Journal of Computer-Aided Molecular Design, 8 (1994) 467–475
ESCOM

J-CAMD 243

Current methods for site-directed structure generation

Richard A. Lewis[a] and Andrew R. Leach[b]

[a]*Dagenham Research Centre, Rhône-Poulenc Rorer Ltd., Dagenham, Essex RM10 7XS, U.K.*
[b]*Department of Chemistry, University of Southampton, Highfield, Southampton SO9 5NH, U.K.*

Received 19 November 1993
Accepted 24 November 1993

Key words: Drug design; De novo design; Protein–ligand complexes; Enzyme inhibitors

SUMMARY

There has been a rapid growth of interest in techniques for site-directed drug design, fuelled by the increasing availability of structural models of proteins of therapeutic importance, and by studies reported in the literature showing that potent chemical leads can be obtained by these techniques. Structure generation programs offer the prospect of discovering highly original lead structures from novel chemical families. Due to the fact that this technique is more-or-less still in its infancy, there are no case studies available that demonstrate the use of structure generation programs for site-directed drug design. Such programs were first proposed in 1986, and became commercially available in early 1992. They have shown their ability to reproduce, or suggest reasonable alternatives for, ligands in well-defined binding sites. This brief review will discuss the recent advances that have been made in the field of site-directed structure generation.

INTRODUCTION

The specificity and affinity of a ligand for a protein receptor is governed by the principles of molecular recognition [1]. X-ray crystallographic studies of receptor–ligand intermolecular complexes have yielded many insights into the factors that influence the formation of a complex. It has long been a goal in medicinal chemistry to use the information contained in the structure of a receptor binding site to design new families of ligands. During the last few years, an increasing number of protein structures have been solved to atomic resolution, using X-ray crystallography and nuclear magnetic resonance spectroscopy. Pharmaceutical companies and academic institutions have recently been able to use these structures to find potent new ligands [2–5]. There has been a parallel growth in research into novel techniques for site-directed drug design. The techniques can be divided into two related families, database searching (reviewed by Martin [6]) and structure generation. The key advantage of database searching is that it requires no synthetic effort to generate new lead compounds. The key advantage of structure generation is that novel compounds, not contained in any database, can be suggested, but these compounds do have to be synthesised. This brief review will focus on the recent advances that have been made in the

field of structure generation, and on the research problems that remain to be solved. We have chosen to discuss the various structure generation programs in roughly historical order, beginning with an overview of site-directed drug design.

CHARACTERISING THE RECEPTOR

It is generally assumed that the specificity of many ligand–receptor interactions is controlled by a few key receptor groups that are found at the contact surface of the receptor binding site. These give rise to regions of space, called ligand points, where ligand atoms that interact favourably with the site may be located. Structure generation programs try to connect the ligand points in the binding site with molecular structures in an exhaustive and unbiased way. Only sterically and electrostatically reasonable molecular structures should be produced. However, structure generation algorithms are particularly prone to combinatorial explosion, so that simplifying assumptions have to be employed. The output from a structure generation program is often referred to as a molecular template. At present, these programs have little knowledge of how easy or difficult a structure may be to synthesise. It would be foolish to put a great deal of effort into making a compound for a binding assay, if there is no prior information that the compound will bind with reasonable affinity to the receptor. Molecular templates should therefore be regarded as suggestions for synthesis, to be inspected and changed where appropriate by the medicinal chemist.

Ligand points in a binding site can be determined computationally, for example by the GRID program [7]. In this program a probe group, such as a methyl or carbonyl group, is systematically positioned at all points on a regular grid that encompasses the binding site. The interaction energy between the probe and the macromolecule is determined using an empirical potential. Many of the programs described below use a variation of the GRID approach to assess the enthalpy of binding between a candidate ligand and the receptor. The interaction energy is derived from an empirical force field function, such as AMBER [8] or CHARMm [9]; such functions typically include van der Waals (vdW) and electrostatic terms, and sometimes a specific hydrogen bonding term.

An alternative approach dispenses with any form of energy function and instead uses rules based on experimental data to determine the locations where a strong interaction would be expected. Such an approach works best for directed interactions; the hydrogen bond is a prime example. Careful analyses of small-molecule crystal structures have yielded parameters that describe the directional nature of hydrogen bonds [10–11]. This data is exploited in the HSITE program [12] for predicting H-bond ligand points in a protein. A particular feature of HSITE is that it takes internal hydrogen bonds within the protein into account. A similar strategy has been used in the LUDI program [13,14], which also identifies hydrophobic and charge ligand points. The use of discrete ligand points is clearly an approximation to the continuous nature of the intermolecular potential in a site. However, the use of force fields is not without problems. The introduction of a ligand into a binding site can strongly perturb the potential. This, at least, is to some extent taken into account in the empirical scoring schemes.

