13 February 1981, Volume 211, Number 4483 SCIENCE

Real-Time Color Graphics in Studies of Molecular Interactions

Robert Langridge, Thomas E. Ferrin, Irwin D. Kuntz
Michael L. Connolly

Summary. Studies of the structures and interactions of large biological molecules require both coordinate data and three-dimensional visualization. Orthodox molecular models often bear a tenuous relationship to the coordinate data. In contrast, computer graphics requires that the display directly and accurately represent the data, and storage of modified configurations and recovery of original structures are simple. Software has been developed that allows real-time display of color line and surface displays of several interacting molecules, while quantitatively monitoring the stereochemistry.

Most problems in molecular structure and function require visualization, and as the size of the molecules increases, the third dimension becomes crucial. Space-filling or wire models are satisfactory up to a certain level of complexity, but purely mechanical problems cause serious difficulties since the model on the bench and the list of coordinates in the computer are not necessarily closely related (especially after the model is degraded by many curious hands). Particularly difficult is the restoration of a structure after simple modifications.

With computer graphics, the display and the data are directly related, storage of prior configurations is simple, and pieces do not fall off. Drawings for machine design or architecture—other fields in which computer graphics is a significant tool—are simple by comparison with large biological molecules. Consider also the need for a variety of forms of display, such as stick models or electron densities, representation of chemical properties, such as charges and solvent interactions, and visualization of the detailed interactions between large molecules, and the complexity of the problem is clear. For studies with these requirements real-time color graphics has unique advantages. Such a system is discussed in this article.

Computer Graphics

We began to apply interactive three-dimensional computer graphics to large biological molecules in late 1964, using the Electronic Systems Laboratory display at Project MAC (now the Laboratory for Computer Science) at the Massachusetts Institute of Technology. The display system was based on an IBM 7094 computer running the CTSS operating system. In spite of the limitations of this early system, real-time three-dimensional rotation of the displayed object was possible for the first time, and applications were made to protein and nucleic acid structures and to direct visualization of energy minimization; the enormous potential was evident (*1*).

Applications of interactive three-dimensional graphics to molecular structure have been reviewed by Feldmann (*2*), who reiterated that "Scientist users working with molecules on a display find it very important to be able to rotate the image of the molecule." To present a three-dimensional object on a two-dimensional screen, a series of projections are calculated as the object is rotated. The coordinate transformations require several multiplications and additions for each end point, and the key to smooth representation of complex objects is to implement this in hardware. This is most efficiently done with homogeneous coordinates, using a 4 by 4 matrix representation for the transformations. Rotation, translation, scaling, and perspective matrices may then be concatenated and handled as a single entity (*3*).

Simple manipulations such as bond rotations can obviously be modeled in a computer; the main advantage is the ability to retain the modified coordinates. In the system discussed here (*4*), the bond rotations are computed very rapidly with the built-in matrix multiplier, and the host computer is called on only to compute derived functions such as interatomic distances. But if a computer merely reproduces a real three-dimensional model, much of the potential power is unused. The simple bond rotations of a wire model can be extended to bond stretching, bond angle changes, and online calculation of contact distances and energies, limited only by the speed and size of the computer, the efficiency of the algorithms, and the imagination of the user.

When building a wire model it is usually necessary to construct the entire assembly and then to concentrate on the parts of special interest. In contrast, with a computer representation, although the complete molecule is available, it is rarely necessary to look at all of it; instead, one can reproduce in detail only those parts of immediate concern. Much of the work in producing a usable com-

Robert Langridge is a professor of pharmaceutical chemistry and director of the Computer Graphics Laboratory, School of Pharmacy, and a professor of biochemistry and biophysics, School of Medicine, University of California, San Francisco 94143. Thomas Ferrin is a specialist and lecturer in pharmaceutical chemistry and facility manager of the Computer Graphics Laboratory, School of Pharmacy, University of California, San Francisco. Irwin Kuntz is a professor of pharmaceutical chemistry, School of Pharmacy, University of California, San Francisco. Michael Connolly is a graduate student in the Department of Biophysics, University of California, Berkeley 94720.

