ELSEVIER

Journal of Molecular Graphics and Modelling 20 (2002) 463–468

Journal of Molecular Graphics and Modelling

www.elsevier.com/locate/JMGM

Structure-based combinatorial library design: methodologies and applications

Mary Pat Beavers*, Xin Chen

Computer Assisted Drug Discovery, R.W. Johnson Pharmaceutical Research Institute, 1000 Route 202, Raritan, NJ 08869, USA

Abstract

Rational design of small focused libraries that are biased toward specific therapeutic targets is currently at the forefront of combinatorial library design. Various structure-based design strategies can be implemented in focused library design when the 3D structure of the target is available through X-ray or NMR determination. This review discusses the major methods and programs specifically developed for the purpose of designing combinatorial libraries under the constraint of the binding site of a biological target, with emphasis on their advantages and disadvantages. Examples of the successful application of these methodologies are highlighted, demonstrating their performances within the practical drug discovery process. © 2002 Published by Elsevier Science Inc.

Keywords: Small focused libraries; Medicinal chemistry; X-ray crystallography

1. Introduction

The advent of combinatorial chemistry [1,2] and high throughput screening [3,4] has dramatically changed the drug discovery process in the pharmaceutical industry, making it possible to synthesize and test a large number of chemical compounds in a short period of time. Currently, hundreds of thousands of compounds are prepared and screened on a regular basis, whereas this number was previously limited to a few hundred compounds. This increase in compound flow in the drug discovery cycle was expected to significantly reduce the time required for finding an initial hit and turning it into a pre-clinical drug candidate. However, although some lead compounds have been identified by this approach, [5,6] people have gradually realized that a sheer increase in the number of compounds produced for screening would not automatically yield good and new leads. Actually, it has been observed that the hit rates obtained from screening general unbiased libraries are usually much lower than hit rates obtained with traditional medicinal chemistry approaches. Recently, with an aim toward increasing hit rates, the emphasis in computational and medicinal chemistry has shifted toward the rational design of small focused libraries that are biased toward one or several specific therapeutic targets.

Focused library design [7] can be conducted in several ways, depending on the initially available information about the therapeutic target, such as lead compound SAR, binding site interactions, etc. Whenever the 3D structure of the therapeutic target has been determined, either by X-ray crystallography or by NMR, structure-based combinatorial library design becomes an excellent tool to use in combination with other chemistry tools in the initial phase of a drug discovery program. Traditional structure-based strategies for drug design and discovery [8–10] have been widely used in the pharmaceutical industry for many years. The earliest emphasis in structure-based design and docking strategies involved searching corporate or commercial databases against a pre-defined binding site of a protein target. Docking methods implemented throughout academic and corporate laboratories usually resulted in low hit rates averaging around 2%, although this number is now approaching 10%, given the progress of new methodology development in recent years [11,12].

Structure-based strategies will have the added advantage of introducing valuable information about the target structure into the combinatorial library design process and may significantly increase the hit rate in the final designed library. Not only are the compounds designed inside a specific binding site, but they are also synthetically feasible through combinatorial or parallel synthesis. This review will highlight the current computational strategies for structure-based combinatorial library design.

2. Methodologies

From the computational standpoint of view, structure-based combinatorial library design is simply an extension of

* Corresponding author.

1093-3263/02/$ – see front matter © 2002 Published by Elsevier Science Inc.
PII: S1093-3263(01)00147-4

traditional structure-based drug design, with the design and screening interest shifted from individual compounds to libraries of compounds that can be easily synthesized by combinatorial chemistry or parallel synthesis. Consequently, all the major computational strategies that have been specifically developed for structure-based library design originate from the field of either molecule docking [13,14] or de novo ligand design [15,16]. These strategies can be generally divided into three classes.

