# Targeted molecular diversity in drug discovery: integration of structure-based design and combinatorial chemistry

Jin Li, Christopher W. Murray, Bohdan Waszkowycz and Stephen C. Young

A powerful new approach emerging in drug discovery research combines computational screening of virtual combinatorial libraries against a therapeutic target and targeted combinatorial library synthesis. This new approach includes positive features from both structure-based design and combinatorial chemistry. It has the potential of producing combinatorial libraries with a high hit rate, and hence accelerates the generation of quality lead compounds. The effectiveness of this novel approach has been shown by the design and synthesis of potent inhibitors for serine and aspartyl proteases.

An important step in drug discovery is the cost-effective identification of lead molecules. In recent years, combinatorial chemical synthesis methods have been developed to the stage where large libraries of compounds can be synthesized and tested with moderate cost and effort[1–3]. In particular, very large libraries of peptides and peptoids have been synthesized, leading to the identification of novel, active compounds[4–7]. As peptide-like molecules often display unsuitable pharmacokinetics for oral dosing, there is considerable interest in extending combinatorial libraries to a wider range of peptoid structures and especially to small organic molecules[8–10].

A combinatorial library of small organic molecules typically consists of a common central template or scaffold onto which different sets of substituents are attached at several substituent positions (Figure 1). The substituents will be derived from available chemical reagents, and the template is either an available reagent, or is easily synthesized, or is formed in the synthesis of library members. Large libraries are necessary in random screening applications because a wide diversity of chemical compounds is needed to ensure that there is an adequate number of high-quality hits. In many cases, the synthesis of full libraries may be problematic because of the enormous numbers of library members. Diversity analysis of the substituents is used to reduce the library to a more manageable size[11–14].

Figure 2 illustrates the basic concepts underlying diversity analysis. Chemical compounds can be classified according to the values of calculated molecular properties (descriptors), such as chemical functionality, shape, physicochemical properties and the pharmacophores a molecule can match[13]. Molecules with similar values of relevant molecular descriptors are likely to have similar biological activities. In

**Jin Li***, **Christopher W. Murray**, **Bohdan Waszkowycz** and **Stephen C. Young**, Proteus Molecular Design Ltd, Beechfield House, Lyme Green Business Park, Macclesfield, Cheshire, UK SK11 0JL. *tel: +44 1625 500555, fax: +44 1625 500666, e-mail: jin@proteus.co.uk

  1359-6446/98/$19.00. PII: S1359-6446(97)01138-0 

practice, it is found that substituents that make up a combinatorial library tend to cluster in molecular descriptor space (Figure 2). If synthetic resource considerations mean that only some of the available substituents can be used in synthesis, the chances of finding good hits will be maximized if substituents encompassing a broad range of molecular descriptor values are preferentially chosen. Similarly, if an active molecule is found in the library, diversity analysis can be used to locate substituents with similar properties to the active substituents, and target or direct subsequent combinatorial synthesis towards molecules more likely to be active.

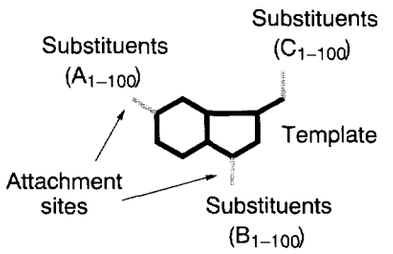


***Figure 1.*** *Schematic illustration of a combinatorial library. Each member of the library consists of a common template or scaffold (black) connected via template attachment sites (grey) to different lists of substituents. In this example, there are 100 possible substituents at three sites, leading to a library of 1 million compounds.*

Practical problems with diversity analysis are that it can still lead to the design of libraries that are too large or in which the molecular property space may be sampled too coarsely or in the wrong areas. To illustrate this, Figure 3 gives the structures of three inhibitors of thrombin, together with their superimposed crystallographically determined binding geometries[15]. It can be seen that they share a common mode of molecular recognition despite having different molecular frameworks. These inhibitors will not be close neighbours in many descriptors of molecular similarity, and, if diversity space is sampled too coarsely or in inappropriate regions, they and their close analogues could be missed.