0036-8075/81/0213-0661$01.50/0

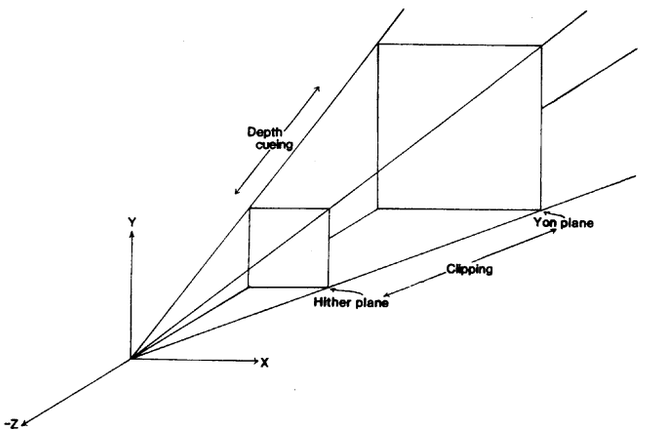
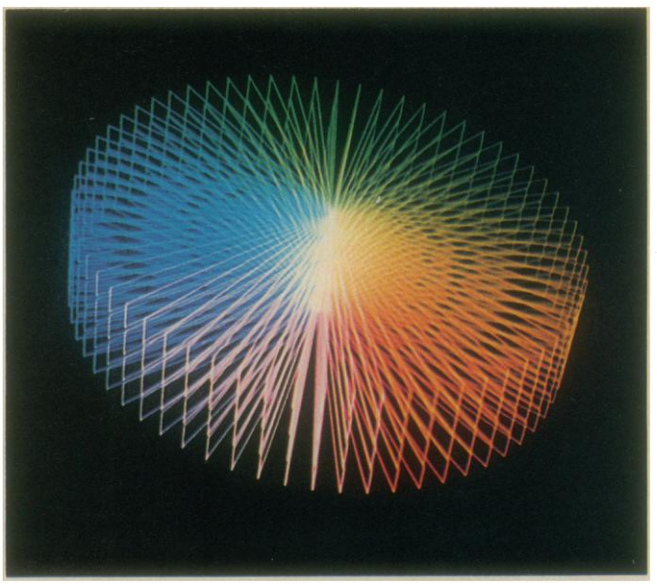


Fig. 1 (left). Pyramid of vision with the three-dimensional clipping planes indicated [see Newman and Sproull (*3*)]. Fig. 2 (right). Direct photograph from the color display of the color wheel illustrating the colors available and the intensity depth cueing.



puter graphics system for molecular modeling involves the development of representations and interactions that would be impossible with space models. Interactively selecting and manipulating only the interesting parts of a molecule may seem easy in principle, but for a system designed for the user unsophisticated in computers this is as important and difficult as producing the correct pictures on the screen.

Line Drawing and Raster Displays

There are two main forms of computer graphic displays—calligraphic, or line-drawing, and raster, or "TV-like" systems. The display we use is a calligraphic one that is capable of rapid transformations of three-dimensional data with arbitrary combinations of rotation, translation, scaling, and perspective. Real-time clipping is performed not only on the top, bottom, and sides of the picture space, but also in the depth (conventionally *Z*) axis. This is associated with a "viewing pyramid," which includes a hither (near) boundary, yon (far) boundary, and eye position (Fig. 1). The intensity of the cathode-ray tube (CRT) electron beam may be controlled in proportion to the distance between the hither and yon planes, giving a strong sense of depth to the picture. The combined effects of real-time perspective, clipping, and depth cueing provide a powerful means of exploring complex structures without the cost of extensive software.

Raster systems, in contrast, build the picture by scanning across the CRT screen, with each picture element, or pixel, being controlled from a "refresh" memory. Since this scanning is similar to that used in standard television receivers, this method of computer display can draw on the large technology base developed from home television, including the production of shaded and fully colored pictures (*3*). Unfortunately, the refresh memory must be large, a quarter-million elements for a typical 512 by 512 pixel grid. Each element of the memory must contain enough information to define the color and intensity of the pixel. Typically, this requires at least 8 bits. Updating this amount of memory in real time in order to create dynamic interaction is a substantial and at present expensive problem (*3*). For this reason, a calligraphic system is still superior for truly interactive studies of complex systems best represented by line drawings.

Even with a line-drawing system there are other drawing modes besides continuous lines, and the use of dots covering a surface in combination with the depth cueing and color gives excellent three-dimensional representations of molecular surfaces. We have developed software (*5*) operating on the Bell Laboratories UNIX time-sharing system (*6*), which allows real-time display of color line and surface displays of several interacting molecules, while quantitatively monitoring the stereochemistry.

Interaction

Display interaction is controlled by manipulation of various input devices, including switches, knobs, joysticks, and a tablet. These devices transmit data to the computer which can be used in any manner, but the user interface demands that their functions be reasonable. Typically, we assign three translations to one of the three-function joysticks and three rotations (usually about fixed axes) to the other joystick. The tablet is used to choose functions from a menu displayed on the screen, and the switches to select molecules or parts of molecules for manipulation.

