for hepatocellular and testicular carcinoma) or carcinoembryonic antigen (monitoring of colorectal cancer) [150,151]. However, none of the currently used marker tests has been issued from metabolomics studies. This may come as a surprise since the reprogramming of cellular energy metabolism has been recognized as a crucial hallmark of cancer. Almost a century ago, Otto Warburg already showed that cancer cells show a ravenous appetite for glucose with a higher glycolysis rate compared with normal cells. This so-called Warburg effect is a consequence of multiple metabolic adaptations caused by altered signaling pathways. For instance, the oncometabolite D-2-hydroxyglutarate accumulates in isocitrate dehydrogenase-mutated cells, and high plasma levels have been reported in patient with intrahepatic cholangiocarcinomas or acute myeloid leukemia. However, this oncometabolite has neither been successfully validated nor approved by the US FDA yet, remaining at the stage of a potent biomarker [152]. Much work thus lies ahead in clinical metabolomics and validation of potential biomarkers, especially for population screening and early diagnosis of cancer, two important areas.

Conclusion

Over the last decade, metabolomics-based strategies have shown their relevance in clinical research, providing innovative tools to better understand pathophysiological processes, disease mechanisms and offering new strategies for improved diagnosis, prognosis and the discovery of new therapeutic targets. However, the results of multiple studies have been influenced by unexpected issues during the analytical process. This review highlights the analytical challenges and pitfalls that may be encountered in every step of the analytical process, from study design to data analysis, and provides some solutions to tackle these challenges in both NMR- and MS-based clinical metabolomics. The whole analytical workflow, the potential pitfalls as well as possible solutions and recommendations are illustrated in Figure 6. The clinical outcome of metabolomics studies strongly relies on the quality and accuracy of the acquired data. A careful study design including all the crucial steps should therefore be planned prior to the actual start of the study and the sample collection, ensuring highest data quality and lowest analytical variability. Even though a successful biomarker discovery does not necessarily mean that the validation stage will be rewarding, it certainly positively impacts the success rate of the latter.

Future perspective

Metabolomics has been showing its interest and promises for improved biomarker discovery for a large range of diseases, notably cardiovascular disorders, diabetes, metabolic syndrome, cancer, gastrointestinal diseases, infectious diseases, as well as neurological and psychiatric disorders. It has been paving the way for personalized health strategies for the development of novel and adequate therapies and the implementation of tailor-made and efficient diagnosis approaches. Nevertheless, despite its potential, the translation of metabolomics findings to personalized diagnostic and prognostic medicine is still at an early stage. Various biomarkers or metabolic patterns have been already highlighted in multiple diseases but not validated and used in clinic applications so far. Massive efforts have been put into the discovery phase, efforts that are now also expected for improving the validation of biomarker candidates. Today's literature is overloaded with novel biomarker candidates that, for the large majority of them, do not reach the validation phase. Cross-disciplinary efforts, translational collaborations between academic institutions, pharmaceutical agencies and diagnostic companies, as well as standardization of procedures are crucial to move forwards. Precision medicine is a translational discipline involving multiple scientific communities; the clinical community remains largely unfamiliar with the field of metabolomics. By better understanding the analytical process underlying metabolomics studies, it will certainly help clinicians to foresee its importance in clinical research.

Acknowledgements

HS Jónasdóttir is acknowledged for acquisition of the data resulting in Figure 2.

Financial & competing interests disclosure

I Kohler is a recipient of a fellowship (P2GEP3_155633) supported by the Swiss National Science Foundation. The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.

Executive summary

Chromatographic & electrophoretic separation techniques

- GC is characterized by a high separation efficiency and in combination with electron ionization allows for straightforward spectral data comparison using highly standardized libraries. However, metabolites are prone to thermal degradation and need to be derivatized prior to analysis; highly standardized procedures are therefore needed.
- LC allows for the separation of rather hydrophobic compounds (RPLC) and more polar analytes (HILIC) at low temperatures without the need for derivatization. LC–ESI–MS suffers from matrix effects and a lower reproducibility between instruments, hampering the use of spectral libraries for compounds identification.
- Capillary electrophoresis (CE) shows a very high separation efficiency, notably for enantiomers, and is very well adapted for volume-limited samples. Its limitations rely on the poorer stability, robustness, sensitivity and repeatability of the commercially available CE–MS interfaces compared with LC- and GC-based techniques.

Ionization techniques

- Compounds identification is facilitated with electron ionization via comparison to spectral libraries, but the hard fragmentation may lead to the loss of the molecular ion. Chemical ionization has the advantage of allowing for detection of the molecular ion but renders the compounds identification more cumbersome due to a lack of exhaustive spectral libraries.
- ESI allows for an efficient ionization of a large range of analytes presenting diverse physicochemical properties and masses, and a straightforward hyphenation to both LC and CE. However, it may suffer from significant matrix effects, in other words, ion suppression or enhancement. Atmospheric pressure chemical ionization is less prone to matrix effects and well adapted for the ionization of hydrophobic compounds but it is a less versatile technique.

Detection techniques

- MS is a universal detection technique that presents a very high selectivity and sensitivity, but leads to the destruction of the sample. Quantitative information can be only gathered via the analysis of calibration standards for the compounds of interest.
- NMR is a nondestructive and highly reproducible detection approach that allows for simultaneous qualitative and quantitative determination without complicated sample pretreatment. However, its sensitivity is rather limited compared with MS.

References

Papers of special note have been highlighted as:
• of interest; •• of considerable interest

1 Gates SC, Sweeley CC. Quantitative metabolic profiling based on gas chromatography. *Clin. Chem.* 24(10), 1663–1673 (1978).

2 Dalgliesh CE, Horning EC, Horning MG, Knox KL, Yarger K. A gas-liquid-chromatographic procedure for separating a wide range of metabolites occurring in urine or tissue extracts. *Biochem. J.* 101(3), 792–810 (1966).

• **The first report describing the application of chromatography combined with MS to the analysis of what is known today as the metabolome.**

3 Oliver SG, Winson MK, Kell DB, Baganz F. Systematic functional analysis of the yeast genome. *Trends Biotechnol.* 16(9), 373–378 (1998).

4 Fiehn O. Metabolomics-the link between genotypes and phenotypes. *Plant Mol. Biol.* 48(1–2), 155–171 (2002).

•• **Lays the foundations of modern metabolomics and defines important terms and concepts related to metabolomics strategies.**

5 Nicholson JK, Lindon JC, Holmes E. 'Metabonomics': understanding the metabolic responses of living systems to pathophysiological stimuli via multivariate statistical analysis of biological NMR spectroscopic data. *Xenobiotica* 29(11), 1181–1189 (1999).

6 Patti GJ, Yanes O, Siuzdak G. Innovation: metabolomics: the apogee of the omics trilogy. *Nat. Rev. Mol. Cell Biol.* 13(4), 263–269 (2012).

•• **A very informative review about metabolomics in general.**

7 Cajka T, Fiehn O. Toward merging untargeted and targeted methods in mass spectrometry-based metabolomics and lipidomics. *Anal. Chem.* 88(1), 524–545 (2016).

8 Naz S, Vallejo M, Garcia A, Barbas C. Method validation strategies involved in non-targeted metabolomics. *J. Chromatogr. A* 1353, 99–105 (2014).

9 Tzoulaki I, Ebbels TM, Valdes A, Elliott P, Ioannidis JP. Design and analysis of metabolomics studies in epidemiologic research: a primer on -omic technologies. *Am. J. Epidemiol.* 180(2), 129–139 (2014).

10 Dettmer K, Aronov PA, Hammock BD. Mass spectrometry-based metabolomics. *Mass Spectrom. Rev.* 26(1), 51–78 (2007).

11 Sands BE. Biomarkers of inflammation in inflammatory bowel disease. *Gastroenterology* 149(5), 1275–1285 (2015).

12 Kell DB. Metabolomics and systems biology: making sense of the soup. *Curr. Opin. Microbiol.* 7(3), 296–307 (2004).

13 Viswanathan CT, Bansal S, Booth B *et al.* Quantitative bioanalytical methods validation and implementation: best practices for chromatographic and ligand binding assays. *Pharm. Res.* 24(10), 1962–1973 (2007).

14 Rozet E, Marini RD, Ziemons E, Boulanger B, Hubert P. Advances in validation, risk and uncertainty assessment

of bioanalytical methods. *J. Pharm. Biomed. Anal.* 55(4), 848–858 (2011).

15 Food and Drug Administration. Guidance for Industry, Bioanalytical Method Validation. www.fda.gov/downloads/Drugs/Guidances/ucm070107.pdf

16 European Medicine Agency. Guideline on Bioanalytical Method Validation. www.ema.europa.eu

17 International Conference on Harmonisation. Validation of Analytical Procedures: Text and Methodology Q2(R1). www.ich.org

18 Thompson M, Ellison SLR, Wood R. Harmonized guidelines for single-laboratory validation of methods of analysis (IUPAC technical report). *Pure Appl. Chem.* 74(5), 835–855 (2002).

19 Dunn WB, Broadhurst D, Begley P *et al.* Procedures for large-scale metabolic profiling of serum and plasma using gas chromatography and liquid chromatography coupled to mass spectrometry. *Nat. Protoc.* 6(7), 1060–1083 (2011).

20 Zhu ZJ, Schultz AW, Wang J *et al.* Liquid chromatography quadrupole time-of-flight mass spectrometry characterization of metabolites guided by the METLIN database. *Nat. Protoc.* 8(3), 451–460 (2013).

21 Yuan M, Breitkopf SB, Yang X, Asara JM. A positive/negative ion-switching, targeted mass spectrometry-based metabolomics platform for bodily fluids, cells, and fresh and fixed tissue. *Nat. Protoc.* 7(5), 872–881 (2012).

22 Want EJ, Wilson ID, Gika H *et al.* Global metabolic profiling procedures for urine using UPLC-MS. *Nat. Protoc.* 5(6), 1005–1018 (2010).

23 Want EJ, Masson P, Michopoulos F *et al.* Global metabolic profiling of animal and human tissues via UPLC-MS. *Nat. Protoc.* 8(1), 17–32 (2013).

24 Dunn WB, Wilson ID, Nicholls AW, Broadhurst D. The importance of experimental design and QC samples in large-scale and MS-driven untargeted metabolomic studies of humans. *Bioanalysis* 4(18), 2249–2264 (2012).

•• **Highlights the importance of the study design in metabolomics studies and provides relevant strategies to significantly improve the quality of metabolomics data.**

25 Goodacre R, Broadhurst D, Smilde AK *et al.* Proposed minimum reporting standards for data analysis in metabolomics. *Metabolomics* 3, 231–241 (2007).

26 Sumner LW, Amberg A, Barrett D *et al.* Proposed minimum reporting standards for chemical analysis Chemical Analysis Working Group (CAWG) Metabolomics Standards Initiative (MSI). *Metabolomics* 3(3), 211–221 (2007).

27 Beisken S, Eiden M, Salek RM. Getting the right answers: understanding metabolomics challenges. *Expert Rev. Mol. Diagn.* 15(1), 97–109 (2015).

28 Hyotylainen T, Oresic M. Bioanalytical techniques in nontargeted clinical lipidomics. *Bioanalysis* 8(4), 351–364 (2016).

29 Johnson CH, Gonzalez FJ. Challenges and opportunities of metabolomics. *J. Cell. Physiol.* 227(8), 2975–2981 (2012).

30 Arkin CF, Wachtel MS. How many patiends are necessary to assess test performance. *JAMA* 263, 275–278 (1990).

31 Bando K, Kawahara R, Kunimatsu T *et al.* Influences of biofluid sample collection and handling procedures on GC-MS based metabolomic studies. *J. Biosci. Bioeng.* 110(4), 491–499 (2010).

32 Liu L, Aa J, Wang G *et al.* Differences in metabolite profile between blood plasma and serum. *Anal. Biochem.* 406(2), 105–112 (2010).

33 Gil A, Siegel D, Permentier H, Reijngoud DJ, Dekker F, Bischoff R. Stability of energy metabolites-An often overlooked issue in metabolomics studies: a review. *Electrophoresis* 36(18), 2156–2169 (2015).

• **Energy metabolites represent an important class of metabolites in clinical metabolomics and systems biology approaches, for instance, in cancer metabolism research. Analyzing such compounds is rather complex and requires specific strategies that are described in this exhaustive review.**

34 Siegel D, Permentier H, Bischoff R. Controlling detrimental effects of metal cations in the quantification of energy metabolites via ultrahigh pressure-liquid chromatography-electrospray-tandem mass spectrometry by employing acetylacetone as a volatile eluent modifier. *J. Chromatogr. A* 1294, 87–97 (2013).

35 Mirnaghi FS, Caudy AA. Challenges of analyzing different classes of metabolites by a single analytical method. *Bioanalysis* 6(24), 3393–3416 (2014).

36 Wisniewski CA, White JA, Michel CE *et al.* Optimal method of collection of first-void urine for diagnosis of Chlamydia trachomatis infection in men. *J. Clin. Microbiol.* 46(4), 1466–1469 (2008).

37 Delanghe J, Speeckaert M. Preanalytical requirements of urinalysis. *Biochem. Med. (Zagreb)* 24(1), 89–104 (2014).

38 Gika HG, Wilson ID, Theodoridis GA. LC-MS-based holistic metabolic profiling. Problems, limitations, advantages, and future perspectives. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* 966, 1–6 (2014).

39 Vigor C, Bertrand-Michel J, Pinot E *et al.* Non-enzymatic lipid oxidation products in biological systems: assessment of the metabolites from polyunsaturated fatty acids. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* 964, 65–78 (2014).

40 Kanani H, Chrysanthopoulos PK, Klapa MI. Standardizing GC-MS metabolomics. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* 871(2), 191–201 (2008).

41 Giera M, Lingeman H, Niessen WM. Recent advancements in the LC- and GC-based analysis of malondialdehyde (MDA): a brief overview. *Chromatographia* 75(9–10), 433–440 (2012).

42 Bernini P, Bertini I, Luchinat C, Nincheri P, Staderini S, Turano P. Standard operating procedures for pre-analytical handling of blood and urine for metabolomic studies and biobanks. *J. Biomol. NMR* 49(3–4), 231–243 (2011).

43 Gika HG, Theodoridis GA, Wilson ID. Liquid chromatography and ultra-performance liquid chromatography-mass spectrometry fingerprinting of human urine: sample stability under different handling and storage

conditions for metabonomics studies. *J. Chromatogr. A* 1189(1–2), 314–322 (2008).

44 Lauridsen M, Hansen SH, Jaroszewski JW, Cornett C. Human urine as test material in 1H NMR-based metabonomics: recommendations for sample preparation and storage. *Anal. Chem.* 79(3), 1181–1186 (2007).

45 Van De Merbel NC, De Vries R. Aging of biological matrices and its effect on bioanalytical method performance. *Bioanalysis* 5(19), 2393–2407 (2013).

46 Roux A, Thevenot EA, Seguin F, Olivier MF, Junot C. Impact of collection conditions on the metabolite content of human urine samples as analyzed by liquid chromatography coupled to mass spectrometry and nuclear magnetic resonance spectroscopy. *Metabolomics* 11(5), 1095–1105 (2015).

47 Saude EJ, Sykes BD. Urine stability for metabolomics studies: effects of preparation and storage. *Metabolomics* 3(1), 19–27 (2007).

48 Khamis MM, Adamko DJ, El-Aneed A. Mass spectrometric based approaches in urine metabolomics and biomarker discovery. *Mass Spectrom. Rev.* doi:10.1002/mas.21455 (2015) (Epub ahead of print).

49 Kohler I, Schappler J, Rudaz S. Microextraction techniques combined with capillary electrophoresis in bioanalysis. *Anal. Bioanal. Chem.* 405(1), 125–141 (2013).

50 Raterink RJ, Lindenburg PW, Vreeken RJ, Ramautar R, Hankemeier T. Recent developments in sample pre-treatment techniques for mass spectrometry-based metabolomics. *TrAC* 61, 157–167 (2014).

51 Bruce SJ, Tavazzi I, Parisod V, Rezzi S, Kochhar S, Guy PA. Investigation of human blood plasma sample preparation for performing metabolomics using ultrahigh performance liquid chromatography/mass spectrometry. *Anal. Chem.* 81(9), 3285–3296 (2009).

52 Want EJ, O'maille G, Smith CA *et al.* Solvent-dependent metabolite distribution, clustering, and protein extraction for serum profiling with mass spectrometry. *Anal. Chem.* 78(3), 743–752 (2006).

53 Fernandez-Peralbo MA, Luque De Castro MD. Preparation of urine samples prior to targeted or untargeted metabolomics mass-spectrometry analysis. *TrAC* 41, 75–85 (2012).

54 Kohler I, Derks RJ, Giera M. The rise of hydrophilic interaction chromatography in untargeted clinical metabolomics. *LCGC Eur.* 30(1), 60–75 (2016).

• **HILIC shows interesting advantages in clinical metabolomics but remains little used in this field. This practical review discusses the possible issues encountered when using HILIC in metabolomics and provides solutions to ensure high-quality and repeatable data.**

55 Kohler I, Schappler J, Rudaz S. Highly sensitive capillary electrophoresis-mass spectrometry for rapid screening and accurate quantitation of drugs of abuse in urine. *Anal. Chim. Acta* 780, 101–109 (2013).

56 Matyash V, Liebisch G, Kurzchalia TV, Shevchenko A, Schwudke D. Lipid extraction by methyl-tert-butyl ether for high-throughput lipidomics. *J. Lipid Res.* 49(5), 1137–1146 (2008).

57 Sarafian MH, Gaudin M, Lewis MR *et al.* Objective set of criteria for optimization of sample preparation procedures for ultra-high throughput untargeted blood plasma lipid profiling by ultra performance liquid chromatography-mass spectrometry. *Anal. Chem.* 86(12), 5766–5774 (2014).

58 Van Eeckhaut A, Lanckmans K, Sarre S, Smolders I, Michotte Y. Validation of bioanalytical LC-MS/MS assays: evaluation of matrix effects. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* 877(23), 2198–2207 (2009).

• **A highly relevant review about matrix effects in bioanalytical method development.**

59 Krumpochova P, Bruyneel B, Molenaar D *et al.* Amino acid analysis using chromatography-mass spectrometry: an inter platform comparison study. *J. Pharm. Biomed. Anal.* 114, 398–407 (2015).

60 Fang M, Ivanisevic J, Benton HP *et al.* Thermal degradation of small molecules: a global metabolomic investigation. *Anal. Chem.* 87(21), 10935–10941 (2015).

61 Borman S. Heated dispute over analytical method. *C&EN* 93(42), 25–26 (2015).

62 Koek MM, Jellema RH, Van Der Greef J, Tas AC, Hankemeier T. Quantitative metabolomics based on gas chromatography mass spectrometry: status and perspectives. *Metabolomics* 7(3), 307–328 (2011).

63 Mastrangelo A, Ferrarini A, Rey-Stolle F, Garcia A, Barbas C. From sample treatment to biomarker discovery: a tutorial for untargeted metabolomics based on GC-(EI)-Q-MS. *Anal. Chim. Acta* 900, 21–35 (2015).

64 Naz S, Moreira Dos Santos DC, Garcia A, Barbas C. Analytical protocols based on LC-MS, GC-MS and CE-MS for nontargeted metabolomics of biological tissues. *Bioanalysis* 6(12), 1657–1677 (2014).

65 Kloos DP, Gay E, Lingeman H *et al.* Comprehensive gas chromatography-electron ionisation mass spectrometric analysis of fatty acids and sterols using sequential one-pot silylation: quantification and isotopologue analysis. *Rapid Commun. Mass Spectrom.* 28(13), 1507–1514 (2014).

66 Fiehn O, Kopka J, Trethewey RN, Willmitzer L. Identification of uncommon plant metabolites based on calculation of elemental compositions using gas chromatography and quadrupole mass spectrometry. *Anal. Chem.* 72(15), 3573–3580 (2000).

67 Kloos D, Lingeman H, Mayboroda OA, Deelder AM, Niessen WM, Giera M. Analysis of biologically-active, endogenous carboxylic acids based on chromatography-mass spectrometry. *TrAC* 61, 17–28 (2014).

68 Halket JM, Zaikin VG. Derivatization in mass spectrometry–1. Silylation. *Eur. J. Mass. Spectrom (Chichester, Eng)* 9(1), 1–21 (2003).

69 Villas-Boas SG, Smart KF, Sivakumaran S, Lane GA. Alkylation or silylation for analysis of amino and non-amino organic acids by GC-MS? *Metabolites* 1(1), 3–20 (2011).

70 Spagou K, Tsoukali H, Raikos N, Gika H, Wilson ID, Theodoridis G. Hydrophilic interaction chromatography coupled to MS for metabonomic/metabolomic studies. *J. Sep. Sci.* 33(6–7), 716–727 (2010).

71 Denoroy L, Zimmer L, Renaud B, Parrot S. Ultra high performance liquid chromatography as a tool for the discovery and the analysis of biomarkers of diseases: a review. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* 927, 37–53 (2013).

72 Fekete S, Veuthey JL, Guillarme D. Comparison of the most recent chromatographic approaches applied for fast and high resolution separations: theory and practice. *J. Chromatogr. A* 1408, 1–14 (2015).

73 Fekete S, Kohler I, Rudaz S, Guillarme D. Importance of instrumentation for fast liquid chromatography in pharmaceutical analysis. *J. Pharm. Biomed. Anal.* 87, 105–119 (2014).

74 Kloos DP, Lingeman H, Niessen WM, Deelder AM, Giera M, Mayboroda OA. Evaluation of different column chemistries for fast urinary metabolic profiling. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* 927, 90–96 (2013).

75 Hemstrom P, Irgum K. Hydrophilic interaction chromatography. *J. Sep. Sci.* 29(12), 1784–1821 (2006).

76 Buszewski B, Noga S. Hydrophilic interaction liquid chromatography (HILIC) – a powerful separation technique. *Anal. Bioanal. Chem.* 402(1), 231–247 (2012).

77 Novakova L, Havlikova L, Vlckova H. Hydrophilic interaction chromatography of polar and ionizable compounds by UHPLC. *TrAC* 63, 55–64 (2014).

78 Periat A, Kohler I, Veuthey JL, Guillarme D. Advances in hydrophilic interaction liquid chromatography for pharmaceutical analysis. *LCGC Eur.* 26(5), 17–22 (2013).

79 Barbas C, Moraes EP, Villasenor A. Capillary electrophoresis as a metabolomics tool for non-targeted fingerprinting of biological samples. *J. Pharm. Biomed. Anal.* 55(4), 823–831 (2011).

80 Ramautar R, De Jong GJ. Recent developments in liquid-phase separation techniques for metabolomics. *Bioanalysis* 6(7), 1011–1026 (2014).

81 Zenobi R. Single-cell metabolomics: analytical and biological perspectives. *Science* 342(6163), 1243259 (2013).

82 Wuethrich A, Haddad PR, Quirino JP. Chiral capillary electromigration techniques-mass spectrometry-hope and promise. *Electrophoresis* 35(1), 2–11 (2014).

83 Britz-Mckibbin P. Capillary electrophoresis-electrospray ionization-mass spectrometry (CE-ESI-MS)-based metabolomics. *Methods Mol. Biol.* 708, 229–246 (2011).

84 Lindenburg PW, Haselberg R, Rozing G, Ramautar R. Developments in interfacing designs for CE-MS: towards enabling tools for proteomics and metabolomics. *Chromatographia* 78, 367–377 (2015).

85 Ramautar R, Somsen GW, De Jong GJ. CE-MS for metabolomics: developments and applications in the period 2012–2014. *Electrophoresis* 36(1), 212–224 (2015).

86 Hirayama A, Wakayama M, Soga T. Metabolome analysis based on capillary electrophoresis – mass spectrometry. *TrAC* 61, 215–222 (2014).

87 Harrison AG. *Chemical Ionization Mass Spectrometry (2nd Edition)*. CRC Press LLC, Boca Raton, FL, USA (1992).

88 Quehenberger O, Armando AM, Dennis EA. High sensitivity quantitative lipidomics analysis of fatty acids in biological samples by gas chromatography-mass spectrometry. *Biochim. Biophys. Acta* 1811(11), 648–656 (2011).

89 Wilm M. Principles of electrospray ionization. *Mol. Cell. Proteomics* 10(7), M111 009407 (2011).

90 Matuszewski BK, Constanzer ML, Chavez-Eng CM. Strategies for the assessment of matrix effect in quantitative bioanalytical methods based on HPLC-MS/MS. *Anal. Chem.* 75(13), 3019–3030 (2003).

• **Presents how to quantitatively assess matrix effects. This quantitative approach has been widely used in bioanalysis for evaluation of matrix effects during method development.**

91 Jonasdottir HS, Nicolardi S, Jonker W *et al.* Detection and structural elucidation of esterified oxylipids in human synovial fluid by electrospray ionization-fourier transform ion-cyclotron mass spectrometry and liquid chromatography-ion trap-MS(3): detection of esterified hydroxylated docosapentaenoic acid containing phospholipids. *Anal. Chem.* 85(12), 6003–6010 (2013).

92 Pacchiarotta T, Derks RJ, Hurtado-Fernandez E *et al.* Online spectral library for GC-atmospheric pressure chemical ionization-ToF MS. *Bioanalysis* 5(12), 1515–1525 (2013).

93 Hommerson P, Khan AM, De Jong GJ, Somsen GW. Ionization techniques in capillary electrophoresis-mass spectrometry: principles, design, and application. *Mass Spectrom. Rev.* 30(6), 1096–1120 (2011).

94 Li DX, Gan L, Bronja A, Schmitz OJ. Gas chromatography coupled to atmospheric pressure ionization mass spectrometry (GC-API-MS): review. *Anal. Chim. Acta* 891, 43–61 (2015).

95 Dahinden CA, Clancy RM, Hugli TE. Stereospecificity of leukotriene B4 and structure–function relationships for chemotaxis of human neutrophils. *J. Immunol.* 133(3), 1477–1482 (1984).

96 Jonasdottir HS, Papan C, Fabritz S *et al.* Differential mobility separation of leukotrienes and protectins. *Anal. Chem.* 87(10), 5036–5040 (2015).

97 Chouinard CD, Wei MS, Beekman CR, Kemperman RH, Yost RA. Ion mobility in clinical analysis: current progress and future perspectives. *Clin. Chem.* 62(1), 124–133 (2016).

98 Baker PR, Armando AM, Campbell JL, Quehenberger O, Dennis EA. Three-dimensional enhanced lipidomics analysis combining UPLC, differential ion mobility spectrometry, and mass spectrometric separation strategies. *J. Lipid Res.* 55(11), 2432–2442 (2014).

99 Basit A, Pontis S, Piomelli D, Armirotti A. Ion mobility mass spectrometry enhances low-abundance species detection in untargeted lipidomics. *Metabolomics* 12, 50 (2016).

100 Lintonen TP, Baker PR, Suoniemi M *et al.* Differential mobility spectrometry-driven shotgun lipidomics. *Anal. Chem.* 86(19), 9662–9669 (2014).

101 Kowlgi NG, Chhabra L. D-lactic acidosis: an underrecognized complication of short bowel syndrome. *Gastroenterol. Res. Pract.* 2015, 476215 (2015).

102 Struys EA. 2-Hydroxyglutarate is not a metabolite; D-2-hydroxyglutarate and L-2-hydroxyglutarate are! *Proc. Natl Acad. Sci. USA* 110(51), E4939 (2013).

103 Armstrong DW, Zhang B. Chiral stationary phases for HPLC. *Anal. Chem.* 73(19), A557–A561 (2001).

104 Westerhuis JA, Hoefsloot HC, Smit S *et al.* Assessment of PLSDA cross validation. *Metabolomics* 4, 81–89 (2008).

105 Hoye TR, Jeffrey CS, Shao F. Mosher ester analysis for the determination of absolute configuration of stereogenic (chiral) carbinol carbons. *Nat. Protoc.* 2(10), 2451–2458 (2007).

106 Daykin CA, Corcoran O, Hansen SH *et al.* Application of directly coupled HPLC NMR to separation and characterization of lipoproteins from human serum. *Anal. Chem.* 73(6), 1084–1090 (2001).

107 Akira K, Mitome H, Imachi M, Shida Y, Miyaoka H, Hashimoto T. LC-NMR identification of a novel taurine-related metabolite observed in (1)H NMR-based metabonomics of genetically hypertensive rats. *J. Pharm. Biomed. Anal.* 51(5), 1091–1096 (2010).

108 Psychogios N, Hau DD, Peng J *et al.* The human serum metabolome. *PLoS ONE* 6(2), e16957 (2011).

109 Bouatra S, Aziat F, Mandal R *et al.* The human urine metabolome. *PLoS ONE* 8(9), e73076 (2013).

110 Tomasi G, Van Den Berg F, Andersson C. Correlation optimized warping and dynamic time warping as preprocessing methods for chromatographic data. *J. Chemometrics* 18(5), 231–241 (2004).

111 Savorani F, Tomasi G, Engelsen SB. icoshift: a versatile tool for the rapid alignment of 1D NMR spectra. *J. Magn. Reson.* 202(2), 190–202 (2010).

112 Veselkov KA, Lindon JC, Ebbels TM *et al.* Recursive segment-wise peak alignment of biological (1)h NMR spectra for improved metabolic biomarker recovery. *Anal. Chem.* 81(1), 56–66 (2009).

113 Staab JM, O'connell TM, Gomez SM. Enhancing metabolomic data analysis with Progressive Consensus Alignment of NMR Spectra (PCANS). *BMC Bioinformatics* 11, 123 (2010).

114 De Meyer T, Sinnaeve D, Van Gasse B *et al.* NMR-based characterization of metabolic alterations in hypertension using an adaptive, intelligent binning algorithm. *Anal. Chem.* 80(10), 3783–3790 (2008).

115 Anderson PE, Mahle DA, Doom TE, Reo NV, Delraso NJ, Raymer ML. Dynamic adaptive binning: an improved quantification technique for NMR spectroscopic data. *Metabolomics* 7(2), 179–190 (2011).

116 Davis RA, Charlton AJ, Godward J, Jones SA, Harrison M, Wilson J. Adaptive binning: an improved binning method for metabolomics data using the undecimated wavelet transform. *J. Chemolab.* 85, 144–154 (2007).

117 Hao J, Astle W, De Iorio M, Ebbels TM. BATMAN-an R package for the automated quantification of metabolites from nuclear magnetic resonance spectra using a Bayesian model. *Bioinformatics* 28(15), 2088–2090 (2012).

118 Wishart DS, Tzur D, Knox C *et al.* HMDB: the Human Metabolome Database. *Nucleic Acids Res.* 35, D521–D526 (2007).

119 Ulrich EL, Akutsu H, Doreleijers JF *et al.* BioMagResBank. *Nucleic Acids Res.* 36, D402–D408 (2008).

120 Cui Q, Lewis IA, Hegeman AD *et al.* Metabolite identification via the Madison Metabolomics Consortium Database. *Nat. Biotechnol.* 26(2), 162–164 (2008).

121 Fonville JM, Maher AD, Coen M, Holmes E, Lindon JC, Nicholson JK. Evaluation of full-resolution J-resolved 1H NMR projections of biofluids for metabonomics information retrieval and biomarker identification. *Anal. Chem.* 82(5), 1811–1821 (2010).

122 Lewis IA, Schommer SC, Markley JL. rNMR: open source software for identifying and quantifying metabolites in NMR spectra. *Magn. Reson. Chem.* 47(Suppl. 1), S123–S126 (2009).

123 Xia J, Bjorndahl TC, Tang P, Wishart DS. MetaboMiner-semi-automated identification of metabolites from 2D NMR spectra of complex biofluids. *BMC Bioinformatics* 9, 507 (2008).

124 Van QN, Issaq HJ, Jiang Q *et al.* Comparison of 1D and 2D NMR spectroscopy for metabolic profiling. *J. Proteome Res.* 7(2), 630–639 (2008).

125 Cloarec O, Dumas ME, Craig A *et al.* Statistical total correlation spectroscopy: an exploratory approach for latent biomarker identification from metabolic 1H NMR data sets. *Anal. Chem.* 77(5), 1282–1289 (2005).

126 Smith R, Ventura D, Prince JT. LC-MS alignment in theory and practice: a comprehensive algorithmic review. *Brief. Bioinform.* 16(1), 104–117 (2015).

127 Lange E, Tautenhahn R, Neumann S, Gropl C. Critical assessment of alignment procedures for LC-MS proteomics and metabolomics measurements. *BMC Bioinformatics* 9, 375 (2008).

128 Warrack BM, Hnatyshyn S, Ott KH *et al.* Normalization strategies for metabonomic analysis of urine samples. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* 877(5–6), 547–552 (2009).

129 Dieterle F, Ross A, Schlotterbeck G, Senn H. Probabilistic quotient normalization as robust method to account for dilution of complex biological mixtures. Application in 1H NMR metabonomics. *Anal. Chem.* 78(13), 4281–4290 (2006).

130 Chen Y, Shen G, Zhang R *et al.* Combination of injection volume calibration by creatinine and MS signals' normalization to overcome urine variability in LC-MS-based metabolomics studies. *Anal. Chem.* 85(16), 7659–7665 (2013).

131 Van Den Berg RA, Hoefsloot HC, Westerhuis JA, Smilde AK, Van Der Werf MJ. Centering, scaling, and transformations: improving the biological information content of metabolomics data. *BMC Genomics* 7, 142 (2006).

132 Teahan O, Gamble S, Holmes E *et al.* Impact of analytical bias in metabonomic studies of human blood serum and plasma. *Anal. Chem.* 78(13), 4307–4318 (2006).

133 Kjeldahl K, Ramsus B. Some common misunderstandings in chemometrics. *J. Chemometrics* 24, 558–564 (2010).

134 Worley B, Powers R. Multivariate analysis in metabolomics. *Curr. Metabol.* 1(1), 92–107 (2013).

135 Christin C, Hoefsloot HC, Smilde AK *et al.* A critical assessment of feature selection methods for biomarker

discovery in clinical proteomics. *Mol. Cell. Proteomics* 12(1), 263–276 (2013).

136 Leichtle AB, Dufour JF, Fiedler GM. Potentials and pitfalls of clinical peptidomics and metabolomics. *Swiss Med. Wkly* 143, w13801 (2013).

137 Lin D, Hollander Z, Meredith A, Mcmanus BM. Searching for 'omic' biomarkers. *Can. J. Cardiol.* 25(Suppl. A), A9–A14 (2009).

138 Pepe MS, Fan J, Seymour CW, Li C, Huang Y, Feng Z. Biases introduced by choosing controls to match risk factors of cases in biomarker research. *Clin. Chem.* 58(8), 1242–1251 (2012).

139 Dona AC, Coffey S, Figtree G. Translational and emerging clinical applications of metabolomics in cardiovascular disease diagnosis and treatment. *Eur. J. Prev. Cardiol.* doi:10.1177/2047487316645469 (2016) (Epub ahead of print).

140 Abu Bakar MH, Sarmidi MR, Cheng KK *et al.* Metabolomics – the complementary field in systems biology: a review on obesity and Type 2 diabetes. *Mol. Biosyst.* 11(7), 1742–1774 (2015).

141 Roberts LD, Koulman A, Griffin JL. Towards metabolic biomarkers of insulin resistance and Type 2 diabetes: progress from the metabolome. *Lancet Diabetes Endocrinol.* 2(1), 65–75 (2014).

142 Priori R, Scrivo R, Brandt J *et al.* Metabolomics in rheumatic diseases: the potential of an emerging methodology for improved patient diagnosis, prognosis, and treatment efficacy. *Autoimmun. Rev.* 12(10), 1022–1030 (2013).

143 Bhargava P, Calabresi PA. Metabolomics in multiple sclerosis. *Mult. Scler.* 22(4), 451–460 (2016).

144 Haas CT, Roe JK, Pollara G, Mehta M, Noursadeghi M. Diagnostic 'omics' for active tuberculosis. *BMC Med.* 14, 37 (2016).

145 Hassan-Smith G, Wallace GR, Douglas MR, Sinclair AJ. The role of metabolomics in neurological disease. *J. Neuroimmunol.* 248(1–2), 48–52 (2012).

146 Gentile G, Remuzzi G. Novel biomarkers for renal diseases? None for the moment (but One). *J. Biomol. Screen.* doi:10.1177/1087057116629916 (2016) (Epub ahead of print).

147 Trezzi JP, Vlassis N, Hiller K. The role of metabolomics in the study of cancer biomarkers and in the development of diagnostic tools. *Adv. Exp. Med. Biol.* 867, 41–57 (2015).

148 Mordente A, Meucci E, Martorana GE, Silvestrini A. Cancer biomarkers discovery and validation: state of the art, problems and future perspectives. *Adv. Exp. Med. Biol.* 867, 9–26 (2015).

149 Shackleton CH. Role of a disordered steroid metabolome in the elucidation of sterol and steroid biosynthesis. *Lipids* 47(1), 1–12 (2012).

150 Lynn Henty N, Hayes DF. Cancer biomarkers. *Mol. Oncol.* 6(2), 140–146 (2012).

151 Duffy MJ, Crown J. A personalized approach to cancer treatment: how biomarkers can help. *Clin. Chem.* 54(11), 1770–1779 (2008).

152 Poinsignon V, Mercier L, Nakabayashi K *et al.* Quantitation of isocitrate dehydrogenase (IDH)-induced D and L enantiomers of 2-hydroxyglutaric acid in biological fluids by a fully validated liquid tandem mass spectrometry method, suitable for clinical applications. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* 1022, 290–297 (2016).

Bioanalysis

Best practices for metabolite quantification in drug development: updated recommendation from the European Bioanalysis Forum

Metabolite quantification and profiling continues to grow in importance in today's drug development. The guidance provided by the 2008 FDA Metabolites in Safety Testing Guidance and the subsequent ICH M3(R2) Guidance (2009) has led to a more streamlined process to assess metabolite exposures in preclinical and clinical studies in industry. In addition, the European Bioanalysis Forum (EBF) identified an opportunity to refine the strategies on metabolite quantification considering the experience to date with their recommendation paper on the subject dating from 2010 and integrating the recent discussions on the tiered approach to bioanalytical method validation with focus on metabolite quantification. The current manuscript summarizes the discussion and recommendations from a recent EBF Focus Workshop into an updated recommendation for metabolite quantification in drug development.

First draft submitted: 26 January 2016; Accepted for publication: 27 April 2016; Published online: 24 May 2016

Keywords: EBF • ICH M3 (R2) • metabolite quantification • MIST • scientific validation

Philip Timmerman*[,1], Stefan Blech[2], Stephen White[3], Martha Green[4], Claude Delatour[5], Stuart McDougall[‡,6], Geert Mannens[1], John Smeraglia[5], Stephen Williams[4] & Graeme Young[3]

[1]Janssen R&D, Beerse, Belgium
[2]Boehringer-Ingelheim Pharma GmbH & Co. KG, Biberach, Germany
[3]GlaxoSmithKline Research & Development, Ware, UK
[4]Charles River, Edinburgh, UK
[5]UCB BioPharma, Braine-L'Alleud, Belgium
[6]Covance, Alnwick, UK (at time of authoring)
*Author for correspondence:
Tel.: +32 1460 3581
ptimmerm@its.jnj.com
[‡]Currently at Arcinova, Alnwick, UK

In 2010, the European Bioanalysis Forum (EBF) published a recommendation on best practices for metabolite quantification in relation to the principles of the tiered approach [1].

This publication focused on the ongoing discussions within the EBF on the tiered approach best practices for metabolite quantification introduced at the Crystal City III meeting [2].

With the ICH M3(R2) guidance becoming effective in 2009 [3], the team temporarily suspended the discussions on the general tiered approach and focused on providing the industry with an initial framework to meet the increasing need for quantification of metabolites in early preclinical and clinical studies as required by the Guidance. The EBF recommendation proved to be useful and was often referred to as a cornerstone in the industry discussions on metabolite quantification. At the same time, it was endorsed by the Health Authorities [4,5]. Gradually, more scientists and professional groups published strategic papers relating to the practical consequences of ICH M3(R2) [6–8] and other emerging guidelines relating to metabolite research such as drug–drug interactions (DDI) [9–11].

Subsequently, the discussion on the tiered approach continued not only within the EBF, but also within other cross-company organizations like the Global Bioanalysis Consortium (GBC) [12] and the Japan Bioanalysis Forum (JBF), as well as through publications of individual scientists [13–15]. As both industry and the heath authorities (HA) developed greater understanding and acceptance of which parts of the drug development program were within the scope of a tiered approach, new opportunities arose to extend and refine this flexibility to include bioanalytical method validation. Introducing more descriptive nomenclature highlighting the scientific validity of tiered approach practices, for example, 'scientific

FUTURE SCIENCE part of fsg

validation', helped the bioanalytical community and their stakeholders understand the many benefits of adopting such procedures. As a consequence, by identifying metabolite quantification as an area within the scope of scientific validation [16] discussions to understand what scientific validation means within this context and the best way to apply it within the tiered approach form the basis of the current proposed recommendation.