The first strategy is conceptually straightforward. All the possible members of a combinatorial library are first enumerated, according to the available reagents and the established synthetic chemistries. Individual members are then separately docked into the binding site of a receptor. Finally, an optimal sub-library is selected for synthesis, based on the ranking of their docking scores and/or diversity measures. Software for fast library enumeration has been developed, including CombiLibMaker in Sybyl [17], Analog_Builder in Cerius2 [18], and the QuaSAR-CombiGen module available in MOE [19], to name a few. Most of these programs can easily generate all of the 2D or 3D structures for a combinatorial library containing millions of compounds, using either fragment-based or reaction-based schemes. Other tools within these software packages are also available for decreasing the size of a virtual library prior to docking. For example, a library enumerated through CombiLibMaker can subsequently be analyzed with diverse solutions (available in Sybyl [17]) to provide a sub-library that adequately samples chemical space. QuaSAR-CombiDesign is another unique combinatorial library design tool available in MOE [19] that provides a non-enumerative method for combinatorial library generation, and can, e.g. test against rule of five filters using statistical sampling techniques during library creation, creating smaller sub-libraries with user-defined property ranges. In principle, the docking step that follows library creation can be conducted using any of the available docking programs like DOCK [20,21] or FlexX© [22], while the diversity selection can be performed using software available from Daylight [23], Tripos [17] (diverse solutions), or BCI [24]. Recently, Diller and Merz [25] reported their study entitled "high throughput docking for library design and library prioritization", providing a good example of this step-wise approach for library construction. With their approach, the flexibility of ligands can be efficiently handled by pre-generating a set of initial conformations, which are then flexibly docked and energy minimized in the binding site of a receptor. They demonstrated that their methodology could generate a ligand binding mode within 2.0 Å of the observed one for nearly 90% of 103 test complexes from the Protein Data Bank, with a speed of less than 5 s per molecule. This implies the capability of processing hundreds or thousands of compounds per week. Generally speaking, this strategy has the advantage of generality and flexibility, and each step can be repeated whenever new algorithms become available. For example, if a new conformational search algorithm is developed, the database can be re-built and used for docking against many different protein targets, making it a general database for electronic screening. Currently, the major limitation of this strategy is still the computational speed. Due to the inherent combinatorial character, most of the combinatorial libraries can generate over millions of possible compounds, easily exceeding the current capability of molecular docking.

The second strategy is essentially a 'divide-and-conquer' approach. With this strategy, all of the product structures in a combinatorial library are viewed as having variable substituents attached through one or multiple sites on a common template. The template is first docked into the binding site and only the top-scoring poses are saved for the further consideration. Individual substituents are then independently attached onto each pose of the template, to assess which substituents can fit well into the binding site. Only those combinations of top-scoring substituents are further considered and scored to identify the whole product structures that can dock really well into the binding site. Many programs that have been specifically developed for structure-based library design follow this strategy, including PRO_SELECT [26], CombiBUILD [27,28], CombiDOCK [29], DREAM++ [30], and FlexX© [31]. This 'build-up' strategy has the obvious advantage of avoiding the combinatorial explosion problem introduced by combinatorial chemistry itself, making it possible to handle over millions of compounds. It also makes the selection of reagents, which are directly related to the variable substituents in most cases, relatively straightforward. However, with this approach, the correct docking of the template becomes extremely critical. If the template is docked incorrectly, the subsequent substituent docking becomes meaningless. Unless the template is large and rigid and/or has strong specific interactions with the binding site, such as covalent bonding, a large number of samplings are still needed to cover all of the promising template binding poses. This will tremendously increase the computational burden of this approach. Moreover, this strategy is based on the pre-assumption that the contributions of the substituents at different attachment sites are additive. This may not be the case under all circumstances, especially with the combinations of so-called 'weak' substituents that may actually have strong binding affinities when taken together as a whole structure.

The third strategy is also a kind of 'divide-and-conquer' approach. However, with this strategy, the docking of the substituents is considered first, followed by a linking together of the docked substituents to satisfy the constraints of combinatorial chemistry. Recently, Leach et al. [32] described an elegant way to address the problem of synthetic feasibility inherent within this strategy. They suggested using the template as a 2D query to search a structure library (such as the ACD database) to find suitable compound derivatives of the template already containing linkages that could connect the substituents together. The resultant product structures containing these commercially available compound fragments would thus satisfy both binding site

constraints and combinatorial chemistry pre-requisites. Ideally, this 'linkage' strategy can address the problem of anchor dependence in the 'build-up' strategy, by focusing more on contributions from the substituents, while treating contributions from the template as a secondary effect. However, this may introduce a new problem of underestimating the influence of the template in some cases, and important binding modes with critical interactions between the template and binding site could be missed. Furthermore, this strategy also relies on the assumption of additivity, as described above.