In addition to these input devices, we have a viewer for generating stereoscopic images. Left and right perspective views are presented alternately. When they are viewed through a synchronized shutter, each of the observer's eyes sees only its associated image, and the result is perceived as a stereoscopic image with a strong sense of depth of field. The actual shutter is adapted from a commercially available stereo image alternator (*7*).

Full utilization of these facilities greatly aids the information processing ability of the user, but one recent addition enormously advances the interpretability of complex objects—color.

Color

The addition of color to a molecular display considerably increases the information content, as illustrated by the work of Feldmann (*8*), Max (*9*), and their co-workers on color surface displays. But these are raster systems used for single-frame displays or the production of animated movies and are not truly interactive, real-time systems.

We made color films of molecular displays from an earlier monochromatic line-drawing system in 1970 to 1974 (*10*), using computer-controlled red, green, and blue filters and exposing appropriately for the desired colors. Real-time color was possible only with a synchro-

nized, but awkward, version of rotating color filters called a lorgnette. In principle an excellent idea, it did not prove satisfactory in use.

Most line-drawing systems are monochromatic. Color monitors based on beam penetration techniques are in use, but they suffer from the disadvantages of poor intensity, low speed, and limited range of color, which make them unsuitable for displays of complex molecular structures. Recently, a color monitor was introduced for a line-drawing system that was based on standard color television shadow mask technology (*4*). The system includes the electronics to operate the color monitor in line-drawing mode. It has the advantages of providing bright, high-resolution, full-color pictures, coupled with well-established techniques for real-time manipulation of line-drawing images.

We applied this system to real-time molecular interaction studies, and although we were confident of the value of color, it has turned out to be greater than our most optimistic estimates. Even though the calligraphic system does not permit shaded surface drawings, the use of the various drawing modes (dots and lines), depth cueing, perspective, clipping, and dynamic rotation permits the equivalent of almost all varieties of display possible with a raster system, but in real time and with immediate interaction.

A new command added to the list of display instructions sets the hue and saturation values so that different portions of the picture may be arbitrarily displayed in different colors. The hue controls the relative amounts of red, green, and blue primary colors in the picture, while different saturation levels are used to create "pastels." The 64 hue levels available and the depth cueing are shown in the color wheel in Fig. 2. This may be rotated and the hither and yon planes manipulated in real time, giving a strong sense of three dimensionality.

To illustrate the value of color in an interactive line-drawing display, we use carboxypeptidase and an inhibitor as an example (*11*) (Fig. 3). Tyrosine 248 is rotated about the bond between the α and β carbons, and the distance from the terminal OH of tyrosine 248 to an atom in the glycine-tyrosine inhibitor is monitored. This bond rotation is directly linked to an interactive knob and the rotation and distance calculation are done in real time. By interactive selection, only the lines connecting the α carbons are displayed, together with the interacting parts of the inhibitor binding site, before and after rotation of tyrosine 248 (see Fig. 3).

Surfaces

Since molecules interact at their surfaces, it is important to display these surfaces, and in particular to display that part of the surface which is accessible to other molecules. This is usually equivalent to defining the part of the surface accessible to a water molecule—that is, the solvent-accessible surface.

When this term was introduced by Lee and Richards (*12*), they defined it as the surface traced out by the center of a sphere of radius 1.4 angstroms, representing a water molecule, as it rolls over the target molecule. This definition was improved on by Richards (*13*) when he defined the molecular surface as the surface traced out not by the center of the probe, but by the inward-facing surface of the probe sphere. When the probe is touching a single atom this is equivalent to the outward-facing van der Waals surface of that atom. Richards calls this the contact surface. The other part of the molecular surface—what Richards calls the reentrant surface—is not part of the van der Waals surface of any atom, but

