Cyrus Levinthal, the Kluge and the origins of interactive molecular graphics

Eric Francoeur

In the mid-1960s, a group of scientists at Massachusetts Institute of Technology, led by Cyrus Levinthal, took hold of one of the early interactive graphics terminals and used it to visualize, study and model the structure of proteins and nucleic acids. From this encounter between cutting-edge computer technology and molecular biology emerged the crucial elements for the development of a research-technology field known today as interactive molecular graphics. The following account is not only about how computer graphics technology has literally changed the way scientists view the molecular realm, but also a look at how an epistemic and institutional space was created to integrate this technology into scientific research.

The capacity to visualize and study the three-dimensional structures of molecules is of central importance in many areas of science, including chemistry, biochemistry, molecular biology and pharmacology. In the 20th century, a variety of graphical and physical means to represent these structures was developed for this purpose[1]. Over the past three decades, scientists have increasingly relied on computer graphics to create dynamic representations of molecular structures, using a technique known as interactive molecular graphics. The advantages of this technique revolve not only around the capacity to display the structure of molecules, but also around the capacity to interact with and transform this display to gain different perspectives on a given structure, highlight specific features or occlude others. Whereas some molecular graphics packages are simple aids to the visualization of molecular structures, others act as a 'front-end' for molecular dynamics simulation systems – as an interface between the user and the simulation algorithms. There are currently dozens of molecular graphics packages available, catering to all levels of user and types of computer platform, and some of the more popular packages are available for free[2].

The earliest efforts to combine the representation of molecular structures with the dynamic properties of a computer display can be traced back to the high-tech environment of the Massachusetts Institute of Technology (MIT) in the mid-1960s. This is where Cyrus Levinthal, professor of biophysics, and his collaborators took advantage of one of the earliest interactive graphics systems to develop a visually oriented, computer-based 'molecular-model building system'.

This peculiar development, in which cutting-edge computer technology and the particular problems of visualizing and modeling the enormous structure of proteins intersected to give rise to a solution, which in a few years became a field of scientific practice in its own right.

Eric Francoeur

Holds a PhD in Sociology from McGill University and is currently a postdoctoral fellow at the Max-Planck Institute for the History of Science in Berlin. He specializes in the study of visualization and modeling practices in the sciences, particularly chemistry, biochemistry and molecular biology. He is currently working on a book on the design and use of molecular models in chemistry.
e-mail: efranc65@yahoo.com

Interactive computing at MIT in the early 1960s

In the 1960s, MIT was a frontrunner in the development of computer technology. In 1963, it became the host site of Project MAC ('Machine-Aided Cognition' or 'Man-and-Computer'), a unique time-shared computer system, which for the first time provided a community of researchers with the possibility of real-time interactive computing, mainly through a series of teletype terminals scattered around the campus[3].

MIT was also at the forefront of computer graphics technology, thanks largely to a Computer-Aided Design research program sponsored by the Air Force[4]. This program notably gave rise to Sketchpad, the first and archetypal computer drawing system, designed by Ivan Sutherland[5], and to the Electronic Systems Laboratory Display Console, otherwise known as the 'Kluge'[6]. The Kluge was the first computer terminal able to show three-dimensional objects on a cathode-ray tube, through axonometric projection. Three-dimensional perception was achieved by rotating the displayed object on the screen, with the user controlling the rate of rotation of the image through a trackball-like device known as the 'globe'. The user could interact with the displayed objects through a variety of interfaces, notably buttons and a light-pen. A vector-based display, it could only represent white lines on a black background. Although limited by today's standards, it had the familiar trappings of a graphics workstation and eventually became one of the terminals available to Project MAC users. One of those users, Cyrus Levinthal, would put the Kluge to tasks that nobody had foreseen.

