STRUCTURE DETERMINATION OF ASYMMETRIC MEMBRANE PROFILES USING AN ITERATIVE FOURIER METHOD

ROBERT M. STROUD AND DAVID A. AGARD, *Department of Biochemistry and Biophysics, University of California, San Francisco, California 94143 U.S.A.*

ABSTRACT An iterative Fourier method is applied to solving and refining the electron density profile projected onto the line perpendicular to a membrane surface. Solutions to the continuous X-ray scattering pattern derived from swelling of multilayer systems or from membrane dispersions can be obtained by this technique. The method deals directly with the observed structure factors and does not rely on deconvolution of the Patterson function. We used this method previously to derive the electron density profile for acetylcholine receptor membranes (Ross et al., 1977). The present paper is an analysis of the theoretical basis for the procedure. In addition, the technique is tested on artificially generated continuous-scattering data, on the data for frog sciatic nerve myelin derived from swelling experiments by Worthington and McIntosh (1974), and on the data for purple membrane (Blaurock and Stoeckenius, 1971). Although the method applies to asymmetric membranes, the special case of centrosymmetric profiles is also shown to be solvable by the same technique. The limitations of the method and the boundary conditions that limit the degeneracy of the solution are analyzed.

INTRODUCTION

The principles of X-ray crystallography were first developed to determine structure in repeating systems. Three-dimensional order in a crystalline specimen permits a great enhancement of the X-ray scattering expected from a single molecule or dispersion of molecules, although the scattering function is consequently sampled only at the positions determined by Bragg's law. The components of biological membranes are inherently closely packed in two dimensions, although crystalline ordering of membrane-bound protein, even in domains containing several hundred molecules, is rare. Most specialized biological membranes are inherently asymmetric structures, i.e. the distribution of protein (or other material) in a direction perpendicular to the plane of a lipid bilayer is asymmetric.

X-ray diffraction has played a major role in establishing the nature of phospholipid bilayers (Levine and Wilkins, 1971). Although phospholipid molecules in the leaflet were shown to possess no regular two-dimensional crystalline order at temperatures above the phase transition, ordering and structural repetition could be obtained in the third dimension, perpendicular to the membrane surface, by close packing of the equivalent and symmetric layers. As a result, the diffraction pattern defined along a reciprocal lattice direction **s**, where **s** is normal to the membrane surface, is contained in discrete Bragg reflections.

Robert M. Stroud is an Alfred P. Sloan Foundation Fellow. David A. Agard is a National Science Foundation Predoctoral Fellow at the California Institute of Technology, Pasadena, Calif. 91125.

Since a small number of biological membranes, generated by folding over a cell envelope such as the myelin sheath, or rod outer segment, contain a preexistent center of symmetry, the theories of diffraction from centrosymmetric structures could be applied. Solution of the phase problem, for the signs of successive reflections in such cases, is often by deconvolution of the Patterson function or by trial, since experimental methods, such as isomorphic replacement of one component of the sample by another, have not generally been successful for membranes. Often the diffraction data can only be phased to relatively low resolution, as the correct phase choice relies to some extent on agreement with a reasonable structural model. There is no precedent for use of the finer modulations of structure (which depend on the higher resolution data) as a means of choosing between the possible signs associated with higher order terms.

In some centrosymmetric membrane structures, such as myelin sheath, the membrane multilayers may be reversibly swollen (Robertson, 1958; Moody, 1963; Blaurock, 1971; Caspar and Kirschner, 1971; Worthington and McIntosh, 1974). If the structural unit remains intact during swelling, i.e., if the function $\rho(\mathbf{r})$ (referred to as the membrane profile) does not change within the limited thickness of the leaflets and only the repeat or multilayer spacing d is assumed to vary, then the observed transform is sampled at different positions defined by $\mathbf{s}_h = \lambda h/d$, and the continuous transform can be mapped. The positions of zeros in the transform, i.e., where $|F(\mathbf{s})|^2 = I(\mathbf{s}) = 0$, provide a powerful constraint useful in defining similarities of signs for adjacent reflections that lie between the zeros in the continuous transform. However, the function $F(\mathbf{s})$ need not change sign at one of these nodes. For frog sciatic myelin (*Rana pipiens*), the mapping of the continuous transform by Worthington and McIntosh (1974) reduced the number of possible sign combinations for diffraction orders $h = 1$–12, from 2,048 to 64.

Difficult as it is to define the phases for a centrosymmetric structure, determination of general phases for asymmetric membranes is much more difficult, and this is a primary concern here. Generally, when asymmetric membranes are closely aligned one against the other, there will be no true repetition in the direction perpendicular to the sheets. The statistical distribution of membranes with alternate surfaces up or down gives rise to large statistical fluctuations in the true repeat distances in the specimen. As a result, there is a pillar of continuous scatter perpendicular to the membrane surface due to the product of the true transform of one membrane sheet, $F(\mathbf{s})$, and the transform of a statistical stacking function. The continuous scatter may show slight evidence for the minimum possible repeat distance in the stacking direction arising from the chance stacking of several layers in the same direction. Where sheets are almost symmetric in terms of the distribution of density through the membrane, the pattern may show a reasonably well-defined minimum repeat distance. This is the case for the asymmetric purple membrane from *Halobacterium halobium* (Blaurock and Stoeckenius, 1971; Henderson, 1975). But in all cases, the pattern obtained from closely packed asymmetric membrane systems is made only more complex by the close packing of successive leaflets, as it is the product of the continuous transform from one layer $F(\mathbf{s})$ and a sampling function, which may not depend on random statistics of stacking for obvious physical reasons. Since the second function is extremely difficult to derive (Burge and Draper, 1967), the first is equally difficult to extract.

A more easily obtained scattering pattern is the one from membrane dispersions or

moderately close-packed membrane sheets, where the scattering pattern $I_{obs}(\mathbf{s}) \sin^2\theta$ more nearly represents the square of the Fourier transform of a single membrane sheet $F(\mathbf{s})$ (Wilkins et al., 1971). It is also possible to orient the membrane sheets centrifugally while maintaining only moderate close packing. This allows for separation of the normal and in-plane diffraction, which is important if there is significant in-plane lattice structure.

The information present in continuous X-ray scattering patterns of these types is in one sense greater, since the transform is a continuous function often observed to moderate resolution. In many cases the transform can be shown to be unsampled by the statistical stacking disorder function (disorder of the second kind of Hosemann and Bagchi, 1962) by comparison with the diffraction pattern from membrane vesicles, for example. The continuity of the observed transform places restrictions on the solution. The main difficulty in interpreting these patterns is again determination of the phases $\phi(\mathbf{s})$ of the observed continuous transform $|F(\mathbf{s})|$.

In the case of sampled diffraction patterns, as in crystallography, Fourier refinement methods generally require a good approximation to the real solution before refinement is at all meaningful. The continuity of the observed Fourier transform partially removes this condition. One further piece of information, namely knowledge that the structure is of finite rather than infinite thickness, is sufficient to restrict the number of possible solutions to a very small number, and often to just one.

The method of refinement discussed in this paper imposes the criterion of limited physical extent (constrained thickness) during successive cycles of Fourier refinement of an electron density model to the observed continuous transform $|F_0(\mathbf{s})|$. The starting trial model used to initiate the procedure is demonstrably unimportant when a unique solution exists, since the method always leads to convergence on a solution.

Likewise, the constraining thickness is theoretically unimportant as regards obtaining a solution, providing it is larger than the actual thickness of the membrane structure. No converged solution can be obtained if the thickness used for refinement is too small. Practically, a correct choice for the constraining thickness assists in achieving more rapid convergence to a solution, and in minimizing problems associated with errors in the data, including the fact that data are necessarily of limited resolution. Resolution is limited theoretically by the wavelength of the radiation to $\lambda/2$, and practically limited by experimental conditions, or by signal-to-noise ratio.

This application of Fourier refinement can determine and refine possible solutions for the electron density function, although absolute certainty that all solutions have been obtained would require testing of a large number of possible alternatives. In practice, the number of real alternatives is limited to those consistent with the maximum and minimum electron density allowed by the chemical constituents of the system, i.e., to chemically reasonable alternatives that can in principle all be tested. The more difficult question is whether a determined solution is necessarily the correct one. If more than one solution is obtainable, identification of the correct one requires experimental evidence of a different type, such as isomorphic replacement. Thus, in reporting membrane profiles, many solutions should be sought and reported where obtainable.