Because of the increased focus on metabolites in safety testing (MIST) [2], the industry rapidly gained expertise on how metabolite profiling and quantification was impacting their projects, both in relation to safety and efficacy decisions taken from the data as well as in relation to the resources required to generate metabolite concentration data in the bioanalytical (BA) or drug metabolism & pharmacokinetics (DMPK) laboratory. As a result of the latter, while the bioanalytical community created awareness of the possibility of quantifying multiple metabolites in multiple studies, it also challenged an industry attitude to quantify metabolites in too many studies just 'because we can' (i.e., the methodology allowed it). In a collaborative effort, they looked at better ways to generate the plasma (or by extension serum, blood, tissue, urine,…) concentration data needed to make project decisions: combining better ways to interpret the different Guidances surrounding the scientific questions related to MIST and to provide exposure data in different phases of clinical development. The goal was to rationalize which metabolites should be quantified, at which phase of development and with what level of scientific and regulatory rigor as a project moves forward in the development pipeline. Essential in this discussion was a correct understanding of the regulatory requirements from an ICH M3 (R2) and DDI perspective, combined with applying the right level of bioanalytical validation in line with the stage of development (i.e., scientific vs regulatory validation) [5,16–18].

One of the challenges was how to implement a new Guideline in an industry where decisions made following this Guideline will only impact a project 5–10 years later. In the absence of rapid feedback from the HA, the resulting behavior is often to do more than is needed to be sure any risk is minimized, irrespective of the added scientific value. However, this is a sustainable approach only if applied in late stage development when attrition has decimated the number of projects to which the work needs to be applied. But, for ICH M3 (R2) the challenge lies in the disjunction between the timing when the data are needed from a regulatory perspective versus the best timing in a project based on the scientific need and availability of the study samples to generate those data. Indeed, ICH M3 (R2) only requires the metabolite coverage (expression and exposure) information to be available at the start of clinical Phase III (*from ICH M3 (2): "Nonclinical characterization of a human metabolite(s) is only warranted when that metabolite(s) is observed at exposures greater than 10% of total drug-related exposure and at significantly greater levels in humans than the maximum exposure seen in the toxicity studies. Such studies should be conducted to support Phase III clinical trials."*) [3]. But the study samples needed to make those decisions related to metabolite coverage across species are typically available in an earlier stage of development, for example, at the first-into-man (FiM) and multiple ascending dose (MAD) studies when samples from 1 or 3 months GLP toxicology studies are available. As a consequence, the industry is tempted to de-risk their projects with respect to metabolite coverage at an earlier stage of development than required for ICH M3 (R2) and is therefore investing a lot of resources in projects that will never reach the clinical Phase III stage.

With this awareness, we deemed it necessary to revisit our 2010 recommendation paper [1]. In the initial recommendation, we identified a significant number of early development studies in scope to be analyzed using regulatory validation where, based on current thinking on the tiered approach/scientific validation, alternative leaner methodologies will be equally valid to make the right project decision. Supported and stimulated by the feedback from experts in the drug metabolism arena, the EBF issued a survey on the topic and organized a Focus Workshop in September 2015 in Brussels to discuss and propose a refined recommendation for metabolite quantification. At the Workshop we also took the opportunity to discuss the current practice of metabolite profiling. The proposed recommendation largely applies to small molecule new chemical entities (NCEs), but is appropriate for other drugs if the tools are available to characterize and quantify their metabolites.

Feedback from the survey

The workshop was preceded by a survey in the EBF community on the topic.

In this survey, we probed the EBF community on their day-to day practices around metabolite profiling and quantification. A total of 29 companies responded. The majority of the responses related to NCEs (26 companies) and peptides (9 companies). The questions related to:

- Availability of a company strategy for metabolite profiling and quantification;
- The Guidance is followed;

- The laboratory in which the analysis is performed (internally, outsourced or DMPK or BA laboratory);
- The technology platform is used;
- The samples (species) and matrices used for profiling and/or quantification;
- The regulatory rigor used for metabolite quantification;
- The timing of the work;
- Reporting.

A summary of the survey data can be found on the Focus Workshop conference website [19].

This allowed us to re-establish our community's best current practice and forward thinking on the subject.

Focus workshop

The Focus Workshop was attended by approximately 80 scientists from over 50 different companies. Both the survey feedback and the good attendance of metabolite profiling (drug metabolism scientists) and metabolite quantification experts (bioanalytical scientists) at the workshop allowed us to consider our current updated recommendation to be representative.

The first half of the workshop was built around feedback on the current status of metabolite profiling and quantification practices. Both the EBF as well as four individual (global) pharma companies presented their current strategy on the subject, followed by a CRO perspective. The second half of the meeting comprised two parallel breakout sessions. The first one discussed the updated EBF recommendation proposal on metabolite quantification; the second one discussed potential ways to strategize metabolite profiling.

Workshop on metabolite quantification

The updated EBF recommendation for metabolite quantification supported by the workshop discussions is visualized in Figure 1.

Discussion

Referring to the initial recommendation of 2010, we removed the difference in strategy for metabolites with known activity/toxicity versus unknown activity/toxicity. In general, our current thinking is that there should be no difference in strategy toward documenting relative exposures in experimental animals versus human based on activity/toxicity prior to evaluating a drug against the ICH M3(R2) Guideline. Indeed, prior to understanding the coverage in human versus experimental animals or the contribution of metabolites to the pharmacological activity, this difference seems to make little sense. In unique cases where the metabolite is known to be equipotent or of similar activity to the dosed drug, it is obvious that quantification of that metabolite should follow the same strategy as that of the dosed drug. However, and as mentioned at the workshop: 'the exception is the exception' and should not be setting the rule. Removing this difference allows a simplified decision tree, which can be more easily connected to the aforementioned discussions on alternative approaches to bioanalytical method validation.

Metabolite quantification in the drug discovery phase

In the discovery phase, we recommend the use of screening methods for metabolite quantification [1]. Also, and connected to the fact that more requests for metabolite quantification are entering the bioanalytical laboratory in this early phase, we suggest to scrutinize these requests to balance the workload of metabolite quantification versus the added value of these additional data when making project decisions. This does not imply we advise against metabolite quantification in general in the discovery phase, but it makes sense to challenge the request to the extent that the added workload is understood by the project representatives.

Metabolite quantification in pre-Phase I & early clinical phase

This is an area of major change to our recommendation. Our update aligns both with the discussions at the workshop and the two major surveys held in 2015, and another related to the areas of application of scientific validation in support of the recent EBF recommendation paper [16]; this latter survey was re-issued after refinement in preparation of the global discussion on the subject at the 2015 AAPS APQ Open Forum. In both surveys, we probed the EBF community on their view as to whether metabolite quantification in early development would require regulatory or scientific validation. More than 80% of the responders agreed that scientific validation is a better approach, both in the EBF community as well as in the AAPS and JBF community.

In the current updated recommendation, we propose to use either screening methods (which includes relative comparison of metabolite abundance – peak area ratios of metabolites/dosed drug across species) or scientifically validated method (see Box 1) for bioanalysis in support of assessing the exposure of metabolites in man relative to experimental animals.

Several strategies can be applied to the timing of the experiments based on the availability of samples. Three

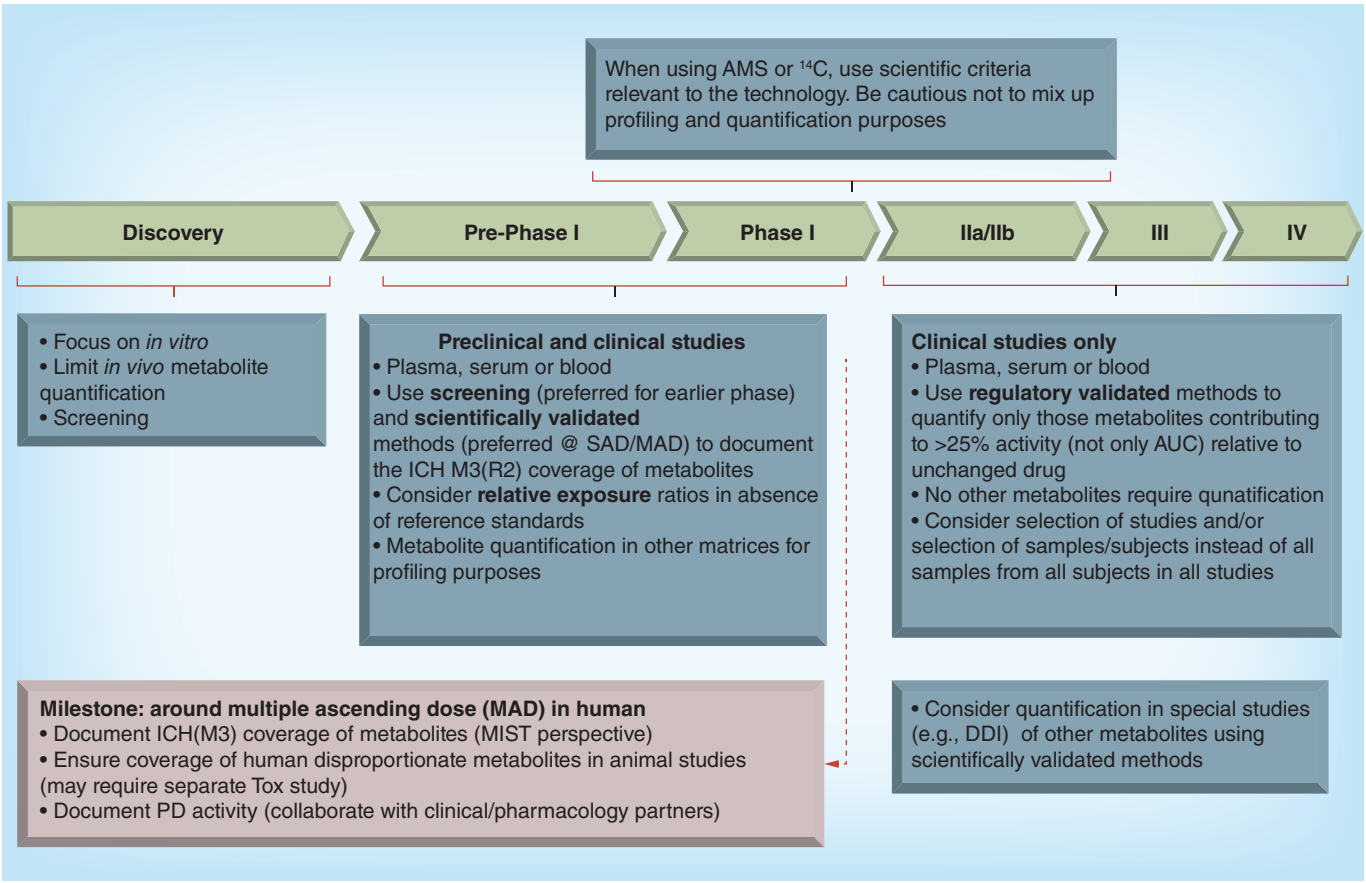


Figure 1. Updated EBF recommendation for metabolite quantification.

major approaches are common in the industry today, the latter being the most common approach:

- Samples from earlier toxicokinetic (TK) studies are stored, with the analysis to be included in this single assay campaign at the time human samples become available.
- Samples from more recent TK studies are included in a single assay campaign when samples from the multiple ascending dose study are available.
- Metabolite concentration data in the different species and human are generated at the time these samples become available.

In any strategy that is followed, the bioanalytical scientist needs to ensure that the matrix concentrations are supported with documented and reconstructable assay performance data on accuracy, precision and stability in line with the purpose of the decision to be made. We suggest using the scientific validation criteria for those metabolites for which a reference standard is available. However, strictly speaking, at this stage, there is no need to have access to a reference standard as long as a relative assessment comparing metabolite/dosed drug peak areas across species can be made with similar scientific rigor and considering relevant matrix stability data.

Although ICH M3(R2) only requires a metabolite strategy at the start of Phase III, it is likely to be more practical for the industry to perform this metabolite coverage assessment at the time the samples from multiple dosing studies are available. Considering a standard study package in the development process of a drug, we usually have access to samples from 1-month or 3-month multiple dose studies in the main toxicology species at the time of the MAD study. So, from a practical perspective and provided stability data can be generated for the metabolite (ratio), it makes sense to perform this experiment around that time frame in a single campaign.

The result of the analysis is either 'coverage' or 'no coverage'.

- The desirable outcome is that there is metabolite coverage in the preclinical species. From here onward, the strategy can become simple: stop all future analysis of metabolites in preclinical species. For man, we still need to factor in the potential contribution to the pharmacological activity and

Box.1 Definitions for regulatory and scientific validation.

- **Regulatory validation:** Assay validations to provide scientifically accurate, reproducible and reconstructable concentration data to allow valid decision making for the intended purpose of the study and comply with regulated BA standards as specified by health authority (HA) guidance documents
- **Scientific validation:** Assay validations to provide scientifically accurate, reproducible and reconstructable concentration data to allow valid decision making for the intended purpose of the study and can withstand independent review – including scientific review from regulators if so required – although not applying all elements specified by HA guidance documents

the DDI potential. This is discussed in the next paragraph, but is irrelevant for preclinical studies.

- For metabolites where there is no coverage in man, we suggest updating the development plan and include studies/experiments to ensure coverage in preclinical species is documented (e.g., separate dosing of the metabolite or higher doses of the investigational drug to reach higher exposure levels of the metabolites, if ethically acceptable). In some cases (e.g., dosing a metabolite in a GLP tox study) under the current paradigm of bioanalytical method validation, a regulatory validated method may be required to quantify these metabolite concentrations. Moving forward, we are encouraging the industry and regulators to consider accepting scientific validation in early development GLP studies, with the view that the proposed scientific rigor, *a priori* set criteria and level of documentation are aligned with scientific and GLP requirements. However, although discussed at the workshop, this is not the subject of the current recommendation paper.

In addition to metabolite coverage from an ICH M3(R2) perspective, it is important in the process to include additional characteristics of the metabolite beyond the exposure (i.e., AUC at steady state) to fully understand the contribution of the metabolite to the overall activity. Often assessments are made on AUC only, but it is recommended to include those parameters which describe the contribution to the activity as well. Although this is an area in need of intensified discussion, a proposed starting point would be to include at least the pharmacological activity (if available) or plasma protein binding of the metabolite(s) and the dosed drug to correct the AUC for activity. It is our recommendation that, in addition to not requiring further quantification of metabolites when we have either coverage in preclinical species or the exposure of a metabolite in human plasma is less than 10% of total drug exposure, also to preclude the requirement for further quantification of metabolites when the relative contribution to the activity in man is less than 25% compared with the dosed drug, (based on steady-state AUC values corrected for pharmacological activity).

In summary, and in contrast to our initial recommendation, the EBF does not see a requirement to use regulatory validated methods for metabolites until after the MAD study, for either preclinical or clinical studies. All metabolite quantification and decisions on coverage of metabolites in relation to ICH M3(R2) can be documented using screening and scientific validation approaches. Also, if metabolite exposure in man is covered in preclinical species, we recommend stopping all further analysis of metabolites in preclinical studies.

Continued metabolite quantification after the multiple ascending dose

This is the second area of change in comparison to our initial recommendation, and is likely one with significant impact on the bioanalytical laboratory. The current regulatory requirements for metabolite coverages in man versus experimental animals, stating: "*Nonclinical characterization of a human metabolite(s) is only warranted when that metabolite(s) is observed at exposures greater than 10% of total drug-related exposure and at significantly greater levels in humans than the maximum exposure seen in the toxicity studies.*" [3] equally provides guidance on how to continue the metabolite quantification in later stages of development.

First, and as recommended above, if the metabolite coverage is documented and complies with ICH M3(R2), there is no need for continued quantification of metabolites in preclinical species. Hence, and since our recommendation is to perform this assessment at the MAD milestone, there is no need to continue the quantification of metabolites in preclinical species after this milestone. At the same time, there is also little value in excessive quantification of metabolites prior to this assessment, since the assessment qualifies the need for metabolite quantification.

On a case-by-case basis there may be a need to include additional toxicology species at later stages in development. In those cases we would recommend analyzing metabolites, if needed, using the screening (including relative peak area ratios of metabolite/dosed drug) or scientifically validated approach as detailed earlier.

With respect to continued quantification in human samples, a similar rationale is proposed. And although

it is very tempting for project teams to continue to request metabolite concentrations in later clinical studies, it is our recommendation to only generate metabolite concentrations for those metabolites that contribute to more than 25% of the activity in relation to the dosed drug (based on AUC corrected for the pharmacological activity) in the continuation of the development of the drug. In essence we do recommend using regulatory validated methods for all metabolites complying with these characteristics. We would however consider not necessarily continuing quantification of the metabolites in all studies but only in a selection of studies and/or selection of samples/subjects.

There are two areas of concern:

- There are programs in which the drug is extensively metabolized (e.g., prodrugs or hormones are good examples). In those cases, the above recommendation can lead to an explosion of metabolites qualifying for continued quantification and it may be warranted to agree on an alternative strategy. Agreeing beforehand on the bioanalytical strategy to be followed with the HA is advised. Both the sponsor and the HA should be open to a strategy which considers understanding the PK/PD relationship while appropriately limiting the resource investments required to generate these data. Strategizing the BA work around a major pharmacological metabolite can be a fair starting point.

- Special population studies (e.g., elderly, hepatic/renal impaired and children) and DDI studies where, depending on differences in clearance and/or metabolism, metabolite exposure could increase above the thresholds mentioned in the ICH Guidance. If known metabolites of the investigational drug would need to be measured in those studies, we would recommend using scientific validation and focus on those metabolites which in an earlier assessment were on the borderline of being covered and for which there may be a concern. Again we would recommend to consider only quantifying metabolites in a selection of samples/subjects. For the analysis of metabolites from the industry standard DDI probes, we recommend regulatory validated methods.

In summary, after the MAD study, identified as the milestone to document metabolite coverage in relation to ICH M3(R2), we recommend to only quantify those metabolites that contribute more than 25% to the pharmacological activity compared with the dosed drug. Quantification of those metabolites should be performed using a regulatory validated method. We recommend evaluating whether the metabolite(s) need(s) to be included in every study from that stage onward, and at which stage enough data are gathered. Any metabolites that require quantification in special clinical or preclinical studies would only require a scientific validation. For an extensively metabolized drug, other strategies may apply.

Conclusion for metabolite quantification

In this updated recommendation on metabolite quantification in drug development, we recommend limiting quantitation of metabolites in all stages of development to those that are needed to understand their contribution to the activity or toxicity of the dosed drug. Also, we recommend using the principles of tiered approach for bioanalytical method validation.

For the different stages of development, this means:

- Drug discovery: minimize the analysis of metabolites in *in vivo* samples. If analysis is required, use screening methods, including relative peak area ratios of metabolites/dosed drug.

- Early development (pre-Phase 1 until MAD): assess metabolite exposure in man and coverage in preclinical species at the timing of the MAD compared with metabolite exposure in multiple dose preclinical studies. Use screening methods, including relative peak area ratios of metabolites/dosed drug, or scientific validation.

- After MAD: no need to quantify any metabolites in preclinical species unless required by ICH M3(R2). In clinical studies, quantify only those metabolites contributing to >25% activity (based on AUC corrected for activity) relative to dosed drug using regulatory validated methods and consider selection of studies and/or selection of samples.

- Treat exceptions as exceptions: any metabolites requiring quantification in special clinical or preclinical studies would only require a scientific validation. For extensively metabolized drug, other strategies may apply. In special cases, agree on a bioanalytical strategy, preferably with the HA based on scientific needs and sustainable use of resources.

A schematic overview is presented in Figure 1.

Workshop on metabolite profiling

The discussions related to metabolite profiling tried to unite the audience around the strategy used for metabolite profiling: "Who does What, When, Why and How?" Then build on the input from the aforementioned survey [19].

From the survey, Figures 2 & 3 were developed. The figures show at what stage the DMPK laboratory gets

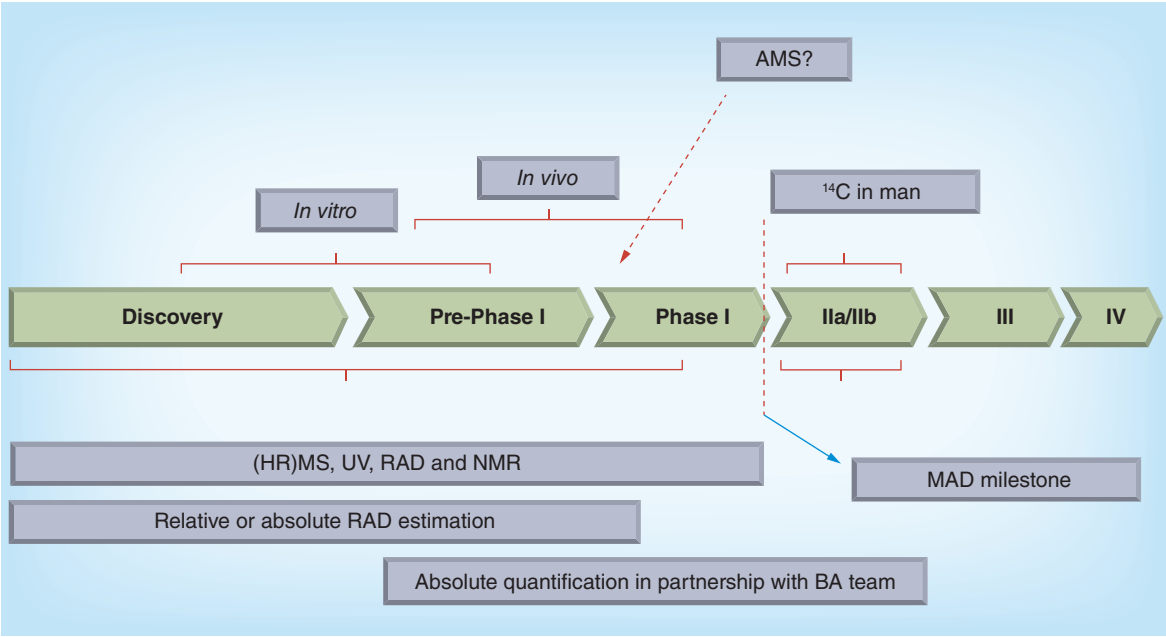


Figure 2. Metabolite profiling strategies in drug development: tools involved.

involved, which analytical tools are used for profiling and which types of *in vivo* and *in vitro* studies are performed for small molecule (NCE) metabolite profiling. This information, together with the presentations from four Pharma companies and a CRO were the starting point of the workshop discussions on metabolite profiling. In contrast to metabolite quantification, where the audience could build on an existing EBF Recommendation on top of many years of intense discussion in the industry, metabolite profiling was identified as an area where it will be difficult to provide a similar recommendation. Every molecule is different and requires scientific freedom in regard to documentation and development of understanding of its metabolic fate. Hence, our expert audience focused on sharing ideas, experience and strategies on how the various technologies in different stages of development add value.

Highlights from the discussions

In addition to sharing individual company strategies as a source of inspiration for metabolite profiling, the workshop discussion covered four main areas:

- General considerations
- Regulatory aspects
- Quality aspects
- Scientific aspects – with focus on the complexity of metabolite stability during sampling storage and sample handling in multiple matrices involved in metabolite profiling.

The delegates rapidly identified that because of the absence of regulatory guidance related to metabolite profiling and the split between the more traditional front loaded approach (i.e., a relatively large package of work conducted preclinically often using radiolabeled drug) versus a 'human first' based approach (with more limited support work conducted pre-FTIH and radiolabel not used until much later in development) it is difficult to provide guidance on this subject. The governance in a company should drive the strategy. Hence, we limit our feedback on the metabolite profiling workshop to state that it is essential to allow scientific freedom and continue to benefit from the rapidly evolving technological innovations.

All the presentations from the meeting and bulleted highlights from the workshop can be found on the EBF Focus Workshop website [20]. They illustrate how companies strategies can differ in their approaches yet are able to generate the required information in time to make adequate project decisions. All roads apparently lead to Rome, although it was not built in a day. The latter is certainly also true for elucidation of the metabolism of many compounds in industry.

In addition, the discussions at the workshop provided an opportunity to connect DMPK experts with the ongoing discussions in the bioanalytical arena on scientific validation. In this respect, it is important to highlight scientists in both functional teams are involved in quantification of metabolites, yet using different quality standards or processes. Often, the analytical acceptance criteria in the DMPK labora-

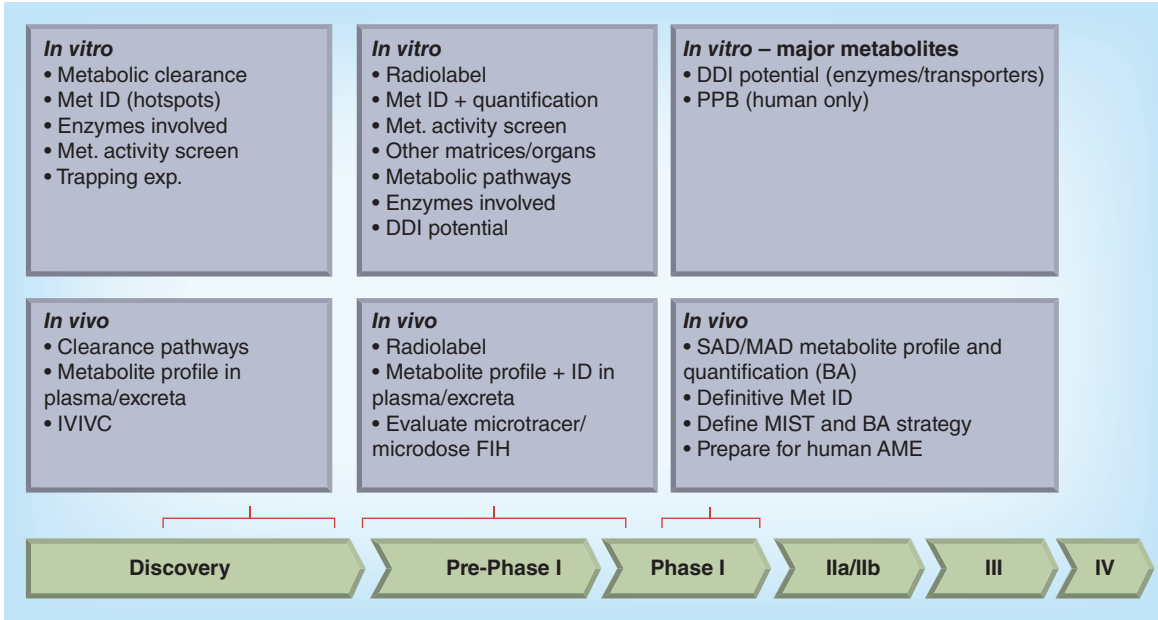


Figure 3. Metabolite profiling strategies in drug development: studies involved.

tory are not governed by internationally agreed quality standards as compared with regular bioanalysis. Nevertheless, these metabolite profiling activities do follow scientifically valid criteria in line with the specific technology used (and often described in company SOPs or procedural documents), be it ultraviolet (UV) detection, accelerator mass spectrometry (AMS), radioactive detection (RA), NMR (nuclear magnetic resonance) or other tools. This manuscript is not intended as a plea for more regulations in the area of metabolite quantification embedded in metabolite profiling activities. We do want to highlight that a close(r) partnership between the metabolite profiling scientists in the DMPK laboratory and the bioanalytical laboratory is essential to understand acceptance criteria and to avoid potentially incorrect interpretation of the accuracy of concentration data originating from RAD, UV or MS which can, if so required, be annotated appropriately.

Conclusion

Over the last 5 years, the industry has gained a lot of experience with the practical aspects, including the use of the data in drug R&D, of metabolite profiling and quantification. This, together with the intensified discussion and added value of the principle of the tiered approach and scientific validation in bioanalysis has stimulated the EBF to take a fresh look at their recommendation on metabolite quantification from 2010. As a result of new targeted surveys and peer discussion at a Focus Workshop in September 2015, the EBF is now refining their 2010 recommendation. In the updated recommendation, we are putting more emphasis on the value of scientific validation approaches as an alternative to regulatory validation for metabolite quantification in the earlier stages of development. At the same time, we recommend the industry reflects on limiting quantitation of metabolites once the contribution of metabolites to the toxicology or pharmacology is understood.

For metabolite profiling, the discussions from the workshop highlighted the importance of scientific freedom, use of state-of-the-art technology and increased partnership between drug metabolism experts and bioanalytical scientists as a good recipe for a result-driven strategy.

Acknowledgements

The authors want to acknowledge the EBF member companies that have contributed to the surveys, workshops and internal discussions, which form the backbone of the current recommendation.

Disclaimer

The views and conclusion presented in this paper are those of the European Bioanalysis Forum and do not necessarily reflect the representative affiliation or company's position on the subject.

Financial & competing interests disclosure

The authors have no relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript. This includes employment, consultancies, honoraria, stock ownership or options, expert testimony, grants or patents received or pending, or royalties.

No writing assistance was utilized in the production of this manuscript.

References

1 Timmerman P, Kall M, Gordon B *et al.* Best practices in a tiered approach to metabolite quantification: views and recommendations of the European Bioanalysis Forum. *Bioanalysis* 2(7), 1185–1194 (2010).

2 Viswanathan CT, Bansal S Booth B *et al.* Workshop/ Conference Report – Quantitative bioanalytical methods validation and implementation: best practices for chromatographic and ligand binding assays. *AAPS J.* 9(1), E30–E42 (2007).

3 European Medicines Agency. CPMP/ICH/286/95 – ICH Topic M 3 (R2) non-clinical safety studies for the conduct of human clinical trials and marketing authorization for pharmaceuticals. London, UK (2009).

4 Booth B. When do you need a validated assay? *Bioanalysis* 3(24), 2729–2730 (2011).

5 Ministry of Health, Labor and Welfare. Guideline on bioanalytical method validation in pharmaceutical development. Tokyo, Japan (2013) www.nihs.go.jp/drug/BMV/250913_BMV-GL_E.pdf

6 Leclercq L1, Cuyckens F, Mannens GS, de Vries R, Timmerman P, Evans DC. Which human metabolites have we MIST? Retrospective analysis, practical aspects, and perspectives for metabolite identification and quantification in pharmaceutical development. *Chem. Res. Toxicol.* 22(2), 280–293 (2009).

7 Ma S, Chowdhury S. Analytical strategies for assessment of human metabolites in preclinical safety testing. *Anal. Chem.* 83, 5028–5036 (2011).

8 Yu H, Bischoff D and Tweedie D. Challenges and solutions to metabolites in safety testing: impact of the International Conference on Harmonization M3(R2) guidance. *Expert Opin. Drug Metab. Toxicol.* 6(12), 1539–1549 (2010).

9 European Medicines Agency. *Guideline on Investigation of Drug Interactions.* London, UK (2012). www.ema.europa.eu

10 Ministry of Health, Labor and Welfare. Guideline on Drug Interactions for Drug Development and Appropriate Provision of Drug Information (final draft.) Tokyo, Japan (2014).

11 US FDA. Guidance for Industry: Drug Interaction Studies — Study Design, Data Analysis, Implications for Dosing, and Labeling Recommendations. www.fda.gov

12 Lowes S, Hucker R, Jemal M *et al.* Tiered approaches to chromatographic bioanalytical method performance evaluation: recommendation for best practices and harmonization from the global bioanalysis consortium harmonization team. *AAPS J.* 17(1) 17–23 (2015).

13 Timmerman Philip. Mini focus issue: tiered approach to bioanalysis. *Bioanalysis* 6(5), 585–682 (2014).

14 Timmerman P. Tiered approach revisited: introducing stage-appropriate or assay-appropriate scientific validation. *Bioanalysis* 6(5), 599–604 (2014).

15 Timmerman P, Lowes S, McDougall S *et al.* Scientific or regulated validation: a tiered approach? Meeting report from a joint EBF/DVDMDG workshop. *Bioanalysis* 7(14), 1703–1710 (2015).

16 Timmerman P, White S, McDougall S *et al.* Tiered approach into practice: scientific validation for chromatography-based assays in early development – a recommendation from the European Bioanalysis Forum. *Bioanalysis* 7(18), 2387–2398 (2015).

17 US. Department of Health and Human Services Food and Drug Administration Center for Drug Evaluation and Research (CDER) Center for Veterinary Medicine (CVM). *Guidance for Industry Bioanalytical Method Validation* (2001). www.fda.gov/downloads/Drugs/Guidance/ucm070107.pdf

18 European Medicines Agency. Guideline on bioanalytical method validation EMEA/CHMP/EWP/192217/2009. www.ema.europa.eu

19 Feedback from the survey – Stuart Mc Dougall on behalf of the EBF. www.focus.europeanbioanalysisforum.eu

20 EBF - Focus Workshop – Bringing Assay Validation and Analysis of Biomarkers into Practice. http://focus.europeanbioanalysisforum.eu/slides/

MedChemNet

Join our free medicinal chemistry community hub!

Connect – Raise awareness of your medicinal chemistry research and build connections with your peers

Collaborate – Join the conversation and provide a catalyst to accelerate progress in this ever-changing field

Discover – Share insights and gain exclusive free access to news, peer-reviewed articles and videos

Research Article
Special Focus Issue – Immunoaffinity MS

Bioanalysis

Quantification Studies

Characterization and quantification of oxyntomodulin in human and rat plasma using high-resolution accurate mass LC–MS

Background: A thorough understanding of the biological role of oxyntomodulin (OXM) has been limited by the availability of sensitive and specific analytical tools for reliable *in vivo* characterization. Here, we utilized immunoaffinity capture coupled with high-resolution accurate mass LC–MS detection to quantify OXM and its primary catabolites. **Results:** Quantification of intact OXM 1–37 in human and rat plasma occurred in pre- and post-prandial samples. Profiles for the major catabolites were observed allowing kinetic differences to be assessed between species. **Conclusion:** A validated assay in human and rat plasma was obtained for OXM 1–37 and its catabolites, 3–37 and 4–37. The value of full scan high-resolution accurate mass detection without selected reaction monitoring for low-abundance peptide quantification was also demonstrated.

First draft submitted: 20 January 2016; Accepted for publication: 1 June 2016;
Published online: 17 June 2016

Keywords: catabolites • high-resolution accurate mass MS • immunoaffinity enrichment • orbitrap • oxyntomodulin • quantification • validation

Jennifer M Cox*[,1], Michael J Berna[1], Zhaoyan Jin[1], Amy L Cox[1], Kyle W Sloop[1], Jesus A Gutierrez[‡,1] & Bradley L Ackermann[‡,1]

[1]Lilly Research Laboratories, Eli Lilly & Company, Lilly Corporate Center, Indianapolis, IN 46285, USA

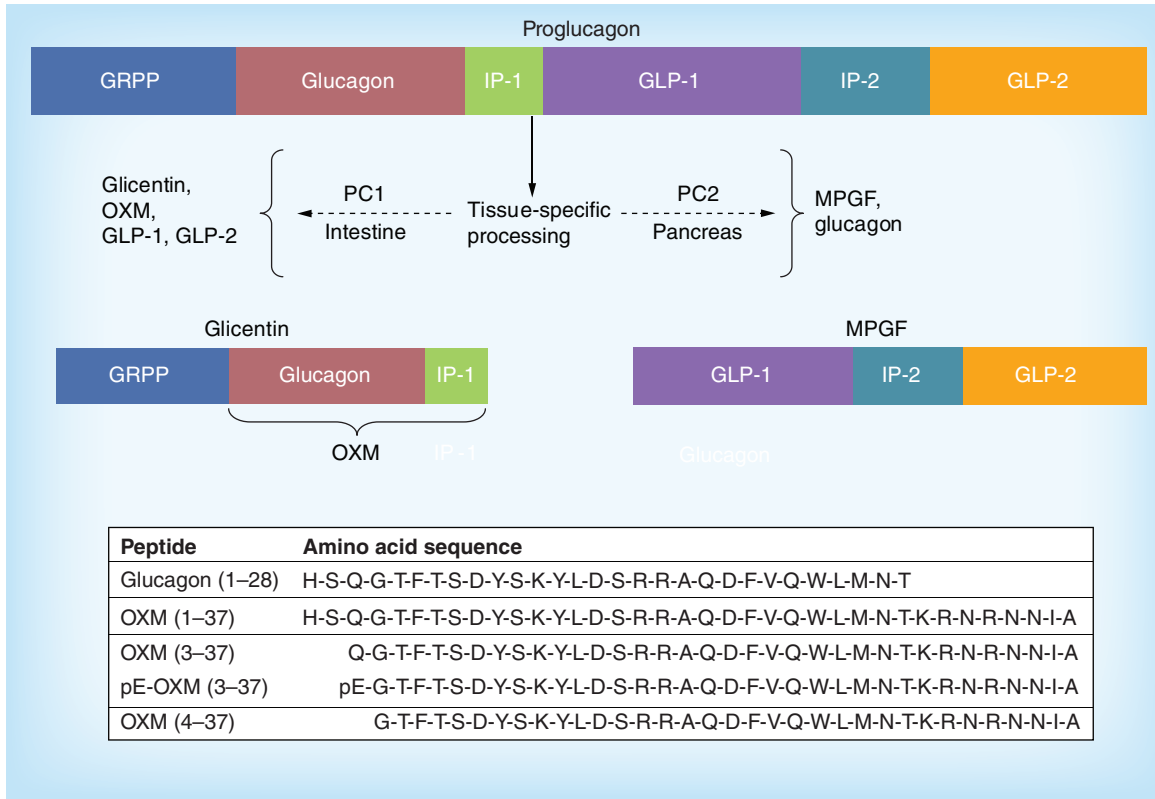
*Author for correspondence:
Tel.: +1 317 651 5746
Fax: +1 317 655 2863
cox_jennifer_marie@lilly.com

‡Authors contributed equally

Since its identification in 1981 [1,2], oxyntomodulin (OXM) has been the focus of extensive clinical and preclinical studies because of its physiological effects on weight loss, food intake and energy balance [3,4]. Studies in humans have shown that acute exposure to OXM reduces food intake resulting in body weight reduction, increased energy expenditure and improved glucose metabolism; traits of high interest for the therapeutic control of obesity and Type 2 diabetes mellitus [5–8]. As shown in Figure 1, OXM is one of several peptide hormones produced from a single gene that codes for the protein proglucagon. Post-translational processing of proglucagon, regulated at the tissue level by proteases known as prohormone convertases (PC), results in the production of specific peptide hormones [9]. In response to glucose and other nutrients, PC1 found in intestinal tissues (L-cells), produces four bioactive peptides: glicentin, OXM, GLP-1 and GLP-2, each having unique metabolic functions related to glucose homeostasis [5,10]. Alternative processing of proglucagon by the pancreas (α-cells) produces glucagon and a larger peptide fragment known as major proglucagon fragment. This action occurs by PC2 under hypoglycemic conditions.

At the cellular level, the physiologic properties of OXM are mediated by its agonist effect on either the GLP-1 or the glucagon receptor [11,12], albeit with lower affinity than the native ligands for each receptor. While the agonist properties of OXM toward these receptors are widely established, there has been speculation in the field about the possible existence of a specific high affinity receptor for OXM [5,10]. Structurally, OXM shares its N-terminal 28 residues with glucagon and its entire sequence with the C-terminal 37 amino acids of glicentin (Figure 1). As with other proglucagon-related peptides, the sequence for OXM is highly conserved

FUTURE SCIENCE part of fsg



Peptide	Amino acid sequence
Glucagon (1–28)	H-S-Q-G-T-F-T-S-D-Y-S-K-Y-L-D-S-R-R-A-Q-D-F-V-Q-W-L-M-N-T
OXM (1–37)	H-S-Q-G-T-F-T-S-D-Y-S-K-Y-L-D-S-R-R-A-Q-D-F-V-Q-W-L-M-N-T-K-R-N-R-N-N-I-A
OXM (3–37)	Q-G-T-F-T-S-D-Y-S-K-Y-L-D-S-R-R-A-Q-D-F-V-Q-W-L-M-N-T-K-R-N-R-N-N-I-A
pE-OXM (3–37)	pE-G-T-F-T-S-D-Y-S-K-Y-L-D-S-R-R-A-Q-D-F-V-Q-W-L-M-N-T-K-R-N-R-N-N-I-A
OXM (4–37)	G-T-F-T-S-D-Y-S-K-Y-L-D-S-R-R-A-Q-D-F-V-Q-W-L-M-N-T-K-R-N-R-N-N-I-A

Figure 1. Structure of proglucagon and the proglucagon-derived peptides resulting from tissue-specific proteolytic processing. The amino acid sequence of glucagon, as well as that of intact OXM 1–37 and the three OXM catabolites (3–37, pE 3–37 and 4–37) measured in the IA–MS assay described here are provided.
MPGF: Major proglucagon fragment; OXM: Oxyntomodulin; pE: Pyroglutamic acid.

across mammalian species. In addition, GLP-1, GLP-2, glucagon and OXM are all substrates of DPP-4 [13], an N-terminal dipeptidase responsible for their deactivation. Because only two amino acids are removed, an immense burden is placed on immunoassay methods to allow differentiation between active and inactive forms. Moreover, since the proglucagon peptides are formed by proteolytic action, antibodies specific for each C-terminal neoepitope are also needed. Despite these challenges, reliable and specific sandwich ELISA methods have been established for most proglucagon hormones. An assay for glucagon by Sloan and coworkers provides a representative example [14]. Unfortunately, obtaining a reliable ELISA assay for OXM has been a more difficult challenge. This problem was highlighted in a recent report by Bak *et al.* who compared three commercial ELISA kits and concluded that none of the methods provided reliable determination of OXM levels [15]. It also accounts for the wide range in reported literature values for circulating OXM levels and why some researchers have chosen the term 'OXM-like immunoreactivity' to describe its concentrations [16].

