NMR ON TARGET

Technological innovations are creating a powerful tool for target identification and chemical proteomics.

BY DANIEL SEM, HUGO VILLAR, AND MARK KELLY

Proteomics is devoted to the study of the entire complement of proteins at once, rather than the traditional one-at-a-time strategy. Ideally, proteomics is used in the context of the biological milieu and in the presence of other proteins to address the network of direct and indirect interactions between proteins. This "holistic approach" to biology is on the horizon, and it is enabled by the full genome sequence and a new set of tools under development. The ideal is not yet realized, however, because the field is still largely dominated by the use of 2-D polyacrylamide gels and more recently MS. But another -omics has entered the field—chemical proteomics—and when coupled with NMR, it plays a central role in drug target identification.

From proteins to ligands

Chemical proteomics is devoted to the parallel characterization of ligand interactions with all the proteins encoded by a genome. In this early stage of the proteomics effort, one of the main goals is to identify the best technologies available for massively converting sequence information into small-molecule probes that

enhance our understanding of biological processes or serve as leads for drug discovery. Such studies should allow characterization of a protein in its biological milieu and provide a sense of the specificity of interactions with ligands.

Traditionally, collections of proteins were studied en masse with denaturing polyacrylamide gels, and larger numbers of proteins were studied with 2-D gels. These analyses have been extended into chemical proteomics by companies such as Serenex (*www.serenex.com*) and ActivX (*www.activx.com*), whose researchers have devised means of studying protein–ligand interactions by preselecting subproteome fractions, extracting the components from 1- and 2-D gels, and doing affinity labeling and in-gel detection or mass spectral analysis. And although 2-D gel electrophoresis and MS have become pivotal in this field as tools for visualizing, isolating, and identifying low-abundance proteins, the need for new technologies and methods is stimulating exciting new research and development in this realm.

Some of the more promising new tools are small-molecule probes that can be used to covalently modify a set of related enzymes and subsequently allow their purification and identification as drug targets (*1*). When researchers use this strategy with known, therapeutically active ligands, the ligands can serve as both target-validation probes and novel drug leads (*2*). This parallel strategy provides significant leverage when combined with the use of combinatorial libraries focused on gene families such as kinases, like those developed at Vertex Pharmaceuticals (*www.vpharm.com*). For this parallel attack on a gene family to work, the pairing of a combinatorial library with a gene family must be based on a solid structural proteomics analysis of the gene family. In other words, you need to understand the locks before you start mass-producing keys.

That was the case when researchers at Triad Therapeutics (*www.triadthera.com*) used chemo- and bioinformatic analysis to parse the entire *Mycobacterium tuberculosis* proteome according to dehydrogenase binding sites (*3*). In this strategy, researchers clustered proteins on the basis of conserved cofactor geometry and binding site features, after analysis of more than 200 crystal structures. They developed a method using "protein key" descriptors that enabled the identification of these pharmacofamilies based on sequence alone. Using these protein keys, the scientists determined that *M. tuberculosis* has a disproportionately large family of single-domain Rossman-fold dehydrogenases (2% of its open reading frames). These proteins are typically involved in cell wall synthesis and related reactions, making them likely anti-infective drug targets. Thus, developments in structural proteomics are crucial to the success of these chemical proteomic strategies because they provide the understanding of binding sites in a gene family needed to attack the family in a proteome-wide manner.

NMR in chemical proteomics

For a growing number of researchers, NMR technology offers many potential applications in different areas of chemical proteomics. Among the better-known examples is the use of NMR

