Proteomics: evolution of the technology

Scott D. Patterson

BioTechniques 35:440-444 (September 2003)

INTRODUCTION

Technologies enable the advancement of science, which is a process that has been repeated many times throughout the history of science. One of these recent enabling technologies is proteomics. It is an “overnight” sensation that was more than two decades in the making, being developed in a similar time frame as this publication, *BioTechniques*. Proteomics is the term applied to parallel protein-based analyses. Originally used to describe two-dimensional gel electrophoresis (2-DE) and mass spectrometric analyses of proteins of interest, the term has been applied to any parallel protein analysis method—and even some that really aren’t. This article focuses on parallel protein analyses that result in the identification of components within a system (profiling), their posttranslational modifications and/or quantitative analyses of the amounts of the proteins, and modifications or activities, rather than other fields that come under the banner of proteomics, such as structural biology methods that have been reviewed elsewhere (1). We will address how these technologies have impacted science to date and how they might have an impact in the future.

WHERE HAVE WE COME FROM?

To get to where we are now, a quick glimpse at the past with a couple of limited examples is warranted. Note that this perspective is not intended as a comprehensive review, so only a few pertinent references will be cited per section. There are several recent review articles in the field that provide more in-depth coverage (2,3).

Reports of 2-DE, as we know it today, were published in 1975 by three groups and provided the first glimpse at the complexity of the protein content of cells and, as such, can be thought of as marking the true beginning of proteomics (4–6), almost 20 years before the term “proteomics” was coined. It is still the highest resolution protein separation method available and provided the first means of interrogating the levels of hundreds of proteins and their isoforms. However, in 1975, the ability to identify the proteins that displayed an interesting regulation proved difficult. Determining the relative intensity of the hundreds of spots between numerous gels was also problematic (as was the reproducibility of the method).

Automated protein sequencers that could analyze the amount of protein in 2-DE spots had not been developed at that time (that occurred during the following decade), algorithms for image analysis were at an early stage, and computers that could handle the large amounts of data that multiple high-resolution images contained did not yet exist. Several problems had to be solved before this technology could provide answers to detailed biological questions. However, even in its nascent form, discoveries were made, such as the oscillating levels of proliferating cell nuclear antigens (PCNAs) during the cell cycle observed from cell biology experiments in the 1980s (7,8). Protein-DNA interaction studies also revealed the fos-jun interaction (9). However, the full impact of the technology was still limited because it was mostly conducted in the laboratories of a relatively few practitioners of the art. For that to change, a number of advances had to be made, including some indirectly associated with the field.

The advances that enabled proteomics can be thought of as 4-fold (Figure 1). (i) Image analysis methods needed to improve, as did the reproducibility of the 2-DE method, both of which continue to this day. (ii) A method for the sensitive identification of gel-separated proteins was required that advances in mass spectrometry (MS) provided in the late 1980s. (iii) The content of the sequence databases exploded initially with the large-scale expressed sequence tag (EST) sequencing efforts in both the public and private domains and subsequently in genomic sequencing efforts. (iv) Computational methods for the correlation of the results of mass spectrometric analyses with the content of sequence databases were developed. These advances all occurred in the early 1990s, and in 1994, at the first 2-DE meeting in Siena, Italy, the term “proteome” was coined, short for the “protein complement of the genome” (10). The time was right to spark renewed interest in this field, which had been so promising for so many years, but had yet to enter the mainstream of biological analyses. During the 1990s, the realization of the genomics revolution was taking place, and the potential for true global-scale analyses was becoming a reality with methods such as yeast two-hybrid analysis coming to the fore, and the improvement of EST and genomic sequencing efforts (11).