To test the efficacy of the iterative refinement method, several transforms were generated from artificial model structures. The character and dimensions of the models were of the

kind that might be expected for genuine membranes, so that discussions of resolution limits, etc., may be compared with real cases. The iterative method was then applied to the generated transform data for hypothetical membrane-like structures to assess the power of boundary conditions and limitations of the method. The method was also applied to solution of the density profiles from the continuous purple membrane data of Blaurock and Stoeckenius (1971) and from the myelin data of Worthington and McIntosh (1974).[1]

The primary motivation for evaluation of the method is to apply the technique to the location of biochemically characterized components in specialized membranes.

THEORY OF THE METHOD FOR ASYMMETRIC AND SYMMETRIC STRUCTURES

Let the electron density perpendicular to the membrane be $\rho'(\mathbf{r})$ and let ρ_s be the constant electron density of the solvent. For simplicity in the ensuing discussion, it is convenient to define a solvent relative profile $\rho_0(\mathbf{r}) = \rho'(\mathbf{r}) - \rho_s$, since the transform of $\rho_0(\mathbf{r})$ is observed in a diffraction experiment. The transform of the constant density ρ_s is contained in the direct beam. For a real asymmetric structure, the choice of origin is arbitrary. Thus if the membrane is of width w, then $\rho_0(\mathbf{r})$ may be nonzero only between the limits $\mathbf{r} = a$ and $\mathbf{r} = b$, where $w = b - a$. The function $\rho_0(\mathbf{r})$ is then said to be constrained to lie within these limits. Although $\rho_0(\mathbf{r})$ will generally be largely positive, it may also be negative where, as in the center of a bilayer for example, $\rho'(\mathbf{r})$ is less than the electron density of the solvent (Levine and Wilkins, 1971).

The Fourier transform of $\rho_0(\mathbf{r})$ will be $F_0(\mathbf{s})$, where the reciprocal space vector $\mathbf{s}$ is parallel to $\mathbf{r}$. Then the observed diffraction pattern, if it can be shown to be an unsampled continuous pattern (or is the continuous data extracted from swelling experiments), will be, after suitable corrections for the experimental arrangement are applied, $I_0(\mathbf{s}) = |F_0(\mathbf{s})|^2$. Having derived the function $|F_0(\mathbf{s})|$ from the diffracted intensity, the problem is to determine solutions for $\rho(\mathbf{r})$ that obey the normal Fourier inversion relationship:

$$\rho(\mathbf{r}) = \int_{-s_1}^{s_1} F_0(\mathbf{s}) \exp(-2\pi i \mathbf{r}\cdot\mathbf{s}) d\mathbf{s}, \tag{1}$$

where

$$F_0(\mathbf{s}) = \int_{-\infty}^{\infty} \rho(\mathbf{r}) \exp(2\pi i \mathbf{r}\cdot\mathbf{s}) d\mathbf{r}, \tag{2}$$

and where s_1 defines the highest resolution ($d_{\min} = \lambda/s_1$) to which the data are measured. Data obscured by the beam stop can be interpolated by using either the sampling theorem or by curve fitting. Since the observed transform is continuous (unlike the crystallographic case), Eqs. 1 and 2 are linear integrals of a continuous function.

Any random phase function $\phi(\mathbf{s})$ associated with $|F_0(\mathbf{s})|$, such that $F_0(\mathbf{s}) = |F_0(\mathbf{s})| \exp(2\pi i \phi(\mathbf{s}))$, when transformed according to Eq. 1, will yield a function $\rho(\mathbf{r})$ that when

[1] It should be clearly noted that the objectives in these cases were to test the method against solutions obtained by these authors by using their particular data sets. The results are therefore entirely dependent on the data set, and the tests were set up this way so that solutions obtained by this method could be directly compared with those obtained by these authors.

transformed by Eq. 2 will necessarily generate the same $F(\mathbf{s})$. This is true only so long as the integration in Eq. 2 is carried out over all space ($\mathbf{r} = -\infty$ to $\mathbf{r} = +\infty$). The resulting $\rho(\mathbf{r})$ will be a function of infinite extent; i.e. $\rho(\mathbf{r})$ may be nonzero for all $\mathbf{r}$ ($-\infty < \mathbf{r} < \infty$). The solution $\rho(\mathbf{r})$ so obtained clearly has no physical meaning. Since the phases $\phi(\mathbf{s})$ were chosen at random, there are an infinite number of such solutions. Without additional information, there is no basis for determining the phases of $F_0(\mathbf{s})$ from the observed values of $|F_0(\mathbf{s})|$.

Since $\rho(\mathbf{r})$ must correspond to a real object, it cannot be of infinite extent or thickness. This alone restricts the number of solutions, no matter how they are obtained, to a small number, all of which are necessarily of the correct actual membrane thickness, without any assumption of what that thickness may be. This can be seen in the following way.

The transform of $F_0(\mathbf{s}) \cdot F_0^*(\mathbf{s})$ is the Patterson function

$$P(\mathbf{r}) = \int_{-s_1}^{s_1} |F_0(\mathbf{s})|^2 \exp(2\pi i \mathbf{r}\cdot\mathbf{s}) d\mathbf{s}, \tag{3}$$

and will always be equal to the convolution $\rho(\mathbf{r}) \widehat{\rho(-\mathbf{r})}$ (see Lipson and Taylor, 1958, for example). This convolution is termed the autocorrelation function,

$$Q(\mathbf{r}) = \int_{-\infty}^{+\infty} \rho(\mathbf{r}')\rho(\mathbf{r}' + \mathbf{r}) d\mathbf{r}'. \tag{4}$$

Thus $P(\mathbf{r}) \equiv Q(\mathbf{r})$, whatever the phase set that is applied to $|F_0(\mathbf{s})|$ in Eq. 1, providing the integrals in Eqs. 2 and 4 are carried out from $\mathbf{r} = -\infty$ to $\mathbf{r} = +\infty$.

Now consider the autocorrelation function computed from a $\rho(\mathbf{r})$ of finite extent, i.e. $\rho(\mathbf{r}) = 0$ for $\mathbf{r} < a$ or $\mathbf{r} > b$. The autocorrelation function

$$Q(\mathbf{r}) = \int_{-\infty}^{+\infty} \rho(\mathbf{r}')\rho(\mathbf{r} + \mathbf{r}') d\mathbf{r}' \text{ will be zero for all values of } |\mathbf{r}| > w,$$

(where $w = b - a$) since at least one of the terms $\rho(\mathbf{r}')$, $\rho(\mathbf{r} + \mathbf{r}')$ must be zero when $|\mathbf{r}| > w$. Conversely, if $Q(\mathbf{r}) = 0$ for $|\mathbf{r}| > w$, then any function $\rho(\mathbf{r})$ that is a solution to Eq. 4, can only be nonzero within an interval $\Delta\mathbf{r} = w$ provided that $\rho(\mathbf{r})$ is of finite extent.

If the actual (true) thickness of the membrane leaflet is w, the autocorrelation function for the structure $Q_0(\mathbf{r})$ [$\equiv P_0(\mathbf{r})$] will be zero for all $|\mathbf{r}| > w$. $P_0(\mathbf{r})$ is determined by the intensity distribution $|F_0(\mathbf{s})|^2$ alone, and not by the phases assigned to $|F_0(\mathbf{s})|$. Any function $\rho(\mathbf{r})$ of thickness $t > w$ will necessarily produce an autocorrelation function $Q(\mathbf{r})$ that has nonzero components[2] for $w < |\mathbf{r}| < t$. Therefore $Q(\mathbf{r}) \neq P_0(\mathbf{r})$, at least for $w < |\mathbf{r}| < t$, and so $\rho(\mathbf{r})$ cannot be a function derived from the observed moduli $|F_0(\mathbf{s})|$ (Eq. 1).

