FROM HITS TO LEADS:

Focusing the eyes of medicinal chemistry

Getting back to basic pharmacological principles throws new light on how to truly value a lead.

BY ROGER CROSSLEY

The now-obligatory chart produced whenever senior executives in the pharmaceutical industry present their vision for the future makes grim reading. Despite the expenditure on R&D increasing almost exponentially over the past decade, the number of new drugs coming onto the market has stubbornly refused to increase. Indeed, if one strips out the few peptide therapeutics based on recombinant technology, which largely replaced the same molecules extracted from human plasma, the number of new drugs arguably has fallen. Furthermore, the time to get a drug onto the market seems to have risen from 12 to 14 years over the past 5 years.

As Robin Carr of Astex Technology (*www.astex-technology.com*) and Mike Hann of GlaxoSmithKline reflected recently, a significant reason for this failure is the difficulty of identifying high-quality leads using current approaches, which produce a success rate of only 25% (*1*). A recent survey shows that in the United Kingdom in 1999, some 22% of the money for new technologies was spent on high-throughput screening and 31% on combinatorial chemistry to feed the screening libraries (*2*). Although these figures might reflect the spend-

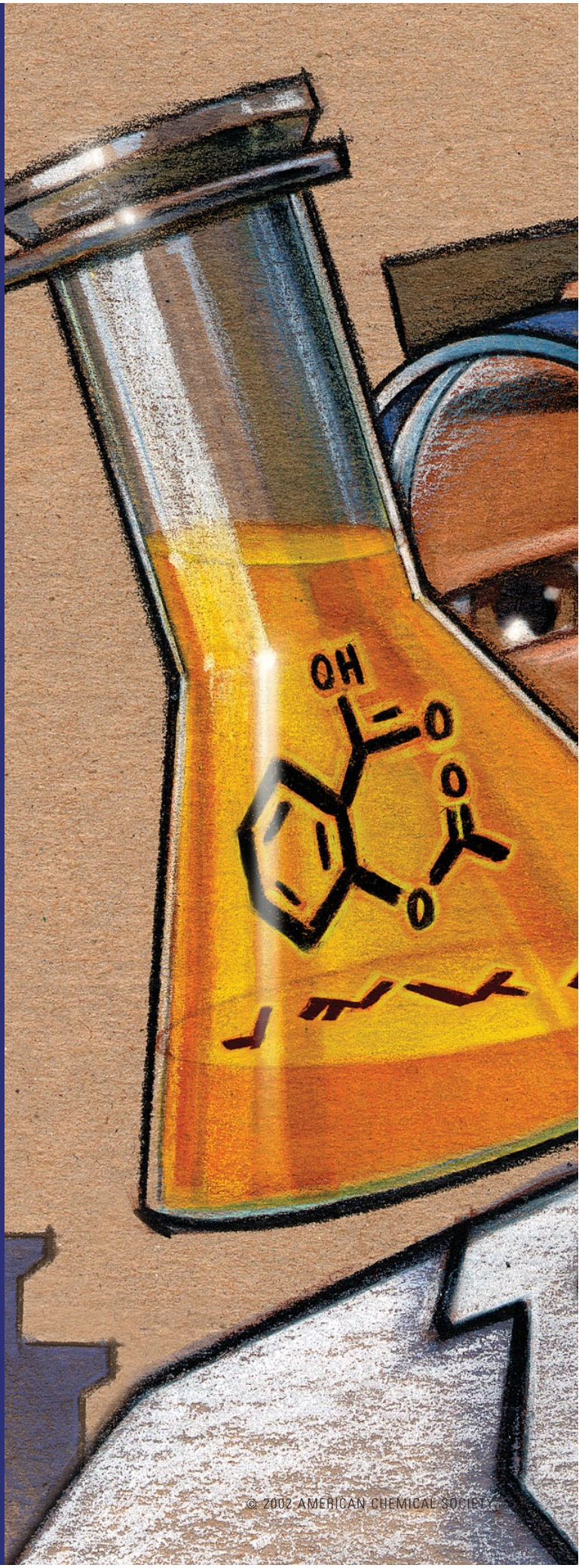
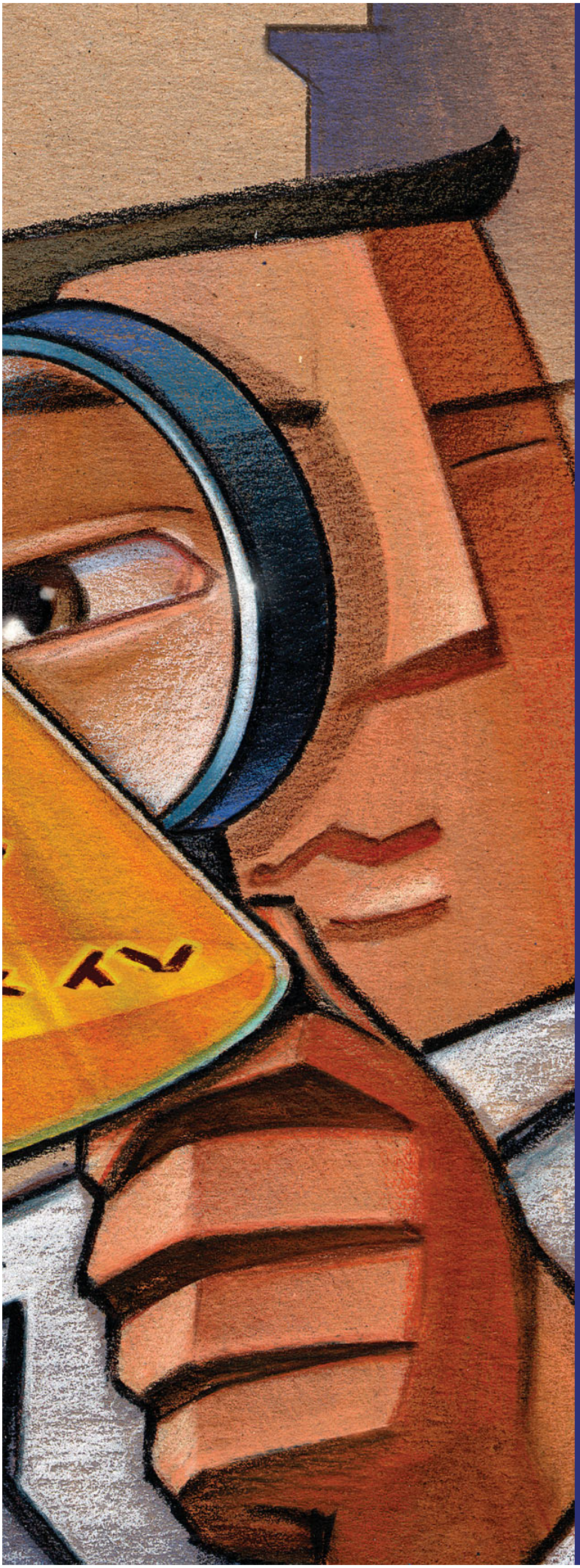


