Computer-aided molecular design

W. Graham Richards and Lance Mangold

Until recently the search for substances having desired biological properties depended largely on systematic screening of very large numbers of compounds, as in the development of the sulphonamide drugs. Today, the alliance of computer graphics with theoretical chemistry is beginning to make it possible to predict the kind of molecule which will produce a particular effect.

In many ways, but most notably in the pharmaceutical and agrochemical industries, chemists are involved in molecular design. It is their job to synthesise molecules having very specific properties which will interfere with biological processes, usually by replacing a molecule which operates the control system of a living animal or plant. The molecules created are thus of the size of neurotransmitters or hormones and biology imposes very strict limitations on the molecules which will mimic the role of a natural molecule or, alternatively, block its action. Changing one atom for another, or replacing a group of atoms by a similar one, may have scant effect on the physical or chemical properties of the molecule, but may nevertheless alter the biological activity by many orders of magnitude. The molecular criteria for successful pharmaceutical compounds are, therefore, exceedingly strict.

In the past many of the most successful new therapeutic agents were discovered by a combination of luck and sharp-eyed clinicians or by systematic screening programmes carried out by researchers in universities or, more often, in private industry. Some of the most recent spectacular successes on the other hand were reached in a more rational way. Projects started with one of the body's natural control molecules, such as adrenaline, and the molecular structure was altered by chemists to produce a series of new but related molecules; a small percentage of these when administered interfered beneficially with the action of the natural molecule.

The problem for the medicinal chemist is deciding which molecules to synthesise and test. Given the vast number of organic chemicals which can be made with only minor differences in atomic constitution, hit-or-miss variations are unlikely to be very successful. This random approach has been referred to, rather unkindly, but accurately, as molecular roulette. More effective and rational procedures involve consciously mimicking the natural molecule with systematic variations and then searching for correlation between activity and functions of measurable parameters, such as the solubility in fat or dipole moment of the molecules. Now theoretical chemists have got in on the act and there is a growing use of computers for the assistance of medicinal chemists: computer-aided molecular design has arrived.

Molecular design

The basis of molecular design is the belief that the small control molecules of biology act at the molecular level by binding to a specific slot in a large molecule (a receptor or enzyme) in order to achieve their effect, rather as a key fits a lock or—perhaps more realistically—as a hand fits a glove.

Designing a molecule to be a suitable replacement for the natural key, or conversely, to jam the lock, will obviously demand creating similarity in shape and electronic properties. Chemists have for many years compared molecular shapes by means of models representing atoms and valency bonds but the computer offers major advantages. Shapes may be superimposed on a television screen to highlight minor differences, and colour can assist the interpretation. In addition, the shape of electron clouds and their properties are amenable to theoretical computation, so that completely new aspects of molecules can be compared. The theoretical techniques are not novel but until the advent of the facilities of computer graphics it was not possible to view and compare the results of computations in a manner readily assimilable by the synthetic chemist. Computer graphics are providing some answers to the question of what molecules to synthesize next.

Nuclear positions

When a chemist talks loosely of 'molecular shape' he means the positions of the atomic nuclei defined in terms of bond lengths and angles or nuclear coordinates: water, H_2O, has a shape defined by the O-H bond lengths and the angle between them. If X-ray crystal structures are available for a series of molecules then the shapes may be compared by displaying the molecules on a screen as three-dimensional line drawings. These may be presented either as pairs of images suitable for stereo-viewing or, more often, in a quasi-three-dimensional form using 'depth cueing', in which parts of the picture supposed to be more distant from the viewer are fainter. Using two colours it is much easier to compare shapes in this way rather than by trying to superimpose physical models.

If the X-ray structures are unknown the likely shape is amenable to theoretical calculation. Purely empirical methods, known as molecular mechanics, treat molecules as collections of balls and springs with potentials which vary with interatomic separation and on making torsional twists around bonds. Very many parameters have to be introduced into these empirical energy functions, but over the years they have evolved to a very satisfactory level and can be used to predict crystal structures [1]. Beyond this, the same potentials can be used to calculate which shapes of flexible molecules could possibly be adopted, as well as the minimum-energy crystal form. Here is something from theory which is not available from experiment, and is of great importance. 'Lock-and-key' is not a good analogy since both partners are frequently flexible; we need to compare not only the shapes of

W. Graham Richards, M.A., D.Phil., C.Chem., F.R.S.C.

Is a lecturer in the Physical Chemistry Laboratory, Oxford, where he was both an undergraduate and graduate student. A second edition of his book 'Quantum Pharmacology' will appear in 1983.