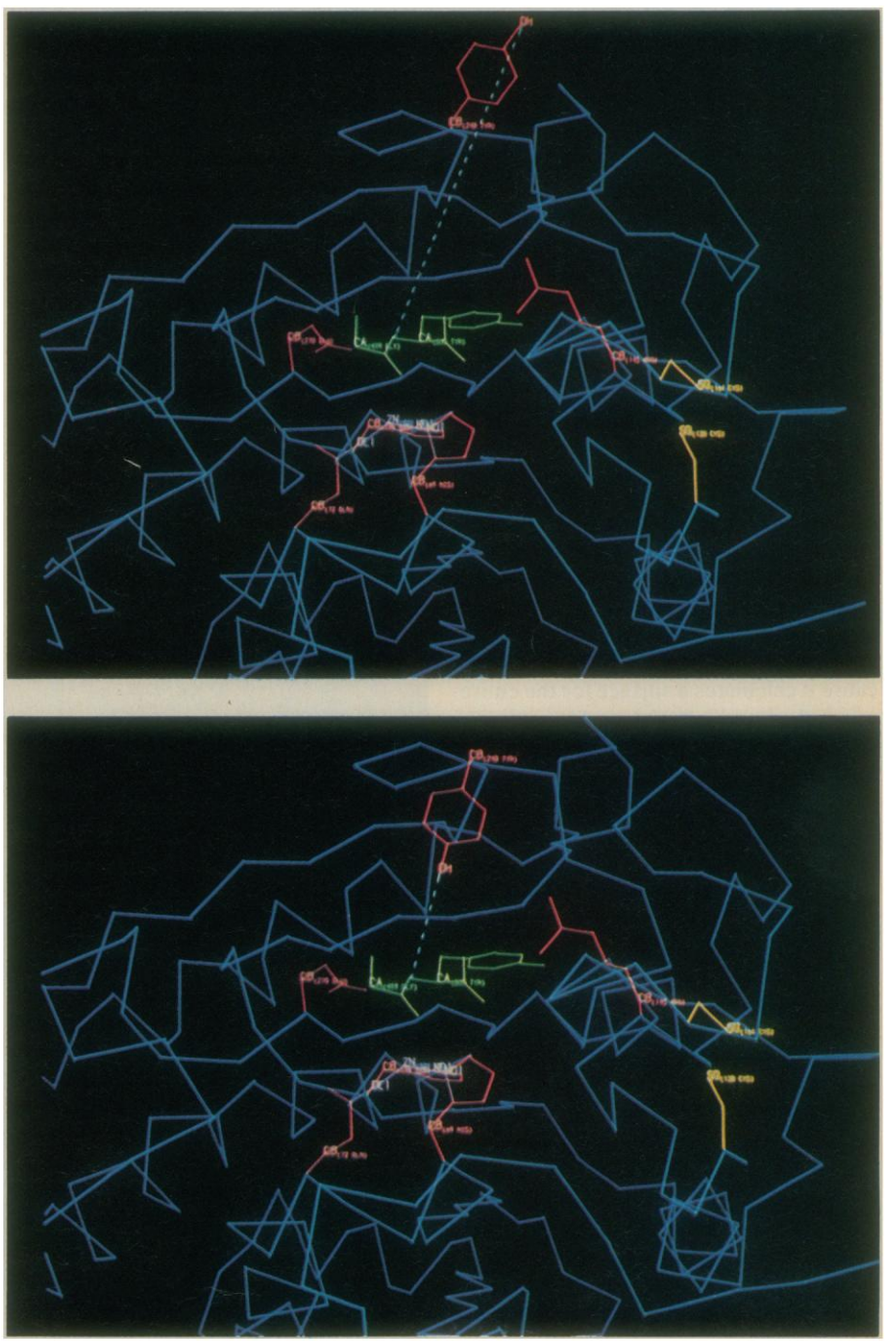


Fig. 3. (Top) Carboxypeptidase A. Color coding: background α-carbon chain, blue; amino acids in the active site, red; cysteine, yellow; zinc and associated atoms, magenta; and inhibitor, green. (Bottom) As above, after rotation of tyrosine 248.

