A molecular fragment with an unsatisfied valency (by which it might later be attached to another such group) would have that feature represented by a 'vanishing virtual' atom which removes the need for any organizational distinction between such fragments and complete molecules. Fragments, such as amino-acid residues, may be assembled from atoms, and molecules may be assembled from atoms and/or from such fragments invoked by name, by the superposition and elimination of the relevant vanishing virtuals. The assembly process includes the development of a connectivity tree for the molecule and provision is made for the 'turning' or reconstruction of such trees if the combination of such fragments redefines the root atom of one or more of the fragments. The system also provides for ring closure. Model building initially uses fixed bond lengths and angles, varying only the dihedral angles in single bonds, but has a library of preformed rings which could not otherwise be modelled on the simple basis. The results of such modelling may then be subjected to an energy-minimization routine using a steepest-descent method in the space of the dihedral angles and referring to the Lennard–Jones potential for non-bonded atom pairs. Atoms are first sorted into contiguous cubes so that all neighbours of any atom may be found by searching not more than 27 cubes.

The system is also capable of modelling by reference to electron density either by the translation and rotation of molecular fragments and the rotation of rotatable bonds within them or by the automatic linking of peaks in an electron-density map which are separated by less than, say, 1.8 Å, which is an important aid to interpretation when the resolution is sufficiently high.

3.3.3.2.2. *GRIP*

This system, developed by Professor F. P. Brooks, Dr W. V. Wright and associates at the University of North Carolina, Chapel Hill, NC, USA, was designed for biopolymers and was the first to enable a protein electron-density map to be interpreted *ab initio* without the aid of mechanical models (Tsernoglou *et al.*, 1977). Girling *et al.* (1980) give a more recent example of its use.

In its 1975 version *GRIP* is a three-machine system. Centrally there is a minicomputer which receives inputs from the user and controls a vector display with high-speed matrix-multiplication capability. The third machine is a mainframe computer with high-speed communication with the minicomputer.

The system develops a polymer chain from a library of monomers and manipulates it through bond rotations or free rotations of fragments explicitly specified by the user, with the aid of dials which may be coupled to bonds for the purpose. Bond rotations in the main chain either rotate part of the molecule relative to the remainder, which may have undesirable long-range effects, or the scope of the rotation is artificially limited with consequential discontinuities arising in the chain. Such discontinuities are removed or alleviated by the mainframe computer using the method of Hermans & McQueen (1974), which treats atomic position vectors, rather than bond rotations, as independent variables.

The system made pioneering use of a two-axis joystick to control orientation and a three-translation joystick to control position.

3.3.3.2.3. *Barry, Denson & North's systems*

These systems (Barry & North, 1971; North *et al.*, 1981; North, 1982) are examples of pioneering work done with minicomputers before purpose-built graphics installations became widespread; examples of their use are given by Ford *et al.* (1974), Potterton *et al.* (1983) and Dodson *et al.* (1982). They have the ability to develop a polymer chain in sections of several residues, each of which may subsequently be adjusted to remove any misfit errors where the sections overlap. Manipulations are by rotation and translation of sections and by bond rotations within sections. These movements are directly controlled by the user, who may simultaneously observe on the screen the agreement with electron density, or calculated estimates of potential or interaction energy, or a volume integral of the product of observed and model densities, or predicted shifts of proton magnetic resonance spectra. Thus models which are optimal by various criteria may be constructed, but there is no optimizer directly controlling the rotational adjustments which are determined by the user.

One of the earliest applications of them (Beddell, 1970) was in the fitting of substrate molecules to the active site of lysozyme using difference electron densities; however, the systems also permitted the enzyme–substrate interaction to be studied simultaneously and to be taken into account in adjusting the model.

3.3.3.2.4. *MMS-X*

The Molecular Modelling System-X (*MMS-X*) is a system of purpose-built hardware developed by Barry, Marshall and others at Washington University, St Louis. Associated with it are several sets of software. The St Louis software consists of a suite of programs rather than one large one and provides for the construction of a polymer chain in helical segments which may be adjusted bodily to fit the electron density, and internally also if the map requires this too. Non-helical segments are built helically initially and unwound by user-controlled rotations in single bonds. The fitting is done to visual criteria. An example of the use of this system is given by Lederer *et al.* (1981).

