

January 28, 2002

BioCentury

BioCentury 100™ Indicators



Week ended 1/25/02

PRICES
1694.06
up 0.2%

VOLUME
478.4M shrs
dn 19%

VOLUME 10, NUMBER 5

THE BERNSTEIN REPORT

ON BIOBUSINESS

PAGE A1 OF 21

Product Development

Drug-likeness rules

BioCentury This Week

Cover Story

Attempts to improve Lipinski's Rule of 5 for predicting drug-likeness are focusing on the physicochemical properties of molecules and computer models based on biological assays.

The metabolic landscape/Page A6

Strategy

ABI's rebuilding year/Page A8

Product Development

Amgen's coup in Germany;
Amarin's LAX-101/Page A9

Technology Briefing

Tissuelnformatics; cyclic peptides/Page A10

Emerging Company Profile

Innate: Teach the immune system/Page A11

Politics & Policy

EPO restricts patent access/Page A12

Washington: PDUFA stutters/Page A13

Europe: Embryo research/Page A14

Ebb & Flow

More ImClone revelations; Ostro resurfaces; et al./Page A15

Focus: Keeping talent at IMNX/Page A19

Online this week/Page A8

Stock charts & tables/Page A20

Company index/Page A13

By Kathryn Calkins
Senior Writer

Drug development would be a lot more successful and much less costly if it were simple to tell which molecules would behave like drugs. In fact, given the number of drugs on the market, it would seem like it should be easy to design new compounds that behave like drugs. But even though the search for drug-likeness predictability has gone on for at least 100 years, medicinal chemistry remains as much an art as a science.

The gold standard is Lipinski's Rule of 5, which was proposed in 1995 (see "Rule of 5," A2). The promulgation of the rule sparked renewed interest in drug-likeness models, which have proceeded along two paths: the physicochemical properties of molecules and computer models based on data from biological assays. While both have a long way to go, the combination of increased volumes of biological and chemical data, and more powerful computing capabilities, should result in continued improvement in models.

Impetus for the Rule

High throughput screening of small molecule compounds focused attention on the drug-likeness issue in the early 1990s, because the new methods tended to identify compounds with more solubility and permeability problems than older methods had produced.

In response to this problem, Christopher Lipinski, a scientist at Pfizer Inc., proposed a set of rules based on physi-

'Off the top of my head, how many people want drug-like compounds versus chemical diversity? I'd say it's 70%-30%.'

— Christopher Lipinski of Pfizer

cochemical properties of molecules that predispose compounds to better solubility and permeability. Industry has embraced the rules, which allow companies to avoid types of chemical structures that are unlikely to have drug-like properties early in drug discovery.

But Lipinski's rule is not infallible, and is a far cry from having a magic machine that takes in the three-dimensional structure of a target and spits out an ideal lead candidate. Some structures that fall outside the parameters can make good drugs, and there is no guarantee that compounds within the parameters will succeed.

Thus researchers have continued to look for new rules, and a few new drug-likeness observations have been published since 1995, such as observations related to polar surface area and the role of active transport.

But the pace of understanding drug-likeness has not accelerated for several reasons. Foremost is a lack of biological data. Predicting absorption, distribu-

See next page

C21 update

The 4th C21 Ventures conference in San Diego issues its call for presenters. Please see announcement following A21.

Drug-likeness Rules,
from previous page

tion, metabolism, excretion and toxicity (ADMET) based on biological assay data is a hot area, but researchers are only beginning to understand how to create assays that will provide an accurate representation of ADMET properties in humans. Without such assays, data cannot be amassed to drive predictive modeling.

What to avoid

Even before Lipinski articulated the Rule of 5, chemists knew from experience to avoid certain structures that will not make good drugs. "There's a set of 'avoid this' rules," said Richard Cramer, chief scientific officer of Tripos Inc. (TRPS, St. Louis, Mo.).

One example is featureless molecules. "Molecules without any features, like long-chain hydrocarbons, don't make good drugs, because they tend not to do anything specific," Cramer said (see "Predicting Drug-Likeness", A7).