Lance Mangold, L.I.I.P.

Is a scientific photographer working in the Department of Molecular Biophysics, Oxford on new methods of molecular display.

Endeavour, New Series, Volume 7, No. 1, 1983.

0160-9327/83/010002-03 $03.00.

molecules found in the solid state, but also those which could possible be adopted without expending too much energy. Ideally, these energy calculations should be performed as the operator manipulates the shape of the molecule on the computer display. For this reason the other method of calculating the energy of a molecule as a function of altering its shape—quantum mechanics—despite being on firmer ground than molecular mechanics in terms of physics, is not so useful for this type of design work. Where quantum mechanics does come into its own is in calculating and comparing electronic properties.

Electronic positions

The relative positions of the atomic nuclei in a molecule, or the line drawings with bonds indicated, represent merely a molecular skeleton. The 'flesh' is provided by the clouds of electron density. The shape of a molecule, from which it may be possible to infer the complementary form of its receptor site, demands a knowledge of the electronic contribution to supplement the nuclear information.

The simplest and most widely used method of fleshing out a molecular skeleton is to assume that each atom is a small sphere of a size dependent on its position in the Periodic Table of elements and on the nature of its bonded neighbours in the molecule. Such radii—van der Waals radii, or Pauling radii—are available in tables. The computer graphic displays of molecules now almost universally can switch from the illustration of the skeleton to the more solid looking form embodying these radii. The actual pictures (see figure 1) look like CPK (Corey-Pauling-Koltun) models, with hidden lines removed, and shading added as if they were illuminated from a point, to give an impression of depth. Atoms of different types may be colour-coded to facilitate interpretation.

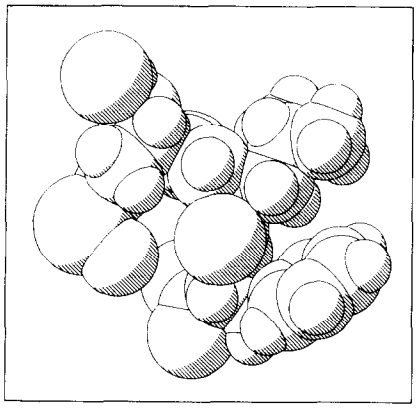

Figure 1 A space-filled model drawn using the PLUTO computer program. The molecule is Formyl-L-Met-L-Leu-D-Phe-OH; picture produced using the program PLUTO 78 written by S. Motherwell, Cambridge Crystallographic Centre and suppled by Dr A. Morffew, IBM UK Scientific Centre.

Atoms are, of course, far more more complex objects than tiny spheres. The electron density round each nucleus may be anything but spherically distributed. More accurate pictures of the electronic flesh on the nuclear skeleton can be derived from quantum mechanical calculations. There are now universally available computer programmes which can solve the Schrödinger equation for a molecule to an almost arbitrary level of accuracy [2]. The data for these *ab initio* quantum mechanical programs are merely nuclear positions, nuclear charges, and the number of electrons in the system. The solution yields the energy, E, on an absolute scale for the molecule in the defined geometry and conformation, and the corresponding wave function, ψ, from which, if the appropriate operator is used, any physically observable property may be calculated. Most notably, the value of ψ^2 (or more generally $\psi^* \psi$ if ψ is a complex function) at any point in the space in or around the molecule gives the electron density at that point.

Pictures of electron density contours for simple molecules calculated in this way are familiar in elementary chemistry texts. For the non-symmetrical, more complicated molecules of biology, contour maps of electronic charge drawn in arbitrary planes through the molecule are not particularly helpful. Computer graphics, on the other hand, does render it possible to display quasi-three-dimensional representations. Figure 2 shows an example where two contours are indicated: a low value with dots and a higher level with a mesh or 'chicken-wire'. The choice of this contour level is of course arbitrary, but logical choice may be made by doing a series of separate quantum mechanical calculations to investigate the energy of interaction of the molecule in question with model neighbours such as hydrocarbons or water. Such calculations show very clearly that what we loosely refer to as the 'size' of a molecule is highly dependent on its molecular context. The energy of interaction between a molecule and water molecule is strongly influenced by the orientation of the water molecule as it approaches. For useful displays some compromise must be made. We have found that molecules are very 'hard'; that is there is strong repulsion when partners get as close as the 0.01 atomic unit of charge