Miller *et al.* (1981) have described an alternative software system for the same equipment. Functions invoked from a keyboard allow dials to be coupled to dihedral angles in the structure. Their system communicates with a mainframe computer which can deliver small blocks of electron density to be contoured and stored locally by the graphics system; this provides freedom of choice of contour level at run time. An example of the use of this system is given by Abad-Zapatero *et al.* (1980).

3.3.3.2.5. *Texas A&M University system*

This system (Morimoto & Meyer, 1976), a development of Meyer's earlier system (Section 3.3.3.1), uses vector-graphics technology and a minicomputer and is free of the timing restrictions of the earlier system. The system allows control dials to be dynamically coupled by software to rotations or translations of parts of the structure, thus permitting the re-shaping or re-positioning of the model to suit an electron-density map which may be contoured and managed by the minicomputer host. The system may be used to impose idealized geometry, such as planar peptides in proteins, or it may work with non-idealized coordinates.

The system was successfully applied to the structures of rubredoxin and the extracellular nuclease of *Staphylococcus aureus* (Collins *et al.*, 1975) and to binding studies of sulfonamides to carbonic anhydrase (Vedani & Meyer, 1984). In addition, two of the first proteins to be constructed without the aid of a 'Richards' Box' were modelled on this system: monoclinic lysozyme in 1976 (Hogle *et al.*, 1981) and arabinose binding protein in 1978 (Gilliland & Quiocho, 1981).

3.3.3.2.6. *Bilder*

This system (Diamond, 1980*a*,*c*, 1982*b*) runs on a minicomputer independent of any mainframe. It builds a polymer chain from a library of residues and adapts it by internal rotations and overall positioning in much the same way as previous systems described in this section. Like them, it can provide user-controlled bond rotations, but its distinctive feature is that it has an optimizer within the minicomputer which will determine optimal combinations of bond rotations needed to meet the user's declared

objectives. Such objectives are normally target positions for atoms set by the user by visual reference to the density, using the method of Section 3.3.1.3.9, but they may include target values for angles. These latter may either declare a required shape that is to take precedence over positional requirements, which are then achieved as closely as the declared shape allows, or they may be in least-squares competition with the positional requests. The optimizer also recognizes the constraints imposed by chain continuity and enables an internal section of the main chain to be modified without breaking its connection to the rest of the molecule. Similar techniques also allow ring systems to adopt various conformations, by bond rotation, without breaking the ring, simultaneously permitting the ring to have target positions. The optimizer is unperturbed by under-determined situations, providing a minimum-disturbance result in such cases. All these properties of the optimizer are generated without recourse to any 'special cases' by a generalization of the subspace section technique which was used to maintain chain continuity in a 'real-space-refinement' program (Diamond, 1971). This is based entirely on the rank of the normal matrix that arises during optimization, which may serve to satisfy a constraint such as chain continuity or ring closure and simultaneously to establish what degrees of freedom remain to be controlled by other criteria. In *Bilder* this is achieved without establishing eigenvalues or eigenvectors. The method is described in outline in Section 3.3.2.2.1 and in detail by Diamond (1980*a*,*b*).

The angular variables used normally comprise all single bonds but may include others, such as the peptide bond with or without a target of 180°. Thus this bond may be completely rigid, elastic, or completely free. Any interbond angles may also be parameterized but at some cost in storage. The normal mode of working is to develop a single chain for the entire length of the molecule, but if cumulative error makes fitting difficult a fresh chain may be started at any stage. *Bilder* may itself reconnect such chains at a later stage.

Construction and manipulation operates on a few residues at a time within the context of a polymer chain, but any or all of the rest of the molecule, or other molecules, may be displayed simultaneously.

Contouring is done in advance to produce a directoried file of contoured bricks of space, each brick containing up to 20 independently switchable elements which need not all be from the same map. Choice of contour level and displayed volume is thus instantaneous within the choices prepared.

