12 December 1975, Volume 190, Number 4219 SCIENCE

Protein Crystal Structures: Quicker, Cheaper Approaches

An economical approach to phase refinement is coupled with a new computer graphics system.

D. M. Collins, F. A. Cotton, E. E. Hazen, Jr.,
E. F. Meyer, Jr., C. N. Morimoto

The use of x-ray diffraction methods to study crystals of biological macromolecules dates from 1934, when Bernal and Crowfoot (*1*) found that the diffraction patterns from crystals of the proteolytic enzyme pepsin were, in Hodgkin's words (*2*), "rich and full of detail." The potential for revealing the precise molecular structure of protein molecules inherent in such diffraction patterns remained unrealized until 1954, when M. F. Perutz, who had devoted his early career to the effort, showed that the "isomorphous" insertion of heavy atoms into the crystal lattice could provide phase information, after which J. C. Kendrew solved the structure of myoglobin and Perutz solved that of hemoglobin. In 1965 Phillips reported the first detailed structure of an enzyme, lysozyme, but perhaps more important was his demonstration that it is possible to study directly the binding of a substrate-like molecule at the active site. For details of the methods and references to the many contributions of others to the field, the authoritative reviews by Phillips (*3*) and North and Phillips (*4*) should be consulted.

Today the crystal structures of more than 50 proteins and transfer ribonucleic acids are under investigation in laboratories throughout the world. However, despite the wide applicability and success of the method, crystal structure determination for biological macromolecules remains time-consuming and expensive. It is particularly arduous to achieve a level of resolution which reveals structural details of direct chemical significance. Therefore, research aimed at improving the state of the art is at least as important as the extension of existing methodology to new molecules. We shall discuss here two steps where improvements would be of particular value and describe some recent advances which have been made, in part by work in our own laboratories.

Refinement

Normal crystal structure analysis leads to an electron density map, which is the basis for an approximate atomic model. In the routine case of a structure at high resolution, a preliminary electron density map reveals the atoms as individual, partially isolated peaks and interpretation of the map to yield semiquantitative atomic positions is straightforward. Since the advent of the digital computer it has become standard practice to refine such semiquantitative models by least squares. This classical nonlinear least-squares refinement is complex and expensive, as it is based on minimizing the difference between the observed and calculated diffraction patterns by variation of up to nine position and thermal motion coordinates per atom (*5*). Whereas least-squares coordinate refinement is remarkably cost-effective for small molecules, similar calculations would be prohibitively expensive for most macromolecules. In addition, the diffraction data generally are of insufficient resolution to define a semiquantitative atomic structure model as a starting point.

It is, nevertheless, possible to refine structural information for biological macromolecules by using approaches which can be classified in three broad categories: (i) coordinate refinement, including classical least squares, difference Fourier methods, and the approach devised by Diamond (*6, 7*); (ii) minimization of potential energy, using approximate interatomic potential functions (*8*); and (iii) phase refinement, including applications of the tangent formula (*9*) and the methods discussed in more detail below (*10*). Although the refinement of more and larger structures is under way, it is indicative of the time and expense involved in such work that refinements have been published for only six proteins. With the exception of chymotrypsin (*6, 11, 12*), where the calculations were deliberately limited to reduce expense, the largest molecule yet refined, lysozyme, has a molecular weight of some 14,600.

While the method of isomorphous replacement has opened up the field of protein crystallography, it is still only a first step in deriving good phases, which can lead to accurate stereochemical descriptions of macromolecules. This has long been apparent to the experimentalist, whose best efforts result in obviously flawed electron density maps. In fact, it is not unusual for electron density maps to have regions of negative density nearly as large in magnitude as the largest positive regions. It is neither practical nor possible to obtain a data set so large or of such high resolution that the corresponding electron density function would be everywhere nonnegative, but it is clear that negative density regions ought to be few and relatively small if the data are carefully measured,

D. M. Collins is Senior Research Associate; F. A. Cotton, Robert A. Welch Distinguished Professor; E. E. Hazen, Jr., Associate Professor; E. F. Meyer, Associate Professor; and C. N. Morimoto, Postdoctoral Fellow at Texas A&M University, College Station 77843. The first three authors are associated with the Department of Chemistry, the latter two with the Department of Biochemistry and Biophysics.

Table 1. Some parameters of phase refinement for rubredoxin. There are 2812 unique structure amplitudes in the 2.0 Å data set; in the 1.5 Å data set there are 5033, and each was assigned a phase when the arbitrary criterion $|F_{calc}| > 0.1|F_{exp}|$ was satisfied (*24*). The average error in phase angle is based on comparison of our refined phases with those of Watenpaugh *et al.* (*7*). The angle parameters are given as modulus-weighted averages. For these calculations $\chi = 0.33$. Abbreviations: N, number of phases; R, disagreement factor.

Iteration	N	R (*25*)	Phase angle (degrees)	
			Average change (*35*)	Average error
0	2812			36
1	4899	0.35	35	35
2	5021	0.21	13	32
3	5030	0.17	8	30
4	5032	0.17	6	30*

*For the starting subset of 2812 data, this value is 27°.

accurately phased, and constitute a reasonably complete set.

There are many powerful and useful algebraic relationships among structure factors based on the assumption of positive electron density (*13*). Sayre (*14*) provided an especially illuminating statement of this idea when he proposed a proportionality between the electron density and its square. Straightforward application of the convolution theorem then gives a complex system of structure factor equations, variations of which are being used routinely in crystal structure analysis of small molecules (*15*). Sayre (*16*) has shown that these same equations can be used for the refinement and extension of phases in protein crystallography. Unfortunately, the calculations are formidable and, at present, very expensive.

Barrett and Zwick (*17*) demonstrated that phase extension for protein crystal structures was possible at reasonable cost. They started with an electron density map of myoglobin at a resolution of 3.0 Å, set all negative density values equal to zero, squared the positive density values, and performed a Fourier inversion of the modified map using the Cooley-Tukey fast Fourier transform algorithm (*18*). From the resulting set of calculated structure factor moduli and calculated (extended) phases, they took the phases they lacked for the 2.0 Å data and calculated the inverse transform to obtain an improved map based on the extended phase set.