Many of the methods for structure generation discussed below assume an essentially static picture of the receptor. This assumption is clearly unsound, but the nature of the conformational changes that occur on complexation cannot be predicted until a ligand has been fully designed. The snapshot model is commonly used, i.e., it is assumed that the uncomplexed conformation

of the receptor is a low-energy state and, as such, will be reasonably well populated in the complex at room temperature. Several conformations of the receptor and the ligand may be examined, but it is not possible at present to examine all the low-energy states. If there is evidence that a substantial conformational change has occurred on complexation, then the receptor model should be refined, new structures generated, and the design process repeated.

STRUCTURE GENERATION BY FRAGMENT JOINING

To our knowledge, DesJarlais et al. [15] published one of the first papers describing structure generation within a protein binding site, using the DOCK program to assemble methotrexate (MTX) within the binding site of dihydrofolate reductase. The ligand molecule was broken into fragments, each fragment was docked into the site (applying the features of the site to constrain the number of structures), and then the fragments were joined back together using energy minimisation to form MTX. This encapsulates the concept of structure generation by fragment joining, in which molecular skeletons are formed by joining together preformed three-dimensional fragments. There are a number of ways in which fragment joining has been employed in structure generation. For example, some approaches construct the entire molecule by joining fragments; other methods use fragment joining as one component, with different algorithms being used in other stages of the design process. All fragment-joining methods employ some form of database or template library that contains the 3D atomic coordinates of each fragment, and a set of joining rules. The library may contain more than one conformation of some fragments, such as ring structures, to allow better sampling of conformational space. The fragment conformations can be taken from crystal structures or generated using theoretical methods. Two fragments are joined by identifying and then superimposing the atoms common to the two fragments. A popular method is to minimise the rms distance between the two sets of common atoms; however, Leach et al. have shown that it may be better to employ other approaches to achieve the lowest energy molecule [16].

The GROW program of Moon and Howe [17] is a good example of a fragment-joining approach. GROW was initially designed to generate peptide inhibitors. It used a library containing a large number of 3D templates of the 20 naturally occurring amino acids. The templates were obtained by randomly generating conformations of the *N*-acetyl-*N'*-methylamide amino acids. Any randomly generated structure that contained two nonbonded heavy atoms closer than 2 Å was discarded; all other structures were minimised to remove minor vdW clashes. This gave between 53 (for proline) and 4987 (arginine) conformations per amino acid. A 'seed' fragment is placed in the site; this can be done manually, or by using a docking algorithm. Positioning and orientation of the seed in the site is perhaps the most subjective part of this method, as it can have a profound effect on the structure generation phase. All the fragments from the library are then fused to the seed, using a joining rule, to form a series of child templates. In this case, a join is effected through the formation of a new peptide bond. Each child template is scored according to the electrostatic and steric interactions between the template and the binding site, the internal energy of the template and solvation terms for the ligand and receptor. The best templates are selected for further generation cycles. The scoring function controls how atoms in the binding site influence the generation of structures. If all the child templates were retained, the number of structures to be stored and manipulated would grow exponentially. For this reason, only the 10

best structures are kept at the end of each cycle. The cycle of joining fragments to an existing template, scoring and selection is repeated until the site is full, or the template has reached a predefined size. This type of algorithm is analogous to some graph-searching algorithms, with the same limitations. Full-grown templates with good scores can easily be missed if the addition of one residue with a poor or neutral score causes the template to be pruned out at an earlier stage. In later versions of GROW, a Monte Carlo/simulated annealing algorithm was introduced at the selection stage to combat this effect. It has been reported that peptides generated by GROW show affinity for their intended target. It is worth noting that the use of peptides as leads means that the generated structures are easy to synthesise, but the task of finding non-peptidic leads remains.

Rotstein and Murcko have described a program called GroupBuild [18], which uses 'organic' fragments such as ethane, methoxy, benzene, etc. As with GROW, the interaction energy of each template is evaluated using molecular mechanics with a solvation contribution based on accessible surface area. The method can be used for complete de novo design or to modify known drugs by building upon a core structure. GroupBuild was evaluated against the FK506 binding protein, the human immunodeficiency virus (HIV) protease and carbonic anhydrase.