All the three strategies have their own advantages and disadvantages, and therefore should be chosen according to the characteristics of different receptor binding sites and combinatorial libraries. Scoring function and conformational sampling still remain the two biggest challenges in the area of structure-based drug design, despite the progress that has been made [33–35]. Any advances in these two directions will definitely improve the performance of structure-based library design. Secondly, diversity selection under the constraint of a binding site is also an important but unsolved problem. Presently, most of the diversity-based selection work is done as a separate process either before or after the docking step, while the inclusion of binding features may provide a better measure of diversity in the case of structure-based library design. Thirdly, library design based on a family of protein structures is definitely an area worthy of further exploration, especially regarding the target specificity issue.

3. Applications

Two striking examples of the power of integrating library building, docking, synthesis, and biological evaluation were demonstrated through the work of Kick et al. [36] and Haque et al. [37], who showed that focused library design and synthesis could result in potent optimized leads for therapeutically relevant targets. They employed the second strategy mentioned above, identifying potent inhibitors of the aspartyl proteases, cathepsin D and plasmepsin II, as potential therapies for cancer metastasis and malaria, respectively.

A non-peptide isostere that mimicked the natural inhibitor pepstatin and consisted of three attachment sites or R groups was chosen as the template for library design of cathepsin D inhibitors, [36] starting from the X-ray coordinates of the cathepsin D/pepstatin complex. The system had the potential for generating almost 1 billion compounds. Using the CombiBUILD software [27,28] that automates the process of library design and docking, building blocks were selected for parallel synthesis. The orientation of the template was fixed relative to that of pepstatin in the X-ray complex. Their library design process incorporated conformational searching and clustering to examine the positioning of the potential R groups independently, then in combination with each other. Combinations that caused severe clashing were discarded from the library. Subsequent scoring with the AMBER force field [38] and hierarchical clustering of the highest scoring compounds provided a first generation 'directed' library of 1000 compounds for synthesis. These compounds were synthesized and tested, and 67 of them inhibited cathepsin D $\geq$ 50% at a 1 μM screening dose. Furthermore, subsequent K_i determinations on the most potent hits revealed a compound with a K_i below 100 nM. In a separate experiment, a diverse library of 1000 compounds was generated as a control. Synthesis and testing of these compounds provided only 26 confirmed hits, inhibiting cathepsin D $\geq$ 50% at a 1 μM screening dose. The diverse library, however, upon subsequent K_i analyses of selected compounds, provided inhibitors that were three to four times less potent than the sub-100 nM hits found from the directed library design. The authors stated that when their screening was performed at concentrations below 1 μM, there were seven times more hits resulting from the directed library than from the diverse library, suggesting that hit rates can be enhanced when protein structural information is used to guide the library design process.

Non-peptide plasmepsin II inhibitors were also identified by Kuntz and co-workers [37] using a library docking approach similar to the one described for cathepsin D inhibitors. Plasmepsin II is another aspartyl protease and the coordinates for plasmepsin II complexed to pepstatin are available. Thse library designed for cathepsin D was screened against plasmepsin II, since there was a reasonable degree of sequence homology between the active sites of the two proteases. The most potent lead identified from screening of this library was a potent inhibitor of cathepsin D (15 nM) as well as plasmepsin II (220 nM). Two leads identified from the cathepsin D library were further optimized for plasmepsin II activity in a second generation library. The template, which had three possible attachment sites, was examined in the active site and each site was individually optimized. This approach was also used in the cathepsin D library design protocol described above, and is very useful because it allows for the exploration of large numbers of substituents individually, without having to build and dock so many structures. If an individual substituent cannot be optimized for a site, it is discarded, so that a smaller number of substituents are selected for the final products, thus making the problem much less computationally intensive. The first library (R1 library) was constructed such that the R1 substituent was explored using DOCK while the other two were held constant as the native substituents derived from the initial leads. The 50 highest scoring substituents resulting from this analysis were added to the template, and screening of this initial library yielded a 100 nM plasmepsin II inhibitor. Next, a small library containing 12 variants of the R1 substituents identified in the first library was examined and prepared, yielding inhibitors in the 3 nM range. Both R2 and R3 libraries were also synthesized, yielding additional single digit nM inhibitors of plasmepsin II. Furthermore, the compounds in this study were designed to have relatively low molecular weights (594–650 Da) and

desirable Clog *P*-values (2.86–4.46). Two inhibitors were 15-fold selective for plasmepsin II over cathepsin D, the closely related homolog. In summary, protein binding site characteristics, biophysical properties, and enzyme selectivity were addressed through a combination of library building and docking to yield very potent and selective inhibitors of both cathepsin D and plasmepsin II in these two studies.