Thus no function $\rho(\mathbf{r})$ of finite width $t > w$ (and by similar arguments $t < w$) can be a solution to Eqs. 1 and 2. It follows that imposition of any constraining thickness, which alters Eq. 2 to

$$F_0(\mathbf{s}) = \int_a^b \rho(\mathbf{r}) \exp(2\pi i \mathbf{r}\cdot\mathbf{s}) d\mathbf{r}, \tag{5}$$

[2]In general $Q(\mathbf{r})$ will similarly be nonzero for all $|\mathbf{r}| < w$. However, there may be particular values of $\mathbf{r}$ for which all terms in Eq. 4 cancel to give a zero in $Q(\mathbf{r})$.

where $b - a > w$, only allows solutions, $\rho(\mathbf{r})$, of thickness w. In the Fourier refinement method, solutions that obey Eqs. 1 and 5 are determined in an iterative process.

By analogy with usual Fourier refinement procedures, an initial structure $\rho_i(\mathbf{r})$ is transformed according to Eq. 5, i.e., between some limits $\mathbf{r} = a$ to $\mathbf{r} = b$ to generate an initial calculated transform $F_c(\mathbf{s})$. The phases $\phi_c(\mathbf{s})$ derived from this model are then associated with the "observed" $|F_0(\mathbf{s})|$ and transformed back by using Eq. 1 to obtain a new function $\rho_i'(\mathbf{r})$. The entire process is cycled until convergence is reached (Fig. 1). All refined solutions, $\rho(\mathbf{r})$, will necessarily be of the same thickness.

If multiple solutions, $\rho(\mathbf{r})$, exist that are compatible with a thickness of w, then these will form a set of possible solutions (strictly a homometric set of solutions) that cannot be eliminated without further information by any method of treating $|F(\mathbf{s})|$. The choice of starting model $\rho_i(\mathbf{r})$ in these instances will tend to select for the solution closest to the starting model. Consequently, this approach can be used to test hypothetical structures. Obviously, if only one $\rho(\mathbf{r})$ of finite thickness exists, this procedure will converge upon that structure. In practice it seems that when multiple solutions do exist, they are often qualitatively quite similar to one another.

The boundary limits $\mathbf{r} = a$, $\mathbf{r} = b$ are unimportant as regards achieving a converged solution. The choice of a constraining thickness w that is close to the correct value for converged solutions greatly speeds up convergence.

The constrained thickness may be estimated in several ways and, therefore, is not difficult to determine accurately enough to be effective. It may be estimated: from electron microscopy of isolated protein molecules, or of thin sections cut perpendicular to close-packed membrane sheets; from the minimum close-packing distance found in X-ray scattering from densely packed specimens; or from the Patterson function, which identifies (within resolution limits) the largest vectors in the structure.

So far in our considerations we have assumed that $|F_0(\mathbf{s})|$ contains no errors. In any real case the values of $|F_0(\mathbf{s})|$ will be extracted with care to try to achieve this condition (see, for example, Ross et al., 1977). If there is independent evidence for the thickness of the leaflet

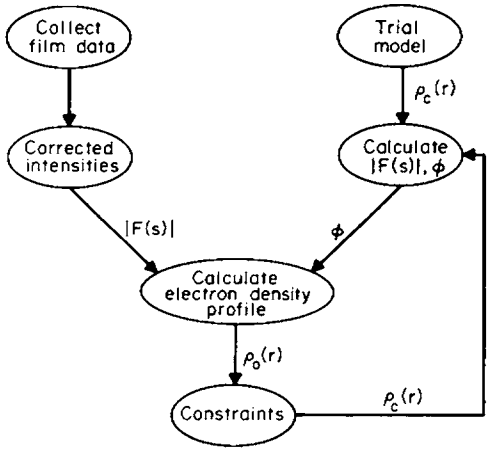

FIGURE 1 Schematic of the refinement procedure.

(of the kind mentioned above), then imposition of the boundary thickness can minimize errors in the profile introduced by errors in $|F_0(\mathbf{s})|$. If such errors in the data exist, they will be assayed by the final agreement between $|F_0(\mathbf{s})|$ and $|F_c(\mathbf{s})|$ for the constrained structure.

Systematic errors in the refined profile arise when the data are cut off at some particular resolution, s_1, such that $|F_0(\mathbf{s})|$ is nonzero for $\mathbf{s} > s_1$. This is to be expected since the "true" resolution limited structure (computed with true phases and limited in resolution to s_1) will contain Fourier series termination ripples of frequency equal to s_1. The presence of termination ripples is inconsistent with the refinement procedure used here [$\rho(\mathbf{r})$ is constrained to be zero outside the boundary thickness]. Therefore the phases of the terms $F(\mathbf{s})$ with $\mathbf{s}$ closest to s_1 will be modified, and so changed away from their true phase to best accommodate the termination ripples. In any real case, data to the highest observable resolution should be included, such that $|F_0(\mathbf{s})| \to 0$ at the resolution cutoff limit. The imposition of an artificial temperature factor $\exp(-B\mathbf{s}^2)$ is useful when this condition is not met.

Criterion for Acceptance of a Solution

The criterion for accepting a refined solution, $\rho(\mathbf{r})$, is that the amplitude of the transform $|F_c(\mathbf{s})|$ computed in Eq. 5 should agree with the observed transform $|F_0(\mathbf{s})|$. The degree to which $|F_c(\mathbf{s})|$ approximates $|F_0(\mathbf{s})|$ characterizes the extent of convergence, and indicates the degree to which the diffraction data are consistent with a structure of width $w < b - a$ in thickness. The agreement index,

$$R = \frac{\int_0^{s_1} ||F_c(\mathbf{s})| - |F_0(\mathbf{s})|| \, d\mathbf{s}}{\int_0^{s_1} |F_0(\mathbf{s})| \, d\mathbf{s}} \tag{6}$$

is typically 0.1%–1% for a solution to artificially created and purple membrane data, or about 1%–5% for measured $|F_0(\mathbf{s})|$ data from acetylcholine receptor membranes, for example, (Ross et al., 1977).

Since for any solution, $\rho(\mathbf{r})$, there will be three other trivial solutions, $-\rho(\mathbf{r})$, $\rho(-\mathbf{r})$, and $-\rho(-\mathbf{r})$, only one solution of the set for which $\int \rho(\mathbf{r}) \, d\mathbf{r} > 0$ will be considered.

TESTS OF THE METHOD

An Asymmetric Structure

A supposed (initial) electron density function shown in Fig. 2 A was used to generate the function $|F(\mathbf{s})|$, which was then regarded as an observed transform. The back transform of $F(\mathbf{s})$ at 10 Å resolution, the correct resolution-limited structure, is also shown in Figure 2 A. Attempts were then made to derive the electron density function $\rho(\mathbf{r})$ starting only from the moduli $|F(\mathbf{s})|$ by application of the refinement method. Numerous different trial models were used to initiate the procedure by using data to 10 Å resolution. These ranged from a single delta function, to several step functions, to various continuous functions. In each case convergence on the same final structure was achieved (Fig. 2 C). The boundary thickness chosen varied from 210 to 250 Å.