Before the introduction of these technologies, reductionist approaches were often employed in the pursuit of the individual proteins or protein families that contributed to an activity/pathway of interest in hypothesis-driven research

projects. Reductionist approaches take a function-to-sequence approach, whereas discovery science approaches go from sequence to function, which is a more difficult task, but one that is hoped to enable a more thorough understanding of the context of the protein. The ability to view a large number of proteins in a given system (as well as transcripts and metabolites) lends credence to the idea that one can begin to undertake discovery science projects aimed at assaying the function and control of biological systems. But for this to work, one has to be able to analyze proteins over an enormous range of concentrations and to understand their interactions. However, these are only some of the issues that are currently being addressed.

WHAT ARE THE CURRENT ISSUES?

The dynamic range of protein concentrations in cells is thought to be 10^5 – 10^6 ; in biological fluids such as serum, it is even worse, with potentially greater than a 10^{10} dynamic range (e.g., from interleukin-6 at 2 pg/mL to albumin at 50 mg/mL) (3). There are no parallel protein analysis methods that can interrogate such a large dynamic range. Therefore, at this point in time, one has to accept that only a portion of the total potential protein complement of a system under analysis is visible to the analyst. To compound the problem, the number of proteins in any biological system is undefined; unlike the genome, which has a defined end point, the size and complexity of the proteome are unknown. The number of potential proteins, including their isoforms, structural or functional (and often transient), is no doubt orders of magnitude greater than the number of genes. Extreme examples of the number of isoforms of individual proteins can be enormous, such as the neurexins, in which more than 1000 isoforms have been defined from analyses of alternative splicing (12). The number of isoforms of glycoproteins can also be very large due the nature of the modifications [e.g., over 400 different glycoforms of the hamster scrapie prion protein PrP 27-30 were estimated from the analysis of released glycans from the two sites of *N*-linked glycosylation (13)]. Thus, the complexity of the proteome is exacerbated by both genetic events (e.g., alternative splicing) and dynamic and static posttranslational modifications. Because these modifications are the results of other genes, a protein may be considered the phenotype of many genes (14). Furthermore, the ratio of these modifications is also quite often substoichiometric but functional (i.e., transient phosphorylation of only 10% of a given protein within a cell may have a dramatic effect on its biological function).

In addition to the dynamic range

and variable modification issues, the proteome is dynamic with respect to the turnover of the proteins within the system. Protein half-lives are under close regulatory control, from seconds to days. Many of the parallel protein-based analysis systems are currently unsuited to the analysis of protein turnover. 2-DE is the exception, but only when pulse-chase labeling with radioactive precursors is used because it allows one to follow the synthesis and degradation of a broad range of proteins. This is a current limitation of the MS-based methods that will be described here.

BRINGING TECHNOLOGIES TOGETHER CONTINUES TO ADVANCE THE FIELD

As stated earlier, the archetypal proteomics method is 2-DE combined with image analysis and the mass spectrometric identification and characterization of digested proteins; essentially the analysis of intact proteins that are not intentionally modified. For example, 2-DE/MS has been successfully applied to the analysis of serum protein homeostasis (15) and the protein substrates of the chaperone GroEL (16). The analysis of intact proteins using liquid chromatography (LC)-MS with high-resolution mass spectrometers is possible, as evident by the description of so-called top-down approaches to protein analysis (17). However, peptides are better suited to MS and can be used as surrogates for the measurement of the proteins from which they were derived. This was first exemplified with the combination of LC coupled to MS for the analysis of peptides bound to the major histocompatibility complex (MHC) (18). This laid the foundation for the analysis of digests of protein mixtures, in which the analysis of the peptide was a surrogate for the analysis of the protein.

Figure 1. The birth of proteomics. Advances in multiple disciplines enabled the field of parallel protein-based analyses, proteomics, to begin to realize its potential. The integration of the results of mass spectrometric analysis with the output of nucleotide sequencing projects [e.g., expressed sequence tag (EST) and genomic sequencing] allowed the bottleneck in protein identification to be removed. Complex mixture analysis using liquid chromatography-mass spectrometry (LC-MS)/MS-based methods for profiling and quantitation were developed at around the same time. 2-DE, two-dimensional gel electrophoresis.

