3

Simplified chain pictures of whole proteins

E J Milner-White
Dept. of Biochemistry, University of Glasgow, Glasgow G12 8QQ, UK

In collaboration with members of the Computing Science Department at Glasgow University, I and colleagues have developed programs for portraying simplified views of the chain structure of proteins. There are three sets of programs that correspond to three levels of computer sophistication. The first type of program runs on an unmodified BBC microcomputer. The second is for a BBC microcomputer with extra hardware in the form of a graphics board. The third currently runs on a Sigmex T5684 graphics terminal. They will be described in turn.

The programs that run on an unmodified BBC microcomputer[1] show the chain structure of a protein, or protein domain, of up to 250 amino acids. The structure is viewed as an averaged α-carbon plot. The averaging procedure, which causes α-helices to appear as wavy lines rather than as helices, and strands of β-sheet to appear as straight lines, has been described previously[2]. The plotting method was developed by Kemp[3]. As well as displaying the chain structure, this package also provides simplified views of the quaternary structure and the positions of any ligand binding sites. The latter pictures can be rotated in real-time by users, but each drawing of the chain pictures is fairly slow, say 1 min.

The programs that operate on a BBC microcomputer with extra hardware are part of a hardware and software package designed for BBC series B, BBC Master, or Apple IIe, microcomputers. It is commercially available from Chemdata Ltd., Grambla, Wendron, Helston, Cornwall TR13 0NQ, UK. This package was initially designed for the real-time rotation, manipulation and examination of molecules containing up to 250 atoms. As well as stick pictures, molecules can be represented as depth-sorted spheres or as dot-surfaces. Many other facilities exist. I have modified these programs so that the data can be transferred from the Protein Data Bank stored on a mainframe computer. In addition, the averaging procedure mentioned above has been utilized to allow whole proteins, or domains, of up to 250 amino acids to be displayed as averaged α-carbon plots. Rather as before, but to much greater effect because of the real-time rotation, the main features of the secondary and tertiary structure of the protein can be visualized.

The third type of molecular graphics display[4,5] illustrates all the intermain-chain hydrogen bonds in a protein. The pictures are on a Sigmex T5684 graphics terminal attached to a mainframe computer and using the Gino graphics package. The main chain is displayed as an averaged α-carbon chain plot rather as before. The intermain-chain hydrogen bonds are portrayed by drawing lines between the appropriate averaged α-carbon atoms, different colours being used to indicate their distance apart in the sequence (thus the hydrogen bonds in α-helices, where the CO of residue i bonds to the NH of residue $i + 4$, are drawn in red). In addition, where both NH and CO groups of one amino acid are bonded to the CO and NH groups of another, as found in antiparallel β-sheet, the pair of bonds is drawn as a thickened line. Hence, as well as displaying the different types of secondary structure, several other hydrogen-bonded motifs can be recognized. One outcome of this work has been the realization[6] that β-hairpins exist as four distinct classes that are of evolutionary and structural significance. Furthermore, certain loop motifs are typically found at the ends of each of the classes of hairpins[5,7].

References

1 **Mohamed, R and Milner-White, E J** Computer Graphics Forum (submitted) (1987)
2 **Milner-White, E J and Poet, R** *Biochem. Soc. Trans.* Vol 13 (1985) pp 793–795
3 **Kemp, G D** *Biochem. Ed.* Vol 14 (1986) pp 22–24
4 **Poet, R and Milner-White, E J** *Computer Graphics Forum* Vol 5 (1986) pp 211–215
5 **Milner-White, E J and Poet, R** *Trends Biochem. Sci.* (in press) (1987)
6 **Milner-White, E J and Poet, R** *Biochem. J.* Vol 240 (1986) pp 289–292
7 **Milner-White, E J** *Biochim. Biophys. Acta* Vol 911 (1987) pp 261–265

4

Computer modelling and materials science

C R A Catlow
Dept. of Chemistry, University of Keele, Staffs ST5 5BG, UK

Modelling methods are playing an increasingly important role in the study of materials. In particular, with the power of modern supercomputers, it is possible to make detailed predictions of the crystal structures of complex materials, of the behaviour of defects and impurities, of surface properties and of dynamical properties such as diffusion. The paper summarizes the methodologies of the current simulation techniques and reviews recent applications. Emphasis will be given to areas in which graphical display is important. Accounts of recent work on zeolite catalysts, electronic ceramics, superionic conductors and reactor fuels are included.