Despite the inherent specificity advantage of MS-based detection, reports using MS to quantify OXM have been slow to appear in the literature attributed to the need for low pg/ml sensitivity. In 2012, Halquist *et al.* described a 2D-LC–MS/MS assay for quantification of OXM in rat plasma [17]. The method, which had a range of 1–1000 ng/ml, was used to enable pharmacokinetic analysis following intravenous injection of recombinant OXM. Although reliable quantification was achieved, the assay lacked sufficient sensitivity to detect endogenous OXM. Recently, the first reported levels for endogenous OXM obtained by MS appeared in a publication by Lee *et al.* who used immunoaffinity (IA) enrichment with nano-electrospray ionization and 2D-LC–MS/MS to deliver an assay for OXM having an LLOQ of 3.5 pg/ml [18]. The method, which requires a sample volume of 0.5 ml, is part of a 4-plex panel combining OXM and glucagon with active and inactive GLP-1. An assay for the latter two analytes was also previously published by this group [19].

In this report, we show a further illustration of IA-MS to quantify endogenous OXM in human plasma and provide data which corroborate the results of Lee and coworkers. An important distinction of the current method is that LC–MS analysis occurred under high-resolution detection (R = 70,000) in

full scan mode using an orbitrap mass spectrometer. Although the method has a slightly elevated LLOQ (7.8 pg/ml), of analytical significance is the fact that operation occurred at a higher flow rate (400 µl/min) using narrow bore chromatography (2.1 mm ID) and without 2D-LC. Because similar LOQs were obtained for identical sample volumes, these results challenge previous beliefs about the comparative sensitivity of high-resolution methods for peptide quantification, particularly without the use of selected reaction monitoring (SRM) detection [20].

To date, a more complete understanding of the role of OXM has been impeded by a lack of analytical tools to sufficiently characterize OXM and its catabolites. An important feature of high-resolution detection is the ease for multi-analyte analysis. In the present method, OXM 1–37 was detected along with three N-terminal catabolites initiated by DPP-4 mediated hydrolysis. Owing to the conserved sequence for OXM across mammalian species, the assay was directly applied to measure OXM and its N-terminal catabolites in both rat and human plasma. The data obtained allow for a kinetic comparison of the rates of catabolism between species and add to the growing base of knowledge for OXM.

Materials & methods

Materials & reagents

Sodium chloride, octyl β-D-glucopyranoside (OGP) and formic acid (88%) were purchased from Sigma-Aldrich (MO, USA). Tris-HCl buffer (1 M, pH 7.5), Invitrosol™ LC–MS protein solubilizer and Dynabeads MyOne streptavidin T1 magnetic beads were obtained from Invitrogen (CA, USA). HEPES buffer (1 M) was from Thermo Scientific (IL, USA). Anhydrous DMSO was purchased from Life Technologies (CA, USA). Phosphate buffered saline was obtained from GE Healthcare Life Sciences (UT, USA). RIPA lysis buffer, DPP-4 inhibitor and acetonitrile were from EMD Millipore (MA, USA). cOmplete, Mini, EDTA-free protease inhibitor cocktail tablets were purchased from Roche Diagnostics (IN, USA). P800 blood collection tubes were obtained from Becton, Dickinson and Company (NJ, USA). Innovative grade US origin guinea pig and Sprague Dawley rat plasma (K_2EDTA) were from Innovative Research, Inc. (MI, USA). With the exception of matrix utilized in the OXM degradation studies, all plasma was thawed and incubated in BD P800 blood collection tubes for 15 min at room temperature with rotation to introduce DPP-4, esterase and other protease inhibitors. The pretreated plasma was then stored at -70°C until use. Unless stated otherwise, all references to guinea pig and rat plasma refer to P800 pretreated K_2EDTA plasma. Human K_2EDTA and K_2EDTA P800 plasma was obtained from subjects enrolled in the Eli Lilly volunteer blood donation program. All subjects gave informed consent. Isopropanol, HPLC grade water containing 0.1% (v/v) formic acid and HPLC grade acetonitrile containing 0.1% (v/v) formic acid were purchased from Fisher Scientific (NJ, USA). The OXM peptides 1–37, 3–37 and 4–37 as well as their corresponding stable isotope-labeled internal standards (each containing one [$^{13}C_9$, ^{15}N]-labeled phenylalanine and two [$^{13}C_6$, ^{15}N]-labeled leucine residues) were synthesized by CPC Scientific, Inc. (CA, USA) and were characterized by LC–MS and amino acid analysis. All solutions were prepared using deionized water purified by an EMD Millipore Milli-Q Plus ultra-pure water purification system (MA, USA).

Preparation of standard curve & quality control samples

OXM 1–37, 3–37 and 4–37 stock solutions were prepared individually in DMSO at concentrations of 710, 600 and 920 µg/ml, respectively and were stored at -70°C. These stock solutions were then used to prepare a combined standard working solution in DMSO consisting of 10 µg/ml of each peptide. Next, the combined standard working solution was used to prepare seven standard curve spiking solutions in DMSO (937.5, 1875, 3750, 7500, 15,000, 30,000 and 60,000 pg/ml for the human plasma assay and 1875, 3750, 7500, 15,000, 30,000, 60,000 and 120,000 pg/ml for the rat plasma assay). For the human plasma assay, 10 µl of standard curve spiking solution was added to 1190 µl of guinea pig plasma surrogate matrix to generate a seven point standard curve ranging from 7.8 to 500 pg/ml of OXM 1–37, 3–37 and 4–37. For the rat plasma assay, 7.5 µl of standard curve spiking solution was added to 892.5 µl of rat plasma control matrix to generate a seven point standard curve ranging from 15.6 to 1000 pg/ml of OXM 1–37, 3–37 and 4–37. Two separate standard curves were then subjected to IA enrichment as described below and analyzed at the beginning and end of each LC–MS run.

The individual peptide stock solutions were also used to prepare a combined quality control (QC) working solution in DMSO consisting of 10 µg/ml each of OXM 1–37, 3–37 and 4–37. The combined QC working solution was then used to prepare six QC spiking solutions in DMSO (1872, 5760, 15600, 28,800, 82,800 and 120,000 pg/ml for the human plasma assay and 3750, 11,280, 30,000, 76,800, 180,000 and 240,000 pg/ml for the rat plasma assay). For the human plasma assay, 15 µl of the 1872, 5760

Table 1. Interday accuracy (%RE) and precision (%CV) statistics for oxyntomodulin 1–37, 3–37 and 4–37 in human plasma.

QC level	LLOQ QC[†]	QC1[†]	END[‡]	QCGM[§]	QC2[§]	QC3[§]	ULOQ QC[†]
1–37:							
Theoretical (pg/ml)	7.81	32.1	8.05	73.1	128	353	500
Interday mean (pg/ml)	8.75	31.5	8.05	74.1	137	368	525
Interday %RE	12.0	-1.7	N/A	1.4	6.9	4.1	5.0
Interday %CV	5.5	4.5	7.1	3.9	3.5	3.3	2.8
n	18	18	18	18	18	18	18
3–37[¶]:							
Theoretical (pg/ml)	7.81	24.0	0	65.0	120	345	500
Interday mean (pg/ml)	9.38	26.8	0	70.0	132	372	526
Interday %RE	20.1	11.6	N/A	7.4	10.1	7.8	5.3
Interday %CV	8.9	5.4	N/A	3.4	2.8	2.4	2.5
n	18	18	18	18	18	18	18
4–37:							
Theoretical (pg/ml)	7.81	24.0	0	65.0	120	345	500
Interday mean (pg/ml)	8.94	26.4	0	71.2	136	374	534
Interday %RE	14.5	10.2	N/A	9.6	13.0	8.5	6.9
Interday %CV	6.6	4.6	N/A	3.1	2.3	2.2	3.2
n	18	18	18	18	18	18	18

[†]Prepared in guinea pig plasma surrogate matrix.
[‡]Pooled human plasma.
[§]Prepared in pooled human plasma.
[¶]Total 3–37 determined by summation of 3–37 and pE 3–37.
END: Endogenous; LLOQ: Lower limit of quantitation; QCGM: QC geometric mean; ULOQ: Upper limit of quantitation.

or 120,000 pg/ml QC spiking solution was added to 3585 µl of guinea pig plasma surrogate matrix to generate the LLOQ QC, QC1 and ULOQ QC samples, respectively. The QCGM, QC2 and QC3 samples were prepared by adding 15 µl of the 15,600, 28,800 or 82,800 QC spiking solutions, respectively, to pooled human K_2EDTA P800 plasma. The pooled human plasma was also used to create an endogenous (END) QC. The theoretical concentration of each analyte at each QC level is provided in Table 1. The assay intra- and inter-day accuracy (percent relative error,% RE) and precision (percent coefficient of variation,% CV) were assessed by subjecting six replicates of each QC sample to IA enrichment followed by LC–MS analysis on each of three days.

For the rat plasma assay, 10 µl of QC spiking solution was added to 2390 µl of rat plasma control matrix to generate the LLOQ QC, QC1, QCGM, QC2, QC3 and ULOQ QC samples. The theoretical concentration of each analyte at each QC level is provided in Table 2. It is important to note that the rat plasma pool used to create the standard curve and QC samples had no detectable levels of intact OXM or its catabolites, therefore the theoretical concentration was equal to the nominal spike concentration. The assay intra-day accuracy and precision was assessed by subjecting six replicates of each QC sample to IA enrichment followed by LC–MS analysis on a single day.

IA enrichment

For the human plasma assay, IA enrichment involved the addition of 500 µl of standards, quality control samples or human plasma unknowns to individual 2 ml microcentrifuge tubes (Costar). A sample volume of 350 µl was used for the rat plasma assay. To each sample, 5 µl of combined internal standard working solution, consisting of 25 ng/ml each of the 1–37, 3–37 and 4–37 stable isotope labeled internal standards prepared in DMSO, was added and the samples were vortex mixed. Next, either 1.3 ml (human plasma assay) or 1.4 ml (rat plasma assay) of IA enrichment buffer (25 mM Tris-HCl, 25 mM HEPES, 300 mM NaCl, 0.1% [v/v] OGP, pH 7.5) was added to each sample. Then, 10 µl of a 200 µg/ml solution of a biotin-labeled in-house generated antiglucagon monoclonal antibody [14] was added to each sample and IA enrichment was allowed to proceed overnight at 4°C with rotation. Following incubation, 50 µl of prewashed Dyna-

beads™ MyOne™ streptavidin T1 magnetic beads was added to each sample and the samples were incubated for 30 min at room temperature with rotation. The beads were then washed with 1 ml each of 1X RIPA lysis buffer, a buffer consisting of 25 mM Tris-HCl, 25 mM HEPES, 500 mM NaCl, 0.1% (v/v) OGP (pH 7.5) and deionized water. After the final water wash, bound analytes were eluted from the beads with the addition of 50 µl of 0.2% (v/v) formic acid/1X Invitrosol/10% (v/v) acetonitrile. The entire 50 µl was then subjected to LC–MS analysis as described below.

Chromatography

Chromatographic separation was achieved with a Leap PAL HTC autosampler (Leap Technologies, NC, USA), two Shimadzu Prominence LC-20AD pumps (MD, USA) and a Shimadzu Prominence CMB-20A controller. The analytical column used was a MACMOD HALO peptide ES-C18 (2.1 × 50 mm, 2.7 µm) held at 45°C and operated at a flow rate of 400 µl/min. Mobile phase A consisted of aqueous 0.1% formic acid and mobile phase B was 0.1% formic acid in acetonitrile. For the first minute following injection, mobile phase B was held at 10% and the effluent was delivered to waste. From 1.00 to 4.75 min, mobile phase B was ramped linearly from 10 to 35% and the effluent was delivered to the mass spectrometer. At 4.76 min, the mobile phase B composition was increased to 80% and was held until 5.85 min to wash the column. The effluent was delivered to waste starting at 5.00 min. At 5.86 min, the mobile phase B composition was returned to 10% for re-equilibration prior to the next injection. The total run time was 8.00 min. The needle was washed once with a 1:1 isopropanol: water (v/v) solution and then once with water. The injector port and sample loop were then washed twice with 100 µl of 1:1 isopropanol: water (v/v) followed by 2 × 100 µl of water.

MS

Mass spectrometric detection was accomplished using a Thermo Scientific Q Exactive mass spectrometer (CA, USA) equipped with an Ion Max source operated in positive ionization mode. The H-ESI vaporizer temperature was set to 400°C and the spray voltage to 3500 V. The sheath and auxiliary gas flow rates were set to 60 and 15 arbitrary units, respectively. The ion transfer tube temperature was 380°C and the S-lens RF level was set to 55. The analytes were detected using full scan MS with a scan range of 575–900 *m/z*, resolu-

Table 2. Intraday accuracy (%RE) and precision (%CV) statistics for oxyntomodulin 1–37, 3–37 and 4–37 in rat plasma.

QC level	LLOQ QC†	QC1†	QCGM†	QC2†	QC3†	ULOQ QC†
1–37:						
Theoretical (pg/ml)	15.6	47.0	125	320	750	1000
Intraday mean (pg/ml)	19.0	49.2	126	324	749	1034
Intraday %RE	21.4	4.8	1.1	1.3	-0.1	3.4
Intraday %CV	8.4	3.1	7.2	2.9	3.8	1.9
n	6	6	6	6	6	6
3–37‡:						
Theoretical (pg/ml)	15.6	47.0	125	320	750	1000
Intra-day mean (pg/ml)	17.6	49.0	126.0	327	760	1034
Intra-day %RE	12.4	4.3	1.1	2.3	1.3	3.4
Intra-day %CV	7.2	3.2	3.9	1.5	1.8	0.8
n	6	6	6	6	6	6
4–37:						
Theoretical (pg/ml)	15.6	47.0	125	320	750	1000
Intraday mean (pg/ml)	18.8	51.4	130	328	750	1042
Intraday %RE	20.3	9.4	4.0	2.4	0.0	4.2
Intraday %CV	10.0	2.7	3.1	1.0	2.0	1.8
n	6	6	6	6	6	6

†Prepared in pooled rat plasma.
‡Total 3–37 determined by summation of 3–37 and pE 3–37.
END: Endogenous; LLOQ: Lower limit of quantitation; QCGM: QC geometric mean; ULOQ: Upper limit of quantitation.

tion of 70,000 FWHM and scan rate of 7 Hz, which resulted, on average, in 12 points being collected across the chromatographic peak. The AGC target was set to 1×10^6 and the maximum inject time to 100 ms. Data were collected using Thermo Xcalibur software version 2.2.

Data analysis

Postacquisition data analysis was accomplished with Thermo Xcalibur Quan Browser software. The exact masses of the A + 1 through A + 4 isotopes for the three predominant charge states ($[M + 5H]^{5+}$, $[M + 6H]^{6+}$ and $[M + 7H]^{7+}$) of OXM 1–37 and its catabolites (3–37, pyroglutamic acid [pE] 3–37, 4–37) were extracted using a mass tolerance of 2 millimass units. These isotopes were chosen as they did not have corresponding interfering ions in blank plasma samples within the selected mass tolerance and were observed at sufficient abundance (>20% relative abundance). The extracted ions were then summed to calculate an extracted ion chromatogram used for quantification. Calibration curves were obtained by plotting the summed extracted ion chromatogram peak area ratio of each analyte to its respective internal standard versus concentration. A linear regression with $1/x^2$ and 1/x weighting was used for all analytes for the human and rat plasma assays, respectively, as they resulted in the best fit. Given that 3–37 was observed in both its non-pE and pE forms, and eluted at different retention times, peak areas were first determined using the Quan Browser software, were summed using Excel and the calibration curve was then constructed using SigmaPlot software version 12.5.

Human mixed meal tolerance test

Blood samples from 17 subjects were collected in BD P800 Vacutainer tubes containing K_2EDTA and a proprietary cocktail of DPP-4, esterase and other protease inhibitors. Subjects were asked to fast for a minimum of 10 h. Blood was then drawn upon their arrival (marked fasting) and at approximately 5, 45 and 90 min after eating a provided breakfast consisting of a 12 oz medium pulp orange juice and a sausage, egg and cheese biscuit sandwich. Subjects were asked to finish their meals within 15 min. The exact time of each blood draw was recorded. Following the separation of plasma from cells, plasma samples were stored at ≤-70°C prior to IA enrichment.

Rat mixed meal tolerance test

Male Sprague Dawley rats (NY, USA), 24–26 weeks old, were surgically cannulated by the vendor (femoral artery [FAC] for blood sampling and femoral vein [FVC] for IV dosing). All rats were housed individually on a 12 h light cycle and maintained on free choice chow (#2014, Harlan Teklad, IN, USA) and house auto water. Treatment and care of the animals was in accordance with regulations outlined in the USDA Welfare Act and conditions specified in "The Guide for Care and Use of Laboratory Animals." All procedures involving the care or use of animals in this study was reviewed and approved by the Eli Lilly Institutional Animal Care and Use Committee prior to the initiation of such procedures.

Eight overnight (~16 h) fasted rats were transported to the study room on the morning of the experiment and weighed for proper dose calculations. The femoral artery and femoral vein cannulas were checked for patency, attached to 18" long tether infusion lines and the rats were acclimated to the study boxes for approximately 20 min. At -15 min rats were injected with 5 mg/kg sitagliptin (n = 4) or vehicle (saline) (n = 4) into the FVC. At 0 min rats were orally (by gavage) challenged with 10 ml/kg of mixed meal (Nepro shake, Abbott Laboratories, OH, USA) just after a blood sample was withdrawn via FAC. All blood samples were collected into cold P800 tubes. Serial blood samples were withdrawn in the same manner via FAC at 5, 10, 15, 20, 30, 45 and 60 min post the oral mixed meal challenge. Plasma was separated from cells by refrigerated centrifugation, placed immediately into individually labeled tubes and stored frozen at -80°C until IA enrichment was performed as described above. Given the plasma volume requirement of 350 µl, the plasma from two rats was combined for each time point, therefore resulting in two plasma pools for each treatment condition (sitagliptin and vehicle).

OXM degradation in human & rat K_2EDTA & P800 plasma

Pooled human K_2EDTA plasma was thawed, warmed to 37°C and then split into two aliquots. One of the aliquots was incubated in a BD P800 tube to introduce DPP-4, esterase and other protease inhibitors. Both aliquots were spiked with 250 pg/ml of OXM 1–37 and then incubated with mixing at 37 °C for 0, 1, 5, 10, 15, 30, 45, 60, 120, 240 and 360 min. At each time point duplicate 500 µl aliquots were removed and placed in 2 ml microcentrifuge tubes containing DPP-4 inhibitor (50 µM final concentration) and complete mini EDTA-free protease inhibitor cocktail (1X final concentration). The aliquots were immediately frozen and then stored at ≤-70°C until IA enrichment was performed the following day as described above. The same procedure was used for rat plasma, with the exception that 350 µl aliquots were removed at each time point.

Results

Method development

Guinea pig plasma was chosen as the standard curve surrogate matrix for the human plasma assay given that guinea pig and human OXM do not share the same amino acid sequence. In addition, throughout the course of assay development, it was observed that the N-terminal glutamine of 3–37 was cyclizing to form pyroglutamic acid (pE), resulting in two forms of 3–37, the nonpyroglutamic acid (3–37) and pyroglutamic acid (pE 3–37) forms. Conversion of 3–37 to pE 3–37 was found to be most similar between guinea pig and human plasma upon addition of exogenous 3–37 compared with other surrogate matrices investigated (data not shown). Representative extracted ion chromatograms depicting both forms of 3–37 as well as 1–37 and 4–37 are shown in Figure 2. These chromatograms correspond to 125 pg/ml of 1–37, 3–37 and 4–37 spiked into guinea pig plasma and then subjected to IA enrichment. Although not shown, the stable isotope labeled versions of each peptide co-eluted with their nonlabeled forms.

The use of high-resolution accurate mass (HRAM) LC–MS for quantification is highlighted in Figure 3 using OXM 1–37 as an example. Figure 3A depicts the extracted ion chromatogram of unfortified guinea pig plasma and Figure 3B depicts the extracted ion chromatogram of guinea pig plasma fortified with 7.8 pg/ml of the 1–37 standard, which corresponds to the LLOQ. The extracted ion chromatograms were constructed by summation of the A + 1 through A + 4 isotopes of the $[M + 5H]^{5+}$, $[M + 6H]^{6+}$ and $[M + 7H]^{7+}$ charge states. The insets to each panel show the region of the mass spectrum containing the $[M + 6H]^{6+}$ ion and the asterisks indicate which isotopes were used for quantification. Figure 3A clearly shows that guinea pig plasma is devoid of endogenous 1–37, while the characteristic isotope pattern depicted in Figure 3B provides solid evidence that 1–37 can in fact be detected and quantified at low pg/ml levels using the HRAM approach described herein.

Assay validation

The interday accuracy (%RE) and precision (%CV) of the human plasma IA-MS assay was assessed by analyzing duplicate calibration curves (7.8–500 pg/ml for all analytes) and six replicates each of six quality control (QC) samples on three separate days (Table 1). An endogenous (END) QC comprised of an unspiked human plasma pool was analyzed over the three accuracy and precision batches and the average concentration of the 18 replicates was used to assign the endogenous analyte concentration. In this pool, only 1–37 was detectable and quantifiable. The average concentration determined for 1–37 (8.05 pg/ml) was then added to the nominal spike concentration for those QCs that were prepared in human plasma matrix in order to assign the theoretical concentration. The interday accuracy ranged from -1.7 to 12.0%, 5.3 to 20.1% and 6.9 to 14.5% for 1–37, 3–37 and 4–37, respectively. The interday precision ranged from 2.2 to 8.9% for all three analytes. Dilutional linearity was also assessed. For this investigation, the plasma from two individual subjects that had sufficiently high levels of OXM 1–37 was diluted two- and three-fold with surrogate matrix. After application of the corresponding dilution factor, the back-calculated concentration of the diluted samples agreed with the calculated concentration of the undiluted pool to within 10% (data not shown).

Chromatographic carryover was also assessed in each accuracy and precision run by injecting two zero samples (unfortified surrogate matrix with internal standard) after the highest standard curve samples. While the mean analyte response for 1–37, 3–37 and 4–37 in the first zero sample was upwards of 60, 20 and 35%, respectively, of the mean response of the LLOQ, the second zero sample resulted in a mean analyte response no greater than 11% for all three analytes.

Within-run selectivity was investigated in all three accuracy and precision batches in order to demonstrate the suitability of the surrogate matrix and internal standard used in the assay. Two control blanks, consisting of unfortified guinea pig plasma without internal standard, and two zero samples, were analyzed in each run prior to injection of the LLOQ standards. Placement of the samples prior to the LLOQ ensured that the results were not impacted by chromatographic carryover. For all three analytes, there was zero contribution of the surrogate matrix and internal standard to the analyte response.

Human mixed meal tolerance test

The validated IA–MS assay was applied to the analysis of human plasma samples from subjects in both the fasted and fed state in order to characterize and quantify the forms of OXM present under these conditions. Briefly, 17 subjects were asked to fast overnight for a minimum of 10 h. On the morning of the study, blood from each subject was drawn upon their arrival into BD P800 tubes and then each subject was asked to finish a provided breakfast within 15 min consisting of a 12 oz orange juice and sausage, egg and cheese biscuit sandwich. Blood was then drawn into BD P800 tubes at approximately 5, 45 and 90 min after completion of the meal. For logistical reasons, the above post-meal collection times are approximate, however the exact time of each blood draw was recorded (data not shown).

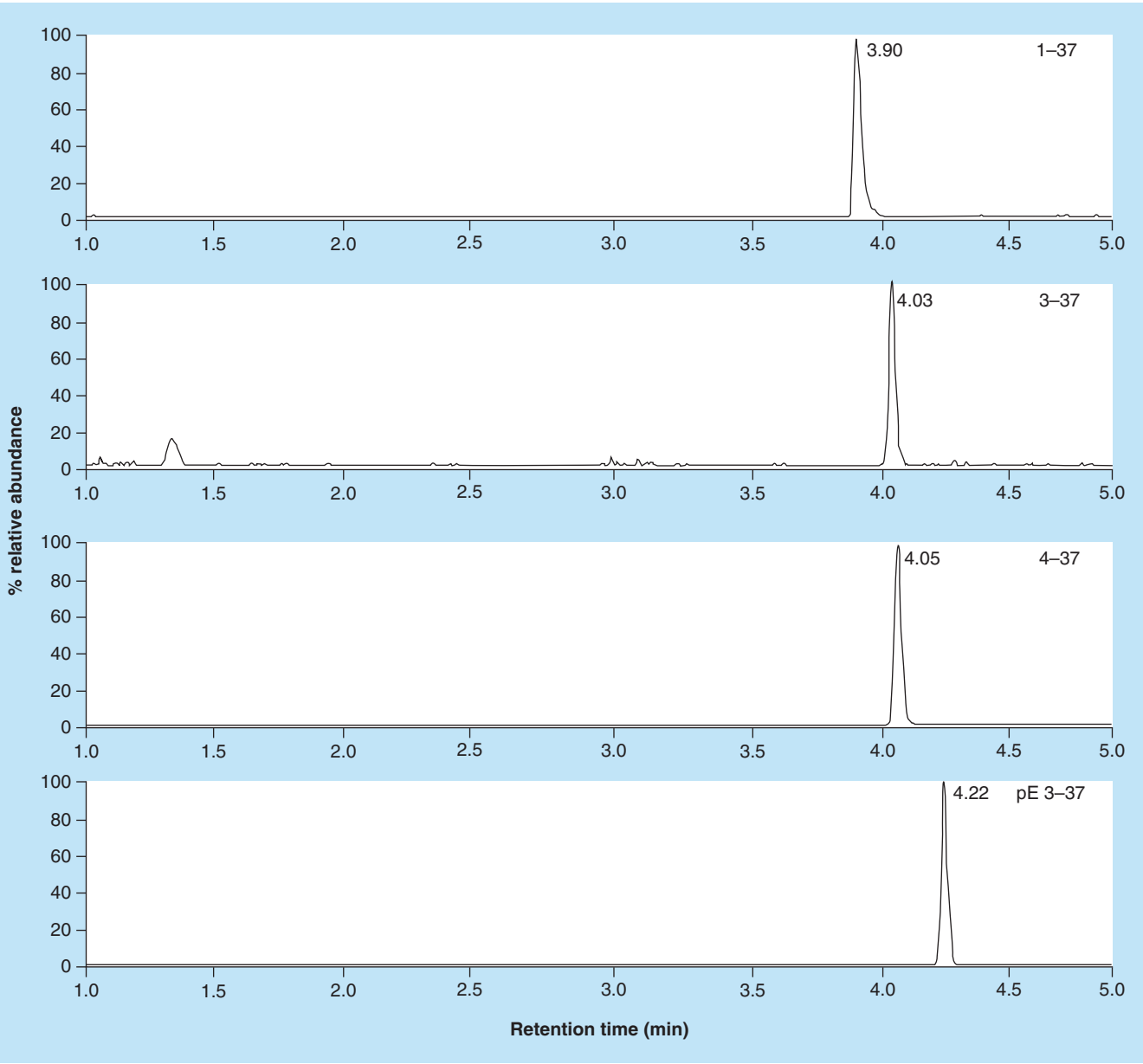


Figure 2. High-resolution full scan LC–MS analysis of OXM 1–37, 3–37, pE 3–37 and 4–37 following immunoaffinity enrichment of 500 µl of guinea pig plasma surrogate matrix fortified with 125 pg/ml each of OXM 1–37, 3–37 and 4–37. The nonenzymatic conversion of 3–37 to pE 3–37 occurred during sample handling.

For all subjects that participated in this study, the only form of OXM that was observed above the LLOQ of the assay was intact 1–37 (Figure 4). For the majority of volunteers, the concentration of 1–37 was lowest in the fasted state and increased upon consumption of the meal. For 10 of the 17 subjects, the fasting concentration was slightly below the LLOQ of 7.8 pg/ml; however these values were still included in Figure 4 as the detection of 1–37 could be confirmed by its characteristic isotope pattern. Subjects 6 and 15 had elevated levels of 1–37 in the fasted state. The results presented in Figure 4 also reveal that the peak concentration of 1–37 did not always occur at the same time post meal and the excursion profile also varied from individual to individual.

Rat mixed meal tolerance test

In addition to characterizing the forms and levels of human plasma OXM following a mixed meal tolerance test (MMTT), an MMTT with a cohort of male Sprague Dawley rats was also performed to determine potential differences between clinical and preclinical species. Prior to performing the rat MMTT, a one day accuracy and precision run was conducted. For the rat plasma assay, a pool of P800 K_2EDTA rat plasma that

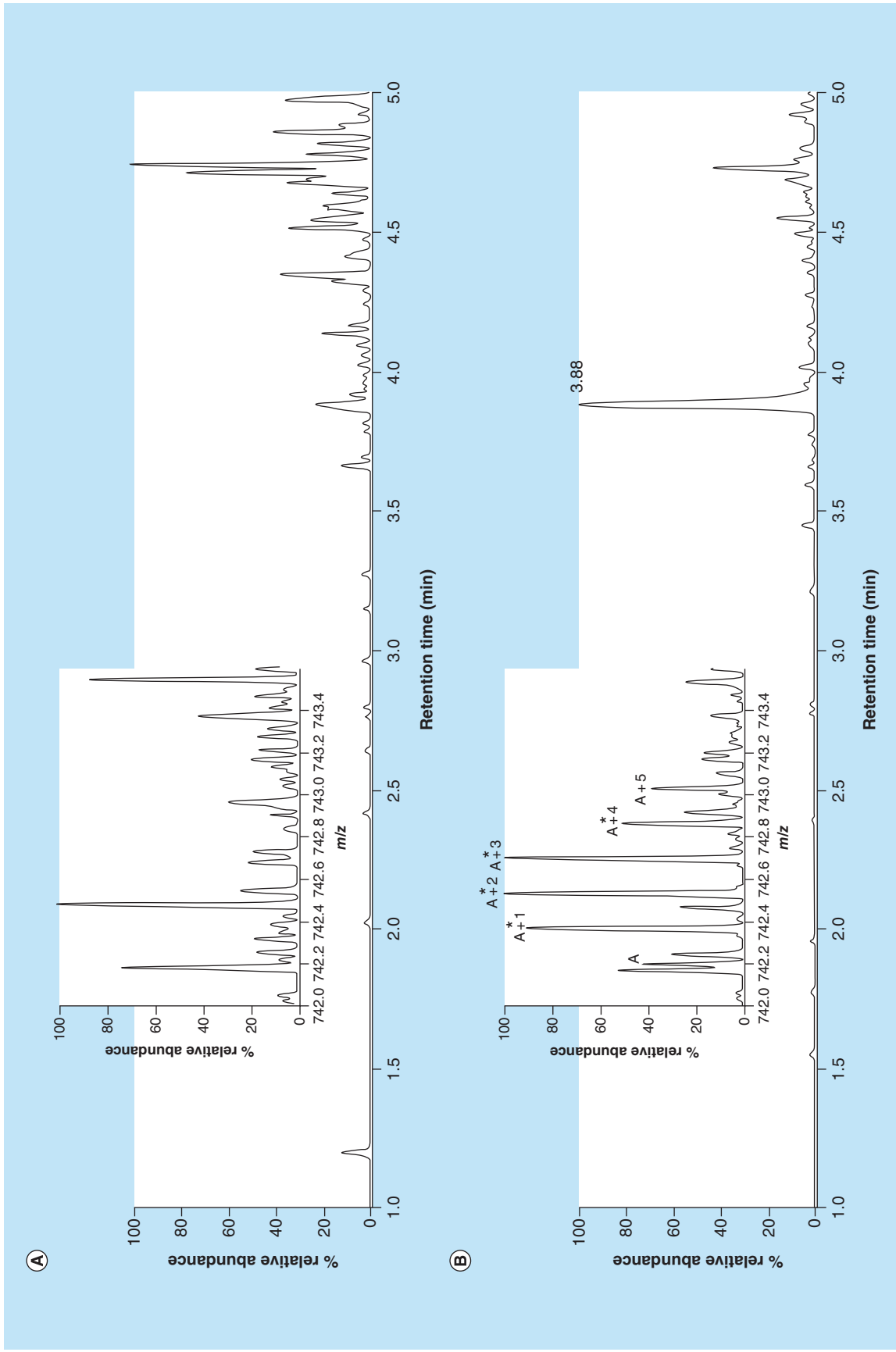


Figure 3. Extracted ion chromatograms for (A) unfortified guinea pig plasma surrogate matrix and (B) guinea pig plasma surrogate matrix fortified with 7.8 pg/ml OXM 1–37, which corresponds to the lower limit of quantification of the IA–MS assay. The insets to each panel show the region of the mass spectrum containing the $[M + 6H]^{6+}$ charge state and are the composite sum of the spectra collected across the chromatographic peak. The asterisks indicate the isotopes that were used for quantification.

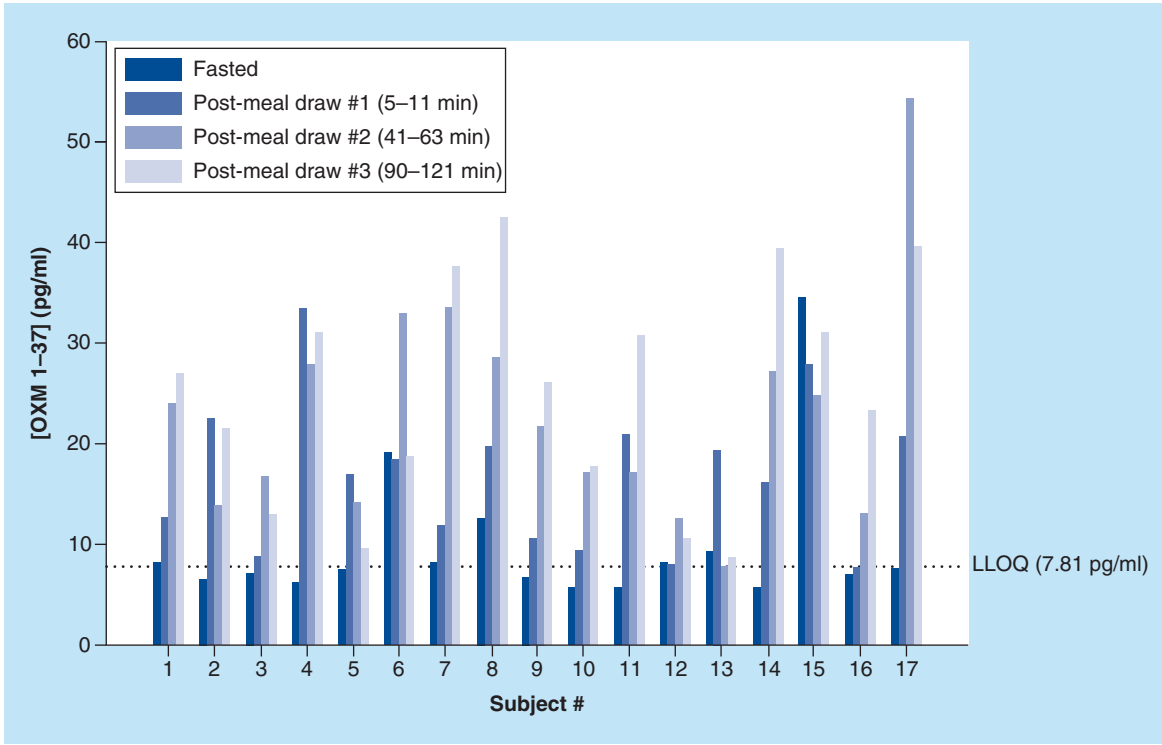


Figure 4. Concentration of OXM 1–37 in human P800 plasma collected from 17 subjects following fasting for a minimum of 10 h and at approximately 5, 45 and 90 min post prandial. The range of collection times are noted in parentheses. All subjects were provided the same meal and were asked to complete the meal within 15 min.

was devoid of endogenous intact OXM and its catabolites was chosen as the control matrix. While guinea pig plasma could have been used, the percent conversion of 3–37 to pE 3–37 was found to vary between the guinea pig and rat plasma and therefore the rat plasma pool was deemed a more appropriate choice. The intraday accuracy (%RE) ranged from 0 to 21.4% and the intraday precision (%CV) ranged from 1.0 to 10.0% for all three analytes (Table 2).

For the MMTT, Sprague Dawley rats were dosed with either vehicle (saline) (n = 4) or 5 mg/kg of the DPP-4 inhibitor, sitagliptin (n = 4) and were then dosed with 10 ml/kg Nepro nutrient mixture. Plasma samples were collected prior to dosing the Nepro nutrient mixture and 5, 10, 15, 20, 30, 45 and 60 min post dose. Ideally, the OXM levels in individual rats would have been determined, however given the requisite volume for IA enrichment (350 µl), the plasma had to be pooled. For both the vehicle and sitagliptin treated animals, two pools were created with each pool consisting of two rats each. This pooling scheme applied to all time points post dose with Nepro, however for the fasting time point the plasma from all four rats for a given condition had to be pooled given the blood volume constraints.

The data from this experiment are shown in Figure 5 with Figure 5A displaying results for the vehicle treated rats. Under these conditions, the major form of OXM that was observed was 3–37, which constituted approximately 70% of total OXM. OXM 1–37 was also observed, albeit to a lesser extent (~30% of total OXM), whereas 4–37 was observed at levels below the LLOQ. Both 1–37 and 3–37 increased following dosing with the Nepro nutrient mixture with the maximum concentration of 3–37 (~75 pg/ml) occurring 20 and 30 min post dose for pool 1 and 2, respectively. The maximum concentration of 1–37 (~25 pg/ml) occurred at 20 min post dose for both pools.

The results for rats that had been treated with the DPP-4 inhibitor, sitagliptin, are depicted in Figure 5B. For both pools, the only form of OXM that was observed above the LLOQ was 1–37, which provides further confirmation that the conversion of 1–37 to 3–37 is mediated by the DPP-4 enzyme. As expected, and consistent with rats that had been treated with vehicle, 1–37 was observed to increase following dosing with the Nepro nutrient mixture and then decreased as a result of clearance. The maximum concentration of 1–37 (~133 pg/ml) occurred at 30 min post dose for both pools.

