Direct determination of crystallographic phases for diffraction data from lipid bilayers

I. Reliability and phase refinement

Douglas L. Dorset
Electron Diffraction Department, Medical Foundation of Buffalo, Inc., Buffalo, New York 14203

ABSTRACT Direct analysis of lipid lamellar packing based on the probabilistic estimate of Σ_1- and Σ_2-triplet phase invariants is evaluated here for a larger variety of bilayer structures than examined in an original study of this problem (Dorset, D.L., 1990. *Biophys. J.* 58:1077–1087). Using x-ray crystal structures of five phospholipids, three glycerides and two cerebrosides, lamellar diffraction data were generated at the ~3 Å resolution often found experimentally from oriented multilayers. For structures where no significant density occurs at the unit cell origin, the ab initio phase determination is successful for six of the ten structures. A seventh structure can be solved if a limited set of Σ_2-triples are used to determine the initial phase set based on the hierarchy of the A_2 values. Bilayers, e.g., with solvent at the origin, can be analyzed if a modified criterion for accepting phase estimates for Σ_1-triples is used, as suggested by the distribution of normalized structure factors and the number of probable single-valued phase domains. In all cases, partial phase determinations can be refined effectively by density modification ("flattening") of the hydrocarbon region in real space. A figure of merit suggested by Luzzati et al. (Luzzati, V., A. Tardieu, and D. Taupin. 1972. *J. Mol. Biol.* 64:269–286) used to evaluate the success of such refinement can be supplemented by an evaluation of density smoothness, which can also detect the presence of near structure homomorphs not identified by the former test for density flatness.

INTRODUCTION

In recent years, we have been interested in developing electron crystallographic techniques for the quantitative structure analysis of lipid bilayers, in order to exploit the use of lamellar data from microcrystalline samples that can be epitaxially-nucleated readily on organic substrates (Dorset et al., 1983; Dorset, 1990*a*). Although there may be some perturbation by secondary scattering (Cowley et al., 1951), the measured electron diffraction intensity data, nevertheless, adequately correspond to the kinematical scattering approximation (i.e., the crystallographic R-factor is ≤ 0.31) so that ab initio structure analyses can be carried out. Initial determinations of such structures were based on an idea proposed by Hitchcock et al. (1975) for the phasing of lamellar x-ray data from a phosphatidylethanolamine, i.e., that a conformational model based in a similar crystal structure be used in a translational search for a minimum of the crystallographic residual in space group $P\bar{1}$. Electron diffraction structure analyses of this kind on a variety of representative phospholipid bilayers (Dorset, 1987, 1988*a*, *b*; Dorset el al., 1987; Dorset and Zhang, 1990) seemed to produce reasonable results, sometimes indicating the presence of headgroup conformations or lamellar packings not found in x-ray crystal structures of close analogues. On the other hand, the significance of the crystallographic R-factor as a figure of merit (Hamilton, 1964) for such analyses is always problematic, especially if one has to distinguish between two possible structures which result in minima with comparable depth (Dorset, 1982; Dorset et al., 1987).

It had also been found that high resolution electron microscope images could be obtained from these bilayer arrays (Fryer and Dorset, 1987) and that the Fourier transform of these images were useful for directly determining crystallographic phases for the low angle diffraction data (Dorset, 1988*a*). Experiments on a cryo-electron microscope demonstrated that the ~16 Å resolution seen in images obtained on a conventional instrument could be extended to at least 6 Å (Dorset et al., 1990) and that the crystallographic phases obtained from the computed Fourier transform of these images exactly matched the results from a structure analysis employing a model derived from the x-ray crystal structure of a very similar lipid (Elder et al., 1977).

While these results from high resolution image data were encouraging, there was still a need fo find an unbiased method for extending the phase determination to the ~3 Å resolution limit found in many electron diffraction patterns from such lipids. Direct methods, based on the probabilistic estimate of structure invariant phase sums (Hauptman, 1972), were employed for this purpose and found to be successful in assigning correct values to the phases not defined by the Fourier transform of the electron microscope image (Dorset et al., 1990). Surprisingly enough, it was also discovered that these direct phase determinations could also be useful, even without supplementary electron microscope im-

0006-3495/91/12/1356/10 $2.00

ages, in many cases finding phase values for enough reflections to permit calculation of a directly interpretable structure map. This was demonstrated for a number of electron diffraction data sets and also x-ray data sets from other laboratories, for which the phase determinations had been originally made by other methods (Dorset, 1990*b*).