A new concept in library design was recently explored by Lamb et al. [39] to determine if multiple proteins could be probed effectively against multiple combinatorial libraries to yield compounds that would inhibit across structurally similar proteins. Their validation study focused on three proteins from the serine protease family: elastase, trypsin, and chymotrypsin, for which experimental binding affinities for the ligand/protein systems were available [40,41]. Three different virtual combinatorial libraries were prepared. One library was composed of peptides and the other two were non-peptide libraries, one containing benzodiazepines and the other tetrahydroisoquinolines. The methods used for library construction within the protein binding sites involved docking the template, selecting the best substituents for each individual attachment point, and then examining the fully built molecules. The DOCK 4.0 program was used to obtain up to 1500 conformations of the isolated template which displayed a methyl side chain at each attachment site. These conformers of the methylated template were subsequently sampled for orientational preference in the binding site, providing 2000 poses. Rigorous vector scoring that included hydrogen-bond weighting and a hydrophobic complementarity term was utilized for the structures that were attached with substituents at a single site, and the results were clustered based on RMS deviation. Only those representative poses from clusters that had good complementarity with the binding sub-sites were retained. Each subsequent attachment site was then examined individually, adding substituents to the template of each library. Poses that did not lead to a good presentation of the substituents within the binding subsites were discarded. The compounds in these single substituent sub-libraries were ranked through a combination of intramolecular scores within the ligand as well as intermolecular scores between the ligand and the protein. As a post-processing filter, the generalized Born solvation model [11] was implemented to improve the binding affinity calculations. Although it was not possible to include this correction early on in the docking because of its computational demand, it provided a means to more thoroughly evaluate the free energy of binding during the later stages of the docking protocol. The libraries were then taken on through subsequent stages of adding substituents to the remaining attachment sites, again utilizing the generalized Born solvation scoring [11] for the fully built molecules. In this way, a final library of fully substituted and optimized molecules was assembled for each of the three combinatorial libraries. The next task was to compare the scores of the docked virtual compounds obtained for the elastase, chymotrypsin, and trypsin systems. The peptide library scored well against all three targets, whereas the benzodiazapine library scored poorly against them. The tetrahydroisoquinoline library scored well for the elastase system (even better than the peptide library did), but it scored worse against the other two targets. In short, the authors were able to compare scaffolds against related targets within a protein family, and to determine which side chains were best able to present themselves to the sub-sites of the enzyme. Although the virtual compounds in the libraries were never synthesized or tested, the authors validated their computational findings by comparing their calculated results to the experimental findings of Lu et al. [40] and Otlewski and co-workers [41], who investigated the thermodynamic parameters within the primary specificity (S1) sub-sites of serine proteases complexed to various peptide inhibitors. This data provided a good basis for comparison to the virtual libraries, since the side chains on the three libraries (peptide, benzodiazapine, and tetrahydroisoquinoline) were all derived from the naturally and non-naturally occurring amino acid side chains.

Bohm et al. [42] demonstrated that a de novo ligand design strategy could be used effectively with the program LUDI [43,44] to design a combinatorial library of low molecular weight, non-peptide thrombin inhibitors. The library they designed is quite simple in terms of chemistry, and the compounds can be generated by a single-step reductive amination reaction, making the computational build-up of library compounds in the thrombin binding site relatively easy. The experimental results from this study were striking: of the 10 compounds that were synthesized, five showed K_i values in the nanomolar range, and the predicted binding mode was confirmed by a subsequent X-ray structure determination.

