Pure & Appl. Chem., Vol. 65, No. 2, pp. 231–234, 1993.
Printed in Great Britain.

Computers in drug design

W. Graham Richards

Physical Chemistry Laboratory, South Parks Road, Oxford OX1 3QZ, England

ABSTRACT

Computational techniques may be used to aid the process of drug discovery both when the target macromolecule is of unknown structure and when it is understood in atomic detail. Methods include measures of molecular shape; molecular similarity; quantum, statistical and molecular mechanics; molecular dynamics; Monte Carlo calculations and molecular graphics.

Amatz Mayer of the Hebrew University came to Oxford to spent a sabbatical year in 1989. He and I had been contemporary post-doctoral researchers in Paris in the mid 1960s but had never met, although I was familiar with his work on molecular shape. Together we devised a problem for him to work on which extended his work on molecular shape and contributed to my research on methods involved in computer-aided drug design. Despite becoming ill, Amatz largely completed that project and his untimely death deprived him of making an even bigger impact in this field as well as causing great grief to his scientific colleagues in Oxford as well as to his friends and family in Jerusalem.

This contribution is dedicated to him and shows the important nature of his work in the field of computers in drug design.

INTRODUCTION

Computer-aided drug design can be sub-divided into two categories, depending on the available knowledge about the macromolecule with which the smaller drug molecule interacts so as to disrupt some biochemical process. Broadly speaking these distinct categories are:

a) Nothing is known in atomic detail about the target receptor.
b) A detailed molecular structure of the target is available from X-ray crystallography, nuclear magnetic resonance or by homology model building.

In the commercial world the former sub-class predominates.

Here consideration will be given to both aspects, starting with the former, where the contribution from Amatz Meyer is seminal (ref. 1).

THE UNKNOWN TARGET

A simple logical procress should lead to a prescription for drug design. We start with the diagram of biochemical pathways and choose which step we aim to block; such as synthesis of a particular reagent. Then the energy profile of the biochemical reaction must be calculated, using quantum chemistry or empirical approaches separately or in combination (Fig. 1).

From these calculations, transition states or intermediates must be specified. Working on the logic that catalytic enzymes lower transition states, then stable transition state analogues should bind to and inhibit the enzyme and hence disrupt the biochemistry. This idea goes back to Linus Pauling (ref. 2) and is essentially the mechanism of action of penicillin.

Once the structure of the transition state is specified then we need to design the stable molecule which 'looks like' that transitory structure. The term 'molecular similarity' has come to be used

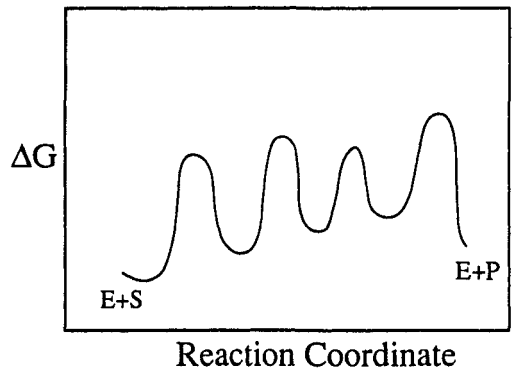

Fig. 1. The free energy profile of a typical enzyme catalysed reaction.

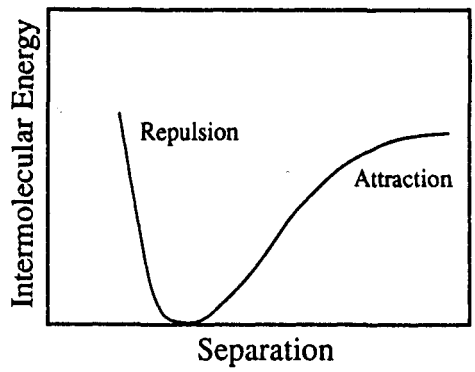

Fig. 2. The general form of the intermolecular interaction between a drug and its receptor.

for quantitative measures of just how similar molecule A is to molecule B. Similarity indices have been devised by Carbo (ref. 3) and by Hodgkin (ref. 4):

$$\text{Carbo, } R_{AB} = \frac{\int p_A p_B \, dv}{\int (p_A^2 \, dv)^{1/2} \int (p_B^2 \, dv)^{1/2}}$$

$$\text{Hodgkin, } R'_{AB} = \frac{\int 2 \, p_A p_B \, dv}{\int p_A^2 \, dv + \int p_B^2 \, dv}$$

Here p_A and p_B are properties of molecules A and B at points in space; the numerator represents an overlap of the properties and the denominator provides a normalizing factor such that the range of similarity is from zero to unity (or identity). In the earliest work the property p_A or p_B was taken as an electron density from a wave function but later other properties have been used.

