Fingerprinting the unusual suspects

Chemoproteomic techniques might help drug investigators tie biological knowledge to chemistry.

BY DAVID BRADLEY

Law enforcers generally fingerprint the usual suspects when a crime has been committed. This age-old forensics method has convicted countless criminals on the grounds that they were at the right place at the right time to have committed the crime. By the same token, taking fingerprints of chemicals—using techniques from the growing field of chemoproteomics—could lead drug investigators to new and unusual suspects, tying complicated biological knowledge directly to chemistry.

Virtual screening has many advantages over conventional synthesis and testing approaches to drug discovery. In silico, any number of compounds can be created from scratch without the fuss of designing a reaction scheme. The molecules can be pulled in from various database sources, and libraries of compounds can be tweaked and fine-tuned to create new diversity that would take many lab hours to build in vitro.

Moreover, postgenomically speaking, we are seeing a burgeoning collection of newly revealed protein targets—whether enzyme or receptor—that might be laid open to new therapeutic agents. There is more than a chance that in vitro screening methods will collapse under the weight of this rapidly growing number of macromolecular targets.

In silico methods, however, might just allow drug discovery scientists to lighten the load, providing a means of cutting the number of potential compounds to a more manageable number that will allow leads to be picked up more quickly.

A certain affinity

The problem, as it stands, is far worse than looking for a needle in a haystack. Jürgen Bajorath, senior director of computer-aided drug discovery at Albany Molecular Research (*www.albmolecular.com*), believes that compound classification and similarity search methods will be key to finding the desired needle. Virtual screening by molecular similarity allows scientists to search beyond just chemistry by following the similar-property principle, which states that similar molecules will have similar physicochemical and biological properties.

One line of attack, developed in the mid-1990s by Palo Alto-based Telik (*www.telik.com*), is affinity fingerprinting. This allows small molecules to be characterized and compared on the basis of their binding profiles to a reference set of proteins. The technique is experimental and "direct" in nature, explains Bajorath. Every molecule in a collection is tested against each of the reference proteins, and the set of binding affinities associated with a given compound represents its affinity fingerprint (Figure 1). The fingerprint provides a unique identifier for each chemical and represents the binding of that "type" of molecule within a protein reference framework. It is analogous to the multidimensional chemical spaces that define compounds in terms of their solubility, logP, pK_a, and other descriptors. In other words, the descriptors in the "protein space" are the binding affinities to the panel of reference proteins.

Affinity fingerprints can be used to pick out likely leads, and these in turn can be used as individual probes to validate the potential therapeutic value of a target. There is little point to focusing on a putative protein target, creating an expensive library of compounds, and using high-throughput screening methods only to discover that the target had no disease relevance in the first place. "The key to identifying these probes is the computational analysis of each compound's affinity fingerprint," says Paul Beroza, director of computational chemistry and informatics at Telik. "Patterns in the affinity fingerprints are associated with biological activity and provide the basis for selecting compounds for assay."

