CHIMIA 2000, 54, No. 1/2

Chimia 54 (2000) 35–36

ISSN 0009–4293

Focal Point: Medicinal Chemistry

ILMAC Congress: October 14, 1999

Molecular Modeling for Drug Design*

Computer-aided molecular modeling (CAMM) has proven to be a highly valuable tool for rational drug design. Main applications of CAMM come from structure-based drug design (if 3D information of the target molecule is available) and QSAR (Quantitative Structure Activity Relationship). More recently new methods for the *de novo* design of ligands have been developed, *e.g.* LUDI, CAVEAT, *etc.* In addition new CAMM tools are now available that enable structure-based design of combinatorial libraries followed by high-throughput computational docking at the target binding site to select the most promising candidate ligands for chemical synthesis. The purpose of the mini-symposium was to illustrate to chemists from both industry and academia the current possibilities, limitations, and future developments in the molecular modeling area.

Keywords: Drug design · ILMAC · Medicinal chemistry · Molecular modeling

Structure-based Combinatorial Ligand Design: Methods and Applications

Prof. Amedeo Caflisch
Institute of Biochemistry, University of Zürich, CH–8006 Zürich, Switzerland

Two methods for structure-based computational ligand design were presented: Hydrophobicity maps are used to quantitatively estimate and graphically display the propensity of nonpolar groups to bind at the surface of a protein target. The approach is based on the calculation of the binding energy, van der Waals interaction, and protein electrostatic desolvation of a nonpolar probe sphere rolled over the protein surface, and on the color coding of this quantity on the molecular surface. The method was validated on ten protein–ligand complexes and was shown to distinguish precisely where polar and nonpolar groups preferentially bind [1]. Comparisons with existing approaches, such as the display of the electrostatic potential or the curvature, illustrate the advantages and the superior predictive power of the present method.

The program SEED (Solvation Energy for Exhaustive Docking) finds optimal positions and orientations of nonpolar fragments using the hydrophobicity maps, while polar fragments are docked with at least one hydrogen bond with the protein [2]. In SEED, an efficient evaluation of the binding energy, including continuum electrostatic solvation, allows a library of 100 fragments to dock onto a 25-residue binding site in about five hours on a personal computer. The SEED continuum electrostatic approach has been successfully validated by a comparison with finite difference solutions of the Poisson equation for more than 2500 complexes of small molecules with thrombin and the monomer of HIV-1 aspartic proteinase. The fragments docked by SEED in the active site of thrombin reproduce the structural features of the interaction patterns between known inhibitors and thrombin. Moreover, the combinatorial connection of these fragments yields a number of compounds that are very similar to potent inhibitors of thrombin.

[1] M. Scarsi, N. Majeux, A. Caflisch, *Proteins: Struct. Funct. Genet.* **1999**, *37*, 565.
[2] N. Majeux, M. Scarsi, J. Apostolakis, C. Ehrhardt, A. Caflisch, *Proteins: Struct. Funct. Genet.* **1999**, *37*, 88.

Highlights of Lessons Learned During the Last Twelve Years of Protein Structure Based Drug Design at Agouron

Dr. Michael D. Varney
Agouron Pharmaceuticals Inc., San Diego, USA

The presentation described many of the strengths, limitations, pitfalls, and other lessons learned using protein structures to design therapeutically useful drug candidates. Where possible, real-life examples were used to illustrate particular points. The talk was organized using small, sometimes provocative phrases or statements of lessons learned followed by an example where the lesson could be clearly seen. One example would be 'Go big early'. This statement was used to make the point that when designing into an active site of a protein, one should try and fill as much of the available space as possible. The example given was in the *thymidylate synthase* project where replacing a hydrogen with a methyl group filled an empty hydrophobic space and gave an increase in potency of 20 times. One of the key points of the presentation was that lessons learned in one project can be generalized to other projects.

Key references describing Dr. Varney's research on *thymidylate synthase*:

– T.R. Jones, S.E. Webber, M.D. Varney, M.R. Reddy, K.K. Lewis, V. Kathardekar, H. Mazdiyasni, J. Deal, D. Nguyen, K.M. Welsh, S. Webber, A. Johnson, D.A Matthews, W.W. Smith, C.A. Janson, R.J.