0160-9327/02/$ – see front matter PII:S0160-9327(00)01428-9

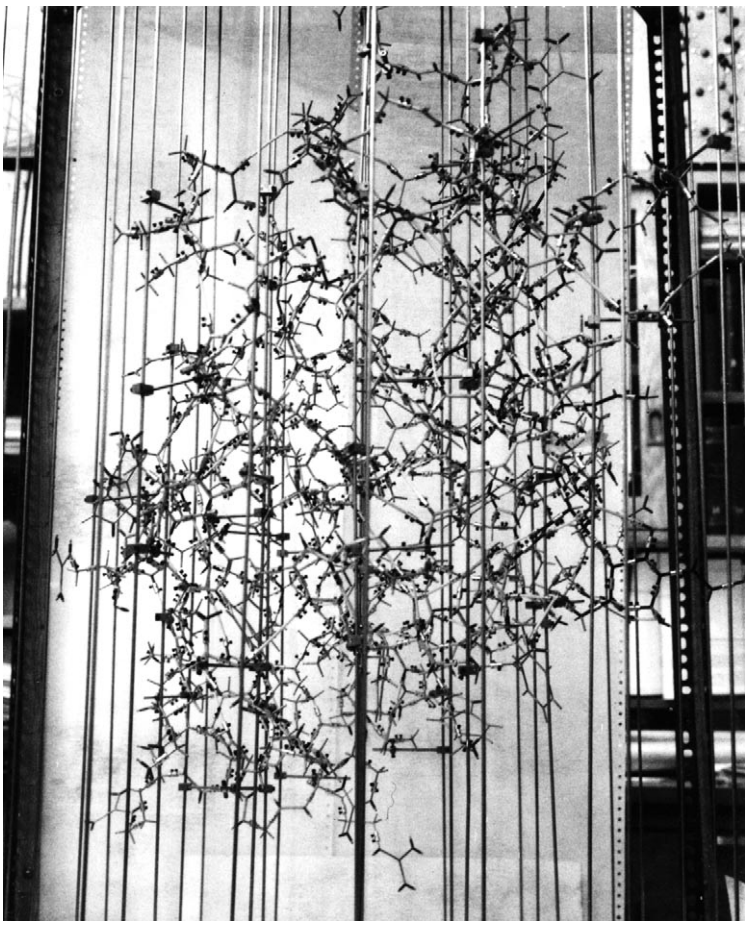

Figure 1. Skeletal model of the structure of the protein lysozyme, attached to a framework of vertical rods (courtesy of C. David Barry).

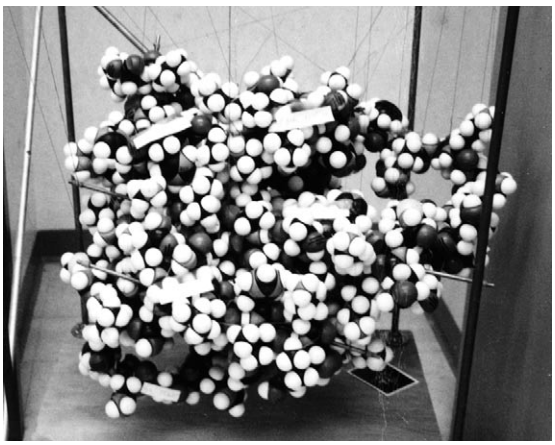

Figure 2. Space-filling model of the structure of the protein myoglobin. The model is hanging from wires that can be barely distinguished against the background. Photograph under the model, in the lower-right corner, gives an idea of the size of the model (courtesy of C. David Barry).

Cyrus Levinthal meets the Kluge

Levinthal was not, at first glance, the dream user for Project MAC. Although originally trained as a physicist and reportedly adept at mathematical modeling, the molecular biologist nevertheless later admitted that he 'started with a rather strong distaste for computers and computing; at that time it did not seem that they could be useful to biology'[7]. In late 1963, Robert Fano, the director of Project MAC, introduced Levinthal to the Kluge. Its capacity to produce images changed Levinthal's mind. He became, in his own words, immediately enthused at the idea of employing this system to create models of protein structures.

To make sense of Levinthal's enthusiasm, it helps to understand the state of the art in the representation of macromolecular structures at the time. Although various forms of symbolic and graphical representations were in use, the study of macromolecular structures often required the use of molecular models (i.e. physical structures assembled from modular components which 'map out' in three-dimensions the structure of a given molecule). By the 1960s, the use of such models was well-established in the chemical and biochemical sciences.

Molecular models were suitable for small molecules, but building the physical model of a macromolecular structure, which might contain hundreds or thousands of atoms, let alone exploring its various possible conformations, was at best very difficult. The models had to be supported by metal rods (Figure 1) or suspended by wires (Figure 2), and even then could easily be displaced (or sag out of place), greatly reducing the accuracy of the resulting structure. Collapsing models were a common occurrence, and the suggestion of building space-filling models underwater to reduce or cancel the effects of gravity was floated once or twice. The problem was widely recognized[8] and in the early 1960s, the National Institutes of Health (NIH), the National Science Foundation and the American Society of Biological Chemists were involved in developing space-filling models more suitable for macromolecular structures. The result of this work, the Corey–Pauling–Koltun (CPK) models, made macromolecular modeling marginally easier[9]. Levinthal had first-hand experience of the problem: 'At that time several MIT associates and I were trying to do molecular modeling as an aid in thinking about intracistronic genetic complementation. Our models kept falling down and we were having all the usual problems associated with models'[10]. In the Kluge, Levinthal saw a possible solution to what was perceived at best as an annoyance and at worst as a serious hindrance to research.