A variant on this strategy has been described by various groups [19–21]. Here, the library of fragments is composed of individual atoms, with different characters. As before, generation starts by placing a seed in the site. GRID contours are used to aid this procedure. Structures are grown atom-by-atom; each new atom is positioned at an appropriate distance from its bonding partner, and at the correct angle from any 1,3 neighbours. Newly added atoms are rotated around a range of torsion angles to find the best position, according to the scoring function employed. Rings are formed when a torsional rotation brings two nonbonded atoms close together. A Monte Carlo process is used to select the next atom to grow from, and to select the best structures. Further refinements involve the use of rigid superatoms, such as benzene rings, to speed up structure growth. The advantage of this method is that non-peptide structures are generated, but they may be very hard to synthesise. A consequence of the method of generation is that acyclic structures are created more frequently than is perhaps desirable. Programs like GENSTAR [20] should therefore be seen as (to quote the authors) 'molecular sketch pads'; a rich source of creative ideas for a medicinal chemist to explore. A similar procedure for building structures in two dimensions (with connectivities but no coordinates) has been described by Nilakantan et al. [22].

Fragment joining has a number of distinct advantages. Structures can be generated quickly, and if care is taken over the joining rules, these structures can represent low-energy conformations. However, as every structure is a composite of the fragments in the template library, the richness of the structures created is limited by the diversity of the database. As a defence to this, the use of very large databases would unacceptably increase the running time of the programs. In the interest of speed, and because the template is constantly changing, the scoring functions used are necessarily crude. This means that the energy of the template derived from the force field might only be a poor approximation to the true free energy change when the template binds to the receptor. Studies of the X-ray structures of ligand–receptor complexes indicate that the conformation of the bound ligand does not always correspond to an in vacuo minimum. MTX is a case in point. One solution to this difficulty is to include conformations in the fragment database that do not correspond to energy minima. An alternative approach, from the work of Leach and Kuntz [23] on conformationally flexible docking, is to use a set of discrete allowed

torsion angles, but to allow these torsions to be changed to eliminate any small steric clashes with the receptor atoms. In this way, the ligand is moulded to the shape of the receptor; an analogous procedure can be used to maximise any important enthalpic interactions, such as hydrogen bonds.

OUTSIDE-IN STRUCTURE GENERATION

The methods discussed thus far use an 'inside-out' approach, in which the ligand is grown within the site, typically using some form of energy function to direct the search. Another strategy for structure generation is to grow the molecular template from the 'outside in'. Here, the ligand points inside the binding site are determined (using the GRID, HSITE OR LUDI programs) and small chemical groups are fitted into these sites so as to maximise the enthalpy of interaction. The structure generation problem is then reduced to finding structures that connect all the chemical groups. The same criticism can be levelled against this approach as against the fragment-joining methods. The position and orientation of the seed chemical groups may not be well defined, leading to a range of starting conditions. This is only a difficulty when no structures are generated from a particular set of seeds.

The first attempt at an automated outside-in structure generator was described by Lewis and Dean [24,25]. It was noted that most drug molecules contain at least one ring and that the cyclic part of the structure is often important for the biological action. A library of ring fragments was created and these fragments were fused into large planar lattices. The lattices can be fitted to any quasi-planar array of ligand points and sterically clipped to provide a site-directed lattice. Structures are then generated by tracing paths through the atoms in the lattice. At each atom in the lattice, a choice is made for the next atom to visit. The generator follows along the bonding pathways contained within the lattice, always trying to fill the site in an energetically sensible way. This strategy can be extended to three dimensions in two ways, i.e., by the use of a regular diamond lattice [26] (to mimic an sp^3 C system), or through irregular lattices.

An irregular lattice can be obtained from a collection of molecules that have been positioned within the site, using a program such as DOCK [27]. The collection of molecules can be analysed to find all pairs of atoms that are separated by approximately one bond length. This will lead to the creation (in computro) of new bonds between atoms from different molecules [28]. These new bonds create a novel pathway for connecting fragments or molecules in different parts of the site. The fused collection of molecules forms a sensible irregular lattice. A program, BUILDER, has been written to use this irregular lattice to generate novel molecular templates [29]. BUILDER will generate templates to fill the site between specified points, using only the atoms in the lattice. The program also recognises sensible collections of atoms in the lattice, such as ring fragments, and will include these features in the template. BUILDER was used to design an inhibitor for the HIV protease. The active site of this enzyme is so large that database searches are very unlikely to find any one structure that fills the site. Use of an irregular lattice allowed ligand points in different parts of the site to be easily connected by a single molecular template.