ILLUSTRATION: GEOFF SMITH



ing that was committed a few years before and that the benefits might yet come through, this is looking less likely. Part of the argument that Carr and Hann use is that the wrong compounds were synthesized and screened, and that they were too druglike and not sufficiently leadlike. In turn, the finger was pointed at the Lipinski Rule of 5 in the definition of screening compounds, and they have explored the possibility of using simpler fragments to generate hits.

Faced with the simpler fragments as hits, the medicinal chemists can do what they are good at and elaborate them, by adding weight and complexity, and still keep within Lipinski limits, thus arriving at the final drugs with high potency and selectivity. The problem is that without some extra guidance as to how to optimize simple fragments, such as that provided by NMR or X-ray techniques (*1*), these techniques might eventually generate drugs but are unlikely to do so soon enough. It is quite likely that piperadine would be a low-affinity starting point for aminergic G-protein coupled receptors (GPCRs), but how useful would it be? Where would you go from there?

Consider the problem as a fiendishly difficult game in which the goal, a drug, is separated from the starting point, the screening set, by a series of mazes and obstacles. There are stages along the way, each separated by its own obstacles and mazes, typically represented by hits, leads, optimized leads, and development candidates (or whatever term you wish to use for each stage). Each maze has several points of entry, but all except one lead to a dead end. The object of the game is to get to the goal as quickly as possible and with minimum cost. Then ask whether a single strategy is appropriate to get through all the stages, and of course, the answer is no.

In the end game, where the aim is to get from some advanced lead stage to development candidate through a successful Phase I clinical trial, Lipinski is right to talk about desirable properties for absorption and distribution so as to maximize the chances of success. In other words, he advocates the use of generalized principles to rank the points of entry to this maze in terms of their likely success. But Carr and Hann also have a point in suggesting the benefits, at the very earliest stage and when technology is available (at a cost), of hovering over the maze and directing the compounds through. What other options do we have? Well, we can cheat, and use some prior knowledge of the game to jump ahead.