In any phase refinement calculation based on map modification [see also (*19*)], the first step is to find f, a modification function such that

$$\rho_{calc}(\mathbf{r}) = f[\rho_{exp}(\mathbf{r})] \quad (1)$$

where ρ_{calc} is a better representation of the true structure than is ρ_{exp}, an electron density function derived directly from the experimental data, and $\mathbf{r}$ is a position vector. Even though any suitable design for f is largely a matter of art, there are four practical criteria (*20*) of particular importance for working with biological macromolecules. First, to minimize computation it must be possible to obtain overlap-free Fourier inversions of ρ_{calc} using coarse-grid maps (*21*), in such a way that correct regions of ρ_{exp} are modified as little as possible. The curve given by Eq. 1 should also be quite smooth. Second, resolution of completely overlapped atom profiles must be encouraged in the course of phase refinement. Third, it must be possible to adjust the modification function to allow for special treatment of heavy atoms or groups (*22*). Fourth, there must be no requirement of information beyond an initial electron density map so that the refinement can be a fully automatic, iterative process.

For a hypothetical structure of equal atoms (*21*) it has been shown that after ρ_{exp} has been scaled to have a maximum value of 1.0, a serviceable modification function (also shown in Fig. 1) is

$$\rho_{calc}(\mathbf{r}) = \begin{cases} 3\rho_{exp}^2(\mathbf{r}) - 2\rho_{exp}^3(\mathbf{r}); & \rho_{exp}(\mathbf{r}) \geq 0 \\ 0; & \rho_{exp}(\mathbf{r}) < 0 \end{cases} \quad (2)$$

Of course, no real biological structure will consist of equal atoms. Moreover, the usual experimental resolution and consequent atom overlapping nearly guarantee that even atoms of the same type will be represented at varying levels of electron density. To deal with these conditions we scale the curve of Fig. 1 to the same shape in the range $\rho_{exp} \leq \chi$ and smoothly piece it with a tangent through the point $(\rho_{exp},\rho_{calc}) = (1.0,1.0)$, as shown in Fig. 2 (*22*). In any particular application (assuming ρ_{exp} is scaled to have a maximum value of 1.0) we take as a first estimate of χ the magnitude of the maximum density in regions consisting solely of oxygen and lighter atoms. During the course of phase refinement the value of χ may be revised downward so that it is not appreciably larger than twice the magnitude of the most negative value of ρ_{exp}. If χ approaches zero, the map modification becomes nothing more than replacement of negative density values by zero.

Unfortunately, the outlined modification procedure is blind and not only suppresses incorrect map features, but alters correct ones as well. In attempting to regain the correct features we take the best electron density map based on refined phases to be the Fourier transform of $(2|F_{exp}| - |F_{calc}|)\exp i\alpha_{calc}$, where F denotes structure factor and α phase angle (*22–24*).

Table 2. Some parameters of phase refinement for staphylococcal nuclease. There are 18,457 unique structure amplitudes in the 1.5 Å data set. The tabulated number of phases is an estimation based on the number of phase assignments in a symmetry-expanded data set; phases were assigned when $|F_{calc}| > 0.1|F_{exp}|$ (*24*). The angle parameters are given as modulus-weighted averages. For these calculations $\chi = 0.60$. Abbreviations: N, number of phases; R, disagreement factor.

Iteration	N	R (*25*)	Average change in phase angle (degrees) (*35*)
0	9,637		
1	17,963	0.29	31
2	18,359	0.22	10
3	18,389	0.20	8

We can summarize our procedure for refinement and extension of phases by the following sequence (*24*).

1) Compute an experimental electron density map, ρ_{exp}, using the best available phases.

2) Calculate ρ_{calc} using a curve of the type shown in Fig. 2.

3) Obtain $|F_{calc}|$ and α_{calc} by Fourier inversion of ρ_{calc}.

4) From $(2|F_{exp}| - |F_{calc}|)\exp i\alpha_{calc}$ (*24*) compute a best experimental electron density map and return to step 2 unless convergence has been achieved.

We follow a refinement by means of the disagreement factor, $R = \Sigma||F_{exp}| - |F_{calc}|| / \Sigma|F_{exp}|$ (*24, 25*), and judge that convergence has been achieved when R has fallen to a stable value; this usually requires four iterations.

We have phase-refined the structures of the rubredoxin of *Clostridium pasteurianum* (*7, 22*) and the extracellular nuclease of *Staphylococcus aureus* (*26*). In both cases we started with phase sets with a resolution of 2.0 Å derived by isomorphous replacement methods and refined and extended the phases to a nominal resolution of 1.5 Å. In Table 1 we assume that the phases from the classical least-squares model refinement carried out by Jensen and co-workers (*7*) are free of error and give several parameters to describe the course of phase refinement for rubredoxin, the objective control for our procedure. Table 2 summarizes the progress of phase refinement for the nuclease. For rubredoxin, each iteration took 10 minutes on an IBM 360/65 computer at a cost of $25; the corresponding figures for nuclease are 20 minutes and $50.

Map to Model: The Richards Box

In the crystallography of macromolecules one cannot go directly from experimental data (structure amplitudes and derived phases) to atomic positions, as is usually the case with small molecules. The immediate, tangible product obtained from the experimental data is a set of numbers representing the relative magnitude of the electron density at points on a three-dimensional grid in the crystallographic unit cell. The meaningful resolution is usually somewhat poorer (approximately 2.0 to 2.5 Å) than the distances between atoms (1 to 2 Å) so that the data are intrinsically incapable of resolving the individual atoms.

In order to convert this three-dimensional grid of numbers into a more recognizable image of the molecule it is helpful to replace it by contour surfaces, thus calling attention to the regions of greatest electron density and to the shapes of these regions. Then by comparing these contours with the known shapes of the amino acid residues, especially their side chains, one can attempt to fit a model to the electron density. The use of three-dimensional contour surfaces as such is a practical impossibility, and until 1968 the procedures employed included simple inspection of stacks of contour sections or placement of color-coded beads on rods to try to convey a visual image of the peaks and valleys of electron density. This was almost unbelievably cumbersome.