If we note that the general form of the interaction between the drug and its target macromolecule will be as in Fig. 2, then it is clear that of all the contributions to intermolecular forces then it is the repulsion which is the chief discriminator. In other words the first prerequisite is that the drug will fit the binding site. This is where Meyer's measure of shape similarity comes in (ref. 1). In order to define a molecular shape the van der Waals volume of a molecule is placed in a three-dimensional grid and the property of shape given a value 1 if it is inside the surface, a zero outside. When molecules are superimposed as in Fig. 3, the numbers of grid points in both molecules, B, or only in molecule 1 (O_1) or only in 2 (O_2) are counted. In these terms the number of points in 1 is ($B + O_1$) and in 2 is ($B + O_2$). The similarity indices now become

$$R = \frac{B}{(T_1 T_2)^{1/2}} \quad \text{and} \quad R' = \frac{2B}{T_1 + T_2}$$

For the attractive part of the intermolecular potential the property of electrostatic potential seems the most promising choice. This can be computed rapidly from point charges as the individual atoms and again treated numerically using a gridded box (ref. 5). Recently (ref. 6) we have achieved a significant improvement over this numerical approach by realising that since electrostatic potential is given by a sum of terms of the type q/r, then the 1/r part may be fitted by gaussians. This proves to have some very important advantages when one remembers that it is the electrostatic potential surrounding the molecule which is of importance in recognition and binding. By using, say, two or three gaussians to represent 1/r we can perform the integrals in R_{AB} and R'_{AB} analytically and avoid any singularities at nuclei (where 1/r goes to infinity). Improvements in speed of calculation are approximately one hundred-fold and much better success is achieved in optimizing the overlap of structures as local minima are avoided. For instance we can take a complex molecule and overlap it upon itself giving a similarity of one, when a grid-based method frequently gets stuck at perhaps R = 0·85.

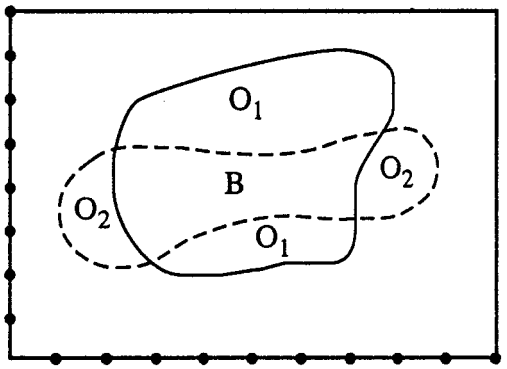

Fig. 3. Schematic representation of a cut through two superimposed molecules and the containing box. The surface of molecule 1 is drawn solid and that of molecule 2 dashed.

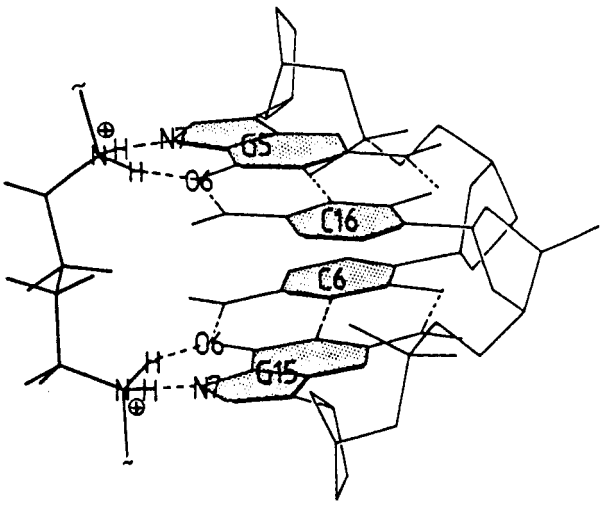

Fig. 4. Proposed model of spermine binding to the major groove of poly (dG-dC).

The similarity of shape and electrostatic potential may be used not only to design compounds resembling transition state structures but also in quantitative structure activity studies (QSAR). Currently we are trying to use the two similarities as inputs to neural networks in this regard. Preliminary results are extremely encouraging.

THE KNOWN TARGET

A known target is likely to be either DNA or an enzyme whose crystal structure is available.

DNA is a target of growing importance as the human genome project yields more sequences. The challenge is to design sequence specific compounds. Much work has been published (ref. 7) on minor groove binders such as netropsin, but we have looked to major grove specific effects. Figure 4 shows the results of modelling work which leads to the notion that spermine may show some preference for G-C regions of DNA.

Although spermine has a wide variety of effects upon DNA, we believe that variants may be good lead compounds for the development of drugs and it may also act as a useful substituent to enhance both drug transport and selectivity. Another line of approach is the use of organometallic compounds. In particular we have made detailed modelling structures of ruthenium triphenanthrolines (ref. 13).

Other exciting possibilities include antisense oligonucleotides which will certainly produce sequence specificity but will need modification if they are to produce blocking effects.

