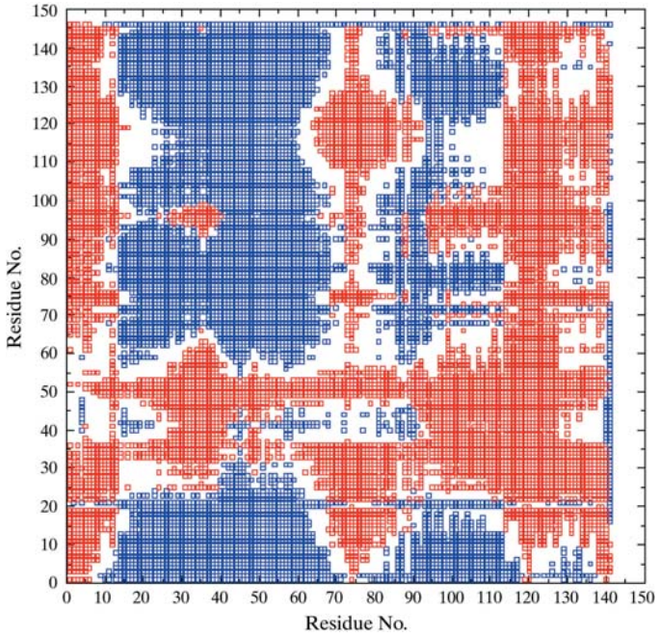

Fig. 17.2.3.9. A difference distance matrix plot of the α_1–β_2 interface of haemoglobin in the T to R transformation. The *x*-axis reperesents the α_1 subunit and the *y*-axis the β_2 subunit. Red points indicate residues that are closer following the transformation and blue points indicate residues that move farther apart. Plot by Raj Srinivasan, Johns Hopkins University.

visualization is that it is coordinate-frame independent. Thus two structures can be compared for features without first superposing their coordinates in the same frame. This approach also works well when comparing two different structures of the same molecule, where there may be some movement between the two. By computing the distance matrix for each structure, and then computing the difference between the two distance matrices, the resulting difference distance matrix will indicate those parts of the structure that stay in the same relative relationships and those that may move relative to each other (Fig. 17.2.3.9).

Animating trajectories of molecular structures and changes in volumetric properties over time is one way to look for trends and patterns in molecular dynamics and other time-course simulations. However, other modes of information visualization can assist analysis and communication of results, sometimes more effectively. Plotting an array of small images showing the time course of key properties can reveal patterns that may be difficult to see in a trajectory. For instance, using the program *MolMol* (Koradi *et al.*, 1996), the time course of the seven nucleic-acid backbone torsion angles during a dynamics simulation of an RNA polynucleotide can be plotted on a circular graph (starting from the centre and progressing outward) to uncover patterns of change and correlation between a large number of variables over time.

In addition to the enormous amount of information generated by computational simulations of molecular dynamics, dockings and other multi-structure, multi-modal techniques, the floodgates of molecular information have opened, gushing data from genomics and high-throughput structure determination. Thus, the need for novel visualization methods has become even more acute. Circle maps defining genomic structure at various levels of detail and annotation have become a common graphical form for organizing and communicating the positional and functional aspects of genome structure. Aligned nucleic acid or amino-acid sequences coded by conservation, chemical property, or any number of other functional relationships have become the *lingua franca* of gene hunters and gatherers. As the protein structure database continues its exponential growth, the opportunities for defining and refining structural family relationships abound. Developing methods for effectively visualizing the relationships that arise from all-by-all computational comparisons of the entire database is an important current challenge in molecular graphics.

17.2.4. Presentation graphics

Much of molecular graphics can be classified as working or 'throw-away' graphics. Typically this involves the interactive creation of graphical representations on screen or paper that are used in the course of research to build, modify and analyse molecular structures, their motions and interactions. Such graphics need only be intelligible to the researchers involved. On the other hand, a presentation or publication graphic must be able to stand on its own to convey information to a broader audience. Thus it requires additional thought and work in its creation. It is unfortunate that many scientists simply capture the working graphics on their computer screens for use in publications or other forms of communication. Both interpretability and intelligibility may suffer badly if a number of issues are not considered in the production of a presentation graphic: What is the medium of publication? What is the main point of the graphic? Who are the target audience? How will reproduction or the viewing environment affect the impact of the graphic? Even seemingly simple issues such as when and how to use colour can in reality be a complex mixture of aesthetics, psychology, technology and economics. While an in-depth discussion of these issues is beyond the scope of this chapter, it is worthwhile to look at two categories of publication graphics, illustration and animation, in this context.

