X-Ray Crystallography: The Past and Present of the Phase Problem[a]

David Sayre[b]

Abstract

The story of the progress thus far made on the phase problem of X-ray crystallography is briefly reviewed.

Keywords

X-ray crystallography • Phase problem

Introduction

In 1912–1913 there occurred (Laue in Munich, W.L. Bragg in Cambridge) one of the key ideas of twentieth-century experimental science: that of placing a small crystal as specimen in an appropriately short-wavelength (∼1 A) X-ray beam, recording the pattern and intensities of the diffraction spots produced, and using those data to accept or reject proposed atomic arrangements inside the specimen. By 1929 (again due largely to Bragg), it was recognized that if the diffraction spot phases could also be supplied, the phased magnitudes of the spots could be regarded as Fourier coefficients, whose Fourier sum would directly image the internal structure of the specimen; this would immediately give the atomic arrangement as well as other important information. From that time to the present, the problem of how to supply the phases has been the principal theoretical problem of X-ray crystallography. Progress has been such that today reasonably accurate phases and, therefore, images can usually be arrived at, even for extremely complex structures, provided that crystals are available and provided that desire is strong enough to produce continued trying of the methods which now exist for handling the problem.

In this paper, I give a brief and I hope illuminating account of the post-1929 history. For this purpose, I have divided the history into two parts: the original and still vigorous approach based upon crystals and Bragg spots, and a newer approach, based—in its major form—on noncrystalline specimens. These call for substantially higher exposure than crystals but give a continuous diffraction pattern, from which intensities can be measured on a finer sampling lattice[1] than that given by Bragg spots, permitting more data to be collected. (The specimen in the noncrystalline case may be as small as an individual protein molecule or as large as an individual biological cell, making possible an enormous increase in the range of specimens available to the crystallographic methodology. Also, in theory, it is possible to obtain a continuous pattern from a sufficiently small crystal.) This increase in data is proving in current testing to be often—perhaps always—sufficient to permit a rapid and easy technique of phasing and imaging. Thus the problems of crystallization and phasing promise to disappear in the newer technique, while the problem of damage, due to the increased exposure, will become more important.

It will be pointed out, as the paper progresses, that two general types of phasing methods exist: methods that are physical in nature and succeed in making actual phase

[a] *Structural Chemistry* 2002, 13(1):81–96.

[b] Department of Physics and Astronomy, SUNY (State University of New York), Stony Brook, NY 11794, USA (deceased)

[1] Specifically, sampling at intervals somewhat finer (for more detail see Part II) than the Nyquist sampling interval. The Bragg sampling interval in diffraction space applies to crystals and is the inverse of the unit cell diameter. The Nyquist sampling interval applies to any specimen and is the inverse of the specimen diameter.

I. Hargittai, B. Hargittai (eds.), *Science of Crystal Structures: Highlights in Crystallography*,
DOI 10.1007/978-3-319-19827-9_1

measurements; and methods that fundamentally operate by specifying constraints that may justifiably be imposed (usually in image space), and through mathematics applying the corresponding constraints which must exist on the phases in diffraction space. Both methods fundamentally aim at finding ways to repair an initial information deficit. The latter methods, in this day of high-speed computation, tend to be fast and convenient, but can also be subject to problems of nonuniqueness and possible unwanted biasing of solutions. The former are often more costly in time and trouble, but when an answer is found it is, within the limits of precision of the measurements, correct.

Part I: Crystals and Bragg Spots

Fourier Refinement (Early 1930s)

The major accomplishment of these first years was to bring out a valuable characteristic of the Fourier-coefficient approach to structure solving: its ability to find the correct phases and electron-density map when provided with merely an initial approximation to those quantities. The technique, called Fourier refinement, is an iterative one, in which current phases are used with experimental magnitudes to give a current image or map; the map suggests revisions in the atomic arrangement; the new atomic arrangement, through inverse Fourier transformation, generates new phases; and the process repeats. This back-and-forth ability immediately lent added strength to the 1913–1929 technique, in which possible atomic arrangements were tested for consistency with the observed magnitudes without, however, a method of corrective feedback. Now, with the feedback described above, if one got reasonably close to the structure, refinement would complete the job. In addition, well-refined electron-density maps, available now for the first time, provided genuine images of structure, giving important new information about bond shapes, presence and location of hydrogen atoms, etc. The technique, of course, is in use to this day.

It will be seen that this technique does not, in itself, contribute information. How then does it help to overcome the underlying information deficit present in the phase problem? The answer is that in each cycle the technique presents the known and guessed information in a form (the guessed map) which allows the investigator to add information to the situation. In practice, what the investigator does is to try to collect the electron density into about the right number of distinct atoms in a reasonable spatial configuration; the method is thus of the constraint-based, rather than phase-measuring, type mentioned above. One may legitimately argue that this may have the fault of biasing the technique toward an atomic theory of matter, or toward an incorrect atomic arrangement, but I will disregard that here. What can be stated is that degree of information insertion usually suffices to reach a satisfactory phasing and structure. An important conclusion thus appears: that there is at least a rough equivalency between the information carried in atomicity and the deficit of information represented by the missing phases.

The Patterson Map (1934), Heavy Atoms (Mid- and Later 1930s)

In 1934 the next step was taken by Arthur Lindo Patterson, a young crystallographer then at M.I.T., who saw a way to obtain benefit from the Fourier-coefficient approach *ab initio*, i.e., even before an initial approximation to phases and image is available. If the symbol F (F complex) designates the correctly phased Fourier coefficients, then even without the phases the quantities $|F|^2$ (i.e., the intensities) are known. It was Patterson's idea to begin by using the $|F|^2$ as Fourier coefficients. This, it is easy to show through the convolution theorem, yields a map which, if the true F-map shows a structure containing N atomic peaks, will show a related but more complex map (the Patterson map) consisting of N shifted and superposed copies of the true N-peak structure; equivalently, the Patterson map displays the N^2 interatomic vectors of the true map.

Disentangling the N copies is a formidable task if N is large, although some crystallographers (e.g., Dorothy Hodgkin at Oxford), gifted with great chemical insight and remarkable spatial perception, became adept at doing just that; some complex structures in the next two decades were, in fact, solved by a heroic process of guessing fragments of structure by partial Patterson disentanglement, and submitting those to Fourier refinement, until at last the true details of the structure began to emerge.

A much more convenient Patterson-based method, however, was discovered quite early. At about this time, the special case of structures in which a small number (say M) of the atoms were markedly heavier (higher atomic number Z) than the rest was beginning to attract attention (J. M. Robertson at the Royal Institution in London); if just the M-atom structure could be found or guessed, submitting it to Fourier refinement would often reveal the remainder of the structure. Now the Patterson method and this "heavy-atom" method could combine beautifully to make, for the first time, a systematic and fairly reliable *ab initio* structure-solving method. Looked at from the interatomic vector point of view of the Patterson map, the weight of a Patterson peak is proportional to the product of the weights of the two atoms giving rise to it; hence, it will usually be possible to recognize in the complicated Patterson map of the full structure the simpler Patterson map of the M heavier atoms, allowing the

M-atom structure to be solved first. Submitting the M-atom structure to Fourier refinement will then often allow the remainder of the structure to be solved. This use of the so-called heavy-atom Patterson method (Huse & Powell) became, for organic structures, the workhorse method of the 1940s and early 1950s. Given an organic molecule which it was desired to study, one employed chemical methods to attach, for example, a bromine atom to it, crystallized it, and then solved it by the heavy-atom Patterson method.

At this time, the era of solving protein structures was still years in the future. Yet in 1935, and directly related to the above developments, there was already recognition of what the future might bring. In 1934–1935, J. D. Bernal and Dorothy Hodgkin (then Dorothy Crowfoot), working in Cambridge, obtained the first good diffraction photographs of a crystalline protein (pepsin); in February, 1935, Crowfoot, now in Oxford, repeated the success using zinc insulin. When she informed Bernal of this, he replied with a brief note:

> Dear Dorothy:
>
> | | |
> |---|---|
> | Zn | 0.52 wt. per cent. |
> | Cd | 0.77 wt. per cent. |
> | Co | 0.44 wt. per cent. |
>
> This gives slightly less than 3 in all cases.

What Bernal was saying was that he had looked up the work of the Canadian chemist D. A. Scott, who had prepared zinc insulin crystals and measured their zinc content, and that for zinc, as well as for cadmium and cobalt, the data indicated the presence of three heavy atoms per insulin molecule. Bernal, in other words, was making the suggestion that despite their low concentrations, the heavy atoms might offer, through the methods outlined above, at some future date, a solution to the phase problem, even for proteins. That insight was ultimately to turn out to be 100 % correct.