Chemists also know to avoid compounds that form covalent chemical bonds, that react easily in a biological system or that have low level activity in many different assays. All are signs that a compound could be toxic. "A chemist can eliminate these compounds from consideration just by looking at them," Cramer said.

Conversely, compounds with OH and NH groups do make good drugs, because they are less lipophilic and therefore better absorbed by the body.

Using modeling and pattern recognition from existing data, chemists also have been able to define other aspects of drug-likeness. For example, a good drug needs to be soluble and permeable, while active transport across membranes can sometimes help and sometimes hurt a compound's drug-likeness.

However, these unwritten rules and definitions proved inadequate to combat the tendency for high throughput screening (HTS) to identify large numbers of molecules that were less soluble and less permeable than those identified through older methods. In addition, Cramer said, combinatorial chemistry emphasized the rapid generation of large numbers of compounds, while de-emphasizing "good" structures such as those with OH and NH groups that are more difficult to make.

The Rule of 5

In 1995, Christopher Lipinski consolidated observations from the literature into a set of rules related to drug absorption. Each observation involves a multiple of 5. Thus, the name, "Rule of 5." ClogP is a measure of a molecule's hydrophobicity.

Absorption is likely to be poor if:

- Molecular weight is >500
- Calculated octanol/water partition coefficient (ClogP) is >5
- Number of hydrogen bond (H-bond) donors on molecule is >5
- Number of H-bond acceptors is >10

Teasing out the rule

Lipinski, senior research fellow at Pfizer Global Research and Development (Groton, Conn.), told BioCentury his Rule of 5 was prompted by his observation that HTS was resulting in an inherent bias toward large, greasy molecules that were less soluble in aqueous media and therefore less well absorbed. As a result, active leads were becoming less attractive in terms of their pharmaceutical properties.

Part of the problem was that while chemists are good at pattern recognition, they are less comfortable with equations, the language of pharmaceutical scientists. "We needed something that would connect for the chemists," Lipinski said. "I had realized from doing database work that there were simple parameters that would be meaningful."

The parameters were based on observations from the literature about physicochemical properties of drug-like molecules. "I got a list of drug-like compounds, all of which had an international designation and a U.S. designation, which means that most were entering Phase II clinical trials and thus were going to be fairly drug-like," Lipinski said. "Initially, I had about 2,500 compounds."

He looked at the distribution of physicochemical properties and saw a plateau effect at certain thresholds. More than 90% had a molecular weight of 500 or less, had lipophilicity below log 5, had 10 or fewer hydrogen bond acceptors, and had five or fewer hydrogen bond donors. Each characteristic included a five or a multiple of five, hence the Rule of 5.

"All these observations were already in the literature," Lipinski said. "But we pulled them together. The timing was right for the industry. People began using these as simple guidelines for medicinal chemistry. And in the last two to three years, you begin to see the impact on what is available commercially. For example, companies marketing combinatorial libraries say, 'we offer compounds screened according to the Rule of 5.'"

Thus he added a warning to PFE's compound database, which would pop up for any compound that fell outside his parameters. "Within 18 months, we had dramatically changed the properties of our compounds. Whereas prior to the warning being added to the database, 25-30% of compounds were out of range, 15% were out of range afterward, and it has been stable since," he said.

One of the benefits of the rule was that it prompted chemists to look at multiple properties up front.

"The Rule of 5 brought forward for the first time the idea that you need more than just target affinity," said Frederic Revah, vice president of drug discovery at Cerep S.A. (NM:Cerep, Rueil-Malmaison, France). "Until then, the chemist did affinity first, then looked at other elements. But Lipinski said 'you should be able to consider all the parameters up front.' At the time the rule was proposed, the chemist was making many compounds as quickly as possible and making them as diverse as possible. Lipinski said that diversity is not the only thing. Quality criteria exist that are necessary for success."