17.2.4.1. *Illustration*

In print media, shaded colour images can present a number of difficulties. In addition to the issue of cost, colour shifting, reproducibility and loss of detail in the half-toning process may lead to less than the desired result. Simple line art is an effective way to bypass many of these complications. Since the advent of printing in the middle ages, artists and scientists have explored the problems of creating illustrations within the limitations of the printing process. Over time, artists have built a vocabulary of outlines, hatched shading and varied textures to simplify and clearly portray an object. While the creation of such illustrations was time consuming and required considerable artistic talent, they effectively portrayed the observational science of the day. The advent of computers and computer graphics removed any requirement for skilled hand draftsmanship in the production of molecular representations, but did not solve all of the problems of good illustration. As mentioned above, prior to the widespread use of interactive computer graphics, molecular structures were often published as outline drawings of ball-and-stick models using programs such as *ORTEP* or *PLUTO* (Motherwell & Clegg, 1978). More recently, programs such as *MOLSCRIPT* (Kraulis, 1991) have re-established the popularity of line-art illustration in the molecular realm. A good *ORTEP* drawing usually took a great deal of preparation time in order to get the best representation and viewpoint to display the structure effectively. As the visual repertoire of molecular structure has expanded to a wide variety of shapes including ribbons, tubes and solvent-based surfaces, the challenge of automating the general illustration process has grown. A number of techniques have been developed in the computer-graphics community to generate images in the style of technical or artistic illustrations (Fig. 17.2.4.1). These approaches use lighting, depth information and geometry to produce black-and-white drawings with shapes defined by silhouette lines and cross-hatched

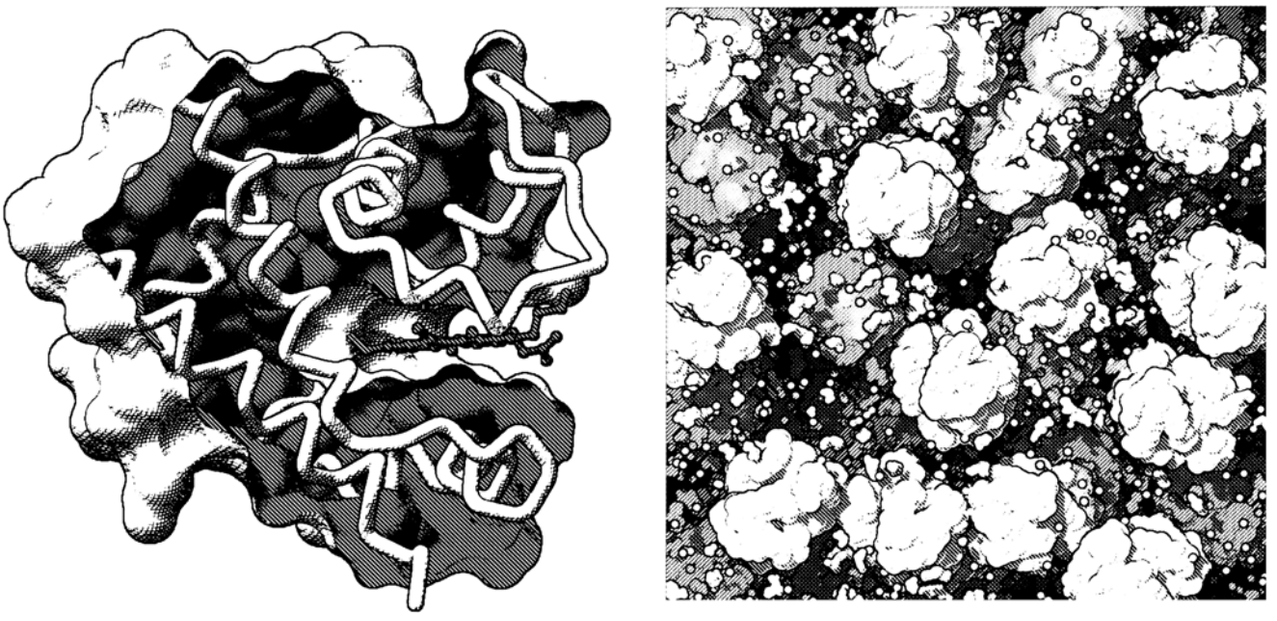

Fig. 17.2.4.1. Line-art molecular illustration. The figure on the left depicts the α-carbon backbone and molecular surface of the α subunit of haemoglobin (PDB code: 2hhb). Outlines define the shapes of the surfaces and tubes, and contour lines enhance their three-dimensionality. Hatch lines, normally used for shadows, are used here to darken the inside of the molecular surface. The figure on the right shows the interior of a human red blood cell (300 $\mathring{A}^3$), showing all molecules except water. The picture is drawn with outlines defining each molecule. Shadows and depth cuing are used to enhance the three-dimensional character of the image. Contour lines are not used. From Goodsell & Olson (1992).

shading, and details shown by a variety of textures. *MOLSCRIPT* has used some of these techniques for ribbon and ball-and-stick renderings. More general applications of these approaches to molecular illustration have also been described by Goodsell & Olson (1992).