It will be seen that, in a manner reminiscent of Fourier refinement, the Patterson method, while not in itself adding information, does assist the investigator in doing so. The adding of the heavy atoms, however, is at least a step toward a true phase-measuring method. For once the arrangement of the heavy atoms is known, they become (in the language of holography) the source of a known reference signal, which mixes coherently with the signal from the remainder of the structure and assists in phasing it. As stated, the early technique relied simply on Fourier refinement to carry out the phasing and imaging. Since then, under the driving force of protein crystallography, that technique has given way to better ones (see sections below on macromolecules), but the idea that this might someday be done was already present at the time being discussed. It may also be noted how very neat an arrangement it is to have the reference signal source (i.e., the M special atoms) repeated in every unit cell of the crystal. Because of the strength of the Bragg spots, a strong reference signal is needed; the repeated arrangement produces that strong coherent signal exactly at the points where it is needed, i.e., at the diffraction spots.

Three Brief Digressions

The State of Affairs at the Start of the Postwar Period (1946)

It may be useful to summarize the overall situation in crystallography as the world prepared to return to more normal conditions after World War II. The field was small but healthy, with a considerable record already of elucidating the atomic structure of matter in inorganic compounds, minerals, metals and alloys, and organic compounds, in the latter area reaching to sterols and (within the next several years) to penicillin. For its time, the field had an adequate X-ray source (the metal target X-ray tube) and an adequate X-ray detector (X-ray film); in the Patterson map with and without heavy atoms, and Fourier refinement, it had a workable system of solving structures; and it had a seemingly inexhaustible supply of interesting crystalline materials to study. It even had (although only a handful were as yet thinking about it) the key to the future imaging of the highly complex structures which make up the living cell.

Crystallography did have a problem, however, and it lay in the area of computation. For the carrying out of Fourier sums (from $|F|^2$'s to Patterson maps and from F's to Fourier maps), the slowness of the prevailing method— the use of hand-operated office-type mechanical adding machines plus Beevers–Lipson or Patterson–Tunell strips (sine–cosine tables formatted for the Fourier operation)—was such that only two-dimensional (2D) maps were normally done, and even those would take a day or two each. As long as crystallography could not freely enter three-dimensional (3D) work, its potential for the future would be severely limited. Fortunately, the wartime advances in electronics would quite soon change this situation.

A Bit of Personal History

In 1943 I emerged from Yale University with an undergraduate degree in physics, and in 1943–1946 I did wartime electronics research in radar at the Radiation Laboratory of M.I.T. In 1946, guessing that biology might become the exciting field of the next half-century, I became a graduate student in biology at Pennsylvania and then Harvard, but did not find myself responding with interest to what I was learning. However, in June 1947, I happened upon J. M. Robertson's review article on X-ray crystallography [1], written shortly before the war began, which changed my life.

I joined the crystallography laboratory of Raymond Pepinsky at Alabama Polytechnic Institute, where I participated in the design of X-RAC (X-Ray Automatic Computer), an electronic analog computer which, once it had been given the $|F|^2$'s or F's, could compute 2D Patterson and Fourier maps in 1 s. (That machine remained in productive operation for some years, until it was supplanted by digital computation). When in 1949 the machine was successfully operating, my wife Anne and I went to Oxford, so that I could be a Ph. D. student with Dorothy Hodgkin.

The Beginnings of Digital Computation (Late 1940s, Early 1950s)

Digital computation, of course, was now just around the corner. In 1949 England was a little in the lead in that field, and so, not long after arriving in Oxford, I went to Manchester to visit Prof. Fred Williams (with whom I had become acquainted through radar work during the war) and to see his computer, MADAM, which had just recently started to work, with the idea that I might perhaps do crystallographic computing on that machine as my thesis work at Oxford. That did not happen and it was in fact Durward Cruickshank in Leeds who beautifully carried out such a program on the Manchester computer a little later (as did Bennett and Kendrew on EDSAC in Cambridge). In one of the major moments of my life, however, Williams handed me over to Alan Turing, who was then part of the Manchester effort, to discuss the idea. Instead, as it turned out, we talked for the better part of 2 days about the phase problem, which clearly intrigued him. What he gave me, in return, was an early look at the Shannon sampling theorem, which led me to some early thoughts on oversampling in crystallography [2] and undoubtedly helped prepare my mind for the Nyquist-basis concept many years later. As for crystallographic programming of a digital machine, I did do some later, but not until 5 years had gone by, when I did it on the IBM 701 machine, by which time the ability of the digital computer to handle crystallography's computing needs was quite well established.

Direct Methods (Late 1940s–Present)

In 1948, 14 years after Patterson, a major new advance in the phase problem occurred when David Harker and John Kasper of the General Electric Research Laboratory applied Schwarz and Cauchy inequalities to the expressions for the F's and derived inequality relationships which must exist between the magnitudes of some F's and the phases of others; 2 years later they used these inequalities to determine the values of enough phases to allow a quite difficult structure, dekaborane, to be solved. These papers [3] caused a major stir in crystallography; for 20 years it had been considered that the phases are lost, but now it appeared that they are not entirely lost, but are in some degree capable of being deduced from the magnitudes.

For several years, it was not really known how best to understand the Harker–Kasper work. Then in January 1950, in an important paper [4], Jerome Karle and Herbert Hauptman at the Naval Research Laboratory in Washington gave a general form of inequality (the Karle–Hauptman inequality) which provides all the inequalities of the Harker–Kasper type that can exist; in addition they showed that what those inequalities actually express is the nonnegativity of the electron density map. To me, now a student at Oxford, this was important news. It said that whereas the positivity condition had been captured in the Harker–Kasper and Karle–Hauptman inequalities, the atomicity condition (remember that the information in atomicity roughly balances the information in the missing phases) had not as yet been caught. To find a way to express mathematically, in terms of the F's, the concept of atomicity, became my goal.

We now know that this mathematization of atomicity has been done essentially twice, in two different ways, I did it first, by adopting an idealization of the problem in which the atoms are equal and resolved; this permits a complete mathematization, in the form of a set of equations to be satisfied by the F's. Solving the phase problem then reduces to inserting the measured |F|-values in the equations (called the squaring-method equations) and solving for the phases. The latter is not a trivial problem, but based on the structure of the equations, I developed a solution process consisting of three main steps.[2] I gave these results orally at the Third International Congress of Crystallography in Stockholm in June 1951 and in a paper [5], which appeared in *Acta Crystallographica* in January 1952. Then, in 1972, for the field of protein crystallography, I developed a second solution process, in the form of finding (by least squares) an extremum of a functional based on the equation system, and applied it successfully to the structure of rubredoxin in 1974 [6]. Finally, in 2000, Richard Rothbauer, still within the framework of exact mathematization, showed [7] that the extremum method can be extended in such a way as to allow the initial idealization to be dropped.

Returning now to the early 1950s, Karle and Hauptman sought to dispense with the idealization which I had used, so as to bring the study closer to real life. Their work extended over numerous papers, beginning [8] with Monograph #3 of the American Crystallographic Association, which appeared in September 1953, and over many years. The mathematics

[2]The steps are: (1) identification and tentative phasings of strong terms of the equations (these terms, which involve three F's whose Fourier indexes sum to zero, are known as "triplets"); (2) extension of these phasings to many or most of the weaker F's through further use of the equations; (3) use of an equation-based figure of merit and choice of a final phasing.

which they chose for this purpose is the mathematics of probability, giving their work a very different character from mine. Nevertheless, when it came to the application of the results in the form of a computational procedure, the Karle–Hauptman approach gradually came to a closely similar three-step process, their version of the second step (which is really the critical one of the three) appearing under the name "tangent formula" in a paper which they published in 1956. Much later, in the 1980s, the introduction of a newer probabilistic machinery (maximum likelihood) to replace the original machinery used by Karle and Hauptman has been carried out by Bricogne and others.[3] Also Hauptman and his co-workers, starting in 1990, have developed an extremum methodology for protein crystallography ("Shake'n Bake"), which resembles the 1972–1974 work which I did, but with the very useful insertion of automated Fourier refinement into the process as a means of securing good final atomicity [10].

Comparing the two lines of work, there is no doubt that the equal-atom idealization disqualifies the Sayre approach in some cases (especially in the field of inorganic structures) where the Karle–Hauptman approach may succeed. With organic structures, however, idealization does not severely penalize the Sayre approach, while its equational form adds significant general strength, so that, in this area, the difference, although there is a shortage of serious testing, is probably not great. In any event, the true choice today, I think, lies between the maximum likelihood approach on the probabilistic side and the methodology of Rothbauer on the exact side.