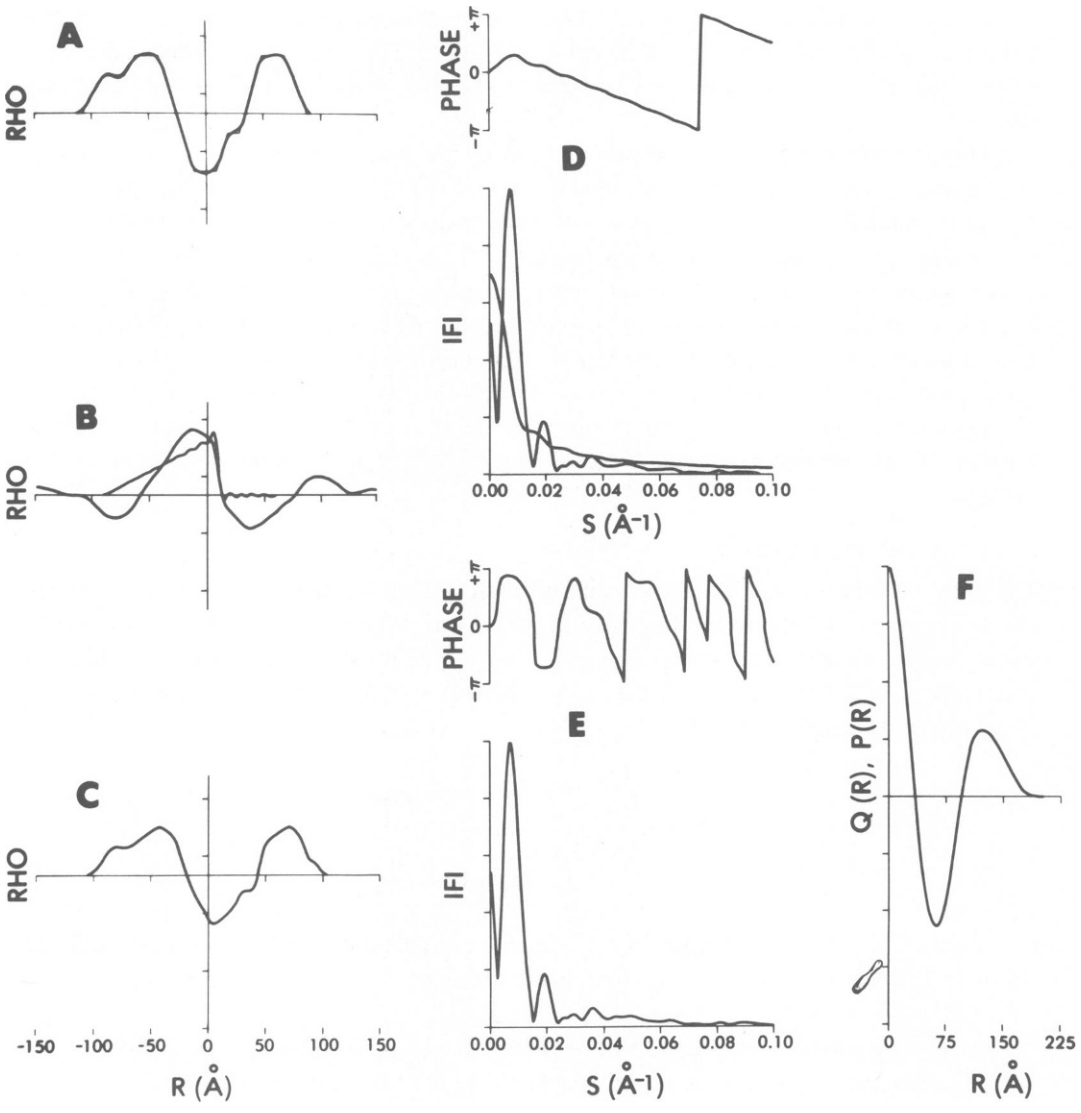

FIGURE 2 Results of refinement of an asymmetric "membrane" profile using synthetic data to 10 Å. The model used to generate the observed structure factors is shown in A. Using a ramp function for the trial structure results in the profiles depicted in B and C (1 and 30 cycles of refinement, respectively). Corresponding observed and calculated Fourier amplitudes and calculated phases are shown in D and E. A typical plot of both $Q(\mathbf{r})$ and $P(\mathbf{r})$ is shown in F. R_F and R_P factors in all refined cases are less than 1.5%.

When the structure was refined at 20 Å resolution, series termination ripples in $\rho(\mathbf{r})$ at each successive stage were forced into the structure by the refinement procedure (which eliminates ripples outside the structure) (Fig. 3 B). Later addition of higher resolution data (20–10 Å) did not immediately lead to further refinement of the 20 Å data, since subsequent refinement first tends to minimize ripples outside the boundary that at this stage are produced only by the additional data (Fig. 3 C). In this situation the additional data from

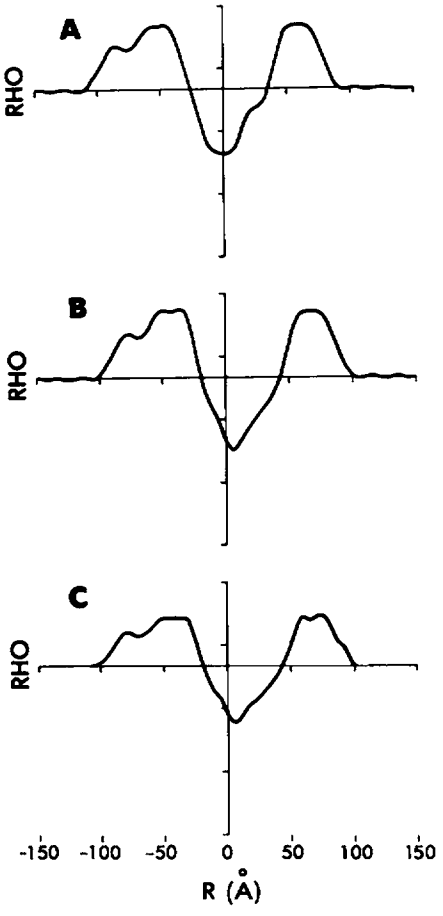

FIGURE 3

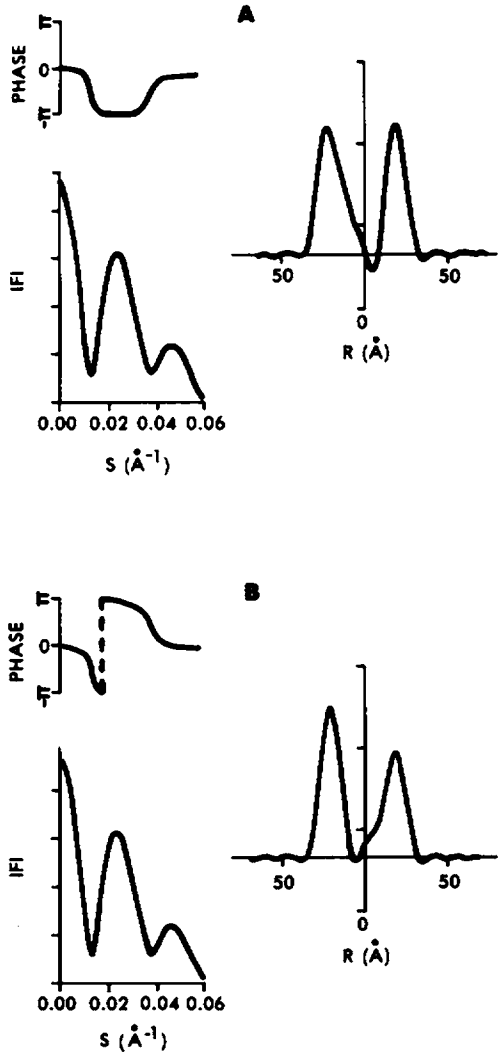

FIGURE 4

FIGURE 3 Refinement of the same profile as in Fig. 2 but at lower resolution. The true structure is depicted in A. The converged solution using 20 Å data after 20 refinement cycles appears in B. Although similar to the profile obtained with 10 Å data (Fig. 2 C), there are distinct differences resulting from forcing the series termination ripples to be accommodated within the structure. These ripples are not immediately removed by a later inclusion of the higher resolution data, as shown in C.

FIGURE 4 The two solutions (A, B) to the purple membrane continuous scattering data of Blaurock and King (obtained after 30 cycles of Fourier refinement). The left side of the figure shows the derived phase functions and the equivalence of $|F_0(s)|$ and $|F_c(s)|$ (shown superimposed). The R-factor in both cases was below 0.15%. On the right are the corresponding electron density profiles, showing the asymmetry in the protein distribution across the membrane. Many different starting models were used; however, only these two solutions could be obtained.

20 to 10 Å resolution are initially being treated as an independent data set. Only as successive changes in structure begin to affect the phases of the lower resolution data does the pressure to refine the lower resolution phases develop.

When the starting trial model $\rho_i(\mathbf{r})$ is centrosymmetric, convergence to the correct structure, which is asymmetric, requires that the centrosymmetric phase relationships (which are signs if the origin is chosen at the symmetric center) be broken. If the refined model is constrained to be concentric with the centrosymmetric starting model, it is impossible to achieve even qualitative agreement with the observed data. Convergence on the final solution is possible when the constraint to symmetry is removed by imposing a boundary size whose limits are not concentric with the starting model. The asymmetric truncation of series termination ripples allows a degree of freedom for the phases and so allows refinement to continue.