In 1968 Frederic M. Richards of Yale University invented the Richards optical comparator, less formally known as the Richards box (or even, affectionately, as Fred's folly), which, in its awesome simplicity, revolutionized the interpretation of protein electron density maps (*27*). This device uses a half-silvered mirror to superimpose optically the image of a stick model on the contoured electron density map, made effectively three-dimensional by stacking a series of two-dimensional sections. When the observer looks into the half-silvered mirror the model is perceived as being embedded in the electron density map. The elements of the model can then be adjusted by hand to match the map.

Although the invention of the Richards box has greatly facilitated the interpretation of protein crystallographic data, it still leaves much arduous work. The "fine tuning" of an atomic model in a Richards box is a long, tedious project; but worse is yet to come. The coordinates of all the atoms must then be measured from the model. This is both highly tedious and inherently inaccurate.

An obvious desideratum is an electronic

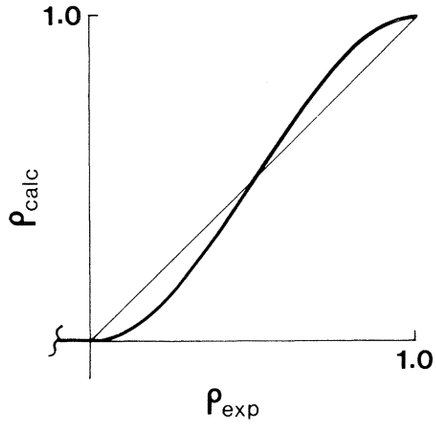


Fig. 1. The heavy line is the curve given by Eq. 2 for map modification in the range $0 < \rho_{exp} \leq 1$; the straight line of slope 1 corresponds to no modification.

Richards box, that is, a device which allows the crystallographer to display a sufficient volume of the electron density map, in stereo, on a television screen and simultaneously superimpose atomic models, also in stereo, in such a way that the latter can be translated and rotated until an optimum fit of the model to the map is achieved. Fitting model to map in this way can be far more convenient and faster than the mechanical operations in the Richards box, but even more important, the reading of coordinates can then be completely automated. Once the model is fine-tuned the push of a button will record the coordinates.

Electronic Richards Box

What we can call electronic Richards boxes have been developed at Washington University (*28*) and at Oxford (*29*). The former, which is based on a set of specially designed and constructed digital modules, has been used to superimpose subunits of

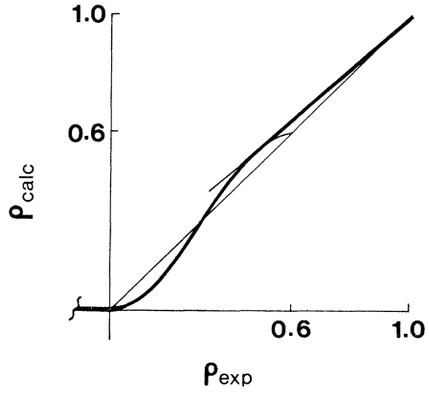


Fig. 2. The heavy line gives the relationship between positive ρ_{exp} and ρ_{calc} when $\chi = 0.60$.

malate dehydrogenase and lactate dehydrogenase (*30*); the latter, based on commercially available hardware, has been used to study the binding of inhibitors to lysozyme (*29*).

The system used in the Department of Biochemistry and Biophysics at Texas A&M is part of CRYSNET (*31*), a prototype network designed to make a range of computer graphics generally available to chemists and biologists. It is based on a Digital Equipment Corp. PDP 11/40 computer with 28 kilowords of core memory and Vector General graphics hardware. Peripherals include a card reader, 1.2 megaword disk, tape drive, and printer. Interactive control of the display may be effected by means of keyboard commands and console switches. However, the most useful input device is a box of ten control dials. Because the number of variables involved in fitting a tripeptide or tetrapeptide greatly exceeds ten, the function of each dial may be reassigned at the convenience of the operator. A major constraint to any graphics system is its flicker rate or the number of vectors it can draw in about 1/30 second. This system can draw models of molecules with the size and complexity of, for example, the alpha and beta subunits of horse methemoglobin (2229 atoms) without flicker. Programs for contouring selected portions of the electron density map, building the model, and fitting the model to the map have been written.

Graphical models can be viewed three-dimensionally by several techniques. The kinetic depth technique permits three-dimensional viewing of a continuously rotating image under operator control via the control dials. This method is effective for surveying vantage points and for motion picture photography, but it is troublesome for model fitting, which requires a more static representation of electron density contours. Side-by-side viewing of stereo images is available as an option, as is the viewing through a half-silvered mirror. Both techniques permit several individuals to view the display simultaneously. Another option allows the use of split screen side-by-side orthogonal views. Two contouring approaches are available and often used in tandem. First, an approximate fit of model to map is obtained by using a three-dimensional grid of single level contours. Then, stacks of multilevel contours (as in the Richards box) are used for fine tuning of the fit.

A variety of three-dimensional interactive manipulations are available for the fitting process: (i) rotation of the entire picture about a selectable vector, whether in the static or continuous rotation mode;

for moving the model or parts of the model with respect to the contours four rotational and four translational modes are available; (ii) complete model translation; (iii) translation of any subunit of the model (for instance, the inhibitor of an enzyme-inhibitor complex) with respect to the rest of the model and the contours; (iv) translation of the model for any selected portion of any subunit; (v) individual atom translation; (vi) complete rotation of the model; (vii) rotation of any subunit of the model; (viii)