Initially, the adaptations of this approach to proteomics studies were aimed at profiling as many components as possible within complex mixtures (19,20), but then the same concept was employed for quantitative studies of two sets of samples (21) through the application of isotope dilution theory (22). Because of the high complexity of the samples and increased complexity generated through proteolytic digestion, strategies to reduce the complexity were being investigated for these “shotgun proteomics” methods [e.g., capturing only cysteinyl peptides (19,23) or posttranslationally modified peptides, such as phosphopeptides (24) or glycopeptides (25,26)]. The goal of such analyses was to isolate one idiotypic peptide per protein in the mixture and/or capture proteins (peptides) with specific characteristics. Thus, as shown in Figure 2, the analysis of proteins can be accomplished at either the level of the protein or, as a surrogate measure, the peptide (following proteolytic digestion, usually with trypsin).

The stable isotope dilution theory mentioned earlier was combined with a complexity reduction method to enable ratios of relative abundance of cysteinyl peptides from paired samples to be interrogated using a reagent referred to as an isotope-coded affinity tag (ICATTM) (21). This overcomes the problem associated with MS being inherently a nonquantitative technique (different peptides at the same concentration do not necessarily ionize to the same degree; that is, give the same signal response). However, when the chemical composi-

tion of the peptides is identical (and co-elute by LC) and differ only by the mass of the isotopes, quantitative ratios of the two molecular species can be measured. The initial embodiment of ICAT consisted of an affinity tag linked to a sulfhydryl reactive moiety via a linker in which the mass difference was encoded (one had eight hydrogens; the other had eight deuteriums). The mass difference and ratios of the two species are then measured in a mass spectrometer. As an extension of the isotope dilution approach described earlier, in which specific groups are introduced into a protein/peptide, one can also draw on methods used in metabolite analysis, in which heavy isotope versions of the molecules are synthesized and used as standards to allow for absolute quantification. The approach has recently been reintroduced for phosphopeptide (and peptide) quantitation under the term AQUA (27). This versatile technique is beginning to be exploited by cell biologists, providing an important means of determining stoichiometry of sites of phosphorylation (28). For the promise of proteomics to be realized, the tools need to be in the hands of scientists in other disciplines and in a format that they are comfortable using. A proteome-wide application of this concept has been proposed that synthesizes representative peptides for every protein and isoform as a means of allowing greater access of this technology to the biological community (29).

Given the limitations of dynamic range in MS of 2–3 orders of magnitude and 2-DE providing up to 5 orders (but

Figure 2. Proteomic analyses at the level of protein or peptide? Technologies for parallel protein analysis can be conducted at either the level of the protein or, as a surrogate, the level of the peptide. The hierarchical nature of protein analysis is shown from the organism to the protein. Proteins or digested proteins can be analyzed from any of these levels, but often greater information is obtained if the analysis is targeted to either subcellular organelles or proteins in specific complexes or specific classes (functional or structural; for example, enzymes or phosphoproteins) using either specific capture methodologies or enrichment strategies. 2-DE, two-dimensional gel electrophoresis; LC-MS, liquid chromatography-mass spectrometry; MS, mass spectrometry; CE, capillary electrophoresis; 1-DE/MS, one-dimensional electrophoresis/mass spectrometry.

only with radiolabeling; 3 orders otherwise), it is no surprise that efforts have gone into the fractionation of proteins into specific spatial or functional classes to both reduce the complexity and to attempt to improve the sensitivity of the analysis. Often, enrichment strategies have employed biochemical methods developed in the past for subcellular fractionation or even chemical entities previously used to inhibit a class of enzymatic activities through covalent interaction (referred to as “activity-based” probes). Numerous proteomics studies have been undertaken to profile the constituents of various subcellular organelles (30,31) or proteins released from them following a specific physiological event (32). For example, the field of chemical proteomics is growing out of this adaptation of previously well-characterized chemical entities being used as affinity reagents, which form covalent complexes with proteins in a target class of interest. One can attach either fluorescent labels or affinity tags (in much the same way as ICAT) to those inhibitors, thereby reducing the complexity, but in a targeted manner

and providing functional (activity) information. In this case, it was the use of suicide inhibitors of serine hydrolases that allowed an activity profile for this target class to be revealed (33); in other cases, gel or capillary electrophoresis with fluorescent labels can be used to reveal the pattern of activity or lack of it in the presence specific inhibitors (Figure 2).