In protein studies, the subject of structural genomics is now producing an urgent desire to see a large productivity increase come into being, similar to the one produced by the atomicity-based direct methods in small-structure crystallography. Atomicity-based methods are thus starting to play a small role in protein crystallography, but their role may be fundamentally limited by the fact that protein data sets in which the |F|'s extend to atomic resolution are as yet rather rare.[4] In anticipation of the Nyquist-based technique, one of its most important potential advantages is that it can use, but does not require, atomic resolution data.

[3]The original machinery suffered from probability distributions (e.g., of phases of strong triplets) being computed too locally (e.g., from the three |F|'s only), and not being recomputed as phasing proceeded. Hauptman introduced strengthenings of the machinery in the late 1970s through his "Neighborhood Principle," but subsequently seems to have ceased from that work. (See Bricogne [9c] for details of the newer techniques.)

[4]In the extremum methodology work which I did in 1974 on rubredoxin, the data set had 1.5 Å resolution, whereas both maximum likelihood and Shake'n Bake today are reported to require somewhat higher resolution data. A Shake'n Bake, but with the direct method phasing being done by a Rothbauer process, might thus prove to be of interest. Another possible direction (Bricogne [9b]), but probably a very difficult one, would be to develop direct methods in which larger entities such as the amino acid residue, rather than the single atom, play the central role.

Finally, as between phase-measuring and constraint-based methods, the direct methods are, of course, in the latter group.

Macromolecules: DNA (1953), Isomorphous Replacement (1953–Present)

In 1953, only a little more than a year after the introduction of the atomicity-based direct method into small-structure crystallography, came the first two decisive events for macromolecules. Both took place in Cambridge, England.

The first, the working-out of the double-helical structure of DNA, did not really mark an advance in the treatment of the phase problem; however it is so important that it cannot be overlooked. It was in fact an outstanding example of the 1913–1929 method of structure solving, in which Crick and Watson utilized several lines of work—the classical chemistry on DNA, the 1950 work on DNA base ratios (Chargaff), and the 1952 theoretical work on the diffraction patterns of helical arrangements of atoms (Cochran, Crick, and Vand)—to propose a structure (the double-helical based-paired DNA structure) which beautifully accounted for the critical 1951 diffraction data obtained at Kings College London on the B-form of DNA (Franklin). Phase-finding, in other words, was not the solution method here (although Wilkins and his colleagues at Kings did subsequently move on to Fourier refinement of the DNA structure).

The second event was the 1953 discovery by Perutz and his co-workers of the phenomenon of isomorphous replacement in protein crystals, i.e., the ability of proteins to attach heavy atoms or heavy-atom complexes to specific protein sites, and to do so with little or no alteration of the crystal structure as a whole. The ability largely rests on the fact that protein crystals contain considerable regions of solvent, meaning that minor rearrangement of solvent can often succeed in maintaining crystal structure, even after the heavy atom addition. An important consequence of this mechanism is that *several* attachment patterns, often involving different types of heavy atoms, can often be prepared; this is the phenomenon of *multiple* isomorphous replacement (MIR).

The important effect of all this is that in the protein setting, the Patterson heavy-atom method is strengthened to the point where it can operate successfully even with the very low heavy-atom concentrations first noticed in insulin by Bernal in 1935.

This strengthening arises in two ways. There is, first, the fact that, by studying the *difference* map between the Pattersons of a native protein structure and of an isomorphous heavy-atom structure, most of the background material disappears, allowing the heavy-atom peaks belonging to the isomorphous structure to show up sufficiently well to permit the heavy-atom portion of that structure to be solved. Then, second, the existence of multiple isomorphous derivatives means

that, in the different diffraction patterns, different reference signals have been mixed holographically with the signal from the native protein; in 1956 it was shown (Harker) that with as few as two isomorphous derivatives, this is sufficient, in principle, to determine uniquely the phases of the native protein. With further derivatives present, the accuracy of the phases can usually be further improved, a subject on which much work has been done (Blow and Crick; Hendrickson and Lattman). There has also been much work on obtaining the coordinates of the heavy atoms in the various derivatives as exactly as possible (Blow and Matthews; Bricogne). Progress in these topics has by no means stopped today.

Perutz and his co-workers were working on hemoglobin, but in fact it was Kendrew (again in Cambridge) and his coworkers, working with MIR on myoglobin, who, in 1957 first succeeded in fully carrying out the technique; this success was then soon repeated with other proteins, including hemoglobin. Thus it was through MIR phasing that a collection of solved protein structures first began to be accumulated.

Macromolecules: Anomalous Dispersion (1949–Present)

Valuable as MIR phasing is, it has the disadvantage that, even when starting with good crystals of the native protein, months or years can be spent in finding good isomorphous heavy-atom derivative crystals. Fortunately, at about the same time as MIR was being developed, an alternative was also coming into view. Isomorphous replacement operates by changing the chemical identity of a few of the scattering atoms; anomalous dispersion accomplishes essentially the same thing (and with perfect isomorphism), using just a single crystal form, by wavelength rather than chemical change. The day would come when, through synchrotron X-ray sources, accurately choosable wavelengths for this purpose would be quite readily available.

The observation (Nishikawa; Knol and Prins) that the diffraction patterns of crystals present additional data through development of Friedel nonsymmetry at wavelengths near X-ray absorption edges goes back to the same period in the late 1920s as the emerging of the phase problem. As nearly as I can tell, the first utilization of this effect in connection with the phase problem occurred 20 years later (Bijvoet) with the settling of the choice of the two possible enantiomorphic phasings for a rubidium salt of tartaric acid. (This established, for the first time, the absolute handedness of an optically active molecule). Then, in the same year that Harker gave the underlying analysis establishing the MIR technique, a sort of parallel analysis (Okaya and Pepinsky) gave some of the underlying ideas of what would later become the multiple-wavelength anomalous dispersion technique (MAD).

Despite that paper, the lack of an appropriate X-ray source delayed the application of the ideas for a quarter-century; interest, however, revived again with the promise of the synchrotron as a suitable source (Karle; papers from the Dutch school following Bijvoet, etc.). Finally the MAD method appeared essentially in its present form (Hendrickson), and is now the leading phasing method in macromolecular crystallography. The method calls for a specimen crystal containing a substructure of a moderate number of atoms of a species having an X-ray absorption edge in the general 1 A wavelength region; such a specimen can often be made by the usual heavy-atom addition methods, but a favorite method for protein studies is to substitute selenium methionine for normal methionine in an organism's diet, causing the organism to produce protein which differs from the normal protein only in the fact that the methionine sulfur atoms have been replaced by selenium. The MAD method is rapid and produces maps of excellent quality, and (as stated above) is now the dominant method in protein crystallography.

Both anomalous dispersion and isomorphous replacement, by providing for the coherent mixing of the diffraction signal from the major structure with known signals from one or several reference structures, fall in the favorable category of actual phase-measuring techniques.

Macromolecules: Additional Methods

In isomorphous replacement and anomalous dispersion we have now treated the key methods in macromolecular crystallography, but large numbers of structures have been worked out by other methods as well. In this section, I will briefly touch on some of these methods.

1. Once a protein structure has been solved, the structure of the same protein after addition e.g. of an inhibitor or substrate molecule becomes of interest. That structure can usually be found quite easily by adding to the original map a further (so-called difference-Fourier) map, which can be calculated with fair accuracy from the diffraction intensity differences between the native and modified proteins. The technique, first used by Stryer, Kendrew, and Watson at Cambridge, is in a sense an inverted version of MIR, and its theory resembles the Harker theory for MIR.
2. Similarly, given that the structure of a particular protein is known, there are usually numerous other proteins with closely related amino acid sequences that are of interest as well. Operating on the assumption that similarity of sequence may imply similarity of structure, it is reasonable to determine whether the known structure can serve as a starting point in solving the new structure. In the computer, the known structure is tried in varying orientations and positions in the unit cell of the new crystal; if, at some

orientation and position, the computed diffraction pattern or its Patterson resembles the new observed pattern or its Patterson, a starting point that will successfully refine to the new structure may have been found. This method, sometimes known as molecular replacement, also originated in Cambridge (Rossmann and Blow).