"Off the top of my head, how many people want drug-like compounds versus chemical diversity? I'd say it's 70%-30%. It has to do with the difficulty in addressing the target," Lipinski said.

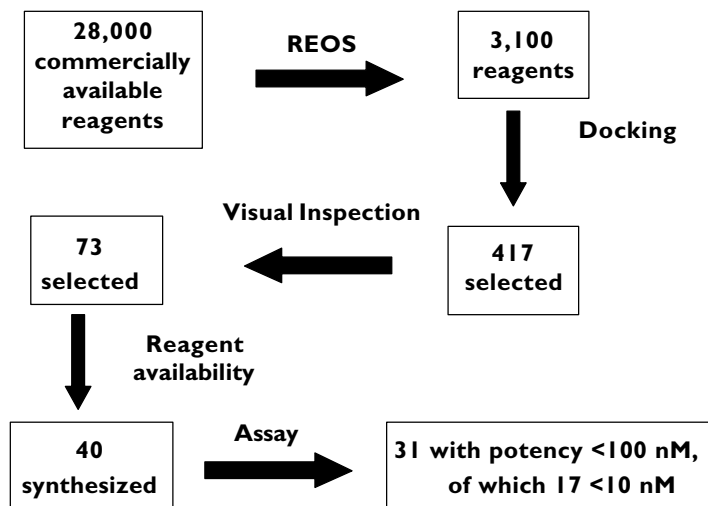
"The rules provide a template that is very easy for all chemists to follow," said Barry Selick, president and CEO of Camitro Corp. (Menlo Park, Calif.), a subsidiary of ArQule Inc. (ARQL, Woburn, Mass.).

Of course, the rules are only guidelines, and certain valuable compounds fall outside the parameters. For example, substrates for biological transporters fall outside the rule, but the transporters can enhance absorption, making these an attractive class of compounds. "This is a hot area now," Lipinski noted. Moreover, biological

See next page

Vertex's in silico approach

Vertex believes that its in silico approach narrows the number of compounds it needs to make and assay in order to get a lead, in this example against ICE (interleukin-1 beta converting enzyme). REOS (Rapid Elimination of Swill) is the company's set of computer programs for drug-likeness. The "Docking" step uses software programs to test binding against the target. "Reagent availability" refers to the availability of commercial reagents. One of VRTX's ICE inhibitors, prlnacasan (VX-740), is in Phase II trials in rheumatoid arthritis with partner Aventis S.A.



Drug-likeness Rules, from previous page

assays are becoming available to screen for transporter activity.

Thus, the rule is imperfect, and since its broad adoption, Lipinski said it has been tailored by many for their own uses. "For example," he said, if you're looking at compounds against central nervous system targets, you need lower molecular weights and tighter parameters on the number of hydrogen-bond donors, etc."

Defining chemical space

Chemistry can be thought of as the universe of possible compounds distributed through three-dimensional space according to their structure. "The search space is enormous," said TRPS's Cramer. But according to a Lipinski paper published in 2000 in the *Journal of Pharmacological and Toxicological Methods*, drug-like compounds are not uniformly distributed through chemical space. Applying the Rule of 5 to a diverse set of chemical structures results in clusters of like structures whose properties fall within the rule's param-

eters.

"The literature emphasis on combinatorial chemistry and the screening of up to million(s) of compounds tends to obscure the fact that the world of drug-like compounds is quite limited and that most of the information content related to desirable drug-like properties is contained in relatively small numbers of compounds," Lipinski wrote.

Moreover, if structures that are active against pharmaceutical targets form one subset of chemical space, the structures that conform to drug-likeness properties form an even smaller subset. And, according to

Revah, the universe shrinks even more. "As you develop models, you restrict more the parts of the space where drug-like molecules exist," he said.

"Maximal chemistry diversity is an inefficient library design strategy, unless there are vast numbers of useful undiscovered targets," Lipinski wrote in the paper. In fact, even though the list of validated targets is growing, it remains fairly limited.