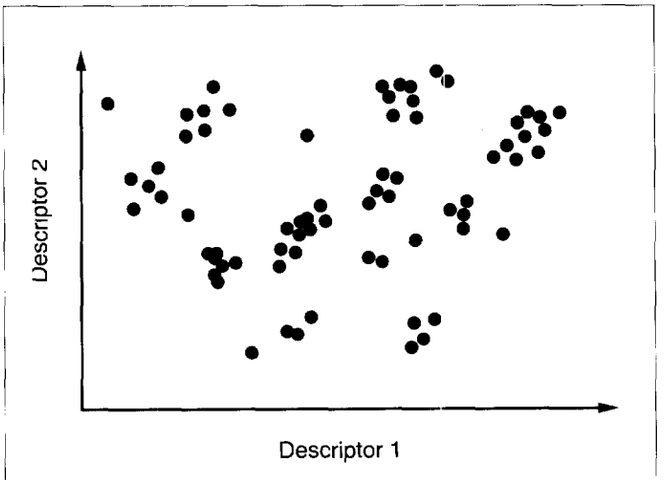


***Figure 2.*** *Schematic illustration of diversity within a combinatorial library. The molecules in the full library are marked as circles according to two computational molecular descriptors. For a molecule to be active, it must lie in a grey area on the graph. The red circles indicate the elements of a diverse sublibrary. The gold circles show how a directed sublibrary can be expanded around an active library member.*

## Structure-based drug design

An alternative paradigm for lead identification is structure-based drug design (SBDD), which is the rational design of molecules based on the 3D structure of the target receptor, usually an enzyme[16,17]. Table 1 gives a selection of enzyme families together with their associated areas of therapeutic interest and the 3D structural information available. The goal of SBDD is to identify or design molecules with 3D complementarity to the target enzyme. A well-known method that has been shown to identify active compounds is DOCK (Ref. 18); this method screens chemical compound databases on the basis of complementarity to the active site. Manual design techniques are also successful – here there is a reliance upon good graphical and interpretative tools such as GRID (Ref. 19). This has spurred a great deal of interest in the automated generation and assessment of molecular designs, as exemplified by the *de novo* design program LUDI (Ref. 20). A major difficulty with automated and manual design is ensuring that designed molecules are relatively easy to synthesize. The problem is exacerbated by the inaccuracies of the computational methods used for the assessment of binding affinity, the prediction of binding modes and the treatment of receptor flexibility. This means that significant synthetic effort can be wasted on weakly active molecules.

Recently, powerful new approaches have emerged that utilize the 3D structure of target enzymes (or enzymes structurally related to the target) to customize the design of combinatorial libraries. Virtual combinatorial libraries are constructed in the computer and are screened computationally for 3D complementarity to the target enzyme[21,22]. The advantage of this is that diversity can be focused to relevant areas of molecular property space. This allows

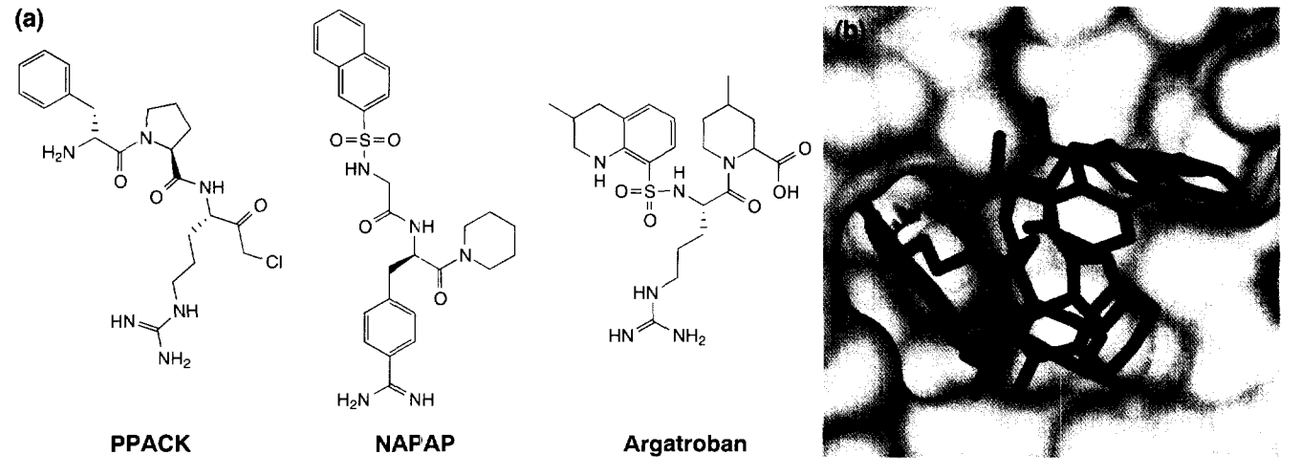


***Figure 3.*** *(a) The chemical structure of three inhibitors of thrombin. (b) The superimposed, crystallographically determined structures of the same three inhibitors. The molecules share a common mode of molecular recognition despite differences in chemical structure. PPACK (green), NAPAP (red), argatroban (blue).*