Johnson and co-workers [30] have followed a similar strategy in the SPROUT program. A template library of all commonly occurring carbocyclic and acyclic fragments has been constructed by analysing the ORAC reaction database. The structure generation phase is started by positioning a template so that one of its atoms can occupy a ligand point. The template may be clipped to remove any steric clashes. This template then becomes the seed for the next cycle of

generation. An interesting feature of SPROUT is the use of probability tables to control the ratio of cyclic to acyclic fragments that are selected from the database. A disadvantage is that there are no heteroatoms in the structures generated; these have to be added at a later stage.

An interesting atom-based variant on the irregular lattice algorithm has been described by Levitt. Here, the binding site is filled with a large number of atomic spheres (microwaters) of radius 0.75 Å, and the system is equilibrated using molecular dynamics (MD). The centres of any touching pair of spheres are 1.5 Å apart, so that the microwaters form an irregular lattice. Structures are generated as before. The idea of using energy minimisation and MD to place atomic probes in the site in their most favourable binding positions has been explored more thoroughly by Miranker and Karplus in the MCSS program [31]. The binding site is filled with a large number of independent probes. During the energy minimisation or the MD simulation, the probes experience the field due to the receptor atoms, but not the field due to the other probes. The advantage over GRID is that the probe atoms are not restricted to fixed grid positions. When used with MD, the receptor atoms are free to move to accommodate the probes, introducing a degree of conformational flexibility. This does of course introduce a considerable computational overhead.

Pearlman and Murcko have described an MD-based approach called CONCEPTS [32]. The receptor site is filled with atom-like particles which interact with the site under the influence of a force field. The particles are also able to make covalent connections with each other in a dynamically reversible fashion. Although still in the early stages of development, the approach is able to suggest interesting structures that resemble reasonable molecules.

The LUDI program [13,14] uses a different strategy to connect fragments that have already been positioned within the site. The fragments that LUDI docks into the site are multi-atom groups or ring systems. The distance between atoms in neighbouring fragments must be relatively short (<6 Å). Connecting groups are found by searching a small database of fragments (in several conformations) for an appropriate template. The database is searched on the basis of distance. It is not clear if this procedure for joining the fragments will be robust when the distances to be bridged are large or the fragments are not ideally positioned with respect to each other. LUDI's rich library of fragments for docking, its speed of interaction and its interface within the INSIGHT program [33] make it a powerful tool for site-directed drug design, and it is now widely used in the pharmaceutical industry.

An alternative approach for finding connecting fragments is to search a database using bond vectors. This is the basis of the CAVEAT program [34], which searches a database of 3D fragments (mainly ring systems). The relative orientation of two fragments positioned within a site can be described by the vector between the two atoms to be joined, and the bond vectors between these atoms and the next atoms in the fragments. The CAVEAT database of structures is coded in terms of these parameters, and can be searched very quickly. A similar approach has been used in the NEWLEAD [35] and HOOK [36] programs. HOOK searches a database of fragments derived from the Cambridge Structural Database (CSD) [37] for groups that can bridge a pair of probes produced by MCSS. Each bridge group is scored according to its complementarity with the site and whether it can also link any other probe groups within the site. Ho and Marshall have also used databases of molecules taken from the CSD. Molecules that can interact with several ligand points are oriented in the site. These fragments are then 'spliced' together [38] to give larger structures.

A final example of outside-in structure generation is a torsion-based method [39]. This uses the geometric requirements of the site to determine the torsion angles in an unbranched chain of atoms. The positions of the atoms in the chain are only dependent on the torsion angles, as all other quantities (bond angles and lengths) are assumed to be fixed by the types of atoms. If the positions of the chemical objects to be joined are also known, it is possible to determine the torsion angles required to bridge the gap and hence to generate a chemical chain to connect the objects in an energetically reasonable manner. In particular, steric clashes with the site atoms would be avoided. This method has been used by Gō and Scheraga [40] and by Bruccoleri et al. [41] for constructing protein loops. The structure generator creates all the possible chains from a palette of atom types and checks whether each chain can be made to fit into the site. Leach and Kilvington [42] have described an alternative approach, using the 'tweak' algorithm of Shenkin et al. [43]. Similar strategies to these are used in some 3D database searching systems.