OXM degradation in human & rat K_2EDTA & P800 plasma

Given that different forms of plasma OXM were observed between human and rat following the

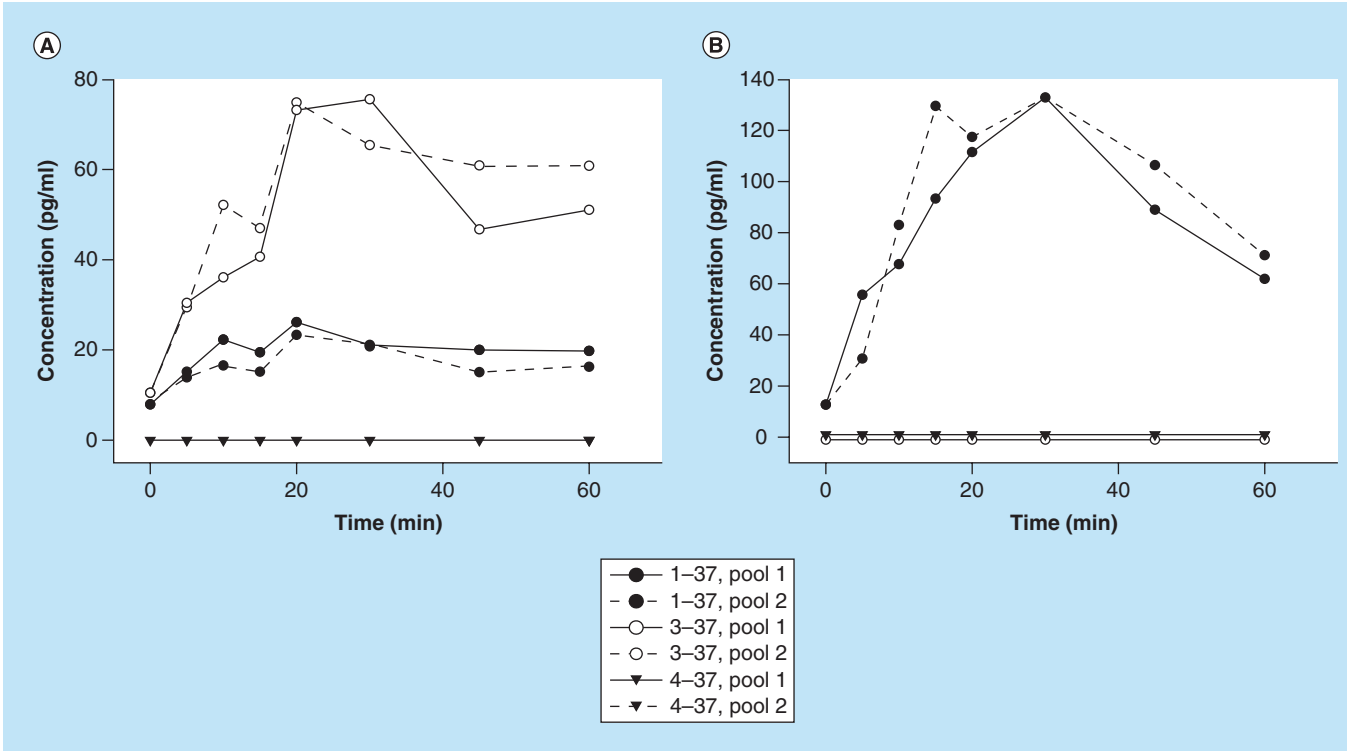


Figure 5. Fasting and postprandial concentrations of OXM 1–37, 3–37 and 4–37 in rat P800 plasma. Sprague Dawley rats were fasted overnight (16 h) and the following morning were dosed intravenously with either **(A)** vehicle (saline) or **(B)** 5 mg/kg sitagliptin (marked -15 min). At 0 min, blood was collected in BD P800 tubes (fasting) and then 5, 10, 15, 20, 30, 45 and 60 min post dose (oral gavage) with 10 ml/kg of Nepro nutrient mixture. The data for two plasma pools are shown with each pool consisting of plasma from n = 2 rats, with the exception of the fasting time point where the plasma from all four rats had to be combined due to volume limitations. The concentration of 3–37 was determined by summation of the pyroglutamic acid and nonpyroglutamic acid forms.

MMTT, we conducted an *in vitro* plasma 1–37 spike-in experiment to enable further investigation of these observations. This experiment utilized pooled human and rat K_2EDTA plasma that were each split into two aliquots. One aliquot was incubated in a BD P800 blood collection tube to introduce DDP-4 and protease inhibitors. Both aliquots were spiked with OXM 1–37 at a concentration of 250 pg/ml and were then incubated with mixing at 37°C. At 0, 1, 5, 10, 15, 30, 45, 60, 120, 240 and 360 min, duplicate aliquots were removed, DPP-4 and protease activity was quenched, and the samples were analyzed by IA-MS to characterize and quantify the forms of OXM.

The results from this experiment are depicted in Figure 6. For the aliquot of human K_2EDTA plasma that had been incubated in the P800 tubes prior to the addition of exogenous 1–37 (Figure 6A), the concentration of 1–37 (summation of 250 pg/ml spike and 12.5 pg/ml endogenous) remained constant throughout the entire 37°C incubation. This is consistent with the fact that the DPP-4 inhibitor introduced by the P800 tubes prevented the conversion of 1–37 to 3–37. It is worth noting that 3–37 was detected and quantified, however this was due to the conversion of endogenous 1–37 to 3–37 prior to incubation of the K_2EDTA plasma in the P800 tubes. The concentration of 3–37 remained constant during the entire course of the experiment.

Under the conditions where DPP-4 and protease inhibitors were not introduced to the human plasma prior to the experiment, conversion of 1–37 to 3–37 began to occur around 30 min post incubation, with approximately 145 pg/ml of 1–37 remaining after 360 min (Figure 6B). Mass balance, as determined by summation of 1–37, 3–37 and 4–37, was maintained until 240 min. Given that 4–37 was not detected suggests that cleave product(s) other than 4–37 may have formed, however their identity was not able to be determined in this particular experiment.

For the rat K_2EDTA P800 plasma, the concentration of 1–37 remained constant during the majority of the 37°C incubation, however in contrast to that observed for human plasma under these same conditions, 1–37 began to decrease between 60 and 120 min (Figure 6C). As expected, 3–37 and 4–37 did not increase in concentration during this time, thus the conversion of 1–37 was likely mediated by a protease not inhibited by the P800 tube inhibitor cocktail.

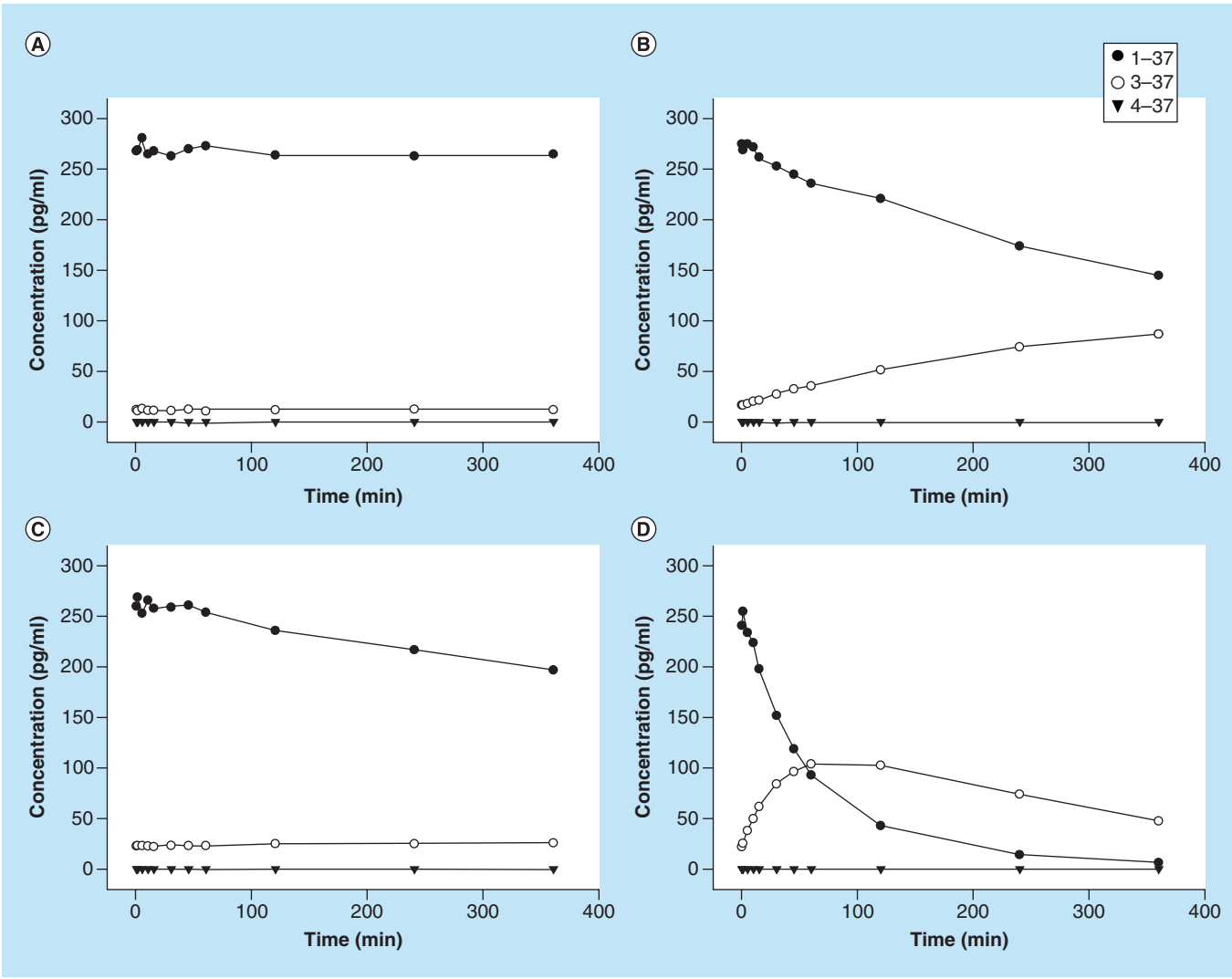


Figure 6. Concentrations of OXM 1–37, 3–37 and 4–37 in (A) human K_2EDTA P800 plasma, (B) human K_2EDTA plasma, (C) rat K_2EDTA P800 plasma and (D) rat K_2EDTA plasma fortified with 250 pg/ml of OXM 1–37 and incubated with mixing at 37°C for 0, 1, 5, 10, 15, 30, 45, 60, 120, 240 and 360 min. Each point represents the average of duplicate measurements. The concentration of 3–37 was determined by summation of the pyroglutamic acid and nonpyroglutamic acid forms.

When the rat plasma was devoid of DPP-4 and protease inhibitors, the conversion of 1–37 to 3–37 occurred almost immediately and continued rapidly (Figure 6D). At 360 min post incubation, there was essentially zero 1–37 remaining and at 120 min the concentration of 3–37 also began to decrease at which time 4–37 did not increase. As was the case with the human K_2EDTA plasma, the identity of any potential cleavage products could not be determined in this particular experiment.

Discussion

OXM biology & interspecies translation

In this work a robust and sensitive IA-MS assay was validated for OXM forms 1–37, 3–37 and 4–37 in both human and rat plasma. Although SRM detection on a triple quadrupole mass spectrometer has been historically used for peptide quantification [21], we opted for full scan HRAM detection using orbitrap MS to allow facile observation of OXM catabolites. As shown herein, OXM 1–37 is converted to its 3–37 form by DPP-IV type cleavage at the N-terminus, a process which was found to occur more readily in rats than in humans. Our data indicate that in rat, plasma OXM 1–37 levels peak roughly 30 min after ingestion of a meal (Figure 5) and that 3–37 conversion appears to be DPP-4 mediated based on the data acquired using sitagliptin treatment. Although 4–37 was not observed above the LLOQ in the study reported here, this catabolite has been observed in other experiments previously conducted in rat (data not shown). In contrast, the results from human volunteers (Figure 4) reveal that the kinetics for OXM 1–37 elevation varied,

with some donors rapidly reaching peak levels within 5–10 min while others had a more delayed response peaking between 90 and 120 min after food ingestion. The physiological relevance of these observations is unknown, but may reflect the various physiological states of the donors in this small study. OXM plays key roles in the control of food intake, energy expenditure and glucose metabolism, and the varying kinetic profiles for OXM 1–37 observed for the donors may reflect the body's need to modulate these responses to nutrient intake. The observation that some subjects have elevated baseline (fasted) levels of OXM 1–37 with modest increases post prandially may also lend support to the application of OXM 1–37 as a biomarker for the stratification of patients who could potentially benefit from OXM-based agonist therapies. However, the elevated baseline levels could also be attributed to these subjects not having fasted as requested. These results are consistent with those reported by Lee *et al.* [18]. In their analysis of eight volunteers, the average pre- and post-prandial concentration of OXM 1–37 was 6 and 36 pg/ml, respectively.

The establishment of this IA-MS approach specific to the various forms of circulating OXM will be useful in studies aimed at increasing the understanding of the various physiological roles for OXM in rodents and humans. As reported by Bak and coworkers [15], there is a paucity of reliable immunoassays for OXM, and in order to support large scale clinical studies, rapid and sensitive assays for OXM are needed. To that end, we recently used this IA–MS assay to cross-validate a new immunoassay with high specificity toward OXM 1–37. Results from these studies showed high correlation between the two assays [Manuscript in Preparation], and further illustrate the utility of a method such as this for orthogonal assay validation.

In the present study a comparison between rat and humans was conducted to more fully understand OXM biology and to assist with clinical translation. A number of significant differences were observed that should be noted. On average, rat concentrations for OXM 1–37 were approximately 2.5-fold higher than found in man. In addition, as mentioned previously, rats exhibited greater DPP-4 activity as documented by the *in vitro* incubation results presented in Figure 6. The reason for this difference is not known, but is likely driven by a greater specific DPP-4-like activity in rats. This difference carried over *in vivo* as OXM 3–37 was the most abundant circulating form in rat. In contrast, only intact OXM was detected in humans. A complicating factor in making this comparison was the greater tendency for OXM 3–37 to cyclize to OXM pE 3–37 in human plasma. Thus, the failure to detect OXM 3–37 in human plasma could have resulted from signal dilution by the two forms. While the origin of this difference is not completely understood, it likely reflects residual, species-specific glutaminyl cyclase activity and/or spontaneous cyclization, known to be favorable for glutamine [22]. The fact that guinea pig plasma gave comparable cyclization to human plasma was a factor taken into account when selecting it as the surrogate matrix for the assay.

IA sample preparation

During the course of assay development, there were several key learnings that enabled low pg/ml sensitivity to be achieved. For example, the wash conditions were optimized to include a 1X RIPA buffer wash prior to the high salt wash, which reduced the protein background and improved assay sensitivity. Preparation of the standard curve was also optimized such that standard curve spiking solutions were first created in DMSO and then spiked into the plasma surrogate matrix as opposed to serial dilutions directly into plasma, which had initially resulted in significant peptide loss. In addition, the introduction of Invitrosol LC–MS protein solubilizer to the elution solvent during IA enrichment helped to achieve a tight magnetic bead pellet following elution, which led to minimal bead carryover into the eluate and improved LC robustness (i.e., eliminated column clogging). We also speculate that the Invitrosol solubilizer helped to stabilize OXM in solution and minimize nonspecific losses in the polypropylene autosampler plates.

An essential part of any biomarker assay involving a surrogate matrix is to demonstrate assay parallelism [23]. Results from two experiments attest to the qualification of guinea pig plasma as an acceptable surrogate matrix. First, the data in Table 1 indicates that good bias and precision were obtained for QC pools prepared in human plasma. In a second experiment, plasma from two human subjects having high OXM 1–37 levels was diluted with the surrogate matrix up to threefold with acceptable results. This latter dilutional linearity experiment provides key evidence for the viability of the assay as it is in the range where most OXM 1–37 levels were observed in humans.

Analytical considerations for IA-MS

Several options exist for MS-based protein quantification, extending well beyond traditional targeted MRM methods on triple quadrupole mass spectrometers (TQMS). Indeed, the utility of HRAM for small molecule bioanalysis has been widely documented [24–26] and HRAM is now being used for peptide quantification with quadrupole time-of-flight (Q-ToF) as the primary tool. Initial comparisons between TQMS and Q-ToF showed the promise of HRAM, but clear supe-

riority in favor of TQMS largely due to an advantage in duty cycle [20,27]. This picture has now changed in part due to a series of publications by Garofolo and coworkers demonstrating that HRAM conducted in association with SRM can outperform TQMS when proper attention is paid to optimizing the extraction window size, summing multiple charge states and isotopes, and avoiding the integration of peaks contaminated with interfering ions [28,29]. The second paper, which performed a head-to-head comparison to TQMS, showed that HRAM performed equal to or better than TQMS for each of the six peptides studied [29]. Interestingly, the fact that SRM detection was needed to achieve competitive performance suggests that matrix background was still limiting despite the use of 30,000 mass resolution (FWHM).

For the present study, orbitrap MS was selected as the instrumental platform for intact analysis of OXM and its N-terminal catabolites. In contrast to the work by Garofolo *et al.*, selectivity was achieved in full scan mode at elevated mass resolution (70,000 FWHM) without MS/MS. Our work builds on initial work by Ruan and colleagues who demonstrated intact quantification of lysozyme (14.3 kDa) by orbitrap MS in the full scan mode, also without MS/MS [30]. In their paper, it was shown that the mass resolution needs to be at least four-times the molecular weight of the target peptide leading to an applied resolution of 60,000 FWHM for lysozyme. The importance of elevated resolution is significant, because it allows adjacent analyte isotopes to be more than baseline resolved thereby providing sufficient space between the isotopes to detect and discriminate against background ions.

As cited previously, LC–MS using HRAM detection in the full scan mode offers several advantages including simplified method development, access to both qualitative and quantitative information, and retrospective analysis [26]. Moreover, because of known inefficiencies associated with MRM-based peptide quantification, including precursor ion signal dilution from multiple charging and further signal loss stemming from the multitude of peptide product ion fragmentation pathways, we believe that sensitivities exceeding TQMS can be achieved provided sufficient resolution is applied in the full scan mode. While the resolution used in this study exceeded the analyte molecular weight by more than 15-fold, it has been our experience that a mass resolution of 70,000 or higher is routinely needed to analyze protein biomarkers in serum or plasma without resorting to MS/MS even when immunoprecipitation is used for sample enrichment. We maintain that this is simply a consequence of plasma complexity and the large volumes (0.1–1 ml) currently required for low pg/ml biomarker analysis. While higher resolution is possible using orbitrap mass spectrometers, the inverse relationship with acquisition rate ultimately limits the number of points that may be acquired across a chromatographic peak. We anticipate that mass resolution exceeding 100,000 will be routinely deployed as further refinements are made to this technology platform.

The viability of HRAM for low pg/ml peptide quantification, particularly without MS/MS, requires careful attention to the three factors cited above by Garofolo *et al.* In the present method, quantification was achieved by summing the A + 1 through A + 4 isotopes for the $[M + 5H]^{5+}$, $[M + 6H]^{6+}$ and $[M + 7H]^{7+}$ charge states. In addition, each isotope was carefully examined to produce customized extraction windows selected to optimize S/N. Blank matrix extracts were used in making these assessments. Utilizing this approach, an LLOQ of 7.8 pg/ml was achieved for the human plasma assay. Although a slightly elevated LLOQ of 15.6 pg/ml was achieved for the rat plasma assay, a reduced sample volume was employed (350 vs 500 µl). Despite this achievement, a drawback to the present method is that it did not allow the fasting levels to be quantified for all human subjects. It is important to recognize that the method employed a 2.1 mm ID column operated at 400 µl/min without 2D LC for on-line sample enrichment and without UHPLC. While this simplified approach to LC favors robustness, future implementation of smaller bore chromatography could be used to achieve LLOQs and/or to reduce sample volume. The fact that OXM 1–37 was detected in the samples from all fasted human subjects under the present conditions speaks to the potential strength of HRAM for future development of IA-MS biomarker methods.

Although the present article emphasizes IA sample preparation, it is important to recognize a growing trend towards non-IA methods for peptide biomarker quantification [31]. A recent example is the work by Howard *et al.* who quantified glucagon in human plasma (LLOQ 25 pg/ml) using an approach based on LC–MS/MS and SPE [32]. Despite the potential viability of such methods, the extraordinary method development needed to minimize the impact of background and ion suppression should not be underestimated.

Conclusion

OXM is a 37 amino acid peptide that is derived from tissue-specific processing of proglucagon in the intestinal L-cells. OXM is a dual agonist of the GLP-1 and glucagon receptors and has garnered interest as a potential therapeutic agent for obesity and Type 2 diabetes. In this paper, IA enrichment coupled with

HRAM on an orbitrap mass spectrometer enabled a robust and sensitive method to be established for the reliable quantification of plasma OXM and its catabolites. Utilizing this approach, key differences between human and rat OXM secretion in response to nutrient ingestion were ascertained, which could potentially further our understanding of OXM biology and assist with clinical translation. Furthermore, the benefits of utilizing HRAM in full scan mode on an orbitrap mass spectrometer at elevated resolution for peptide quantification have been highlighted.

Future perspective

From the biological perspective, the current ability to quantify OXM and its breakdown products will lead to new understandings about this peptide hormone and will assist current efforts aimed at exploiting the properties of OXM for therapeutic benefit. Specifically, the ability to measure OXM concentrations will enable studies to better characterize entero-endocrine regulatory mechanisms that control the response to nutrient ingestion. For instance, it will be important to determine differences in the abilities of various macronutrients to stimulate OXM secretion and whether different stimuli always lead to a comparable release of GLP-1. Additional studies evaluating the capability of established GLP-1 secretagogues to also stimulate OXM secretion may help prioritize which pharmacological mechanisms to pursue as potential new therapies for metabolic disease.

From the analytical side, the current IA-MS method for OXM demonstrates the ability to perform full scan HRAM for low pg/ml peptide quantification, without resorting to nanoelectrospray ionization, capillary LC or MS/MS. Importantly, this will lead to method simplification and reduce the expertise required to perform high end IA-MS analysis, including multiprotein quantification. At present, an unfortunate downside to the current approach is that extensive data manipulation is required both in the development and execution of HRAM methods. This issue was previously cited as a concern for HRAM methods performed by Q-ToF [28]. We believe this to be a temporary issue which can be addressed by software tools customized for quantitative HRAM applications. Finally, we believe that full scan HRAM methods will be increasingly deployed as larger and larger peptides are measured intact owing to the plurality of charge states encountered as size increases.

Financial & competing interests disclosure

The authors of this manuscript are all employees of Eli Lilly and Company and have no financial ties or connection to any company or party supplying equipment or reagents used in the experiments conducted. The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.

Ethical conduct of research

The authors state that they have obtained appropriate institutional review board approval or have followed the principles outlined in the Declaration of Helsinki for all human or animal experimental investigations. In addition, for investigations involving human subjects, informed consent has been obtained from the participants involved.

Executive summary

- Oxyntomodulin (OXM) is one of several peptide hormones resulting from post-translational processing of proglucagon. Due to the high sequence similarity of OXM with other proglucagon-derived peptides, MS is an attractive tool for quantification.
- A validated method is presented which couples immunoaffinity enrichment with high-resolution accurate mass (HRAM) on an orbitrap mass spectrometer to quantify intact OXM 1–37 as well as its catabolites, 3–37, pE 3–37 and 4–37 in human and rat plasma both pre- and post-prandially.
- Data were obtained providing the first evidence for differences in the levels and forms of OXM between preclinical and clinical species.
- Full scan HRAM using elevated mass resolution (R = 70,000 FWHM) offers a competitive approach to existing SRM-based methods for low-abundance peptide biomarker quantification.
- The relative simplicity of full scan HRAM offers several advantages for biomarker analysis including simplified method development, ease of multi-analyte detection and the ability for postanalysis data mining.

References

Papers of special note have been highlighted as:
• of interest; •• of considerable interest

1 Bataille D, Tatemoto K, Gespach C, Jörnvall H, Rosselin G, Mutt V. Isolation of glucagon-37 (bioactive enteroglucagon/oxyntomodulin) from porcine jejuno-ileum: Characterization of the peptide. *FEBS Lett.* 146(1), 79–86 (1982).

2 Dubrasquet M, Bataille D, Gespach C. Oxyntomodulin (glucagon-37 or bioactive enteroglucagon): a potent inhibitor of pentagastrin-stimulated acid secretion in rats. *Biosci. Rep.* 2(6), 391–395 (1982).

3 Troke RC, Tan TM, Bloom SR. The future role of gut hormones in the treatment of obesity. *Ther. Adv. Chromic Dis.* 5(1), 4–14 (2014).

• **Reviews the development of gut peptides as current and emerging therapies for the treatment of obesity.**

4 Bagger JI, Holst JJ, Hartmann B *et al.* Effect of oxyntomodulin, glucagon, GLP-1, and combined glucagon + GLP-1 infusion on food intake, appetite and resting energy expenditure. *J. Clin. Endocrinol. Metab.* 100(12), 4541–4552 (2015).

5 Drucker DJ. The role of gut hormones in glucose homeostasis. *J. Clin. Invest.* 117(1), 24–32 (2007).

6 Pocai A. Action and therapeutic potential of oxyntomodulin. *Mol. Metabol.* 3, 241–251 (2014).

7 Pocai A. Unraveling oxyntomodulin, GLP1's enigmatic brother. *J. Endocrinol.* 215, 335–346 (2012).

8 Tasyurek HM, Altunbas HA, Balci MK, Sanlioglu S. Incretins: their physiology and application in the treatment of diabetes mellitus. *Diabetes Metab. Res. Rev.* 30, 354–371 (2014).

9 Mojsov S, Heinrich G, Wilson IB, Ravazzola M, Orci L, Habener JF. Preproglucagon gene expression in pancreas and intestine diversifies at the level of post-translational processing. *J. Biol. Chem.* 261(25), 11880–11889 (1986).

10 Spreckley E, Graeme Murphy K. The L-cell in nutritional sensing and regulation of appetite. *Front. Nutr.* 2(23), 1–17 (2015).

11 Fehmann HC, Schweinfurth J, Wheeler MB, Boyd AE 3rd, Göke B. Stable expression of the rat GLP-1 receptor in CHO cells: activation and binding characteristics utilizing GLP-1(7–36)-amide, oxyntomodulin, exendin-4, and exendin (9–39). *Peptides* 15(3), 453–456 (1994).

12 Baggio LL, Huang Q, Brown TJ, Drucker DJ. Oxyntomodulin and glucagon-like peptide-1 differentially regulate murine food intake and energy expenditure. *Gastroenterology* 127(2), 546–558 (2004).

13 Zhu L, Tamvakopoulos C, Xie D *et al.* The role of dipeptidyl peptidase IV in the cleavage of glucagon family peptides: *in vivo* metabolism of pituitary adenylate cyclase activating polypeptide-(1–38). *J. Biol. Chem.* 278(25), 22418–22423 (2003).

14 Sloan JH, Siegel RW, Ivanova-Cox YT, Watson DE, Deeg MA, Konrad RJ. A novel high-sensitivity electrochemiluminescence (ECL) sandwich immunoassay for the specific quantitative measurement of plasma glucagon. *Clin. Biochem.* 45, 1640–1644 (2012).

15 Bak MJ, Wewer Albrechtsen N, Pedersen J *et al.* Specificity and sensitivity of commercially available assays for glucagon and oxyntomodulin measurement in humans. *Eur. J. Endocrinol.* 170(4), 529–538 (2014).

16 Le Quellec A, Kervran A, Blache P, Ciurana AJ, Bataille D. Oxyntomodulin-like immunoreactivity: diurnal profile of a new potential enterogastrone. *J. Clin. Endocr. Metab.* 74(6), 1405–1409 (1992).

17 Halquist MS, Sakagami M, Karnes HT. Determination of oxyntomodulin, an anoretic polypeptide, in rat plasma using 2D-LC–MS/MS coupled with ion pair chromatography. *J. Chromatogr. B* 903, 102–111 (2012).

18 Lee AYH, Chappell DL, Bak MJ *et al.* Multiplexed quantification of progluagon-derived peptides by immunoaffinity enrichment and tandem mass spectrometry after a meal tolerance test. *Clin. Chem.* 62(1), 227–235 (2015).

•• **First assay of any platform type to reliably quantify endogenous oxyntomodulin in human plasma.**

19 Chappell DL, Lee AYH, Castro-Perez J *et al.* An ultrasensitive method for the quantitation of active and inactive GLP-1 in human plasma via immunoaffinity LC–MS/MS. *Bioanalysis* 6(1), 33–42 (2014).

20 Dillen L, Cools W, Vereyken L *et al.* Comparison of triple quadrupole and high-resolution TOF-MS for quantification of peptides. *Bioanalysis* 4(5), 565–579 (2012).

21 Anderson L, Hunter CL. Quantitative mass spectrometric multiple reaction monitoring assays for major plasma proteins. *Mol. Cell. Proteomics* 5(4), 573–588 (2006).

22 Seifert F, Schulz K, Koch B *et al.* Glutaminyl cyclases display significant catalytic proficiency for glutamyl substrates. *Biochemistry* 48(50), 11831–11833 (2009).

23 Jones BR, Schultz GA, Eckstein JA, Ackermann BL. Surrogate matrix and surrogate analyte approaches for definitive quantitation of endogenous biomolecules. *Bioanalysis* 4(19), 2343–2356 (2012).

24 Ramanathan R, Jemal M, Ramagiri S *et al.* It is time for a paradigm shift in drug discovery bioanalysis: from SRM to HRMS. *J. Mass Spectrom.* 46(6), 595–601 (2011).

25 Xia YQ, Lau J, Olah T, Jemal M. Targeted quantitative bioanalysis in plasma using liquid chromatography/high-resolution accurate mass spectrometry: an evaluation of global selectivity as a function of mass resolving power and extraction window, with comparison of centroid and profile modes. *Rapid Commun. Mass Spectrom.* 25(19), 2863–2878 (2011).

26 Huang M-Q, Lin Z, Weng N. Applications of high-resolution MS in bioanalysis. *Bioanalysis* 5(10), 1269–1276 (2013).

27 Plumb RS, Fujimoto G, Mather J *et al.* Comparison of the quantification of a therapeutic protein using nominal and accurate mass MS/MS. *Bioanalysis* 4(5), 605–615 (2012).

28 Ramagiri S, Garofolo F. Large molecule bioanalysis using Q-TOF without predigestion and its data processing challenges. *Bioanalysis* 4(5), 529–540 (2012).

• **Highlights some of the data processing challenges associated with high-resolution accurate mass (HRAM) methods.**

29 Morin LP, Mess JN, Garofolo F. Large-molecule quantification: sensitivity and selectivity head-to-head comparison of triple quadrupole with Q-TOF. *Bioanalysis* 5(10), 1181–1193 (2013).

• **Challenges previous belief that triple quadrupole MS outperforms HRAM with respect to sensitivity.**

30 Ruan Q, Ji QC, Arnold ME *et al.* Strategy and its implications of protein bioanalysis utilizing high-resolution mass spectrometric detection of intact protein. *Anal. Chem.* 83(23), 8937–8944 (2011).

- **Describes first attempt at high-resolution full scan MS for the analysis of an intact protein.**

31 Chambers EE, Lame ME, Bardsley J *et al.* High sensitivity LC–MS/MS method for direct quantification of human parathyroid 1–34 (teriparatide) in human plasma. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* 938, 96–104 (2013).

32 Howard JW, Kay RG, Tan T *et al.* Development of a high-throughput UHPLC–MS/MS (SRM) method for the quantitation of endogenous glucagon from human plasma. *Bioanalysis* 6(24), 3295–3309 (2014).

Boehringer Ingelheim ranks among the world's 20 leading pharmaceutical corporations. Our vision drives us forward. It helps us to foster value through innovation in our company and to look to the future with constantly renewed commitment and ambition.

Boehringer Ingelheim

Value through Innovation

Our family's science, your family's health.

For more than 125 years Boehringer Ingelheim has been committed to the research and development of innovative medicines that help improve the lives of patients and their families.

Boehringer Ingelheim at their Ridgefield, CT site is actively developing novel protein drug candidates addressing major therapeutic areas. We provide a stimulating scientific environment for bioanalytical scientists with both ligand binding and large molecule LC-MS expertise.

Learn more about us at: **us.boehringer-ingelheim.com**

Research Article

Bioanalysis

Characterization of matrix effects in developing rugged high-throughput LC–MS/MS methods for bioanalysis

Aim: There is an ever-increasing demand for high-throughput LC–MS/MS bioanalytical assays to support drug discovery and development. **Results:** Matrix effects of sofosbuvir (protonated) and paclitaxel (sodiated) were thoroughly evaluated using high-throughput chromatography (defined as having a run time ≤1 min) under 14 elution conditions with extracts from protein precipitation, liquid–liquid extraction and solid-phase extraction. A slight separation, in terms of retention time, between underlying matrix components and sofosbuvir/paclitaxel can greatly alleviate matrix effects. **Conclusion:** High-throughput chromatography, with proper optimization, can provide rapid and effective chromatographic separation under 1 min to alleviate matrix effects and enhance assay ruggedness for regulated bioanalysis.

Fumin Li*[,1], Jun Wang[1] & Rand Jenkins[2]

[1]PPD® Laboratories, 3230 Deming Way, Middleton, WI 53562, USA

[2]PPD® Laboratories, 2240 Dabney Rd, Richmond, VA 23230, USA

*Author for correspondence: fumin.li@ppdi.com

First draft submitted: 8 January 2015; Accepted for publication: 17 March 2016; Published online: 15 April 2016

Keywords: high-throughput LC–MS/MS methods • matrix effects • paclitaxel • regulated bioanalysis • sofosbuvir • sub-minute chromatography

LC–MS/MS is an established methodology for quantitative measurement of small molecule drugs and/or their metabolite(s) in biological matrices to support drug discovery [1,2] and development [3]. With the ever-rising cost to develop and obtain marketing approval for a drug, there is great demand for faster and more cost-effective bioanalytical assays from early lead optimization to preclinical and clinical studies.

Various strategies have been applied to increase throughput of LC–MS/MS methods. Columns packed with sub-2 µm particles (generally termed ultra-high performance LC or UHPLC columns) can improve assay sensitivity, selectivity and throughput over traditional larger 5–10 µm particle columns for LC–MS/MS [4,5]. Kate *et al.* showed that up to threefold increase in throughput and one order of magnitude increase in detected peak height were realized with chromatography using sub-2 µm particles [6]. Columns packed with fused-core particles have been demonstrated to increase assay throughput [7,8]. Each fused-core particle, 2.7 µm in diameter, comprises a 0.5 µm porous shell and a solid 1.7 µm silica core particle. High-throughput is realized with faster separation due to shorter mass transfer or diffusion path. high-throughput can be achieved via the use of high-flow chromatography. The inherent merit of low pressure drop offered by monolithic columns facilitates the use of high flow. For instance, an eluting flow rate of 3.2 ml/ml was applied on a monolithic C18 column compared with 1.2 ml/min from a traditional C18 column, where twofold gain in throughput was realized with comparable chromatographic separation and assay performance [9]. A flow rate up to 6 ml/min was used with a monolithic column for analysis of four analytes extracted from plasma, where baseline separation was achieved in 1 min [10]. Likewise, HILIC chromatography can be conducted at high flow rate due to the use of high organic content (low column back pres-

FUTURE SCIENCE part of fsg

sure) [11]. High throughput can be accomplished via the use of parallel columns or HPLC systems with only one mass spectrometer [12,13] to increase the utilization of a mass spectrometer, allowing more samples to be run within a given time.

High-throughput bioanalysis can be realized without sample clean up and/or chromatographic separations. Ambient desorption ionization techniques in combination with MS have been utilized for high-throughput bioanalysis [14,15]. Direct analysis in real time (DART) coupled with a triple-quadrupole mass spectrometer has been developed and applied for direct analysis of samples from biological matrices [15]. DESI MS [16,17] was employed for high-throughput analysis of pharmaceutical samples, where sampling rates as high as two samples/s were achieved in the ambient environment [16]. Laser diode thermal desorption (LDTD) MS/MS has been utilized to increase assay throughput to support *in vitro* drug discovery bioanalysis of chemically diverse compound libraries [18]. Methodologies using an automated high-throughput solid-phase extraction tandem MS (HT-SPE/MS), for example, Rapid Fire™, were demonstrated with analysis times of 5–10 s per sample [19,20]. HT-SPE/MS integrates fast online SPE with MS detection. It should be noted, however, that the above-mentioned high-throughput approaches are accomplished with the sacrifice of sample preparation and/or chromatographic separation, where severe matrix effects are often accompanied. For example, significant matrix suppression, indicated by matrix factors ranging from approximately 0.05 to 0.7 [15], was encountered when untreated raw plasma samples were analyzed. Matrix factor (MF) is the quantitative measure of matrix effect and is defined as the peak response of an analyte in the presence of matrix components over that in the absence of matrix components. Internal standard (IS) normalized matrix factor (IS-MF) is the MF of an analyte divided by the MF of its corresponding internal standard. MF of 1 is most desirable, indicating no matrix suppression or enhancement. The high-throughput methods cited here are mostly aimed to support discovery bioanalysis, where the assay performance requirements are fit-for-purpose and typically much less stringent than those expected of assays intended to support regulatory filing.

High-resolution MS (HRMS) [21,22,23] has emerged as an alternative detection technique for bioanalysis over triple quadrupole MS – the current standard platform. HRMS is capable of post-acquisition data processing, full-scan acquisition for multiple compounds and no need of compound-specific optimization. Jie *et al.* used UHPLC-HRMS to simultaneously measure 19 bile acids that exhibited poor fragmentation [24]. A high-resolution MS-LC multiplexing method was used to analyze phosphorylated peptides at sampling rates of up to 18 s per sample [25]. Despite its advantages in mass selectivity, HRMS is still subject to the same ESI-related matrix effect issues as low-resolution systems.

Striving for increased assay throughput must be balanced with assay integrity, especially for regulated bioanalysis. The speed of analysis should not undermine the quality of the data [26]. For example, results from the HT-SPE/MS approach showed severe matrix effects, with matrix factors (MF) ranging from 0.20 to 0.68 for compounds tested, due to the lack of any chromatographic separations [20]. Per regulatory guidance requirements [27,28], bioanalytical assays intended to support regulatory filings should evaluated for matrix effects, which are one of the most important considerations for LC–MS/MS-based assays [29]. Matrix effects typically refer to either suppression or enhancement of ionization efficiency of an analyte due to the presence of co-extracted matrix components such as dosing vehicles and phospholipids [30]. It is well documented that matrix effects can adversely impact assay sensitivity, accuracy and precision [31–36]. If uncompensated, inaccurate PK results and incorrect estimation of drug exposure could impact critical development and regulatory decisions.

As such, matrix effects should be thoroughly evaluated for bioanalytical methods intended regulated analysis, especially with the use of high-throughput chromatography [37]. Assay throughput in LC–MS/MS bioanalysis should be achieved with minimal matrix effects [38]. Here we conducted comprehensive characterization of matrix effects using high-throughput chromatography (HTC) with run time ≤1 min. Sofosbuvir (protonated) and paclitaxel (sodiated) were chosen as model compounds. Their matrix effects, with or without correction from their corresponding stable labeled internal standards, were evaluated using HTC under 14 elution conditions with extracts from protein precipitation (PPT). Sofosbuvir's matrix effects were further assessed using samples obtained from liquid–liquid extraction (LLE) and SPE. Matrix effects for sofosbuvir were compared between a HALO C18 column (2.1 × 20 mm, 5 µm) and an XBridge C18 column (2.1 × 20 mm, 5 µm) using 14 HTC conditions. Last but not the least, a total of 36 lots of plasma, six species with six different lots per species, was investigated for sofosbuvir's matrix effects using an HTC with a run time of 0.6 min.

Experimental

Chemicals & reagents

Sodium acetate (≥99%), HPLC-grade acetonitrile (ACN), Methyl-t-butyl ether (MtBE, 99.9%) was from

Sigma-Aldrich. Formic acid (ACS grade) was from EMD. 2-Propanol (IPA) with HPLC grade was from J.T. Baker. Human, rat, dog, monkey, rabbit and mouse plasma containing K_2EDTA; human plasma containing sodium fluoride/potassium oxalate; and human plasma containing sodium heparin as the anticoagulant were purchased from Biological Specialties.

Chromatographic conditions

The HPLC system was a Nexera (Shimadzu, MD, USA). Either an XBridge C18 column (2.1 × 20 mm, 5 µm, Waters, MA, USA) or a HALO C18 column (2.1 × 20 mm, 5 µm, Advanced Materials Technology) was employed for the HTC evaluations. Chromatography was performed at room temperature, and unless otherwise noted, the following screening conditions were used. Mobile phase A and B were 0.1% formic acid in water and 0.1% formic acid in acetonitrile, respectively. A total of 14 conditions were tested, where the initial organic content varied from 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 70, 80 to 95% of mobile phase B (i.e., acetonitrile). The isocratic elution lasted for 0.78 min followed by a 0.1-min wash at 95% acetonitrile. The column was then equilibrated at the initial isocratic condition at 0.9 min and the chromatography stopped at 1 min. The flow rate was set at 1.5 ml/min.