With enzymes as targets an increasing variety of possibilities are emerging. We have worked for some time with dihydrofolate reductase as this enzyme to block with anti-tumour activity being the aim.

There already exist excellent inhibitors of this enzyme such as methotrexate or trimethoprim. Our angle (ref. 8) has been to attempt to design anti-cancer drugs which are bioreductive. This means having compounds which can exist in oxidized or reduced forms with only the latter form inhibiting the enzyme. If the redox potential is appropriate the lethal reduced form will be available in anoxic tumour cells while the oxidized harmless form would be prevalent in normal cells. Design will then go beyond mere model fitting to include calculation of binding free energy and redox potential. Both have proved possible using the free energy perturbation technique.

BUILDING ENZYME STRUCTURES BY HOMOLOGY

If we were restricted to looking at enzymes whose crystal structure had been determined, the range of targets would be severely limited. On the other hand, were it possible to go from gene sequence to tertiary protein structure computationally, there would be a huge range of possibilities. In general that leap is not possible, but in the limited sense of building protein structures using homology with those known in the protein structural database is very promising.

We have made two distinct approaches to homology modelling based on quite different philosophies, one embracing subjectivity and the capacity of the human eye to detect patterns and the other being purely non-subjective.

The homology studies which harness the pattern recognition qualities of the human mind have been incorporated into computer graphics software by Morris (ref. 9) as the CAMELEON program. This takes amino acid sequences from a known and an unknown structure. By using interactively colour grouping files the user is aided in matching bits of the sequence of one protein to sections of the other. Gaps are inserted by eye. This method has permitted the prediction of interesting cytochromes P-450 on the basis of the crystallographic structure of $P\text{-}450_{camphor}$ which is itself of no major biological importance (ref. 10).

The alternative strategy of Gilbert (ref. 11) seeks homologous regions of protein sequence using the properties of residues rather than their identity. Similarities of short lengths of sequence can be stored in a similarity matrix where low scores of root mean square difference are important. A matrix of the low- scoring mutation may be made for window lengths varying from one to the length of the shorter protein. Averages of these indicate regions of similarity. Using this software Menziani et al (ref. 12) managed to shown an homology between the structure of human big endothelin (big ET) and a section (residues 13-51) of a scorpion neurotoxin (ISN 3). The important structure of the big endothelin was then obtained by superimposition of its sequence on the known toxin structure followed by minimization and molecular dynamics including solvent to yield a predicted tertiary shape for the 38-residue polypeptide. This molecule is believed to be cleaved by proteolysis to yield the 21 amino acid endothelin which is said to be the most potent and long-lasting vasoconstrictor yet identified. The structure of the big ET should facilitate the design of inhibitors of endothelin release.

CONCLUSION

Both when the therapeutic target is known and when it is unknown, computational chemical techniques can help in drug discovery.

Where we have a crystal structure of the target we are on safe ground. Homologies promise to extend this.

If the target is unknown, then molecular similarity is an important tool in finding either quantitative structure-activity relations or preferably molecules which resemble transition states in properties or in shape.

REFERENCES

1. A.Y. Meyer and W.G. Richards, J. Comput. Aided Mol. Design 5, 427-439 (1991).
2. L. Pauling, American Scientist 36, 51-58 (1947).
3. R. Carbo, L. Leyda and M. Arnau, Int. J. Quantum Chem. 17, 1185-93 (1980).
4. E.E. Hodgkin and W.G. Richards, Int. J. Quantum Chem. Quantum Biol. Symp. 14, 105-109 (1987).
5. C. Burt, 'ASP', Oxford Molecular Ltd., Magdalen Centre, Oxford Science Park, Oxford OX4 4GA, UK.
6. A.C. Good, E.E. Hodgkin and W.G. Richards, J. Chem. Inf. & Comput. Sci. in press.
7. R. Lavery, K. Zakrzewska and B. Pullman, J. Biomol. Struct. Dyn. 3, 1155-1170 (1986).
8. C.A. Reynolds, P.M. King and W.G. Richards, Nature 334, 80-81 (1988).
9. G.M. Morris, 'CAMELEON', Oxford Molecular Ltd., Magdalen Centre, Oxford Science Park, Oxford OX4 4GA, UK.
10. G.M. Morris and W.G. Richards, Biochem. Soc. Trans. 19, 793- 795 (1991).
11. R.G. Gilbert, D.Phil. Thesis, Oxford University (1992).
12. M.C. Menziani, M. Cocchi, P.G. de Benedetti, R.G. Gilbert, W.G. Richards, M. Zamai and V.R. Caiolfa, J. Chim. Phys. 88, 2687-2695 (1991).
13. I.S. Haworth, A.H. Elcock, A. Rodger and W.G. Richards, J. Biomol. Struct. & Dynam. 9, 533-569 (1991).