Several different initial structures were used to generate observed values of $|F(\mathbf{s})|$. The experience with the case described above was shared in other similar cases. As the complexity of the profile increased to the point where variations in $\rho(\mathbf{r})$ were of the same order as the resolution cutoff limit applied to the data, multiple solutions were obtained. These solutions were similar with differences accountable as alternative ways of phasing the lower intensity data at higher resolution while maintaining a width $\simeq w$. This is a consequence of the errors introduced by using data of limited resolution (see Theory section), and the lower pressure to phase weak-intensity data, whose contribution to $\rho(\mathbf{r})$ is least. Since the lowest intensities are found in general at highest resolution, phase errors will correspondingly be greater at higher resolution.

As another test, data from suspensions of purple membrane from *Halobacterium halobium*, kindly provided by Allen Blaurock (California Institute of Technology), were refined by using this method. Various starting models were used and only two different solutions could be obtained (Fig. 4). These solutions are identical to the two reported by Blaurock and King, 1977. The two solutions correspond to different senses of phase change $(d\phi/ds)$ between maxima in the observed transform $|F_0(\mathbf{s})|$.

Centrosymmetric Structures

A more difficult question is whether an initially centrosymmetric model $\rho(\mathbf{r})$ used to generate a transform or a centrosymmetric membrane structure will also be determined correctly, since the final phase angles must accommodate discontinuities between regions of opposite sign in the transform $F(\mathbf{s})$. This was tested first by calculation of observed transforms from symmetric models (see Fig. 5). Refinement against the data shown in Fig. 5 always led to convergence on the correct structure (Fig. 5 D). The phases changed most rapidly at the positions of sign change in the observed transform (where the computed $|F_0(\mathbf{s})|$ are smallest) and accommodated a phase change of 180° within a small range of $\Delta\mathbf{s}$.

In a second test of the method to determine centrosymmetric structures, the continuous transform for frog sciatic nerve myelin mapped from a series of swelling experiments by Worthington and McIntosh (1974) was used. The data were extracted from their Fig. 9 and digitized at intervals of 0.001 $Å^{-1}$ in reciprocal spacing to a resolution of $\mathbf{s} = 1/14$ $Å^{-1}$. Different data sets reported elsewhere, or different interpretations of what constitutes the continuous transform of myelin (considered extensively by Blaurock, 1976), would clearly give rise to different electron density profiles. Our objective was to see how closely the

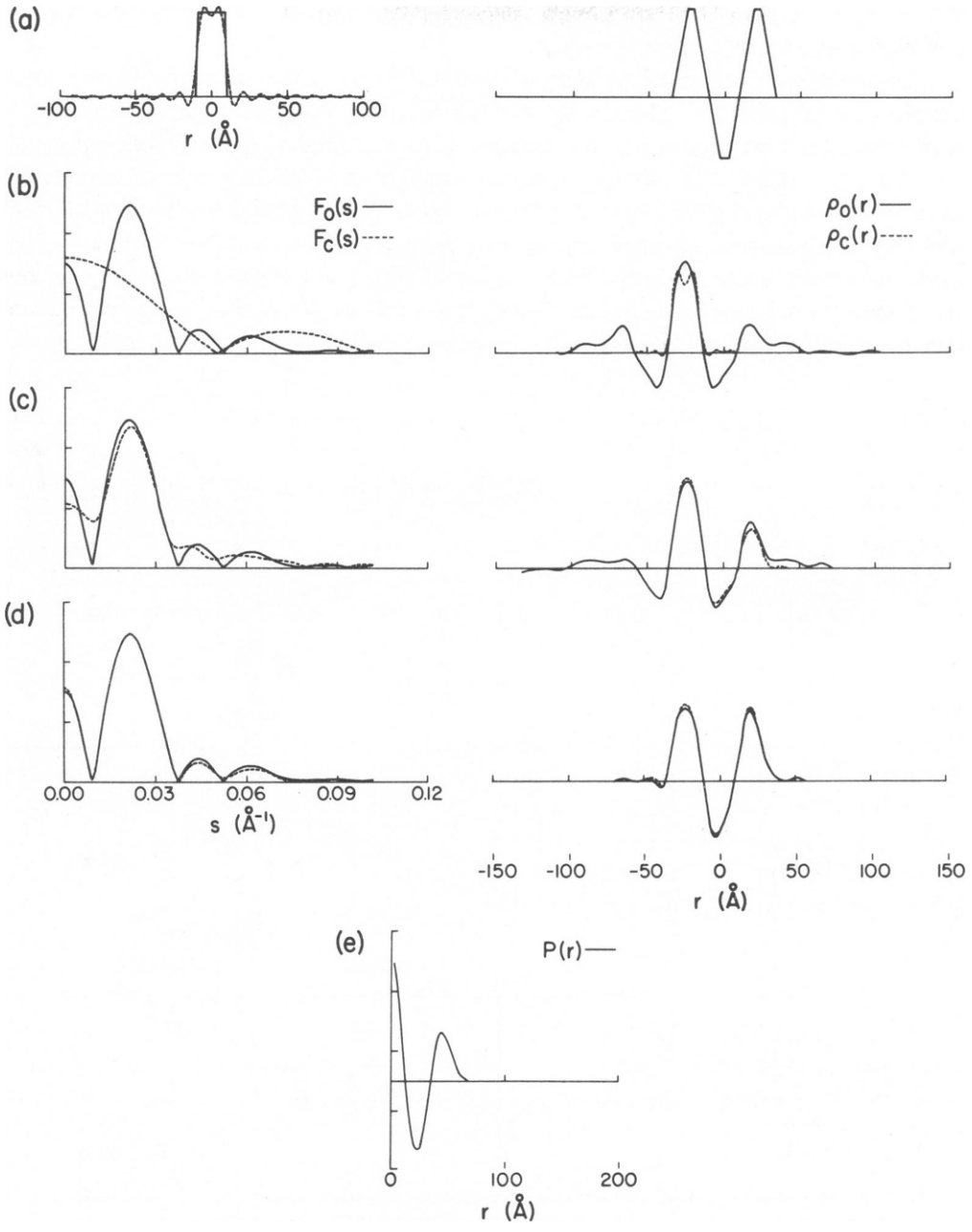

FIGURE 5 Test of solution at 10Å resolution for data generated from a centrosymmetric structure is shown in part 2 of A. The trial model used to initiate refinement was a box function (solid line in part 1 of A). Fourier transformation of the $|F_c(s)|$, $\phi_c(s)$ generated from the trial model (A), and cut off at 10 Å (- - - - - - - - in A) contains series termination ripples. At the first cycle (B) observed $|F_0(s)|$ (——) were coupled with phases calculated from the trial model (A) to generate a new model $\rho_0(r)$. This model was used to generate amplitudes $|F_c(s)|$ (- - - - - -) and phases $\phi_c(s)$. ($\rho_c(s)$ is computed from $|F_c(s)|$, $\phi_c(s)$). The procedure was iterated, and results after 4 cycles (C), and after 14 cycles (D), show convergence onto an essentially correct structure, which may be compared with A above. The agreement between the Patterson and autocorrelation function is essentially perfect (E).

solution using the Fourier method fitted the conclusions and the ambiguities derived by Worthington and McIntosh from these data.

The myelin data also represent an ideal system for illustrating the theoretical basis for the multiple solution problem. The normal frog sciatic nerve myelin structure gives rise to Bragg reflections that lie on maxima determined by the minimum multilayer repeat of **d** = 171 Å. A width of 180–190 Å was chosen as an estimate of the thickness of a single repeated unit. A width slightly larger than the true repeat Bragg spacing was chosen so as to allow a smooth truncation of series termination ripples within the boundary thickness. The Fourier refinement method was applied to several different starting models bearing no relationship to the solutions obtained by Worthington and McIntosh (1974). One solution summarized in Fig. 6 again indicates that convergence on an almost centrosymmetric struc-