3. Protein crystallography presents a few special situations in which, even within the constraint of observing at Bragg spots only, a limited degree of extra sampling of a diffraction pattern can occur. Crystals may contain considerable solvent, separating the molecules by more than their natural diameters, and causing the Bragg spots to sample at higher fineness than would otherwise be the case; in addition, there may be swelling and shrinking of a crystal form with changing amounts of solvent, causing the Bragg spots to shift their positions along the pattern, and in that way providing further sampling; or a protein may crystallize in several different crystal forms, thus providing several different Bragg-spot samplings of its pattern; or a protein, and especially a virus, may be large, with correspondingly fine-spaced Bragg spots, and yet really consist of N identical subunits, meaning that the *subunit* pattern is being sampled N times more frequently than would otherwise occur. Methods by which the increased information obtained in these situations can be used to assist in handling the phase problem are the subject of a considerable literature, once again due largely to Cambridge, or formerly Cambridge, personnel, including such names as Main, Rossmann, Blow, Bricogne, and Crowther. (Terms applied to these methods include solvent methods, change of space group methods, and noncrystallographic symmetry methods; a unified maximum-entropy treatment of many of these methods has been given by Bricogne [9a].) The methods bear a relationship to that which will be discussed in Part II, but are limited and made more complicated by the Bragg-spot constraint. Somewhat similarly, Szoke [11a] has developed a number of additional structure-solving techniques based on the viewpoint that a diffraction pattern can always be regarded as a hologram, with the signal from one part of the structure serving as reference signal for the remainder of the structure.
4. In addition to the above, there are still other worthwhile concepts which are attracting effort (see summary of a conference [12] on new phasing approaches held May 17–19 2001 at the Lawrence Berkeley Laboratory).

One More Method: Three-Beam Diffraction (1977–Present)

A very attractive conceptual approach to the phase problem in crystals is to excite two different reflection wavefields which will propagate in the same direction, allowing them to interfere coherently; then, from the resultant observable intensity, it may be possible to infer phase information about the crystal. As was pointed out by Lipscomb in 1949, the excitation can be realized by having two reflections h and g simultaneously on the Ewald sphere; these will generate reflected wavefields in directions K(h) and K(g); examination will show that through geometrical identities the reflections g–h and h–g will be on the Ewald spheres for the K(h) and K(g) wavefields; these will produce two additional, doubly reflected, wavefields in the original directions K(h) and K(g), bringing about the desired situation in both of those directions. Experimental evidence for this was then provided by Hart and Lang in 1961. It remained for Post and Colella, working independently in 1977, to suggest the details of extracting the phase information. They suggested fixing h while slowly causing g to move through the Ewald sphere; as this occurs the doubly reflected (g, h–g) wavefield passes through a resonance, which, in turn, produces an intensity trace for the combined K(h) signal; from the shape of this trace the sum of the phases of the $-h$, g and h–g reflections can be read off. But $-h+g+h-g=0$, so the three reflections constitute a "triplet" in the sense described in the section on Direct Methods (this paper). Thus three-beam diffraction allows, in principle, the experimental measurement of the same quantities (triplet phases) as can be estimated from direct-method arguments in cases where atomicity is present. From that point on, the details merge with those of direct methods. The method falls, however, in the true phase-measurement category.

Practical experience with the method is fairly encouraging in that triplet phases have been successfully measured for a number of small molecules and proteins. Proteins, however, pose several problems: their large unit cell means that care must be taken to avoid having several g's contributing overlapping interference traces; mosaicity is desired to be low for the three-beam method, but tends to be high in proteins; and cryoprotection, which is needed because of the long exposures involved, tends to be ruled out because of its effect in increasing mosaicity still further. Nevertheless, groups in Karlsruhe (Weckert et al.) and Cornell (Shen et al.) have, in several cases, succeeded in overcoming these difficulties.

The Sparse or "Dream" Crystal

Should it ever become possible to stabilize a crystal by forces other than contact forces, it might be possible to create the ideal for X-ray crystallography; this would be a crystal in which the structure is regularly repeated at distances somewhat (perhaps 30–40 %) greater than contact distances.

This would provide oversampling for phasing (see Part II), while retaining the large amplification enjoyed by Bragg spots (permitting high resolution at moderate exposure levels).

An apparatus for virtually instantaneous formation and stabilization of 2D crystals in a standing-wave field has been described by Golovchenko and his colleagues [13]; in addition, orientational alignment of molecules by the use of elliptically polarized laser fields has recently been demonstrated by Larsen and his colleagues [14]. If such apparatus could be developed to generate accurate 3D crystals, then in situations like protein crystallography, where large numbers of exact copies of a structure are available, a high-throughput (i.e., free of slow crystallization, free of phasing problems) and high-resolution crystallography would immediately become possible.

Part II: Finer Sampling (1980–Present)

However, although the apparatus just noted would thus be extremely desirable, it may never come about. More serious still, many of the most important structures do not exist in multiple exact copies (e.g., major cellular constituents and whole cells in biology, amorphous structures and structures with imperfections in materials science). Fortunately, however, it is beginning to appear that there exists an adaptation of present-day crystallography which can operate successfully even in those situations. This is the subject of the remainder of this article.

Overview and Brief History

In Part I, the attitude taken was that the diffraction pattern consists solely of spatially discrete Bragg spots. In Part II we drop this attitude and consider instead that diffraction intensity may exist at every point of diffraction space. This view is mathematically correct whenever the diffracting specimen is of finite size (and when "point" in the sentence above is replaced by "elementary volume"). It is nearest to being incorrect (i.e., Bragg spots are most dominant) when the finite specimen is a sizable and well-formed crystal such as is assumed in Part I; it becomes more correct as we move to imperfect or very small crystals; and it is most correct (to the point where Bragg spots may give way entirely to continuous pattern) when the specimen is noncrystalline. Thus, in Part II, we are specifically considering specimens which are *not* excellently crystalline. Three major things result:

1. Our base of usable specimens expands enormously and we are relieved of the necessity of crystallization;
2. The diffraction features we now seek to record include many which are much weaker than those in Part I, making necessary a much higher level of X-ray exposure, with consequent specimen damage and danger of lowered image quality; and
3. The patterns, being continuous and capable of being finely sampled, yield more information than the Bragg-spot-only patterns of Part I; current evidence is that the added information allows the phase problem to largely, if not fully, disappear.

Points (1) and (3) are major gains for X-ray crystallography, point (2) is a major, but probably not disabling, problem.

These perceptions, although they seem obvious now, came slowly and with difficulty. I have mentioned (see section on Three Brief Digressions) a paper [2] which grew out of my conversation with Turing; it pointed out that if sampling could be carried out at twice the Bragg fineness, the phase problem (at least for centrosymmetric crystals) would effectively go away. The paper did not, however, suggest an effective way of obtaining such sampling; the situation was still conceived of as referring to crystalline matter and sampling at Bragg positions only. The suggestion that crystallographers might put a specimen like a single biological cell in an X-ray beam and nevertheless treat it essentially as if it were a crystal—collecting, phasing and Fourier inverting its *continuous* diffraction pattern—was not made[5] until a paper [15] written in 1980.

Starting, therefore, in 1980, the question became whether that process could actually be carried out. Here the principal issues were whether the diffraction patterns, which would be very weak, could be successfully observed, and whether they could be phased. At that time the answer to both questions was probably no; X-ray sources and detectors, and, similarly, the computers of the time, were not adequate to the tasks. Fortunately, however, the Brookhaven National Laboratory, only 20 miles distant from Stony Brook, decided shortly before 1980 to build a major synchrotron X-ray facility, with the result that it has been possible for us to enjoy the full advantage of the subsequent advances in

[5]From 1955 to 1970, I did not do much work in crystallography, working instead on early computer projects at IBM. In 1970, however, still at IBM, I returned to crystallography, and did the work on the extremum formulation of direct methods mentioned earlier [see section on Direct Methods]. Following that, and feeling that an extension to noncrystalline objects would be an important future need in crystallography, I began to work in that direction, soon forming a close collaboration with Janos Kirz at SUNY Stony Brook. Our initial approach, which has been excellently realized by Janos and his students, was for an X-ray microscope, based on a Fresnel zone plate as an image-forming device. It was not until about 1979 that it occurred to me to turn to crystallography itself for a methodology.

needed technologies. Stages along the way have included: the demonstration, using early bending-magnet radiation, that the diffraction pattern from a large and contrasty single cell (diatom) could be recorded [16]; demonstration, using undulator radiation, that the pattern from a small typical single cell could be recorded [17] and proposal[6] of oversampling of the continuous pattern as a possible phasing method [17]; mathematical treatment of the exposure levels required for pattern observation and first tests of oversampling as a phasing method [18]; general article expressing the probability that this extended version of crystallography could exist [19]; and improved theory of oversampling [20]. A decisive point was achieved with the successful demonstration of the complete process in a 2D nonbiological situation [21]. Following that, there have been three pieces of work emphasizing potential applications: demonstration, with simulated data, of oversampling phasing for amorphous specimens, very small crystals, and imperfect crystals [22]; preparation of a detailed grant proposal for a full 3D project on a cryoprotected yeast cell; and [23] successful acquisition (simulated) of a 3D high-resolution large-protein data set without need for crystallization by the method of femtosecond-scale FEL laser pulses [24], followed by phasing and imaging by oversampling.

With that as background, there follows a brief discussion of some of the issues involved in collecting the diffraction pattern (next Section), and in phasing it (see section on Phasing the Pattern).