MS

An API 4000 triple quadrupole mass spectrometer (AB Sciex) with Turbo Ion Spray was set up in positive ESI mode. The following conditions were optimized for sofosbuvir. The collision gas flow (CAD) was 8; curtain gas flow was 25; gas supply 1 (GS1) was 50; gas supply 2 (GS2) was 80. Ion spray voltage (IS) was 5000 V; source temperature (TEM) was 550°C; declustering potential was 40 V; collision energy was 30 eV; entrance potential was 10 V; collision cell exit potential was 15 V. The multiple reaction monitoring (MRM) transitions for sofosbuvir and sofosbuvir-d6 were *m/z* 530.2>243.1 and 536.1>243.1. The following parameters were optimized for paclitaxel. The collision gas flow (CAD) was 10; curtain gas flow was 25; gas supply 1 (GS1) was 30; gas supply 2 (GS2) was 80. Ion spray voltage (IS) was 5500 V; source temperature (TEM) was 550°C; declustering potential was 52 V; collision energy was 40 eV; entrance potential was 10 V; collision cell exit potential was 20 V. The MRM transitions for paclitaxel and paclitaxel-d5 were *m/z* 876.4 > 308.3 and 881.3> 313.3. The sodium adduct was monitored for both paclitaxel and paclitaxel-d5.

Analyst v 1.6.2 software (Applied Biosystems) was used in data analysis. When appropriate, a 1/concentration^2 weighted linear regression was applied in quantitation.

Samples preparation

Matrix-containing test samples were prepared by spiking sofosbuvir or paclitaxel into blank matrix extracts, while matrix-free test samples were prepared by spiking sofosbuvir or paclitaxel into a neat solution. The spiking solutions for sofosbuvir or paclitaxel were prepared at 100 ng/ml with their corresponding internal standards at 1000 ng/ml, in diluent/reconstitution buffer (Tween-20/acetonitrile/water, 0.1:10:89.9, v/v/v). Matrix-containing samples were prepared by extracting blank plasma from PPT, LLE or SPE, followed by postspiking with the neat analyte/IS solutions.

PPT procedures for blank & CAL/QC samples

A 50 µl plasma sample aliquot was added to a 96-well plate, followed by addition of 25 µl of working internal standard in ACN/water (1:1, v/v) or ACN/water (1:1, v/v) for blank samples. Samples were mixed for 1 min. Each well was treated with 300 µl of ACN to precipitate plasma proteins. The plate was vortexed and centrifuged. Either a 100 µl aliquot of supernatant (matrix-containing) or a 100 µl aliquot of reconstitution solvent (matrix-free) was transferred to clean wells containing 100 µl of a neat spiking analyte/IS solution for matrix factor analysis.

LLE procedures for blank samples

A 50 µl of plasma sample aliquot was transferred to a 96-well plate, followed by addition of 25 µl of ACN/water (1:1, v/v). The plate was mixed. Each sample well received 800 µl of MtBE (100%). The plate was vortexed and centrifuged at 2500 rpm for 5 min at room temperature. Seven hundred microliters of the upper organic layer were transferred to a clean 96-well plate. The organic layers were evaporated to dryness under a stream of nitrogen. Finally, 200 µl of a neat reconstitution solution with analyte/IS was added to either wells of dried extracts from plasma blanks (matrix containing) or clean wells (matrix free).

SPE procedures for blank samples

SPE extractions were conducted using OASIS HLB (Waters), in other words, RP-SPE. The HLB plate was conditioned with 1 ml methanol and 1 ml water. A 50 µl sample aliquot was transferred to a 96-well plate, followed by addition of 25 µl of ACN/water (1:1, v/v). The plate was mixed. The entire samples were then loaded to the equilibrated HLB plate. The HLB plate was washed with 1 ml water and 1 ml 10% methanol in water, followed by elution with 1 ml methanol. The eluate was evaporated to dryness. Finally, 450 µl of a reconstitution solution with analyte/IS was added to either wells of dried extracts from plasma blanks (matrix containing) or clean wells (matrix free).

Results & discussion

Matrix effects under different high-throughput chromatographic conditions

One of the central questions to apply high-throughput chromatography (HTC) in bioanalysis, in other words, with a run time ≤ 1 min, is how to prevent adverse matrix effects. To answer this question, sofosbuvir and paclitaxel were employed as model compounds. Their matrix factors (MFs) and internal standard normalized matrix factors (IS-MFs) were measured under HTC in triplicate with extracts obtained via PPT, considered to be the least effective sample preparation method at removing (nonprotein) matrix components [31]. A total of 14 conditions were tested using an XBridge C18 column (2.1 × 20 mm, 5 µm).

Figure 1 shows sofosbuvir's MFs (black bars, left axis) and IS-MFs (red bars, left axis) under 10 to 95% of acetonitrile isocratic elution. The line with blue rectangles depicts the corresponding retentivity or k' (right axis). Retentivity, in other words, retention factor or k', refers to the time an analyte stays in the stationary phase relative to the time it resides in the mobile phase [39]. A retention factor of at least 2 is generally recommended for bioanalysis to alleviate matrix effects. From 80 to 95% ACN, sofosbuvir was not retained on the column and was eluted at the void volume (i.e., 0.06 min). Its MFs were 0.143 ± 0.023 and 0.212 ± 0.014 (n = 3) at 95 and 80% acetonitrile, respectively. Sofosbuvir MS signal was greatly suppressed likely due to coelution of matrix suppressing components at the void. Decreasing ACN from 95 to 80%, certain matrix ion suppressing components could be retained on the column, leading to a slightly reduced suppression. Sofosbuvir's IS-normalized matrix factors (IS-MFs) were 1.113 ± 0.072 and 1.255 ± 0.122 (n = 3), where sofosbuvir-d6 largely corrected sofosbuvir's signal variation. Nevertheless, sofosbuvir-d6 did not seem to completely compensate for sofosbuvir's matrix effects as expected given that sofosbuvir-d6 is the stable isotope-labeled version of sofosbuvir, both of which co-eluted chromatographically. This is possibly related to ESI mechanisms [40,41]. Molecules are distributed within ESI droplets depending on their hydrophobicity [40]. The more hydrophobic a molecule, the closer it stays to the surfaces of ESI droplets, and it is more likely ionized. In comparison, polar compounds, staying close to an ESI droplet's center, are less efficiently ionized during ESI. The addition of six deuterium atoms makes sofosbuvir-d6 slightly more polar than sofosbuvir. It is postulated that sofosbuvir and sofosbuvir-d6 were ionized a little differently in ESI, leading to a slight difference in MFs for sofosbuvir and sofosbuvir-d6. Therefore, the IS-MF for sofosbuvir is not 1. Specifically, sofosbuvir-d6 was more suppressed by matrix components than sofosbuvir was.

When acetonitrile was reduced from 70 to 45%, sofosbuvir's MFs increased from 0.275 ± 0.010 to 0.901 ± 0.022 (n = 3). Its retention time increased slightly from 0.07 to 0.09 min. The results suggested that a small but adequate separation between underlying matrix suppression components and sofosbuvir greatly alleviates its matrix effects. Hence, sofosbuvir's MF was increased to 0.901 even with a k' of 0.33. Sofosbuvir's IS-MFs changed from 1.202 (70% B), 1.192 (60% B), 1.070 (55% B), 1.013 (50% B), to 1.043 (45% B). The improvement in IS-MFs is likely due to the increased chromatographic separation between sofosbuvir/sofosbuvir-d6 and matrix components. With further reduction in acetonitrile strength (40–25% ACN), the retention time of sofosbuvir was increased from 0.11 to 0.55 min (k' from 0.83 to 8.17). Sofosbuvir's MFs (IS-MFs) are 1.021 (0.962) at 40%, 0.945 (0.990) at 35%, 0.970 (0.914) at 30% and 1.049 (1.043) at 25%, all of which are within the range of 0.9–1.1 (bracketed by black lines). The results demonstrated that HTC can provide near matrix-free conditions for optimum detection of compounds such as sofosbuvir.

Majority of small molecules have sufficient inherent basicity to be ionized via protonation in the positive ESI mode. Others have electronegative sites that may readily attract cationic species, including alkali metal ions and ammonia, to form adduct ions. Thus it would be valuable to evaluate how matrix effects are impacted when analyte ions are generated as adducts. Paclitaxel was chosen as a model compound, as it is readily measured in LC–MS as adducts of sodium, which is common in solvents used for bioanalysis. The same systematic experiments for sofosbuvir were conducted for paclitaxel prepared in PPT extracts. Figure 2A shows paclitaxel's MFs (black bars, left axis) and IS-MFs (red bars, left axis) as a function of acetonitrile strength from 10 to 95%. The line with blue rectangles is the corresponding retentivity (right axis). From 80 to 95% ACN, paclitaxel showed no retention on the column and eluted at the void volume (~0.06 min). When acetonitrile was reduced to 70%, its retention increased slightly from 0.06 to 0.09 min. Paclitaxel's MFs were 5.469 ± 0.513 (n = 3), 7.440 ± 0.297 (n = 3) and 7.843 ± 0.700 (n = 3) at 95, 80 and 70% acetonitrile, respectively, indicating significant ionization enhancement for paclitaxel sodium adduct signals. Under these chromatographic conditions, sodium, from plasma extracts and reagents, co-eluted with paclitaxel at the void volume. Sodium is excessively abundant in ESI droplets, making its positive charges readily available for paclitaxel and paclitaxel-d5. Hence, the apparent ionization efficiency for paclitaxel, as sodium adducts, from plasma extracts is much higher than that from a neat or matrix-free solution. Paclitaxel's MFs were compared

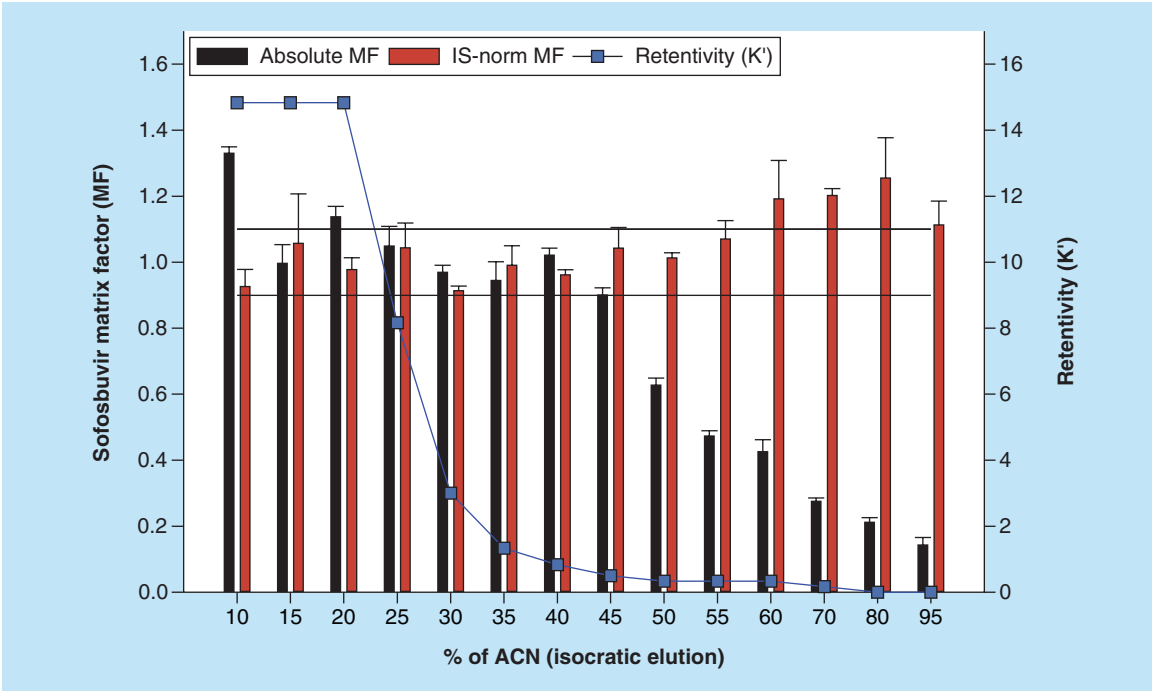


Figure 1. Shows the matrix factors, internal standard normalized matrix factors and retentivity of sofosbuvir using high-throughput chromatography. Extracts were from protein precipitation.

using extracts obtained from LLE and SPE as well. The results (not shown) illustrated that LLE and SPE resulted in reduced ionization enhancement than PPT when eluted at void. The observation suggests that an excess of sodium ions at the void contributes to the enhancement.

In comparison, the paclitaxel's IS-MFs were 0.983 ± 0.121 (n = 3), 0.918 ± 0.061 (n = 3) and 1.053 ± 0.083 (n = 3) at 95, 80 and 70% acetonitrile, respectively. The IS-MFs are all within 0.90–1.10, demonstrating that paclitaxel-d5 compensated well for the variability of paclitaxel response even when paclitaxel eluted at the void. Compared with sofosbuvir-d6, paclitaxel-d5 seemed to track paclitaxel better at void volume. This could be, in part, attributed to the nature of sodium adducts. In ESI droplets, there are excessive amounts of sodium available and there is negligible competition to form adducts with other positively charge species. When acetonitrile is further reduced to 60%, the retention time of paclitaxel is increased to 0.11 min with a k" of 0.57. Its MF became 1.095, indicating little matrix effects. This drastic change in MF (i.e., MFs reduced from 7.843 at 0.09 min to 1.095 at 0.11 min) suggests that paclitaxel moved away from the void zone (0.06 min), where sodium was especially abundant.

Outside the void, sodium ions are most likely from mobile phases instead of plasma extracts. Therefore, matrix effects driven by the availability of excessive sodium should no longer be related to matrix. Rather, matrix effects were derived from other matrix components than sodium. Paclitaxel showed increasing retention on the column when acetonitrile was reduced from 60 to 40%. Paclitaxel's MFs were 1.014 (55% B), 1.172 (50% B), 1.141 (45% B) and 1.181 (40% B), while its IS-MFs were 1.011 (55% B), 0.984 (50% B), 0.951 (45% B) and 0.951 (40% B) respectively. Apparently, paclitaxel-d5 compensated paclitaxel's response better than sofosbuvir-d6 did for sofosbuvir. This may be explained by the fact that the polarity difference between paclitaxel and paclitaxel-d5 is smaller than that between sofosbuvir and sofosbuvir-d6. In other words, paclitaxel-d5 ionized more closely to paclitaxel than sofosbuvir-d6 to sofosbuvir. When the percentage of acetonitrile is further reduced to and below 35%, paclitaxel retained on the column during the isocratic elution and did not elute until acetonitrile was increased to 95% during column flush. This region was not recommended for bioanalysis. Figure 2B shows paclitaxel's IS-MFs encompassing the entire range of chromatographic conditions, in other words, isocratic elution from 10 to 95% acetonitrile. The top and bottom lines correspond to values of 1.1 and 0.9. The results show that paclitaxel-d5 tracked paclitaxel sufficiently under all 14 high-throughput chromatographic conditions evaluated.

Matrix effects under different sample preparation methods

Sample preparation plays a key role in reducing matrix effects in LC–MS/MS-based bioanalysis [31]. Matrix

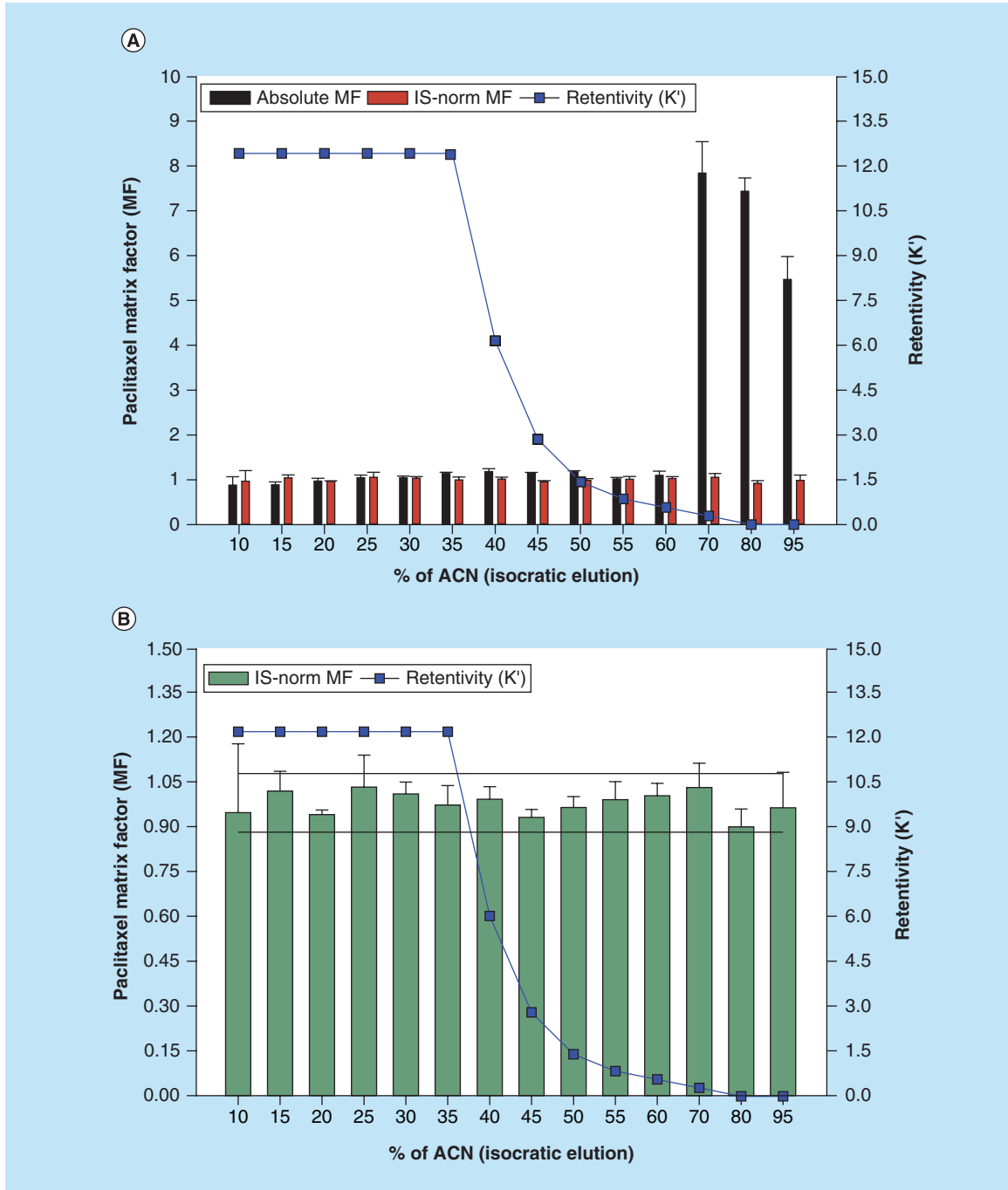


Figure 2. Shows the matrix factors, internal standard normalized matrix factors and retentivity of paclitaxel using high-throughput chromatography. Extracts were from protein precipitation.

effects were systematically evaluated for sofosbuvir under the same 14 HTC conditions using extracts obtained from three common sample preparation techniques, PPT, LLE and SPE. For each sample preparation technique, sofosbuvir's MFs were measured in triplicate (n = 3) at each isocratic elution conditions, ranging from 10 to 95% acetonitrile. Figure 3A shows sofosbuvir's matrix factors as a function of acetonitrile strength, using extracts from PPT (black bars), LLE (red bars) and SPE (green bars). At 95% acetonitrile, sofosbuvir was not retained on the column, co-eluting along with the majority of co-extracted matrix components. As anticipated, severe matrix ion suppression was observed regardless of what sample extraction method was employed. Specifically, sofosbuvir MFs were 0.143, 0.404 and 0.350 from PPT, LLE and SPE, respectively. The injection volume was the same under each condition. Although the amount of matrix com-

ponents injected may differ among PPT, SPE and LLE, sofosbuvir's MFs may be used to qualitatively gauge the effectiveness of extraction methods in removing matrix-effect-causing materials. The results indicated increasing effectiveness in the order of PPT, SPE and LLE, where LLE provides the cleanest sample extracts for sofosbuvir, similar to observations from a previous study [31].

The 14 chromatographic conditions can be divided into three regions. In the first region from 50 to 95% acetonitrile, sofosbuvir showed no or little retention on the column. Severe matrix suppression, approximately 37 to 86%, was observed from PPT extracts. From 50 to 80% acetonitrile, sofosbuvir's MFs were all within 15% of the nominal value for SPE and LLE, demonstrating that HTC effectively separated

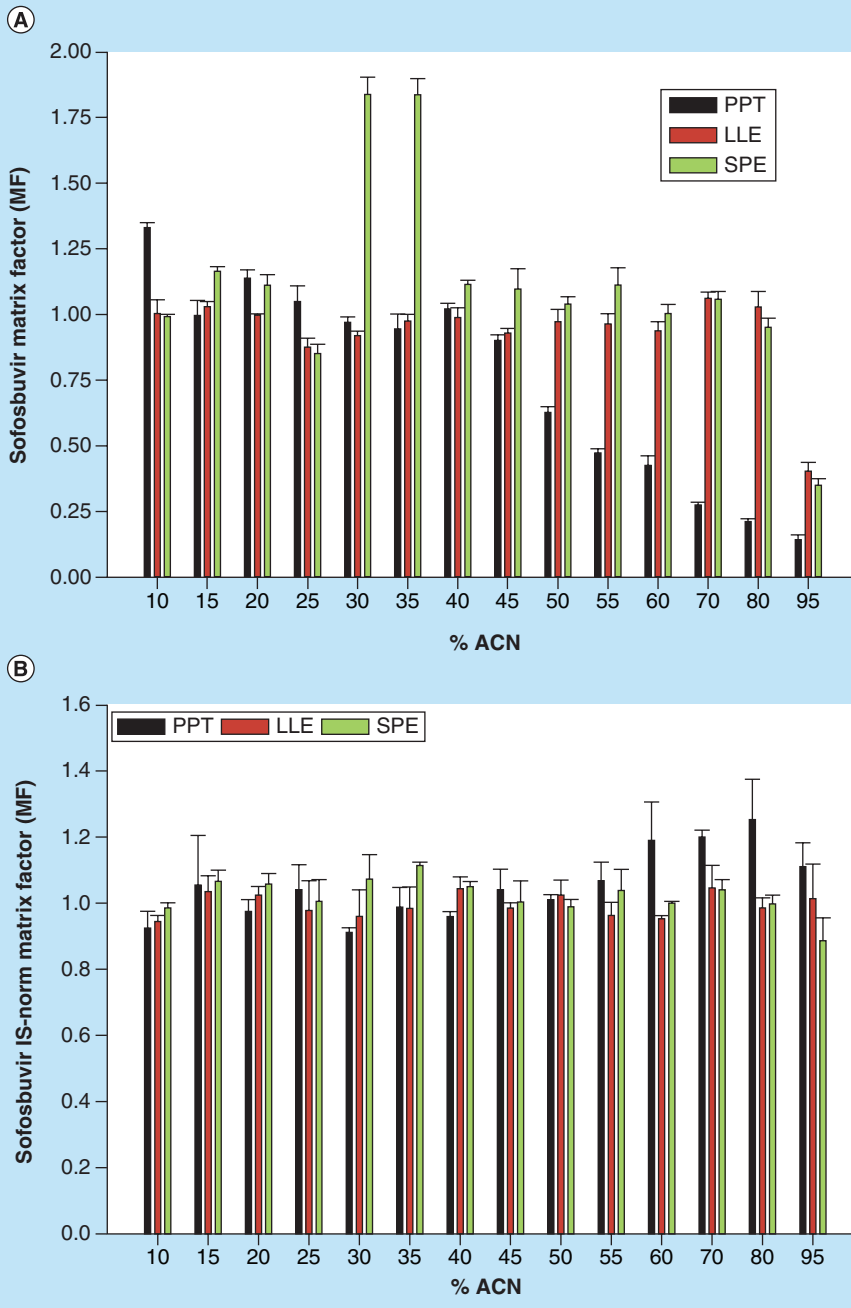


Figure 3. Shows sofosbuvir's matrix factors (A) and sofosbuvir-d6 IS-normalized matrix factors (B) under different isocratic elution, from 10 to 95% acetonitrile, using extracts from three commonly used sample extraction methods – protein precipitation, liquid–liquid extraction and solid phase extraction.

sofosbuvir from matrix components derived from SPE and LLE. In the second region from 25 to 45% acetonitrile, sofosbuvir was increasingly retained on the column. Chromatography in this region is ideal for bioanalysis. Sofosbuvir's MFs were within ±15% of 1 for PPT and LLE. Given that significant matrix suppressing components were present in the PPT extracts, the results supported that HTC could effectively separate sofosbuvir from matrix components. At 30 and 35% acetonitrile, approximately 80% of matrix enhancement was observed from SPE. This may be due to the specific chromatography conditions that resulted in co-elution of matrix enhancement components from SPE. The results showed that matrix effects in LC–MS/MS-based bioanalysis are a combined effect of sample preparation and chromatography. The third region, from 10 to 20% acetonitrile, was unusable for bioanalysis, as sofosbuvir was completely retained on the column.

Figure 3B shows sofosbuvir's IS-MFs using extracts from PPT (black bars), LLE (red bars), and RP-SPE (green bars) under the same 14 chromatographic conditions. The IS-MFs were mostly within ± 15% of the nominal value for all three sample preparation methods, demonstrating that sofosbuvir-d6 sufficiently compensated for matrix effects impacting sofosbuvir. One exception was that the IS-MFs from PPT in the region of 60–80% acetonitrile were about 1.2. In the practical chromatographic region, 25–45% acetonitrile, IS-MFs were from 0.914 to 1.043 for PPT, from 1.008 to 1.116 for SPE and from 0.963 to 1.043 for LLE. Overall, the results showed that LLE provides cleaner extracts for sofosbuvir than SPE, while PPT is the least effective approach in removing matrix components.

Matrix effects with the use of different columns

Chromatography plays a key role in mitigating matrix effects. To evaluate how columns impact matrix effects, two commonly used columns were evaluated, in other words, a fully porous column – BEH ethylene bridge hybrid silica particle XBridge C18 (20 mm × 2.1 mm, 5 µm) and a superficially porous – fused core particle HALO C18 (20 mm × 2.1 mm, 5 µm). Sofosbuvir's MFs were measured from samples prepared from PPT extracts using the same 14 conditions as described previously. Figure 4A depicts sofosbuvir's MFs under various isocratic elution conditions, from 10 to 80% acetonitrile, using superficially porous C18 column (the red line with solid triangles) and the porous XBridge C18 column (the dark blue line with solid rectangles), respectively. All other parameters such as mobile phases, flow rates and gradient profiles were kept the same. The void volume for the HALO column corresponded to a retention time of approximately 0.04 min, while that for the XBridge column was approximately 0.06 min. That the fused-core column has a smaller void than the porous one may be attributed to more uniform particle size distributions [7]. This leads to a more tightly packed HALO column and a smaller void volume.

In the region of 50–95% acetonitrile, sofosbuvir was almost unretained and eluted at or near the void volume for either column. Sofosbuvir was considerably more suppressed from the HALO column than the XBridge one. For instance at 95% acetonitrile, sofosbuvir's MFs are 0.143 ± 0.023 (n = 3) and 0.054 ± 0.020 (n = 3) for the porous and superficially porous column, respectively. Sofosbuvir's IS-MFs are 1.727 ± 0.776 (n = 3) for the HALO column and 1.113 ± 0.072 (n = 3) for the XBridge column. Due to high variability especially from the HALO column, sofosbuvir's MF and IS-MF at 95% acetonitrile elution were not included in Figure 4. At 50% acetonitrile, sofosbuvir's MFs are 0.627 ± 0.021 (n = 3) from the XBridge column and 0.136 ± 0.016 (n = 3) from the HALO column, indicating a 4.6-fold more suppression from the superficially porous column. One potential explanation is the column loading capacity difference between the two sorbent types. Column loading capacity can be defined as the maximum amount of sample load in a column that leads to less than 10% loss of chromatographic efficiency [42]. The total surface area of sorbent materials in a column defines its loading capacity. The surface area and the bulk density of packing materials determine the total surface of a column. The surface areas of a C18 porous column and a superficially porous column are 300 m^2/g and 150 m^2/g, respectively [43]. In practice, the loading capacity of the porous C18 column was ~53% more than the superficially porous column (of the same dimension) [43]. The relative difference in loading capacity of the two columns used in the study (20 × 2.1 mm) should be similar to those (100 × 4.6 mm) reported previously [43]. Therefore, it is reasonable to expect that the porous XBridge C18 column has ~53% more loading capacity than the superficially porous Halo column. As a result, the porous XBridge column can handle a larger amount of matrix components that would otherwise overload the Halo column and coelute/suppress sofosbuvir's signal.

When the percentage of acetonitrile was reduced to 40%, the difference in terms of sofosbuvir suppression between the two columns was lessened, where the MFs of sofosbuvir were 1.021 ± 0.022 (n = 3) and 0.314 ± 0.009 (n = 3) for the porous and

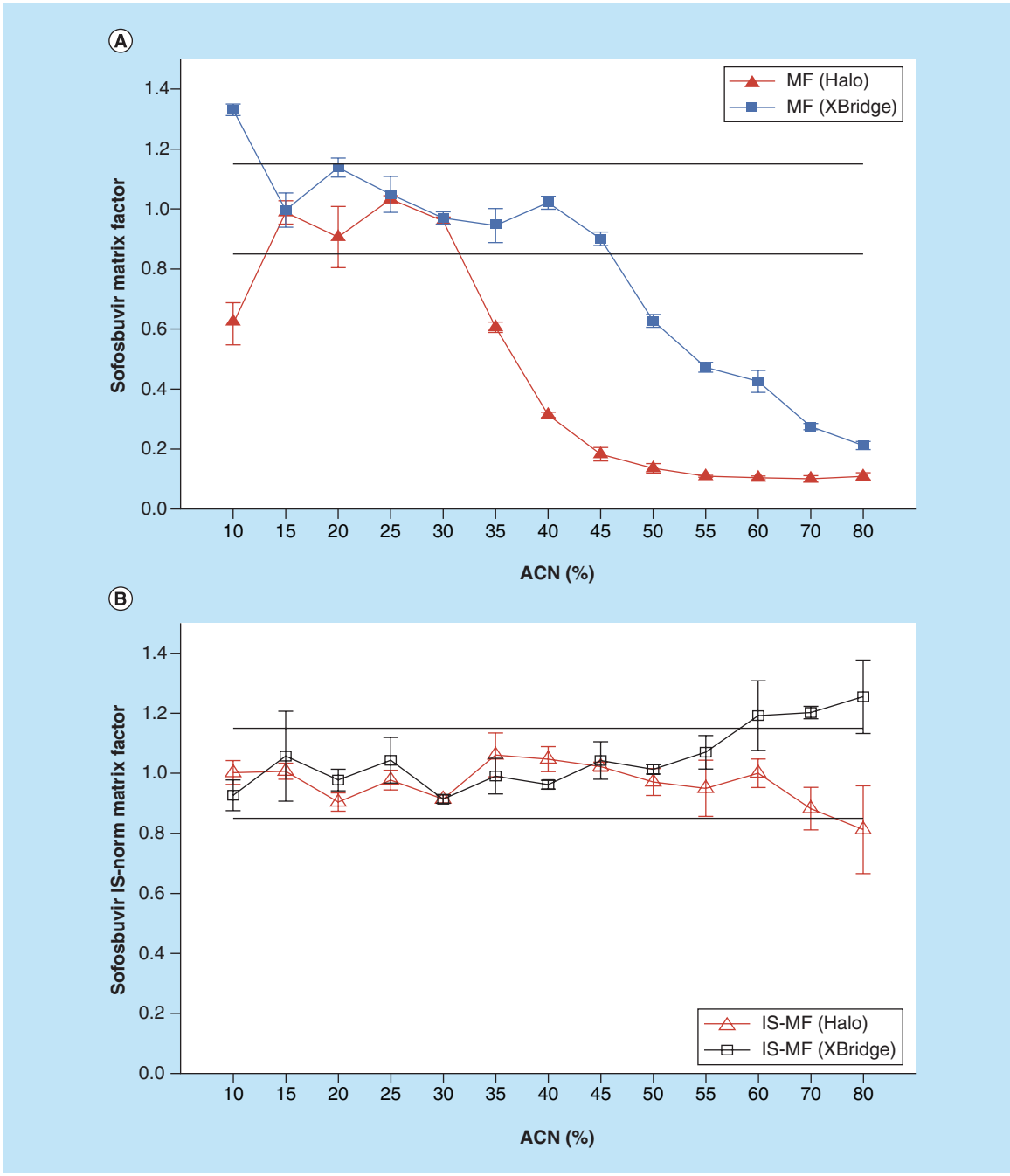


Figure 4. Sofosbuvir's matrix factors (A) and sofosbuvir-d6 IS-normalized matrix factors (B) under various isocratic elution (10–80% acetonitrile) using two different columns, an Xbridge C18 column (2.1 × 20 mm, 5 μm) and a Halo C18 column (2.1 × 20 mm, 5 μm).

superficially porous columns correspondingly. There was almost no matrix suppression from the porous column. At 30% acetonitrile, no matrix suppression was seen for both porous and superficially porous columns. From 10 to 20% acetonitrile, sofosbuvir was fully retained on both columns and eluted only at the flush stage with 95% acetonitrile. Therefore, the optimal region of elution with minimal matrix effects is from 25 to 30% ACN for the Halo column and from 25 to 45% ACN for the XBridge column. This significant difference in the matrix effects-free region of operation can, at least partly, be ascribed to the difference in column loading capacity. Figure 4B depicts sofosbuvir's IS-MFs. Sofosbuvir-d6 tracked sofosbuvir well overall. For instance, the IS-MFs are all within ±15% of the nominal value in the region 10–55% acetonitrile. The variation is slightly higher when acetonitrile is at or above 60% acetonitrile.

Matrix effects of plasma from different species & lots

To thoroughly evaluate matrix effects under high-throughput chromatography, sofosbuvir's MFs and IS-MFs were measured from six different lots of plasma from rat, mouse, rabbit, monkey, dog and human – six common species in drug discovery and development [32,44]. There were three replicates per measurement. A 0.6-min high-throughput chromatography was used. Mobile phase A and B are 0.1% formic acid in water and acetonitrile, respectively. The XBridge C18 column (2.1 × 20 mm, 5 µm) was used at a flow rate of 1.5 ml/min. The gradient is: 0 min (20% B), 0.05 min (40% B), 0.3 min (40% B), 0.31 min (95% B), 0.41 min (95% B), 0.42 min (20% B) and 0.60 (20% B/Stop). Under this chromatographic conditions, sofosbuvir eluted at 0.24 min (k′ = ~3).

The left panel of Figure 5 shows the sofosbuvir's MFs from six different lots of plasma from six common species, in other words, dog (solid red circles), human (open red circles), mouse (solid blue triangles), monkey (open blue triangles), rabbit (solid black rectangles) and rat (open black rectangles). Overall, sofosbuvir's matrix factors for all 36 lots of plasma are in the range of 0.9–1.1. The matrix factor variations, in terms of different lots of the same species or different lots across six different species, are all within ± 10%. The results clearly showed that the 0.6-min HTC effectively minimizes matrix effects for sofosbuvir and can be reliably applied for regulated bioanalysis. This level of chromatography throughput, 36 s/sample, essentially rivals other approaches without sample preparation and/or chromatography [14–20,45] but offers the benefit of having matrix effects under control. With a close examination of the results, the MFs from 33 plasma lots are all within ± 5% of the nominal value. Lot #5 from mouse, lot #1 from monkey and lot #5 from human had matrix factors of 1.094, 0.9345 and 0.943, slightly outside of ± 5% but within ± 10%.

The right panel of Figure 5 shows sofosbuvir's IS-MFs from the same six different lots of plasma from dog (solid red circles), human (open red circles), mouse (solid blue triangles), monkey (open blue triangles), rabbit (solid black rectangles) and rat (open black rectangles). In all, the IS-MFs for all 36 lots of plasma are in the range of 0.97–1.03. The variations, in terms of different lots (same species) or different lots across six different species, are all within ± 3%. The results showed that sofosbuvir-d6 tracked sofosbuvir well under this sub-minute high-throughput chromatography. Based on results from Figure 1, sofosbuvir MF and IS-MF at 40% acetonitrile is 1.021 and 0.962. This observation agreed well with results from this sub-minute chromatography, where sofosbuvir eluted at 40% acetonitrile. The fact that sofosbuvir's MFs and IS-MFs are all within 0.9–1.1 from different lots and different species strongly supports that matrix effects can be successfully managed with the use of high-throughput, sub-minute chromatography.

The performance of the sub-minute chromatography was further assessed. Figure 6 shows representative chromatograms of sofosbuvir from an LLOQ sample (blue/top trace) and a matrix blank (black/bottom trace) using the sub-minute chromatography. Human plasma samples were processed using PPT as described in the Experimental section. Sofosbuvir's retention time was 0.24 min corresponding to a k' of 2.4. Sofosbuvir was baseline separated from an interfer-

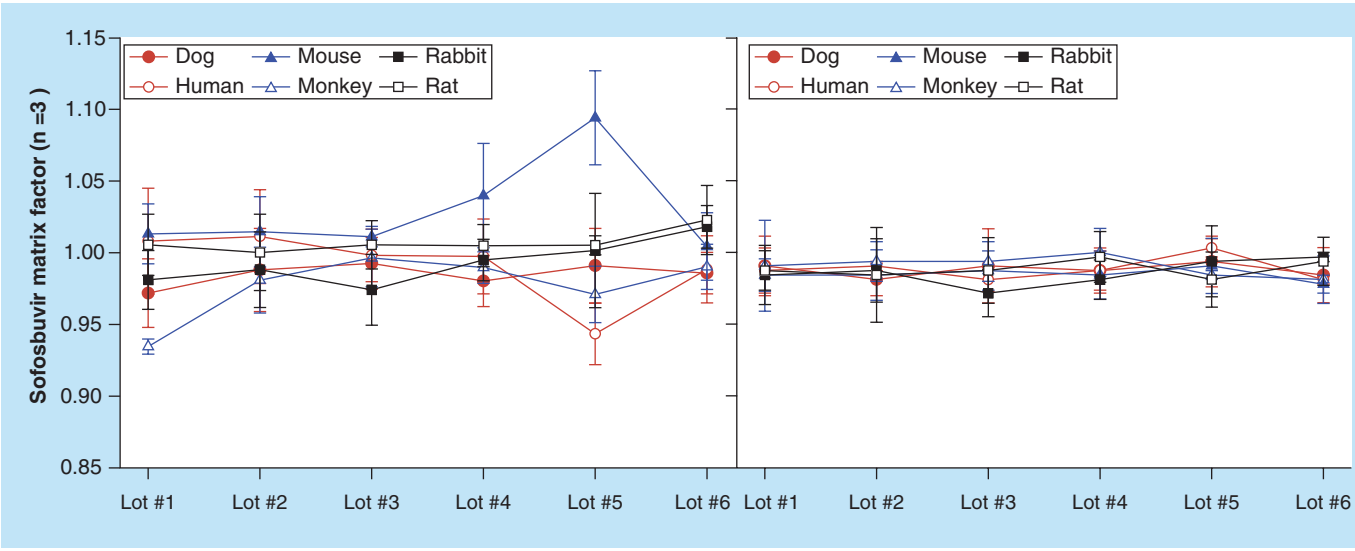


Figure 5. Matrix factors and internal standard normalized matrix factors for sofosbuvir from six common species (six different lots each) used in drug discovery and development under high-throughput chromatography.