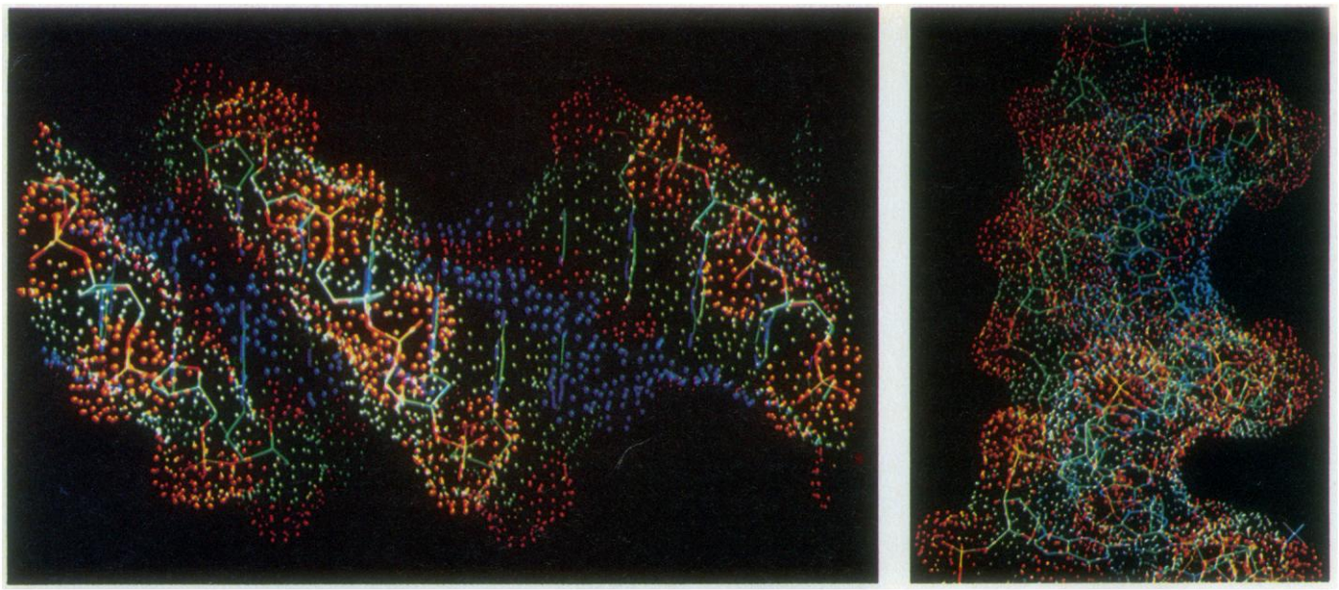


Fig. 4. The B form of DNA, as shown on the cover, but rotated to view down the major and minor grooves (left) and with the helix axis vertical (right). Color coding: carbon, green; nitrogen, blue; oxygen, red; and phosphorus, yellow.

corresponds to the inward-facing surface of the probe when it is simultaneously in contact with more than one atom. The molecular surface thus resembles the van der Waals surface of the atoms, except that interstices too small to accommodate the probe are eliminated, and clefts between atoms are smoothed over.

An algorithm to calculate the molecular surface (*14*), originally developed as part of a protein docking study, was adapted for use with interactive color graphics. We find it of greater value than an earlier effort (*15*) to implement Richards' molecular surface definition, because it calculates a surface for the entire molecule, not just the surface of one face, and also because undersurfaces of overhangs, deep invaginations, and internal cavities are properly represented.

The algorithm produces coordinates for several thousand points distributed over the molecular surface. These are precomputed, which may take 30 to 60 minutes on our computer (*4*) for a small protein. Each point is associated with a specific atom, so the point can be colored according to atom type or a related property. The dots are distributed at a density of 2 to 5 per square angstrom and can be plotted on paper, but the most striking display, as with other three-dimensional objects, is made with a fully equipped computer graphics system with interactive real-time depth cueing, clipping, rotation, stereo, and color. Given a system limitation of about 15,000 dots, this permits the display, at atomic resolution, of molecules up to the size of small proteins or of selected portions of larger structures. Although the two-dimensional pages of a journal cannot present this adequately, the cover picture and Fig. 4 provide views of DNA (*16*) with both the surface and the skeleton displayed. On the cover, the molecule is viewed directly down the axis. In Fig. 4 (left), the model is turned, by using the interactive real-time controls, to an angle corresponding to a view down the major and minor grooves, which are clearly seen. In Fig. 4 (right), the DNA axis is vertical and the interactive depth cueing turned to a maximum, with the hither and yon planes moved to contact the front and back of the molecule. This gives a strong sense of depth; with the real-time display it is still more striking.

The hardware clipping planes also allow interactive removal of the front surface of a molecule to reveal the back surface as seen from the inside. This kind of interactive sectioning also reveals numerous internal cavities, corresponding to the inner molecular surface of the pro-

tein. Our survey indicates that the arrangement of these cavities is nonrandom and that many of them are occupied by water molecules (*14*). This system provides a remarkably clear visualization of the molecular surfaces of these interior cavities; it is not possible to study them easily on a system lacking hither and yon clipping planes, because the outer surface gets in the way. Similarly, many binding sites are partially obscured because they lie in angled invaginations, between domains, or between subunits. By using real-time rotation, scaling, and clipping, one can obtain a close-up view of any part of a binding site, from any angle, and either from a position in the solvent looking inward at the protein or from a position in the protein looking outward toward the solvent.