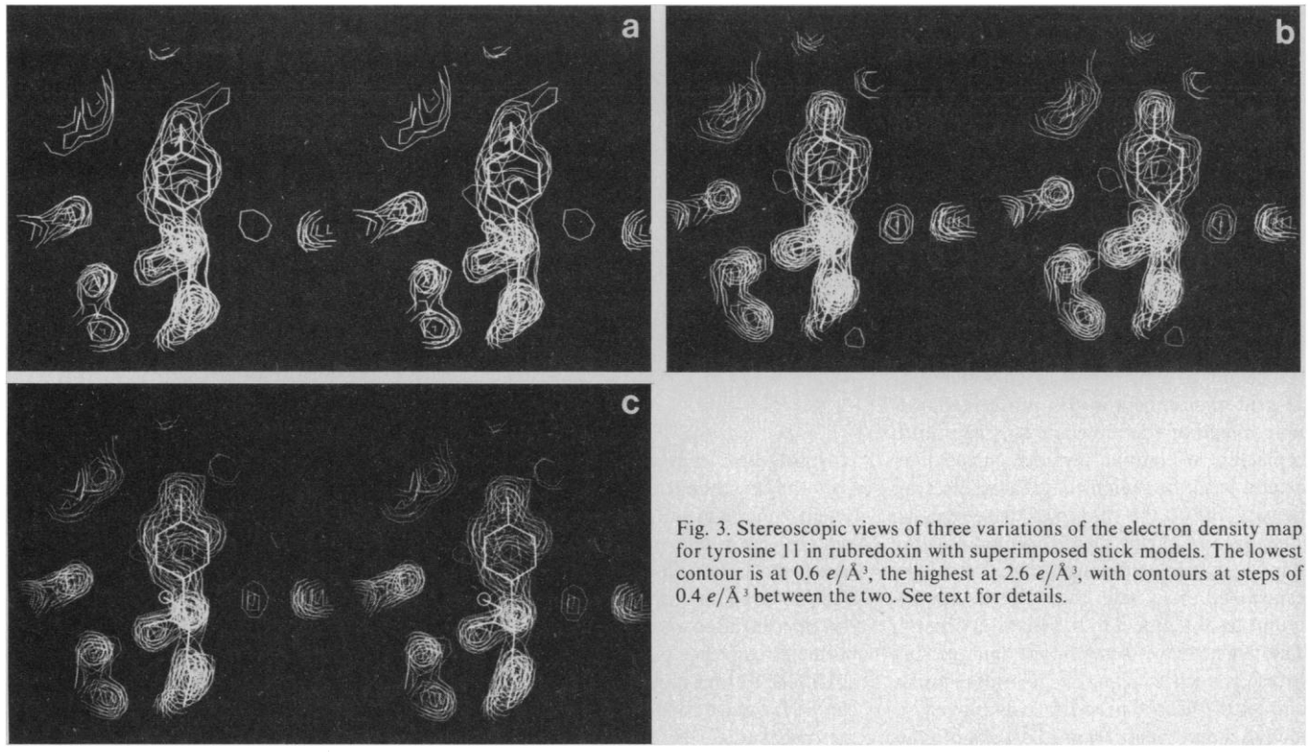


Fig. 3. Stereoscopic views of three variations of the electron density map for tyrosine 11 in rubredoxin with superimposed stick models. The lowest contour is at 0.6 $e/Å^3$, the highest at 2.6 $e/Å^3$, with contours at steps of 0.4 $e/Å^3$ between the two. See text for details.

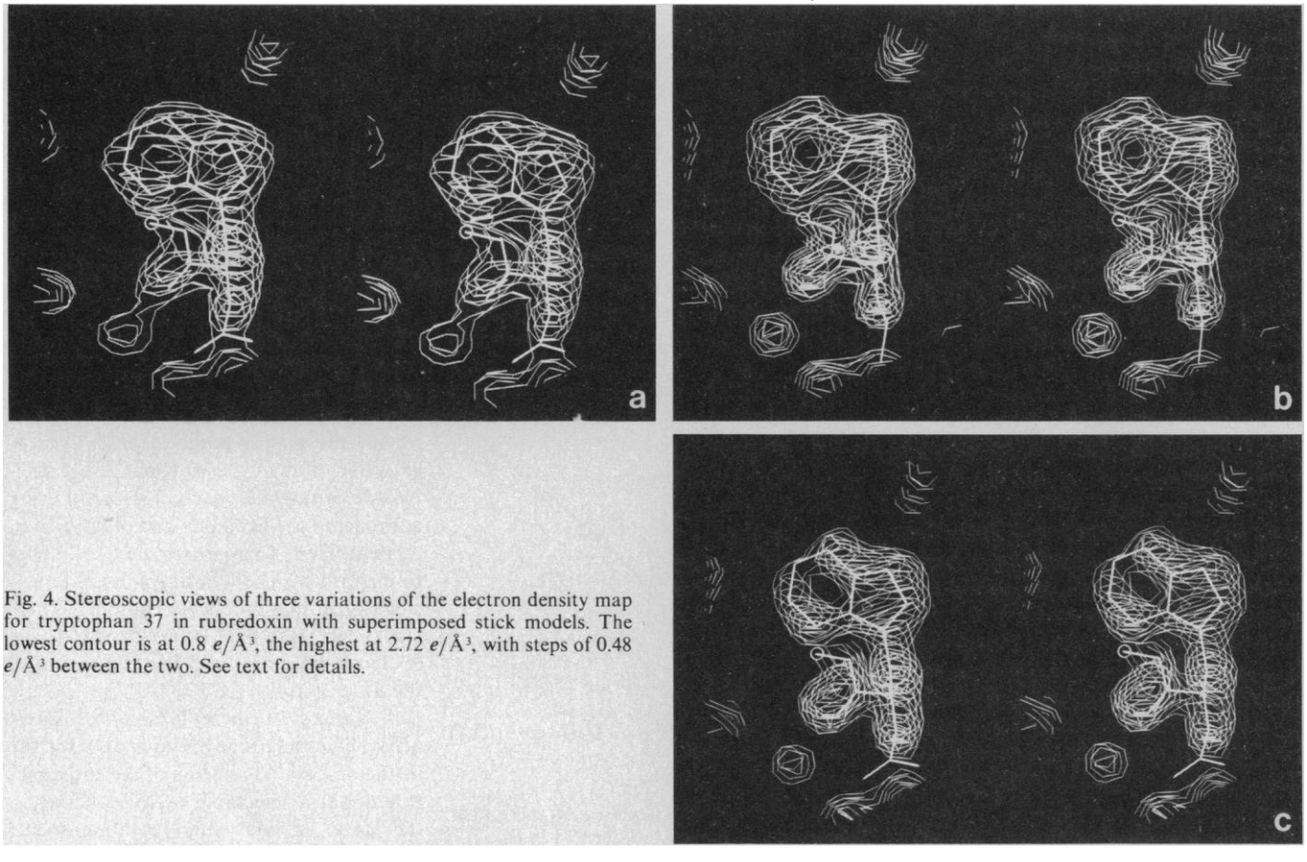


Fig. 4. Stereoscopic views of three variations of the electron density map for tryptophan 37 in rubredoxin with superimposed stick models. The lowest contour is at 0.8 $e/Å^3$, the highest at 2.72 $e/Å^3$, with steps of 0.48 $e/Å^3$ between the two. See text for details.