Novel factor Xa inhibitors were recently identified by Jones et al. [45] using the PRO_SELECT [26] module within the software suite Prometheus. A benzamidine-3-carboxamide containing a glycine amino acid linkage was selected form a small designed library and found to have a K_i of 12 mM against factor Xa. A model of this initial lead was examined in the active site of factor Xa, suggesting that the central glycine be replaced with lipophilic D amino acids. According to the model, this change on the structure could create an interaction between the central lipophilic residue of the ligand and the amino acid side chains comprising the disulfide pocket in the enzyme. A small library of 32 compounds with central lipophilic side chains was synthesized and tested, and a racemate containing a phenylglycine side chain had a K_i of 300 nM against factor Xa. Based on this finding, a second library of 20 lipophilic glycines was designed and prepared. A D-cyclohexylglycine derivative had a K_i of 500 nM, whereas the corresponding L-cyclohexylglycine derivative had a K_i of 7.8 mM, supporting the modelling hypothesis that only D-configuration side chains could interact within the disulfide pocket. A third generation library afforded even more potent arylglycine-derived inhibitors, with activities as low as 28 nM. Finally, one of the inhibitors with a K_i of 10 nM against factor Xa and 540 nM against trypsin was bound

to trypsin. The solved structure from the crystallographic study of the complex indeed supported several key interactions proposed by the model. These included a bidentate interaction between the benzamidine amidino of the ligand and Asp189 of trypsin, as well as the proposed lipophilic interaction within the disulfide pocket. The benzamidine inhibitor of the modeled complex overlapped extremely well with the inhibitor from the experimentally derived complex, especially where the key interactions were made in the S1 and lipophilic pockets.

4. Conclusions

Unlike the compounds in corporate and commercial libraries that are available for testing, the compounds in a structure-based designed library are not available and the chemist must develop a synthetic plan addressing reagent selection, efficiency of synthesis, and number of compounds which could be made in a reasonable time. Some of the reagents that are commonly used in library design such as amines, alkylating agents and carboxylic acids are abundantly available. Consequently, it is possible to generate up to a billion compounds in one library, even with a template containing only three points of attachment, as illustrated by the cathepsin D application described earlier. In that example, computational docking methods were used successfully to reduce these large libraries to smaller tractable libraries for synthesis. A medicinal chemistry team will hopefully commit to synthesizing between 100 and 300 compounds. Making only a handful of compounds will reduce the probability of finding an optimized lead, and will not provide adequate information about structure activity relationships.

Combinatorial chemistry provides us with efficient ways to explore chemical space while structure-based design guides this exploration to a distinct region of a therapeutic target, creating a powerful synergy between the two fields. With the explosive increase in the number of solved X-ray and NMR structures of therapeutically relevant targets, and with continuous innovation occurring in combinatorial chemistry, a significant impact on the drug discovery process in the pharmaceutical industry should be realized in the near future. Rational design of combinatorial chemistry libraries in binding sites of proteins could lead to success stories in the industry, with lead compounds being identified and progressed further into pre-clinical trials within a rapid period of time.

From the technical perspective, structure-based library design brings us more challenges, on one hand, due to the large size of most combinatorial libraries. On the other hand, it also provides us with ways to address some of the current difficulties in traditional structure-based drug design, like the inaccuracies of scoring functions and the inefficiencies of pose sampling. We might also reason that a library of ‘many’ designed compounds should afford better chances of finding strong binding hits than one or two compounds designed in the binding site. During the past several years, new computational methods for structure-based library design have been developed and applied to reduce the sizes of large combinatorial libraries to smaller, reasonably sized libraries for synthesis. These endeavors have successfully lead to the discovery of several novel lead compounds, and it is expected that this promising field will soon realize even more progress and success.