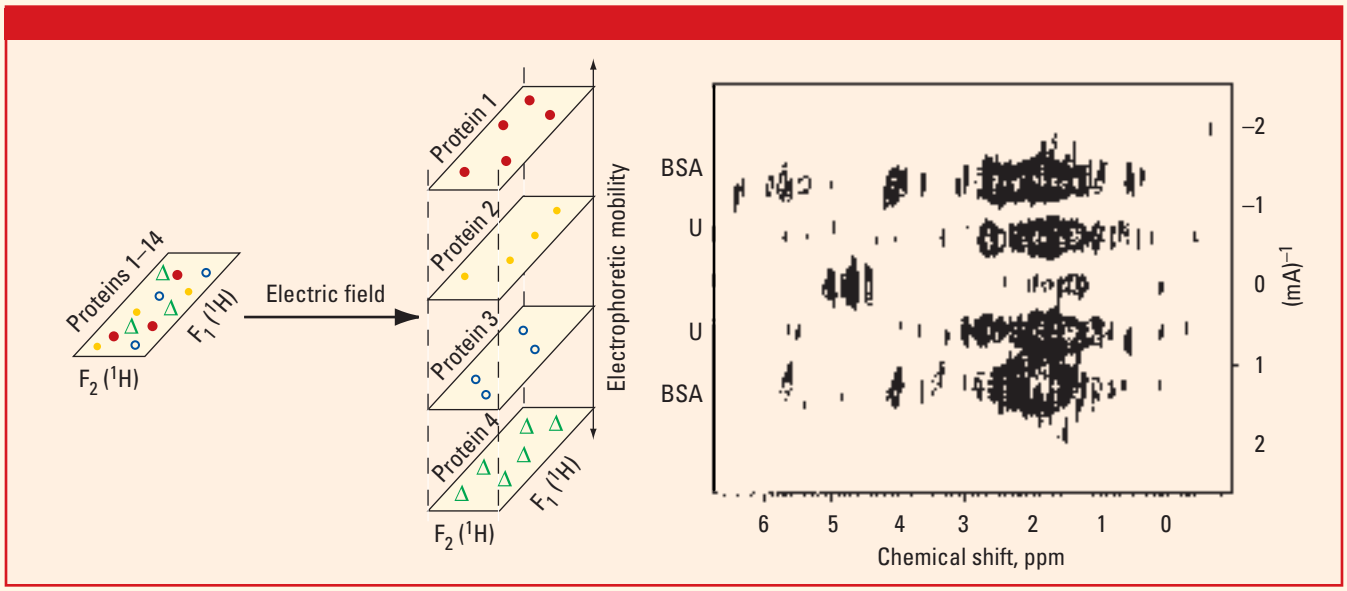

Figure 1. Electrophoretic NMR. Four 2-D NMR spectra of various proteins in a mixture, separated on the basis of electrophoretic mobility (left), and a 2-D spectrum showing the separation of ubiquitin (U) and bovine serum albumin (BSA), with electrophoretic mobility as one of the two axes (right). (Adapted with permission from Ref. 7.)

technology for characterizing protein structures on a large scale, which could be used to establish structure–function relationships and to aid structure-guided drug design. Many prominent NMR groups such as Stephen Fesik's at Abbott Labs (*www.abbott.com*) are working in this arena of high-throughput structure determination and are seeing exciting results.

For example, researchers at several institutes used a variety of methods to determine structures for more than 500 proteins from 5 different microorganisms (*4*). Approximately 20% of these proteins were amenable to NMR. In a similar endeavor, researchers involved in the Southwest Collaboratory for Structural Genomics used residual dipolar couplings to increase the speed of protein fold determination (*5*). They noted that proteins that show evidence of low structural order tend to be difficult to crystallize, and this is where NMR might be most helpful. In one example, they obtained an accurate structure of rubredoxin using only residual dipolar coupling data, which requires ~10% of the data acquisition time of traditional nuclear Overhauser effect strategies. The scientists think that the technique might also be applicable to much larger proteins (20–80 kDa) if there is a need only to validate or verify homology-modeled or ab initio-calculated structure.

NMR is also well suited to the study of weak protein–protein interactions, which are critical to understanding the effects elicited by protein cascades and which also play a role in drug discovery efforts. All of these NMR-based structural proteomic studies might provide the clues needed to assign functions to proteins, a first step toward target identification. They also enable better focusing of combinatorial libraries of chemical proteomic probes to a particular gene family.

NMR IS WELL SUITED TO THE STUDY OF WEAK PROTEIN–PROTEIN INTERACTIONS, WHICH PLAY A ROLE IN DRUG DISCOVERY EFFORTS.

To this end, researchers at Triad Therapeutics developed the NMR SOLVE technology to guide the design of gene-family-focused libraries (*6*). In this strategy, binding sites of even large proteins are mapped with well-characterized ligands such as cofactors. This mapping process allows the assignment of NMR cross-peaks for atoms in the protein binding site, which then serve as spectroscopic beacons for orienting cofactor-binding inhibitors or antagonists in the binding site.

Scientists can use this information to generate a bi-ligand library composed of the antagonist covalently bound to other molecules (diversity elements) in such a manner that the latter are directed to an adjacent specificity pocket. This method works on the principle that molecules bound weakly at two sites on a protein will bind more tightly when joined than a compound bound at only one.

Thus, even in the absence of a full characterization of protein structure, by combining techniques from NMR, chemistry, and bioinformatics, it is possible to design collections of bi-ligand antagonists that interact with many members of a protein family. Such molecules can then serve as chemical proteomic probes of function.