subunit rotation about a particular bond; and (ix) dihedral rotation of any part of a subunit at either the head or the tail of the flashing bond pointer (*32*). The coordinates of the atom at the pointer's head, which also flashes, can be printed on command. The pointer can be moved forward or backward through 15 atoms by a dial on the control box. When the dial is turned fully, the computer moves the starting point, allowing the entire model to be traced.

As indicated above, the system is entirely based on commercially available hardware. The programming has been done in FORTRAN and has proved to be usable on other minicomputer-based systems. The display and fitting programs are completely general; that is, they are not limited to proteins but can be used for any molecule, large or small. Views of a particular fitting can be stored for later retrieval. A list of atomic coordinates can be produced simply by a single command calling for a record to be created.

Although the CRYSNET system has a number of other features (*31*), the practical work described below is the first extensive trial of this interactive computer graphics system as an electronic Richards box. In these particular fittings of the electron density maps of the phase-refined rubredoxin and staphylococcal nuclease data, only idealized parameters for the peptide bond and the amino acid residues have been used, but the system is not restricted to this. As the use of the system in this work has progressed, the interaction among the individuals involved in the programming and model-fitting has led to modifications and improvements in the system. As this type of interaction continues and expands, the system should evolve into a practical and efficient tool for those studying the structure and function of biological macromolecules.

Illustrative Practical Results

Armed with both the phase refinement and extension method and the electronic Richards box we have been able to achieve some remarkable improvements in the clarity of the electron density maps and accuracy of the atomic models for the two structures, rubredoxin and staphylococcal nuclease, with expenditures of time, effort, and money which are comparatively very modest. The efficiencies achieved represent, we believe, an order of magnitude improvement in the state of the art.

In the case of rubredoxin, accuracy can be verified by a change from the approximate starting phases to phases that compare more favorably with the authoritative results of the classical least-squares model refinement by Jensen and co-workers (*7*). It is important to recognize that the validation of our work depends largely on their refinement. Note should also be taken of

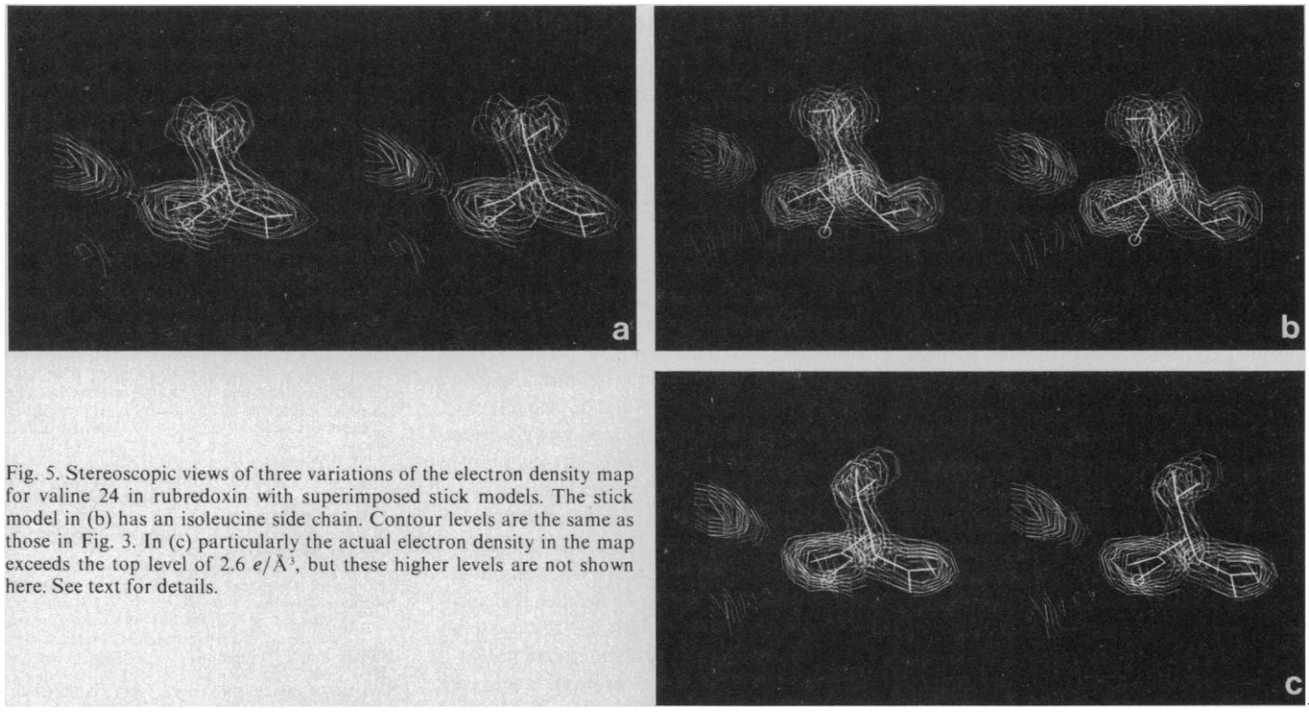


Fig. 5. Stereoscopic views of three variations of the electron density map for valine 24 in rubredoxin with superimposed stick models. The stick model in (b) has an isoleucine side chain. Contour levels are the same as those in Fig. 3. In (c) particularly the actual electron density in the map exceeds the top level of 2.6 $e/Å^3$, but these higher levels are not shown here. See text for details.