FUTURE DEVELOPMENTS

Second-generation structure generation programs are now being developed and commercialised, prime examples being Weininger's Genetic Algorithm (GA) approach [44] and the LEAPFROG program [45]. The GA program generates structures from SMILES codes. Each linear SMILES code can be converted into a 3D structure within the site, using distance geometry, and scored against the receptor. The results are used to select the best individuals in the population as parents for the next generation of structures. Crossover (for example, exchange of ring systems) and mutation (changes in atom type) are incorporated. The appealing feature is that the structure generation is entirely controlled by the scoring function, which can be changed to include biases for ring formation etc. While the LEAPFROG program is not a GA program, it uses many of the same techniques of crossover, mutation and weeding to remove poor structures. In addition, it can perform local minimisation of groups within the growing template, and refinement via minor changes (for instance, in atom type) to enhance the score of a structure. The choice of which operation to perform is stochastic and is controlled via data tables of relative probabilities. These two programs pull together many of the ideas described above, in a very interesting and exciting manner.

CONCLUSIONS

All the algorithms for structure generation discussed in this review will generate different flavours of structures, reflecting the biases introduced into each method. No one program is clearly better than another, and in a drug design project it is advisable to try out several strategies. The process of structure generation is still too slow to be interactive, unless simplifying assumptions are made. These assumptions speed up the algorithms, but at the cost of rigour and of limiting the diversity of the molecular templates that can be produced. A sensible compromise must be found between the need to provide a good choice of templates for the designer and the speed with which the scaffolds can be generated. This does not appear to be an insuperable difficulty. The difficulties of assessing the affinity and the synthetic accessibility of a designed ligand are less tractable. At the end of the design process, whatever method has been used, the generated structure should be re-evaluated critically, and possibly modified.

This review has concentrated on the presentation of all the common approaches towards structure generation, dwelling perhaps too long on some of the disadvantages of the current strategies. By way of contrast, it is clear that many groups are actively involved in the development or use of site-directed structure generation programs, and that these techniques are giving rise to much excitement in the medicinal chemistry community. We apologise to those whose work we have not included. Many of the programs discussed in this review are being constantly updated and improved. We think that it will not be long before many of the difficulties outlined above are solved, and the first reports of leads generated by these methods are published.

ACKNOWLEDGEMENTS

A.R.L. acknowledges support from the SERC under the Advanced Fellowship scheme.

REFERENCES

1 Dean, P.M., Molecular Foundations of Drug–Receptor Interaction, Cambridge University Press, Cambridge, 1987.
2 Lam, P.Y.S., Jadhav, P.K., Eyermann, C.J., Hodge, C.N., Ru, Y., Bacheler, L.T., Meek, J.L., Otto, M.J., Rayner, M.M., Wong, Y.N., Chang, C.-H., Weber, P.C., Jackson, D.A., Sharpe, T.R. and Erickson-Viitanen, S., Science, 263 (1994) 380.
3 Appelt, K., Bacquet, R.J., Bartlett, C.A., Booth, C.L.J., Freer, S.T., Fuhry, M.A.M., Gehring, M.R., Herrmann, S.M., Howland, E.F., Janson, C.A., Jones, T.A., Kan, C.-C., Kathadekar, V., Lewis, K.K., Marzoni, G.P., Matthews, D.A., Mohr, C., Moomaw, E.W., Morse, C.A., Oatley, S.J., Ogden, R.C., Reddy, M.R., Reich, S.H., Schoettlin, W.S., Smith, W.W., Varney, M.D., Villafranca, J.E., Ward, R.W., Webber, S., Webber, S.E., Welsh, K.M. and White, J., J. Med. Chem., 34 (1991) 1925.
4 Shoichet, B.K., Stroud, R.M., Santi, D.V., Kuntz, I.D. and Perry, K.M., Science, 259 (1993) 1445.
5 Kuntz, I.D., Science, 257 (1992) 1078.
6 Martin, Y.C., J. Med. Chem., 35 (1992) 2145.
7 Goodford, P.J., J. Med. Chem., 28 (1985) 849.
8 Weiner, S.J., Kollman, P.A., Case, D.A., Singh, U.C., Ghio, C., Alagona, G., Profeta, S. and Weiner, P., J. Am. Chem. Soc., 106 (1984) 765.
9 Brooks, B.R., Bruccoleri, B.E., Olafson, B., States, D.J., Swaminathan, S. and Karplus, M., J. Comput. Chem., 4 (1983) 187.
10 Murray-Rust, P. and Glusker, J.P., J. Am. Chem. Soc., 106 (1984) 1018.
11 Taylor, R. and Kennard, O., Acc. Chem. Res., 17 (1984) 320.
12 Danziger, D.J. and Dean, P.M., Proc. R. Soc. London, Ser. B, 236 (1989) 115.
13 Böhm, H.-J., J. Comput.-Aided Mol. Design, 6 (1992) 61.
14 Böhm, H.-J., J. Comput.-Aided Mol. Design, 6 (1992) 593.
15 DesJarlais, R.L., Sheridan, R.P., Dixon, J.S., Kuntz, I.D. and Venkataraghavan, R., J. Med. Chem., 29 (1986) 2149.
16 Leach, A.R., Prout, K. and Dolata, D.P., J. Comput.-Aided Mol. Design, 2 (1988) 107.
17 Moon, J.B. and Howe, W.J., Protein Struct. Funct. Genet., 11 (1991) 314.
18 Rotstein, S.H. and Murcko, M.A., J. Med. Chem., 36 (1993) 1700.
19 Rotstein, S.H. and Murcko, M.A., J. Comput.-Aided Mol. Design, 7 (1993) 23.
20 Nishibata, Y. and Itai, A., J. Med. Chem., 36 (1993) 2921.
21 Bohacek, R. and McMillen, C., Abstract OC-08.4, XIIth International Symposium on Medicinal Chemistry, Basel, 1992.
22 Nilakantan, R., Bauman, N. and Venkataraghavan, R., J. Chem. Inf. Comput. Sci., 31 (1991) 527.
23 Leach, A.R. and Kuntz, I.D., J. Comput. Chem., 13 (1992) 730.
24 Lewis, R.A. and Dean, P.M., Proc. R. Soc. London, Ser. B, 236 (1989) 125.
25 Lewis, R.A. and Dean, P.M., Proc. R. Soc. London, Ser. B, 236 (1989) 141.