*Organized by:
Dr. Rainer Metternich
Novartis Pharma AG
Research, WKL-136.5.22
CH–4002 Basel
Tel.: +41 61 696 62 60
Fax: +41 61 696 62 83
E-Mail: rainer.metternich@pharma.novartis.com

Bacquet, E.F. Howland, C.L.J. Booth, R.W. Ward, S.M. Herrmann, J. White, C.A. Bartlett, C.A. Morse, 'Structure-Based Design of Substituted Diphenyl Sulfones and Sulfoxides as Lipophilic Inhibitors of Thymidylate Synthase', *J. Med. Chem.* **1997**, *40*, 677–683.
- T.R. Jones, M.D. Varney, S.E. Webber, K.K. Lewis, G.P. Marzoni, C.L. Palmer, V. Kathardekar, K.M. Welsh, S. Webber, D.A Matthews, K. Appelt, W.W. Smith, C.A. Janson, J.E. Villafranca, R.J. Bacquet, E.F. Howland, C.L.J. Booth, S.M. Herrmann, R.W. Ward, J. White, E.W. Moomaw, C.A. Bartlett, C.A. Morse, 'Structure-Based Design of Lipophilic Quinazoline Inhibitors of Thymidylate Synthase', *J. Med. Chem.* **1996**, *39*, 904–17.
- M.D. Varney, C.L. Palmer, J.G. Deal, S. Webber, K.M. Welsh, C.A. Bartlett, C.A. Morse, W.W. Smith, C.A. Janson, 'Synthesis and Biological Evaluation of Novel 2,6-Diaminobenz[cd]indole Inhibitors of Thymidylate Synthase Using the Protein Structure as a Guide', *J. Med. Chem.* **1995**, *38*, 1892–903.

Applications of Cheminformatics in Drug Discovery

Prof. Peter Willett
Dept. of Information Studies, University of Sheffield, Sheffield, UK

Cheminformatic methods for searching and processing databases of chemical structures play an increasingly important role in the discovery of novel drugs and agrochemicals. Many of the problems that are encountered in molecular design are inherently combinatorial in nature, and are thus amenable to processing by genetic algorithm (GA) methods that identify good, but suboptimal, solutions to combinatorial optimization problems. The talk described GAs for field-based similarity searching and for the selection of structurally diverse combinatorial libraries.

A molecular field is conventionally represented by embedding a molecule at the center of a 3D grid, and then calculating the value of the field at each point in the grid. While appropriate for QSAR studies, where only small numbers of molecules need to be aligned, current alignment procedures are far too slow for use in similarity searching of chemical databases, where a user-defined target structure needs to be aligned with each of the molecules in the database that is to be searched. Prof. Willett and his group have developed a GA-based method for rapidly aligning pairs of molecular fields. The chromosome in this algorithm encodes the translations and (torsional) rotations that are needed to align the target structure for a search with a database structure, and the fitness function for the GA is the similarity value that results from the alignment specified by the encoded translations and rotations. An extended series of similarity searches on the Standard Drug File and BIOSTER databases has been carried out to determine the effectiveness of the GA-based approach for identifying bioactive molecules. The results suggest that searches using representations of the electrostatic, shape, and lipophilic fields do not retrieve as many bioactive molecules as do conventional similarity searches using 2D fragment bit-strings. However, the active molecules that are retrieved by the field-based searches are far more disparate than those resulting from 2D searches, which suggests that the GA may provide a means for suggesting novel structural types in lead discovery programs that complements existing database searching tools.

The second project that was presented involved the design of combinatorial libraries. Given the sets of reagents involved in a combinatorial synthesis, it is possible to enumerate all of the possible reaction products; a so-called virtual library. Efficiency considerations dictate that only some of the members of the full virtual library should be synthesized and tested, and the GA developed in Prof. Willett's group seeks to identify a subset of this library that is maximally diverse and that can be synthesized in a combinatorial manner. The chromosome in this GA encodes possible sets of reagents from amongst those available, and the fitness function is the diversity of the resulting library when calculated using a quantitative index of structural diversity. Experiments with the GA demonstrate that its effectiveness can be increased by taking account of the physical properties of the molecules that are suggested for synthesis and testing. Specifically, the algorithm has been extended so that the suggested libraries have property profiles (such as molecular weight, numbers of donors or acceptors, or logP) that are as similar as possible to the profiles of known bioactive compounds.

Key references describing Prof. Willett's research:

- S.K. Drayton, K. Edwards, N.E. Jewell, D.B. Turner, D.J. Wild, P. Willett, P.M. Wright, K. Simmons, 'Similarity searching in files of three-dimensional chemical structures; identification of bioactive molecules', *Internet Journal of Chemistry*, at URL *http://www.ijc.com/articles/1998vl/37/*.
- V.J. Gillet, P. Willet, J. Bradshaw, 'The effectiveness of reactant pools for generating structurally diverse combinatorial libraries', *J. Chem. Inf. Comput. Sci.* **1997**, *37*, 731–740.
- V.J. Gillet, P. Willet, J. Bradshaw, D. Green, 'Selecting combinatorial libraries to optimize diversity and physical properties', *J. Chem. Inf. Comput. Sci.* **1999**, *39*, 169–177.
- V.J. Gillet, A. Schuffenhauer, P. Willet, 'Similarity searching in files of 3D chemical structures: analysis of the BIOSTER database using 2D fingerprints and molecular field descriptors', *J. Chem. Inf. Comput. Sci.* in press.