This article highlights one approach to cheating—the use of focused libraries—in which the added value built into the libraries enables rapid lead generation and optimization. This approach assumes that the closer you start to the finishing line, the faster you are likely to get there. It illustrates the options available for GPCRs and suggests that the way forward is to build in your medicinal chemistry up front.

Screen the right compounds

The ideal hit should be capable of being validated as a lead in a minimum number of steps, and the ideal lead should be capable of optimization to a development candidate, again in the minimum number of steps.

One way to generate a hit molecule with a better chance is to remove from the screen all those that would not be expected to gen-

Advertisers

erate leads. There are a number of approaches to maximize the chances of getting a good hit rate, although the most successful use some knowledge of the target. For a general screening approach, a method of filtering an established compound collection by virtual screening might be worthwhile. Sometimes, preselection can lead to excellent results, such as with PTP1b phosphatase inhibitors reported by Thompson Doman and colleagues at Pharmacia Corp. (Skokie, IL, and Chesterfield, MO) and Northwestern University (Chicago), where high-speed docking combined with high-throughput screening led to a 1700-fold increase in hit rate and, surprisingly, produced more druglike compounds (*3*). With phosphatases, one can opt to use X-ray soaking or docking into a crystal structure or homology model to aid further development, but this option is not available for GPCRs. There are, however, means of preselection that use a more general knowledge of the target. Thus, if it were possible to establish that you wanted to have a basic group in the final molecule, as you might decide if you were to screen an aminergic GPCR for an agonist, it might not make sense to screen molecules without this functional group.

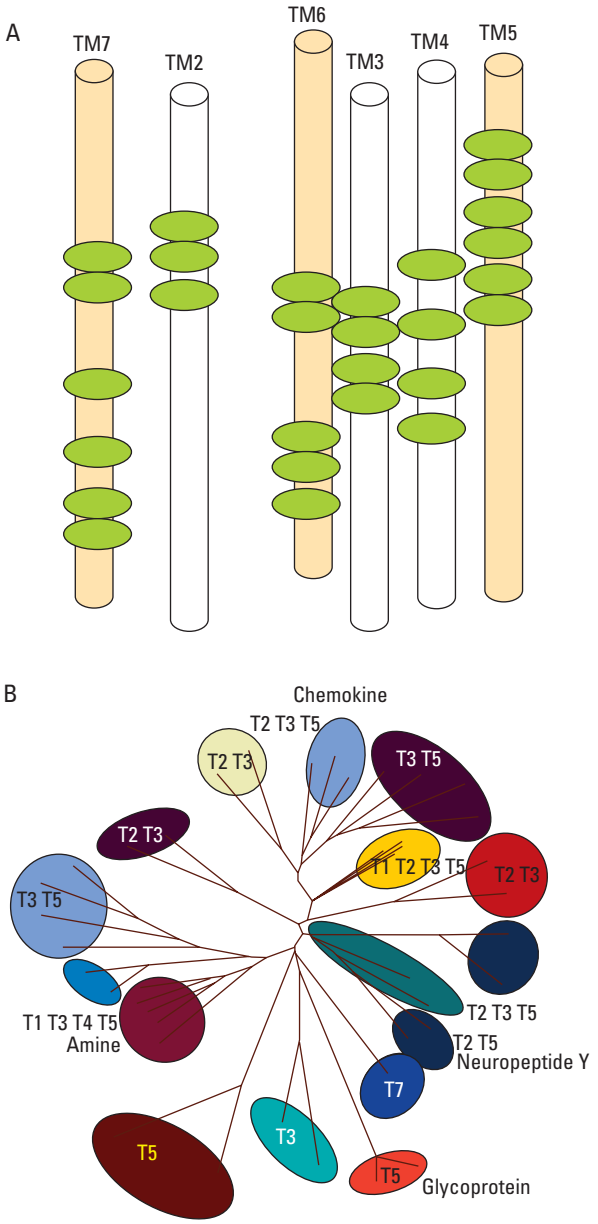


Figure 1. The thematic approach. (A) A model of GPCR transmembrane segments highlights the residues principally involved in the binding of small molecule drugs, which can be grouped together to produce Themes (Tn). (B) Superimposition of Themes onto a phylogenetic tree of GPCRs indicated their distribution across GPCR drug space, enabling a reclassification of receptors according to their ability to recognize drugs. (Courtesy of BioFocus plc.)