Efforts to define chemical space by drug-likeness fall into at least two categories. First are criteria like the Rule of 5 based on physicochemical properties of molecules or their structures. Second – and the hottest area in the field – is development of computer models based on data from biological assays. Both, said Cramer, "are generally designed to answer the question whether a molecule can get to where it can do something useful."

Non-biological efforts

Scientists have added general drug-likeness criteria to the literature since the Rule of 5. For example, K. Palm and

colleagues at Uppsala University, looking at the ratio of polar to non-polar surface area of a molecule, have defined certain parameters within which the ratio should fall. They published their work in the *Journal of Pharmaceutical Science* in 1996.

"This says something more exact about permeability. But it requires a 3-D model of the molecule to do," Cramer said.

Similarly, Gabriele Cruciani, of the University of Perugia, has used 3-D models to develop a computer program called VolSurf that describes molecules in terms of their volume, surface area and relationship to the molecules surrounding them to predict solubility and permeability. "This seems to be working fairly well," Cramer said.

Some drug-likeness criteria have been formulated by applying new tools to existing data, e.g., using computational programs to tease out possible reasons why some molecules became drugs and others did not based on their documented structures and physical properties rather than performance in biological assays.

Mark Murcko, vice president and chief technology officer at Vertex Pharmaceuticals Inc. (Cambridge, Mass.), said that VRTX has had an active program producing guidelines based on the historical record. The company has about 300 such guidelines to help it understand drug-likeness.

Murcko noted that the biology underlying such guidelines need not be elucidated for the guidelines to prove useful. For example, Murcko and his colleagues have analyzed the elements of compound shape that are commonly found among commercially available drugs. Depending on the way that Murcko analyzed the structures, he found that about 32 structural frameworks accounted for half of the drugs analyzed, suggesting that the diversity of shapes is low among drugs that work.

"The drugs which possess these topological shapes are quite different in polarity, conformation, hydrogen-bonding potential, and other properties; they bind to different classes of receptor; and they serve different pharmacological needs," Murcko wrote in the *Journal of Medicinal Chemistry*. The paper suggests that although the structural themes do not account for all drugs, they could be used as a basis for constructing compound libraries.

See next page

Drug-likeness Rules,
from previous page

ies.

"It's possible to tease out the biology sometimes," Murcko said. For example, drug molecules with large numbers of hydrogen bonds have more trouble crossing the cell membrane. "We can make sense of this," he said. "It is because the compound is surrounded by a shell of water molecules."

Some areas are more difficult to understand, because they involve more than one mechanism. "It's still useful to do the work, however, because you still know what molecules to look at even without knowing the reason why at the atomic level," Murcko said.

Use of such guidelines, he said, "is a tree pruning exercise." The historically derived rules are pruning saws that allow the chemist to cut off certain branches because statistically they are unlikely to yield a satisfactory drug. "You always know, however, that there could be a drug down the branch you cut off. It's a matter of choice."

Modeling from biological data

The biggest current efforts to understand drug likeness use biological assay data to predict ADMET in computer models. But the data are limited. Shelia DeWitt, director of business development at ARQL, noted that "the physico-chemical data were there for the Rule of 5, but for ADMET, new data must be generated."

However, the pace should be accelerated by increasing knowledge about what assays should be conducted to elucidate ADMET properties. Indeed, Cramer noted that assays for absorption, distribution and excretion are becoming more reliable, even if assays for metabolism and toxicity are still in the ad hoc stage.

Lipinski also noted that all the big pharma companies now have in silico prediction programs that have reached operationally useful levels, depending on the size of the compound database being screened. Larger libraries can be screened usefully with the models available now.

With computer models, "the idea is to improve over a random process," Lipinski said. In a model that predicts whether a compound's water solubility is good or bad, "a random process —

'Drug-likeness has everyone's attention right now, which was not true three or four years ago.'

— Mark Murcko of Vertex

pure chance — is correct 50% of the time. So, if we can predict to 70-80% of correctness, we have a much better library that overall will be more soluble."