Although the success of direct methods is unexpected (because the data do not extend to atomic resolution), a number of questions remain to be answered before they can be used with confidence for any lipid data set. Thus, how reliable is this phase determination for the analysis of other bilayer structures? In the initial work on a variety of different phospholipids, one can nevertheless argue that all examples have very similar molecular architectures (especially headgroup conformations) and layer structures. Is it possible to determine correctly the profiles of structurally different lipid bilayers, i.e., composed of phospholipids with different headgroup geometries? Additionally, if such analyses are possible, they will often be incomplete. For example, in the ab initio determination of the 1,2-dihexadecyl-*sn*-glycerophosphoethanolamine bilayer structure from electron diffraction data (excluding electron microscope images), 13 of 16 phase values were found (Dorset et al., 1990). Is it possible, therefore, to refine the phases after the initial determination to produce a more complete set and thus a more accurate representation of the layer packing?

This paper attempts to answer these questions by carrying out direct structure determinations for phospholipid bilayers with different headgroup geometries than considered before and also for several other glycerol- and sphingosine-based compounds that also pack in bilayers. In many cases, it will be found that the analysis leads to the correct structure. A possible method for phase refinement will also be discussed which is based on the "solvent flattening" techniques used in protein crystallography (Wang, 1985).

DIFFRACTION DATA

In order to evaluate the applicability of the direct phase determinations to diffraction data from the lipid multilayers, model lamellar x-ray diffraction data were calculated from the atomic coordinates, given in the x-ray crystal structures of a number of bilayer structures listed in Table 1, and employing the structure factor expression

$$F_{\vec{h}} = 2 \sum_j f_j \cos 2\pi(\vec{h} \cdot \vec{r}_j). \quad (1)$$

Here, $\vec{h}$ is a reciprocal lattice vector for the row with index 00ℓ and r_j and z-coordinates of atomic positions along the normal to the bilayer. As can be appreciated from the examination of individual crystal structures, the centrosymmetry implied by Eq. 1 (hence space group $P\bar{1}$) can be justified for all of the lamellar data used in this study (Table 2). This restricts the phase choice to 0, π for any reflection. For the structure factor calculations, Doyle-Turner (1968) scattering factor values, f_j, (also found in the International Tables for X-Ray Crystallography, Vol. 4 [Ibers and Hamilton, 1974]), were used without an isotropic temperature factor. (As it turns out, this assumption that $B_{iso} = 0.0$ Å^2 is not a bad approximation for the low angle diffraction data.) The data resolution used (e.g., 3 Å) is about the same as that often encountered experimentally (e.g., see examples treated earlier [Dorset, 1990*b*]). One of the compounds listed in Table 1, i.e., GDLT, is based on the crystal structure of its ω-brominated analogue, (Watts et al., 1972), using atomic coordinates reported by Pangborn (1973). From the calculated structure factor magnitudes $|F_{\vec{h}}|$, normalized values $|E_{\vec{h}}|$ are computed from the structure factor magnitudes in

TABLE 1 Bilayer lipid structures used for direct phase determination

Compound	Lamellar spacing (Å)*	No. of reflections	Reference
1,2-dilauroyl-*rac*-glycero-phosphoethanolamine · acetic acid (DLPE · Ac)	47.66	15	Elder et al. (1977)
1,2-dimyristoyl-*sn*-glycero-phosphocholine · H_2O (DMPC)	54.95	15	Pearson and Pascher (1979)
1,2-dimyristoyl-*sn*-glycero-phospho-*rac*-glycerol, sodium salt (DMPG)	45.31	16	Pascher et al. (1987)
1,2-dilauroyl-*rac*-glycero-phospho-N,N-dimethylethanol-amine ($DLPEM_2$)	39.73	13	Pascher and Sundell (1986)
1,2-dimyristoyl-*sn*-glycero-phosphate, monosodium salt (DMPA)	40.11	14	Harlos et al. (1984)
rac glycerol 1,2(dilaurate) 3(p-toluene sulfate) (GDLT)	40.63	13	Watts et al. (1972)
1,2-dilauroyl-*sn*-glycerol (GDL)	34.15	13	Pascher et al. (1981)
1,2-dipalmitoyl-*sn*-glycerol (GDP)	43.25	16	Dorset and Pangborn (1988)
cerebroside · EtOH (CER)	45.89	17	Pascher and Sundell (1977)
methyl cerebroside (MCER)	44.05	16	Nyholm et al. (1990)