Molecular model building by computer

Assisted by Project MAC personnel, Levinthal set out to learn programming, and within a few months was able to develop a set of programs to construct, display and analyze macromolecular structures in real time on the Kluge. The result was described as a new, exciting way of looking at molecular structures; Levinthal and some of his collaborators 'went on to study proteins, protein crystals and whatever structural data we could get our hands on'[11]. The representation of the structures was very schematic, composed uniquely of white lines representing the bond between the atoms (Figure 3). The light-pen and buttons were used to interact with the displayed structure. The Kluge was hardwired to the Project MAC computer (a modified IBM 7904), which took care of all the calculations. This original set-up was changed in 1966 when the Kluge was replaced by a DEC 340 display – a commercial product it had inspired – connected to a 'satellite'

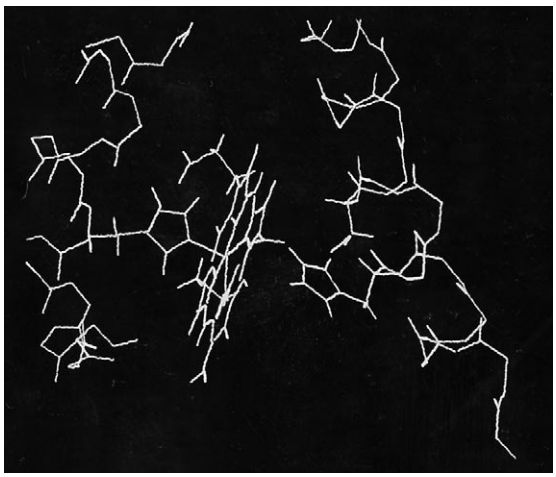

Figure 3. Photograph of the Kluge display showing a detail of the myoglobin structure, namely the heme group of myoglobin viewed sideways with portions of the surrounding polypeptide chain. Such photographs (mainly Polaroid snapshots) were the principal means of obtaining a 'hard copy' of the displayed image (courtesy of Martin Zwick).

Figure 4. The second generation of hardware for Levinthal's molecular model building system: in the back, the PDP-7 mini-computer; in the middle, the standard teletype programming console; in the foreground, the DEC 340 graphics display console, with the 'globe' that served to control the direction and rate of rotation of the image. The space-filling model of an unspecified molecule stands next to the globe (courtesy of Martin Zwick).

PDP-7 mini-computer, which took over some of the computational tasks from the IBM 7904 (Figure 4).

Over the next few years, a dozen researchers, postdocs and graduate students gravitated around the computer-based molecular model building system, working on specific projects or contributing to its maintenance and improvement. These collaborators included Martin Zwick, a graduate student in biology with a background in physics; Robert Langridge (then at the Harvard Medical School and the Children's Cancer Research Foundation), an x-ray crystallographer with computing experience; C. David Barry and Edgar F. Meyer, postdoctoral fellows with backgrounds in physics and x-ray crystallography, respectively; and Stephen A. Ward, William R. Brody and David E. Avrin, all graduate students in Electrical Engineering.

The visual and the numeric were closely intertwined in the combination of real-time computing and graphic interface: the user was able not only to witness but also, through the display, to interact with the computations, and if necessary to control, correct and tweak them. Although this approach to computing is common now, it was radically new at a time when batch processing was still by far the norm[12]. The display was particularly useful for debugging: a single glance at the display could reveal a badly programmed structure, which would have been hard to detect in a list of numbers. For Robert Langridge, who used the system to refine crystallographic data of DNA structures, it meant faster processing times and doing away with cumbersome models to visualize the resulting structures[13].

