For reprint orders, please contact: reprints@future-science.com

Nuclear magnetic resonance structure-based drug design

Felix Torres¹ & Julien Orts^{*,1}

¹Laboratory of Physical Chemistry, ETH Zürich, Vladimir-Prelog-Weg 2, 8093 Zürich, Switzerland

*Author for correspondence: Tel.: +41 0 44 632 28 64; julien.orts@phys.chem.ethz.ch

“NMR has been shown to be relevant in all four stages of drug discovery, but we will focus on NMR methods for structure-based lead development.”

First draft submitted: 4 May 2018; Accepted for publication: 8 August 2018; Published online: 16 October 2018

Spectroscopy in drug design

Drug discovery is the first step in pharmaceutical development and involves four stages: target identification, high-throughput screening of hit compounds, hit validation and lead optimization. Although lead compounds can emerge from thousands of hits during high-throughput screening, chemical modifications are still necessary to improve their affinity and/or specificity. Structural insights into the ligand-target complex is a necessary step of the structure-activity relationship that leads the medicinal chemist to rational modifications. An x-ray crystallography plays an important role in providing the structure of complexes, but many are not readily crystallizable especially at the early stage when the ligands are still weak binders ($K_D \sim \text{mM}-\mu\text{M}$). However, progresses in the field may alleviate the need for well-diffracting crystals. Recently, the very challenging system *Trypanosoma brucei* cysteine protease cathepsin B in interaction with an inhibitor was solved by x-ray free electron laser [1]. Moreover, *in vivo* crystallization lead to structures at physiological conditions. Simultaneously, the outstanding progresses in cryo-EM allowed the description of the hitherto unknown structure-activity relationship of mefloquine bound to the *Plasmodium falciparum* 80s ribosome [2]. While structures below 100 kDa are not reachable yet, advances in hardware technology should overcome these limitations in the upcoming decade. Nuclear magnetic resonance (NMR) is a well-established technique for the study of target-ligand structures. It operates under near physiological conditions and provides information on both structure and dynamics. However, the classic protocol for complex structure refinement is time consuming and is limited regarding the size of the target (~ 60 kDa for a protein). Consequently, methods using sparse NMR data are actively developed as they provide fast and specific structural restraints for lead optimization [3]. NMR has been shown to be relevant in all four stages of drug discovery, but we will focus on NMR methods for structure-based lead development.

The versatility of nuclear magnetic resonance

NMR benefits from a variety of distance restraints that originate from different physical interactions. The nuclear overhauser effect (NOE) provides cross-relaxation rates (σ) that are proportional to r^{-6} , where r is the distance between two nuclei. Hence, NOE is rather suitable for relatively short distance restraints (1 to 5–6 Å) and thus perfectly adapted for ligand-binding pocket restraints [4,5]. Saturation transfer difference experiment also originates from the Overhauser effect. The saturated receptor protons transfer their magnetization to the ligand's protons in the binding site and modulate their NMR resonance signal intensities. The most straightforward way to obtain proximity restraints is through chemical shift perturbations (CSP) that originate from changes in the surrounding magnetic shielding of a spin. Such perturbations are often caused by ligand-binding or conformational changes [6,7]. In most cases, ^{15}N -HSQC spectra are recorded for two samples of protein (e.g., with and without ligand). However, ^{13}C -HSQC are more specific to ligand binding since H_N CSP are often related to allosteric effects or conformational changes [8]. Yet another way of obtaining restraints is using paramagnetic effect, which arises from unpaired electrons, that have a three orders of magnitude higher magnetic moment leading to strong interactions with the nuclear spin visible by NMR. It generates two types of restraints: PseudoContact Shifts (PCS) and Paramagnetic Relaxation Enhancement (PRE). PseudoContact shifts is observed for unpaired electrons in anisotropic orbitals such as for

newlands
press

lanthanides in oxidation state III (Ln^{3+}). It depends on the distance (r^{-3}) and the angle between the inter-nuclei axis and the anisotropic tensors of the paramagnetic nucleus. Several distance restraints relative to the lanthanide can provide the 3D structure of the complex, knowing the position of the paramagnetic lanthanide on the protein [9]. PRE causes signal broadening and intensity lowering in an isotropic dependence, and thus is only function of distance (r^{-6}). While this can be used for structural purpose, the anisotropic properties of PCS make these restraints more desirable. Hence, Ln^{3+} type and position regarding the binding pocket are chosen in such a way that the PCS that are suitable for longer range are not mixed with PRE restraints. Others restraints are observable in NMR, such as J-coupling or cross-correlated relaxation, but were not used in the context of sparse restraints and will generally not provide direct information about protein–ligand interfaces since they originated from through bound interactions.