*The d_{001} spacing of the original crystal structure where c is taken to be the longest unit cell axis. (A shift of a and c values is required for the DLPE · Ac structure to conform to this assignment.)

TABLE 2 **Structure factors used in direct phase analysis**

ℓ	DLPE · Ac	DMPC	DMPG	$DLPEM_2$	DMPA	GDLT	GDL	GDP	CER	MCER
1	255.6	259.0	357.3	122.9	158.6	123.5	83.8	88.3	311.2	219.9
2	16.3	−43.7	122.0	18.0	64.0	3.0	15.3	26.6	80.1	−10.6
3	−12.4	−60.2	87.9	23.2	57.0	−13.6	21.0	41.3	14.5	−12.0
4	−67.5	−117.7	−37.3	−7.6	4.1	−39.6	−29.0	−15.9	−186.2	−125.7
5	0.2	−9.2	44.2	40.5	26.9	−28.7	−12.2	−0.7	−155.3	28.3
6	−14.8	−15.2	19.7	1.4	26.2	−48.0	−28.0	−30.8	−175.6	−40.4
7	−46.9	−7.5	50.3	−2.6	34.4	−64.0	−19.8	−11.2	−39.3	−25.2
8	−33.7	−14.8	−51.2	−70.9	36.1	−45.2	−30.9	−23.8	−18.3	−178.9
9	−38.9	−18.4	−101.3	−69.0	22.9	−61.1	−49.8	−19.6	−22.9	−143.8
10	41.5	−0.4	−180.1	−83.2	46.7	−52.7	−46.0	−27.5	−55.5	−79.1
11	37.7	−23.3	−201.2	−43.9	38.4	−87.2	−59.2	−41.6	−53.2	56.9
12	123.0	−28.3	−185.1	−35.8	67.9	−26.5	−18.9	−42.5	−65.1	66.5
13	151.5	−34.1	−162.3	3.7	38.8	−2.7	−20.0	−59.6	52.1	16.5
14	198.6	−58.7	−88.8		31.7			−35.9	100.4	−83.4
15	107.9	−6.2	−86.1					−46.8	166.2	−105.7
16			−46.2					−3.8	130.2	−47.8
17									69.1	
18									−44.9	

Table 2 by

$$|E_{\vec{h}}|^2 = |F_{\vec{h}}|^2 \Big/ \sum_i f_i^2. \tag{2}$$

The normalization assumes that $\langle |E_{\vec{h}}|^2 \rangle_{\vec{h}} = 1.0$.

DIRECT PHASE DETERMINATION

As applied in our earlier work, direct determination of crystallographic phases is based on the probabilistic estimate of three phase invariants in space group $P\bar{1}$ (Hauptman, 1972). That is, one predicts the value of the structure invariant phase sums

$$\phi = \psi_{\vec{h}_1} + \psi_{\vec{h}_2} + \psi_{\vec{h}_3}, \tag{3}$$

where the Miller indices $\vec{h}_i \equiv 00\ell_i$ have the constraint $\vec{h}_1 + \vec{h}_2 + \vec{h}_3 \equiv 0$. (Hence the ϕ sums the phase values for three individual reflections $\psi_{\vec{h}_i}$.) For the general case where $\vec{h}_1 \neq \vec{h}_2 \neq \vec{h}_3$ the phase invariant is called a Σ_2-triple and the probability of its predicted value being correct is directly proportional to the value of $A_2 = k|E_{\vec{h}_1} E_{\vec{h}_2} E_{\vec{h}_3}|$, where $k = 2/\sqrt{N}$ is a scaling constant based on the number of nonhydrogen atoms in the unit cell. When $\vec{h}_1 = \vec{h}_2 = -\frac{1}{2}\vec{h}_3$, one has a Σ_1-triple and probability of predicting its value is based on

$$A_1 = (|E_{\vec{h}}|^2 - 1)|E_{2\vec{h}}| / \sqrt{N}. \tag{4}$$

In x-ray crystallography, as well as other applications of direct phasing methods to electron crystallography, the prediction of three phase invariant sums is most reliable for the Σ_2-triplets. In general, only a few phase values from Σ_1-triples can be found with confidence. The surprising success of the Σ_1-triples for predicting phases for these lamellar data is subject to constraints that will be discussed below.