5

Geometric intersection problems in computational chemistry

M Connolly
Dept. of Chemistry, New York University, 4 Washington Place, New York, NY 10003, USA

Computational chemistry has two branches. The first is concerned with physics and includes quantum chemistry, electrostatics, molecular mechanics and molecular dynamics. The second is concerned with solid geometry and includes protein-packing studies (cube algorithms, Voronoi polyhedra), crystallographic symmetry, molecular volume computation, solvent-accessibility and computer graphics. Geometric methods

often require the computation of the intersections of geometric objects, such as spheres and polyhedra.

Three such methods are discussed in this presentation: (1) a hidden-line elimination algorithm for stacked planar contours, (2) a protein surface shape measurement algorithm which intersects a sphere with a polyhedron and (3) an analytical molecular volume algorithm for the union of a set of intersecting spheres (atoms). Generalizations about geometric intersection algorithms are inferred from these three examples.

6

Transparallel applications in molecular graphics

D N J White
Dept. of Chemistry, University of Glasgow, UK

There has been a considerable growth of interest in parallel processing by the computational chemistry community over the last couple of years. Despite this interest, almost all the practical laboratory experience so far has been confined to Intel Hypercubes. The performance of early Hypercubes has been disappointing, although later versions incorporating a vector processor at every node should lead to increased performance. Although machines like the Hypercube are inexpensive when compared with supercomputers, they are still expensive in absolute terms (c. £200 000) and offer no performance advantages over conventional array processors such as the FPS-164.

Since these early developments, an exciting new contender has appeared in the form of the Inmos transputer. The transputer is cheap (approximately £350 for the 1.5 Mflop T800-20), ubiquitous, and very easy to design into large arrays. At the time of writing, only 0.15 Mflop T414-20 integer transputers are readily available, and the only fully supported language is OCCAM, although β-releases of FORTRAN 77, C and PASCAL are available.

The most interesting question regarding the transputer is whether the manufacturers' publicity and the undeniable promise can be translated into a system suitable for large-scale molecular graphics calculations.

Our experiments with parallel energy minimization calculations indicate that the advent of the transputer is a major landmark in the history of computational chemistry. A 64 node T800-20 system fits comfortably inside a 19 × 16 × 18 in desktop rack and offers computational performance equivalent to a Cray-1S for energy minimization and other codes. Such a system can cost well under £100 000.

The authors will discuss commercially available hardware, the programming of transputer arrays for molecular-graphics calculations, and the efficiencies of alternative interconnect topologies.

7

Model building studies in supramolecular chemistry

G Wipff
Dept. of Chemistry, University of Strasbourg, France

Supramolecular chemistry, at the frontier of organic and inorganic chemistry, biology and physics, involves intermolecular interactions and conformational processes of flexible units, leading to the recognition of 'substrates' by 'receptors'. From a theoretical point of view, this field raises challenging problems, such as the design and construction of structures suitable for complexation (e.g. with adequate cavities), energy representation and optimization, and the treatment of solvent effects on structure and selectivity.

First, we recall the limits of two basic working hypotheses used for empirical force-field calculations: transferability of fragment properties (structural and electronic features, reactivity), and the additivity of energy components. Next, different types of empirical-energy calculations in the field of macrocyclic receptors will be presented. Molecular mechanics is used for 18-crown-6 as a flexible receptor for alkali cations, and as an anchoring site for ammonium derivatives; here the graphics system is essential for building starting structures.

The graphics system brings a major contribution in the analysis of Monte Carlo simulations on 18-6 in a water cluster: not only the hydration energy, but also the structure of water around the 'solute' depend on the conformation of the solute. Such a pattern has been confirmed since by X-ray crystallography.

Visualization of dynamic structures, from molecular dynamics simulations and normal modes of vibration, is also illuminating. Particularly, low frequency modes of vibration for cryptands and cryptates lead to the opening of the cavity. A similar process has been found recently for the active site of several enzymes. It thus provides a pathway for substrate inclusion.

There is no way of predicting computationally whether or not a substrate will bind to a receptor in solution. However, a new technique, involving stepwise perturbations from one system to another, may give relative free energies of binding.

8

Microcomputer graphics in biochemical education

A G Booth
Dept. of Biochemistry, The University, Leeds, UK

It is virtually impossible to study any biological subject at the undergraduate level without coming into contact with biochemistry at some stage. Biochemistry classes therefore tend to be large and contain students with very different backgrounds. Many students chose biological subjects simply 'to get away from maths' and these students often find such subject areas as enzyme kinetics difficult to grasp.

Other problems are posed by trying to teach a rapidly evolving subject at a time of dwindling resources. In particular, many modern techniques in biochemistry are too time-consuming, too expensive, too dangerous or simply impractical to actually carry out in the teaching laboratory.

For these reasons, we are developing simulation programs that are being integrated into the courseware. These programs rely heavily on graphics for their effectiveness. For example, in teaching enzyme kinetics, animated displays show enzyme substrate, product and