much smaller libraries to be constructed with the potential to deliver improved hit rates and hence accelerate the process of drug discovery. Another advantage over existing SBDD methods is that designed molecules are chosen to be members of a virtual combinatorial library and are therefore amenable to synthesis. Furthermore, because it is possible to synthesize a considerable number of designs quickly to explore different aspects of molecular recognition, the inherent inaccuracies of the computational methods become less of a liability. Also, the process can efficiently yield QSAR sets to stimulate further design.

This review describes receptor-targeted combinatorial methods and in particular focuses on those that have been applied to real-life drug design applications. The issues raised by the methods are discussed and a perspective on future development is provided.

## Virtual combinatorial libraries

Figure 1 schematically illustrates a virtual combinatorial library. In combinatorial chemistry, the choice of library is dictated by practical synthetic constraints. For example:

- What chemical reagents are available for constructing the library?
- What reactions are amenable to high-throughput synthesis?
- How difficult are the reactions?

In receptor-targeted applications, the choice of virtual library is also constrained by knowledge of the 3D structure of the receptor. The definition of the template or scaffold is an important consideration. A variety of computational tools can be used to identify templates that make good contacts with the enzyme. The most reliable method is to define some portion of a known inhibitor as the template. This is especially useful if the crystal structure for the enzyme (or a member of the enzyme family) with the inhibitor (or a related inhibitor) is available. Otherwise, molecular modelling can be used to suggest more novel templates. Important criteria in the choice of templates include:

- The template should make favourable contacts with the receptor (not always possible).
- Substituent attachment points should target binding pockets of interest.
- The mode of binding of an ideally functionalized template should be stable when subjected to molecular dynamics.
- The virtual library associated with the template should be capable of delivering relevant diversity.

Some of these issues are illustrated schematically in Figure 4. As is generally true within combinatorial chemistry, the choice of template can to a large extent determine the success of the library.

**Table 1. A representative list of enzyme families and the related targets[a]**

| Families | Therapeutic areas | Available 3D structures |
|---|---|---|
| Serine proteases | Thrombosis | Thrombin, Factor Xa, Factor VIIa, etc. |
| | Emphysema | Elastase |
| | Cancer | Urokinase |
| | Hepatitis C | HCV protease |
| Aspartic proteases | HIV treatment | HIV protease |
| | Hypertension | Renin |
| | Cancer | Cathepsin D |
| Cysteine proteases | Inflammation | ICE |
| | Osteoporosis | Cathepsin K |
| | Cancer | Cathepsin B |
| Metalloproteases | Cancer | Various MMPs |
| Tyrosine kinases | Cancer | FGFR1 |
| MAP kinases | Alzheimer's disease | ERK2 |
| | Cancer | P38 |

[a]Targets homologous to available 3D structures are excluded, as well as many enzymes that do not belong to well-known homologous families (e.g. thymidylate synthase, purine nucleoside phosphorylase).
ERK, extracellular-signal-regulated kinase; FGFR, fibroblast growth factor receptor; HCV, hepatitis C virus; ICE, interleukin 1β-converting enzyme; MAP kinases, mitogen-activated protein kinases; MMP, matrix metalloprotease.

## The problem of combinatorial explosion

Combinatorial methods give rise to an explosion in the number of library members as more reactions, and more chemicals in those reactions, are used. For example, there are about 9,000 primary amines in the *Available Chemicals Directory*[23] and 4,000 carboxylic acids. If three reactions are used, two involving amines and one using carboxylic acids, the full virtual library will contain $3 \times 10^{11}$ molecules. Even if only 10% of each list of starting materials is relevant, the virtual library will still contain $3 \times 10^{8}$ molecules. To put this in perspective, if a program on a powerful computer (or set of computers) spends 1 s assessing each molecule in this smaller library, it will still take over 10 years to screen it. Furthermore, it is very difficult to see how current computational methods could reliably reject one molecule over another, if only 1 s is available for the assessment. This is especially true, because the analysis has already assumed that a generous 90% of the starting materials can be rejected on the basis of simple considerations such as molecular weight, duplicates or synthetic chemistry constraints. Clearly some approximations are required, but what should they be?