The system is menu driven from a tablet, only file assignments and the like requiring the keyboard, and it offers dynamic parallax as an aid to 3D perception (Diamond *et al.*, 1982). Bloomer *et al.* (1978), Phillips (1980), and Evans *et al.* (1981) give examples of its use.

3.3.3.2.7. *Frodo*

This system, due to Jones (Jones, 1978, 1982, 1985; Jones & Liljas, 1984), in its original implementation was a three-machine system comprising graphics display, minicomputer and mainframe, though more recent implementations combine the last two functions in a 'midi'. Its capabilities are similar to those of *Bilder* described above, but its approach to stereochemical questions is very different. Where *Bilder* does not allow an atom to be moved out of context (unless it comprises a 'chain' of one atom) *Frodo* will permit an atom or group belonging to a chain to be moved independently of the other members of the chain and then offers regularization procedures based on the method of Hermans & McQueen (1974) to regain good stereochemistry. During this regularization, selected atoms may be fixed, remaining atoms then adjusting to these. A peptide, for example, may be inverted by moving the carbonyl oxygen across the peptide and fixing it, relying on the remaining atoms to rearrange themselves. (*Bilder* would do the equivalent operation by cutting the chain nearby, turning the peptide explicitly, reconnecting the chain and optimizing to regain chain continuity.) The *Frodo* approach is easy to use especially when large displacements of an existing structure are called for, but requires that ideal values be specified for all bond lengths, angles and fixed dihedrals since the system may need to regain such values in a distorted situation. *Bilder*, in contrast, never changes such features and so need not know their ideal values.

Frodo may work either with consecutive residues of a polymer chain, useful for initial building, or with a volume centred on a chosen position, which is ideal for adjusting interacting side chains which are close in space but remote in sequence.

In recent implementations *Frodo* can handle maps both in density grid form and in contour form and permits on-line contouring. It has also been developed (Jones & Liljas, 1984) to allow the automatic adjustment of the position and orientation of small rigid groupings by direct reference to electron density in the manner of Diamond (1971) but without the maintenance of chain continuity, which is subsequently reintroduced by regularization.

Horjales and Cambillau (Cambillau & Horjales, 1987; Cambillau *et al.*, 1984) have also provided a development of *Frodo* which allows the optimization of the interaction of a ligand and a substrate with both molecules being treated as flexible.

3.3.3.2.8. *Guide*

Brandenburg *et al.* (1981) have described a system which enables representations of macromolecules to be modified with reference to electron density. Such modifications include rotation about single bonds under manual control, or the movement with six degrees of freedom, also under manual control, of any part or parts of the molecule relative to the remainder. The latter operation may necessitate subsequent regularization of the structure if the moved and unmoved parts are chemically connected, and this is done as a separate operation on a different machine. The system also has the capability of displaying several molecules and of manually superimposing these on each other for comparison purposes.

3.3.3.2.9. *HYDRA*

This program, due to Hubbard (1985) (and, more recently, to Molecular Simulations) has several functional parts, referred to as 'heads', which all use the same data structure. The addition of further heads may be accomplished, knowing the data structure, without the need to know anything of the internal workings of existing heads.

The program contains extensive features for the display, analysis and modelling of molecular structure with particular emphasis on proteins. Display options include dotted surfaces, molecular skeletons, protein cartoons and a variety of van der Waals, ball-and-stick, and other raster-graphics display techniques such as ray tracing and shaded molecular surfaces. Protein analysis features include the analysis of hydrogen bonding, and of secondary and domain structure, as well as computational assessment of deviations from accepted protein structural characteristics such as abnormal main-chain or side-chain conformations and solvent exposure of hydrophobic amino acids. A full set of protein modelling facilities are provided including homology modelling and the 'docking' of substrate molecules. The program contains extensive tools for interactive modelling of structures from NMR or X-ray crystallographic data, and provides interfaces to molecular-mechanics and dynamics calculations. There are also database searching facilities to analyse and compare features of protein structure, and it is well suited to the making of cine films.

references