The combination of numeric and the visual in the study of molecular structures comes forth clearly in Levinthal's project of using the interactive graphic display to address the problem of protein folding[14]. In the early 1960s, it had become widely accepted that the linear sequence of amino acids in a protein was sufficient to predict its native (functional) three-dimensional structure. The prediction of protein structure from amino-acid sequence was conceived as a potentially interesting alternative to determining these structures by x-ray crystallography, then a long and complex process. The thermodynamic hypothesis of protein folding held that of all the possible conformations of its polypeptide chain, the native structure of a protein had the lowest internal energy. Checking systematically for all the possible conformations of the polypeptide chain of a protein to find the one with the lowest energy term was not an option, as the number of such possible conformations was (and still is) simply beyond the means of any available computing power. By the mid-1960s, some researchers had started to devise computer programs in which short virtual polypeptide chains were submitted to energy minimization algorithms[15]. Such algorithms produced conformations with lower total energy, but they could only be expected to alter the structure to the bottom of a local energy minimum, rather than reach the global energy minimum corresponding to the conformation of the protein in its native state. Levinthal believed that using the graphic display to combine the skills and knowledge of a user with the analytical capacity of the computer might provide a solution to this specific problem. The key was to have the user identify why the conformation was stuck at a particular local minimum and make alterations to the structure so as to return the algorithm to a 'downhill' path. Levinthal characterized this interactive work as a combination of 'manual manipulation and energy minimization'[16]. He attempted to predict the structure of cytochrome C, a protein whose structure Richard Dickerson was in the process of determining crystallographically at the California Institute of Technology.

This prediction attempt failed completely. A 'plausible' structure was proposed, but it did not prove to be unique – several equally plausible structures could be postulated, rendering the whole process dubious. When Dickerson finally solved the structure of cytochrome C, it became clear that the predicted structure was completely wrong. In retrospect, one can point to many reasons for this failure, not least the fact that Levinthal (like many of his peers) had underestimated by a wide margin the complexity of the problem. Although some progress has been made, the results of protein-structure prediction methods are today still far from the accuracy provided by experimental methods[17]. Yet, all this work was not in vain: from Levinthal's cogitations on protein folding came two short and very influential notes on the concept of 'folding pathways', the idea that specific local interactions between atomic groups guide the folding process of proteins[18].

Interactive molecular graphics after MIT

In 1967, Levinthal left MIT to become chair of the Department of Biology at Columbia University. At that point, any effort to use the computer graphics resources of MIT for macromolecular analysis effectively ceased. At the time of the move, Levinthal abandoned his work on protein structure, because, as he commented later, 'it seemed to me that most of the activities which were directed at predicting structure from sequence data were, at the time, more game playing than serious science'[19]. This particular failure did not dampen his enthusiasm for the application of computer graphics to scientific research. It would, on the contrary, become his main center of interest in the following years.

There could be several reasons for this. First, the simple capacity to visualize macromolecular structures with an ease and flexibility that other means did not afford was in itself an achievement, which, in the eyes of many, warranted further development. Second, although it had failed in the ambitious project of protein-folding prediction, the molecular model building system had proved useful in more mundane and straightforward analytical tasks. Third, molecular graphics held many promises, the most prominent of which was to make the whole process of structure refinement in protein x-ray crystallography much easier, by providing a means for virtual model building. The fourth, and perhaps most important reason, was simply the availability of research funding for molecular graphics work.

As the work at MIT was taking place, at the NIH William Raub was in charge of developing what eventually became the PROPHET system – a time-shared, interactive, graphics-oriented database-management system for laboratory and clinical scientists studying the relationship between molecular structure and biological function[20]. Told of Levinthal's work, Raub paid him a visit at MIT. After a demonstration of the molecular graphics system, Raub became 'entranced' with the technology[21], and the idea of including the capacity to display molecular structures to the nascent PROPHET system was formed. The NIH entered into a contract with Levinthal to extend the capabilities for three-dimensional representation and display of macromolecules. The contract financed the acquisition of an Adage interactive graphics terminal[22], which formed the core of Levinthal's computer graphics facility at Columbia, a facility that eventually became a NIH-sponsored National Research Resource. By the early 1970s, Levinthal had put aside 'wet' molecular biology to focus on the application of computer graphics to scientific research. He continued his work on molecular graphics[23], and went on to develop a computer graphics system for the three-dimensional reconstruction of biological structures from serial section microphotography[24].