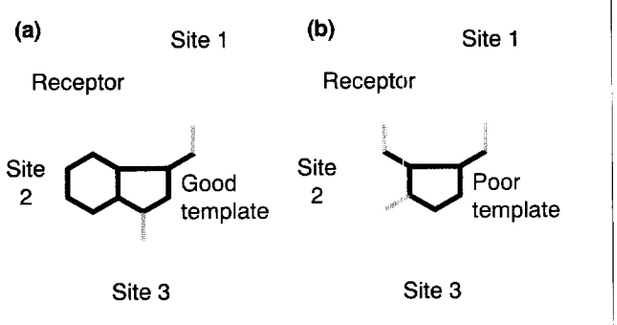


***Figure 4.*** *(a) The well-positioned template binds well to site 2 of the receptor, and the substituent attachment sites (in grey) are well situated to allow substituents to enter easily into sites 1 and 3 of the receptor and make favourable interactions. (b) The poorly positioned template does not make favourable interactions with the receptor and does not have its substituent attachment sites well positioned with respect to the receptor. Substituents attached to this template will not be able to enter site 3 easily.*

Two groups have applied and experimentally validated a receptor-based screening of virtual libraries[21,22], and both have adopted essentially the same approximation: that the template is assumed to be fixed in the active site and each substituent position is evaluated in isolation. In this case, it is possible to perform a reasonably detailed conformational search of each substituent. If there are substantial correlations between substituents attached at one site and substituents attached to a different site, these cannot be taken into account properly until a later stage. Using this approximation, the scaling of the computational cost is given by the sum of the sizes of the substituent lists rather than by their multiplication. In the earlier example, this means that about 20 s can be spent analysing each substituent in a 12 h run instead of 1 s per library member in a 10-year run. It is worth adding that it is the separate treatment of the substituent lists, rather than the fixing of the template, that leads to this time saving. It is possible to attach the template to each substituent and screen the template substituent molecule against the receptor without fixing the template position. This is an important strategy when the template makes little specific contact with the receptor. An alternative strategy might be to concentrate entirely on the substituents without considering the template but, as yet, there are no practical demonstrations of this.

### Selection of synthesis candidates

In receptor-based screening of virtual libraries, those substituents that cannot fit into the targeted region of the active site are rejected. Those that pass this test will comprise lists of potential candidates for inclusion in a synthesis program. However, the lists may still be too large for full combinatorial synthesis. The following criteria have been used to diminish the substituent lists:

- Empirical estimates of binding affinity;
- Forcefield scores;
- Diversity analysis;
- Practical synthetic chemistry considerations (cost, feasibility).

Accurate scoring of substituents is a major problem, which is why it is important to use diversity analysis to sample 2D and 3D chemical structures. In the PRO_SELECT program (Ref. 21; S.C. Young *et al.*, unpublished), two types of scoring are used. One uses a molecular mechanics forcefield to assess the strain on a substituent in the conformation predicted by the conformational searching algorithm. Using molecular mechanics in this way to score substituents is only an approximation, especially without a good treatment of solvation. The second scoring is based on an empirical function that has been calibrated to give the experimental binding affinity for a large database of 82 ligand–receptor crystallographic complexes[24,25]. The scoring function is accurate to about 1.5 orders of magnitude in binding affinity for the training set and related test sets of complexes. Given a putative binding mode for the complex, five simple, physically based terms are calculated to give an estimate of the binding affinity:

$$\Delta G_{\text{binding}} = \Delta G_0 + \Delta G_{\text{H-bond}} + \Delta G_{\text{metal}} - \Delta G_{\text{lipo}} + \Delta G_{\text{flex}}$$

These terms are, respectively, a constant term, a weighted count on hydrogen bonds (and their quality), a weighted count on contacts with metals, an estimate of the lipophilic contact energy, and a term to punish ligand flexibility. The scoring function implicitly takes into account solvation and entropy effects but is clearly a severe approximation to the complicated molecular recognition process. However, the quality of the scoring function can be improved in response to assay data on synthesized molecules.