To evaluate these simultaneous equations in phase, one is permitted to define the phase of one reflection a priori in order to fix the unit cell origin (Dorset, 1990*b*). This reflection must be of odd index ($h = 00\ell$, where $\ell = 2n + 1$) and, generally, the phase of the 001 reflection is used for this purpose.

RESULTS

In the direct phase determination with data generated from the 10 crystal structures listed in Table 1, six ab initio analyses were carried out with no difficulty, but three others required a revision of the criterion used to accept estimates of Σ_1-triples. One other structure analysis needed to be initiated with one erroneous Σ_2-triplet in the starting set of phase invariants. The details of these analyses are given in the following discussion.

1. Correct structure determinations

(a) DMPC. Unlike the lecithin bilayers considered in earlier x-ray (Torbet and Wilkins, 1976) and electron diffraction (Dorset, 1987*a*) studies, the crystal structure described by Pearson and Pascher (1979) contains two distinct headgroup conformations. Nevertheless, as schematized in Fig. 1 *a*, 11 Σ_2-quartets with $A_2 \geq 0.01$ can be used to link the values of 13 of the 15 reflections after $\psi_{001} = 0$ is chosen to define the origin. Additionally, the values of four Σ_1-triples can be used where A_1 has a negative value, to predict that $\psi_{004} = \psi_{006} = \psi_{0012} = \psi_{0014} = \pi$, thus assigning phases to all the reflections linked by the Σ_2-triples. Only the value of ψ_{0010} remains undefined. A one-dimensional electron density map is shown in Fig. 2 *a*.

(b) DMPG. The values of nine Σ_2-triples where $A_2 \geq 0.13$ are found to be correct, linking reflections as shown in Fig. 1 *b*. A large $A_1 = 0.30$ is taken to indicate that $\psi_{002} = 0$, whereas three negative $A_1 \leq -0.04$ determine

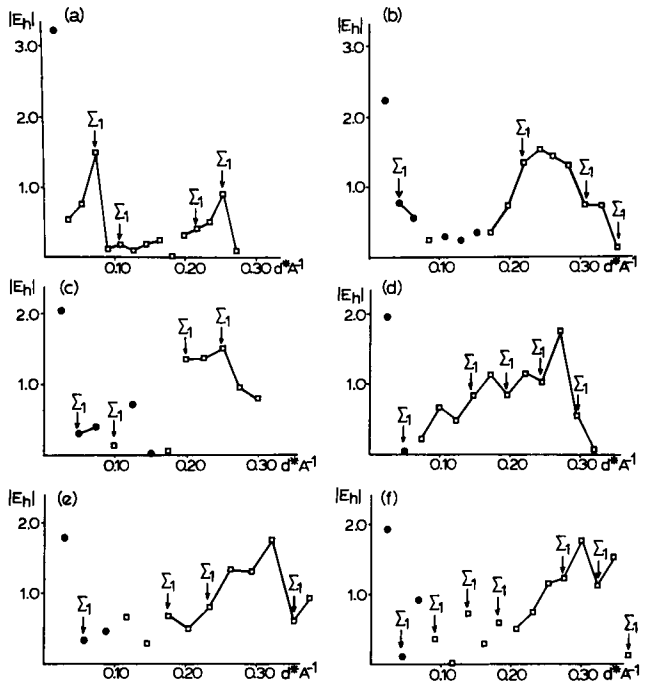

FIGURE 1 Normalized structure factor values $|E_h|$ vs. $d^*_{00\ell}$ for bilayer lipid model structures successfully analyzed by direct methods. The designation "●" signifies that the reflection has phase $\psi_h = 0$, whereas "□" denotes $\psi_h = \pi$. Reflections joined by a line (—) are those correctly related by a Σ_2-triple through the origin defining phase $\psi_{001} = 0$. Those reflections assigned phase values by Σ_1-triples are also indicated. (*a*) DMPC, (*b*) DMPG, (*c*) $DLPEM_2$, (*d*) GDLT, (*e*) GDL, (*f*) GDP.