Another interesting application of MS to proteomics is the direct imaging of tissue sections using matrix-assisted laser desorption-mass spectrometry (MALDI-MS) (34). This method ionizes peptides as well as proteins and has shown specific signals that correlate with specific cell types. Although only a subset of the protein constituents of the cell types in the tissue section is analyzed, the maintenance of the spatial context is very important, and the success to date indicates that this is certainly a field to watch with interest.

HOW CAN PROTEOMICS ASSIST BIOLOGICAL DISCOVERY?

An ultimate aim of proteomics is to contribute to biological knowledge. This could be achieved through the integration of hypothesis-driven and discovery science-driven projects that bring together analyses of other molecular species (e.g., mRNA and metabolites) within a cell and organism together with proteomics analyses, thus enabling a systems biology approach. This is beginning to be achieved in defined systems in simple organisms such as yeast (35), but it is a more difficult task in mammalian systems. The most rapid impact that proteomics is likely to make in the near future is through the discovery and validation of biomarkers of disease and disease progression or surrogate markers of drug efficacy. For a biomarker to be useful, it should be obtained from readily available sources such as serum and urine. For serum, the problem is that more than 50% of the protein load consists of a single protein, albumin, and several other proteins that may or may not be useful biomarkers are also present at high concentrations. Methods to remove albumin have been developed, but one should consider that even with the removal of 99% of the protein, the remaining 1% will still be one of the most abundant species in the sample. This is a vexing problem that continues to be addressed. In addition, it is likely that patterns of changes in protein expression will more specifically define a disease than individual protein markers (36,37). As the results of these approaches are validated in larger patient sets, the question of how they will be translated into true diagnostic tests needs to be answered. It is not necessarily obvious that the research discovery tool employed in these analyses (MALDI-MS) is the most appropriate tool in its current format for the diagnostics laboratory. The development of protein chips, which can take many forms, may provide an answer. To date, the commercially available protein chips are essentially miniaturized versions of standard ELISAs [although some use unique amplification methods (38)]. However, advances in this field using different affinity reagents and surfaces may well provide tools for the measurement of the levels of multiple proteins from small sample quantities (39). Interestingly, it has also been shown that peptides derived from proteolytic digests can be used in antibody-based arrays

with a mass spectrometric readout (40).

Although global analysis is an aim of proteomics, this is not yet feasible for mammalian systems. As such, the studies that focus on specific classes of proteins of interest to the biology under study will no doubt have a significant impact. For example, a recent publication of a method to capture cell surface proteins will assist in the study of this important class of proteins (26). The use of activity-based methods (described earlier) will also prove of value in drug target screening and lead identification (41). The analysis of protein-protein interactions has been demonstrated on a large scale biochemically by two groups, but the overlap between their data sets was not large (42,43). This brings up an important point with respect to proteomic data analysis (as with other parallel methods): the validation of results is critical before substantive conclusions can be drawn.

Lastly, as almost all proteomics methods use MS at some stage, and mass spectrometers generate large quantities of data, software that can handle the reduction of these vast data sets into knowledge is required (44). Although some of the analysis is conceptually simple and, as previously described, the original software enabled the birth of proteomics, the interpretation still requires some skill and experience on the part of the analyst. Clearly, the software tools need to be improved, and this can best be accomplished through biochemists/biologists and computer scientists working in concert. This is required not only for the analysis of mass spectrometric data but also for the intelligent collation of data from all aspects of scientific programs (45).