Once the substituent lists have been reduced to a manageable number, it is possible to enumerate the full library (or a selected portion of it) and subject individual library members to further computational screening (Ref. 21; S.C. Young *et al.*, unpublished).

### Applications and results

The following two examples apply a receptor-based screening of virtual combinatorial libraries and give experimental validation of the approach. Methods that target diversity using QSAR information will be discussed below.

#### *Serine protease inhibition*

By applying the PRO_SELECT program to the inhibition of thrombin and other serine proteases, three templates have been explored in an initial validation exercise of the receptor-targeted method.

The first template and its position in the active site were based on the crystallographic structure of PPACK (D-Phe-Pro-Arg-$CH_2Cl$) in thrombin[15]. The structure and the key interactions are shown schematically in Figure 5. In this example, reversible inhibitors were designed based around proline as the central template. Primary amines were to be attached to the C-terminus of proline and it was required that these amines also form hydrogen bonds in the S1 subsite (see Figure 5). Carboxylic acids or sulphonic acids were to be attached to the N-terminus of proline and it was required that these substituents form contacts with the hydrophobic D pocket of thrombin.

The full technical details of the working of PRO_SELECT are given elsewhere[21]. In brief, the definition of a template and the design criteria define a virtual library, and then a 3D database searching program is used to identify appropriate substituents at each position. The virtual library can be extended through the specification of virtual molecular transformations that correspond to synthetically simple chemical steps on available chemicals. Substituents are then screened thoroughly against the receptor using a fast conformational searching strategy. Substituent conformations that pass this screen are minimized in the receptor using a forcefield and then empirically scored, as described earlier. More recent versions of PRO_SELECT allow direct minimization of the empirical scoring function, taking into account substituent and template flexibility.

Table 2 describes the efficiency of the screening process for one of the virtual libraries constructed with the proline template. It can be seen how the use of 3D structural constraints rapidly reduced the size of the library in this case. The final library was clustered on 2D and 3D diversity and high scoring representatives of different clusters were put forward for synthesis. Table 3 shows a summary of results from the proline template for inhibition against thrombin. The method yields good hit rates and the SAR from these

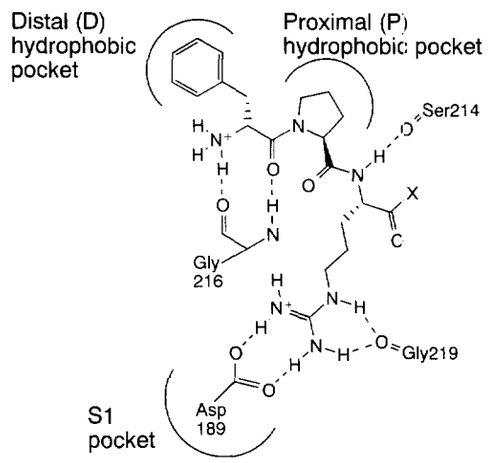


***Figure 5.*** *Schematic illustration of the binding of PPACK (D-Phe-Pro-Arg-$CH_2Cl$) to thrombin. Note that PPACK is covalently attached to thrombin at X.*

results could have been used to target related chemistries to yield improved activity and selectivity profiles.

The next stage in validating receptor-targeted diversity was the design of libraries based on novel templates (B. Waszkowycz *et al.*, unpublished). This is important because the generation of novel chemical entities is the goal of most drug discovery programmes. Two templates, which are novel in the context of serine-protease inhibition, have been explored. The first (Figure 6a) was based around diketopiperazines (DKPs). This template occupies part of the P pocket and forms a hydrogen-bonding network (β-sheet) with Gly216 (Figure 5). The $R_1$ substituent binds to the D pocket the $R_2$ substituent is targeted towards the S1 subsite and the $R_3$ substituent explores the P pocket. The other novel template was an imidazolinone (Figure 6b), which explores a similar region of the active site forming a β-sheet with Gly216. ($R_1$, $R_2$ and $R_3$ substituents are aimed at the D, S1 and P sites, respectively.) The templates were identified by searching reaction databases for heterocycles that could form the desired β-sheet, had suitable substituent positions, and were relatively easy to synthesize. Subsequent modelling of candidate templates with idealized substituents identified templates that retained good interactions with the enzyme during simulation.