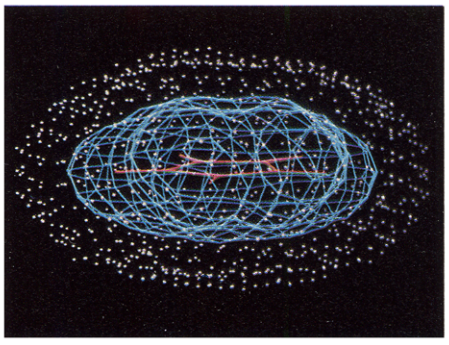

Figure 2 A quasi-three-dimensional electron density map of pyrrole. The contours are drawn at 0.01 a.u. (blue) and 0.001 a.u. (white dots).

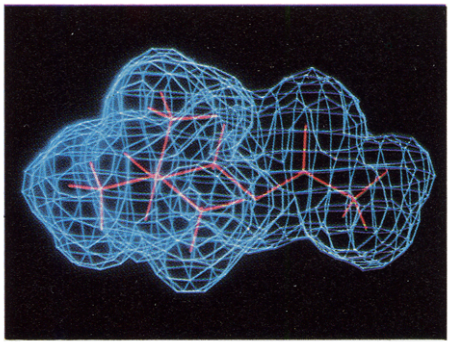

Figure 3 A quasi-three-dimensional electron density map of acetylcholine—contour at 0.01 a.u. of charge.

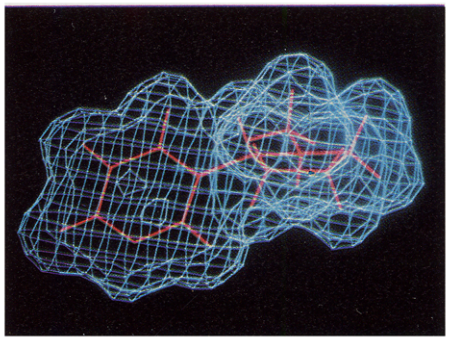

Figure 4 A quasi-three-dimensional electron density map of nicotine—contour at 0.01 a.u. of charge.

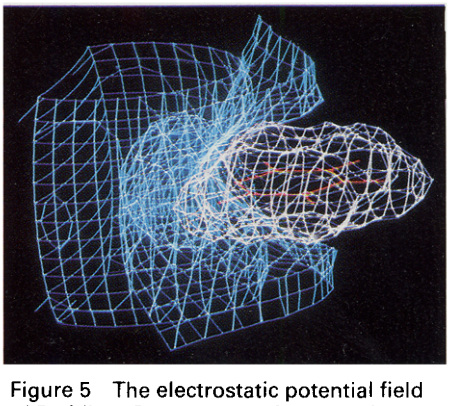

Figure 5 The electrostatic potential field of Pyridine. Contours at −0.0001, −0.005, −0.03 a.u. (blue) and +0.05 a.u. (white). [1 a.u. = 2625.4 kJ mol^{-1}].

contour (1 a.u. of charge density = 67.5 electrons per cubic Ångstrom). Figures 3 and 4 show a comparison of acetylcholine and nicotine displayed in this way. Some similarities in shape are indicated and we have found that this extends to other molecules which act biologically at the same centre. This provides useful clues for the synthetic chemist.

Electrostatic potentials

The fact that the electronic picture of a molecule is not fixed but is strongly influenced by its molecular neighbour is met in small part by using the quantum mechanical wave function to compute not the simple charge density but rather the electrostatic potential generated by the molecule. This potential is the energy of interaction of the molecule in terms of both nuclei and electrons density, with an isolated proton [3]. Energy will vary with the location of the proton, being attractive in some regions and repulsive in others. Yet again contour diagrams may be calculated but like electron densities they are difficult to appreciate unless displayed using computer graphics, and preferably using colour to distinguish attractive and repulsive areas. Figure 5 gives an example. Pictures such as this are of value in highlighting similarities in molecules which go beyond mere shape. If two molecules bind to the same receptor they may have to present closely similar electronic aspects to that site.