A significant advantage of digital black-and-white illustration is the efficiency of representation. Since each picture element takes only a single bit of information (black or white), and since there are typically large areas that are of constant value, these images can be compressed, stored, transmitted and printed very efficiently. Thus, with the advent of electronic web publication, such illustrations represent an attractive alternative to full colour. These same characteristics represent significant advantages for the digital transmission and use of animated sequences as well.

17.2.4.2. *Animation*

Computer-graphic molecular animations began to appear in the late 1960s. Recording directly off their vectorscope, Levinthal and colleagues in Project Mac produced a record of an interactive molecular modelling session in 1967. In the early 1970s, a number of molecular animations were produced to convey new scientific results. Wilson at UC San Diego showed vibrational modes of small molecules in a film produced frame-by-frame on a vectorscope. Parr & Polyani painstakingly filmed pen plotter drawings of space-filling diatomic molecules to animate a bimolecular chemical reaction. Sussman & Seeman produced a black-and-white vector animation of the dinucleotide UpA structure in 1972 by recording directly off a vectorscope. Seeman, Rosenberg & Meyserth produced a more ambitious molecular animation in 1973 entitled *Deep Groove*, which depicted the structure of double helical segment ApU CpC and its implications for more extended DNA geometry. This film was shot in colour, using a monochrome vectorscope and multiple exposures through a colour filter wheel. Around this time, Knowlton, Cherry & Gilmer at Bell Labs used early frame-buffer devices to display and animate patterns of crystal growth based upon aggregation of spheres. In the mid-1970s, Porter & Feldman had developed a scan-line based CPK representation for raster displays and had animated molecular structures, and Langridge and co-workers had taken up recording off the black-and-white vector displays then available. By the end of the 1970s, Max had produced high-quality animations of DNA using a high-resolution Dicomed

Fig. 17.2.4.2. A stereolithographic model of betalactamase inhibitory protein (BLIP). The model depicts the shape of the molecule as represented by a spherical-harmonic approximation of the solvent-excluded surface. Inside the surface is a hollow tube which follows the α-carbon trace of the protein backbone. The model was designed by Michael Pique and Arthur Olson, and fabricated on a 3D systems SLA 1000 by Beryl Hodgson.

film writer (Max, 1983) and Olson had used an early colour vector display from Evans and Sutherland to produce an eight-minute animation depicting the structure of tomato bushy stunt virus (Olson, 1981). By the early 1980s, animation projects became more ambitious. Olson produced large-screen OmniMax DNA and virus animation segments for Disney's EPCOT center in 1983. Max produced a red–blue stereo OmniMax film for Fujitsu entitled *We Are Born of Stars*, which included a continuous scene depicting the hierarchical packaging of DNA from atoms to chromosomes, based on the best current model of the time.

Computer-graphics animation has presented both great potential and significant challenges to the molecular scientist wishing to communicate the results of structural research. Animation can not only enhance the depiction of three-dimensional structure through motion stereopsis, it can show relationships through time, and demonstrate mechanism and change. The use of pans, zooms, cuts and other film techniques can effectively lead the viewer through a complex scene and focus attention on specific structures or processes. The vocabulary of film, video and animation is familiar to all, but can be a difficult language to master. While short animations showing simple rotations or transitions between molecular states, or dynamics trajectories, are now routinely made for video or web viewing, extended animations showing molecular structure and function in depth are still relatively rare. The time, tools and expertise that are required are not generally available to structural researchers.