NMR has also taken a nod from the traditional 2-D protein gels. A method referred to as 2-D electrophoretic NMR allows for the separation of proteins in an electric field, but in solution rather than a gel (*7*). This NMR method provides an electrophoretic dimension for protein mobility, as shown in Figure 1. The method has tremendous potential as a tool in studies of multiple targets in parallel and might someday be a powerful adjunct to gel-based strategies for studying protein target function.

Crossing families of ligands and proteins

Of course, drug discovery would be accelerated if there were a way of visualizing protein–ligand interactions with large numbers of compounds proteome-wide. To some extent, this is now feasible in the context of protein gels with the use of fluorescent chemical proteomic probes.

Companies such as Amersham Biosciences (*www.amershambiosciences.com*) provide machines that scan gels for fluorescently tagged proteins, while others like Molecular Probes (*www.molecularprobes.com*) offer large toolboxes of fluorescent probes. But can we move beyond gels? Although studies in solution using electrophoretic NMR offer one alternative, in vivo studies are preferred.

Fluorescence microscopy can provide views of protein–ligand interactions in organisms such as zebrafish (see "Imaging: Portraits from life", *Modern Drug Discovery*, April 2003, p 30), and molecular imaging using magnetic resonance imaging (MRI) or magnetic resonance spectroscopy (MRS) will likely provide a complementary noninvasive look within organisms (see "The anatomy of metabolism", *Modern Drug Discovery,* April 2003, p 37).

Such a holistic approach to biological studies suffers mostly from a lack of sensitivity. Although imaging methods like positron emission tomography do not suffer from this limitation, they have their problems, such as the need for appropriately labeled positron-emitting biomarkers. MRI and MRS can still provide this inward glimpse of protein–ligand interactions if sensitivity can somehow be boosted. Indeed, such a boost might be provided by the noble gas xenon. Hyperpolarized xenon is used to enhance the sensitivity of MRI images, and Alex Pines and colleagues at the University of California, Berkeley (*www.berkeley.edu*), and the Lawrence Berkeley National Laboratory (*www.lbl.gov*) have performed several in vivo imaging studies that demonstrate its potential (Figure 2, *8*).

NMR has played a prominent role in the growing field of metabonomics.

Most methods that use NMR to obtain information from living specimens (ranging from cell suspensions to human beings) have focused on small molecules that can be distinguished from all other molecules in the cell by their abundance or because they have been isotopically labeled. Studies of changes in proteins prompted by similar means are starting to appear in the literature. They are opening new avenues to characterizing the dynamics of proteins and other biological macromolecules in their natural environment.

From target to antitarget space

Although the pursuit of noninvasive, highly parallel ligand–protein interaction studies continues, it must be realized that such analyses have two goals: to define targets to understand biology better, and to design better and more selective drugs with fewer side effects. To accomplish the latter requires a consideration of interactions with enzymes that degrade the drug or molecule—the so-called xenobiotic metabolizing enzymes. To this end, NMR has also played a prominent role in the growing field of metabonomics. As shown by the efforts of Jeremy Nicholson and colleagues (Figure 3, *9*), who developed the concept of metabonomics and founded Metabometrix (*www.metabometrix.com*), NMR is used to characterize the metabolic profile of a drug to better predict, understand, and avoid pathophysiological side effects.

Drugs are metabolized primarily by the cytochrome P450 (CYP) enzymes, a large family of drug antitargets. Researchers try to keep their new drugs from binding to these proteins—that is, to antitargets rather than targets. NMR also provides glimpses of how drugs bind to the drug-metabolizing CYP enzymes; by using T1 relaxation studies, scientists have obtained distances of drug atoms from the reactive heme iron. This allows researchers to dock drugs in the CYP