CONCLUSION

In summary, proteomics and parallel protein analysis are powerful tools that can be utilized to provide insight into many biological processes. Ultimately, proteomics will be an invaluable aid to true systems biology approaches in which the flux of mRNA, protein, and metabolites are being examined in toto. However, they are only tools, and if they are not used in the appropriate context, they will not necessarily provide useful information. Too often new tools are thought to provide answers to many previously unsolvable questions, and often they are hyped to give that impression. In the field of proteomics, there are multiple approaches that can be employed [e.g., 2-DE/MS, LC-MS/MS, protein chips, etc], and they are all valid for different experimental scenarios. One has to carefully consider the biological question to determine which of the available tools is best suited to potentially provide the answer. Just as genomics provided the ultimate blueprint but not a clear path for all new drug targets (as some had claimed/hoped), proteomics will not provide all the new drug targets or biomarkers in simple experiments. However, both of these fields will continue to reveal more interesting biology as they are further examined and applied in different ways. Ultimately, the more scientists from different disciplines understand the utility of proteomics, the more the technology will be used in appropriate experimental settings, and the more the field will expand to play an important role in many aspects of biological science, both industrial and academic.

REFERENCES

1. Thornton, J. 2001. Structural genomics takes off. *Trends Biochem. Sci.* 26: 88-89.
2. Marte, B. (Ed.) 2003. Nature insight: proteomics. *Nature* 422 (Suppl.):191-237.
3. Patterson, S.D. and R.H. Aebersold. 2003. Proteomics: the first decade and beyond. *Nat. Genet.* 33:311-323.
4. Klose, J. 1975. Protein mapping by combined isoelectric focusing and electrophoresis of mouse tissues: a novel approach to testing for induced point mutations in mammals. *Humangenetik* 26:231-243.
5. O'Farrell, P.H. 1975. High resolution two-dimensional gel electrophoresis of proteins. *J. Biol. Chem.* 250:4007-4021.
6. Scheele, G.A. 1975. Two-dimensional gel analysis of soluble proteins. Characterization of guinea pig exocrine pancreatic proteins. *J. Biol. Chem.* 250: 5375-5385.
7. Celis, J.E. and R. Bravo. 1984. Synthesis of the nuclear protein cyclin in growing, senescent and morphologically transformed human skin fibroblasts. *FEBS Lett.* 165:21-25.
8. Mathews, M.B., R.M. Bernstein, B.R. Franza, Jr., and J.I. Garrels. 1984. Identity of the proliferating cell nuclear antigen and cyclin. *Nature* 309:374-376.
9. Rauscher, F.J., III, D.R. Cohen, T. Curran, T.J. Bos, P.K. Vogt, D. Bohmann, R. Tjian, and B.R. Franza, Jr. 1988. Fos-associated protein p39 is the product of the jun proto-oncogene. *Science* 240:1010-1016.
10. Wilkins, M.R., J.-C. Sanchez, A.A. Gooley, R.D. Appel, I. Humphrey-Smith, D.F. Hochstrasser, and K.L. Williams. 1995. Progress with proteome projects: why all proteins expressed by a genome should be identified and how to do it. *Biotech. Gen. Eng. Rev.* 13:19-50.
11. Tucker, C.L., J.F. Gera, and P. Uetz. 2001. Towards an understanding of complex protein networks. *Trends Cell Biol.* 11:102-106.