Many of those who collaborated with Levinthal at MIT maintained an interest in computer graphics. In 1969, Robert Langridge established, with Todd Wipke, the Computer Graphics Laboratory at Princeton as another NIH National Research Resource Facility. The initial grant of over US$1 million allowed the acquisition of an Evans and Sutherland LDS-1 graphic terminal, by far the most powerful equipment of its kind at the time. Langridge moved this facility to the University of California in San Francisco in 1976, and managed its operations until his retirement in 1994. C. David Barry brought the experience he had acquired at MIT to Washington University and Oxford, UK, where he helped to develop molecular graphics facilities[25]. Edgar Meyer took on a faculty position at Texas A & M and became a research collaborator at the Brookhaven National Laboratory, where he used the Brookhaven Raster Display to produce three-dimensional images of molecular structures and contributed to the development of the Protein Data Bank[26]. He later established the Biographics Laboratory at Texas A & M.

Molecular graphics activity was also spreading outside this inner circle. William V. Wright, a PhD student at the Department of Computer Science at the University of North Carolina (Chapel Hill), developed, in the early 1970s, an interactive computer graphics system for molecular studies[27], and spent the subsequent years at UNC directing a NIH-funded project for the development of computer graphics systems to study molecular structures. By 1971, the NIH had its own molecular graphics facility, developed by Richard Feldmann, from its Division of Computer Research and Training[28].

From the association of molecular biology and cutting-edge computer technology at MIT in the 1960s emerged a new research technology, along with a group of scientists who in various institutions played a key role in establishing and developing interactive computer graphics as an approach to the study of molecular structures – pioneers of a field now known as 'scientific visualization'[29]. For them, molecular graphics was not simply a means to an end, but had become a field of research and development in its own right. The subsequent development of the field was by no means spectacular, but was nevertheless steady. By 1974, 19 computer modeling systems with molecular graphics had been described in the literature[30], a fair achievement considering the cost of each of these facilities, not only simply in terms of hardware, but also in terms of software development – there was no

off-the-shelf molecular graphics software, and basic hardware incompatibilities made it difficult to 'export' software from one platform to another. By then, the field was gaining a certain autonomy: it had its practitioners; its centers, funding sources – the nucleus of a culture centered around computers and computer graphics technology. Research funding agencies, not surprisingly, were a key factor in this development. The NIH, in particular, financed the creation and maintenance of many early molecular graphics and modeling facilities.

The development of a computer graphics industry, which provided the necessary hardware, also proved crucial. The nascent molecular graphics community was not involved directly in the technological and commercial development of computer graphics technology, but constituted a group of software developers and 'power users' who sought the state-of-the-art – whenever they could afford it – and pushed the available technology to its limit. The interpretation and display of x-ray crystallographic data (in particular for elucidating protein structure) remained the principal area of research and application for interactive molecular graphics throughout the 1970s[31]. By the early 1980s, the sustainable use of color had been introduced in the field[32] and drug design had become another prominent application[33]. The Molecular Graphics Society was created in 1982 and its journal, the *Journal of Molecular Graphics*, first appeared in 1983, marking a turning point in the institutionalization of the field[34]. By the late 1980s, developments in computing had reached a stage that made molecular graphics available on the desktop of every researcher who could afford a PC. By the 1990s, molecular graphics had come of age, a widespread, polyvalent and taken-for-granted research technology.

The digitized version of a 16mm film showing the graphic capacity of Levinthal's molecular-model building system at MIT can be seen on the Web, at the Early Molecular Graphics Movie Gallery (http://purl.org/efranc65/movie).

References and notes

1 Francoeur, E. (1997) The forgotten tool: the design and use of molecular models. *Social Studies of Science* 27, 7–40

2 See, for example, Protein Explorer, a web-based protein-structure visualization system (http://molvis.sdsc.edu/protexpl/index.htm) and DeepView, a package developed and maintained by researchers for GlaxoSmithKline (http://www.expasy.ch/spdbv/)

3 Fano, M. and Corbató, F.J. (1966) Time-sharing on computers. *Sci. Am.* 215, 128–140

4 Coons, S.A. (1963) An outline of the requirements for a computer-aided design system. *AFIPS Conference Proceedings* 23, 299–304

5 Sutherland, I.E. (1963) Sketchpad: a man-machine graphical communication system. *AFIPS Conference Proceedings* 23, 329–346

6 Stotz, R. (1963) Man-machine console facilities for computer-aided design. *AFIPS Conference Proceedings* 23, 323–328. In hacker culture, the term 'kluge' refers to an inelegant yet effective solution to a software or hardware problem. It is not clear how this epithet became attached to the Electronic Systems Laboratory Display Console.