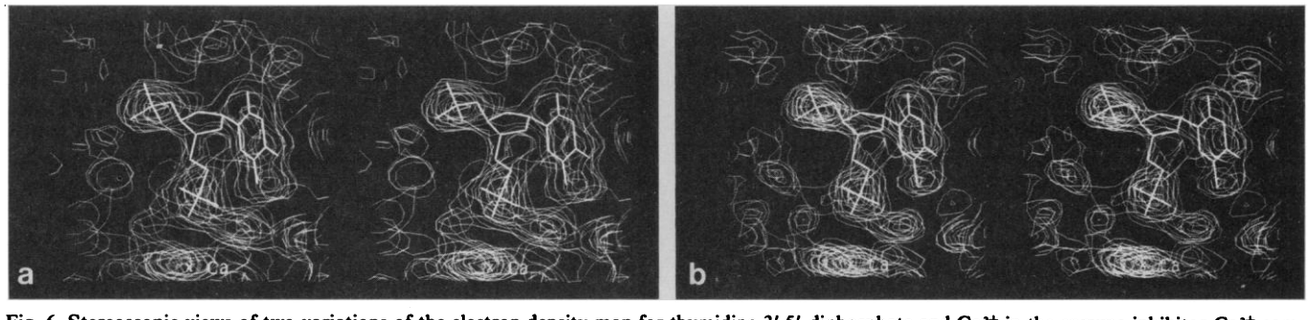
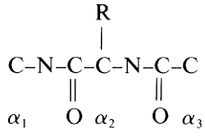


Fig. 6. Stereoscopic views of two variations of the electron density map for thymidine-3′,5′-diphosphate and Ca^{2+} in the enzyme-inhibitor-Ca^{2+} complex of staphylococcal nuclease with superimposed stick models. The lowest contour is at 0.4 $e/Å^3$, the highest at 1.90 $e/Å^3$, with steps of 0.5 $e/Å^3$ between the two. See text for details.

their caveat that their current results for rubredoxin should not be regarded as final but only as a test of the applicability of the least-squares method to a large molecule.

Figures 3 to 5 compare the electron density maps from the 2 Å multiple isomorphous replacement (MIR) data (Figs. 3a, 4a, and 5a), the 1.5 Å least-squares (LS) refined data of Jensen's group (*7*) (Figs. 3b, 4b, and 5b), and the 1.5 Å data subjected to the direct space refinement (DSR) procedure described here (Figs. 3c, 4c, and 5c) for three of the amino acid residues of rubredoxin, namely tyrosine 11, tryptophan 37, and valine 24, respectively. The coordinates for the superimposed stick models for the LS maps are those of Jensen and co-workers (*7*), while in the MIR and DSR maps rigid models comprising standard bond lengths and angles have been fitted to the electron density by using the interactive computer graphics system. In the latter two cases the peptide bonds have not been allowed to deviate from planarity, although in Fig. 3b particularly it appears a much better fit would be obtained by allowing such a deviation. In each figure the amino acid side chain (R) is toward the top and a section of the main peptide chain is shown as diagrammed below.



Our general impression from viewing these sections and others in the refined rubredoxin maps is that the LS map is of somewhat better quality than the DSR map, but it also seems quite clear that the DSR map represents a very significant improvement over the MIR map. Note that in the LS maps for tyrosine 11 (Fig. 3b) and tryptophan 37 (Fig. 4b) there appear to be structurally unreasonable distortions of the aromatic rings.

Figure 6a shows a stick model—again with standard bond lengths and angles—of thymidine-3′,5′-diphosphate (pdTp) interactively fitted to its electron density as it appears in the 2 Å MIR map of the staphylococcal nuclease-pdTp-Ca^{2+} complex (*26*). The position of the calcium ion is indicated with an X. Figure 6b is the 1.5 Å DSR map for the same area of the enzyme-inhibitor complex. Here again our refinement procedure appears to have significantly improved the quality of the map, this time in an area of the map that was surprisingly clear in a 4 Å map based on a single heavy-atom derivative (*33*). Note that the calcium position in the MIR map (Fig. 6a) is poorly resolved and unreasonably low in electron density. However, in the DSR map (Fig. 6b) the calcium position is sharply defined and has an electron density in reasonable accord with its atomic number. Thus, our refinement procedure seems able to selectively intensify real structural features. Conversely, the refinement also appears able to remove incorrect features from the MIR map. For example, study of Fig. 5 reveals that for residue 24, which was originally identified as an isoleucine (*7*) and later correctly identified as valine (*34*), the DSR refinement has removed the electron density that led to the confusion.

Conclusion

Although a number of protein structures and at least one transfer RNA structure have been determined by x-ray crystallography, and several of the protein structures have been extensively refined, the methods used have been lengthy, costly, and inefficient. We have described here major improvements at two key stages in the process of macromolecular structure determination. One has to do with refining and extending phase information by a procedure which is both effective and affordable. The extrapolation aspect of this procedure is of major importance, since it allows resolution to be improved merely by collecting more data on the native structure (say from 2.0 to 1.5 Å) without the necessity of obtaining MIR phase information in the same range as well.

The other has to do with building and fine-tuning explicit structure models by means of an interactive computer graphics system based on commercially available components. The computer graphics system has been developed and used to fit molecular models efficiently to experimentally determined electron density sections at a resolution of 1.5 to 3.0 Å. Besides the convenience and speed of fitting, this system and others like it offer the enormous advantages of providing prompt, accurate readout of atomic coordinates and publishable stereoscopic diagrams.

