BIOACTIVE DESIGN: FORWARD TO NEW FRONTIERS

The contents of the five articles in this special issue were presented at the Molecular Graphics Society Annual Meeting on 'Molecular Graphics and the Design of Bioactive Molecules' which was held in Interlaken (Switzerland), June 7–11, 1993. They were selected from the main lectures introducing the topics of the conference, i.e., molecular modeling methods, use of databases in molecular design, template-based design methods using X-ray and NMR structure data, and progress in QSAR methods. This issue is not intended to be the proceedings of the conference; for more information, the reader may refer to the collected set of abstracts published in J. Mol. Graphics, 11 (1993) 248–283.

The intention of the organizers of this conference was to provide a forum for the presentation of the current state-of-the-art in the field of computer-based bioactive molecular design, with respect to both methodology and results, and to present to a larger scientific community what has been achieved in this area so far, where progress is visible, and in which directions the field is developing. For both the experts and the interested participant of the meeting, the scope and wealth of oral and poster presentations clearly demonstrated that the field is fully alive and expanding – perhaps even a bit chaotic – and the large exhibition of hardware and software on display by commercial and noncommercial groups also proved that many tools – expensive and inexpensive, simple and sophisticated – are available for use or are being developed.

For the purpose of provocation, picture the notorious naive meeting attendant who, a little lost among all the cryptic results and mystic methods presented, asks an expert to explain to him what, in fact, are the most important achievements in this area of biomolecular design, and which are the revolutionizing discoveries and design results? The not-so-naive expert might hope to get away with referring him to the impressive series of successful design stories that have been reported in the literature and at meetings of this kind in the past 10 to 15 years, since biomolecular design started seriously in a few industrial research laboratories. But the naive attendant may insist on some details – just the two or three most important achievements, please – and the expert might finally be pushed to admit that, apart from the initial idea of simulating molecular systems in a computer model (the foundations of which incidentally go back much further than the start of biomolecular design), most of what followed were in fact just extensions and refinements of existing theories and methods, putting them into productive computer programs, and applying these to biomolecular design projects, with varying success. And what about the most significant design exploits? Well, here the expert may feel that he is getting into deep water and may prefer to excuse himself.

To be fair, no one would seriously criticize the immense effort and work that has gone into the development and refinement of the modeling methods and design projects. But, to be realistic, the progress made in 15 years of developing design methods was more of a quantitative nature than of a qualitative one. In other words, today we handle large molecular systems (say, a receptor, a ligand and perhaps some solvent, altogether easily several thousand atoms) with

our powerful arsenal of software and hardware without any difficulties, but we do it practically in the same way as 15 years ago, when a computer simulation of a small organic molecule in vacuo was indeed an achievement, on hardware a thousand times less powerful than today's. And, of course, the deficiencies and inaccuracy of the methods remained the same, and so did the quality of the results.

This may sound like a rather pessimistic view on the state-of-the-art of biomolecular design. However, although no real new breakthroughs have been made in the past years, a number of promising new lines of development can clearly be distinguished:

(1) The 'old' refined, but presently still almost uniquely used, methods of molecular mechanics are quite adequate to study small molecular systems in vacuo. If the subject of investigation, however, is a quantitative simulation of biological systems (i.e., proteins with ligands and solvent), which are dominated by hydrogen bonding and charged interactions, the treatment of the electrostatic interactions in our 'old' molecular mechanics calculation has to be improved by a quantum leap. Some avenues in this direction have been tentatively explored with respect to polarizability of charged groups, but they are still in development.

(2) For realistic biomolecular design, the complexity of the model system has to match the complexity of the biological system under study, i.e., the above-mentioned 'classical' system consisting of a protein, ligand and solvent cannot reasonably be reduced to some 20 atoms of interest and still remain predictive. Handling such large classical systems requires more efficient programs than presently available, making full use of the architecture of modern parallel supercomputers. Fortunately a few projects in this direction are under way.

(3) A new area called 'lead finding', i.e., the search in chemical databases for bioactive compounds which match a defined structural profile, has emerged and a start has been made to build large 3D chemical databases of known organic molecules and to develop algorithms to subsequently dock molecules as ligands into a defined receptor-binding pocket, calculating for each docked ligand a figure of fit. As a first approach, the presently used 'rigid–rigid' fitting of a (usually) randomly chosen ligand conformation into a protein binding site with rigid side chains may be interesting to start with, but for serious 'lead finding' this is inadequate; only by chance one can hope to find a useful new ligand in this way. In order to overcome the 'rigid–rigid' limitation, novel algorithms and very efficient programs have to be developed. These may improve lead finding from a process depending on pure chance to one with an acceptable probability of success. An interesting development in this direction is described in this issue in the article by G. Lauri and P. Bartlett.

(4) In the related area of 'ligand design' some preliminary results and methods are emerging, called 'de novo' design procedures, which do not depend (heavily) on databases, but are generating new molecules into a binding pocket out of a random distribution of atoms, typically using genetic algorithms. The road to productive programs in this area is still long, but the directions of development have been laid out and there is much potential in these methods.

(5) Due to the accessibility of (almost) any protein by modern biotechnology, and the high efficiency of modern X-ray and NMR structure determination methods, biomolecular design can sometimes be based on experimental models. These, of course, offer the best possible starting models for computer simulation and biomolecular design, but they also challenge the modeler to produce design results that match the quality of the starting model. As pointed out in (1) above, this needs improved molecular mechanics methods.

(6) Biotechnology and experimental structure determination are, and will remain, slow and laborious (and therefore expensive) procedures. So, another challenge for the biomolecular designer has emerged: to predict with high accuracy the structure of novel, not yet experimentally determined, proteins on the basis of homology and family identification (see the article by S.D. Rufino and T.L. Blundell in this issue), making use of the structural data of all the experimentally determined proteins. As the number of experimental protein structures is increasing, this goal may eventually be reached, but again, provided that the computer simulation techniques for proteins are improving as pointed out above.

In summary, the Interlaken meeting of the Molecular Graphics Society showed a very lively biomolecular design scene, but it also made clear that the 'wild and easy' start-up period for this field is over; it has reached a mature state where higher quality is required. Further growth will be slow and difficult, but quite a few promising alleys of progress have been found and are being explored, and methodological breakthroughs are to be expected. A lot still remains to be done – bioactive design, forward to new frontiers.

H.P. Weber