7 Levinthal, C. (1976) Introduction: a brief history of computer graphics. In *Workshop on Computer Graphics in Biology*, pp. 1–5, Verve Research Corporation

8 Platt, J.R. (1960) The need for better macromolecular models. *Science* 131, 1309–1310

9 Koltun, W.L. (1965) Precision space-filling atomic models. *Biopolymers* 3, 665–679

10 Levinthal, C. op. cit. note 7

11 Levinthal, C. (ca. 1989) *The Beginnings of Interactive Molecular Graphics*, unpublished manuscript

12 It was possible to produce 'non-interactive' drawings of molecular structures using off-line computing, as demonstrated by Carroll K. Johnsons' popular ORTEP program, created in the same period. See Johnson, C.K. (1965) *ORTEP: A FORTRAN Thermal Ellipsoid Plot Program for Crystal Structure Illustrations*, Oak Ridge National Laboratory

13 Langridge, R.A. and MacEwan, A.W. (1965) The refinement of nucleic acid structure. In *IBM Scientific Computing Symposium on Computer Aided Experimentation*, IBM

14 Levinthal, C. (1966) Molecular model-building by computer. *Sci. Am.* 214, 42–52

15 Scheraga, H.A. *et al.* (1967) Calculations of polypeptide structures from amino acid sequence. In *Conformation of Biopolymers* (Vol. 1) (Ramachandran, G.N., ed.), pp. 43–60, Academic Press

16 Levinthal, C. (1966) Computer construction and display of molecular models. In *IBM Scientific Computing Symposium on Computer Aided Experimentation*, IBM

17 Baker, D. and Sali, A. (2001) Protein structure prediction and structural genomics. *Science* 294, 93–96

18 Levinthal, C. (1968) Are there pathways for protein folding? *Journal de Chimie Physique* 65, 44–45. Levinthal postulated the folding pathways in view of the fact that proteins fold relatively rapidly into their native state, despite the astronomical number of possible conformations that their polypeptide chains could theoretically take. This discrepancy became rapidly and widely known as the 'Levinthal paradox' in the protein-folding literature, a formulation which Levinthal did not particularly care for.

19 Levinthal, C. (ca. 1989) op. cit. note 11

20 Raub, W.F. (1975) The prophet system. In *Computers in Life Science Research* (Siler, W. and Lindberg, W.A.B., eds), pp. 189–194, FASEB and Plenum Press

21 William Raub, interview with the author, 20 June 2001

22 The Adage graphics terminal was one of the many commercial interactive graphics systems that appeared on the market in the mid-to-late-1960s. See Machover, C. (1978) A brief, personal history of computer graphics. *Computer* 11, 38–45

23 Katz, L. and Levinthal, C. (1972) Interactive computer graphics and the representation of complex biological structures. *Annu. Rev. Biophys. Bioeng.* 1, 465–504

24 Levinthal, C. *et al.* (1974) Computer-aided reconstruction from serial sections. *Fed. Proc.* 33, 2336–2340

25 Barry, C.D. and North, A.C.T. (1971) The use of a computer-controlled display system in the study of molecular conformations. *Cold Spring Harbor Symp. Quant. Biol.* 36, 577–584

26 Meyer, E.F. (1997) The first years of the Protein Data Bank. *Protein Sci.* 6, 1–7

27 Wright, W.V. (1972) An interactive computer graphics system for molecular studies, PhD Thesis, University of North Carolina-Chapel Hill

28 Feldmann, R.J. *et al.* (1972) An interactive, versatile, three-dimensional display, manipulation and plotting system for biomedical research. *Journal of Chemical Documentation* 12, 234–236

29 McCormick, B. *et al.* (1987) Visualization in scientific computing. *Computer Graphics* 21, 1–15

30 Marshall, G.R. *et al.* (1974) Computer modeling of chemical structures: applications in crystallography, conformational analysis, and drug design. In *Computer Representation and Manipulation of Chemical Information* (Wipke, W.T. *et al.*, eds), pp. 203–237, John Wiley & Sons

31 Jones, A.T. (1978) A graphic model building and refinement system for macromolecules. *J. Appl. Crystallogr.* 11, 268

32 Langridge, R.A. *et al.* (1981) Real-time color graphics in studies of molecular-interactions. *Science* 211, 661–666

33 Hopfinger, A.J. (1984) Computational chemistry, molecular graphics and drug design. *Pharmacy International* 5, 224–228

34 Morffew, A.J. (1983) Editorial. *J. Mol. Graph.* 1, 3–4