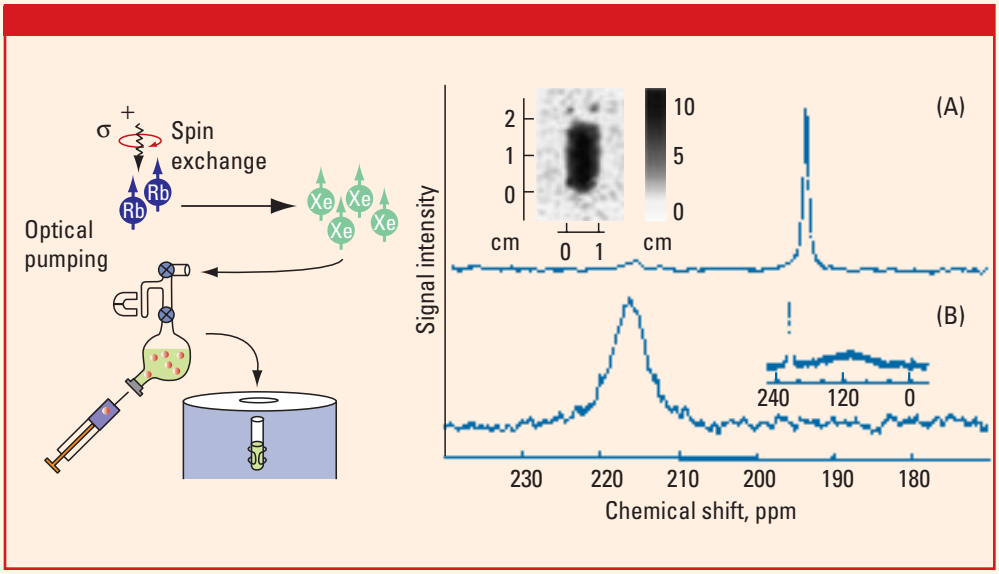

Figure 2. Enhancing NMR imaging with hyperpolarized xenon. Researchers prepare hyperpolarized ^{129}Xe by optical pumping of Ru gas with a laser (left). ^{129}Xe NMR spectra of human red blood cell samples using hyperpolarized xenon in an (A) Intralipid or (B) Fluosol carrier solution (right). The peak at 194 ppm is for ^{129}Xe in a lipid environment, and that at 216 ppm is for ^{129}Xe in red blood cells, with broadening due to interactions with hemoglobin. (Adapted with permission from Ref. 8.)