$\psi_{0010} = \psi_{0014} = \psi_{0016} = \pi$. In all, 12 of 16 reflections are assigned phases and the initial map is shown in Fig. 2 *b*.

(c) $DLPEM_2$. As found in the crystal structure (Pascher and Sundell, 1986), the headgroup conformation of this *racemic N,N*-dimethylphosphatidylethanolamine is somewhat unusual, differing from the one determined for the bilayer structure of a chiral homologue (Dorset and Zhang, 1990). The estimates of five Σ_2-quartets with $A_2 \geq 0.06$ are used to find relationships between the lamellar reflections, as shown in Fig. 1 *c*. A positive $A_1 = 0.11$ value indicates that $\psi_{002} = 0$ and three additional negative $A_1 \leq -0.02$ determine from the Σ_1-triples that $\psi_{004} = \psi_{008} = \psi_{0010} = \pi$. Thus, phase values are found for 9 of 13 reflections to produce the map in Fig. 2 *c*.

(d) GDLT. The 3-tosylate of 1,2-dilaurin has a glycerol conformation like that of $DLPEM_2$ (Pascher and Sundell, 1986). As outlined in Fig. 1 *d*, direct phasing can be used to find values for all reflections. There are 10 Σ_2-triples, where $A_2 \geq 0.02$, which are correct. A positive A_1 value determines that $\psi_{002} = 0$ in addition to $\psi_{001} = 0$ which is used for origin definition. Four additional negative A_1 values indicate that $\psi_{006} = \psi_{008} =$