The Two Probable Experimental Regimes

The radiation dosage needed to obtain a statistically accurate diffraction pattern rises steeply (roughly 8th-power law) with the desired imaging resolution, rough numbers for organic materials being 10^8, 10^{16}, and 10^{24} rads for 20, 2, and 0.2 nm resolution, respectively [18]. Especially at the higher resolutions, these dosages can vastly exceed the specimen's ability to tolerate dosage. Our major defenses today against this problem are crystallinity, replaceability (ability to replace the specimen with an identical fresh one), cryo, and flash. Crystallinity, in conjunction with Part I phasing techniques, works excellently; it reduces the dosage for Bragg spot collection by some 12 orders of magnitude, i.e., to about 10^{12} rads for 0.2-nm resolution work, which, with the assistance of replaceability and cryo, is acceptable; this is the basis for present-day protein crystallography. Where crystallinity does not exist or is intentionally bypassed, one can turn to Part II phasing, but it has only replaceability, cryo, and flash to work with. The result for Part II is the probable emergence of two distinct experimental regimes. With replaceability, flash can be used and it promises to allow the very large dosage needed for atomic or near-atomic resolution, accomplishing the exposure before damage becomes evident [24]. Without replaceability there is only cryo, which should, however, allow dosages to roughly 10^9 rads, or roughly 15 nm resolution, on micrometer size-range specimens (such as the 3000-nm frozen-hydrated yeast cell of our planned project), hopefully making it possible for crystallography to contribute to cell as well as molecular biology. There follows a little more detail on these two regimes.

The Free-Electron Laser-Flash Technique for Molecules

As mentioned above, a recent paper [24] reports computer simulation studies indicating that if the needed dosage for obtaining the diffraction pattern from a protein molecule is delivered to the molecule in a time period of 5×10^{-15} s or less, significant structural alterations in the molecule will not have occurred by the time the pattern has been generated. Furthermore, it appears [25] that it should, in time, be possible to build a free electron laser (FEL) of that pulse width and of the needed photon flux at 1.5 A wavelength. As a result, a consortium of research teams (including our own) under the leadership of Prof. Hajdu has submitted a research proposal for a flash-technique project on high-resolution imaging of large biomolecules at the Stanford Linac Coherent Light Source, based on the possibility that such a laser may in the future be constructed. The research proposal envisages the delivery of single molecules (or very small molecular clusters), in random orientation and with no other matter nearby to confuse the pattern, into the laser beam. In further support of this, in a recent paper [23] it was assumed that the 2D diffraction patterns from a succession of 10^6 such randomly oriented molecules of the protein rubisco delivered into the beam of an assumed 1.5 A wavelength FEL laser had been assembled into a 3D oversampled data set at 2.5 A resolution; that dataset was then successfully phased by the method to be described in the section on Phasing the Pattern. At the planned repetition rate of 120 Hz for the flashes, the 10^6 2D patterns would be acquired in about 2.3 h.

A Synchrotron/Cryoprotection Technique for Supramolecular Specimens

This technique, as currently employed by us on the X1B undulator beamline at the National Synchrotron Light Source, is diagrammed in Fig. 1. The undulator, entrance slit, grating, and exit slit deliver tunable monochromatic photons to the diffraction experiment, the tuning range

[6]This proposal was developed in a conversation in the late 1980s with Gerard Bricogne.

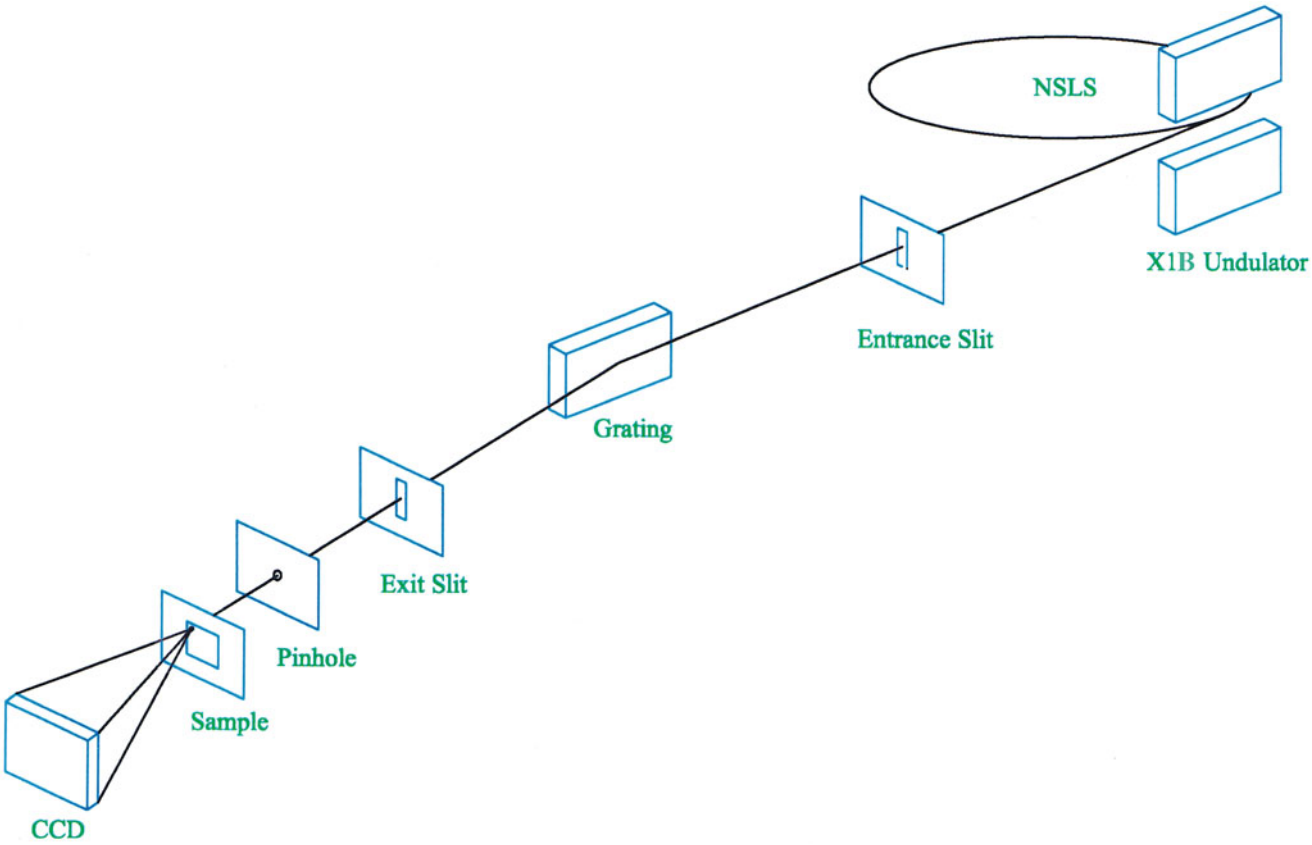

Fig. 1 Simplified schematic of the synchrotron-based experiment at NSLS for supramolecular specimens.

being approximately 1.5–5 nm wavelength.[7] The experiment begins with a 10-μm pinhole approximately 10 cm upstream from the specimen, narrowing the beam so that a single specimen can be illuminated on the specimen holder, and also ensuring that all photons reaching the specimen are monodirectional to 0.1 mrad. (The monochromaticity and monodirectionality produce a sharp Ewald sphere, which will allow future software to place each pixel of the observed 2D patterns accurately in the final 3D pattern). In Fig. 1, the specimen is shown as being close to a corner of the specimen holder; this effectively screens three quadrants of the detector from any edge-scatter or Airy pattern arising from the pinhole; i.e., in those three quadrants, photons emanating from the specimen are essentially falling into a totally dark area of the detector, allowing the detection of very weak pattern. A few inches further downstream is the detector, a back-thinned liquid-nitrogen cooled CCD detector. To catch a 2.34-nm pattern extending to, e.g., 15-nm resolution, the detector must subtend a full angle of about 18° at the specimen. Not shown in the diagram are a beam catch just forward of the detector (to protect the central area of the CCD from the intense central beam), and a silicon nitride window just upstream of the pinhole (to protect the UHV of the beamline from the moderate vacuum of the experiment). Also not shown is the apparatus for positioning and setting the specimen to its successive orientations and, for cryo work, for cooling it. A variant arrangement (but as yet untried) would use long-throw focusing instead of the pinhole to accomplish the beam narrowing.