12. Ullrich, B., Y.A. Ushkaryov, and T.C. Sudhof. 1995. Cartography of neurexins: more than 1000 isoforms generated by alternative splicing and expressed in distinct subsets of neurons. *Neuron* 14:497-507.
13. Endo, T., D. Groth, S.B. Prusiner, and A. Kobata. 1989. Diversity of oligosaccharide structures linked to asparagines of the scrapie prion protein. *Biochemistry* 28:8380-8388.
14. Klose, J. 1999. Genotypes and phenotypes. *Electrophoresis* 20:643-652.
15. Lee, S.J., S. Evers, D. Roeder, A.F. Parlow, J. Risteli, L. Risteli, Y.C. Lee, T. Feiziet al. 2002. Mannose receptor-mediated regulation of serum glycoprotein homeostasis. *Science* 295:1898-1901.
16. Houry, W.A., D. Frishman, C. Eckerskorn, F. Lottspeich, and E.U. Hartl. 1999. Identification of *in vivo* substrates of the chaperonin GroEL. *Nature* 402:147-154.
17. Kelleher, N.L., H.Y. Lin, G.A. Valaskovic, D.J. Aaserud, E.K. Fridriksson, and F.W. McLafferty. 1999. Top down versus bottom up protein characterization by tandem high resolution mass spectrometry. *J. Am. Chem. Soc.* 121:806-812.
18. Appella, E., E.A. Padlan, and D.F. Hunt. 1995. Analysis of the structure of naturally processed peptides bound by class I and class II major histocompatibility complex molecules. *EXS* 73:105-119.
19. Patterson, S.D. 1998. Using MS fragment-ion data to identify proteins from large sequence databases, p. 127-135. In L.M. Savage (Ed.), *Proteomics, Integrating Protein-based Tools and Applications for Drug Discovery*. International Business Communications, Southborough, MA.
20. Yates, J.R., III, A.L. McCormack, D. Schieltz, E. Carmack, and A. Link. 1997. Direct analysis of protein mixtures by tandem mass spectrometry. *J. Prot. Chem.* 16:495-497.
21. Gygi, S.P., B. Rist, S.A. Gerber, F. Turecek, M.H. Gelb, and R. Aebersold. 1999. Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat. Biotechnol.* 17:994-999.
22. De Leenheer, A.P. and L.M. Thienpont. 1992. Application of isotope dilution-mass spectrometry in clinical chemistry, pharmacokinetics, and toxicology. *Mass Spectrom. Rev.* 11:249-307.
23. Spahr, C.S., S.A. Susin, E.J. Bures, J.H. Robinson, M.T. Davis, M.D. McGinley, G. Kroemer, and S.D. Patterson. 2000. Simplification of complex peptide mixtures for proteomic analysis: reversible biotinylation of cysteinyl peptides. *Electrophoresis* 21:1635-1650.
24. Ficarro, S.B., M.L. McClelland, P.T. Stukenberg, D.J. Burke, M.M. Ross, J. Shabanowitz, D.F. Hunt, and F.M. White. 2002. Phosphoproteome analysis by mass spectrometry and its application to *Saccharomyces cerevisiae*. *Nat. Biotechnol.* 20:301-305.
25. Kaji, H., H. Saito, Y. Yamauchi, T. Shinkawa, M. Taoka, J. Hirabayashi, K. Kasai, N. Takahashi, et al. 2003. Lectin affinity capture, isotope-coded tagging and mass spectrometry to identify N-linked glycoproteins. *Nat. Biotechnol.* 21:667-672.
26. Zhang, H., X.J. Li, D.B. Martin, and R. Aebersold. 2003. Identification and quantification of N-linked glycoproteins using hydrazide chemistry, stable isotope labeling and mass spectrometry. *Nat. Biotechnol.* 21:660-666.
27. Gerber, S.A., J. Rush, O. Stemman, M.W. Kirschner, and S.P. Gygi. 2003. Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. *Proc. Natl. Acad. Sci. USA* 100:6940-6945.
28. Ayad, N.G., S. Rankin, M. Murakami, J. Jebanathirajah, S. Gygi, and M.W. Kirschner. 2003. Tome-1, a trigger of mitotic entry, is degraded during G1 via the APC. *Cell* 113:101-113.