Probable screening

Another way to maximize the chance of getting to a good lead is to examine the features of those pharmaceutical compounds that have succeeded before. So-called privileged structure approaches (*4*) can be applied, but on their own, they are essentially unguided attempts to combine fragments that are commonly found in drugs that hit a particular receptor class. One way of increasing their relevance is to use them to produce a pharmacophore model for a particular subset of receptors. This approach was applied to the design of a library based on Ugi chemistry by Jonathon Mason and colleagues at Rhône-Poulenc Rorer (now Aventis) in Collegeville, PA, (*5*) and resulted in a 43% enrichment of relevant molecules. Essentially, what has been done here is to increase the probability of success by working closer to that area of drug space that is relevant to GPCRs and, as long as the approach keeps within the same class of receptors, that is, aminergic GPCRs, it seems to work well (*6*). Another approach to this problem—discrete substructural analysis—has been proposed by Dennis Church at Serono Pharmaceutical Research Institute (Geneva, Switzerland) and has led to efficiency gains of up to 40-fold and has also been applied to library design (*7*).

It is also possible to use more formal statistical treatments in the design of libraries and at the same time start to include some information about the receptor interactions. Structure–profile relationships devised by a team at Cerep (*www.cerep.com*) consider the activity of a molecule in terms of a vector in a space of pharmacological profiles against a panel of 70 receptors, and although the effort is somewhat daunting, it does start to tie the activities of a whole receptor family together (*8*).

Focus on the right compounds

The probabilistic methods described above represent an extension to the fragment-based approaches by the inclusion of extra knowledge about the receptor family. High-tech experimental methods have been replaced by more computational ones for a receptor family to which the experimental approaches are less applicable. We have started to cheat by using some knowledge of our destination. Can this be extended by using more information about the target receptor? As every schoolchild knows, the best way to cheat when you're solving a maze is to start at the end and trace the route back to the beginning. The docking approach certainly adds value here and can be used to sift libraries, both real and virtual, and it also can help to design them from scratch. What else? Is there information in the receptor sequence that alone will enable us to do this?

In work at Novartis (*9*), Edgar Jacoby and colleagues have extended a previous study (*10*) defining consensus binding sites for aminergic GPCRs to link overall drug structure with the sequence similarities of 50 receptors, enabling them to be reclas-

sified in terms of their ability to bind drug molecules.

A separate approach, thematic analysis, at BioFocus (*www.biofocus.com*) has incorporated a fragment-binding approach and has tied this in with sequence information from around 300 GPCRs, leading to a reclassification (*11*) of GPCRs in terms of drug space (Figure 1). The practical applications of this work have been validated in lead optimization, in which it has produced up to 1500-fold increases in potency and up to 270-fold selectivity for hitherto difficult targets (*12*), and in the production of focused libraries for specific subsets of GPCRs (*13*) providing 1–13% hit rates.

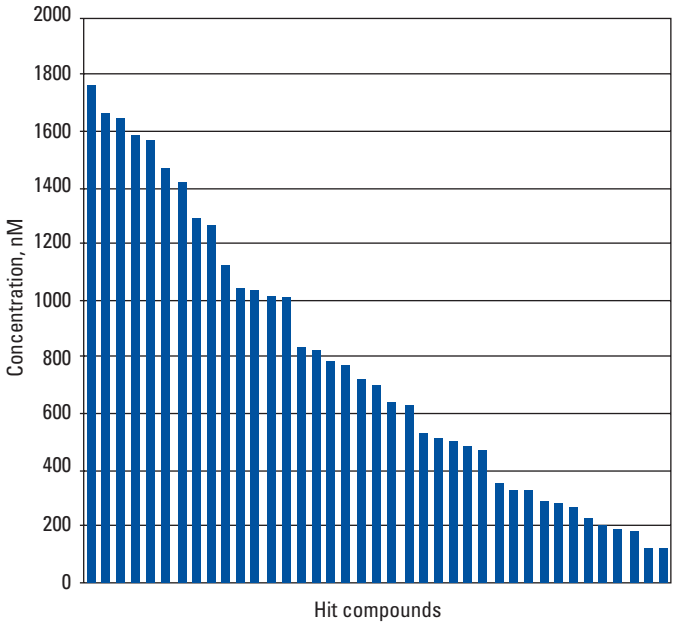


Figure 2. Focused on success. IC_{50} values for compounds in a focused library that was screened against a potential target receptor. The compounds selected for IC_{50} determination came from a 1000-compound focused library and produced 100% inhibition of signal at 10 nM as measured by FLIPR with the natural ligand as agonist. (Courtesy of BioFocus plc.)