Ambiguity is an important concept for structural restraints [10]. An NMR restraint between a pair of defined atoms is said to be ‘unambiguous’ when each atom belongs to a group assigned to a single resonance. Inversely, ‘ambiguous’ restraints are defined between two atoms from not assigned groups, thus allowing several combinations. The CSP-based structure calculations use ambiguous restraints [6,7]. If this ambiguity seems to be acceptable for protein–protein interactions [11], for protein–ligand complexes it is difficult to obtain a unique orientation due to the low resolution of the CSP restraints. Remarkably PCS restraints are not ambiguous since they originate from an identified nucleus, the Ln^{3+} and provide accurate restraints similar to the NOE [12]. While unambiguous restraints require the time-consuming assignment step, semi-ambiguous restraints seem to be a good compromise. In this case, NOE restraints are between an assigned ligand resonance and an unassigned, but unique, protein resonance. This protocol is used in NMR^2 or partly in NOE-matching, thus avoiding signal assignment for protein and limiting the level of ambiguity [4,5]. Furthermore, quantitative analysis of NOE restraints is well established for distance calculations, as well as for PCS and PRE.

Overcoming size limitations

The NMR protein observables are limited by signal broadening and overlapping for protein larger than 60 kDa. However, several methods that rely on ligand’s signals, referred as ligand-observed techniques, do not suffer from this size limitation. For example, INPHARMA determines the relative orientation of a weak binder to a protein from protein mediated inter-ligand NOE [13]. SOS-NMR uses saturation transfer difference where signal intensities of the ligand are increased when in proximity with specifically labeled protein methyl groups, while the receptor’s other protons are deuterated [14]. Complete relaxation and conformational exchange matrix interprets intra and intermolecular transferred-NOE arising from the complex and visible on the ligand population averaged resonances [15]. Furthermore, ligand observed techniques require a low amount of target (20–30 μM). It is worth noting that transferred experiments are only possible for weak binder in fast exchange with the target [16]. On the other hand, methods such as CSP, NMR^2 , NOE-matching or PCS that also observe the receptor signals are limited by its size. But progresses in protein-labeling techniques enable to reach higher molecular weights [17]. Specifically, background deuteration reduces signal broadening and methyl specific labeling enables ^{13}C -edited experiments, thus simplifying the NMR spectra [18].

However, for many proteins, the number of residues reaches several hundreds. Consequently, since a signal assignment procedure needs a complete set of 2D and 3D spectra, it requires days for acquisition and several weeks for analysis. While some ligand-observed methods such as INPHARMA, complete relaxation and conformational exchange matrix or SOS-NMR do not require protein residue resonance assignments, they rely on a pure *in silico* docking step prior to experimental scoring. On the other hand, methods relying on CSP such as j-surface or PCS need resonance assignment. Still, automated resonance assignment procedures might speed up this step. Two recent methods, NMR^2 and NOE-matching derive the complex structure with semi-ambiguous restraints between an assigned ligand and a non-assigned-protein. The difference between the two algorithms is that NMR^2 calculates each time the complex structure, while NOE-matching is a docking-scoring method [4,5]. Furthermore, these methods rely on NOE and thus provide quantitatively good distance restraints.

Moving beyond docking & scoring approaches

All but one of the techniques discussed previously refine the complex structure by docking and scoring approaches. Even if some docking software enable to run data driven docking such as HADDOCK, a force-field will bias the generated poses [19]. Even though it was shown that protein–protein interactions are well defined by restraint docking even with ambiguous restraints thanks to their large interaction surface, for small molecule the energy

penalty due to a wrong orientation is often not large enough to exclude these structures since they do not originate from high resolution restraints. Some other methods such as CLOUDS that builds the proton's isosurface based on unassigned NOE restraints have been developed [20]. However, they still lack application for structure determination of protein–ligand complexes. NMR² is currently the only technique, to the best of our knowledge, that uses semi-ambiguous distance restraints in conjunction with a simulated annealing structure calculation protocol [4]. The advantage of this approach is that it finds the solution structure of the complex driven by experimental data, minimizing the influence of *in silico* parameters and avoiding conformational traps.

The future of NMR structure-based drug design

NMR benefits from a wide toolbox for structure-based drug discovery. However, these tools are in large part not routinely used. The derivation of complex structures should be done as simple as possible and quantitative restraints such as NOE or PCS should be prioritized over semi-quantitative ones like CSP. Moreover, since ligand is most likely small molecules, low resolution restraints should be avoided as they will not provide sufficient structural information and will be dampened by the modeling. Semi-ambiguous restraints are promising as they conveniently do not require receptor assignment but retain discrimination capabilities. Ranking of docking poses is the most common way to derive protein–ligand structures but docking tools still suffer from force field inaccuracies or incorrect scoring functions. Therefore, docking-scoring methods should be avoided as much as possible. We think that effort should be put into developing structure calculation methods integrating current available structural data from the Protein Data Bank with quantitative NMR restraints such as PCS and semi-ambiguous NOE. NMR is a versatile technique that when synergistically combined with other approaches provides new structures of very high quality. We are therefore convinced that integrating available structural data, with quantitative NMR restraints and specific labeling schemes will open a new avenue for structure-based drug discovery.