There were concerns that both these templates were too polar for thrombin, but it was felt that they would be good for other serine proteases and that they were particularly easy to synthesize. Table 3 shows an overview of the results for the two templates against trypsin. The results against thrombin were poor, although a few DKP molecules had activity in the low micromolar range. Only a small number of compounds were made for each template, yet in both cases micromolar hits were identified and, in the case of the DKP template, a submicromolar compound was discovered. This molecule was subsequently crystallized in trypsin and shown to have the same binding mode to that predicted by the modelling. These results demonstrate that active molecules can be located around novel templates, but highlight the need to identify good templates. Nevertheless, better affinities would have probably resulted if more molecules had been synthesized.

### *Aspartic protease inhibition*

Kick *et al.*[22] have reported the application of receptor-targeted diversity to the inhibition of cathepsin D. Figure 7 gives the virtual library used in this work. It is based around the crystal structure for pepstatin[26] and known SAR on statine-like inhibitors. The template shows reliability in both its position and contribution to activity, and is comparable to the proline template above.

Initial 2D database screening for suitable acylating agents and amines led to lists of 1,900 and 700 compounds,

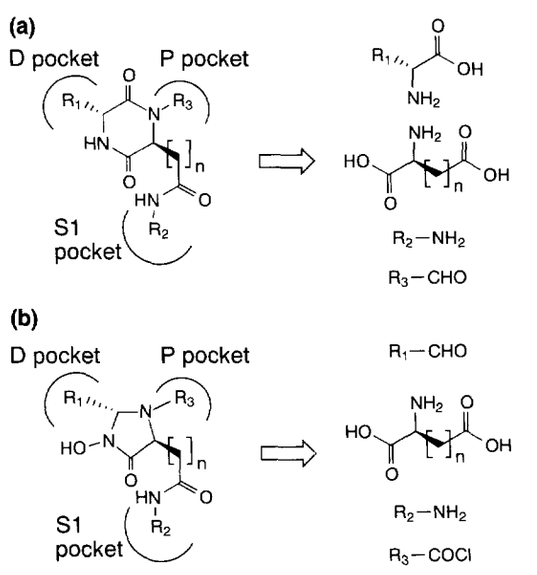


***Figure 6.*** *The novel templates used for serine protease inhibition. (a) The diketopiperazine library and (b) the imidazolinone library, with their associated starting materials illustrated in a retrosynthetic analysis.*

**Table 2. Number of molecules remaining after application of successive screens to substituent lists A and B**

| Filtering stage | No. of substituents at A position[a] | No. of substituents at B position[b] | Size of full library[c] |
|---|---|---|---|
| 2D database screen | 8,803 | 4,262 | 37,518,386 |
| 3D database screen | 437 | 894 | 390,678 |
| PRO_SELECT receptor screen | 145 | 144 | 20,880 |
| Binding affinity filter | 81 | 65 | 5,265 |
| Strain energy filter | 71 | 53 | 3,763 |
| Selected synthesis candidates | 8 | 9 | 72 |

[a]Bisamines targeted at the S1 subsite and attached to the C-terminus of proline.
[b]Carboxylic acids targeted at the D pocket and attached to N-terminus of proline.
[c]Corresponding theoretical virtual library.

respectively, and to a full virtual library containing over 1 billion compounds. Ten reagents at each position were chosen using receptor-based combinatorial screening. The methodology used a rule-based conformational search algorithm to identify substituent conformations not clashing with the receptor. Since the $R_1$ and $R_2$ positions are quite close, $R_1$ conformations that significantly encroached the $R_2$ space were removed and vice versa. Substituent conformers were minimized and scored using a forcefield grid[27]. The substituents were clustered on the grounds of diversity, and the ten best-scoring substituents in different clusters were chosen as synthesis candidates. A thousand molecules were made in the library.

Kick *et al.*[22] compared the performance of this library against a diverse library chosen from the original list without regard to the 3D structure of the enzyme. The results are shown in Table 4. It can be seen that the targeted library has a substantially better hit rate, especially in the high potency range. A highly directed library based on the SAR from the receptor-targeted library was also constructed. This contained only 39 compounds, of which seven inhibited at 100 nM. The most active compound was in the low nanomolar range. This follow-up library indicates how the SAR from the first library, and a knowledge of the binding site interactions, can be used to generate very potent compounds.