References and Notes

1. J. D. Bernal and D. Crowfoot, *Nature (Lond.)* **133**, 794 (1934).
2. D. M. C. Hodgkin, *Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci.* **257**, 65 (1970).
3. D. C. Phillips, *Adv. Struct. Res. Diffr. Meth.* **2**, 75 (1966).
4. A. C. T. North and D. C. Phillips, *Prog. Biophys. Mol. Biol.* **19**, 1 (1969).
5. G. H. Stout and L. H. Jensen, *X-ray Structure Determination* (Macmillan, New York, 1968), p. 385.
6. R. Diamond, *Acta Crystallogr. Sect. A* **27**, 435 (1971); *J. Mol. Biol.* **82**, 371 (1974) (lysozyme).
7. K. D. Watenpaugh, L. C. Sieker, J. R. Herriott, L. H. Jensen, *Acta Crystallogr. Sect. B* **29**, 943 (1973).
8. M. Levitt and S. Lifson, *J. Mol. Biol.* **46**, 269 (1969); M. Levitt, *ibid.* **82**, 393 (1974) (lysozyme); P. K. Warme and H. A. Scheraga, *Biochemistry* **13**, 757 (1974) (lysozyme); D. Rasse, P. K. Warme, H. A. Scheraga, *Proc. Natl. Acad. Sci. U.S.A.* **71**, 3736 (1974) (rubredoxin); H. A. Scheraga, in *Peptides, Polypeptides and Proteins*, E. R. Blout, F. A. Bovey, M. Goodman, N. Loton, Eds. (Wiley, New York, 1974), p. 60; J. Hermans, Jr., and J. E. McQueen, Jr., *Acta Crystallogr. Sect. B* **30**, 730 (1974). Birktoft and Blow (*12*), in their work on chymotrypsin, also made use of an energy refinement procedure.
9. W. A. Hendrickson and J. Karle, *J. Biol. Chem.* **248**, 3327 (1973); W. A. Hendrickson, *J. Mol. Biol.* **91**, 226 (1975) (carp parvalbumin); J. E. Weinzierl, D. Eisenberg, R. E. Dickerson, *Acta Crystallogr. Sect. B* **25**, 380 (1969) (cytochrome c).
10. For some early work see W. Hoppe and J. Gassmann, *Ber. Bunsenges. Phys. Chem.* **68**, 808 (1964); G. Kartha, *Acta Crystallogr. Sect. A* **25**, S87 (1969). It is useful to distinguish two aspects of phase refinement and define it as the improvement of initial phases by either (i) direct use of algebraic relationships among structure factors or (ii) indirect use of electron density modification. Although we shall be dealing with the second aspect, the two are related by Fourier transformation.
11. P. C. Moews and R. H. Kretsinger, *J. Mol. Biol.* **91**, 201 (1975) (carp parvalbumin); J. Deisenhofer and W. Steigemann, *Acta Crystallogr. Sect. B* **31**, 238 (1975) (trypsin inhibitor); S. T. Freer, R. A. Alden, C. W. Carter, J. Kraut, *J. Biol. Chem.* **250**, 46 (1975) (high potential iron protein).
12. J. J. Birktoft and D. M. Blow, *J. Mol. Biol.* **68**, 187 (1972) (chymotrypsin).
13. D. Harker and J. S. Kasper, *Acta Crystallogr.* **1**, 70 (1948); J. Karle and H. Hauptman, *ibid.* **3**, 181 (1950); M. M. Woolfson, *An Introduction to Crystallography* (Cambridge Univ. Press, Cambridge, 1970), pp. 294–303.
14. D. Sayre, *Acta Crystallogr.* **5**, 60 (1952).
15. W. Cochran, *ibid.* **8**, 473 (19➡ J. Karle and H. Hauptman, *ibid.* **9**, 635 (1956); J. P. Declerq, G. Germain, P. Main, M. M. Woolfson, *Acta Crystallogr. Sect. A* **29**, 231 (1973).
16. D. Sayre, *Acta Crystallogr. Sect. A* **28**, 210 (1972); *ibid.* **30**, 180 (1974).
17. A. N. Barrett and M. Zwick, *ibid.* **27**, 6 (1971).
➡ J. W. Cooley and J. W. Tukey, *Math. Comput.* **19**, 297 (1965).
19. W. Hoppe and J. Gassmann, *Acta Crystallogr. Sect. B* **24**, 97 (1968).
20. The four criteria give practical expression to the probable boundaries of design for *f*. For more detailed discussion of these points see D. M. Collins, M. D. Brice, T. F. M. la Cour, M. J. Legg, in *Proceedings of the International Summer School on Computing, 1975*, F. Ahmed, Ed. (Munksgaard, Copenhagen, in press).
21. D. M. Collins, *Acta Crystallogr. Sect. A* **31**, 388 (1975). The problem of transform overlap (or aliasing) is treated in detail for crystallographic applications in D. Sayre, *Acta Crystallogr.* **4**, 362 (1951).
22. D. M. Collins, F. A. Cotton, E. E. Hazen, Jr., M. J. Legg, in *Proceedings of the Fourth Annual Harry Steenbock Symposium*, M. Sundaralingam and S. T. Rao, Eds., in press.
23. W. Hoppe, R. Huber, J. Gassmann, *Acta Crystallogr.* **16**, A4 (1963).
24. Before $|F_{calc}|$ is compared or combined with $|F_{exp}|$ they are placed on the same scale by multiplicative adjustment of $|F_{calc}|$. If, by symmetry, α_{calc} is limited to certain discrete values, then $|F_{exp}|$ is used as the corresponding structure amplitude in the computation of a "best" electron density map. For these calculations we have modified the computer programs reported in C. R. Hubbard, C. O. Quicksall, R. A. Jacobson, *The Fast Fourier Algorithm and the Programs ALFF, ALFFDP, ALFFPROJ, ALFFT and FRIEDEL* (Ames Laboratory Report, Ames, Iowa, 1971).
25. The disagreement factor, *R*, although defined here in the customary way, corresponds to the normal crystallographic figure of merit only in a formal sense. As we use *R*, the changes in its value are useful for judging convergence of phase refinement, but the level of its value is of uncertain significance. It is clear, for example, that the trivial modification $\rho_{calc}(\mathbf{r}) = \rho_{exp}(\mathbf{r})$ would lead to $R = 0.0$, although the resulting phase angles could not differ from the initial angles, however good or bad they might be.
26. A. Arnone, C. J. Bier, F. A. Cotton, V. W. Day, E. E. Hazen, Jr., D. C. Richardson, J. S. Richardson, A. Yonath, *J. Biol. Chem.* **246**, 2302 (1971).
27. F. M. Richards, *J. Mol. Biol.* **37**, 225 (1968).
28. C. D. Barry, H. E. Bosshard, R. A. Ellis, G. R. Marshall, *Fed. Proc.* **33**, 2368 (1974).
29. C. D. Barry and A. C. T. North, *Cold Spring Harbor Symp. Quant. Biol.* **36**, 577 (1972).
30. T. H. Jacobi, R. A. Ellis, J. M. Fritsek, *J. Mol. Biol.* **72**, 589 (1972).