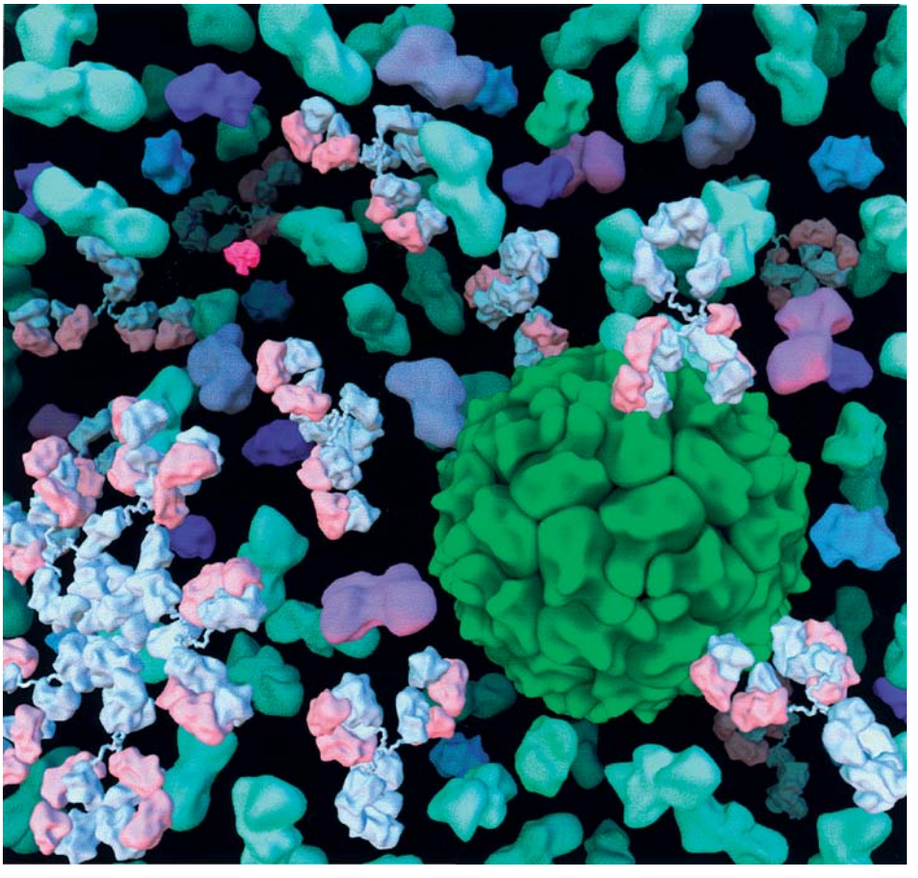

Fig. 17.2.5.1. This image represents a volume of blood plasma 750 Å on a side. Within the three-dimensional model, antibodies (Y- and T-shaped molecules in light blue and pink) are binding to a virus (the large green spherical assembly on the right), labelling it for destruction. It shows all macromolecules present in the blood plasma at a magnification of about 10 000 000 times. This model is composed of over 450 individual protein domains, ranging in size from the 60 protomers making up the poliovirus to a single tiny insulin molecule (in magenta). The model was constructed using atomic level descriptions for each molecule, for a total of roughly 1.5 million atoms. Detailed surfaces were computed for each type of protein using *MSMS* by Michel Sanner and then smoothed to a lower resolution using the *HARMONY* spherical-harmonic surfaces developed by Bruce Duncan. The model geometry contains over 1.5 million triangles.

17.2.4.3. *The return of physical models*

While the use of physical models of molecules has largely been replaced by computer graphics, new computer-driven rapid-prototyping technologies which originated in the manufacturing sector have begun to be utilized in the display of molecular structure. A number of 'three-dimensional printing' methods have been developed to build up a physical model directly from a computational surface representation of an object (Burns, 1954). One of the earliest methods, stereolithography, uses a resin which is polymerized when exposed to laser light of a given wavelength. The laser is passed through a vat of the liquid resin and is lowered, layer-by-layer as it plots out the shape of the object (Fig. 17.2.4.2). Other approaches build up layers of paper or plastic through lamination or deposition. These methods have been used by a number of scientists to produce various representations of molecular structure (Bailey *et al.*, 1998). The ability to hold an accurate representation of a molecular surface in one's hand and feel its shape can give great insight, not only to people with visual impairments, but to anyone. Moreover, when one is dealing with processes such as docking and assembly, these physical models can add a haptic and manipulative appreciation of the nature of the problem. While at this point colour has not been implemented in these technologies, there remains the promise that such automated production of molecular models will enhance the communication and appreciation of molecular structure.

17.2.5. Looking ahead

Moore's law has already delivered on the promise of three-dimensional graphics capability for the desktop and laptop. The internet and World Wide Web have made molecular structure data and display software available to the masses. Have molecular graphics reached a stage of maturity beyond which only small incremental changes will be made?

The Human Genome Initiative and high-throughput structure determination are beginning to change the scope of the questions asked of molecular modelling. Prediction of function, interactions, and large-scale assembly and mechanism will become the dominant domain of molecular graphics and modelling. These tasks will challenge the capabilities of the hardware, software and, particularly, the user interface. New modes of interacting with data and models are coming from the computer-graphics community. Molecular docking and protein manipulation using force-feedback devices have been demonstrated at the University of North Carolina (Brooks *et al.*, 1990). The same team has developed a 'nanomanipulator' which couples a scanning atomic force microscope with stereoscopic display and force-feedback manipulation to control and sense the positioning and interactions of the probe

references