A 10,000-compound library, he said, can tolerate a 70-80% correct model. However, he said, "the accuracy of a computer model has to improve when dealing with smaller numbers of compounds, because important decisions are made on the properties of just a few compounds. An incorrect prediction may send the chemist down the wrong pathway and waste a lot of time. So, in dealing with small numbers of compounds, chemists tend to rely much more heavily on experimental data and tend to demand much higher accuracy in a computer model."

One barrier to accumulating enough data to achieve good predictive models is the fact that biological processes contain many steps and must be broken down to their component parts to be assayed. Because laboratory endpoints are measures of single elements of metabolism, toxicity and other properties, no one knows how predictive they are of an entire property, Cramer noted.

Thus the state of the art for in silico models seems to be to break down ADMET properties into component assays and gather reams of data. And the key to gathering the volumes of data required to train the models will be designing assays with easy-to-measure outcomes.

By contrast, Lipinski noted, existing biological activity screens already have easy-to-measure outcomes that do not depend on the compound being tested. "They provide only one outcome, they are easy to measure, and you can get very high throughput," he said. "This is not true of ADMET screening now. You get outcomes that depend on the compound being tested instead of a single outcome, which slows the process down. It's not high throughput."

Applying the rules

While the drug-likeness effort is still early, it has nevertheless had a profound impact on chemistry. DeWitt said building the models has changed the way ARQL does discovery. "With collaborators, we have generated ADMET data even before doing potency studies," he said. "It has redirected our synthetic effort to avoid known liabilities."

Cerep's BioPrint database contains data from 2,000 marketed drugs, including their performance on 100 in vitro tests of specificity, bioavailability, pharmacokinetics and other properties. The database also contains in vivo data and correlates both in vivo and in vitro data with the compounds' molecular structures.

"This is different from the Rule of 5, which just provides a red or green light," Revah said. "It could not establish a comparison between compounds."

The BioPrint database and models can use structure to predict a molecule's effects not only on its primary target, but also on secondary protein targets potentially responsible for side effects. Revah also said that Cerep now can model certain substeps of bioavailability quite well.

"For example," he said, "GI permeation is becoming better and better, whereas the blood-brain barrier needs work. Also, we cannot do the whole thing. We must add issues like compound stability and accumulation in different tissues."

To tackle toxicity, modelers will need to address components such as exposure of a tissue to a compound and individual genetic variation.

VRTX uses a range of approaches to develop screening guidelines, including historical data, information from its own experiments, and genetic algorithms, neural networks and other types of machine learning. For a given project, the company starts with binary or tertiary classification systems — those that provide yes/no or greater-than/less-than answers — which increases the system's throughput. "This is easier than building systems to predict a precise number, and for most problems, it is more effective," Murcko said.

For example, a guideline that predicts compound solubility based on structure can be very precise, or it can provide a range (i.e., the compound will be more or less soluble than X). Most often, a chemist can make a decision to pursue the compound further based on a range.

See next page

BIOCENTURY

PUBLICATIONS INC.

Subscriber Services

Members of the BioCentury Community have exclusive access to these areas of BioCentury's Web site:

The News Center

Contains *BioCentury Extra*, the concise, intelligent and authoritative newspaper for the industry. Published 3 times most business days. Also provides access to stock quotes and charts and other online resources, and online access to the current week's *BioCentury* and *BioCentury Part II* in PDF format.

The Financial Center

Tools to access, sort and compare more than 8 years of financial data about life science companies in North America and Europe, including more than \$85 billion in fundraising deals.

The BioCentury Archives

A growing library of over 50,000 articles and news briefs published by BioCentury since 1992.

Obtaining Access

Access is via password. BioCentury subscribers who have not yet established their passwords should go to the Home page, click on "Registration/Assistance for Current Subscribers" and follow the 3 easy steps.

Online Communications

Use "Contact Us" at BioCentury Online to reach:

- Editorial: Communicate directly with our News Rooms in the U.S. and Europe.
- Subscriber Services: Establish service, order reprints, request retransmission of an email or fax.
- Conferences & Research Services: Inquire about research, conferences and other special projects.