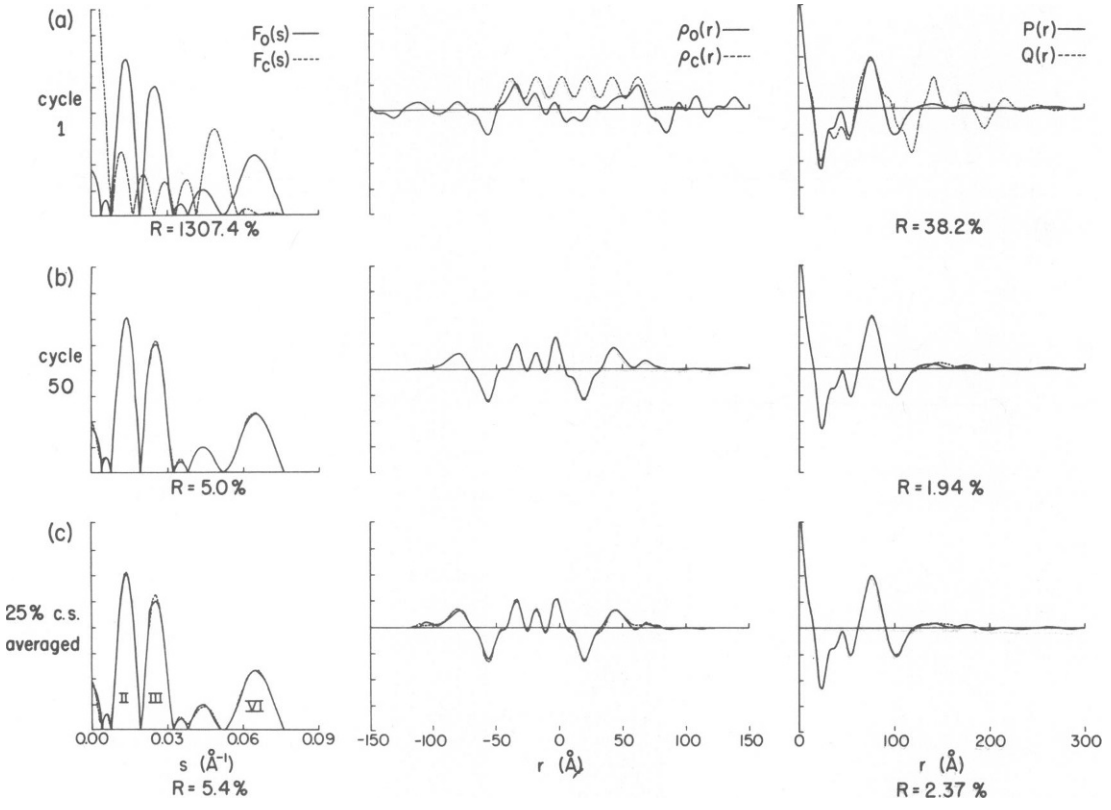

FIGURE 6 Test of solution using the continuous data $|F_0(\mathbf{s})|$ for frog sciatic nerve myelin (Worthington and McIntosh, 1974) to a resolution of $s = 1/14$ Å^{-1}. The trial model for initiating refinement was $\rho_c(\mathbf{r})$ (- - - - - - - - in A); and at cycle 50 (B), are shown (left). The central figures show the calculated model generated by transformation of $|F_0(\mathbf{s})|$, $\phi_c(\mathbf{s})$ $[\rho_0(\mathbf{r})]$ and of $|F_c(\mathbf{s})|$, $\phi_c(\mathbf{s})$ $[\rho_c(\mathbf{r})]$ at corresponding cycles. Agreement between $P(\mathbf{r})$ and $Q(\mathbf{r})$ is indicated at the right side. Values for the residual R at each stage are indicated below each data summary. The residual values underneath the Patterson summary are R_P values. The structure $\rho_0(\mathbf{r})$ shown in B is not completely centrosymmetric, and the two sides of the structure were subjected to weighted averaging to achieve pressure toward centrosymmetricity. Weighting was 0.75 of $\rho_0(\mathbf{r}) + 0.25$ of $\rho_0(r_0 - \mathbf{r})$ with respect to a pseudocenter at the center of mass (r_0). Results are shown in C. The signs indicated for the zero order, and the six regions of the transform between zeros are (+) − + − + − + (the "alternate" structure of Worthington and McIntosh, 1974).

ture is achieved. Further refinement cycles tend to generate more perfectly centrosymmetric structures. The resulting structure (Fig. 6 B) was similar to one of the solutions obtained by Worthington and McIntosh (1974)—their "alternate" structure.

It was possible to incorporate knowledge of centrosymmetry in the final stages of refinement by defining a new $\rho(\mathbf{r})$ composed of 75% of the original value and 25% of the symmetrically related value of $\rho(\mathbf{r})$ (Fig. 6 C). This constraint allows the phases to refine to the correct values, whereas forcing them to be signs, + or −, would not.

In some cases (see Fig. 7) the myelin refinement yielded noncentrosymmetric solutions. This is to be expected in light of the greater degrees of freedom available with general phases. Application of the real-space averaging procedure (above) resulted in an alternative symmetric solution to that of Fig. 6. In all, four different centrosymmetric solutions for the myelin structure were obtained that matched our criteria for an acceptable solution. They correspond to the four possible permutations of − or + sign choices for regions V and VI of the observed transform (see below), and were similarly identified as possible ambiguities by Worthington and McIntosh (1974); two of these solutions were favored. This raises the question as to why there should be four homometric solutions, and ultimately shows that two of these structures cannot be distinguished without more information than is contained in the one continuous transform. The other two structures can in theory be

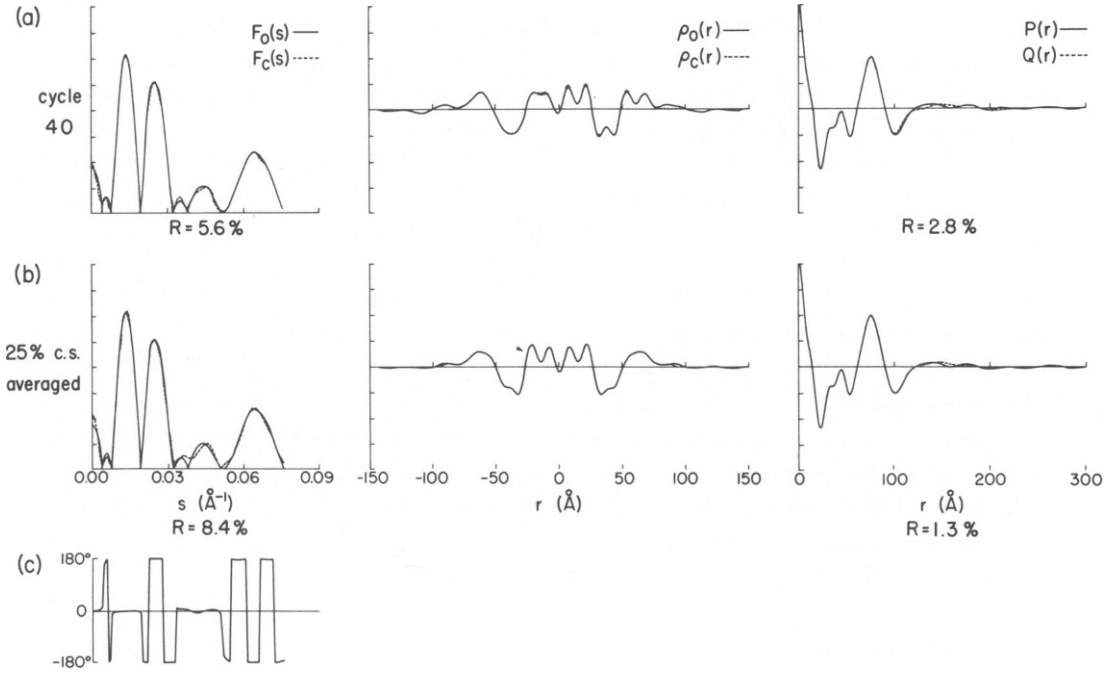

FIGURE 7 Another of the four solutions to the myelin transform, (*a*), after 40 cycles of refinement, is noncentrosymmetric, although the structure retains symmetric features of the structure permuted in cyclic order. Pressure toward centrosymmetricity (using 0.75/0.25 weights) leads to a symmetric structure with somewhat better agreement between $P(\mathbf{r})$ and $Q(\mathbf{r})$ (B), (R_P = 1.3%). The phases derived from refinement (C) indicate a sign choice (+) − + − + + − for the zero-order component, and the six transform regions between zeros.

eliminated, but this requires essentially perfect data, or data from different species as was used by Worthington and McIntosh in their analysis.