Receptor influence

If the presence of a proton close to a molecule influences its electronic distribution then how much more will the molecular surroundings of a receptor binding site distort the properties of a molecule? The answer must be, considerably. This casts doubt on calculations performed on isolated molecules. To counter this defect in theoretical work on biological molecules the traditional posture has been to look at series of similar species and to trust that the influence of the receptor surroundings will be a constant feature for all of them. Attention is then focused on differences in biological activity and this is related to differences in conformation or electronic properties. This procedure is generally unavoidable, because details of the receptors at an atomic level are not available.

The exception to this state of affairs is in enzymes where X-ray crystal structures are available and, together with NMR experiments, provide considerable atomic detail concerning binding sites. Since quantum mechanical molecular orbital programs require only nuclear coordinates, charge, and the number of electrons as data, it is, in principle, possible to incorporate the binding site into the calculation. In practice such an extended system too large to be computed with realistic computer budgets. However, following a suggestion of P. Kollman [4] an approximation to such a calculation is possible. The small binding molecule is treated fully quantum mechanically, but the atoms in the enzyme are replaced by dummy effective point atoms which have non-integral nuclear charges and no electrons. The increase in computer time involved when adding some 200–300 enzymic atoms to a substrate is only perhaps 5 per cent over that required for the substrate alone. The partial charges are found by separate computations on fragments, but do not seem to be critical providing each enzyme residue has the correct overall charge, generally neutral. Handling calculations of this type, which do seem capable of predicting differences in binding energies quantitatively [5], is again dependent on the use of computer graphics. It is necessary to view the site and to locate the small molecule in an approximate binding position before utilising calculations to optimize the binding. Once more, this part of the procedure is greatly facilitated if colour is available to distinguish the site from the binding partner.

Current developments

An advantage of the use of molecular orbital methods which has as yet been little exploited is that all calculated properties, such as charge distribution or electostatic potential, can be broken down into contributions from individual orbitals. In particular, the properties of the highest occupied molecular orbital may be particularly relevant if the molecule is invovled in donating electrons, while those of the lowest unoccupied orbital dominate electron acceptor behaviour. Orbital charge densities coupled to the binding energy of electrons in the orbital can be used in a quantitative form of frontier orbital theory [6].

The computer graphics displays of molecules, their surfaces, and properties, seem ideally suited to a new form of representation which will show a molecular surface of either a small molecule or an enzyme-binding site with a colour coded electrostatic potential on that surface. This should make it possible to highlight complementarity between binding partners. If similar displays are produced demonstrating the polarizability of external electrons then areas of molecules involved in binding by dispersion forces can be assessed quantitatively.

Throughout industry the combination of the tried techniques of theoretical chemistry allied to computer graphics displays are having an increasing impact. Theoreticians are no longer teased with questions of how many drugs they have developed, even though the answer is still none. On the other hand, theoreticians have in many instances contributed sufficiently to be included on patents alongside synthetic chemists.

References

[1] Hopfinger, A. J. 'Conformational Properties of Macromolecules' Academic press, 1973.
Alinger, N. L., *Adv. Phys. Org. Chem.* **13**, 1976.

[2] Quantum Chemistry Program Exchange Catalog., Guide and Index to Volume XIII 1981. (Available from QCPE, Department of Chemistry, Indiana University, Bloomington, Indiana 47405, USA.)

[3] Weinstein, H. L. In ACS Symposium No. 112, 'Computer-Assisted Drug Design' E. C. Olsen and R. E. Christoffersen American Chemical Society, 1979.

[4] Hayes, D. M., and Kollman, P., *J. Am. Chem. Soc.* **98**, 3335, 1976.

[5] Robins, S. A., Marchington, A. F., and Richards, W. G., *J. Am. Chem. Soc.* (in press).

[6] Elliott, R. J. Sackwild, V., and Richards, W. G., *J. Molec, Struct.* THEOCHEM. **86**, 301, 1982.
Elliott, R. J. and Richards, W. G., *J. Molec, Struct.* THEOCHEM. **87**, 247, 1982.
Elliott, R. J. and Richards, W. G., *J. Chem. Soc.* Perkin II. 943, 1982.

The photographs illustrating this article were produced at the Department of Molecular Biophysics, Oxford, on an Evans and Sutherland monochromatic display system. A multiple exposure and colour filtration technique was used to introduce selective colours into the image.