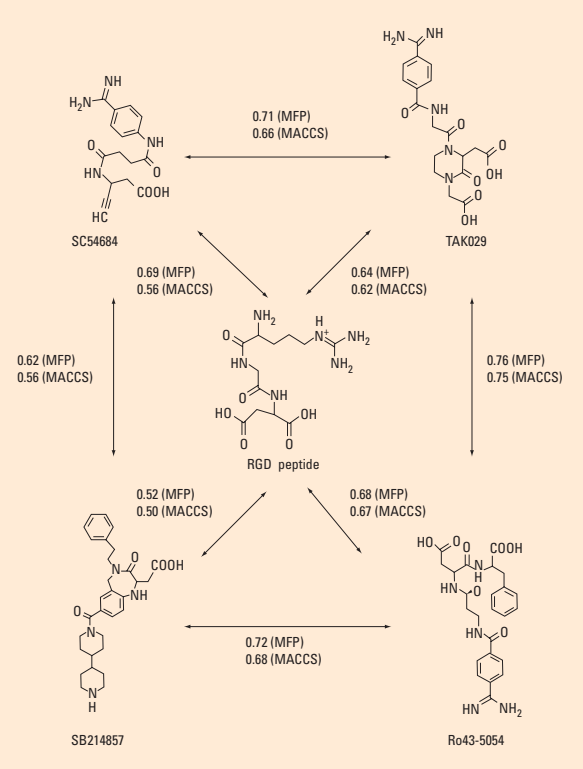

Figure 1. Looking for clues. By using two different pharmacophore fingerprinting methods (MACSS and MFP), researchers score the similarities of the Arg-Gly-Lys (RGD) peptide that naturally binds the fibrinogen receptor and four receptor agonists. (Adapted with permission from Bajorath, J. *J. Chem. Inf. Comput. Sci.* **2001,** *41,* 233–245.)

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

Beroza and colleagues have pointed out that there is almost a glut of methods for classifying molecular similarity and diversity. They say this state has emerged from a desire to redress the imbalance between the sheer numbers of small molecules in libraries worldwide and those with real druglike qualities. Structure and physicochemical characteristics have remained at the forefront until now, but chemoproteomics—a newly emerging field that is likely to build the bridge between genomics, proteomics, and chemistry—has surfaced as an alternative, or at least as a complementary tactic that allows drug teams to make the most of biological information to guide the chemistry. Chemoproteomics could "offer a highly efficient alternative to small-molecule characterization that can accelerate drug discovery in the postgenomic era," according to Beroza's team.

Apply and demand

An interesting application of affinity fingerprints is in expanding a small-molecule compound library to include potentially useful candidates that would otherwise not be found. As new compounds are synthesized or acquired, Beroza explains, their fingerprints are taken, but they are added to the library lineup only if they contribute some novel descriptor information to the library. Compounds that behave toward the set of reference proteins in much the same way as the earlier suspects are not pursued. Compounds are thus assessed computationally by calculating the number of other compounds in their fingerprint neighborhoods. A fairly solitary compound is relatively rare and worth keeping a close watch over.

Beroza further explains that chemists can then synthesize or acquire compounds with structures similar to the rarities and so populate sparse neighborhoods in a database. The point is that small structural changes usually cause small changes in biological activity, which might provide a similarly fingerprinted compound, for example, one that has high affinities for particular proteins in the reference set but with potentially useful yet subtle differences in physicochemical properties.

An attractive feature of this technique is that successive iterations can be used to refine the affinity fingerprint model that is used to mine databases for active compounds. Beroza's team has used this approach to find compounds that are potent in micromolar concentrations against given targets after just three rounds of compound selection and assay. Thus, he explains, only 200 or so compounds need be tested in vitro to provide a sample of leads, whereas the high-throughput screening method of looking at random compounds may require researchers to work with tens or hundreds of thousands of molecules.

The affinity fingerprint approach used by Beroza and colleagues recently resulted in two new drug candidates currently in clinical development. The first is Telcyta (TLK286), which is in a Phase III clinical trial for the treatment of ovarian cancer. The second, known only as TLK199, is in a Phase I/IIa trial for the treatment of a precancerous blood disorder, myelodysplastic syndrome. "In addition," says Beroza, "Telik is applying this technology in collaboration with academic cancer centers to identify drug leads for a number of therapeutic targets, including human intestinal carboxylesterase, laminin-5, specific integrins, and poly(ADP–ribose) glycohydrolase."

Target proteases

Meanwhile, Matthew Bogyo and colleagues at the University of California, San Francisco (*www.ucsf.edu*), recently described a pertinent example of chemoproteomics in action. They generated a large data set of small-molecule affinity fingerprints for a group of closely related enzymes, the papain family of cysteine proteases. These compounds include mammalian enzymes such as cathepsins B and L, which are involved in cancer growth and metastasis, and cathepsin K, which is implicated in osteoporosis. They also include several parasitic enzymes that allow parasite and host to interact in *Trypanosoma cruzi*, the pathogen in Chagas' disease, and *Plasmodium falciparum*, the malaria parasite.