Worthington and McIntosh (1974) identified six zeros ($|F(\mathbf{s})|^2 = 0.0$) in the observed transform out to a resolution of 7 Å ($\mathbf{s} = 1/14$ Å). They identified seven regions of the transform separated by the six zeros, which were labeled regions 0, I, II, III, IV, V, and VI. The ambiguities in solution by the iterative Fourier method are directly associated with the phases (or signs) of regions V and VI. Sign changes in the latter region cannot be discriminated without additional information beyond that contained in a single continuous transform from one species.

Each region of the continuous transform contributes a Fourier component to the electron density synthesis, and in each case, the contribution of just one region is a wave packet that resembles a cosine curve of frequency determined by the s value where the local value of $|F(\mathbf{s})|$ is maximum multiplied by a smooth envelope that is qualitatively like either a Gaussian or a zero-order Bessel function. Looked at another way, each region of the transform $F(\mathbf{s})$ is a delta function at the s value of the center of the peak, convoluted with the profile of the region. The contribution to electron density synthesis is thus the product of a cosine wave with a Gaussian-like envelope. The width of the envelope is inversely proportional to the width of the transform region. Each one of the regions 0–V of the transform is sufficiently narrow (<0.013 Å^{-1}) so that its contribution to the density synthesis extends beyond the constrained thickness of the structure. Region VI, on the other hand, is at least 0.02 Å^{-1} in width, and its contribution to the resulting electron density synthesis is confined well within the thickness (180 Å) of the profile (Fig. 6). There is no pressure on the refinement process to phase this region, since its contribution to the profile outside the boundary is too small to be significant. As a direct consequence of this fact, the autocorrelation function is the same as the Patterson synthesis for both structures when computed using either sign for the data of region VI. Worthington and McIntosh (1974) presented two structures for frog sciatic nerve myelin that differed in the same way—namely, in the choice of sign for region VI of the profile. Their results were obtained by direct methods of deconvoluting the Patterson function, based upon the theory of Hosemann and Bagchi (1962) and reconstruction methods based on the Shannon (1949) sampling theory applied to the sampled transform data. The transform of region V is approximately the same size as the total width of the double bilayer structure and because it is rather small, its sign cannot practicably be determined from a single specimen scattering pattern alone. Different sign choices from those we found (the same as discussed by Worthington and McIntosh, 1974) have been proposed earlier by Caspar and Kirschner (1971), and by Blaurock and Worthington (1969) using different data. Our analysis (which would not allow these solutions) applies strictly to the data set used.

DISCUSSION

From the applications of the Fourier method to continuous scattering data, the following generalizations can be made:

(*a*) The starting model for initiating refinement is arbitrary, and its only function is to generate a set of phases, $\phi(\mathbf{s})$, which roughly correspond to a model centered somewhere near the arbitrary $\mathbf{r}_0$. The initial model may be useful in testing different types of structure.

(*b*) Successive cycles of refinement from the starting model always lead to convergence on a solution, usually after 15–40 cycles.

(*c*) In every case we tested from artificially synthesized data, $|F(\mathbf{s})|$, one solution of a very small number was correct, suggesting that in general the number of solutions will be small. The difference between different solutions was often small and probably reflects the lower pressure to phase the lowest amplitudes in the transform $F(\mathbf{s})$, since they contribute least to the $\rho(\mathbf{r})$ synthesis (Eq. 1). Often a unique solution is found.

(*d*) If the assigned boundary condition, w, is chosen to be too small, refinement can never adequately lead to fitting even the first minimum of the transform. If w is too large (several times the correct thickness) then refinement is slower, since the pressure to produce alterations in the structure at each cycle is diminished.

(*e*) Alternate structures can be identified by the Fourier method where they exist, whereas they cannot necessarily be found by direct methods (King, 1975), or by normal methods of deconvolution of the Patterson function. In those cases where more than one solution exist, one cannot be certain that all possible solutions have been obtained when using only a limited set of starting models. It is possible, however, to test various classes of solutions by appropriate choice of starting models. Since most asymmetric membranes are based on a bilayer structure, chemical reasonableness may be a basis for preference rather than proof. Nevertheless, our concern was primarily with the location of asymmetric protein or other components at relatively low resolution in a simple structure based on a bilayer. Is the protein localized only on one side of a membrane, for example, or does the protein traverse the lipid bilayer in any particular case? The method can be extremely powerful in determining the range of possible solutions in these cases.

The essential component of the Fourier refinement method applied to continuous diffraction is the assumption that the scattering object is of finite, rather than infinite, dimension. This condition is applied by the imposition of a finite size to the density profile. This restraint allows convergence. It also explains the pressure to converge on a definite solution, since it forces all Fourier components used to calculate $\rho(\mathbf{r})$ at each cycle to account correctly for a constant solvent density outside the boundary limits. Since the calculated amplitudes $|F_c(\mathbf{s})|$ must match the observations $|F_0(\mathbf{s})|$, the phases must account for the featureless solvent region. Thus, the method is essentially refinement of the solvent density for $\mathbf{r} = -\infty$ to a; b to ∞.

Alternative Approaches to Solution

With the exception of our own studies of acetylcholine receptor membranes (Ross et al., 1977) and the work on purple membrane that is only slightly asymmetric (Blaurock and King, 1977), solutions to membrane profiles, $\rho(\mathbf{r})$, have so far only been attempted for centrosymmetric structures, i.e., cases in which the repeating unit consists of a pair of membranes generated by the folding over of a cell envelope or artificial lipid bilayers without protein.

The problems associated with phasing of lamellar diffraction from symmetric structures have been approached in several different ways. A number of methods for deriving the electron density profile have been proposed, and most of these depend on a deconvolution of the Patterson function $P(\mathbf{r})$. Several developments of the deconvolution method of Hosemann and Bagchi (1962) have been used. These and other methods have been critically analyzed by Worthington, et al. (1973). Since then Moody (1974) has proposed other methods for

achieving deconvolution in the particular case where swelling experiments can be used to map the continuous transform from multilayer systems. Luzzati et al. (1972) also proposed a pattern recognition approach based on interpretation of even higher order autoconvolutions of $F(\mathbf{h})$ in terms of higher powers of $\rho(\mathbf{r})$.

In the case of asymmetric membrane profile analysis, general deconvolution of the Patterson function has been proposed, but to this point remains untested. A direct method based on the Hilbert transform (Papoulis, 1962; King, 1975) has recently been applied successfully to the purple membrane profile, but only one of the two possible solutions could be found (Blaurock and King, 1977). A general and severe problem with these and other methods providing a single solution from a single data set is the uncertainty that always remains as to whether other solutions exist, and if so, how many might equally well fit the data.

Blaurock and King (1977) also presented a trial and error method which, as with our method, uses knowledge of the membrane thickness. The continuous transform was sampled at intervals determined by a hypothetical unit cell. By using all possible phase combinations for these "reflections," the continuous transform was reconstructed by the Shannon method (1949) and compared to the observed transform. The observed and reconstructed functions would be identical (within data errors) for a correct phase choice. This process is equivalent to choosing a combination of phases that will generate a $\rho(\mathbf{r})$ contained entirely within the "unit cell" (our constraining thickness, w). As our and their approaches employ the same physical criterion for determining a solution, they would provide identical solutions if the trial phases were sampled at infinitesimally small intervals. The method of Blaurock and King (1977) requires that all possible phase choices be tried, and thus in principle, it can obtain all solutions. Within the bounds of practicality, all phase combinations cannot be tried. In an effort to circumvent this problem, the authors chose the rather coarse interval of $\pi/8$ between trial phases to limit the number of tests. Even so, this results in 16^n Shannon reconstructions for n reflections. For purple membrane where a 50 Å "unit cell" was chosen, only two reflections were involved at the observed resolution (196 trials). When larger structures such as the acetylcholine receptor (110 Å actual thickness, Ross et al., 1977) are considered, where higher resolutions are required, or where smaller phase intervals are used, the computation involved would be prohibitive. Application to the acetylcholine receptor, for example, would require $16^9 = 6.9 \cdot 10^{10}$ trials and is clearly impractical. In our Fourier refinement technique, consistent sets of phases are jointly developed in a convergent stochastic process, and are not restricted to be multiples of $\pi/8$. In our approach it is also possible to incorporate other knowledge about the structure (e.g. limits on allowable densities, centrosymmetricity, etc.) whenever it is available.