In addition to subtending a sufficient angle at the specimen, the detector must have enough recording pixels to allow for sufficiently fine sampling of the pattern falling on it. If, in

[7] This region, falling in the soft X-ray region, is not mandatory (there is no reason why X-rays down to 1 Å wavelength cannot be used), but it gives adequately short wavelengths for the resolution expected in this technique, and offers valuable contrast mechanisms for biological work at the absorption edges of low-Z atoms. Thus, for the yeast cell project, we plan to use 2.34-nm photons—just on the low-energy side of the oxygen absorption edge—for best transparency through the ice portion of the frozen-hydrated cell.

the example above, the specimen has a diameter of 3000 nm, sampling at twice the Nyquist fineness—this is usually somewhat finer than necessary; see next section—means sampling at intervals of 1/6000 nm, while the extent of the pattern is 2/15 nm. Thus, 800×800 pixels on the CCD will suffice. Back-thinned CCDs meeting these requirements are available today. The full 3D data set ultimately assembled, in this case, could approach 800×800×800 numbers.

A problem of systematically missing data arises with the technique as described. The need for a beam catch to protect the CCD produces missing data near the central beam; the loss could be reduced if one or a few pixels at the center of the CCD were removed or made transparent. In addition, in 3D work the planar specimen mount produces in the data the undesirable effect of a missing double cone normal to the plane of the mount. At present (see next section), we rely on the phasing algorithm to overcome these data losses.

Phasing the Pattern

It remains to explain how the finer sampling can be used in a phasing process. Turning back to the start of Part II, we pass by the first case of the finite specimen (the sizable well-formed crystal) as providing no opportunity for finer sampling[8] and as having been the subject-matter of Part I. We next defer, for a time, the second case (the imperfect or very small crystal), and turn directly to the major case (the non-crystal). Fourier theory then tells us that if the diffraction pattern sampling is at the Nyquist fineness, and the sampled values (correctly phased) are Fourier summed, the result will be correct images of the specimen indefinitely repeated and just in contact with each other. If the sampling is finer than Nyquist, and the values (again correctly phased) are Fourier summed, the images are still correct and repeated, but are now not in contact, thus giving rise to zero regions between them. It is the presence of these regions, which should be zero, which drives the phasing. The sampling is made sufficiently fine that the number of voxels in image space between the specimen envelopes somewhat exceeds the number of voxels inside the specimen envelopes. Then, starting with e.g. random phases, an iterative algorithm, repeatedly moving via Fourier transformation between the two spaces, alternately (1) in image space pushes the voxels between specimen envelopes toward zero and (2) resets pattern magnitudes in diffraction space to their experimentally observed values. Our experience is that usually, after several hundred to several thousand iterations, the interenvelope voxels are essentially zero and the intraenvelope voxels and the phases are essentially correct (see Figs. 2 and 3). Arriving at the correct phasing can be assisted by adding appropriate positivity constraints on the real and imaginary parts of intraenvelope voxels [20]. The algorithm is basically a Fienup-type algorithm [26] and is attractively simple and fast, being an alternation of simple processing in one space followed by high-speed Fourier transformation to the other space.

This process clearly falls in the constraint-based category of phasing methods, and is thus subject to the concern that it may produce incorrect, or near-correct but inaccurate, solutions; this concern is perhaps here further heightened by the fact that in all real cases there can be data inadequacies, such as data error, too much missing data, possibly insufficiently fine sampling, and inaccuracies in the assumed specimen envelope. I think a fair summary of what has been learned thus far—mainly through computer simulation studies[9]—is that with mild data inadequacies (levels that can probably be reached with careful experimentation) the probability of correct and reasonably high-quality phasing and imaging is high, but that there may not be a large margin of safety[10] in this respect. At this point, however, these matters require more study and experience.

Finally, in a recent paper [22], it is shown that the phasing process also works for the second case, the imperfect or very small crystal. (Insofar as this case has a difficulty, it lies in the experimental portion of the work, and arises from the Bragg spots still being strong and taking intensity away from the non-Bragg Nyquist points, making the latter more difficult to measure.) Here the process could be of considerable

[8]However, see Szoke [11b] who suggests that the use of partially coherent radiation in the crystal case will cause the Bragg spots to become broader and allow additional information to be observed.

[9]The simulation studies of which I am aware at present are slightly flawed in that they assume pointwise sampling at the sampling points, whereas experimental values will normally involve integration over a detector pixel size.

[10]The smallness of the margin of safety may be important in our planned yeast cell work, where we have significant missing data in both the central region and double cone (see section on Two Probable Experimental Regimes: A Synchrotron/Cryoprotection Technique for Supramolecular Specimens). We have shown through simulation studies that the algorithm can start with arbitrary magnitudes as well as phases in those regions and refine both successfully, but that trouble starts to develop as the amount of missing data approaches that which we may have to face in practice. With this in mind, we are considering an alternative algorithm, resembling a crystallographic least-squares model refinement but with values of Nyquist-fineness pixels (rather than atom positions) being the quantities refined in image space, and with observed finer-than-Nyquist diffraction magnitudes (unobserved magnitudes omitted) being the data to be fitted.

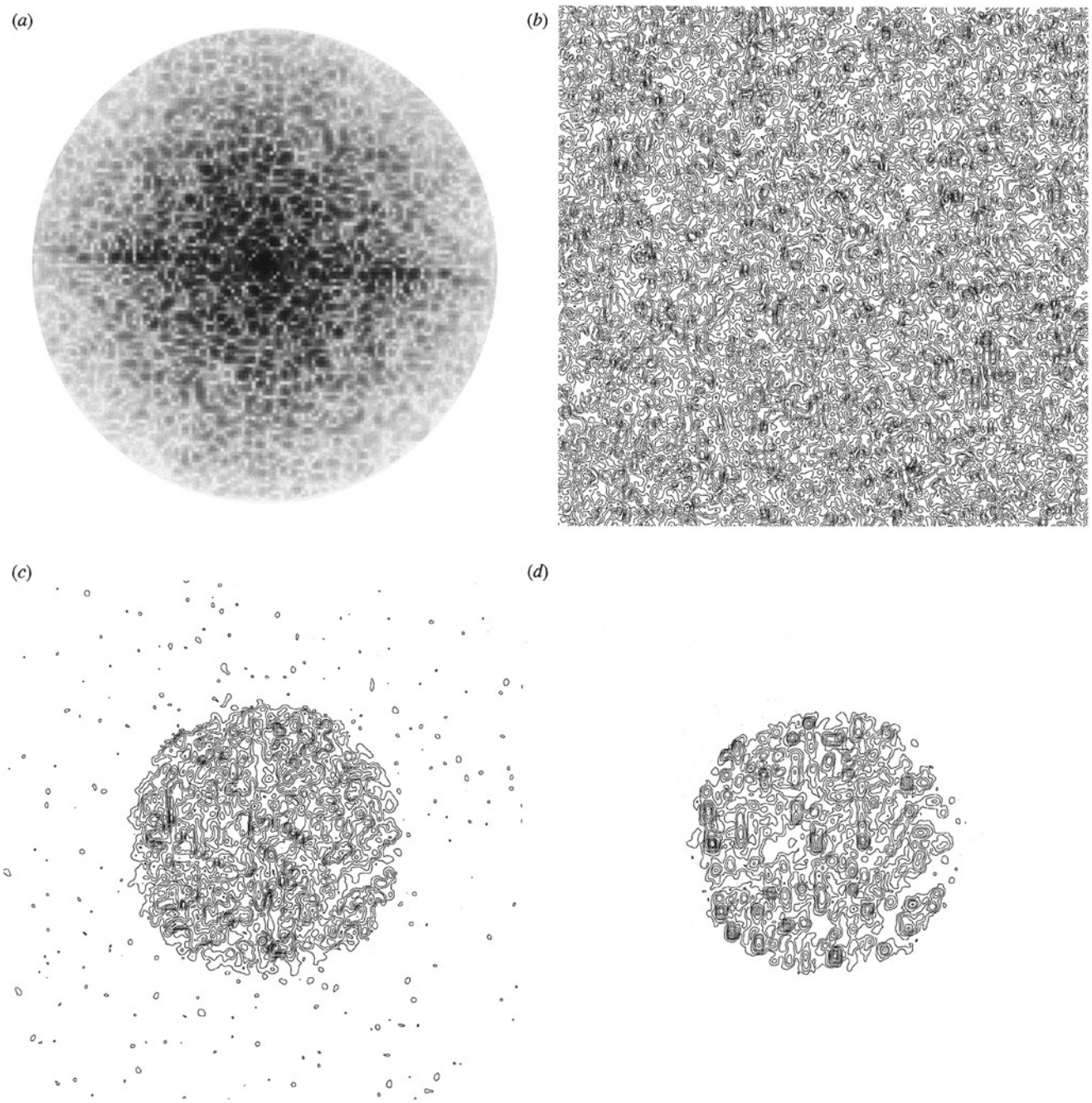

Fig. 2 An example of phasing by oversampling. (**a**) A set of 13,965 discrete magnitudes of a 2D computer-generated diffraction pattern. The magnitudes were oversampled by 2 times in each dimension. (**b**) The structure image corresponding to the initial random phasing. (**c**–**e**) The images after 50, 100, and 200 phasing cycles, the cycles primarily devoted to pushing the electron density outside the roughly known specimen envelope to zero. Note that after 50 cycles, the zeroing has advanced but is not complete. After 100 cycles, the contents of the envelope are taking form, but are not complete. After 200 cycles, the content has reached a final form. (**f**) The actual content on which (**a**) was based. Note that (**e**) has correctly found the structure. (**g**, **h**) Like (**e**, **f**) except that 17 % noise (peak-to-peak) was added to the data in (**a**) and 425 cycles were needed to complete the phasing (Taken from Ref. [19]).