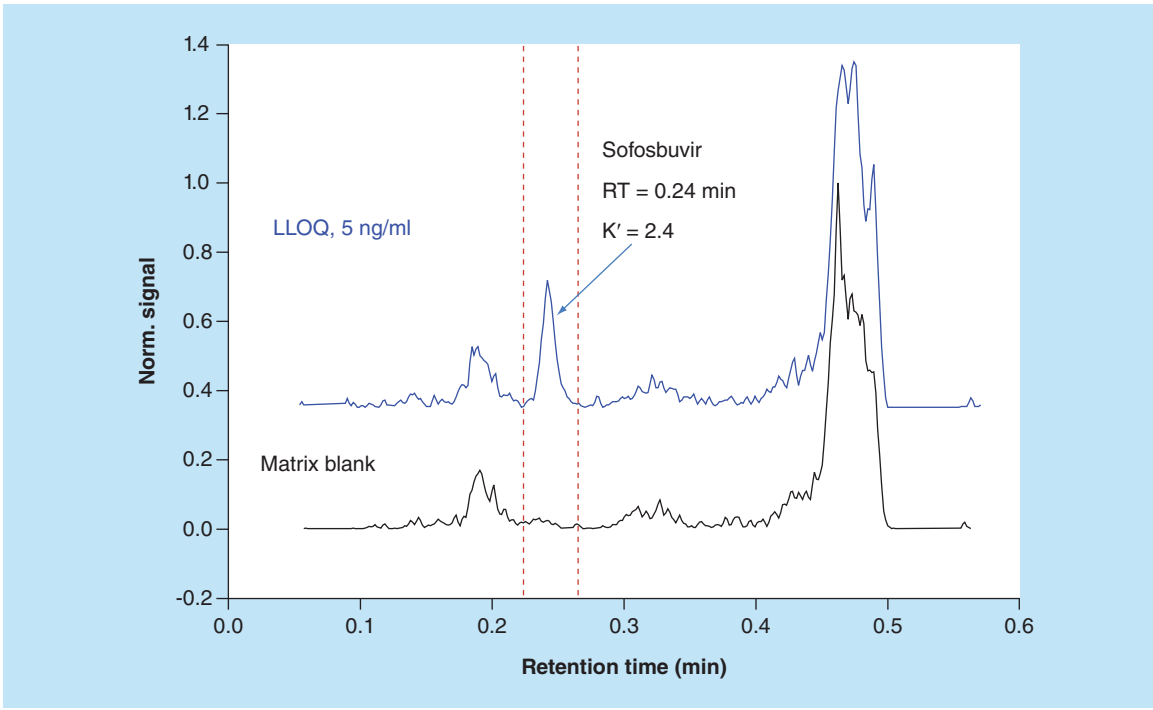


Figure 6. Representative chromatograms of sofosbuvir from an LLOQ sample (blue/top trace) and a matrix blank (black/bottom trace) using a high-throughput chromatography. Protein precipitation was employed for sample cleanup.

ence peak, presumably from PPT – the least effective sample cleanup method. PPT was chosen to challenge the performance of the sub-minute chromatography. The signal-to-noise ratio for sofosbuvir at LLOQ is greater than 10. Table 1 lists the precision and accuracy of six levels of sofosbuvir QCs measured by the sub-minute HTC (n = 6, 1 run) and a previously validated 2-min chromatography (n = 18, 3 runs). The intra-assay accuracy for the six QCs was between 98.3 and 107.4% (n = 6), and between 3.0 and 9.9% (n = 6) in terms of precision for the sub-minute chromatography (0.6 min). For the 2-min chromatography, the inter-assay accuracy for the six QC levels was between 96.9 and 101.9% (n = 18), and between 2.3 and 3.9% (n = 6) in terms of precision. In all, the results demonstrate that the bioanalytical method using the sub-minute chromatography performs similarly to the one with approximately threefold longer in run time.

Conclusion

High-throughput chromatography, defined as having a run time of ≤1 min, was systematically evaluated for matrix effects of sofosbuvir (protonated) and paclitaxel (sodiated) using extracts derived from PPT. The results demonstrated that a small but adequate separation between underlying matrix suppression components and sofosbuvir/paclitaxel can greatly alleviate matrix effects. A stable isotope (i.e., deuterium) labeled internal standard may not always fully compensate for matrix effects of its nonlabeled analyte. LLE generally provides cleaner extracts than RP-SPE, while PPT is the least effective tool in removing matrix components. With the same dimensions, columns packed with superficially porous particles have a narrower window of matrix effects free region than those packed with porous particles. Column loading capacity could be a potential explanation for the difference, where columns packed with fully porous particles can handle larger amounts of matrix components without being overloaded. Last, a sub-minute high-throughput chromatography (e.g., 36 s) was employed to evaluate matrix effects of sofosbuvir prepared from PPT extracts of 36 lots of plasma from six common species (six different lots per species) used in drug discovery and development. Essentially no matrix effects were observed for sofosbuvir from all 36 lots of plasma under the sub-minute chromatography. In all, results amply demonstrated that matrix effects can be effectively managed under high-throughput chromatography that is practical and amenable for regulated bioanalysis.

Future perspective

Quantitative bioanalysis is an integral part of the drug discovery and development continuum and plays a vital role throughout the entire process. High-throughput bioanalysis can help shorten the time to bring therapies to market. High-throughput bioanalysis has been widely employed in drug discovery.

Table 1. Precision and accuracy of six levels of sofosbuvir from a sub-minute chromatography (n = 6, 1 run) and a regular chromatography (n = 18, 3 runs).

QC level	Conc. (ng/ml)	Accuracy (%)† (n = 6)	Accuracy (%)‡ (n = 18)
QC 1	5.00	105.2 ± 9.9	98.2 ± 3.9
QC 2	12.5	103.1 ± 3.6	101.9 ± 2.3
QC 3	30.0	107.2 ± 5.5	101.5 ± 2.3
QC 4	120	105.7 ± 3.0	99.5 ± 2.5
QC 5	400	107.4 ± 3.1	99.7 ± 2.5
QC 6	1900	98.3 ± 5.5	96.9 ± 2.7

†Results are obtained one run using a 0.36-min high-throughput chromatography.
‡Results are obtained from three accuracy and precision runs from a fully validated assay using a 2-min chromatography.

The high-throughput in discovery is often realized with minimal or no sample preparation and/or chromatographic separation. Substantial matrix effects are typically inevitable as a result. Nevertheless, the high-throughput bioanalytical methods can provide results adequate for decision making appropriate for the intended purposes. High-throughput bioanalysis can be practical and amenable to regulated bioanalysis as well. It is important to note that high throughput in this context shall be achieved without sacrifice of quality. High-throughput LC–MS methods can be realized via high-throughput chromatography and/or in combination with effective sample clean up, where matrix effects can be minimal. High-throughput bioanalytical methods need to be comprehensively evaluated to ensure method reliability and robustness.

Financial & competing interests disclosure

The authors have no relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript. This includes employment, consultancies, honoraria, stock ownership or options, expert testimony, grants or patents received or pending, or royalties.

No writing assistance was utilized in the production of this manuscript.

Ethical conduct of research

The authors state that they have obtained appropriate institutional review board approval or have followed the principles outlined in the Declaration of Helsinki for all human or animal experimental investigations. In addition, for investigations involving human subjects, informed consent has been obtained from the participants involved.

Executive summary

- A slight but adequate separation, in terms of retention time, between underlying matrix components and compounds of interest, for example, sofosbuvir and paclitaxel, can greatly alleviate matrix effects.
- Matrix effects in LC–MS/MS-based bioanalysis are largely a combined effect of sample preparation and chromatography.
- A deuterium-labeled internal standard may not always fully compensate for matrix effects of its nonlabeled analyte.
- The column packed with superficially porous particles exhibited a narrower matrix effects free elution window than that packed with porous particles of the same size. Column loading capacity may play a role.
- High-throughput chromatography, with proper optimization, can provide rapid and effective chromatographic separation under 1 min to alleviate matrix effects and enhance assay ruggedness for regulated bioanalysis.

References

Papers of special note have been highlighted as:
• of interest; •• of considerable interest

1 Ackermann BL, Berna MJ, Murphy AT. Recent advances in use of LC/MS/MS for quantitative high-throughput bioanalytical support of drug discovery. *Curr. Top. Med. Chem.* 2(1), 53–66 (2002).

2 Balani SK, Miwa GT, Gan LS, Wu JT, Lee FW. Strategy of utilizing *in vitro* and *in vivo* ADME tools for lead optimization and drug candidate selection. *Curr. Top. Med. Chem.* 5(11), 1033–1038 (2005).

3 Jemal M, Xia YQ. LC–MS Development strategies for quantitative bioanalysis. *Curr. Drug. Metab.* 7(5), 491–502 (2006).

4 Swartz ME. UPLC™: an introduction and review. *J. Liquid. Chromatogr. Relat. Technol.* 28(7–8), 1253–1263 (2005).

5 Churchwell MI, Twaddle NC, Meeker LR, Doerge DR. Improving LC–MS sensitivity through increased in

chromatographic performance: comparisons of UPLC-ES/MS/MS to HPLC-ES/MS/MS. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* 825(2), 134–143 (2005).

6 Yu K, Little D, Plumb R, Smith B. High-throughput quantification for a drug mixture in rat plasma-a comparison of Ultra Performance liquid chromatography/tandem mass spectrometry with high-performance liquid chromatography/tandem mass spectrometry. *Rapid Commun. Mass Spectrom.* 20(4), 544–552 (2006).

7 Hsieh Y, Duncan CJ, Brisson JM. Fused-core silica column high-performance liquid chromatography/tandem mass spectrometric determination of rimonabant in mouse plasma. *Anal. Chem.* 79(15), 5668–5673 (2007).

8 Mallett DN, Ramirez-Molina C. The use of partially porous particle columns for the routine, generic analysis of biological samples for pharmacokinetic studies in drug discovery by reversed-phase ultra-high performance liquid chromatography-tandem mass spectrometry. *J. Pharm. Biomed. Anal.* 49(1), 100–107 (2009).

• **Utility of high flow monolithic chromatography in high-throughput bioanalysis.**

9 Deng Y, Wu JT, Lloyd TL, Chi CL, Olah TV, Unger SE. High-speed liquid chromatography/tandem mass spectrometry using a monolithic column for high-throughput bioanalysis. *Rapid Commun. Mass Spectrom.* 16(11), 1116–1123 (2002).

10 Huang MQ, Mao Y, Jemal M, Arnold M. Increased productivity in quantitative bioanalysis using a monolithic column coupled with high-flow direct-injection liquid chromatography/tandem mass spectrometry. *Rapid Commun. Mass Spectrom.* 20(11), 1709–1714 (2006).

11 Xu RN, Fan L, Rieser MJ, El-Shourbagy TA. Recent advances in high-throughput quantitative bioanalysis by LC–MS/MS. *J. Pharm. Biomed. Anal.* 44(2), 342–345 (2007).

12 Korfmacher WA, Palmer CA, Nardo C *et al.* Development of an automated mass spectrometry system for the quantitative analysis of liver microsomal incubation samples: a tool for rapid screening of new compounds for metabolic stability. *Rapid Commun. Mass Spectrom.* 13(10), 901–907 (1999).

13 Jian W, Edom RW, Xu Y, Weng N. Recent advances in application of hydrophilic interaction chromatography for quantitative bioanalysis. *J. Sep Sci.* 33(6–7), 681–697 (2010).

• **Bioanalytical applications of ambient desorption ionization mass spectrometry (DART, DESI).**

14 Chernetsova ES, Morlock GE. Ambient desorption ionization mass spectrometry (DART, DESI) and its bioanalytical applications. *Bioanalyt. Rev.* 3(1), 1–9 (2011).

15 Harris GA, Galhena AS, Fernández FM. Ambient sampling/ionization mass spectrometry: applications and current trends. *Anal. Chem.* 83(12), 4508–4538 (2011).

16 Chen H, Talaty NN, Takáts Z, Cooks RG. Desorption electrospray ionization mass spectrometry for high-throughput analysis of pharmaceutical samples in the ambient environment. *Anal. Chem.* 77(21), 6915–6927 (2005).

17 Manicke NE, Kistler T, Ifa DR, Cooks RG, Ouyang Z. High-throughput quantitative analysis by desorption electrospray ionization mass spectrometry. *J. Am. Soc. Mass Spectrom.* 20(2), 321–325 (2009).

• **Evaluation of laser diode thermal desorption (LDTD)-MS/MS for high-throughput discovery bioanalysis.**

18 Beattie I, Smith A, Weston DJ, White P, Szwandt S, Sealey L. Evaluation of laser diode thermal desorption (LDTD) coupled with tandem mass spectrometry (MS/MS) for support of *in vitro* drug discovery assays: Increasing scope, robustness and throughput of the LDTD technique for use with chemically diverse compound libraries. *J. Pharm. Biomed. Anal.* 59, 18–28 (2012).

19 Wagner AD, Kolb JM, Ozbal CC *et al.* Ultrafast mass spectrometry based bioanalytical method for digoxin supporting an *in vitro* P-glycoprotein (P-gp) inhibition screen. *Rapid Commun. Mass Spectrom.* 25(9), 1231–1240 (2011).

• **Evaluation a high-throughput SPE-MS/MS strategy for bioanalysis with analysis time of approximately 10 s/sample.**

20 Jian W, Romm MV, Edom RW, Miller VP, LaMarr WA, Weng N. Evaluation of a high-throughput online solid phase extraction–tandem mass spectrometry system for in vivo bioanalytical studies. *Anal. Chem.* 83(21), 8259–8266 (2011).

21 Ramanathan R, Jemal M, Ramagiri S *et al*. It is time for a paradigm shift in drug discovery bioanalysis: from SRM to HRMS. *J. Mass Spectrom.* 46(6), 595–601 (2011).

22 Fung EN, Jemal M, Aubry AF. High-resolution MS in regulated bioanalysis: where are we now and where do we go from here? *Bioanalysis* 5(10), 1277–1284 (2013).

23 Xia YQ, Lau J, Olah T, Jemal M. Targeted quantitative bioanalysis in plasma using liquid chromatography/high-resolution accurate mass spectrometry: an evaluation of global selectivity as a function of mass resolving power and extraction window, with comparison of centroid and profile modes. *Rapid Commun. Mass Spectrom.* 25(19), 2863–2878 (2011).

24 Ding J, Lund ET, Zulkoski J, Lindsay JP, McKenzie DL. High-throughput bioanalysis of bile acids and their conjugates using UHPLC coupled to HRMS. *Bioanalysis* 5(20), 2481–2494 (2013).

25 Murphy K, Bennett PK, Duczak N Jr. High-throughput quantitation of large molecules using multiplexed chromatography and high-resolution/accurate mass LC–MS. *Bioanalysis* 4(9), 1013–1024 (2012).

26 Heudi O. High-throughput quantitative analysis of pharmaceutical compounds in biological matrices. *Bioanalysis* 3(8), 819–821 (2011).

27 European Medicines Agency. Guideline on bioanalytical method validation. www.ema.europa.eu

28 US FDA Guidence for Industry. Bioanalytical Method Validation. www.fda.gov/downloads/Drugs/./Guidances/ucm070107.pdf

29 Bergeron A, Garofolo F. Importance of matrix effects in LC–MS/MS bioanalysis. *Bioanalysis* 5(19), 2331–2332 (2013).

•• **Comprehensive review of matrix effects in LC–MS-based bioanalysis.**

30 Li F, Ewles M, Pelzer M *et al.* Case studies: the impact of non-analyte components on LC–MS/MS based bioanalysis – strategies for identifying and overcoming matrix effects. *Bioanalysis* 5(19), 2409–2441 (2013).

•• **Useful strategies to reduce matrix effects in LC–MS-based bioanalysis.**

31 Chambers E, Wagrowski-Diehl DM, Lu Z, Mazzeo JR. Systematic and comprehensive strategy for reducing matrix effects in LC/MS/MS analyses. *J. Chromatogr. B* 852(1–2), 22–34 (2007).

32 Xia YQ, Jemal M. Phospholipids in liquid chromatography/mass spectrometry bioanalysis: comparison of three tandem mass spectrometric techniques for monitoring plasma phospholipids, the effect of mobile phase composition on phospholipids elution and the association of phospholipids with matrix effects. *Rapid Commun. Mass Spectrom.* 23(14), 2125–2138 (2009).

33 Taylor PJ. Matrix effects: the Achilles heel of quantitative high-performance liquid chromatography–electrospray–tandem mass spectrometry. *Clin. Biochem.* 38(4), 328–334 (2005).

34 Gosetti F, Mazzucco E, Zampieri D, Gennaro MC. Signal suppression/enhancement in high-performance liquid chromatography tandem mass spectrometry. *J. Chromatogr. A* 1217(25), 3929–3937 (2010).

35 Trufelli H, Palma P, Famiglini G, Cappiello A. An overview of matrix effects in liquid chromatography-mass spectrometry. *Mass Spectrom. Rev.* 30(3), 491–509 (2011).

• **Useful information on the causes of ion suppression in LC–MS analysis and strategies to address it.**

36 Furey A, Moriarty M, Bane V, Kinsella B, Lehane M. Ion suppression: a critical review on causes, evaluation, prevention and applications. *Talanta* 115, 104–122 (2013).

•• **Demonstration of high-throughput (e.g., sub-minute) LC–MS methods to support regulated bioanalysis.**

37 Badman ER, Beardsley RL, Liang Z, Bansal S. Accelerating high quality bioanalytical LC/MS/MS assays using fused-core columns. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* 878(25), 2307–2313 (2010).

38 Ye Z, Tsao H, Gao H, Brummel CL. Minimizing matrix effects while preserving throughput in LC–MS/MS bioanalysis. *Bioanalysis* 3(14), 1587–1601 (2011).

39 Chromatography Online. Glossary of HPLC/LC Separation Terms. www.chromatographyonline.com

40 Hawkridge AM. Chapter 1: practical considerations and current limitations. in quantitative mass spectrometry-based proteomics. In: *Quantitative Proteomics.* Claire E Eyers, Simon Gaskell (Eds). Royal Society of Chemistry, UK (2014).

41 Tang K, Page JS, Smith RD. Charge competition and the linear dynamic range of detection in electrospray ionization mass spectrometry. *J. Am. Soc. Mass Spectrom.* 15(10), 1416–1423 (2004).

42 Golay MJE. The hetp of an unretained component of fluid flow in a coiled round pipe. *J. Chromatogr. A* 186, 341–351 (1979).

43 Zheng J, Patel D, Tang Q, Markovich RJ, Rustum AM. Comparison study of porous, fused-core, and monolithic silica-based C18 HPLC columns for Celestoderm-V ointment analysis. *J. Pharm. Biomed. Anal.* 50(5), 815–822 (2009).

44 Jiang H, Ouyang Z, Zeng J *et al.* A user-friendly robotic sample preparation program for fully automated biological sample pipetting and dilution to benefit the regulated bioanalysis. *J. Lab Autom.* 17(3), 211–221 (2012).

45 Yu S, Crawford E, Tice J, Musselman B, Wu JT. Bioanalysis without sample cleanup or chromatography: the evaluation and initial implementation of direct analysis in real time ionization mass spectrometry for the quantification of drugs in biological matrixes. *Anal. Chem.* 81(1), 193–202 (2009).

• **Demonstration of high-throughput DESI MS in pharmaceutical analysis.**

Bioanalysis

LC–MS quantification of protein drugs: validating protein LC–MS methods with predigestion immunocapture

Jeffrey Duggan*[,1], Bailuo Ren[1], Yan Mao[1], Lin-Zhi Chen[1] & Elsy Philip[1]

[1]Boehringer–Ingelheim Pharmaceuticals, Ridgefield, CT, USA

*Author for correspondence: jeffrey.duggan@boehringer-ingelheim.com

A refinement of protein LC–MS bioanalysis is to use predigestion immunoaffinity capture to extract the drug from matrix prior to digestion. Because of their increased sensitivity, such hybrid assays have been successfully validated and applied to a number of clinical studies; however, they can also be subject to potential interferences from antidrug antibodies, circulating ligands or other matrix components specific to patient populations and/or dosed subjects. The purpose of this paper is to describe validation experiments that measure immunocapture efficiency, digestion efficiency, matrix effect and selectivity/specificity that can be used during method optimization and validation to test the resistance of the method to these potential interferences. The designs and benefits of these experiments are discussed in this report using an actual assay case study.

First draft submitted: 27 May 2016; Accepted for publication: 20 July 2016; Published online: 17 August 2016

Keywords: assay matrix effects • bioanalytical validation • digestion efficiency • GXP bioanalysis • hybrid bioanalysis • immunocapture • protein bioanalysis • protein drug validation • regulated bioanalysis • selectivity

Background

The use of LC–MS/MS to quantify protein therapeutics via a surrogate peptide following proteolytic digestion (PrD-LC–MS) is fast becoming a mainstream bioanalytical technique [1–3]. Among the advantages it offers are: rapid method development, the use of established techniques and instrumentation, and significant resistance to interference from specific matrix components such as antidrug antibodies (ADAs) and circulating ligands. While one of the shortcomings of PrD-LC–MS is lower sensitivity as compared with ligand-binding assays (LBAs), coupling PrD-LC–MS assays with a predigestion affinity or immunoaffinity capture step can result in significantly increased sensitivity [4]. Assays which employ a drug-specific immunocapture reagent (PrD-LC–MS IC), such as an antidrug monoclonal antibody (mAb) or a recombinant drug target molecule, can remove many of the potentially interfering matrix components prior to digestion, resulting in lower assay LLOQs [4,5]. PrD-LC–MS IC assays are truly hybrid techniques, as they depend upon ligand-binding interactions to separate the drug from matrix components prior to the digestion and LC–MS/MS quantification steps [6]. As such, PrD-LC–MS IC methods should be carefully tested both before and during validation to ensure that they are substantially free from interference from components such as circulating ligands, ADAs and unknown factors in patient matrices.

In a typical LBA the protein drug is captured from matrix using an immunoaffinity reagent bound to a plate or other solid substrate. After wash steps, a second reagent then specifically binds to the drug–reagent complex for the purpose of detection. The immunocapture (IC) step in a PrD-LC–MS IC assay is similar to that of ELISA, but the drug is first

FUTURE SCIENCE part of fsg

eluted followed by digestion and LC–MS detection (in a solid state method with off-bead digestion, as will be described here). Because of the decoupling of the detection steps and added selectivity of LC–MS/MS, it is generally recognized that PrD-LC–MS methods show significant resistance to matrix interference [3]; nevertheless, it is very important to carefully test these methods both before and during validation to detect any unknown matrix effects (MEs) that might block or inhibit the ligand-binding phase during IC or otherwise interfere with the assay.

General validation criteria for PrD-LC–MS assays

As PrD-LC–MS becomes an accepted bioanalytical technique, a primary concern is the use of these methods in regulated studies, GLP Toxicology studies and GCP Clinical studies. In the current regulatory landscape there exist no official regulatory guidances for the range of validation experiments and acceptance criteria that should be used for PrD-LC–MS (IC) assays, but there have been recent papers reflecting the opinions of expert users in both the US/Canada [7,8] and in Europe [9]. These authors recommended that LBA acceptance criteria be used for the validation of PrD-LC–MS (IC) assays in most cases. For precision and accuracy in both validation and bioanalysis, this would mean that the 20/25 criteria would apply: 20% CV and accuracy to the nominal concentration be applied to the quality controls (QCs) and standards, using 25% at the LLOQ concentration [7,9]. Experiments dealing with analyte stability, chromatographic carry over, run size determination and other assay parameters are discussed elsewhere [6–9]. The above authors recommend that those additional parameters that have quantitative acceptance criteria should also be subject to the 20/25% criteria [7,9], in alignment with criteria for large molecule LBAs set by guidances from US FDA [10] and EMA [11].

As mentioned above, the hybrid nature of PrD-LC–MS IC gives it characteristics resembling both small molecule LC–MS/MS quantification and LBA methods. Therefore it is important that these methods be carefully examined both during development/optimization and also during validation [6–9]. A separate discussion of the most important of these parameters for IC methods will follow below. It should also be noted that the additional validation experiments discussed here have recently been used successfully in our labs, but they are only recommendations. We will provide case study examples using the system and validation results that we recently obtained with one of our PrD-LC–MS IC methods and subsequently applied to a clinical study. Because the drug we discuss here is still in development we cannot reveal structural details, but we will provide elements of the assay and drug molecule necessary to understand the application.

The experimental PrD-LC–MS IC procedure

Analytical peptides & capture reagent biotinylation

The protein drug assayed here, MW <100 KDa, consists of three protein subunits connected covalently by linkers. Two unique tryptic peptides were identified, each residing on different subunits of the molecule. These peptides were shown to have minimal matrix interference and to exhibit sufficient sensitivity and stability for the assay. A stable isotope labeled internal standard for each peptide (SIL IS) (N-terminal ARG $^{13}C^{15}N$) was synthesized by GeneScript (Piscataway, NJ, USA) for each peptide and is described in more detail below. One of the peptides was used as the surrogate peptide for quantification and the other as a monitoring peptide for verification purposes. Study sample peptide monitoring was intended to be performed, when needed for verification or trouble shooting, by measuring the IS-corrected peak area ratio of the quantitation peptide to that of the monitoring peptide under the premise that any large change in this ratio could be indicative of *in vivo* cleavage or metabolism of the drug molecule. The capture reagent was an internally developed, nonblocking, non-neutralizing mAb against BI 123ABC. This reagent was used for the assay in biotinylated form. Biotinylation of the capture reagent was performed using an EZ-Link Sulfo-NHS-LC biotinylation kit, Thermo Fisher Scientific, MA, USA. The biotinylation ratio was approximately 5 biotins to each protein molecule. Biotinylated reagent solutions were stored at -80°C prior to use. One micron diameter streptavidin coated magnetic beads were received from Thermo Fisher Scientific as a 10 mg/ml suspension. Beads were prepared freshly for each assay by washing the commercial suspension with Tris-buffered saline buffer (prepared from premeasured powder from Thermo Fisher) with 0.1% Tween 20 (Sigma Aldrich) added after preparation (TBS-T). The washed magnetic beads were diluted with TBS-T to a final concentration of 5 mg/ml beads.

Immunoaffinity capture

In the assay, a 20 µl volume of drug-spiked calibrator or QC, blank, zero or patient plasma was first subjected to IC by incubation with the biotinylated capture reagent (50 µl containing 5 µg of biotinylated antidrug mAb in phosphate buffered saline) for 2 h at room temperature with mixing, using a Glass Col (IN, USA) Pulse Vortex mixer at a setting of 45 (about 70 pulses per min). These same mixing conditions and apparatus were used for all mixing operations described in this paper but with the defined duration. After the drug capture incubation period the mixture containing matrix and biotinylated reagent with bound drug was then incubated with streptavidin-coated magnetic

beads, 50 µl (5 µg beads/µl) per sample, for 1 h at room temperature with mixing as described above. This incubation quantitatively bound the biotinylated capture reagent with any associated drug molecules.

Automated washing & elution of captured drug from magnetic beads

After the incubation used to bind the magnetic beads to the biotinylated antidrug mAb (above), the Thermo Kingfisher Flex bead handler was used to wash the beads and perform the elution in 96-well plates. The process of IC, immobilization on streptavidin coated magnetic beads and 96-well automation of the washing and elution steps is pictured in Figure 1. Each wash or elution step performed with the Kingfisher Flex utilized a different plate containing the indicated wash or elution solution. After completion of each step, the beads were captured magnetically then transported to the plate to be used for the next step. Details of washing and elution are as follows: the beads, with bound biotinylated drug, were washed three-times with 300 µl of TBS-T buffer and then washed once with 300 µl water. For each wash step, the beads were mixed for 4 min in the Kingfisher Flex at medium speed then collected magnetically using a 10 s collection period for three cycles. During method development, these washing procedures were determined to sufficiently remove extraneous matrix and nonspecifically adsorbed materials. The captured drug was eluted with 100 µl of 25 mM HCl. For elution, the beads were mixed for 4 min at medium speed then collected magnetically using a 10 s period for five cycles. This elution procedure was determined to be sufficient to quantitatively remove bound drug over the full range of concentrations tested, 0.5–100 mg/l and at higher concentrations during method development.

Sample digestion

After the HCl elution described above, 20 µl of 1 M Tris-HCl, pH 8.0 was added to each sample well, followed by 5 µl of the SIL IS solution, containing SIL IS for both the surrogate and monitoring peptides (0.1 µg/ml each in water). The internal standards contained the exact same amino acid sequence as the surrogate and monitoring peptides except they contained N-terminal ARG $^{13}C_6$,$^{15}N_4$, allowing each to be separated in the mass spectrometer from its unlabeled counterpart molecule by 10 mass units (5 mass units for doubly charged species). Five microliters of water was added to blank wells. Next, 5 µl of 100 mM Tris(2-carboxyethyl)phosphine (TCEP) was added to each sample (to cleave disulfide bonds) and the plates were incubated at room temperature for 20 min with mixing as described above. After this incubation, samples were alkylated by adding 5 µl of freshly prepared 200 mM iodoacetamide, the plate was covered and incubated with mixing for 30 min at room temperature in darkness. To each well, 5 µl of trypsin solution, 6 mg/ml (TCPK grade, Thermo) and 5 µl of 100 mM $CaCl_2$ were added. The plate was incubated at 37°C overnight for digestion with mixing at medium speed (described above). Digestion was quenched by adding 10 µl of 10% formic acid. The samples were mixed and then clarified by centrifugation for about 2 min at 1000 × *g* prior to LC–MS/MS analysis.

LC & MS

A Waters Acquity UPLC system coupled with a Sciex 6500 triple quadrupole mass spectrometer was used for LC–MS/MS analysis. The UPLC injection volume was 10 µl. Chromatographic separation was achieved using a Waters Acquity Peptide BEH C18 column

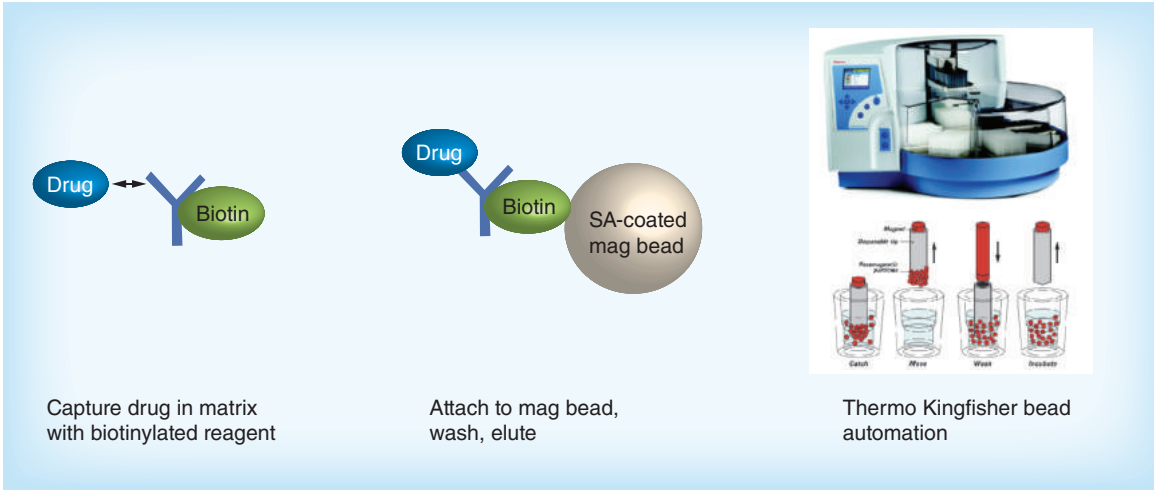


Figure 1. Design of the BI 123ABC immunocapture system using magnetic beads with 96-well automation.
Images at right used with permission of Thermo Fisher Scientific, Inc. [12,13].

(2.1 mm × 50 mm, 1.7 µm) with a Waters Acquity protein BEH C18 1.7 µm VanGuard precolumn and using a column temperature of 45°C. Mobile phases consisted of 0.1% formic acid in water and 0.1% formic acid in acetonitrile. The liquid chromatographic gradient was from 5% B (initial -1.00 min, flow rate 0.3 ml/min), 30% B (1.00–3.00 min, flow rate 0.3 ml/min) to 95% B (3.01 to -3.50 min, flow rate 0.6 ml/min) and back to 5% B (3.51–4.00 min, flow rate 0.6 ml/min; 4.01 min, flow rate 0.3 ml/min). The mass spectrometer was operated in positive ESI multiple reaction monitoring mode. The specific Q1 ion types used for analysis of the quantification and monitoring peptides are 2/Y5 and 2/Y7, respectively. Those for the quantification ion IS and for the monitoring ion IS are 2/Y4 and 2/Y7, respectively. Note that doubly charged peptide ions, yielding Q1 *m/z* < than Q3 *m/z*, are greatly preferred for MS/MS selectivity. Key instrument parameters used for the mass spectrometer were as follows: +5500 V ion spray voltage, 60 nebulizer gas units, 60 axillary gas units, 450°C ion source temperature, 35 curtain gas units, 8 collision gas units and unit resolution on both Q1 and Q3.

PrD-LC–MS IC validation experiments: IC efficiency, recovery, specificity/selectivity & ME

Explanation of validation spiking experiments

Figure 2 illustrates the various experiments that we have used to investigate the factors influencing method efficiencies and recoveries for PrD-LC–MS IC methods. This figure uses graphic references to the entire PrD-LC–MS IC process showing the points where the protein drug, surrogate peptide standard or peptide internal standard are spiked in, to yield the needed comparison. In all cases, the amount of protein drug and any surrogate peptide standard (peptide, red color) or SIL IS (IS, green color) are spiked in at equimolar concentrations.

A note about PrD-LC–MS internal standards: different types of SIL IS have been used successfully with PrD-LC–MS assays [7]. Some SIL IS types that have been utilized include: analog peptides, SIL forms of the surrogate peptide itself [14], SIL surrogate peptides with additional amino acid residues attached (winged SIL IS) [15] and intact drug protein SIL IS molecules. The latter type, intact SIL Protein IS, is sometimes favored because it can be used to monitor the entire assay pro-

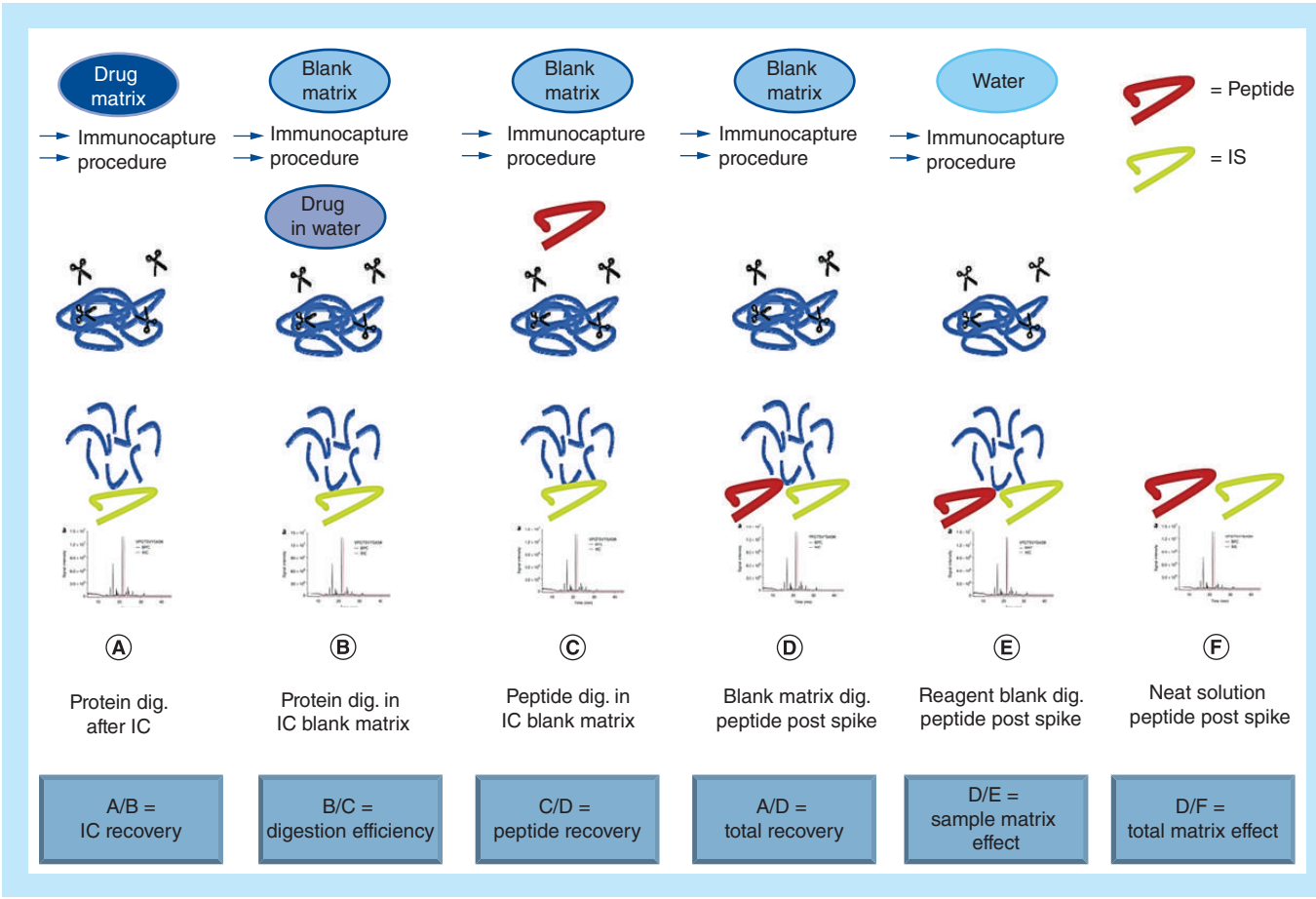


Figure 2. Experiments yielding immunocapture efficiency, digestion efficiency, recovery and matrix effect.

cess [14], but some recent work showed that there was no advantage to using the SIL Protein IS in their assay as compared with the SIL surrogate peptide when used in a transporter protein quantification assay [16]. Further, because of availability and manufacturing issues, SIL proteins are not often available and SIL surrogate peptides or modified versions of these peptides have often been used with good assay results [14,16]. In the studies discussed here, the IS employed is a SIL version of the surrogate peptide containing ^{15}N and ^{13}C with no other modifications or changes to the amino acid sequence.

Immunocapture efficiency experiments

For PrD-LC–MS IC methods, the IC efficiency is calculated from the amount of surrogate peptide recovered from an equimolar amount of protein drug spiked into matrix before (Figure 2A) and after (Figure 2B), the IC step. The comparative results for IC efficiency can provide useful indications of any potential effects that matrix interference might have on the IC part of assay. If ADAs, circulating ligands/targets or other interfering matrix components interrupt the ligand-binding process of drug capture, then the IC efficiency with be expected to decrease with matrix containing these components. This can be tested by spiking matrix with the suspected interferent, performing IC efficiency measurement, then comparing the results to the same experiment done with the appropriate control. For example, if ADA containing plasma samples are available, then the comparative IC efficiency test can be used to gauge the sensitivity of the IC step to interference by ADAs, in the same manner that spike-recovery experiments using ADA containing plasmas are used with LBA PK assays [17]. IC efficiency can also be tested with any potentially interfering ligands that might be present in study matrices, such as shed targets. If a clinical history is available for such circulating ligands, they should be spiked into the blank matrix at concentrations that equal and exceed those from previous clinical determinations, and then the IC efficiency compared with that of control matrix. In addition, healthy matrix can be compared with various individual disease matrices in order to detect any potential interference issues due to unidentified components present in patient matrices. Useful information can also be obtained with simple QC spiking experiments (i.e., selectivity experiments described below using spiked LLOQs), but performing these IC efficiency comparisons using high and low QC concentrations will help to identify any of these possible interferences, specifically linking any detected interferences with the IC step of the assay.

In the case presented here, baseline IC efficiency values were first determined with the same pooled plasma from healthy patients used to test accuracy and precision in the core validation runs. Here, it was necessary to add 0.1% bovine serum albumin (BSA) to the aqueous therapeutic protein spiking solution used for Figure 2, experiment B in order to promote solubility. It was later determined that the 0.1% added BSA had no effect on the IC efficiency when added to either experiment A and B (Figure 2). The data from these experiments (not shown) indicated that capture efficiency was very high (85–95%) and reproducible at both low and high QC levels.

For this method we expected that there might be some interference inherent in the three indicated disease state plasmas. Pre-existing anti-drug antibodies (PEAs) for BI 123ABC were known to exist in some individuals. We selected plasmas from these individuals that were measured with comparatively low and high titers of PEAs in order to determine if they would impact the IC efficiency of this assay as compared with control plasmas with no measurable PEA titer. The results are shown in Table 1. In these experiments every treatment was run in three replicates for both experiments A and B (Figure 2). The mean and %RSD of the calculated A/B ratios is reported in the table. In this case the IC recoveries for ADA plasmas are generally high and very consistent. In comparison to those for pooled healthy plasma. Samples from the pre-existing ADA plasmas were titered in the standard LBA ADA assay used for BI 123ABC studies, and test samples were chosen with relatively low, medium and high titers in order to detect any ADA titer-related effect, and neither such reduction nor trend was observed (Table 1). It should be noted that pre-existing ADAs may have different characteristics from those of emergent ADAs, so the results should be viewed with caution [17,18]. If possible, results should be retested with confirmed emergent ADA samples from the study when they become available.

In the case of the cancer patient plasma, there were three individual plasmas from each of three cancer types. Equal aliquots of these individual plasmas were pooled and used to generate the test plasma for each type. Of the three pooled samples, it can be seen that the breast cancer patient plasma may have slightly reduced the IC efficiency as compared with controls, as the average IC efficiency is decreased by more than 10% for both the low and high QC (Table 1). While it is not practical to place a quantitative specification on IC efficiency, reduced IC efficiency in specific patient samples should be tested further, if possible. We planned tests with predose patient samples at the beginning of the clinical study (data not available). These tests would entail comparing the IC efficiency from actual predose plasmas with that of pooled control plasma.