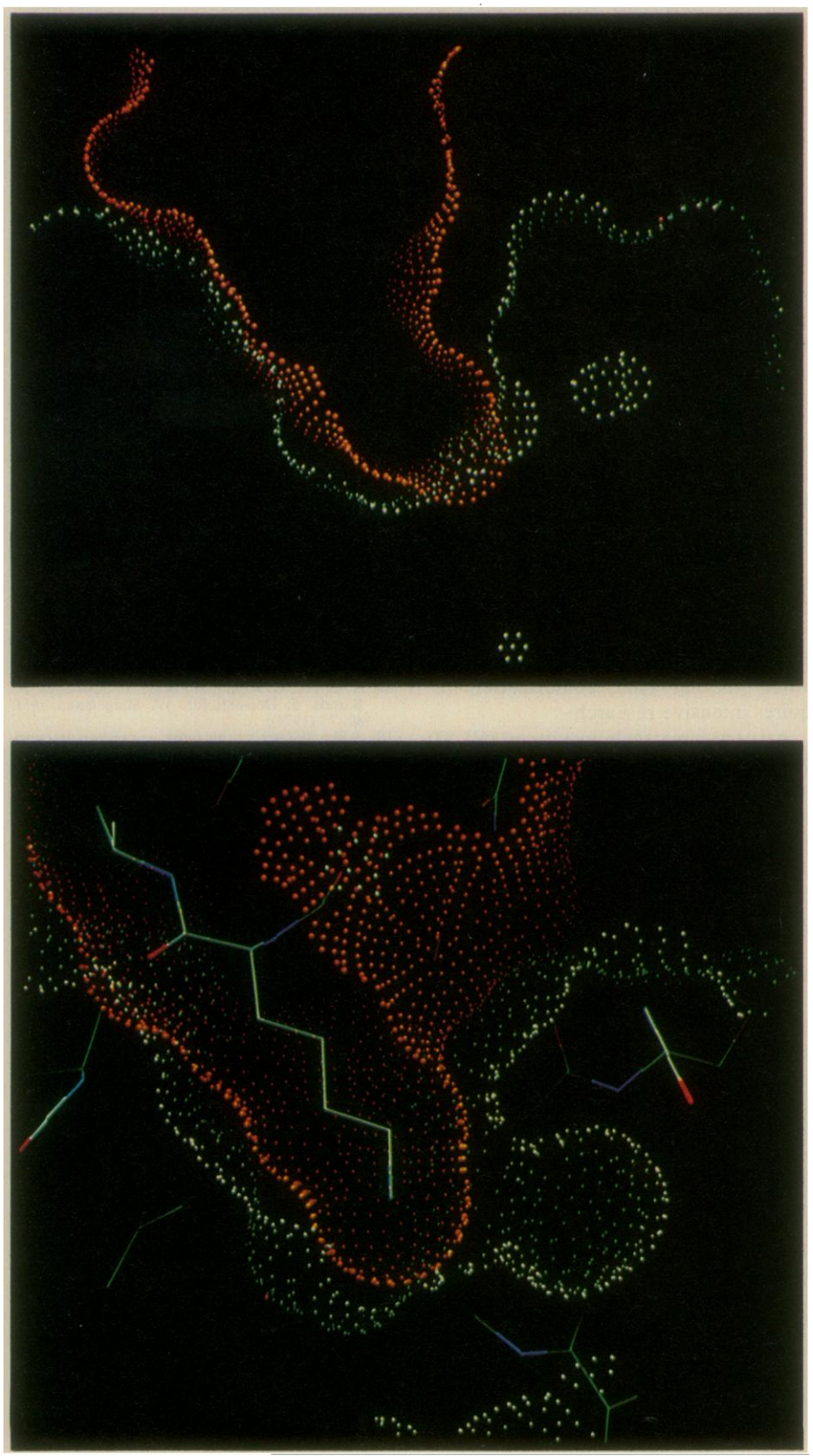


Fig. 5. (Left) Bovine trypsin inhibitor-trypsin complex. Surfaces only are shown. The inhibitor surface is in red, the trypsin surface in green. (Top) As in the picture on the left, but with front and rear surface clipping. (Bottom) As in the top picture, but with color-coded skeleton.

Color greatly enhances the display of the surfaces of two interacting molecules. With a black-and-white display it is difficult to distinguish one surface from the other in any kind of close binding. With color, one can, for example, make one molecule red and the other green. If two surfaces are close in regions of high complementarity, salt bridges and hydrogen bonds appear as small overlaps of the molecular surfaces and are immediately visible; this precludes the need for interatomic distance calculation to identify them, though quantitative calculations may be made at any time.

To illustrate the value of color in the study of molecular interactions, we use the bovine trypsin inhibitor-trypsin system (*17*). The portions of each protein in the region of interface are shown in Fig. 5. It is said that one picture is worth 10,000 words, but by altering the image interactively one can produce thousands of different pictures in a single session with the graphics system. Three such pictures are shown in Fig. 5, each highlighting a different aspect of the interaction, as described below.