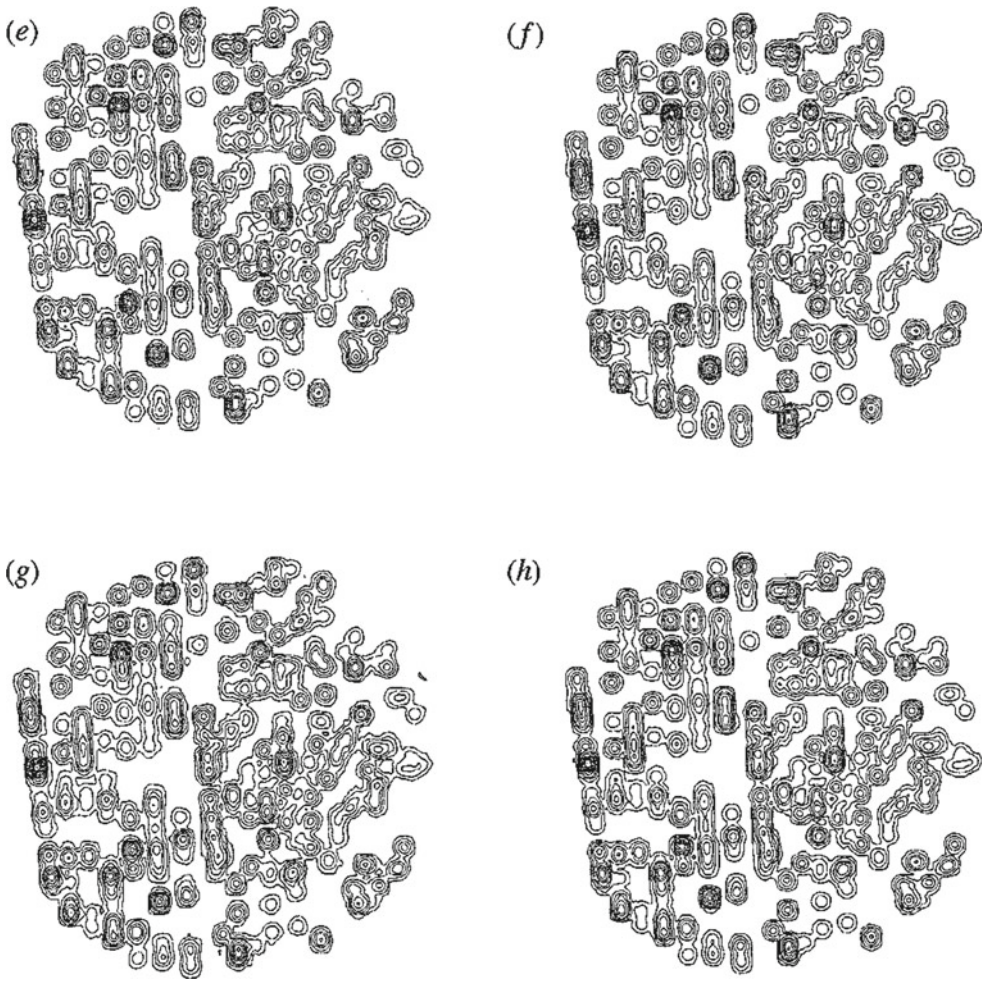

Fig. 2 (continued)

value in the study of crystal imperfections (Fig. 4) and possibly also in protein crystallography, where it is not infrequent for crystallization attempts to result in the production of numerous very small crystals.

Summary

The phase problem, which might have stunted the growth of crystallography, and held it to not much more than its 1929 dimensions, did not do so, and the subject has (through the work of many) been allowed to grow freely to wherever the ability to grow crystals may take it. One path to yet further growth is thus through more powerful crystal-growing techniques, and one version of this, which would be a highly attractive one, is mentioned as a future possibility in the last section of Part I. However, a second path also may exist and that is the use of new technology to extract copious diffraction data from arbitrary (including noncrystalline) specimens. This path, only now approaching its first major tests, is briefly set forth in Part II.

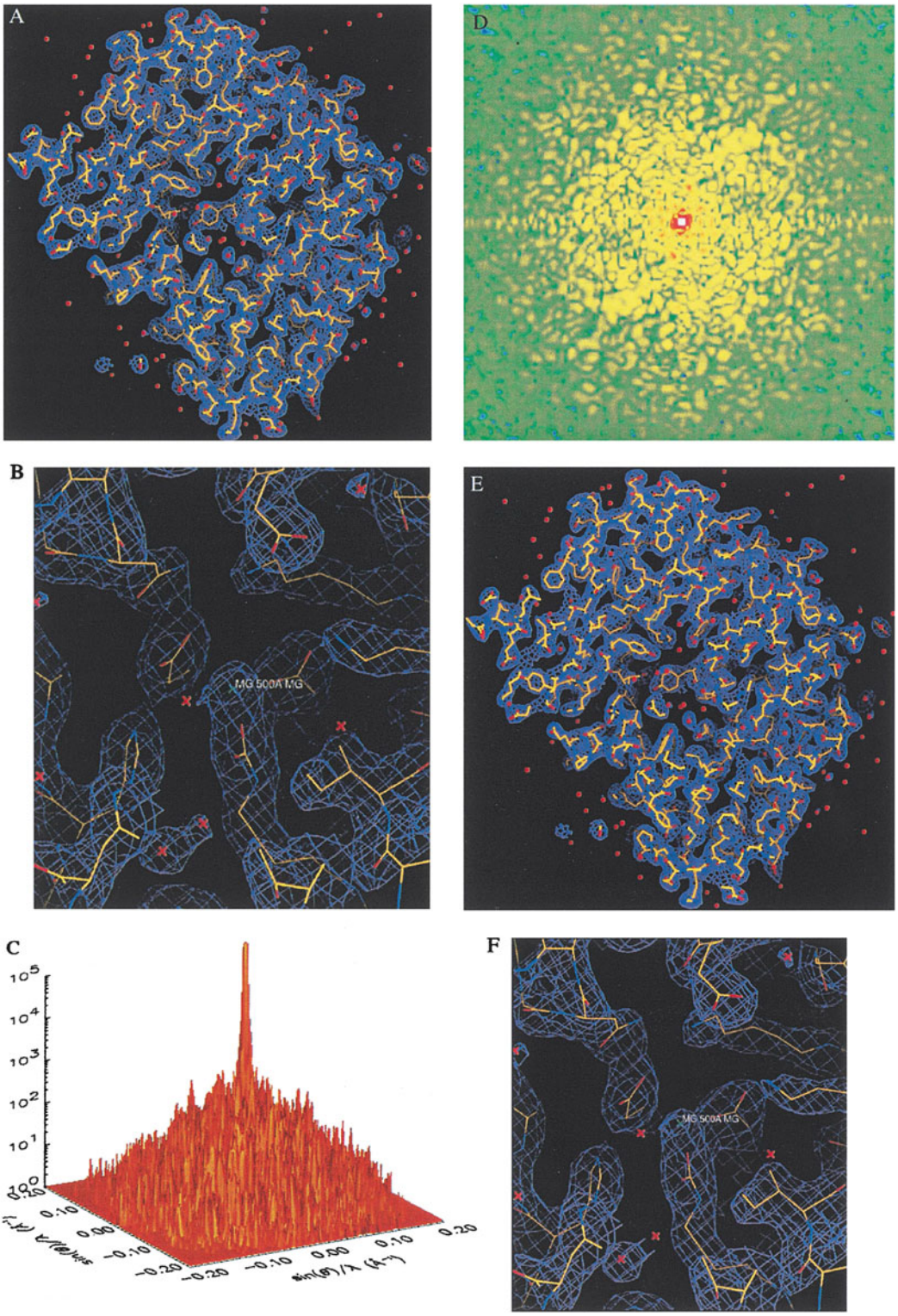

Fig. 3 Simulation of a possible future method of protein structure determination. (**c**) and (**d**) Simulation of one section of the oversampled 3D dataset of the protein rubisco, which could be obtained using 10^6 individual randomly oriented rubisco molecules successively illuminated (and destroyed) by 10^6 fs-scale 1.5 Å wavelength FEL X-ray pulses. The cumulative intensity of the pulses is sufficient to yield a 3D pattern to 2.5 Å resolution with reasonable Poisson noise statistics. The 10^6 exposures could reasonably be carried out in approximately 2.3 h. (**e**) and (**f**) show the molecule and its active site as found after oversampling phasing of the dataset, while (**a**) and (**b**) show the corresponding views of those structures as found in the original normal protein crystallographic investigation [27] (Taken from Ref. [23]).