Table 1. Effect of cancer patient plasma and antidrug antibody containing plasmas on immunocapture efficiency for BI 123ABC.

	Nominal concentration (mg/l)	
	1.50	80.0
Pooled healthy plasma		
Average%	84.7	86.7
RSD%	3.8	3.0
Breast cancer plasma		
Average%	77.5	75.6
RSD%	5.9	3.2
Lung cancer plasma		
Average%	79.6	84.0
RSD%	4.0	22.0
Pancreatic cancer plasma		
Average%	88.4	87.0
RSD%	5.0	4.1
0.1% BSA		
Average%	87.1	85.1
RSD%	4.7	0.8
High ADA plasma		
Average%	83.9	88.3
RSD%	1.6	2.2
Medium ADA plasma		
Average%	86.2	87.2
RSD%	4.1	7.4
Low ADA plasma		
Average%	88.2	86.0
RSD%	2.3	5.2

ADA: Anti-drug antibody; BSA: Bovine serum albumin.

Peptide recovery

Experiment C (Figure 2), where the surrogate peptide is spiked in prior to digestion then its SRM signal ratioed with that from Experiment D, (both the surrogate peptide and SIL IS peptide are added post digestion) yields the surrogate peptide recovery. This is a measure of the ability of the peptide to survive the digestion process. It is important to perform these tests for low, mid and high QC values in replicate (we used n = 3), looking for reasonable and reproducible values in order to both understand and troubleshoot the digestion and post digestion phases of the analysis across the targeted concentration range. The peptide recovery results obtained during validation of BI 123ABC ranged from 85 to 95%, and they were reproducible for both low, mid and high QCs (data not shown).

Digestion efficiency

Digestion efficiency provides a measure of how much of the surrogate peptide is released via proteolytic digestion (PrD), recovered, then measured by LC–MS/MS (Figure 2B), ratioed to that from a post-digestion spike of an equimolar amount of surrogate peptide standard (Figure 2C). Our goal was to determine digestion efficiency at low, mid and high QC levels in triplicate with the same objective of reasonable and consistent efficiency values across the assay concentration range. In our labs, this parameter is retested in a partial validation if there is change in the proteolytic enzyme lot, digestion time or any other parameter that might impact the digestion process. As stated in the previous section, one reason it is important to determine peptide recovery is that it is convoluted with digestion efficiency. One problem that can occur with digestion efficiency determinations can arise from the batch and purity of the synthetic surrogate peptide standard. In our experience, synthetic peptide standards can arrive with a significant but unreported content of salt or other inert impurities. In such cases, the weighing does not properly reflect the compound mass, and the actual molar concentration of peptide is lower than that of the spiked protein standard, resulting in artificially elevated digestion efficiency. Peptide adsorption is another potential issue that may affect results when compare spiked peptide data to spiked protein data (refer to the specificity/selectivity section of this paper). We have experienced some efficiency values measured by this technique of up to and even exceeding 100%. See Table 2, where digestion efficiency values are reproducible, but exceed 100% at all three QC levels. We do not, therefore, specify numeric criteria for digestion efficiency, since the key factor in assessing this parameter is that the values be high and reproducible within the validation with consistency across concentrations and between batches [7]. Although conceptually somewhat different, this approach is consistent with the comments for method recovery in small molecule chromatographic assays stated in the FDA [10] and EMA [11] bioanalytical validation guidances.

Matrix effect

The ME in these experiments was measured by comparing the MS/MS signal (peak area or peak area ratio) from a post digestion spiked processed matrix blank (Figure 2D) to that from a post processed spiked reagent blank (Figure 2E). This experiment essentially measures the LC–MS/MS response contributed only from the biological matrix, since all of the components added to the digestion mixture are passed on to the matrix blank and are also present in the reagent blank. In this respect, these ME experiments are analogous to those performed for small molecules in accordance with the EMA small

molecule guidance [11]. The guidance recommends that ME be determined using individual matrix samples as well as lipemic and hemolysis samples spiked at low and high QC levels [11]. In our lab we run ME for a total of eight individual samples using four individual plasmas, two individual lipemic and two individual hemolyzed samples. These samples are spiked at the low (3× LLOQ) and high QC levels for the assay. Both absolute and IS-normalized matrix factors (MF) can be calculated from these data using the surrogate peptide and IS peak areas as per section 4.1.8 of the EMA bioanalytical validation guidance [11]. In alignment with this guidance, but allowing for the 20/25% PrD-LC–MS IC assay performance specifications [7,9], our acceptance criteria are that the %CV for MF should less than or equal to 20% across all samples tested. In Table 3, we show the results obtained for absolute ME experiments with BI 123ABC spiked at the low and high QC levels into various individual plasmas, including two hemolyzed and two lipemic samples. The peak area for the samples derived from experiment (Figure 2D) (a digested blank, spiked with peptide post digestion) is divided by that from experiment (Figure 2E) (a reagent blank spiked post digestion). In this example the absolute matrix factor ranged from 0.801 to 1.063 with an overall average of 0.968, indicative a low ME. The calculated overall %RSD (Table 3) is 7.07%, less in this particular case, than the acceptance criterion stated for the overall %CV in the EMA guidance [11] Sec 4.1.8 for small molecules (15%). IS-normalized MFs yielded comparable data (not shown). It is also possible to learn the ME contribution of the added reagents and components that are part of the digestion steps by preparing a comparator with the surrogate peptide and IS spiked into water. This experiment is not essential, but it may be done within or outside the validation for information purposes.

Table 2. Digestion efficiency results for BI 123ABC in K_2-EDTA human plasma.

	Nominal concentration (mg/l)		
	1.50	20.0	80.0
Area ratios of postspiked protein samples (see Figure 2B)			
1	0.0753294	1.16110	4.51633
2	0.0710791	1.17483	4.78991
3	0.0717379	1.20771	4.13091
Area ratios of prespiked peptide samples (see Figure 2C)			
1	0.0609257	0.888143	4.45837
2	0.0569186	0.823665	3.59458
3	0.0621229	0.816978	3.89522
Average	0.0599891	0.842929	3.98272
Digestion efficiency%			
1	125.6	137.7	113.4
2	118.5	139.4	120.3
3	119.6	143.3	103.7
Average%	121.2	140.1	112.5
RSD%	3.2	2.0	7.4

ME experiments such as these may be complicated by solubility or adsorption issues when hydrophobic peptides are dissolved in water or solutions with low ionic strength [19]. This can result in ME results that are biased due to a preferential loss of peptide from 'neat solutions' or reagent blanks as opposed to more 'matrix rich' extracts. Similar effects can impact digestion efficiency results, yielding higher than expected values (see above). One recommended solution we have for dealing with this problem is to rely more heavily on selectivity/specificity samples which are run for individual matrices and, without providing direct numeric ME results, they

Table 3. Absolute matrix effect determination from individual blank, hemolyzed and lipemic plasmas.

Sample ID	Area counts					
	Matrix Low	Rgt Blk Low	AMF Low	Matrix High	Rgt Blk High	AMF High
Individual blank plasma 1	4159	4183	0.994	245,843	241,706	1.017
Individual blank plasma 2	4416	4575	0.965	239,999	230,537	1.041
Individual blank plasma 3	4383	4435	0.988	245,327	240,822	1.018
Individual blank plasma 4	4407	4143	1.063	254,286	250,425	1.015
Hemolyzed blank plasma 1	4318	4348	0.993	223,247	278,656	0.801
Hemolyzed blank plasma 2	4204	4659	0.902	240,353	250,880	0.958
Lipemic blank plasma 1	4340	4857	0.894	228,188	240,220	0.949
Lipemic blank plasma 2	4409	4397	1.001	217,855	247,165	0.881
Average	4329	4449	0.975	236,887	247,551	0.960
RSD%	2.3	5.4	5.7	5.3	5.7	8.6

AMF: Absolute matrix factor; Rgt Blk: Reagent blank.

contain the contribution of any ME to the assay and so can act as a reliable indicator of the impact that ME has on quantification at the LLOQ level in the variety of matrices tested.

Selectivity/specificity

It is important to reiterate here that PrD-LC–MS IC assays are hybrid assays containing elements of both ligand-binding and chromatographic methods, and that the approach to selectivity/specificity determination should be thorough. The ligand-binding aspect of a PrD-LC–MS IC assay, that of IC, can be assessed in part by evaluating IC efficiency results in that they are derived from individuals, disease state matrices and QCs containing potential interferents such as ADAs or target molecules known to be in circulation (see 'Immunocapture efficiency' section, Table 1). For implementation of PrD-LC–MS IC selectivity/specificity experiments we have been conducting a combination of the selectivity testing that is done for small molecules and the specificity testing recommended for large molecule LBA assays. For these experiments we use ten individual healthy matrix samples, analyzing both a blank from the individuals and an individual sample spiked at the LLOQ. We have also added ten individual disease state matrix samples where applicable, analyzed as both blanks and as LLOQ spiked QCs. The criterion we have applied for blank evaluation is that there should be no integratable chromatographic signal at the retention time and multiple reaction monitoring channel of the analyte greater than 25% of the LLOQ in eight out of ten blanks. The individual LLOQs should produce eight out of ten results that are within the quantitative acceptance criteria for the LLOQ (within 25% accuracy and precision). We also calculate the total %RSD for the determined concentration across all of the LLOQ selectivity/specificity samples measured with the goal that this RSD should be <30%. This last specification of 30% is not an absolute acceptance criterion, but it is evaluated on a case by case basis with particular attention paid to the nature of any outliers. Overall, this dataset provides good evidence that the method can quantify at the LLOQ with accuracy and precision across the range of matrices that we plan to analyze in the course of the clinical study. We can also determine from these experiments whether there are any significant potentially interfering chromatographic signals or similar trends in any of the matrices tested by examination of blank and spiked chromatograms. This extent of selectivity/specificity testing encompasses any ME in that they will contribute most significantly to observed errors in concentration at the LLOQ. This is the intended purpose to making ME measurements for all bioanalytical studies.

Table 4 shows the results we obtained for the selectivity/specificity testing of BI 123ABC. In these experiments the range of individual plasmas tested for selectivity was increased to include plasmas from three to four individuals for each of the three cancer types within the intended therapeutic scope of the drug, ten individual plasmas from the three disease states that will be studied clinically and ten plasmas from healthy individuals, making a total of 20 individuals tested. Each sample was spiked at the nominal LLOQ of the assay (0.5 mg/l). The determined concentration for each assay was then used to calculate individual relative error values, applying the acceptance criterion that 16 out of 20 individuals must have %RE (percentage relative error) values less than or equal to 25%. In this validation all but one individual passed acceptance criteria, with one sample of healthy subject plasma failing at 36.8% RE.

In addition to the 8/10 passing criteria we also tested the overall mean and %CV as a measure of the of the method's ability to determine the LLOQ accurately and reproducibly across the range of matrices that the assay will be applied to. The nonbinding criterion that we are setting for the overall %CV is 30%. This target is 5% higher than the general LLOQ %CV criterion, but has been relaxed to account for some failures and increased variation that would be expected when quantifying variable matrices at the LLOQ. This has been discussed with colleagues in the field (see section Acknowledgements) who have experienced difficulty in performing ME experiments due to solubility and absorption issues with peptides in aqueous solution [19]. The overall %RE and %CV were 5.3 and 12.5 (Table 4), which indicated very good method specificity for this assay. Collectively, we feel that the above types of selectivity experiments can provide the necessary information, since reliable detection at the LLOQ depends on both selectivity and MEs [6–8], and MEs are integrated into these determinations.

In addition to data such as that derived from Table 4, the chromatographic data from each run can be directly examined for any evidence of interferences or changes in peak shape that may be related to the sample matrix. In any case where a selectivity/specificity problem has been observed in individuals, examination of both the LLOQ spiked chromatogram and the corresponding blank (Figure 3A & B) can provide evidence of the contribution of any chromatographic interference. Examination of the double blank chromatogram form the same run can (Figure 3C) reveal any signal on the IS channel that may have been hidden under the IS peak in the blank (Figure 3B).

Studying dilution effects

PrD-LC–MS IC methods are LBA methods and, as such, there is a possibility that they can exhibit non-

linear dilution effects that might be revealed through dilution and parallelism testing. In the absence of any incurred clinical samples we have performed multiple dilutions in specific matrices (disease and ADAs) and compared the results of the data derived from these dilutions to those from healthy pooled matrix. The EMA guidance requires that higher concentration incurred samples from a patient or animal study samples be diluted at least three-times, then analyzed with the criteria that deviation from the expected concentrations should not exceed 30% ([11] Section 7.1.1.10). This type of experiment is expected to reveal any non-linear dilution effects arising from tested matrices. Parallelism studies as described here could not be performed during this assay validation because patient samples were not available, but would be recommended for PrD-LC–MS IC methods when incurred samples become available. In the absence of incurred samples, we did perform multiple dilutions of a high concentration of drug (2000 mg/l) in disease patient plasma down to the assay QC levels of 1.5, 20 and 80 mg/l and realized agreement of back calculated concentrations within the 20% assay dilution criteria (data not shown).

An additional dilution experiment was performed with our method to test it for possible rapid quantitative drop off above the ULOQ; similar to the prozone effect testing that is performed for LBA methods. This method uses a biotinylated capture reagent with a solid-state magnetic bead system, and the reagent and bead amounts are carefully optimized for the specified assay range. Using more beads or more densely reagent-populated beads might extend the range of the assay more than necessary for its purpose and result in wasted reagents. For this reason the optimized assay becomes nonlinear soon after the plasma concentration exceeds the ULOQ. While hook effects are more likely to occur in LBAs, where the capture and detection reagents are added together [20], the response of our optimized PrD-LC–MS IC assay was tested at several analyte concentrations up to 6× the ULOQ. Plasma samples were spiked with nominal concentrations exceeding the ULOQ (100 mg/l), up to 600 mg/l, to cover and exceed the expected C_{max} of 400 mg/l. The results are shown in Figure 4. This figure shows the fit of 'above the curve' samples prepared at 120, 150, 200 and 600 mg/l to the method's calculated standard curve. The regression was calculated to the ULOQ of 100 mg/l, and then linearly extrapolated. The bias of each data point becomes progressively more negative with increasing concentration, but the response never collapses to near the ULOQ, even at 600 mg/l. This indicates that method is not subject to prozone effect and that plasma concentration exceeding the LLOQ up to at least 600 mg/l will recognized as 'above the curve' during sample assay.

Table 4. Selectivity/specificity experimental results from patient and healthy plasmas from the validation of BI 123ABC.

Sample ID	Nominal concentration 0.500 mg/l	RE%	Result
Cancer patient individual plasmas (conc. mg/l)			
BC 1	0.506	1.2	Pass
BC 2	0.608	21.6	Pass
BC 3	0.579	15.8	Pass
LC 1	0.452	-9.6	Pass
LC 2	0.505	1.0	Pass
LC 3	0.524	4.8	Pass
PC 1	0.497	-0.6	Pass
PC 2	0.511	2.2	Pass
PC 3	0.531	6.2	Pass
PC 4	0.616	23.2	Pass
Overall result			Pass
Healthy individual plasmas (conc. mg/l)			
Ind 1	0.433	-13.4	Pass
Ind 2	0.601	20.2	Pass
Ind 3	0.554	10.8	Pass
Ind 4	0.489	-2.2	Pass
Ind 5	0.684	36.8	Fail
Ind 6	0.513	2.6	Pass
Ind 7	0.499	-0.2	Pass
Ind 8	0.509	1.8	Pass
Ind 9	0.485	-3.0	Pass
Ind 10	0.416	-16.8	Pass
Overall result			Pass

Overall mean = 0.526 (%RE = 5.2%), Overall CV = 12.5%.
BC: Breast cancer; LC: Lung cancer; PC: Pancreatic cancer; RE: Relative error.

Conclusion

PrD-LC–MS IC methods are hybrid LBA–LC–MS methods that require special considerations during method development and validation. In this article we have presented some concepts that go beyond those discussed in recent white papers and reviews toward these considerations for PrD-LC–MS IC methods [6–9]. These procedures have been used in our labs to generate additional data to establish the reliability of our assays when applied to the varying matrices and possible interferents that we expect to encounter in our clinical studies. Among the experiments we discussed here are IC efficiency experiments to help us ensure that various matrix types and potentially interfering ligands will not inhibit the binding of our drug molecule during capture. Other experiments using the techniques illustrated in Figure 2 are used to measure Digestion efficiency, peptide recovery, selectivity/speci-

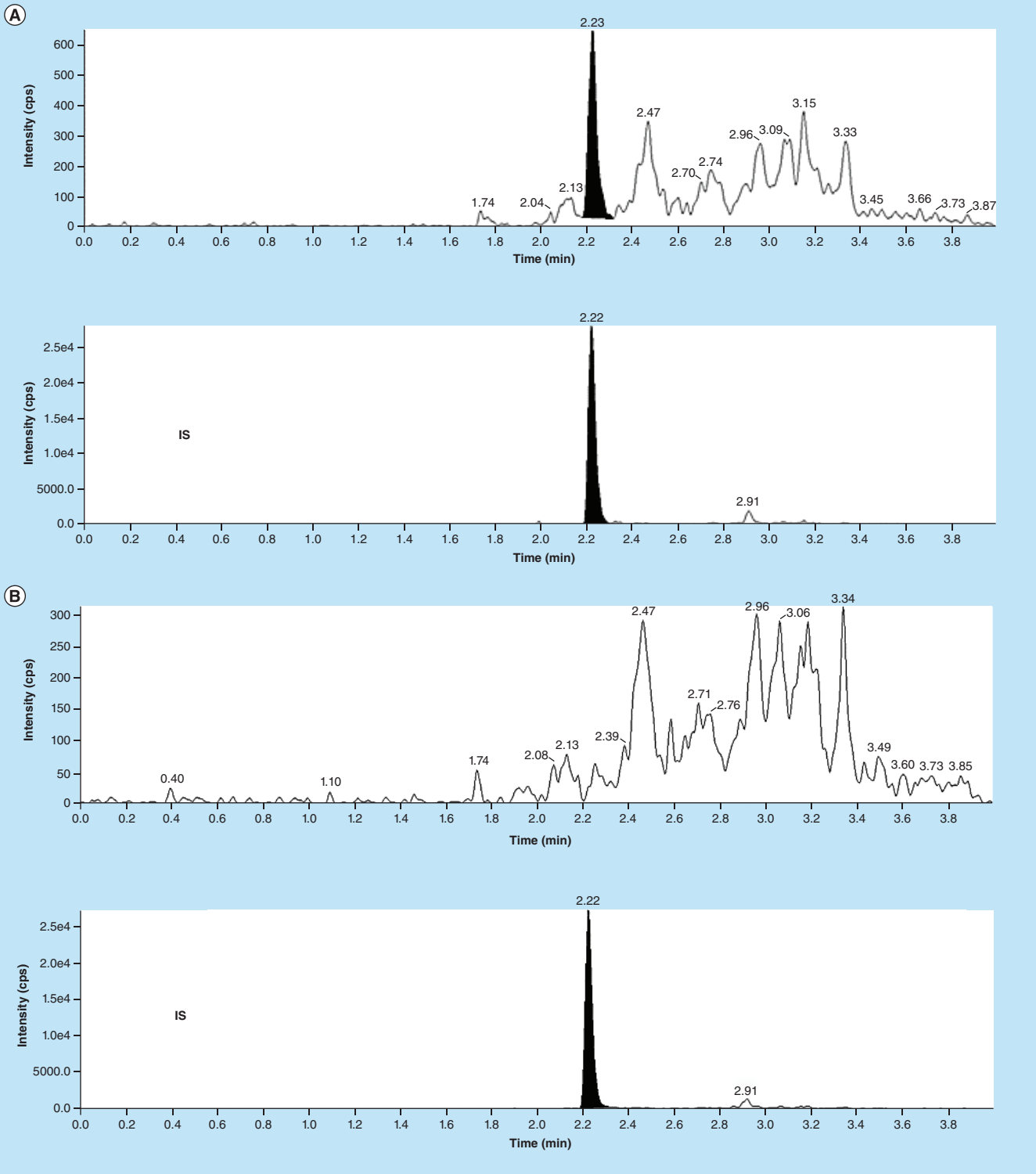


Figure 3. (A) Chromatogram of an LLOQ sample derived from specificity/selectivity testing showing both IS and analyte integrated SRM chromatograms. Upper frame: Analyte channel; lower frame: IS channel. **(B)** Chromatogram of a blank sample derived from specificity/selectivity testing showing IS analyte integrated SRM chromatograms and blank signal. Upper frame: Analyte channel; lower frame: IS channel.
IS: Internal standard.

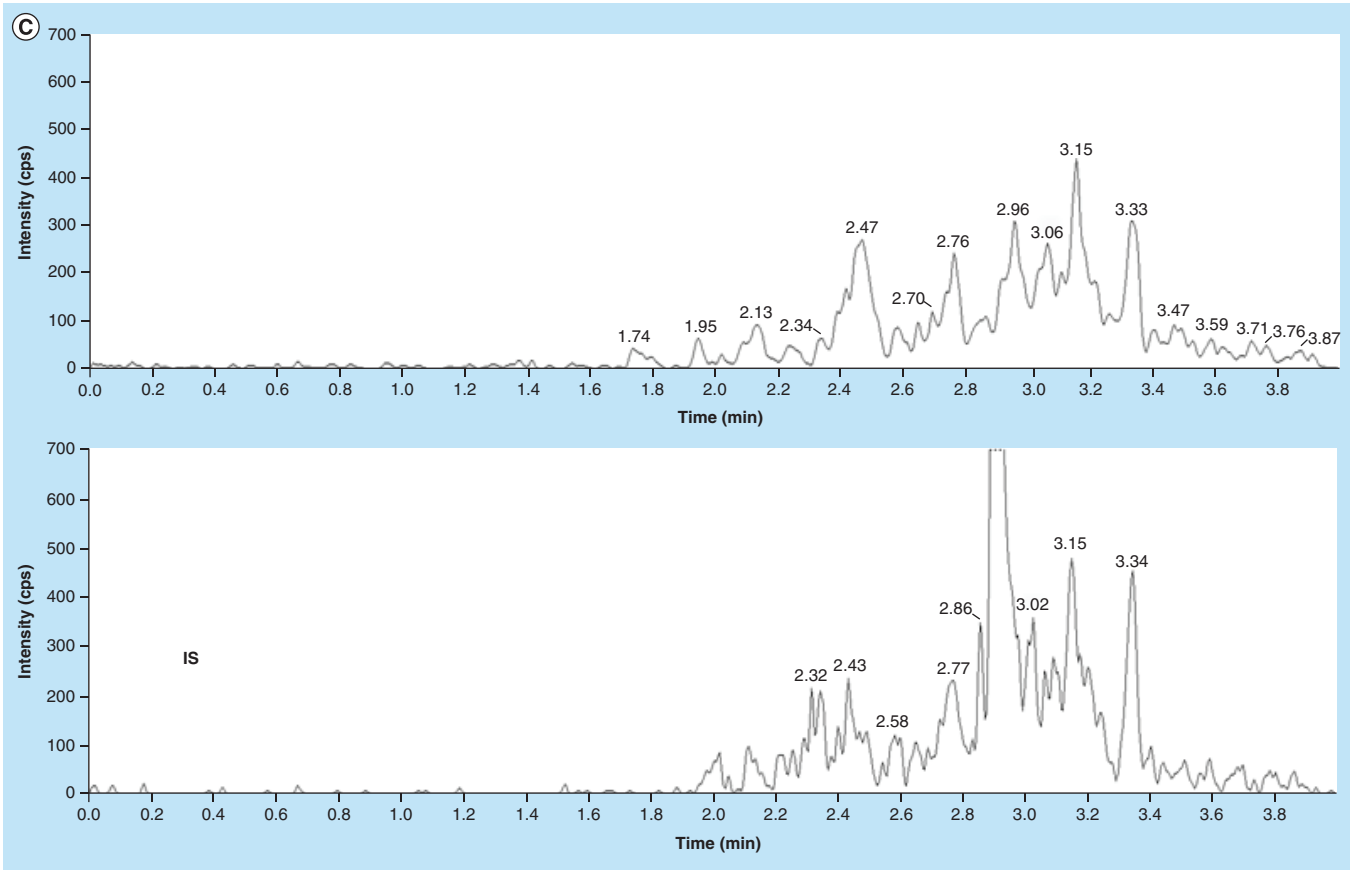


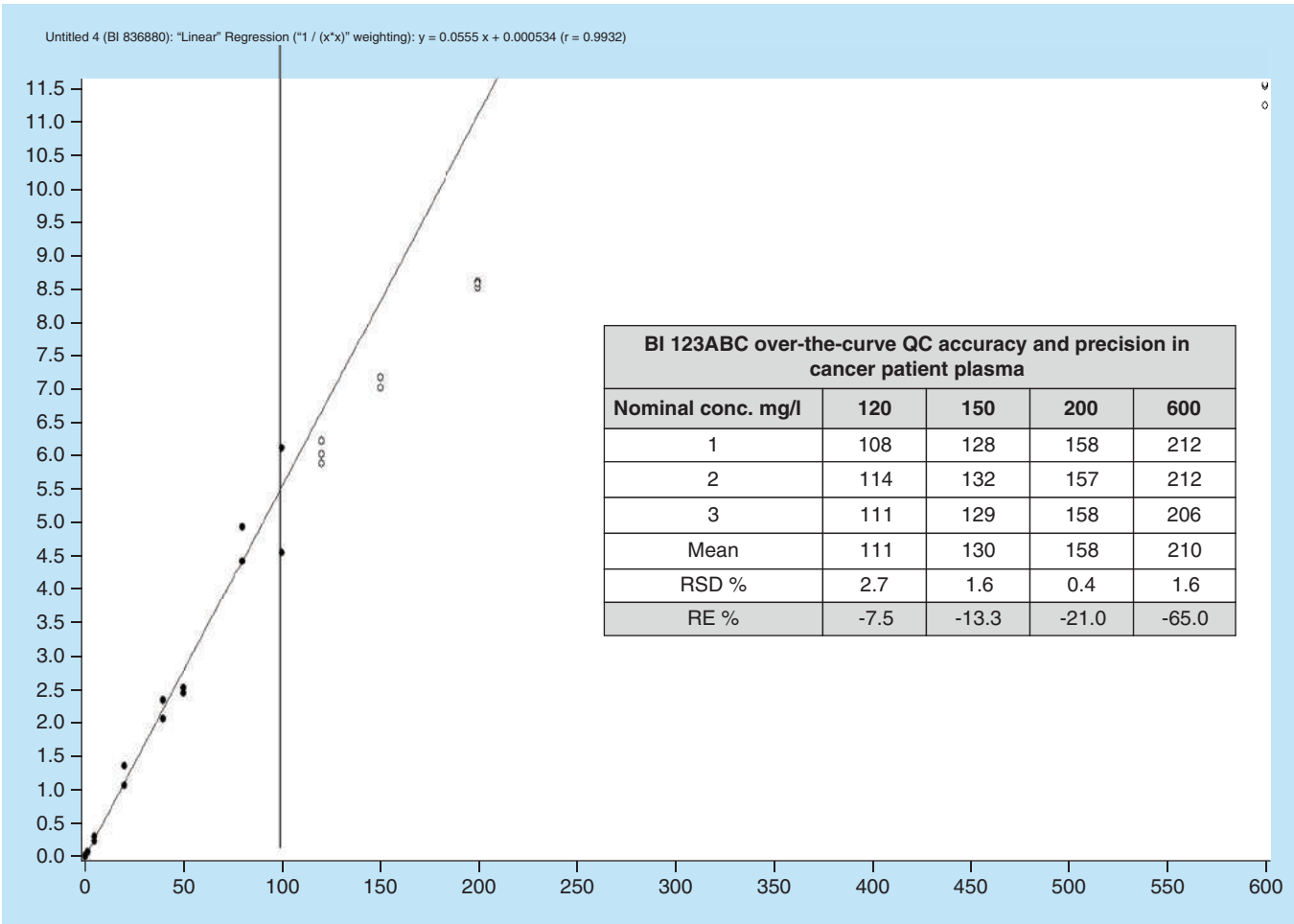
Figure 3 (continued from previous page). (C) Chromatogram of a double blank sample derived from a specificity/selectivity testing run showing IS and analyte SRM chromatograms from the blank extract. Upper frame: Analyte channel; lower frame: IS channel. IS: Internal standard.

ficity, ME and special dilution experiments. The use of these experiments serves both to support assay reliability and to understand how the method is performing at various critical stages of analysis. In the last section on selectivity/specificity experiments we described tests of method performance in individual healthy and disease state matrices both in blanks and in samples spiked at the LLOQ. These experiments can provide valuable insight as to the reliability of the method when applied to matrices that will be analyzed in the clinical study, and they also can provide reliability statistics (overall mean and RSD) that we believe can be indicative of ME, especially in cases where experimental conditions [19] prohibit proper performance of ME tests in accordance with the small molecule EMA guidance [11].

As the technology of PrD-LC–MS continues to progress, more and more of these assays will be used along with LBAs to measure plasma drug concentrations in regulated clinical and nonclinical studies. The regulators have not yet issued an official guidance describing the way in which these assays should be validated, but recent papers have collected the opinions from experts with the aim of providing some preliminary direction for validating these methods [6–9]. PrD-LC–MS IC methods are hybrid in character, and they must be tested and validated with great care to take into account the LBA (IC), digestion and the LC–MS/MS stages of analysis. For the clinical assays we are conducting, it was very important to try to understand the effects of disease state matrices, ADAs and circulating ligands upon our assays. As such we extended the design and scope of our validation experiments to examine these factors. The techniques we used along with some case study examples were described in this paper. We believe that each assay should be approached scientifically, and experiments should be carefully developed to test those specific factors that are likely impact the assay and the studies to which it will be applied.

Future perspective

The use of PrD-LC–MS for protein quantification has grown tremendously over the past 5 years. Recent successful efforts at coupling immunoaffinity purification techniques to PrD-LC–MS have helped to achieve much better sensitivity and method performance than those of previous reagent-free methods. The application of these hybrid IC (PrD-LC–MS IC) techniques is



BI 123ABC over-the-curve QC accuracy and precision in cancer patient plasma

Nominal conc. mg/l	120	150	200	600
1	108	128	158	212
2	114	132	157	212
3	111	129	158	206
Mean	111	130	158	210
RSD %	2.7	1.6	0.4	1.6
RE %	-7.5	-13.3	-21.0	-65.0

Figure 4. Quantitation of samples prepared above the ULOQ of the BI 123ABC assay.

expected to grow greatly over the coming years. As these techniques will be used increasingly in regulated bioanalysis, the importance of additional validation procedures to ensure their in study reliability will also become increasingly important. It will be particularly important to test hybrid LC–MS methods for the impact of circulating ligands and other matrix components that might affect the immunopurification step. We expect that the testing and validation procedures for PrD-LC–MS IC methods will evolve to include most of the experiments described in this paper as well as additional procedures specially developed for individual assay needs. It is our belief that the regulatory guidances now applicable only to traditional small molecule LC–MS and LBAs will be expanded to cover regulate assays using PrD-LC–MS IC, and that these guidelines will contain special provisions to test the IC steps of the assays.

Acknowledgements

The authors would like to thank R Jenkins and W Mylott of PPD for the useful discussion we had, and for shared concepts we discussed related to specificity, matrix effect and peptide solubility/adsorption during assays.

Financial & competing interests disclosure

The authors have no relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript. This includes employment, consultancies, honoraria, stock ownership or options, expert testimony, grants or patents received or pending, or royalties.

No writing assistance was utilized in the production of this manuscript.

Ethical conduct of research

The authors state that they have obtained appropriate institutional review board approval or have followed the principles outlined in the Declaration of Helsinki for all human or animal experimental investigations. In addition, for investigations involving human subjects, informed consent has been obtained from the participants involved.

Executive summary

Background

- LC–MS methods for protein therapeutics that are coupled with immunocapture steps (proteolytic digestion (PrD)-LC–MS IC) are hybrid methods with characteristics of both LC–MS and ligand-binding assay (LBA) methods.
- Because of the potential impact of antidrug antibodies (ADAs), circulating ligands, disease matrix (factors that may directly affect the immunocapture [IC] step) PrD-LC–MS IC methods require special considerations during method development and validation.

The experimental PrD-LC–MS IC procedure

- Description of all experimental aspects of the case study assay.

PrD-LC–MS IC validation experiments: IC efficiency, recovery, specificity/selectivity, matrix effect:

- Explanation of validation spiking experiments
  - Spiking/analysis experiments have been designed to try to determine how each stage of the assay is working with special reference to those that impact the drug capture (IC) step.
- Immunocapture efficiency (IC efficiency) experiments
  - Immunocapture efficiency measures the yield of the assay with analyte spiked matrix containing potential IC inhibitors (ADAs, or circulating ligands) versus the yield from an equimolar amount of drug subjected to direct digestion. IC efficiency can also be compared between spiked disease state matrices and healthy pooled matrix.
- Peptide recovery, digestion efficiency
  - Peptide recovery and digestion efficiency measure the completeness of protein digestion and the ability of the peptide to withstand the digestion process and should be retested if the enzyme or digestion system is changed.
- Matrix effect
  - Matrix effect in PrD-LC–MS systems can determined in the same way as described in the current EMA chromatographic Bioanalytical Method Validation guidance.
- Selectivity/specificity
  - Selectivity/specificity is an important validation parameter because it measures the ability of the assay to perform adequately in various individual matrices that are expected in sample analysis. Special extensions of the selectivity/specificity test includes: the use of ten healthy and ten disease plasmas, measurement of assays for each sample spiked at the LLOQ and extraction/analysis of blanks from each sample.
  - The acceptance criteria we apply is eight out of ten LLOQs must be within 25% RE and CV, and eight out of ten blanks must have no signal at the analyte channel greater that 25% of the average LLOQ.
- Studying dilution effects
  - Dilution experiments were designed to measure any signal drop or possible prozone effect beyond the ULOQ, and we recommend performing a parallelism experiment once ISR samples become available.
  - For each assay and intended sample type, we recommend individually determining the most important factors that may influence the assay, and prepare validation experiments accordingly.

Conclusion

- The described procedures generate additional data to test the resistance of assays to potentially interfering ligands and matrix components.
- These tests: IC efficiency, digestion efficiency, peptide recovery, selectivity/specificity, matrix effects and dilution experiments monitor method performance with potential interferent-spiked and disease state matrices at various stages of the analysis.
- A specific case study was used to exemplify the experiments performed and to discuss interpretation of results.
- Experiments should be specifically tailored to the individual assay and the studies to which it will be applied.

References

Papers of special note have been highlighted as:
• of interest; •• of considerable interest

1 Eweles M, Goodwin L. Bioanalytical approaches to analyzing peptides and proteins by LC–MS/MS. *Bioanalysis* 3(12), 1379–1397 (2011).

2 Van Den Broek I, Van Dongen W. LC–MS-based quantification of intact proteins: perspective for clinical and bioanalytical applications. *Bioanalysis* 7(15), 1943–1958(2015).

3 Zheng J, Mehl J, Zhu Y, Xin B, Olah T. Application and challenges in using LC–MS assays for absolute quantitative analysis of therapeutic proteins in drug discovery. *Bioanalysis* 6(6), 859–879 (2014).

4 Ramagiri S, Moor I. Hybridizing LBA with LC–MS/MS: the new norm for biologics quantification. *Bioanalysis* 8(6), 483–486 (2016).

5 Ackerman B, Hendrik N, Hughes N *et al.* White Paper on recent issues in bioanalysis: focus on new technologies and biomarkers (Part 2 – hybrid LBA/LCMS and input

from regulatory agencies). *Bioanalysis* 7(23), 3019–3034 (2015).

6 Duggan J, Vazvaei F, Jenkins R, Bioanalytical method validation considerations for LC–MS/MS assays of therapeutic proteins. *Bioanalysis* 7(11), 1389–1395 (2015).

• **Provides additional perspectives on the validation of LC–MS protein methods, defining some of the terms used in the technique and paying special attention to considerations for hybrid methods, selectivity and critical reagents.**

7 Jenkins R, Duggan J, Aubry A *et al.* Recommendations for validation of LC–MS/MS bioanalytical methods for protein biotherapeutics. *AAPS J.* 17(1), 1–16 (2015).

•• **Provides a valuable perspective from US and Canadian industry experts on recommendations for validating protein LC–MS methods for use in regulated studies.**

8 Vazvaie F, Duggan J. Validation of LC–MS/MS bioanalytical methods for protein therapeutics. *Bioanalysis* 6(13), 1739–1742 (2014).

• **Provides additional perspective on the validation of protein LC–MS methods, expanding on the original White Paper by Jenkins R, Duggan J, Aubry A *et al.*** [7].

9 Knutsson M, Schmidt R, Timmerman P. LC–MS/MS of large molecules in a regulated bioanalytical environment – which acceptance criteria to apply? *Bioanalysis* 5(18), 2211–2214 (2013).

•• **Provides an interesting and useful discussion of the primary validation parameters and acceptance criteria to use for protein LC–MS methods from the perspective of the European bioanalytical scientists.**

10 *Guidance for Industry: Bioanalytical Method Validation.* US FDA, Rockville, MD, USA (2001).

11 European Medicines Agency, Committee for Medicinal products for Human Use (CHMP). *Guideline on Bioanalytical Method Validation.* European Medicines Agency, London, UK (2011).

12 ThermoFisher Scientific. KingFisher™ Flex Purification System. www.thermofisher.com/order/catalog/product/5400610

13 THERMO-LABSYSTEM KINGFISHER Magnetic Particle Processor. www.labotal.co.il/products/1346/1754/1822.aspx

14 Bronsema K, Bischoff R, Van de Merbel NJ. Internal standards in the quantitative determination of protein biopharmaceuticals using liquid chromatography coupled to mass spectrometry. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* 893–894, 1–14. (2012).

• **Provides both perspective and data for the advantages and disadvantages of a number of protein LC–MS internal standard types, ranging from intact stable isotope-labeled proteins to peptide analogs to cleavable peptides.**

15 Faria M, Halquist M, Yuan M *et al.* Comparison of a stable isotope labeled (SIL peptide and an extended SIL peptide as internal standards to track digestion variability of an unstable signature peptide during quantification of a cancer biomarker, human osteopontin, from plasma using capillary microflow LC–MS/MS. *Chromatogr. B* 1001, 156–168 (2015).

16 Prasad Bhagwat, Unadkat Jashvant D. Comparison of heavy labeled (SIL) peptide versus SILAC protein internal standards for LC–MS/MS quantification of hepatic drug transporters. *Int. J. Proteomics* 2014, 451510 (2014).

17 Sailstad J, Armaravadi L, Clements-Egan A. A White Paper – consensus and recommendations of a global harmonization team on assessing the impact of immunogenicity on pharmacokinetic measurements. *AAPS J.* 16 (3), 488–498 (2014).

18 Gorovitis B, Clements-Egan C, Birchler M *et al.* Pre-existing antibody modality-based review. *AAPS J.* 18(2), 311–320 (2016).

19 Kristensen K, Henriksen J, Andersen T. Adsorption of cationic peptides to solid surfaces of glass and plastic. *PLoS ONE* 10(5), e0122419 (2015).

20 Tate J, Ward G. Interferences in immunoassay. *Clin. Biochem. Rev.* 25, 105–108 (2004).

Bioanalysis

Validation of an LC–MS/MS method for the simultaneous quantification of dabigatran, rivaroxaban and apixaban in human plasma

Aim: Novel oral anticoagulants are characterized by a wide therapeutic window, yet the determination of their plasma–drug concentrations may be useful in some clinical conditions. **Results:** An LC–MS/MS method for the analysis of dabigatran, apixaban and rivaroxaban in human plasma has been successfully developed and validated. The analysis of plasma samples from patients given other concomitant drugs revealed no significant interference. By reanalysis of samples from patients on anticoagulant therapy, we found the percentage difference in results between the concentration of repeat and the original sample to be within the threshold limit of 20% in 60 of 63 specimens. **Conclusion:** The developed LC–MS/MS assay is easily applicable in the clinical management of patients on anticoagulation therapy.