29. Aebersold, R. 2003. Constellations in a cellular universe. *Nature* 422:115-116.
30. Mintz, P.J., S.D. Patterson, A.F. Neuwald, C.S. Spahr, and D.L. Spector. 1999. Purification and biochemical characterization of interchromatin granule clusters. *EMBO J.* 18:4308-4320.
31. Rout, M.P., J.D. Aitchison, A. Suprpto, K. Hjertaas, Y. Zhao, and B.T. Chait. 2000. The yeast nuclear pore complex: composition, architecture, and transport mechanism. *J. Cell Biol.* 148:635-651.
32. Patterson, S.D., C.S. Spahr, E. Daugas, S.A. Susin, T. Irinopoulou, C. Koehler, and G. Kroemer. 2000. Mass spectrometric identification of proteins released from mitochondria undergoing permeability transition. *Cell Death Diff.* 7:137-144.
33. Liu, Y., M.P. Patricelli, and B.F. Cravatt. 1999. Activity-based protein profiling: the serine hydrolases. *Proc. Natl. Acad. Sci. USA* 96:14694-14699.
34. Caprioli, R.M., T.B. Farmer, and J. Gile. 1997. Molecular imaging of biological samples: localization of peptides and proteins using MALDI-TOF MS. *Anal. Chem.* 69:4751-4760.
35. Ideker, T., V. Thorsson, J.A. Ranish, R. Christmas, J. Buhler, J.K. Eng, R. Bumgarner, D.R. Goodlett, et al. 2001. Integrated genomic and proteomic analyses of a systematically perturbed metabolic network. *Science* 292:929-934.
36. Adam, B.L., Y. Qu, J.W. Davis, M.D. Ward, M.A. Clements, L.H. Cazares, O.J. Semmes, P.F. Schellhammer, et al. 2002. Serum protein fingerprinting coupled with a pattern-matching algorithm distinguishes prostate cancer from benign prostate hyperplasia and healthy men. *Cancer Res.* 62: 3609-3614.
37. Petricoin, E.F., G.B. Mills, E.C. Kohn, and L.A. Liotta. 2002. Proteomic patterns in serum and identification of ovarian cancer. *Lancet* 360:170-171.
38. Schweitzer, B., S. Roberts, B. Grimwade, W. Shao, M. Wang, Q. Fu, Q. Shu, I. Laroche, et al. 2002. Multiplexed protein profiling on microarrays by rolling-circle amplification. *Nat. Biotechnol.* 20:359-365.
39. Borrebaeck, C.A.K. and C. Wingren (Eds.) 2002. High-throughput proteomics: Protein arrays. *BioTechniques* 33 (Suppl.):4-69.
40. Scrivener, E., R. Barry, A. Platt, R. Calvert, G. Masih, P. Hextall, M. Soloviev, and J. Terrett. 2003. Peptidomics: a new approach to affinity protein microarrays. *Proteomics* 3:122-128.
41. Leung, D., C. Hardouin, D.L. Boger, and B.F. Cravatt. 2003. Discovering potent and selective reversible inhibitors of enzymes in complex proteomes. *Nat. Biotechnol.* 21:687-691.
42. Gavin, A.C., M. Bosche, R. Krause, P. Grandi, M. Marzioch, A. Bauer, J. Schultz, J.M. Rick, et al. 2002. Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature* 415:141-147.
43. Ho, Y., A. Gruhler, A. Heilbut, G.D. Bader, L. Moore, S.L. Adams, A. Millar, P. Taylor, et al. 2002. Systematic identification of protein complexes in *Saccharomyces cerevisiae* by mass spectrometry. *Nature* 415:180-183.
44. Patterson, S.D. 2003. Data analysis—the Achilles heel of proteomics. *Nat. Biotechnology* 21:221-222.
45. Ryan, T.E. and S.D. Patterson. 2002. Proteomics: drug target discovery on an industrial scale. *Trends Biotechnol.* 20:S45-S51.

Address correspondence to:

Scott D. Patterson
 Farmal Biomedics, LLC
 129 North Hill Ave., Suite 107
 Pasadena, CA 91106, USA
 e-mail: scott@farmalbiomed.com