## Discussion

The two examples above illustrate the power of receptor-targeted combinatorial approaches and demonstrate the general applicability of these methods. Kick *et al.*[22] used automated synthetic methods to synthesize full sublibraries and made an important comparison with a conventional diverse library. The work demonstrates the improved hit rates that receptor-targeted methods provide. In the serine protease example, molecules were synthesized individually. Although this approach has the disadvantage that relatively few compounds can be made, it has the important advantage that a multiplicity of synthetic protocols can be employed so that diversity in the original virtual library is not hampered by the need to conform to a rigid synthetic protocol. In practice, manual and automated synthetic methods are best used in a complementary manner. The serine protease example is also interesting because it examines the use of novel templates. The results are successful, but indicate the difficulty of designing novel templates. The two templates presented do inhibit serine proteases but are not particularly good for thrombin, which was the main target of the design. Current research is focusing on strategies for the identification of good templates.

**Table 3. Summary of biological activities of molecules synthesized based on different templates**

| | Proline[a] | DKP[b] | Imidazolinone[b] |
|---|---|---|---|
| Number of compounds synthesized | 34 | 25 | 11 |
| $K_i$ <100 μM | 27 | 15 | 3 |
| $K$ <10 μM | 18 | 7 | 1 |
| $K_i$ <1 μM | 9 | 1 | 0 |
| Best $K_i$ (nM) | 41 | 800 | 4,000 |

[a]Activity against thrombin.
[b]Activity against trypsin.
DKP, diketopiperazine.

**Table 4. Results of biological activities against cathepsin D for the receptor-targeted library and the diverse libraries**

| | Receptor-targeted library | Diverse library |
|---|---|---|
| Number of compounds synthesized | 1,000 | 1,000 |
| $IC_{50}$ <1 μM | 67 | 26 |
| $IC_{50}$ <0.33 μM | 23 | 3 |
| $IC_{50}$ <0.1 μM | 7 | 1 |
| Best $K_i$ (nM) | 73 | 356 |

This review has concentrated on methods that explicitly use the structural information on the enzyme to screen virtual libraries computationally and that have been validated in realistic applications. A related set of methods has been used to focus combinatorial libraries using only QSAR information. Sheridan and Kearsley[28] examined the use of a genetic algorithm (GA) to analyse a virtual library of peptoids using QSAR information to control the analysis. In separate works, Weber *et al.*[29] and Singh *et al.*[30] have applied a GA to direct the synthesis of sublibraries of a large virtual library. Here the GA uses the activities of compounds in previous sublibraries to define which compounds should be synthesized in the next sublibrary. Weber's group have demonstrated progressive improvement in inhibitory activity against thrombin, using a library formed from the Ugi reaction, and Singh *et al.* have showed the potential of this approach by optimizing the substrate potency of hexapeptides against stromelysin, a metalloprotease. An article by Salemme *et al.*[31] describes the use of QSAR to direct combinatorial libraries iteratively. They also mention the use of database searching, employing receptor-derived pharmacophores to select substituents. There is no detailed description of their approach or data on their method, but the article does indicate the interest and potential of receptor-targeted combinatorial libraries. A brief description of structure-based library design has been written by Martin *et al.*[13] while, Zheng and Kyle[32] have described the use of the simulation method – multiple copy simultaneous sampling – to dock peptides into receptors, and have noted its potential in combinatorial design. It is expected that more papers on theoretical approaches to receptor-targeted library design will appear in the future and that there will be more practical applications of these methods.

The use of QSAR on a target can be regarded as complementary to the use of 3D structural information. QSAR provides knowledge of which interactions are important to activity and can be used to improve the scoring function for future rounds of synthesis and design. One of the reasons receptor-targeted combinatorial chemistry is attractive as an accelerated method of lead discovery is that it can produce QSAR sets quickly. In the same way, crystallographic analysis of molecules found by combinatorial synthesis can be used in future cycles of structure-based design and/or combinatorial chemistry[31,33].