Figure 5 (left) shows two empty, bulbous pockets belonging to trypsin lying on either side of the tip of the inhibitor. When trypsin binds an arginine side chain, the two terminal nitrogens of the guanidino group fill these side pockets. As we follow the surface out of the bottom of the substrate pocket, the fit improves on one side but becomes worse on the other side. The visual display thus immediately localizes the regions of strong interaction. Alternatively, a calculation listing the pairs of atoms of the two molecules that interact strongly with each other still leaves the task of assembling this list into some kind of understandable presentation. For complex interactions it is often advantageous to move the hither and yon clipping planes together so that the image is simplified to a thin section [Fig. 5 (top)].

While pure surface displays have a substantial value, to study real chemistry one needs to include the skeleton [Fig. 5 (bottom)]. Without the "wire" model one cannot estimate the absolute magnitude of overlaps and gaps, since surfaces alone do not provide a sense of stereochemical scale.

Advantages of Real-Time Display

There is no precise definition of the terms real-time and interactive. The difference between interactive and noninteractive uses of computer graphics depends on how long you are willing to wait to see a result. The interval between the user's input and the computer graphic reaction should be immediate for simple tasks such as global rotation or rotation about a bond. A much longer wait can be tolerated for a conformational calculation such as an energy minimization. But even this has an upper limit, and the transition between satisfactory and unacceptable interactive use is a function not only of the system but of the user.

Flicker-free interactions require that the picture on the screen either remain unchanged (as in a storage tube) or be updated more rapidly than the flicker fusion frequency of the human observer (about 20 to 30 cycles per second). The ability to present a changed picture at each frame update obviously depends on the complexity of the picture, the type of display, and the hardware available. The classic problem is global rotation of a figure. Software to do the necessary matrix multiplication is simple, and for a simple figure need not be too time-consuming, but as figures become more complex, special-purpose hardware is necessary to keep up with the frame update rate and to present a smoothly rotating three-dimensional object.

Ideally, a hierarchical system answers these requirements in the most cost-effective manner. The most important calculations are done locally, possibly in hardware, with successively less immediate computations transferred to other machines (*4*).

Future

We now have hardware and software for selective display of complex objects in color and stereo with real-time rotation, depth cueing, and interaction. Both skeletal and surface representations can be displayed simultaneously, molecules can be moved relative to one another, and stereochemical parameters can be computed. These capabilities provide most of the features needed for molecular displays (*18*).

There are many other uses to which real-time color displays can be put. Crystallographic temperature factors have been interpreted in terms of internal protein motions and flexibility (*19*). This problem has also been approached through molecular dynamics calculations (*20*). Not only may these motions be modeled, but differential coloring for the more flexible atoms and the more rigid atoms provides a convenient way to examine the results. Color is also useful in traditional crystallographic work such as electron density fitting ("Richards boxing") (*21*). When building a model into electron density, the model is given one color and the electron density contours a different color, or several different contour levels are displayed simultaneously, each in a different color.

In our own molecular surface work and in the surface pictures of Feldmann (*8*), Max (*9*), and co-workers, the color scheme is based on atom type or a related variable. Since the actual location in space of the molecular surface points is known, the color of each point may represent the electric potential at that point. Charge and dipole interactions between molecules may then be explicitly examined. The difficult problem of manipulating these molecular surfaces while varying their conformation requires intensive research.

As very large scale integrated (VLSI) circuit technology is developed and hardware costs are reduced, interactive raster graphics systems will become more generally available, although whether they will soon compete with the color calligraphic display in speed and resolution is not clear.

Calculation of derived functions such as interaction energy will proceed more rapidly and be extended to higher degrees of structural complexity as computers increase in power, and although computations of greater and greater complexity will become possible, blinding speed is of little use without illuminating interaction. Real-time three-dimensional color graphics provides that interaction.