Open access

This work is licensed under the Attribution-NonCommercial-NoDerivatives 4.0 Unported License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>

Financial & competing interests disclosure

This work was financially supported by the ETH Zürich. 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. Redecke L, Nass K, Deponate DP *et al.* Natively inhibited *Trypanosoma brucei* cathepsin B structure determined by using an X-ray laser. *Science* 339(6116), 227–230 (2013).
2. Wong W, Bai XC, Sleebs BE *et al.* Mefloquine targets the *Plasmodium falciparum* 80S ribosome to inhibit protein synthesis. *Nat. Microbiol.* 2, 17031 (2017).
3. Gossert AD, Jahnke W. NMR in drug discovery: a practical guide to identification and validation of ligands interacting with biological macromolecules. *Prog. Nucl. Magn. Reson. Spectrosc.* 97, 82–125 (2016).
4. Orts J, Walti MA, Marsh M *et al.* NMR-based determination of the 3D structure of the ligand-protein interaction site without protein resonance assignment. *J. Am. Chem. Soc.* 138(13), 4393–4400 (2016).
5. Constantine KL, Davis ME, Metzler WJ, Mueller L, Claus BL. Protein–ligand NOE matching: a high-throughput method for binding pose evaluation that does not require protein NMR resonance assignments. *J. Am. Chem. Soc.* 128(22), 7252–7263 (2006).
6. McCoy MA, Wyss DF. Alignment of weakly interacting molecules to protein surfaces using simulations of chemical shift perturbations. *J. Biomol. NMR* 18(3), 189–198 (2000).
7. Cioffi M, Hunter CA, Packer MJ, Pandya MJ, Williamson MP. Use of quantitative (1)H NMR chemical shift changes for ligand docking into barnase. *J. Biomol. NMR* 43(1), 11–19 (2009).
8. Williamson MP. Using chemical shift perturbation to characterise ligand binding. *Prog. Nucl. Magn. Reson. Spectrosc.* 73, 1–16 (2013).
9. Otting G. Prospects for lanthanides in structural biology by NMR. *J. Biomol. NMR* 42(1), 1–9 (2008).
10. Schieborr U, Vogtherr M, Elshorst B *et al.* How much NMR data is required to determine a protein-ligand complex structure? *Chembiochem* 6(10), 1891–1898 (2005).

11. Stratmann D, Boelens R, Bonvin AMJJ. Quantitative use of chemical shifts for the modeling of protein complexes. *Proteins* 79(9), 2662–2670 (2011).
12. Pintacuda G, Park AY, Keniry MA, Dixon NE, Otting G. Lanthanide labeling offers fast NMR approach to 3D structure determinations of protein-protein complexes. *J. Am. Chem. Soc.* 128(11), 3696–3702 (2006).
13. Sanchez-Pedregal VM, Reese M, Meiler J, Blommers MJJ, Griesinger C, Carlomagno T. The INPHARMA method: protein-mediated interligand NOEs for pharmacophore mapping. *Angew. Chem. Int. Ed. Engl.* 44(27), 4172–4175 (2005).
14. Hajduk PJ, Mack JC, Olejniczak ET, Park C, Dandliker PJ, Beutel BA. SOS-NMR: a saturation transfer NMR-based method for determining the structures of protein-ligand complexes. *J. Am. Chem. Soc.* 126(8), 2390–2398 (2004).
15. Moseley HN, Curto EV, Krishna NR. Complete relaxation and conformational exchange matrix (CORCEMA) analysis of NOESY spectra of interacting systems; two-dimensional transferred NOESY. *J. Magn. Reson B* 108(3), 243–261 (1995).
16. Ni F. Recent developments in transferred NOE methods. *Prog. Nucl. Mag. Res. Sp.* 26, 517–606 (1994).
17. Tugarinov V, Kanelis V, Kay LE. Isotope labeling strategies for the study of high-molecular-weight proteins by solution NMR spectroscopy. *Nat. Protoc.* 1(2), 749–754 (2006).
18. Breeze AL. Isotope-filtered NMR methods for the study of biomolecular structure and interactions. *Prog. Nucl. Mag. Res. Sp.* 36(4), 323–372 (2000).
19. Dominguez C, Boelens R, Bonvin AM. HADDOCK: a protein–protein docking approach based on biochemical or biophysical information. *J. Am. Chem. Soc.* 125(7), 1731–1737 (2003).
20. Grishaev A, Llinas M. CLOUDS, a protocol for deriving a molecular proton density via NMR (vol 99, pg 6707, 2002). *Proc. Natl Acad. Sci. USA* 99(16), 10941–10941 (2002).