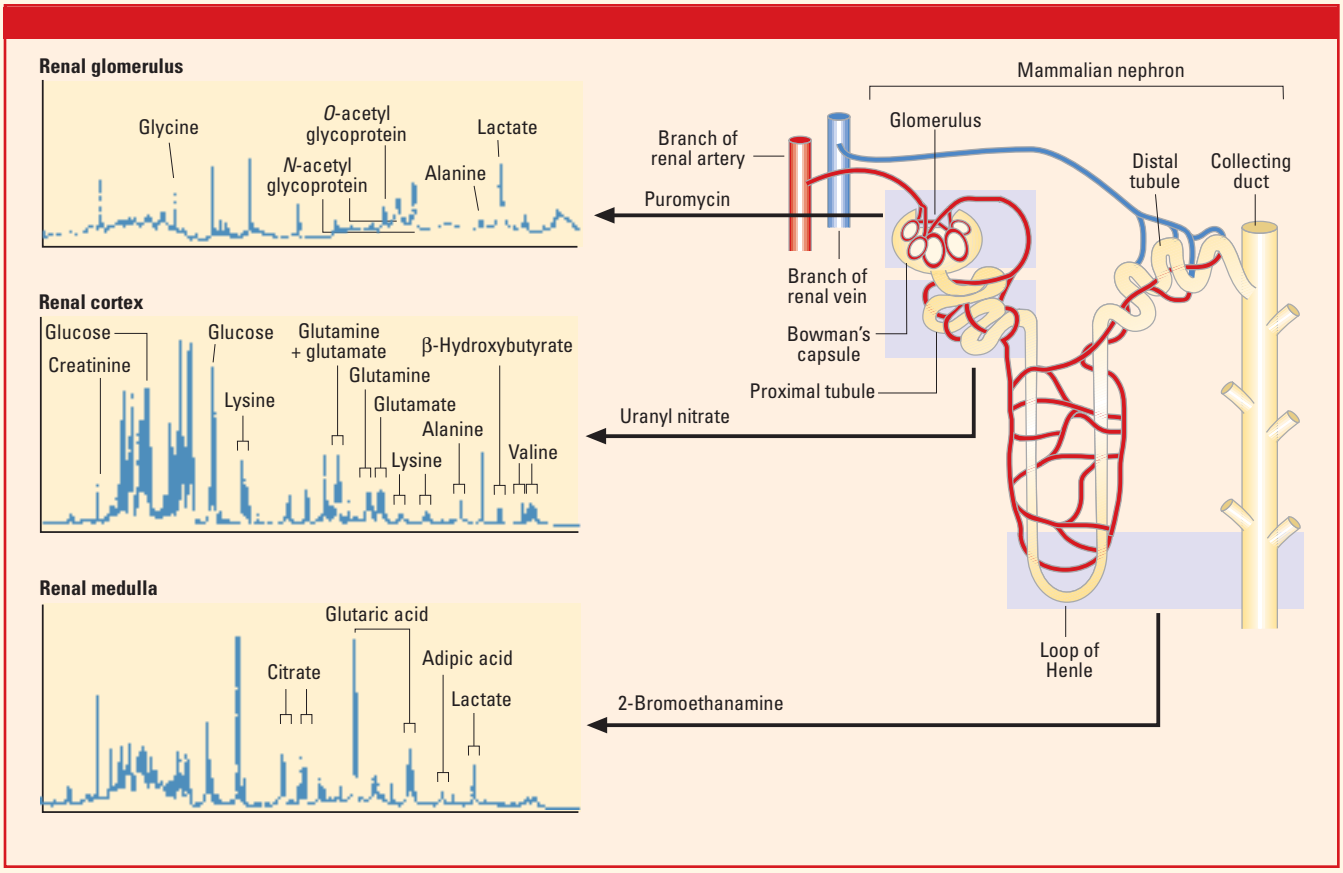

Figure 3. Metabonomic studies of kidney toxicity. Researchers generated kidney damage using the site-selective xenobiotic agents Puromycin, uranyl nitrate, and 2-bromoethanolamine. Each produces a unique profile in the ^{1}H NMR spectrum of biofluid samples. (Adapted with permission from Ref. 9.)

binding site, because T1 values have sixth-power (r^6) dependence on the distance between its protons and the unpaired iron electron. Gordon Roberts and colleagues at the University of Leicester (*www.le.ac.uk*) and the University of Dundee (*www.dundee.ac.uk*) used this approach to provide the first pictures of codeine bound to the CYP enzyme that metabolizes it (*10*).

Future prospects

Chemical proteomics will play a vital role in target identification for both basic and applied studies. Although studies are beginning to extend to whole organisms using fluorescence microscopy and molecular imaging, hope is now growing that NMR imaging might someday provide us with the most comprehensive, noninvasive, and holistic picture of proteome-wide interactions with ligands.

Such is the vision of the Chemical Proteomics Facility at Marquette University (*www.marquette.edu/chem/personnel/faculty/sem.html*), which will be devoted to the use of NMR, including xenon-enhanced and electrophoretic NMR, for the proteome-wide study of protein–ligand interactions. And the use of zebrafish as a model system with parallel fluorescence microscopy studies will also figure prominently. This latter work is being done in collaboration with the National Institute of Environmental Health Sciences Marine and Freshwater Biomedical Sciences Center at the University of Wisconsin, Milwaukee (*www.uwm.edu/Dept/MFB/foundation.html*). These strategies, complemented by metabonomic profiling, will hopefully lead to the identification of new drug targets as well as drugs that are more specific for their intended targets, and will also provide a better understanding of certain aspects of biology. This is, in any case, the promise of chemical proteomics, and NMR technology will likely be a vital part of that toolbox.

References

(1) Jeffery, D. A.; Bogyo, M. *Curr. Opin. Biotechnol.* **2003,** *14,* 87–95.
(2) Lenz, G. R.; Nash, H. M.; Jindal, S. *Drug Discovery Today* **2000,** *5,* 145–156.
(3) Kho, R.; et al. *Proteins* **2003,** *50,* 589–599.
(4) Yee, A.; et al. *Proc. Natl. Acad. Sci. U.S.A.* **2002,** *99,* 1825–1830.
(5) Adams, M. W. W.; et al. *Acc. Chem. Res.* **2003,** *36,* 191–198.
(6) Sem, D. S.; et al. *J. Cell. Biochem.* **2001,** *37,* S99–S105.
(7) He, Q.; Liu, Y.; Nixon, T. *J. Amer. Chem. Soc.* **1998,** *120,* 1341–1342.
(8) Bifone, A.; et al. *Proc. Natl. Acad. Sci. U.S.A.* **1996,** *93,* 12932–12936.
(9) Nicholson, J. K.; et al. *Nat. Rev. Drug Discovery* **2002,** *1,* 153–161.
(10) Modi, S.; et al. *Biochemistry* **1996,** *35,* 4540–4550.

Daniel Sem is an assistant professor at Marquette University (*www.marquette.edu*). **Mark Kelly** is a scientist and **Hugo Villar** is vice president of chemoproteomics at Triad Therapeutics (*www.triadthera.com*). Send your comments or questions about this article to mdd@acs.org or to the Editorial Office address on page 3. ■

KEY TERMS: automation, combinatorial chemistry, high throughput, informatics, medicinal chemistry, modeling, proteomics, screening, technique