26 Lewis, R.A., J. Comput.-Aided Mol. Design, 4 (1990) 205.
27 Kuntz, I.D., Blaney, J.M., Oatley, S.J., Langridge, R. and Ferrin, T.E., J. Mol. Biol., 161 (1982) 269.
28 Lewis, R.A., In Karjalainen, E.J. (Ed.) Proceedings of Scientific Computing and Automation, Elsevier, Amsterdam, 1990, pp. 117–132.
29 Lewis, R.A., Roe, D.C., Huang, C., Ferrin, T.E., Langridge, R. and Kuntz, I.D., J. Mol. Graphics, 10 (1992) 66.
30 Gillet, V., Johnson, A.P., Mata, P., Sike, S. and Williams, P., J. Comput.-Aided Mol. Design, 7 (1993) 127.
31 Miranker, A. and Karplus, M., Protein Struct. Funct. Genet., 11 (1991) 29.
32 Pearlman, D.A. and Murcko, M.A., J. Comput. Chem., 14 (1993) 1184.
33 Biosym Technologies Inc., San Diego, CA.
34 Bartlett, P.A., Shea, J.T., Telfer, S.J. and Waterman, S., In Roberts, S.M. (Ed.) Molecular Recognition: Chemical and Biological Problems, Royal Society of Chemistry, London, 1989, pp. 182–196.
35 Tschinke, V. and Cohen, N.C., J. Med. Chem., 36 (1993) 3863.
36 Eisen, M.B. and Hubbard, R.E., personal communication.
37 Allen, F.H., Bellard, S.A., Brice, M.D., Cartwright, B.A., Doubleday, A., Higgs, H., Hummelink, T., Hummelink-Peters, B.G., Kennard, O., Motherwell, W.D.S., Rodgers, J.R. and Watson, D.G., Acta Crystallogr., B35 (1979) 2331.
38 Ho, C.M. and Marshall, G.R., J. Comput.-Aided Mol. Design, 7 (1993) 623.
39 Lewis, R.A., J. Mol. Graphics, 10 (1992) 66.
40 Gō, N. and Scheraga, H.A., Macromolecules, 3 (1970) 178.
41 Bruccoleri, R.E. and Karplus, M., Macromolecules, 18 (1985) 2767.
42 Leach, A.R. and Kilvington, S.R., J. Comput.-Aided Mol. Design, 8 (1994) 283.
43 Shenkin, P.S., Yarmush, D.L., Fine, R.M., Wang, H. and Levinthal, C., Biopolymers, 26 (1987) 2053.
44 Weininger, D., Dixon, J.S. and Blaney, J.M., (1993) Daylight Chemical Information Systems Inc., Santa Fe, NM.
45 Cramer, R.D. and DePriest, S., implemented in the Sybyl program (1993), Tripos Associates, St. Louis, MO.