The Bogyo team generated binding data for a library of inhibitors based on the ability of each compound to block active-site labeling of the target proteases using a covalent-activity-based probe (ABP). Computation in the form of clustering algorithms then allowed the team to automatically classify a reference group of proteases into subfamilies based on their affinity fingerprints. In this way, the team identified cysteine protease targets modified by the ABP.

With this information in hand, they could then couple the experimental data computationally with reported crystal structures to allow them to predict small-molecule inhibitors. The team says that the method might be generalized for large enzyme families to help design selective inhibitors based on limited structure–function information.

The players

Several companies are pinning their hopes on chemoproteomics and molecular fingerprinting for high-throughput screening. These companies include:

ActivX Biosciences	www.activx.com
Albany Molecular Research	www.albmolecular.com
Cengent Therapeutics	www.cengent.com
Crystal Genomics	www.crystalgenomics.com
Graffinity Pharmaceuticals	www.graffinity.com
Kyorin Pharmaceutical	www.kyorin-pharm.co.jp
Neogenesis	www.neogenesis.com
Serenex	www.serenex.com
Telik	www.telik.com
Triad Therapeutics	www.triadthera.com
Vertex Pharmaceuticals	www.vertex.com

Chemical probes

In a more commercial vein, researchers at ActivX (*www.activx.com*) are developing what they describe as the first high-throughput chemistry platform for directly measuring protein activity on a global scale. The aim is to convert the masses of raw data from genomics into drug discovery targets. Information about small-molecule–protein interactions will allow the scientists to build chemical probes that monitor the functional state of large protein families by binding to the common structural element in all members of the family. This means that the small molecules avoid proteins bound to an inhibitor or in an inactive state.

The approach is based on the work of company co-founder Benjamin Cravatt and his team at the Scripps Research Institute (*www.scripps.edu*), who focused on tracking the activity of the serine hydrolase family of enzymes. The method has now been used to differentially detect active and, moreover, novel membrane-bound serine hydrolases in breast cancer cell lines.

The bottom line

From an academic perspective, of course, it is easy to screen the various approaches to drug discovery, as companies rarely highlight exactly how they arrived at a particular product on the market. There are a few notable exceptions of standalone compounds, such as Viagra (sildenafil citrate) and Taxol (paclitaxel), which took intriguing routes to discovery.

The bottom line in drug discovery is to find new drugs and bring them to market. The industrialization of the drug discovery process, with high-throughput, combinatorial, and parallel synthesis techniques, physicochemical profiling, and, more recently, affinity fingerprinting have been put into the service of pharmaceutical companies. The aim is to multiply the number of new drugs available to medicine without concomitantly raising the cost of R&D. Chemoproteomics may sound like yet another aphorism with an -omic suffix, but it could provide a much-needed link between biological information and chemistry, helping drug discovery specialists crack the case.

Further reading

Bajorath, J. Affinity fingerprints—Leading the way? *Drug Discov. Today* **2002,** *7,* 1035.

Bajorath, J. Selected concepts and investigations in compound classification, molecular descriptor analysis, and virtual screening. *J. Chem. Inf. Comput. Sci.* **2001,** *41,* 233–245.

Beroza, P.; et al. Chemoproteomics as a basis for post-genomic drug discovery. *Drug Discov. Today* **2002,** *7,* 807–814.

Greenbaum, D. C.; et al. Small-molecule affinity fingerprinting: A tool for enzyme family subclassification, target identification, and inhibitor design. *Chem. Biol.* **2002,** *9,* 1085–1094.

David Bradley is a freelance science writer at sciencebase.com based in Cambridge, England. Send your comments or questions about this article to mdd@acs.org or to the Editorial Office address on page 3. ■