Asymmetric Membranes and Continuous Diffraction

It is possible in many cases to prepare asymmetric membrane fractions from specialized tissue, which contain a highly enriched biochemical function. Membranes containing acetylcholine receptors, for example, have been prepared from the electric organ of the electric fish *Torpedo californica*. Other fractions prepared from the same organ show that separate patches contain other biochemical functions such as acetylcholine esterase and ATPase. In this system, at least, it would seem that specialized functions are contained within specialized regions of the synapse (Cohen et al., 1972; Duguid and Raftery, 1973). In the case of acetylcholine receptor membranes, it is clear that there is a structural asymmetry

to the membrane and that receptor molecules are all aligned in the same way, and sometimes organized into an array within each sheet (Ross et al., 1977). This type of situation is also known in bacterial systems, where the purple membrane of *Halobacterium halobium* has been shown to contain an organized, directionally polar array of bacteriorhodopsin molecules (Henderson, 1975). Thus we expect there will be a large number of systems such as these where the Fourier method of structure determination will initially provide valuable information on the distribution of macromolecules in a membrane.

The necessary X-ray data may be generated from vesicle dispersions. Such X-ray scattering patterns have been acquired in several cases for membranes isolated from cells or organelles (Engleman, 1971; Wilkins et al. 1971; Lesslauer, et al., 1971; Lesslauer and Blasie, 1972; and Blaurock, 1973). In some cases centrifugation or drying of vesicles has been used to record lamellar diffraction (Finean and Burge, 1963; Dupont, et al., 1973; and Worthington and Liu, 1973), and some of the complex problems associated with the statistical arrangement of stacked asymmetric structures were detailed in the latter presentation. The iterative Fourier method can be simply applied to continuous scattering data and avoids the complex problems of stacking disorder. It is a simple procedure and can identify and refine possible solutions for the electron density profile from either asymmetric or symmetric membranes.

Since this Fourier method deals directly with $|F_0(\mathbf{s})|$, rather than higher powers of $|F_0(\mathbf{s})|$, as in deconvolution of the Patterson function by the method of moments or recursion, the solution is less sensitive to errors in the data.

In a much broader context it is clear that any solution to structural problems using diffraction techniques can be improved when physical and chemical constraints or noncrystallographic symmetry relationships can be applied. Such constraints can best be applied in the electron density domain. Working in the Patterson synthesis, or with intensities $|F(\mathbf{s})|^2$, or with higher order convolutions of $\rho(\mathbf{r})$ inevitably makes the analysis more complex and does not easily allow for iterative improvements in the result.

This work was supported by National Institutes of Health grant GM-24485 and National Science Foundation grant PCM77-25407.

Received for publication 13 May 1977 and in revised form 18 November 1978.

REFERENCES

BLAUROCK, A. E. 1971. Structure of the nerve myelin membrane: proof of the low-resolution profile. *J. Mol. Biol.* **56**:35–52.

BLAUROCK, A. E. 1973. The structure of a lipid-cytochrome c membrane. *Biophys. J.* **13**:290–298.

BLAUROCK, A. E. 1976. Myelin X-ray patterns reconciled. *Biophys. J.* **16**:491–501.

BLAUROCK, A. E., and KING, G. I. 1977. Asymmetric structure of the purple membrane. *Science (Wash. D.C.).* **196**:1101–1104.

BLAUROCK, A. E., and W. STOECKENIUS. 1971. Structure of purple membrane. *Nat. New Biol.* **233**:152–155.

BLAUROCK, A. E., and C. R. WORTHINGTON. 1969. Low-angle X-ray diffraction patterns from a variety of myelinated nerves. *Biochim. Biophys. Acta* **173**:419–426.

BURGE, R. E., and J. C. DRAPER. 1967. The structure of the cell wall of the Gram-negative bacterium *Proteus vulgaris*. III. A lipopolysaccharide "unit membrane." *J. Mol. Biol.* **28**:205–210.

CASPAR, D. L. D., and D. A. KIRSCHNER. 1971. Myelin membrane structure at 10 Å resolution. *Nat. New Biol.* **231**:46–52.

COHEN, J. B., M. WEBER, M. HUCHET, and J. P. CHANGEUX. 1972. Purification from *Torpedo marmorata* electric tissue of membrane fragments particularly rich in cholinergic receptor protein. *FEBS (Fed. Eur. Biochem. Soc.) Lett.* **26**:43–47.

DUGUID, J. R., and M. A. RAFTERY. 1973. Fractionation and partial characterization of membrane particles from *Torpedo californica* electroplax. *Biochemistry.* **12**:3593–3597.

DUPONT, Y., S. C. HARRISON, and W. HASSELBACH. 1973. Molecular organization in the sarcoplasmic reticulum membrane studied by X-ray diffraction. *Nature (Lond.).* **244**:555–558.

ENGLEMAN, D. M. 1971. Lipid bilayer structure in the membrane of Mycoplasma laidlawii. *J. Mol. Biol.* **58**: 153–165.

FINEAN, J. B., and R. E. BURGE. 1963. The determination of the Fourier transform of the myelin layer from a study of swelling phenomena. *J. Mol. Biol.* **7**:672–682.

HENDERSON, R. 1975. The structure of the purple membrane from *Halobacterium halobium:* analysis of the X-ray diffraction pattern. *J. Mol. Biol.* **93**:123–138.

HOSEMANN, R., and S. N. BAGCHI. 1962. Direct analysis of diffraction by matter. North-Holland Publishing Company, Amsterdam. 300–500.

KING, G. I. 1975. Direct structure determination of asymmetric membrane systems from X-ray diffraction. *Acta. Crystallogr. Sect. A.* **31**:130–135.

LESSLAUER, W., and J. K. BLASIE. 1972. Direct determination of the structure of borium stearate multilayers by X-ray diffraction. *Biophys. J.* **12**:175–191.

LESSLAUER, W., J. E. CAIN, and J. K. BLASIE. 1971. On the location of 1-anilino-8-naphthalene-sulfonate in lipid model systems. An X-ray diffraction study. *Biochim. Biophys. Acta.* **241**:547–566.

LEVINE, Y. K., and M. H. F. WILKINS. 1971. Structure of oriented lipid bilayers. *Nat. New Biol.* **230**:69–72.

LIPSON, H., and C. A. TAYLOR. 1958. Fourier transforms and x-ray diffraction. G. Bell & Sons Ltd., London. 20–27.

LUZZATI, V., A. TARDIEU, and D. TAUPIN. 1972. A pattern recognition approach to the phase problem: application to the X-ray diffraction study of biological membranes and model systems. *J. Mol. Biol.* **64**:269–286.

MOODY, M. F. 1963. X-ray diffraction pattern of nerve myelin: a method for determining the phases. *Science (Wash. D.C.).* **142**:1173–1174.

MOODY, M. F. 1974. Structure determination of membranes in swollen lamellar systems. *Biophys. J.* **14**:697–703.

PAPOULIS, A. 1962. The Fourier integral and its applications. McGraw-Hill Book Company, New York. 192–211.

ROBERTSON, J. D. 1958. Structural alterations in nerve fibers produced by hypotonic and hypertonic solutions. *J. Biophys. Biochem. Cytol.* **4**:349–364.

ROSS, M. J., M. KLYMKOWSKY, D. A. AGARD, and R. M. STROUD. 1977. Structural studies of a membrane-bound acetylcholine receptor from *Torpedo californica. J. Mol. Biol.* **116**:635–659.

SHANNON, C. E. 1949. Communication in the presence of noise. *Proc. IRE.* **37**:10–21.

WILKINS, M. H. F., A. E. BLAUROCK, and D. M. ENGLEMAN. 1971. Bilayer structure in membranes. *Nat. New Biol.* **230**:72–76.

WORTHINGTON, C. R., G. I. KING, and T. J. MCINTOSH. 1973. Direct structure determination of multilayered membrane-type systems which contain fluid layers. *Biophys. J.* **13**:480–495.

WORTHINGTON, C. R., and S. C. LIU. 1973. Structure of sarcoplasmic reticulum membranes at low resolution (17Å). *Arch. Biochem. Biophys.* **157**:573–579.

WORTHINGTON, C. R., and T. J. MCINTOSH. 1974. Direct determination of the lamellar structure of peripheral nerve myelin at moderate resolution (7Å). *Biophys. J.* **14**:703–729.