References and Notes

1. R. Langridge and A. W. MacEwan, in *Proceedings, IBM Scientific Computing Symposium on Computer Aided Experimentation* (IBM, Yorktown Heights, N.Y., 1965), p. 305; C. Levinthal, in *ibid.*, p. 315; *Sci. Am.* **214**, 42 (June 1966).
2. R. J. Feldmann, *Annu. Rev. Biophys. Bioeng.* **5**, 477 (1976).
3. W. M. Newman and R. J. Sproull, *Principles of Interactive Computer Graphics* (McGraw-Hill, New York, ed. 2, 1979).
4. The hardware is an Evans and Sutherland Picture System 2 calligraphic color display interfaced to a Digital Equipment Corporation PDP-11/70 computer, with a remote link to a batch-mode CDC 7600 computer at Lawrence Berkeley Laboratory.
5. T. E. Ferrin and R. Langridge, *Comput. Graphics* **13**, 320 (1980).
6. D. M. Ritchie and K. Thompson, *Commun. ACM* **17**, 7 (1974). UNIX is a registered trademark of Bell Laboratories.
7. Bausch & Lomb, Rochester, N.Y.
8. R. J. Feldmann, D. H. Bing, B. C. Furie, D. Furie, *Proc. Natl. Acad. Sci. U.S.A.* **75**, 5409 (1978). The algorithm was described by T. Porter [*Comput. Graphics* **12**, 282 (1978)].
9. N. L. Max, *Comput. Graphics* **13**, 165 (1979); ———, D. Malhotra, A. Hopfinger, *Comput. Chem.*, in press.
10. R. Langridge, *Fed. Proc. Fed. Am. Soc. Exp. Biol.* **33**, 2332 (1974); in *Computers in Life Sciences Research*, W. Siler and D. A. B. Lindberg, Eds. (Plenum, New York, 1975), p. 53. The display was an Evans and Sutherland LDS-1 interfaced to a Digital Equipment Corporation PDP-10 (KA-10), and the film, "Life and the Structure of Hemoglobin," directed by B. Shapiro, was made for the American Institute of Physics with support from the National Science Foundation. It was shown on National Educational Television in October 1974. (The film is available from the American Institute of Physics, 335 East 45 Street, New York 10017.)
11. W. N. Lipscomb, G. N. Reeke, J. A. Hartsuck, F. A. Quiocho, P. H. Bethge, *Philos. Trans. R. Soc. London Ser. B* **257**, 177 (1970).
12. B. Lee and F. M. Richards, *J. Mol. Biol.* **55**, 379 (1971).
13. F. M. Richards, *Annu. Rev. Biophys. Bioeng.* **6**, 151 (1977).
14. M. L. Connolly, I. D. Kuntz, T. E. Ferrin, R. Langridge, in preparation.
15. J. Greer and B. L. Bush, *Proc. Natl. Acad. Sci. U.S.A.* **75**, 303 (1978).
16. R. Langridge, D. A. Marvin, W. E. Seeds, H. R. Wilson, C. W. Hooper, M. H. F. Wilkins, L. D. Hamilton, *J. Mol. Biol.* **2**, 38 (1960).
17. R. Huber, D. Kukla, W. Bode, P. Schwager, K. Bartels, J. Deisenhofer, W. Steigemann, *ibid.* **89**, 73 (1974).
18. A film illustrating this work was shown at a variety of meetings and symposia, including the American Chemical Society Burger Award Symposium, Houston, Texas, 26 March 1980. Persons interested in a film should write to R.L.
19. H. Frauenfelder, G. A. Petsko, D. Tsernoglou, *Nature (London)* **280**, 558 (1979); P. J. Artymiuk, C. C. F. Blake, D. E. P. Grace, S. J. Oatley, D. C. Phillips, M. J. E. Sternberg, *ibid.*, p. 563.
20. M. Karplus and J. A. McCammon, *ibid.* **277**, 578 (1979).
21. F. M. Richards, *J. Mol. Biol.* **37**, 225 (1968); R. Diamond, *BILDER: An Interactive Graphics Program for Biopolymers, User's Guide* (MRC Laboratory of Molecular Biology, Cambridge, England, 1979); A. Jones, *FRODO: An Electron Density Fitting Program, User's Guide* (University of Uppsala, Uppsala, Sweden, 1980).
22. F. C. Bernstein, T. F. Koetzle, G. J. B. Williams, E. F. Meyer, Jr., M. D. Brice, J. R. Rodgers, O. Kennard, T. Shimanouchi, M. Tasumi, *J. Mol. Biol.* **112**, 535 (1977).

23. Coordinates were obtained from the investigators cited; from R. J. Feldmann, Division of Computer Research and Technology, National Institutes of Health; and from the Protein Data Bank (*22*). We thank M. J. Pensak, O. Jones, C. C. Huang, and R. M. Abarbanel for programming; P. Weiner for photography; J. E. Apodaca for assistance in preparing this article; P. A. Kollmann for discussion; and J. E. Goyan, E. C. Jorgensen, and M. E. Wolff for support and encouragement. This research was supported by NIH biotechnology resource grant RR-1081-03 (R.L.), NIH grant GM 19267 and an NIH research career development award (I.D.K.), an NSF graduate fellowship (M.L.C.), and the University of California.