BioCentury Conference Center

BioCentury organizes 3 major financial meetings and sponsors several other events during the year. See the Conference Center for current meeting activity.

Drug-likeness Rules,
from previous page

VRTX is experimenting with another method called NMR Shapes. "Using nuclear magnetic resonance, you can look at how drug fragments bind to a protein of interest," Murcko said. Compounds based on known drug-like scaffolds are placed in solution with the protein target and subjected to a burst of radiation. The NMR machine reads the resulting spectra, which change depending on whether the compound is binding.

"If nothing binds, it is something of a red flag that the target does not bind small molecules easily," he said.

VRTX's Combidock program to screen large virtual libraries against target crystal structures is also applied to compound selection. The company first creates a library of possible small molecule structures based on a scaffold and three to four side chains. The scaffold could come from the literature or from a competitor's compound, from which VRTX wants to create a patentably distinct molecule, said spokesperson Michael Partridge.

Next, the company runs its REOS (Rapid Elimination of Swill) computer-based filtering process, which weeds out compounds based on drug-likeness. Finally, using Combidock, VRTX docks the remaining compounds computationally to the structure of a protein target and selects the highest scoring compounds.

"With these models, we shoot for being 80% accurate, and we have succeeded in every case to date," said Murcko. "We can use them to eliminate chemicals from the chemist's plate and simplify the task."

He said the company has a series of such filters, each of which generally eliminates about half the compounds under consideration, "so we can whittle down the number enormously in a matter of hours."

But Murcko cautioned that achieving this degree of accuracy using models is no simple task. "There is no cook book," he said. "You need a collaborative mind set, with pharmacokinetics and toxicology people working with modelers, cell biologists and screening people. It is not just a single tool approach."

As more ADMET properties are associated with small molecule structures, companies can use the knowledge during compound optimization. For example, Richard Soll, vice president of chemistry at

'Lipinski said that diversity is not the only thing. Quality criteria exist that are necessary for success.'

— Frederic Revah of Cerep

3-Dimensional Pharmaceuticals Inc. (Yardley, Penn.), said that DDDP has been able to combine its understanding of the structural basis of a compound's selectivity with its understanding of how to influence solubility in designing thrombin inhibitors.

"We introduced more hydrophilic groupings at positions on the molecules that do not interfere with selectivity," he said. "The addition of hydrophilic groups enhances solubility and retention of the compound in plasma."

Such groups can be added at many positions on a molecule and still improve solubility, so choosing where to put them is key. "Our use of crystallography allowed us to home in on the parts of the molecule we should modify to improve pharmacokinetics," Soll said. "We were therefore addressing the ADME issue by knowing where to change the molecule."

Genomics

The influence of genomics on drug-likeness has only begun to be felt. According to Lipinski, the wide variety of protein targets that chemists now must address has slowed progress in defining drug-likeness — a problem that will continue to grow as more targets are validated. Thus, the pressure to define drug-likeness for different targets has only begun.

Some companies have focused on defining drug-likeness for specific target classes. For example, VRTX tracks the characteristics of inhibitors of kinases, its targets of choice. The data gathered provide a good starting point for chemistry programs addressing new kinases, Murcko said.

However, TRPS's Cramer expressed skepticism about focused compound libraries intended for one target class, saying that he is not aware that their efficacy

See next page

Drug Likeness

Painting the metabolic landscape

By Kathryn Calkins
Senior Writer

At least one company believes that biological models have limited utility for predicting metabolism, a big part of the drug-likeness equation. Camitro Corp. has focused on understanding the workings of cytochrome p450 enzymes, which metabolize about 95% of pharmaceutical compounds, according to President and CEO Barry Selick. Specifically, three p450s are responsible: CYP3A4, CY2D6 and CY2C9.

Camitro, a subsidiary of ArQule Inc. (ARQL, Woburn, Mass.), has learned that p450s do not behave in the traditional lock-and-key manner of most enzymes. They recognize far too many substrates to have that kind of specificity, Selick said.