References

[1] E.M. Gordon, M.A. Gallop, D.V. Patel, Strategy and tactics in combinatorial organic synthesis. application to drug discovery, Acc. Chem. Res. 29 (1996) 144–154.
[2] F. Balkenhohl, C. von dem Bussche-Hunnefeld, A. Lansky, C. Zechel, Combinatorial synthesis of small organic molecules, Angew Chem. Int. Ed. Endl. 35 (1996) 2289–2337.
[3] G.S. Sittampalam, S.D. Kahl, W.P. Janzen, High-throughput screening: advances in assay technologies, Curr. Opin. Chem. Biol. 1 (1997) 384–391.
[4] K.R. Oldenburg, Current and future trends in high throughput screening for drug discovery, Ann. Rep. Med. Chem. 33 (1998) 301–311.
[5] R.N. Zuckermann, E.J. Martin, D.C. Spellmeyer, G.B. Stauber, K.R. Shoemaker, J.M. Kerr, G.M. Figliozzi, D.A. Goff, M.A. Siani, et al., Discovery of nanomolar ligands for 7-transmembrane G-protein-coupled receptors from a diverse *N*-(substituted) glycine peptoid library, J. Med. Chem. 37 (1994) 2678–2685.
[6] N.S. Gray, L. Wodicka, A.-M.W.H. Thunnissen, T.C. Norman, S. Kwon, F.H. Espinoza, D.O. Morgan, G. Barnes, S. LeClerc, L. Meijer, S.-H. Kim, D.J. Lockhart, P.G. Schultz, Exploiting chemical libraries, structure and genomics in the search for kinase inhibitors, Science 281 (1998) 533–538.
[7] R.A. Lewis, The design of small- and medium-sized focused combinatorial libraries, in: P.M. Dean, R.A. Lewis (Eds.), Design of Focused Combinatorial Libraries, Mol. Diversity Drug Des. (1999) 221–248.
[8] I.D. Kuntz, Structure-based strategies for drug design and discovery, Science 257 (1992) 1078–1082.
[9] R.E. Babine, S.L. Bender, Molecular recognition of protein–ligand complexes: applications to drug design, Chem. Rev. 97 (1997) 1359–1472.
[10] M.A. Murko, P.R. Caron, P.S. Charifson, Structure-based drug design, Ann. Rep. Med. Chem. 34 (1999) 297–306.
[11] X. Zou, Y. Sun, I.D. Kuntz, Inclusion of solvation in ligand binding free energy calculations using the generalized Born model, J. Am. Chem. Soc. 121 (1999) 8033–8043.
[12] P.S. Charifson, J.J. Corkery, M.A. Murcko, W.P. Walters, Consensus scoring: a method for obtaining improved hit rates from docking databases of three-dimensional structures into proteins, J. Med. Chem. 42 (1999) 5100–5109.
[13] C. Corwin, I.D. Kuntz, Database searching: past, present and future, Des. Bioact. Mol. (1998) 1–16.
[14] S.J. Dixon, M.J. Blaney, Docking: predicting the structure and binding affinity of ligand–receptor complexes, Des. Bioact. Mol. (1998) 175–197.
[15] H.J. Bohm, Current computational tools for de novo ligand design, Curr. Opin. Biotechnol. 7 (1996) 433–4366.
[16] D.E. Clark, C.W. Murray, J. Li, Current issues in de novo molecular design, Rev. Comput. Chem. 11 (1997) 67–125.
[17] Sybyl Software, Tripos, St. Louis, MO 63144, USA.
[18] Cerius2 Software, Accelrys, San Diego, CA 92121, USA.
[19] MOE Software, Chemical Computing Group, 1010 Sherbrooke Street W., Suite 910, Montreal, Canada H3A 2R7.