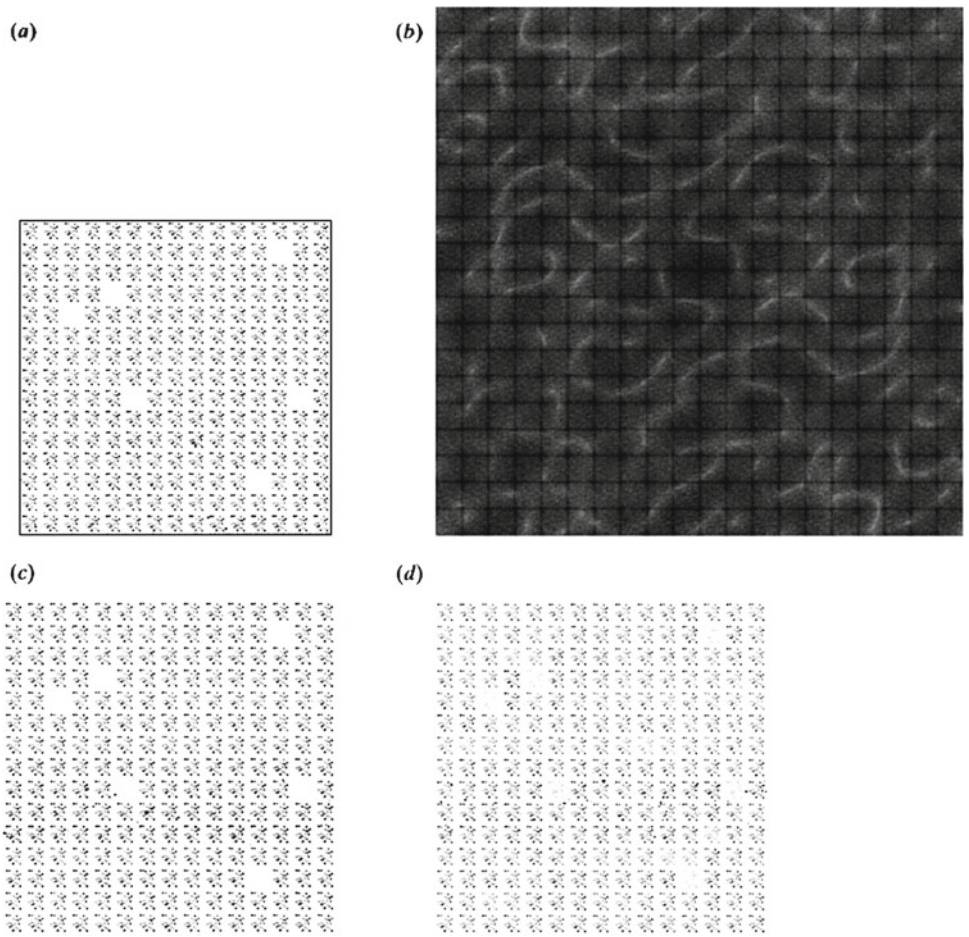

Fig. 4 Example of the possible value of finer sampling in the study of crystal imperfections. (**b**) A simulated strongly exposed diffraction pattern of a small 2D crystal. The pattern is continuous and permits oversampling. (**c**) The result of phasing and imaging. The crystal is found to be imperfect, with six missing unit cells. (**a**) The actual structure from which (**b**) was generated. (**d**) The result with random noise added to (**b**). The imaging is less clear (Taken from Ref. [22]).

Acknowledgments It is evident that the progress briefly recorded in this article is the work of very many people. I hope that I have given a brief but fair indication of that. At the same time I do not wish to close this article without acknowledging the special help given to me by certain individuals in the course of my lifetime. Gerard Bricogne of Cambridge has been, for many years, a companion and often guide in thinking about the phase problem. John Miao, graduate student at Stony Brook and now friend at Stanford, played a vital part in advancing the work reported in Part II of this article. Eaton Lattman of Johns Hopkins has often been my advisor in crystallographic matters, and was so in the writing of this article. Last, but also first, is Janos Kirz of Stony Brook, who more than any other made possible the work of Part II, with untiring support and accurate advice. Finally, my wife Anne Sayre was everything to me in the work, as well as the living, of life.

References

1. Robertson, J. M. *Rep. Progr. Phys.* **1937**, *4*, 332.
2. Sayre, D. *Acta Crystallogr.* **1952**, *5*, 843.
3. Harker, D.; Kasper, J. S. *Acta Crystallogr.* **1948**, *1*, 70; Kasper, J. S.; Lucht, C. M.; Harker, D. *Acta Crystallogr.* **1950**, *3*, 436.
4. Karle, J.; Hauptman, H. *Acta Crystallogr.* **1950**, *3*, 181.
5. Sayre, D. *Acta Crystallogr.* **1952**, *5*, 60.
6. Sayre, D. *Acta Crystallogr.* **1972**, *A28*, 210; *Acta Crystallogr.* **1974,** *A30*, 180.
7. Rothbauer, R. *Z. Kristallogr.* **2000**, *215*, 157.
8. Hauptman, H.; Karle, J. *ACA Monograph #3*; Edwards: Ann Arbor, **1953**.
9. (a) Bricogne, G. *Acta Crystallogr.* **1988**, *A44*, 517; (b) *Trans. ACA* **1994**, *30,* 163; (c) *Methods Enzymol.* **1997**, *276*, 361.
10. Weeks, C. M.; DeTitta, G. T.; Hauptman, H. A.; Thuman, P.; Miller, R. *Acta Crystallogr.* **1994**, *A50*, 210.
11. (a) Szoke, A. *Acta Crystallogr.* **1993**, *A49*, 853.; Somoza, J. R.; Szoke, H.; Goodman, D. M.; Beran, P.; Truckses, D.; Kim, S.-H.; Szoke, A. *Acta Crystallogr.* **1995**, *A51*, 691. (b) Szoke, A. *Acta Crystallogr.* **2001**, *A57*, 587.
12. Spence, J. C. H. *Ultramicroscopy*, in press.
13. Burns, M. M.; Fournier, J.-M.; Golovchenko, J. A. *Science* **1990**, *249*, 749.
14. Larsen, J. J.; Hald, K.; Bjerre, N.; Stapelfeldt, H.; Seideman, T. *Phys. Rev. Lett.* **2000**, *85*, 2470.

15. Sayre, D. In *Imaging Processes and Coherence in Physics*; Schlenker, M.; Fink, M.; Goedgebuer, J. P.; Malgrange, C.; Vienot, J. Ch.; Wade, R. H., Eds.; Springer-Verlag: Heidelberg, **1980**; p. 229.
16. Yun, W. B.; Kirz, J.; Sayre, D. *Acta Crystallogr.* **1987**, *A43*, 131.
17. Sayre, D. In *Direct Methods of Solving Crystal Structures*; Schenk, H., Ed.; Plenum Press: New York, **1991**; p. 353.
18. Sayre, D.; Chapman, H. N. *Acta Crystallogr.* **1995**, *A51*, 237.
19. Sayre, D.; Chapman, H. N.; Miao, J. *Acta Crystallogr.* **1998**, *A54*, 232.
20. Miao, J.; Sayre, D.; Chapman, H. N. J. *Opt. Soc. Amer.* **1998**, *A15*, 1662.
21. Miao, J.; Charalambous, P.; Kirz, J.; Sayre, D. *Nature* (*London*) **1999**, *400*, 342.
22. Miao, J.; Sayre, D. *Acta Crystallogr.* **2000**, *A56*, 596.
23. Miao, J.; Hodgson, K. O.; Sayre, D. *Proc. Natl. Acad. Sci.* **2001**, *98*, 6641.
24. Neutze, R.; Wouts, R.; v. d. Spoel, D.; Weckert, E.; Hajdu, J. *Nature* (*London*) **2000**, *406*, 752.
25. Arthur, J.; Bane, K.; Bharadwaj, V.; Bowden, G.; et al. *Linac Coherent Light Source (LCLS) Design Study Report*; Stanford Linear Accelerator Center, **1998.**
26. Fienup, J. R. *Appl. Opt.* **1982**, *21*, 2758.
27. Lundqvist, T.; Schneider, G. *Biochemistry* **1991**, *30*, 904.