First draft submitted: 7 September 2015; Accepted for publication: 18 December 2015; Published online: 25 January 2016

Keywords: anticoagulants • apixaban • bioanalytical method validation • dabigatran • LC–MS/MS • rivaroxaban

Sara Baldelli*[,1], Dario Cattaneo[1], Pasquale Pignatelli[2], Valentina Perrone[1], Daniele Pastori[2], Sonia Radice[1], Francesco Violi[2] & Emilio Clementi[3,4]

[1]Unit of Clinical Pharmacology, L. Sacco University Hospital, Milano, Italy
[2]Department of Internal Medicine & Medical Specialities, Sapienza University of Rome, Roma, Italy
[3]Clinical Pharmacology Unit, CNR Institute of Neuroscience, Department of Biomedical & Clinical Sciences, L. Sacco University Hospital, Università di Milano, Milano, Italy
[4]Scientific Institute IRCCS Eugenio Medea, Bosisio Parini, Italy
*Author for correspondence:
Tel.: +39 0250319619;
Fax: +39 0250119646;
baldelli.sara@hsacco.it

Background

Novel oral anticoagulants (NOACs) have emerged as an alternative therapy to vitamin K antagonists (VKAs) for the treatment of venous thromboembolism (VTE) and stroke prevention in patients with nonvalvular atrial fibrillation (AF) and for the prevention of VTE after major orthopedic surgery.

Inconclusive but encouraging data from case reports are also available providing preliminary evidences that these drugs may be effective for the prevention and treatment of hepatin-induced thrombocytopenia [1].

NOACs act directly at different sites of the coagulation cascade by inhibiting coagulation pathway factors: dabigatran blocks the transformation of fibrinogen to fibrin through thrombin (factor IIa) inhibition, while rivaroxaban and apixaban target factor Xa prevent the conversion of prothrombin (factor II) to thrombin (factor IIa) [1]. These drugs are characterized by a therapeutic window wider than that of VKAs, therefore they are administered at fixed doses and no routine monitoring is usually required [2]. However, data from PK studies of dabigatran, rivaroxaban and apixaban as well as from large clinical trials have shown a slight intraindividual, but a marked interindividual variability in drug concentrations [3,4]. A limitation of the fixed-dose therapeutic regimen is that some patients may be potentially exposed to inadequate drug concentrations, a condition lasting for all the extent of the treatment [5], this would eventually lead to suboptimal or excessive anticoagulation with possible therapeutic failure or development of adverse drug reactions including major bleeds [6]. Individual monitoring of blood concentrations of NOACs could represent a useful tool to improve drug management in selected clinical scenarios provided that specific, accurate and feasible methods for the quantification of NOACs are available.

The standard laboratory coagulation tests of prothrombin time (PT), activated partial thromboplastin time (APPT) and throm-

FUTURE SCIENCE part of fsg

bin time (TT) used for VKAs, cannot be adequate for NOACs monitoring as these assays are not reproducible and linear throughout a wide range of drug concentrations [2,7]. In addition, the development of PD assays is challenging because NOACs have different targets and, even those with the same target have variable effects on routine coagulation assays.

An alternative approach is the therapeutic monitoring of NOAC plasma concentrations, at least in selected clinical conditions, by taking advantage of LC–MS/MS, which is now a standard tool in many routine clinical laboratories for the determination of small molecule concentrations in biological fluid. Therapeutic drug monitoring has indeed been proposed as a way to minimize issues of under- or over-dosing of NOACs [6]. The assessment of NOAC concentrations may be useful also in the case of emergency bleeding events due to overdose or bioaccumulation, such as in patients with deteriorating renal or hepatic function or when a drug-to-drug interaction is suspected, and also to assess compliance of patients to therapy [3,4]. To date procedures to assess plasma concentrations of all NOACs in a single analytical run easily applicable in the clinical practice are limited [8–10].

Here, we have developed an LC–MS/MS method for quantification of dabigatran, rivaroxaban and apixaban in human plasma samples from patients given NOAC as part of their maintenance anticoagulation therapy. As an added value to available literature, the proposed method evaluated the matrix effect and incurred sample reanalysis, as specifically requested by the recently released EMA guidelines for Bioanalytical Method Validation [11].

Materials & methods

Chemicals & reagents

Apixaban and its internal standard [^{13}C,$^{2}H_3$]-apixaban (apixaban IS) were kindly provided by BMS (Milan, Italy, purity: 99.83 and 99% respectively), while Rivaroxaban powder was kindly provided by Bayer (Milan, Italy, purity 100%). Dabigatran and [$^{2}H_6$]-dabigatran (dabigatran IS) were purchased from Spectra (Rome, Italy, purity: 95 and 96%, respectively). The chemical structure of the NOACs and their internal standard is represented in Figure 1.

Stock solution of apixaban IS (in dimethyl sulfoxide [DMSO]/methanol) and dabigatran IS (in acetonitrile/HCl 0.1N) at 1 mg/ml were diluted with methanol to obtain working solutions at 10 µg/ml.

Stock solutions containing 1 mg of apixaban, and rivaroxaban were prepared in appropriate volumes of DMSO and methanol according to their solubility product, while for dabigatran the concentration was of 0.2 mg/ml in methanol and HCl 0.1N (80:20) and was aliquoted. An intermediate solution was prepared at the concentration of 20 µg/ml by mixing appropriate volumes of stock solutions with methanol. By dilution of this intermediate solution with methanol working solutions of NOAC of 20 ng/ml, 200 ng/ml and 2 µg/ml were prepared. All the solutions were stored at -20°C.

Acetonitrile and methanol (VWR, Milan, Italy) were HPLC grade, while formic acid was LC-MS grade. Ammonium acetate, analytical grade, was purchased from Fluka and a 2 mmol/l solution was prepared in Milli-Q type water obtained from a Milli-Q water purification system (Millipore, Milan, Italy).

Drug-free blank plasma used for the assessment of matrix effect and for the preparation of calibration and control samples was obtained from healthy volunteers participating in local blood donation programs after written informed consent collection.

Calibration curves & quality control samples

Quantitative analysis of the NOACs was performed using the internal standard method by plotting the ratio of the peak area of the analyte to that of the internal standards versus the nominal drug concentration. Apixaban IS was used for both apixaban and rivaroxaban quantification. Each level of the calibration curve was prepared in each chromatographic run and analyzed twice: one at the beginning and one at the end of the run.

Calibrator samples were prepared by mixing appropriate volumes of NOACs working solutions with plasma to achieve different concentrations from 1 to 500 ng/ml (namely 1, 2, 5, 10, 40, 100, 200, 500 ng/ml) so as to cover the range of clinically relevant concentrations expected in patients.

Quality control (QC) samples were prepared to obtain four sets of NOAC concentrations: at the lower limit of quantification (1 ng/ml), at low (3 ng/ml), medium (30 ng/ml) and high (300 ng/ml) drug levels. The QC samples were stored frozen at -20°C until analysis.

Sample preparation

Plasma samples were obtained after centrifugation of blood samples collected in EDTA tubes at 1730 × *g* for 10 min at 4°C and, after separation, stored at -20°C into labeled disposable polypropylene Eppendorf tube.

Solid phase extraction was applied on C18 Bond Elute Cartridge (1 ml, 100 mg, Agilent Milan, Italy) previously activated with methanol and water. To 200 µl plasma based calibrators, quality control samples and study samples, 10 µl IS solutions were added. Samples were centrifuged at 4°C at 10,000 × *g* and the supernatant was diluted with 350 µl of water and loaded on the cartridge. Then the cartridges were washed with water and elution was performed with 500 µl methanol. An aliquot of 100 µl was

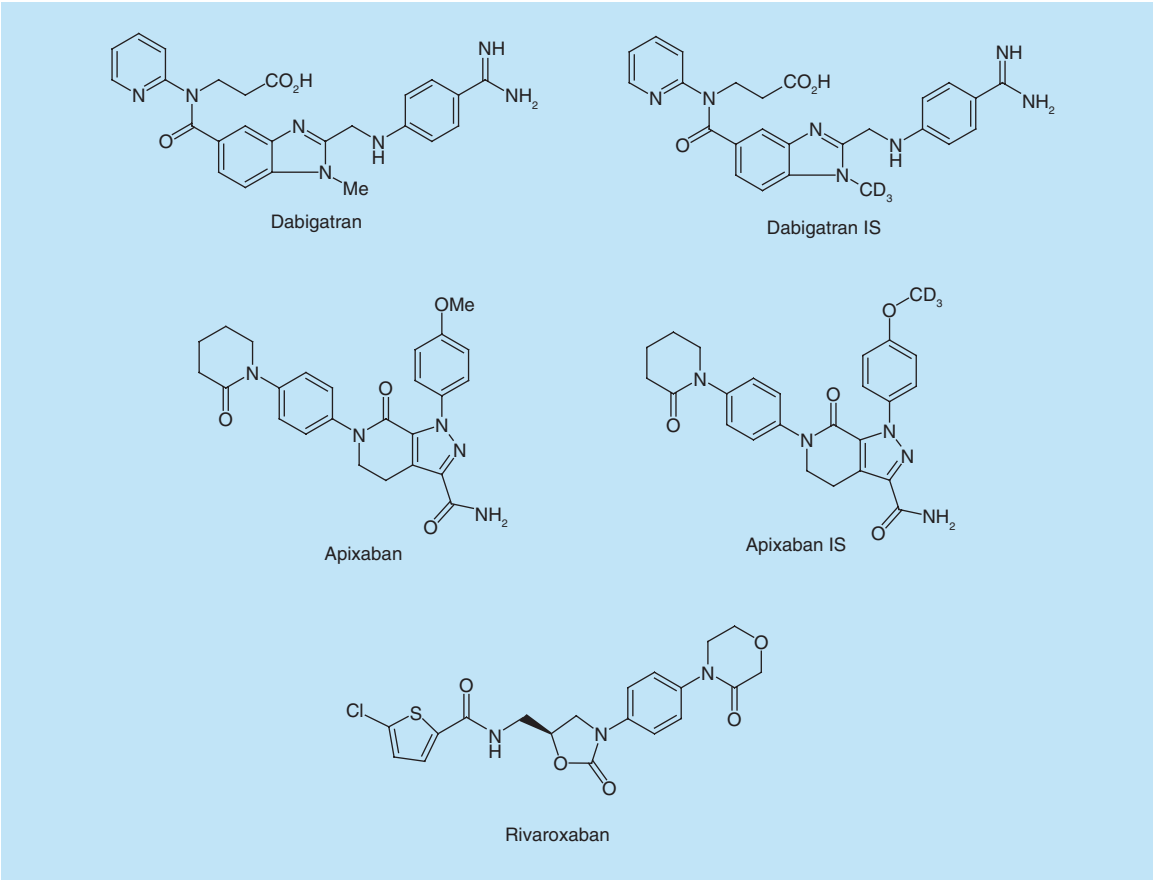


Figure 1. Chemical structure of dabigatran, apixaban, rivaroxaban and their internal standards.

diluted with 100 µl of 0.1% formic acid, filtered and transferred into an HPLC glass vial.

HPLC apparatus & conditions

Samples were analyzed using a Quattro Premier XE triple quadrupole (Waters, Milan, Italy) system coupled with a Waters Alliance 2695 HPLC. Chromatographic separation was performed under gradient conditions with a mobile phase composed of $CH_3COONH_4$2 mmol/10.05% formic acid (solvent A) and 0.1% formic acid in methanol (solvent B). The mobile phase was delivered at 0.3 ml/min using a linear gradient elution program from 15 to 85% of B in 5 min and a re-equilibration step to the initial solvent composition up to 10 min. Solvents were regularly prepared prior to each series of analysis. A reversed-phase C18 column (Zorbax Eclipse plus, 2.1 × 100 mm, 3.5 µm particle sizes, Agilent, Milan, Italy) set to 40°C was employed. The autosampler was maintained at 10°C and the injection volume was of 20 µl.

Positive electron spray ionization (ESI) mode was applied with nitrogen serving as the desolvation gas and high-purity argon as the collision gas. All analytes were quantified by a multiple-reaction monitoring (MRM) mode with the transitions *m/z* 472.05>289.2 for dabigatran (collision energy 30 eV), *m/z* 460.1>443 for apixaban (collision energy 25 eV) and 436>144.9 for rivaroxaban (collision energy 27 eV) while for apixaban IS 464.1>447 (collision energy 25 eV) and for dabigatran IS 475.05>292.2 (collision energy 30 eV). Optimal mass tuning parameters and MS/MS transitions were determined by direct infusion into the MS/MS detector of methanol/water 50/50 solutions of all drugs at a concentration of 1 g/ml using an infusion pump connected directly to the electrospray source.

Method validation

The method validation procedure was based on the recommendations of first guidance from EMA [11]. For the intra-assay precision and accuracy determination replicate analysis (n = 5) of QC samples at the four concentrations used were analyzed. Interassay accuracy and precision were determined by repeated analyses performed on five different days. The precision was calculated as the within and between days coefficient of variation (CV%), and the accuracy as the percentage of deviation between nominal and measured concentrations. The lowest identifiable discrete and reproducible concentration that showed an S/N ratio of 10, an imprecision of 20% and accu-

racy of 80–120% was accepted as the lower limit of quantitation (LOQ).

The selectivity of the method was evaluated as a lack of matrix or other concomitant medications interference by analysis of human drug-free plasma samples from different healthy volunteers (n = 6) and from patients under therapy with most common antifungal, antihypertensive, hypolipidemic and antiepileptic drugs. Matrix effects (MF) were quantitatively investigated at low (3 ng/ml) and high concentrations (300 ng/ml) by calculating the ratio of the peak area in presence of matrix (spiked blank plasma) to the peak area in absence of matrix (standard solution at the same concentration). IS normalized matrix factors (NMF) were calculated by dividing the MF of the analytes by the MF of ISs using six different plasma batches. The CV of the NMF should be below 15%. Furthermore, ion suppression was also investigated by simultaneously post-infusing a standard solution of the analytes and the ISs. [12] Carry-over was estimated by the ratio between the area of the peaks in the blank sample (n = 5) after the injection of the highest calibrator (500 ng/ml) and the area of the first calibrator (1 ng/ml).

To determine the extraction efficiency, the peak area ratios of spiked plasma samples were compared with those obtained from direct injections of the same amount of NOACs and IS.

The robustness of the method was verified by using different batches of analytical columns and by assessing the impact of little variations (i.e., 1%) in mobile-phase composition and column temperature (±5°C).

To ascertain whether the dilution of samples, in case of concentrations above the determined range of quantification could affect the accuracy of drug determination, two- and five-fold dilutions of QC samples at medium and high concentrations with blank plasma were performed.

The stability of the analytes was investigated on QC samples after storage at room temperature for 24 h, at +4°C for 36 h, and after three freeze–thaw cycles from -20°C to room temperature.

Post-preparative stability in the autosampler was also assessed by leaving the extracted sample at +4°C for 36 h. Finally, the long-term stability was also assessed by keeping one set of aliquots at 20°C for 2 months. For all experiments, percent difference in analyte concentrations was determined by comparison to the nominal levels. The relative standard deviations of the set of three samples were calculated.

To test the applicability of the method, samples from 42 AF patients treated with NOACs were analyzed: 16 were men and 26 were women, mean age was 75.2 ± 7.9 years. Mean body mass index was 27.8 ± 5.6 kg/m^2, mean serum creatinine 1.0 ± 0.3 mg/dl. Thirty-one (73.8%) patients were affected by arterial hypertension, and six had Type 2 diabetes. All patients provided written informed consent in accordance with the guidelines of the Helsinki Declaration.

Incurred sample reanalysis was performed on patient samples at different concentrations (n = 14 for apixaban, n = 25 for dabigatran and n = 24 for rivaroxaban). Samples were analyzed in two different analytical runs, the bias of repeat ([result 2 – result 1]/result 1 × 100) was calculated and had to be within 20% in at least 67% of the repeats as requested [11].

The acceptability of re-analyzed samples was also investigated using the method recently proposed by Rocci *et al.* [13].

Results

Under the chromatographic conditions described, retention times were 5.2 min for dabigatran and its IS and 6.6 min for rivaroxaban and apixaban (Figure 2). The calibration curves were satisfactorily fitted by quadratic regression with a weighting factor of 1/nominal concentration for all drugs. The mean correlation coefficients for regression equations, generated for six different days, were 0.9998 (SD: ±0.0002; range: 0.9995–9991.0000) for dabigatran, 0.9996 (SD: ±0.0004; range: 0.9991–9991.00000) for apixaban and 0.9995 (SD: ±0.0006; range: 0.9987–9981.0000) for rivaroxaban. The percentage deviations from nominal values determined for mean back calculated concentrations for each calibrator, ranged from -6.1 to +9.8%, from -2.2 to 3.0% and from -3.7 to +5.9% for dabigatran, apixaban and rivaroxaban, respectively, indicating a good fit of the data to the weighted quadratic regression equation for the three analytes.

The equations obtained were used to calculate the unknown concentrations in patient samples and in QCs. Deviations of the back calculated concentrations in QCs were within ±15% of the nominal concentrations as reported in Table 1. A concentration of 1 ng/ml was set as LOQ for all analytes.

The analysis of plasma samples from healthy volunteers and patients also receiving other concomitant drugs revealed no significant interference with the MRM ion transitions of the analytes.

No ion suppression was observed in correspondence of analytes and IS retention times. Mean NMF for apixaban was 0.86 ± 0.08 (with a value of CV% of 8.9%) for dabigatran 1.08 ± 0.16 (with a CV% of 14.0) and for rivaroxaban was 0.88 ± 0.13 (with a CV% of 14.0). Post-infusion of analytes showed minimal influence of the matrix analyte retention time (Figure 3).

The ratio between the area of the peaks in the blank sample after the injection of the highest calibrator and the area of the first calibrator was 8.3% of the signal

measured with the lowest calibrator, set at 1 ng/ml (carry-over effect = 0.083 ng/ml). The extraction efficiencies demonstrated no concentration dependency, and were similar (102.2, 82.2 and 98.4% for apixaban, dabigatran and rivaroxaban, respectively) between each parent and its isotope compounds.

The retention time of the analyte peaks showed minimal fluctuation (<0.5 min) by using different batches of analytical columns as well as by changing solvent composition or column temperature.

Sample dilution effects were also not observed, as a twofold dilution of plasma samples containing concentrations of 30 and 300 ng/ml as well as a fivefold dilution of plasma at the concentration of 300 ng/ml gave concentrations that were <±7% of their nominal concentration for apixaban, <±6.1% for dabigatran and <±4.3% for rivaroxaban.

The stability of drugs in plasma QCs samples left at room temperature (RT) and at +4°C was ascertained up to 48 h. The variation over time of each drug levels was mostly comprised within the ±10% of nominal concentrations indicating that, taking into account the analytical variability, these drugs can be considered generally stable. The variation of drug concentrations after three freeze–thaw cycles revealed no significant loss for apixaban and rivaroxaban (< 8.7%), on the contrary at all concentration tested, for dabigatran there was an increase of 20%, that was not present after a single freeze–thaw cycle. Quality control samples were prepared in batches, aliquoted and stored at -20°C for use up to 2 months in our routine monitoring laboratory. No evidence of decomposition was found during plasma samples storage in the freezer at -20°C for at least 9 weeks.

This analytical method was applied for the determination of plasma concentration in almost 150 samples from patients on anticoagulant therapy (52% on dabigatran, 33% on rivaroxaban, 15% apixaban). Measured dabigatran concentrations ranged from 1.8 to

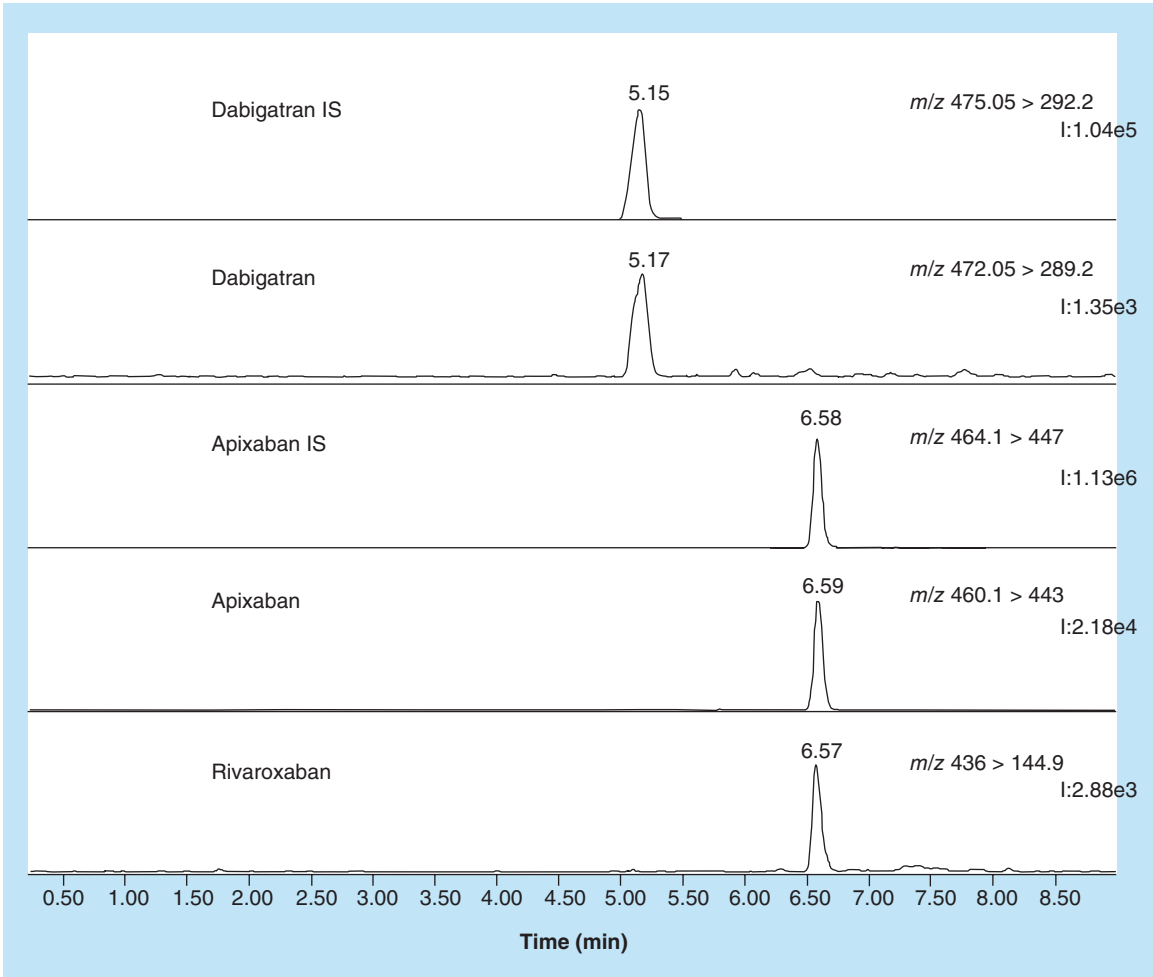


Figure 2. Extracted ion current chromatograms. Positive multiple reaction monitoring for the single-run detection of dabigatran, rivaroxaban, apixaban and their isotopically labelled standards (13C, 2H3-apixaban, 2H6-dabigatran) in a quality control at the limit of quantitation.
IS: Internal standard.

Table 1. Performance of the LC–MS/MS method for the simultaneous determination of dabigatran, rivaroxaban, apixaban concentrations in human plasma.

Parameters	Dabigatran	Apixaban	Rivaroxaban
Within day			
LOQ (1 ng/ml):			
– Mean ± SD	0.94 ± 0.1	1.05 ± 0.0	1.11 ± 0.2
– CV (%)	9.9	3.2	14.2
– Accuracy (%)	93.6	104.6	110.8
Low conc (3 ng/ml):			
– Mean ± SD	2.6 ± 0.2	3.1 ± 0.2	2.8 ± 0.4
– CV (%)	7.1	7.5	14.2
– Accuracy (%)	88.8	104.5	94.4
Medium conc (30 ng/ml):			
– Mean ± SD	26.0± 1.9	30.8 ± 1.6	31.5 ± 1.9
– CV (%)	7.1	5.3	5.9
– Accuracy (%)	86.5	102.6	105.0
High conc (300 ng/ml):			
– Mean ± SD	334.6 ± 7.2	334.6 ± 14.9	339.0 ± 4.2
– CV (%)	2.2	4.3	4.3
– Accuracy (%)	89.7	89.7	87
Between day			
LOQ (1 ng/ml):			
– Mean ± SD	1.07 ± 0.1	1.02 ± 0.1	1.04 ± 0.2
– CV (%)	9.0	5.1	14.6
– Accuracy (%)	106.8	102.0	104.0
Low conc (3 ng/ml):			
– Mean ± SD	2.8 ± 0.1	3.1 ± 0.1	2.9 ± 0.4
– CV (%)	4.4	3.9	13.7
– Accuracy (%)	92.3	102.3	96.4
Medium conc (30 ng/ml):			
– Mean ± SD	29.1± 2.9	31.8 ± 2.1	30.8 ± 2.4
– CV (%)	9.8	6.5	7.9
– Accuracy (%)	96.9	106.1	102.7
High conc (300 ng/ml):			
– Mean ± SD	302.8 ± 14.9	339.2 ± 11.3	327.2 ± 22.0
– CV (%)	4.9	3.3	6.7
– Accuracy (%)	101.0	86.9	90.9

CV: Coefficient of variation; LOQ: Limit of quantitation; SD: Standard deviation.

318.5 ng/ml, with an interindividual variability of 58%, while rivaroxaban and apixaban ranged from 1.0 to 307.4 ng/ml (coefficient of variation 100%) and from 77.4 to 302.7 ng/ml (coefficient of variation 53%), respectively.

Sample reanalysis was performed on 14 samples of apixaban, 25 for dabigatran and 24 for rivaroxaban. The percentage difference in results between the concentration of repeat and the original sample was within the threshold limit of 20% specifically recommended by the EMA guidelines in 60 of 63 samples (in a sample of dabigatran the difference was 25% and in two samples of rivaroxaban differences were of 24 and 23%). Considering the more conservative approach by Rocci *et al.* [13],

the mean ratio obtained for apixaban was 0.98 with ratio limits of 0.96–1.01 and limits of agreement 0.94–1.03 that contain all the determination. For dabigatran the mean ratio was 1.06 with ratio limits of 1.01–1.11 and limits of agreement 0.94–1.19, while for rivaroxaban the mean ratio was 1.05 with ratio limits of 1.00–1.10 and limits of agreement 0.93–1.18. For these two molecules, one determination was out of the limits of agreements.

Discussion

Testing of coagulation parameters is a common practice in clinical medicine during the administration of VKAs, whereas NOACs do not usually require routine coagulation monitoring. However, important differences in the interindividual PK and PD profiles of NOACs have been recently underlined [3]. It has also been suggested that a better understanding of the effect of clinical variables such as age, renal function, dosing interval, genotyping (inherited) differences in drug metabolism/transport and co-medications may allow a better prediction of the risk for sub- and over-therapeutic anticoagulation response and hence help individualizing NOACs selection and dosing. This task, however, cannot be addressed by traditional PD coagulation tests, such as PT and APTT [2,7,14]. As alternative strategies, functional coagulation tests that estimate plasma drug concentrations by assessing a PD response have been developed. These tests are based on the assessment of diluted thrombin time (Hemoclot Thrombin Inhibitors® [HTI], Hyphen-BioMed, Neuville-sur-Oise, France), ecarin clotting assay (ECA-T HaemoSys® [ECA]; Diagnostica Stago, Asnieres, France), or chromogenic anti Xa assays (Biophen Direct Factor Xa inhibitor® [DiXal]; Hyphen-BioMed) [15,16]. Although HTI, ECA and DiXal test are already used in clinical practice to monitor exposure of patients to NOACs, comparative studies have recently documented essential interindividual variability and poor correlation with drug concentrations measured using LC–MS/MS taken as the golden standard reference method; interestingly these biases resulted particularly evident in the low range of NOACs concentrations [15–18].

For the above-mentioned limitations of available chromogenic or clotting-based assay, here we developed and validated an LC–MS/MS method for the simultaneous quantification of dabigatran, rivaroxaban and apixaban concentrations in human plasma designed to be easily applicable in a clinical setting. Although these drugs are not simultaneously in clinical practice, the

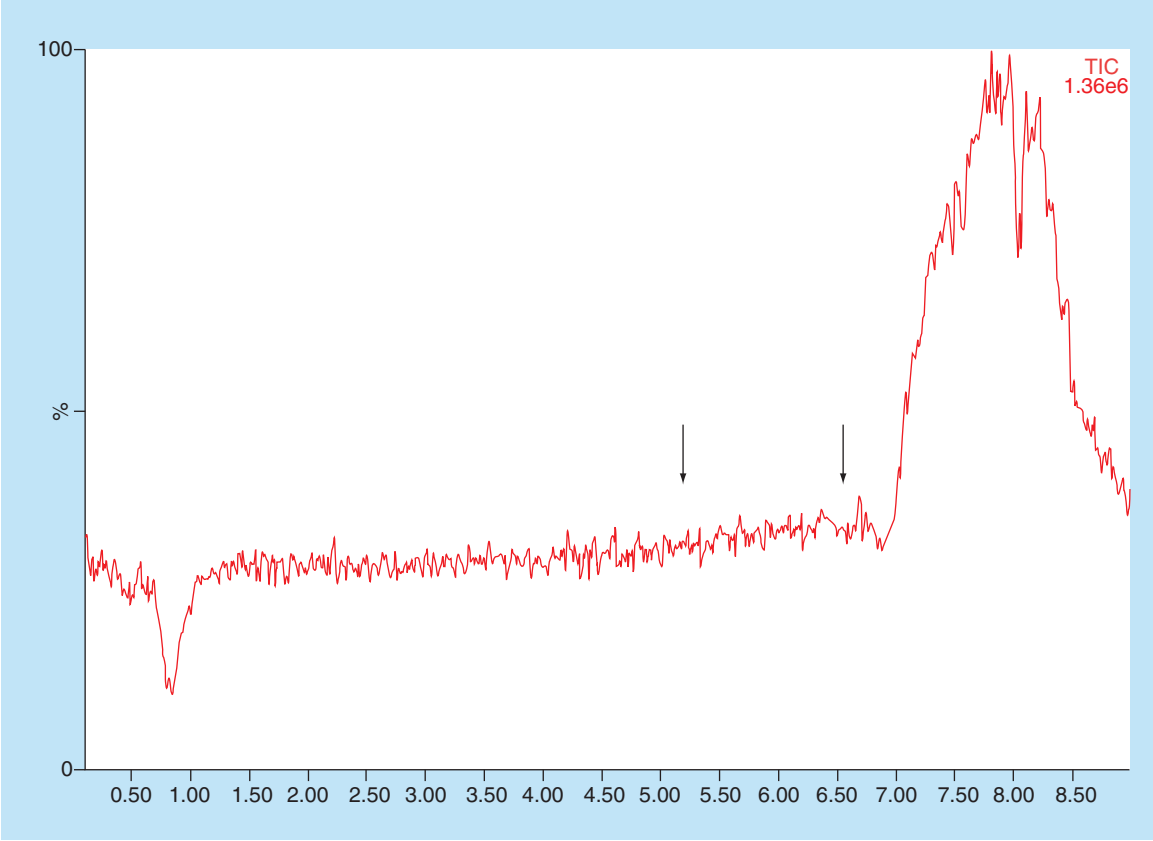


Figure 3. Monitoring of ion suppression by postcolumn infusion method. Total ion current trace of novel oral anticoagulants steadily infused into the LC eluent of a blank plasma extract. Vertical arrows represent the analyte retention time.

availability of a single methodology to assess different NOACs would simplify and facilitate the laboratory testing of these drugs rendering it accessible also to small-size laboratories than cannot afford to maintain multiple analytic procedures in parallel.

The method was precise, accurate, sensitive, specific and robust, fully accomplishing the recently updated guidelines for bioanalytical method validation [11]. Worthy of mention, we have successfully applied the method for routine monitoring of patients on anticoagulation therapy and confirmed that such patients, treated with fixed drug doses, experienced wide variability in the NOACs exposure.

To the best of our knowledge, only three papers have previously dealt with the development of methods able to concomitantly measure different NOACs, with comparable overall performance with our method [8–10]. The main novelty of our work compared with available literature is the conduct of incurred sample reanalysis which is a mandatory requisite in the recently released EMA guidelines on bioanalytical method validation [11]. This is an essential analysis to unmask reproducibility problems in the processing of *in vivo* samples. Using this approach we have provided evidence that of our method fully complies with the requirements dictated by both the EMA guidelines [11] and the more conservative and stringent criteria proposed by Rocci *et al.* [13]. Moreover a potential advantage of our method as compared with the one recently proposed by Gous *et al.* [9] is the use of an instrumentation that nowadays is more likely to be found in clinical laboratories.

As potential limitations of the present study we acknowledged the lack of an isotope-labeled IS for rivaroxaban and the use of solid-phase extraction which is a time-consuming sample preparation step that, however, provides a better clean up of plasma samples and reduced matrix effect allowing better method performance. An additional potential limitation of our method relies on its unproven application for the assessment of dabigatran acylglucuronides, as recently described by Saffian *et al.* [19]. Indeed, as these metabolites are labile, a potential overestimation of dabigatran due to the hydrolysis of the acyl-glucuronide metabolites cannot be definitively ruled out [19]. However, we are confident that this bias could be limited if samples are immediately centrifuged at 4°C and stored at -20°C until analysis.

In conclusion, an accurate and sensitive LC–MS/MS assay for quantification of NOACs in human plasma was developed and validated fully accomplishing the EMA guidelines. This method was shown to be easily applicable for the routine management of patients on anticoagulation therapy in selected clinical scenarios.

Future perspective

We believed that this method would be a specific and precise tool in clinical laboratory to assess patient drug concentration in order to improve the benefit–risk ratio for these drugs.

Many patients could benefit from a precise measurement of plasma concentrations: elderly, patients with hepatic or renal impairment, patient with extreme bodyweight, patient at risk of drug–drug interactions. The availability of method that simultaneously quantify all NOACs in an accurate manner even at low concentration is mandatory for a hospital laboratory. In future, a powdered clinical trial to confirm the benefit of a dose adjusted respect to a standard dose regimen in these and other categories of patients would be necessary.

In addition, a range of concentrations associated with optimal response and low toxic effects will be defined.

Financial & competing interests disclosure

The study was supported by a fellowship grant from ViiV Healthcare (to D Cattaneo) and the Ministry of Health Ricerca

Executive summary

Background

- To develop an LC–MS/MS method for quantification of dabigatran, rivaroxaban and apixaban in human plasma samples from patients given NOAC as part of their maintenance anticoagulation therapy.

Experimental

- The proposed method was validated in agreement with the recently updated EMA guidelines for Bioanalytical Methods Validation.

Results & discussion

- The proposed method fully accomplished the criteria dictated by EMA guidelines.
- Coefficient of variation of interpatient concentrations confirm the large variability in plasma concentrations of the fixed dose therapy.
- The method has been satisfactorily applied to plasma samples of patients in therapy with these drugs.

Conclusion

- The developed method is accurate and precise and free from matrix effects and is considered suitable to quantity novel oral anticoagulants in human plasma samples.

Corrente 2015 grant (to E Clementi). The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.

Ethical conduct of research

The authors state that they have obtained appropriate institutional review board approval or have followed the principles outlined in the Declaration of Helsinki for all human or animal experimental investigations. In addition, for investigations involving human subjects, informed consent has been obtained from the participants involved.

References

Papers of special note have been highlighted as:
• of interest; •• of considerable interest

1 Jung M, Rice L. Treatment of heparin-induced thrombocytopenia: what are the contributions of clinical trials? *Clin. Invest.* 3, 359–367 (2013).

2 Harder S, Graff J. Novel oral anticoagulants: clinical pharmacology, indications and practical considerations. *Eur. J. Clin. Pharmacol.* 69, 1617–1633 (2013).

3 Gong IY, Kim RB. Importance of pharmacokinetic profile and variability as determinants of dose and response to dabigatran, rivaroxaban, and apixaban. *Can. J. Cardiol.* 29(Suppl. 7), S24–S33 (2013).

4 Hellwig T, Gulseth M. Pharmacokinetic and pharmacodynamic drug interactionswith new oral wanticoagulants: what do they mean for patients with atrial fibrillation? *Ann. Pharmacother.* 47, 1478–1487 (2013).

5 Ten Cate Hugo. New oral anticoagulants: discussion on monitoring and adherence should start now! *Thromb. J.* 11(1), 8 (2013).

6 Cohen D. Concerns over data in key dabigatran trial. *BMJ* 349, g4747 (2014).

7 Eriksson BI, Quinlan DJ, Weitz JI. Comparative pharmacodynamics and pharmacokinetics of oral direct thrombin and factor Xa inhibitors in development. *Clin Pharmacokinet.* 48, 1–22 (2009).

8 Schmitz EM, Boonen K, van den Heuvel DJ *et al.* Determination of dabigatran, rivaroxaban and apixaban by ultra-performance liquid chromatography–tandem mass spectrometry (UPLC-MS/MS) and coagulation assays for therapy monitoring of novel direct oral anticoagulants. *J. Thromb. Haemost.* 12, 1636–1646 (2014).

9 Gous T, Couchman L, Patel JP, Paradzai C, Arya R, Flanagan RJ. Measurement of the direct oral anticoagulants apixaban, dabigatran, edoxaban, and rivaroxaban in human plasma using turbulent flow liquid chromatography with high-resolution mass spectrometry. *Ther. Drug Monit.* 36, 597–605 (2014).

10 Harenberg J, Du S, Wehling M *et al.* Measurement of dabigatran, rivaroxaban and apixaban in samples of plasma, serum and urine, under real life conditions. An international study. *Clin Chem Lab Med.* doi:10.1515/cclm-2015-038. (2015) (Epub ahead of print).

11 European Medicines Agency. Guideline on bioanalytical method validation, 2011. www.ema.europa.eu/docs/en_GB/document_library

•• **Illustrates the recently updated EMA guidelines for bioanalytical method validation.**

12 Taylor P. Matrix effects: the Achilles heel of quantitative high-performance liquid chromatography-electrospray–tandem mass spectrometry. *Clin. Biochem.* 38, 328–334 (2005).

13 Rocci ML Jr, Devanarayan V, Haughey DB, Jardieu P. Confirmatory reanalysis of incurred bioanalytical samples. *AAPS J.* 9, E336–E343 (2007).

• **Provides an alternative approach for the reanalysis of incurred biological samples.**

14 Lindhoff-Last E, Samama MM, Ortel TL, Weitz JI, Spiro TE. Assays for measuring rivaroxaban: their suitability and limitations. *Ther. Drug Monit.* 32, 673–679 (2010).

15 Douxfils J, Tamigniau A, Chatelain B *et al.* Comparison of calibrated chromogenic anti-Xa assay and PT tests with LC-MS/MS for the therapeutic monitoring of patients treated with rivaroxaban. *Thromb. Haemost.* 110, 723–731 (2013).

•• **Provides important limitations of the actual PD-based methods for the indirect assessment of rivaroxaban plasma concentrations.**

16 Antovic JP, Skeppholm M, Eintrei J *et al.* Evaluation of coagulation assays versus LC–MS/MS for determinations of dabigatran concentrations in plasma. *Eur. J. Clin. Pharmacol.* 69, 1875–1881 (2013).

• **Provides important limitations of the actual PD-based methods for the indirect assessment of dabigatran plasma concentrations.**

17 Gosselin RC, Adcock D, Hawes EM, Francart SJ, Grant RP, Moll S. Evaluating the use of commercial drug-specific calibrators for determining PT and APTT reagent sensitivity to dabigatran and rivaroxaban. *Thromb. Haemost.* 113, 77–84 (2015).

18 Gosselin R, Hawes E, Moll S, Adcock D. Performance of various laboratory assays in the measurement of dabigatran in patients receiving therapeutic doses: a prospective study based on peak and trough plasma levels. *Am. J. Clin. Pathol.* 141, 262–267 (2014).

19 Saffian SM, Zhang M, Leong Chin PK, Jensen BP. Quantification of dabigatran and indirect quantification of dabigatran acylglucuronides in human plasma by LC–MS/MS. *Bioanalysis* 7, 957–966 (2015).

Contact us

Bioanalysis Zone

Sponsorship and Advertising Sales
Sarah Meyers
✉ s.meyers@future-science.com

Subscriptions and Institutional Trials
Dominik March
✉ d.march@future-science.com

Editorial
Hannah Makin
✉ h.makin@future-science.com

Sankeetha Nadarajah
✉ s.nadarajah@future-science.com

www.bioanalysis-zone.com
☎ +44 (0)20 8371 6090
Bioanalysis Zone
@BioanalysisZone
Bioanalysis Zone
Bioanalysis Zone

Future Science Group
Unitec House
2 Albert Place
London, N3 1QB
United Kingdom

fsg CONNECT