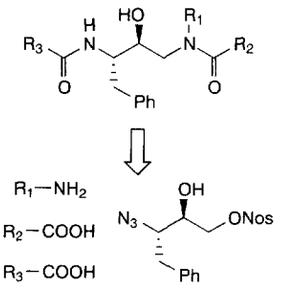


***Figure 7.*** *The library and associated starting materials, illustrated in a retrosynthetic analysis, used for the inhibition of cathepsin D.*

This review has considered the computational screening of virtual libraries against enzyme targets using the 3D structural information on the enzyme. The method has been applied successfully to serine proteases and aspartic proteases. More examples of the method are needed to establish fully the usefulness of the approach; however, we believe that it has great potential in the discovery of potent small organic molecule inhibitors against a wide variety of enzymes.

## REFERENCES


1 Gallop, M.A. *et al.* (1994) *J. Med. Chem.* 37, 1233–1251
2 Terret, N.K. *et al.* (1995) *Tetrahedron* 51, 8135–8173
3 Thompson, L.A. and Ellman, J.A. (1996) *Chem. Rev.* 96, 555–600
4 Koppel, G. *et al.* (1995) *Chem. Biol.* 2, 483–487
5 Murphy, M.M. *et al.* (1995) *J. Am. Chem. Soc.* 117, 7029–7030
6 Wang, G.T. *et al.* (1995) *J. Med. Chem.* 38, 2995–3002
7 Pirrung, M.C., Chau, J-L. and Chen, J. (1995) *Chem. Biol.* 2, 621–626
8 Balkenkohl, F. *et al.* (1996) *Angew. Chem., Int. Ed. Engl.* 35, 2288–2337
9 DeWitt, S.H. and Czarnik, A.W. (1996) *Acc. Chem. Res.* 29, 114–122
10 Carell, T. *et al.* (1995) *Chem. Biol.* 2, 171–183
11 Martin, E.J. *et al.* (1995) *J. Med. Chem.* 38, 1431–1436
12 Brown, R.D. and Martin, Y.C. (1996) *J. Chem. Inf. Comput. Sci.* 36, 572–584
13 Martin, E.J. *et al.* (1997) in *Reviews in Computational Chemistry* (Vol. 10) (Lipkowitz, K.B. and Boyd, D.B., eds), pp. 75–100, VCH
14 Ashton, M.J., Jaye, M.C. and Mason, J.S. (1996) *Drug Discovery Today* 1, 71–78
15 Banner, D.W. and Hadvary, P. (1991) *J. Biol. Chem.* 266, 20085–20093
16 Kuntz, I.D. (1992) *Science* 257, 1078–1083
17 Bohacek, R.S., McMartin, C. and Guida, W.C. (1996) *Med. Res. Rev.* 16, 3–50
18 Desjarlais, R.L. *et al.* (1988) *J. Med. Chem.* 31, 722–729
19 Goodford, P.J. (1985) *J. Med. Chem.* 28, 849–857
20 Böhm, H-J. (1992) *J. Comput.-Aided Mol. Design* 6, 61–78
21 Murray, C.W. *et al.* (1997) *J. Comput.-Aided Mol. Design* 11, 193–207
22 Kick, E.K. *et al.* (1997) *Chem. Biol.* 4, 297–307
23 Available Chemicals Directory (1995) MDL Information Systems Inc.
24 Eldridge, M.D. *et al.* (1997) *J. Comput.-Aided Mol. Design* 11, 425–445
25 Böhm, H-J. (1994) *J. Comput.-Aided Mol. Design* 8, 243–256
26 Baldwin, E.T. *et al.* (1993) *Proc. Natl. Acad. Sci. U. S. A.* 90, 6796–6800
27 Meng, E.C., Shoichet, B.K. and Kuntz, I.D. (1996) *J. Comput. Chem.* 13, 505–524
28 Sheridan, R.P. and Kearsley, S.K. (1995) *J. Chem. Inf. Comput. Sci.* 35, 310–320
29 Weber, L. *et al.* (1995) *Angew. Chem., Int. Ed. Engl.* 34, 2280–2282
30 Singh, J. *et al.* (1996) *J. Am. Chem. Soc.* 118, 1669–1676
31 Salemme, F.R., Spurlino, J. and Bone, R. (1997) *Structure* 5, 319–324
32 Zheng, Q. and Kyle, D.J. (1997) *Drug Discovery Today* 2, 229–234
33 Rockwell, A. *et al.* (1996) *J. Am. Chem. Soc.* 118, 10337–10338