An example of how the hits from a focused library demonstrate a depth of structure–activity relationship (SAR) is illustrated in Figure 2. The compounds in the library were designed to target about 40 GPCRs with a related SAR, and the receptor in question is close to the focus of the library. Because the design process for a focused library can accommodate the need to satisfy Lipinski's rules, the compounds are well within the range of established GPCR drugs. The hits demonstrate the general abilities of such libraries to rapidly establish a lead series with a good SAR and with the best compounds within reach of target potencies. Arguably, such screening has produced a fully validated lead series with enough information to establish patent estate and is only a short distance away from producing development candidates.

The final focus

Focused libraries, which bring together knowledge of the target with knowledge of the sorts of molecules that interact with the target family, are a form of combinatorial medicinal chemistry. They rationally incorporate the instinctive feeling that such a molecule *looks like* a kinase inhibitor (or a protease inhibitor or a GPCR ligand) with knowledge of the target. At the same time, they incorporate the variations required to establish initial SAR, future direction, and those properties needed for absorption and distribution. The leads they produce have cheated the early obstacles and have positioned themselves for a rapid optimization and easy passage through the end game. Perhaps gamesmanship rather than cheating represents a highly cost-effective holistic solution to the drug game.

If this is so, then focused libraries may open up the possibility of using medicinal chemistry up front, before the screening stage, to increase the chances of progressing from target to lead series. Certainly, in view of the failure rates in this phase of drug discovery, something is badly needed to stop the losses and to rapidly get into lead series with some future. Such thinking lies behind the recently announced collaboration between Biovitrum (Stockholm, *www.biovitrum.com*) and BioFocus, where, by identifying receptor function and selective lead molecules in parallel by using focused libraries, the time to develop new drug candidates can be considerably shortened. It is anticipated that some of the leads will be developed further by Biovitrum and that the remainder will be provided as out-licensing opportunities where more traditional methods have failed.

The ability of focused libraries to shortcut the game could hasten changes in the way that at least some target families are tackled. For other receptor classes, the use of high-volume screening will continue for some time despite their comparatively poor performance in producing lead series—indeed, there are few alternatives. However, the techniques used to design focused libraries are developing fast, as their value is increasingly recognized, and the inclusion of good medicinal chemistry principles will be critical in the success of these techniques.

References

(1) Carr, R.; Hann, M. *Modern Drug Discov.* **2002,** *5,* 45–48.
(2) *SCRIP*, May 31, 2002, *4,* 2751.
(3) Doman, T. N.; et al. *J. Med. Chem.* **2002,** *45,* 2213–2221.
(4) Patchett, A. A.; Nargund, R. P. Priviledged Structures—An Update. *Ann. Rep. Med. Chem.*; Doherty, A. M., Ed.; Academic Press: San Diego, 2000; Vol. 35, pp 289–298.
(5) Mason, J. S.; et al. *J. Med. Chem.* **1999,** *42,* 3251–3264.
(6) Lamb, M. L.; et al. Iterative Design of Gene-Family Directed Screening Libraries. Presented at the 222nd ACS National Meeting, Chicago, Aug 26–30, 2001.
(7) Domine, D.; Colinge, J; Church, D. High-Throughput Lead Discovery Using Predictive Substructural Analysis. Presented at the 222nd ACS National Meeting, Chicago, Aug 26–30, 2001.
(8) Poulain, R.; et al. *J. Med. Chem.* **2001,** *44*, 3391–3401.
(9) Jacoby, E. *Quant. Struct.-Act. Relat.* **2001,** *20,* 115–123.
(10) Jacoby E.; et al. *Quant. Struct.-Act. Relat.* **1999,** *18,* 561–572.
(11) Slater, M. Reclassification of G-Protein Coupled Receptors Using Thematic Analysis. Presented at the 17th International Symposium on Medicinal Chemistry, Barcelona, Sept 2002.
(12) Crossley, R. GPCR Libraries—Variations on a Theme. Presented at the 7th Annual G-Protein Coupled Receptors Conf., San Diego, Oct 2002.
(13) Crossley, R. Family Activity Relationships and the Use of Logical Modeling in Library Design. Presented at Drug Discovery by Design, Boston, Nov 2001.

Roger Crossley is director of new product development at BioFocus (Sittingbourne, U.K.). Send your comments or questions about this article to mdd@acs.org or to the Editorial Office address on page 3. ■