Thus, using biological data to try to predict whether a p450 will metabolize a compound is frustrating. Observing the enzymes' effects on a thousand compounds only produces trends in the compounds' global properties, such as size and charge. It does not help to predict whether certain compound structures will be metabolized and at what rate. Instead, metabolism seems to depend on individual compound features.

Camitro (Menlo Park, Calif.) has turned to quantum chemical computations that allow it to predict the relative metabolic rates among a set of related compounds. "We look at the energetics required to extract a hydrogen from

'Most people think you want a drug that doesn't get metabolized, but that isn't true.'

— Camitro's Barry Selick

every site of metabolism on the molecule," Selick said. Energetics can be calculated on a computer, and the calculations show locations on the molecule where hydrogens are easiest to extract. These positions are the ones the p450 enzymes will metabolize.

The map showing hydrogen extractability over the entire molecule is the metabolic landscape. "It allows us to see the sites that are most labile, and from the numbers of labile sites, you can get an idea of the time it will take for the compound to be metabolized, whether it's three seconds or 30 hours," he said.

When considering a compound series, Camitro makes metabolic landscapes for each of the three p450 enzymes, provided the compounds' global characteristics suggest that they will be metabolized by a p450. "This answers several questions, including the relative rate of metabolism compared with related compounds in a series," Selick said.

The calculation also answers whether compounds bind the metabolic enzymes

and how tightly. "Most people think you want a drug that doesn't get metabolized, but that isn't true. What you want to see is a controlled rate of metabolism, because if they aren't metabolized, many drugs can accumulate and become toxic," Selick said.

This is why tight binding of a drug to a p450 affects drug-drug interactions. Bound enzyme is not free to metabolize other drugs, leaving those substances to accumulate and potentially become toxic, Selick said.

He noted that Prozac binds tightly to CY2D6 and therefore requires a label listing all the compounds for which it is contraindicated, which include any that are metabolized by 2D6. "You could say that Prozac has been successful despite this problem, but I would argue that if you developed a similar compound that did not bind 2D6, you would have a bigger drug," Selick said.

Finally, Camitro's energetics calculations can tell it not only where the compound will be metabolized, but what its metabolites will look like. "We have a database of toxicophores, and we can compare the metabolites to identify potential toxicity problems. We're developing this capability now," Selick said.

Recent progress in studying cytochrome p450s also has been made by Astex Technology Ltd. (Cambridge, U.K.), which last month announced it had solved the crystal structure of these enzymes (see *BioCentury*, Dec. 17, 2001).

Drug-likeness Rules, *from previous page*

has been demonstrated. "Given the choice between a more expensive targeted library with fewer compounds and a less expensive one containing more and more diverse drug-like compounds, I would choose to screen the less expensive one," he said.

Putting it together

Since Lipinski published his rules in the mid-90s, chemists have learned to consider drug-likeness much earlier in compound selection and have begun to

define new ways of predicting what makes a compound drug-like based on both physical and biological properties.

"Drug-likeness has everyone's attention right now, which was not true three or four years ago," said Murcko. "It will get more important. But there's also a lot of hype around now about in silico biology. It's like combichem was in the early 90s, when it was 'The Solution.' Like combichem, this trend will mature. The only thing that gets you drugs is putting all this together in the right way."

This requires a multidisciplinary effort. "You need to bring in the right disciplines to work together," Lipinski said. "You need high throughput people work-

ing with computational people working with absorption, metabolism and pharmaceutical scientists."

And you need data, the gathering of which will remain a challenge for some time. Lipinski believes that big pharmas like PFE have an advantage here because they have been gathering and storing data on compounds for so long. However, if Camitro's experience in metabolism is any indication (see "Painting the Metabolic Landscape", above), there remain undiscovered ways to define ADMET that do not rely on volumes of biological data.

"You have to know what to look for," said Selick.