[20] I.D. Kuntz, J.M. Blaney, S.J. Oatley, R. Langridge, T.E.A. Ferrin, Geometric approach to macromolecule–ligand interactions, J. Mol. Biol. 161 (1982) 269–288.
[21] T.J.A. Ewing, S. Makino, A.G. Skillman, I.D. Kuntz, DOCK 4.0: search strategies for automated molecular docking of flexible molecule databases, J. Comput. Aided Mol. Des. 15 (2001) 411–428.
[22] M. Rarey, B. Kramer, T. Lengauer, G. Klebe, A fast flexible docking method using an incremental construction algorithm, J. Mol. Biol. 261 (1996) 470–489.
[23] Daylight Chemical Information Systems, Santa Fe, NM 87501, USA.
[24] Barnard Chemical Information Ltd., Sheffield, UK.
[25] D.J. Diller, K.M. Merz, High throughput docking for library design and library prioritization, Proteins Struct. Funct. Genet. 43 (2001) 113–124.
[26] C.W. Murray, D.E. Clark, T.R. Auton, M.A. Firth, J. Li, R.A. Sykes, B. Waszkowycz, D.R. Westhead, S.C. Young, PRO_SELECT: combining structure-based drug design and combinatorial chemistry for rapid lead discovery. Part 1. Technology, J. Comput. Aided Mol. Design 11 (1997) 193–207.
[27] D.C. Roe, Application and development of tools for structure-based drug design, Ph.D. Thesis University of California, San Francisco, USA, 1995.
[28] D.C. Roe, Molecular diversity in site-focused libraries, Mol. Diversity Drug Des. (1999) 141–173.
[29] Y. Sun, T.J.A. Ewing, A.G. Skillman, I.D. Kuntz, CombiDOCK: structure-based combinatorial docking and library design, J. Comput. Aided Mol. Design 12 (1998) 597–604.
[30] S. Makino, T.J.A. Ewing, I.D. Kuntz, DREAM++: flexible docking program for virtual combinatorial libraries, J. Comput. Aided Mol. Des. 13 (1999) 513–532.
[31] M. Rarey, T. Lengauer, A recursive algorithm for efficient combinatorial library docking, Perspect. Drug Discovery Design. 20 (2000) 63–81.
[32] A.R. Leach, R.A. Bryce, A.J. Robinson, Synergy between combinatorial chemistry and de novo design, J. Mol. Graph. Modelling 18 (2000) 358–367.
[33] M. Stahl, M. Rarey, Detailed analysis of scoring functions for virtual screening, J. Med. Chem. 44 (2001) 1035–1042.
[34] N. Majeux, M. Scarsi, A. Caflisch, Efficient electrostatic solvation model for protein-fragment docking, Proteins Struct. Funct. Genet. 42 (2001) 256–268.
[35] W.C. Still, A. Tempczyk, R.C. Hawley, T. Hendrickson, Semianalytical treatment of solvation for molecular mechanics and dynamics, J. Am. Chem. Soc. 112 (1990) 6127–6129.
[36] E.K. Kick, D.C. Roe, A.G. Skillman, G. Liu, T.J.A. Ewing, Y. Sun, I.D. Kuntz, J.A. Ellman, Structure-based design and combinatorial chemistry yield low nanomolar inhibitors of cathepsin D, Chem. Biol. 4 (1997) 297–307.
[37] T.S. Haque, A.G. Skillman, C.E. Lee, H. Habashita, I.Y. Gluzman, T.J.A. Ewing, D.E. Goldberg, I.D. Kuntz, J.A. Ellman, Potent, low-molecular weight non-peptide inhibitors of malarial aspartyl protease plasmepsin II, J. Med. Chem. 42 (1999) 1428–1440.
[38] W.D. Cornell, P. Cieplak, C.I. Bayly, I.R. Gould, K.M. Merz, D.M. Ferguson, D.C. Spellmeyer, T. Fox, J.W. Caldwell, P.A. Kollman, A second generation force field for the simulation of proteins, nucleic acids and organic molecules, J. Am. Chem. Soc. 117 (1995) 5179–5197.
[39] M.L. Lamb, K.W. Burdick, S. Toba, M.M. Young, A.G. Skillman, X. Zou, J.R. Arnold, I.D. Kuntz, Design, docking, and evaluation of multiple libraries against multiple targets, Proteins Struct. Funct. Genet. 42 (2001) 296–318.
[40] W. Lu, I. Apostol, M.A. Qasim, N. Warne, R. Wynn, W.L. Zhang, S. Anderson, Y.W. Chiang, E. Ogin, I. Rothberg, Binding of amino acid side chains to S1 cavities of serine proteinases, J. Mol. Biol. 266 (1997) 441–461.
[41] D. Krowarsch, M. Dadlez, O. Buczek, I. Krokoszynska, A.O. Smalas, J. Otlewski, Interscaffolding additivity: binding of P1 variants of bovine pancreatic trypsin inhibitor to four serine proteases, J. Mol. Biol. 289 (1999) 175–186.
[42] H.-J. Bohm, D.W. Banner, L. Weber, Combinatorial docking and combinatorial chemistry: design of potent non-peptide thrombin inhibitors, J. Comput. Aided Mol. Des. 13 (1999) 51–56.
[43] H.J. Bohm, The computer program LUDI: a new method for the de novo design of enzyme inhibitors, J. Comput. Aided Mol. Des. 6 (1992) 61–78.
[44] H.J. Bohm, LUDI: rule-based automatic design of new substituents for enzyme inhibitor leads, J. Comput. Aided Mol. Des. 6 (1992) 593–606.
[45] S.D. Jones, et al., The design of phenylglycine containing benzamidine carboxamides as potent and selective inhibitors of factor Xa, Bioorg. Med. Chem. Lett. 11 (2001) 733–736.