31. E. F. Meyer, Jr., *et al.*, *Fed. Proc.* **33**, 2402 (1974).
32. L. Kätz and C. Levinthal, *Annu. Rev. Biophys. Bioeng.* **1**, 465 (1972).
33. A. Arnone *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **64**, 420 (1969).
34. J. R. Herriott, K. D. Watenpaugh, L. C. Sieker, L. H. Jensen, *J. Mol. Biol.* **80**, 423 (1973).
35. We use the average change in phase angle, $<\Delta\alpha>$, to corroborate the trend in R. Consequently, the ideal contribution to $<\Delta\alpha>$ from data without previous phase angles would be based on comparison of α_{calc} with random angles. The tabulated values of $<\Delta\alpha>$ include the approximately equivalent comparisons of α_{calc} with 0.0 for these data.
36. It is a pleasure and a duty to acknowledge contributions by several other persons. M. Brice and M. Legg have contributed to the development of the phase refinement procedure and to the practical application of this and the computer graphics to the rubredoxin and nuclease structures. S. M. Swanson and H. J. Bernstein (of Brookhaven National Laboratory) have participated in developing the computer graphics system. Financial support for this work came from the National Institutes of Health, Institute of General Medical Sciences (grant GM-19455 to F.A.C. and E.E.H., Jr.), the Robert A. Welch Foundation (grant A494 to F.A.C.), and the National Science Foundation (grant GJ33248X to E.F.M., Jr.).

The World Food Prospect

The worldwide food shortages of recent years, assumed to be temporary, could become more or less chronic.

Lester R. Brown

As we make the transition from the third to the final quarter of this century, the world food economy appears to be undergoing a fundamental transformation. Two developments stand out. One, the comfortable reserve of surplus stocks and excess production capacity which the world has enjoyed over the past generation may now be a passing incident in its history. Two, the world is becoming overwhelmingly dependent on North America for food supplies. These two changes point to a new role and responsibility for North America.

Within a span of a few years the world's surplus stocks and excess production capacity have largely disappeared. Today the entire world is living hand to mouth, trying to make it from one harvest to the next.

Grain exports from North America, a measure of growing worldwide food deficits, have doubled during the 1970's, expanding from 56 million tons in 1970 to nearly 100 million tons during the current fiscal year. Of the 115 countries for which data are readily available, all but a few now import grain. Of the countries that remain significant exporters, two dominate: the United States and Canada. During the current fiscal year the two together will export enough grain to feed the 600 million people of India.

The reasons for growing dependence on North American food supplies include ecological deterioration of food systems because of growing population pressure, mismanagement of agriculture, soaring population-induced demands, and sharp increases in demand as a result of newfound wealth, as in the Organization of Petroleum Exporting Countries (OPEC). The causes of the growing deficits vary, and often a combination of factors is responsible, but the effects are the same—ever greater pressure on North American food supplies.

As a result of these trends, North America today finds itself with a near monopoly of the world's exportable grain supplies. In a world of food scarcity, where there may not be enough food to go around, North America must decide who gets how much food and on what terms. The governments of the United States and Canada have not consciously sought this responsibility, any more than the countries of the Middle East have planned their geographical location astride the world's richest oil fields.

In recent years shortages of food have contributed to global double-digit inflation and to severe nutritional stress among low-income people everywhere. In some of the poorer countries, shortages have led to a rise in death rates, reversing postwar trends. National political leaders in the food-deficit countries, rich and poor alike, are becoming uneasy over future access to food supplies. Profound changes in the world food economy have brought into question the basic assumptions underlying North American food policies, particularly at the international level.

The author is president of the Worldwatch Institute, 1776 Massachusetts Avenue, NW, Washington, D.C. 20036. He is author of the forthcoming book *In Transition* which deals with future economic and population growth prospects.

New Sources of Global Food Insecurity

Throughout much of the period since World War II, the world has had two major food reserves: stocks of grain held by the principal exporting countries and cropland idled under farm programs in the United States. During the 1960's and early 1970's some 50 million acres out of a total U.S. cropland base of 350 million acres was held out of production to support prices (*1*). Stocks of grain held by the exporting countries were readily available for use when needed. Cropland idled under farm programs in the United States could be brought back into production within a year. Together grain stockpiles and cropland reserves provided security for all mankind, a cushion against any imaginable food disasters.

As recently as early 1972, it seemed likely that surplus stocks and cropland idled under farm programs would be part of the landscape for the foreseeable future. Then, suddenly, the global demand for food, fueled by the relentless growth of population and by rising affluence, began to outstrip the productive capacity of the world's farmers and fishermen. The world fish catch, which had tripled between 1950 and 1970 and had moved to a new high each year, turned downward for three consecutive years. Although most of the idled U.S. cropland was released for use in 1973 and the remainder thereafter, food reserves have not been rebuilt.

In 1961, the combination of reserve stocks of grain in exporting countries and idle cropland in the United States amounted to the equivalent of 105 days of world grain consumption. In 1972 stocks still equaled 69 days of world consumption. Then reserves began to drop rather abruptly—to 55 days in 1973 and still further to 33 days in 1974. The 1975 carry-over stocks remain precariously low, and all hopes for rebuilding them to safe levels have vanished with the poor 1975 Soviet harvest. Current U.S. Department of Agriculture estimates of carry-over stocks in 1976, already largely determined by the 1975 harvest, indicate an even lower level than in 1975 (*2*) (Table 1).

A third factor leading to global food insecurity and instability in the mid-1970's is