Predicting drug-likeness

Important discoveries influencing scientific understanding of compound solubility, permeability, active transport, metabolism and toxicity. "External Trends" shows independent events that influenced the field. Key events are described in more detail below. The chart was created in cooperation with Richard Cramer, chief scientific officer of Tripos Inc. (TRPS, St. Louis, Mo.).

| Year | External Trends | Trends in the Field | | | | |
|------|---|---|---|---|---|---|
| | | Solubility | Permeability | Active transport | Metabolism | Toxicity |
| 1900 | | Meyer, Overton describe oil solubility's relation to activity | | | | |
| 1964 | QSAR computing EPA founded | ← Hansch, Leo describe measurement system for relating properties to activity → | | | | |
| 1980 | Caco II cell lines established 3-D models of receptors developed HTS begins | Yalkowsky describes crystal packing problems | Abraham, Raesky discover that H-bonding is critical | | Darvas, Johnson develop ad hoc computer models of metabolism | Enslein, Klopman develop ad hoc computer models of toxicity |
| 1996 | CombiChem comes into wide use Genomics developed | ← Lipinski proposes "Rule of 5" → | | | | |
| | | Palm uses 3-D models to describe polar surface area | | Cruciani uses 3-D models to develop and predict cephalosporin oral activity | Murcko, Kubinyi develop neural net-based drug-likeness models | |

1900

- **Solubility:** Looking at the effects of alcohols on tadpoles, Meyer and Overton discover that biological activity is related to oil solubility, a non-biological property of a molecule. This is widely considered the first observation related to drug-likeness of a compound.

1964

- Hansch and Leo contribute a standard system for measuring a property related to drug activity called Octanol/water log P. Their work also describes the first use of a computer to relate several properties simultaneously to drug activity. This becomes QSAR (quantitative structure activity relationship) computing.

1970

- EPA is founded, bringing toxicity issues to the attention of non-drug company chemists and providing funding for research in the area.
- **Solubility/Permeability:** Hansch, Kubinyi, and MacFarland find that solubility and permeability properties form curves in relation to drug activity. For example, increasing oil solubility reaches a peak in relation to activity, beyond which activity decreases.
- Caco II cell lines are established and techniques developed for measuring permeability and active transport. This provides a way to measure directly the movement of molecules across a gastrointestinal barrier, which is also faster and cheaper than using animal models.

1980

- **Solubility:** Yalkowsky discovers there are many ways a molecule can pack itself into crystals. This makes predicting its solubility difficult, because different crystal forms have different solubilities. This remains a challenge for the industry.
- **Permeability:** Abraham and Raesky discover that although a molecule needs to have a few positions that form hydrogen bonds, having relatively few such features is important for good permeability. This

adds precision to the understanding of permeability.

- **Metabolism:** Darvas and Johnson develop computer programs (e.g., DEREK) that correlate structural fragments of molecules with animal model data on metabolism. The models score each fragment on metabolism but do not attempt to describe the underlying reasons for the scores.
- **Toxicity:** Enslein and Klopman formulate similar models that relate structural fragments to toxicity data.
- 3-D computer models of receptors enable receptor-based drug design.
- High throughput screening begins, creating a demand for more molecules to screen.

1995

- Lipinski proposes the "Rule of 5."
- Combinatorial chemistry comes into wide use, increasing the need to understand drug-likeness.
- **Permeability:** Using 3-D models, Palm describes parameters for the ratio of a molecule's polar vs. non-polar surface area to optimize permeability.
- **Active transport:** Using 3-D models, Cruciani develops a computer program called VolSurf that describes molecules in terms of volume, surface area and their relationship to the molecules surrounding them to predict solubility and permeability. Cruciani creates a separate program, called ALMOND, that predicts whether certain cephalosporin antibiotics will be orally active. This is a classically difficult problem, because cephalosporins depend on active transport to reach their targets.
- **Metabolism:** Murcko and Kubinyi use neural nets to identify structural differences between drugs and non-drug chemicals.
- Genomics begins to have an impact on target validation.