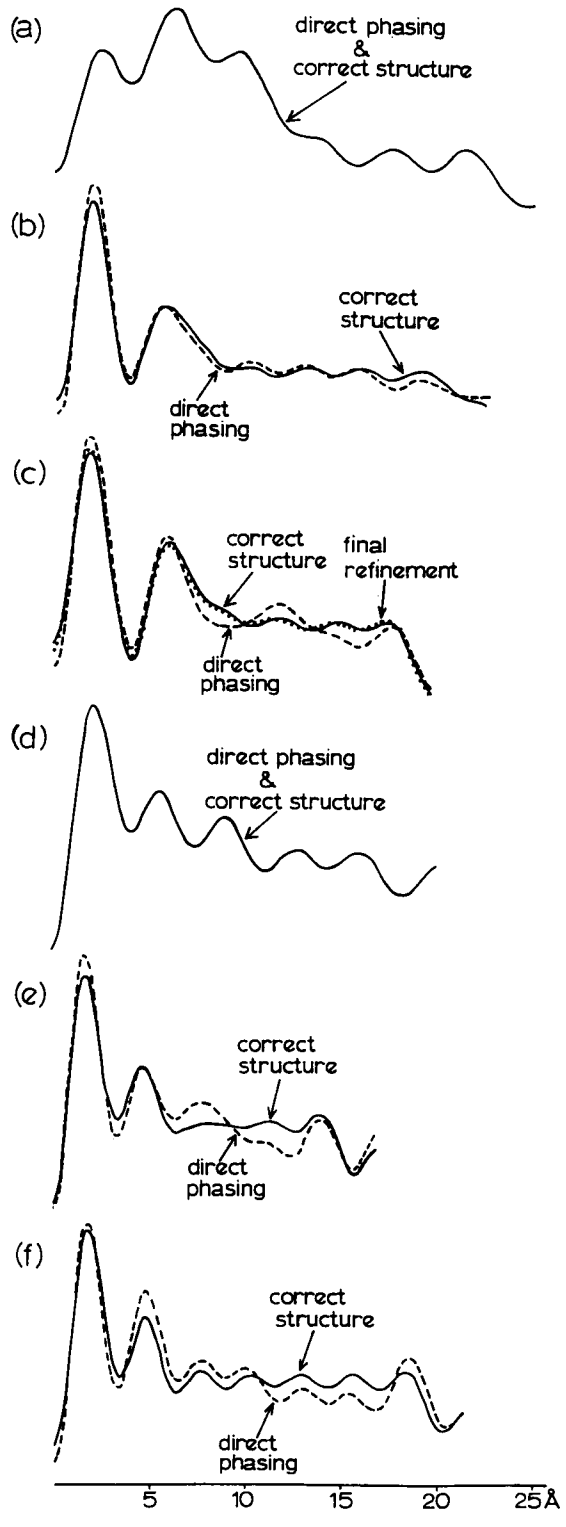

FIGURE 2 Electron density maps for lipids analyzed as in Fig. 1. The density profile is compared with that of the correct structure, which often is reached by phase refinement. When phase errors remain after the refinement, the final profile is also shown. (*a*) DMPC, (*b*) DMPG, (*c*) $DLPEM_2$, (*d*) GDLT, (*e*) GDL, (*f*) GDP. Only one-half of the bilayer is represented.

$\psi_{0010} = \psi_{0012} = \pi$. Thus, the phases that are found for the 13 reflections produce the map in Fig. 2 *d*.

(e) 1,2-diglycerides. Despite the chain length differences, crystal structures of two chiral 1,2-diglycerides demonstrate that the layer packings are strictly homologous (Pascher et al., 1987; Dorset and Pangborn, 1988). For the shorter chain compound, GDL, relationships between reflections can be found from seven Σ_2-triples where $A_2 \geq 0.14$. A positive A_1 value is used to signify that $\psi_{002} = 0$ and three negative values indicate that $\psi_{006} = \psi_{008} = \psi_{0012} = \pi$. As outlined in Fig. 1 *e*, phases are found for 10 of 13 reflections to produce the one-dimensional electron density map in Fig. 2 *e*.

The structure analysis of GDP is no less successful. Relationships between reflections are found from seven Σ_2-triples with $A_2 \geq 0.16$ and additional assignments are made from one positive and six negative A_1 values as outlined in Fig. 1 *f* to give phase values for 13 of 16 reflections. The structure map is shown in Fig. 2 *f*.

2. Problematic structure analyses

(a) DLPE-Ac. Although direct phase determination is quite successful for experimental x-ray and electron diffraction data from phosphatidylethanolamines (Elder et al., 1977), the analysis of the acetic acid solvate is not so straightforward. The problem is not due to the Σ_2-triples. It is found that only one error exists for 10 of these where $A_2 \geq 0.02$. The phase relationships for $A_2 \geq 0.25$ are depicted in Fig. 3 *a*. However, the high angle estimates of Σ_1-triples are incorrect, even though the values $\psi_{002} = 0$ and $\psi_{004} = \psi_{008} = \pi$ are correctly predicted. Only if $\psi_{0012} = \psi_{0014} = 0$ are accepted (hence the inverse of the expected value), can the map in Fig. 4 *a* be calculated.

(b) DMPA. The problem with this structure determination is very similar to the previous one. Twelve Σ_2-triples can be accepted where $A_2 \geq 0.02$ to link up the lamellar reflections as shown in Fig. 3 *b*. Although a larger positive Σ_1-triple where $A_1 = 0.70$ correctly determines that $\psi_{002} = 0$, four other values erroneously predict that higher angle phases have the value π. Only when these are reset to $\psi = 0$ can the map in Fig. 4 *b* be calculated.

(c) CER. As before, there are seven correct Σ_2-triples where $A_2 \geq 0.18$ which can be used to link reflections as shown in Fig. 3 *c*. For the Σ_1-triple, correct estimates are made for $\psi_{002} = 0$ and $\psi_{004} = \pi$ at relatively low angle, but a value for ψ_{0014} is misassigned. On the other hand, the value for $\psi_{0018} = \pi$ is correct. If the ψ_{0014} is reset to 0 to give phases for 12 of 18 reflections the map of Fig. 4 *c* can be computed.

(d) MCER. The problem with this structure analysis is in the prediction of Σ_2-triple phase sums. For example, of seven invariants with $A_2 \geq 0.14$, there are two errors,

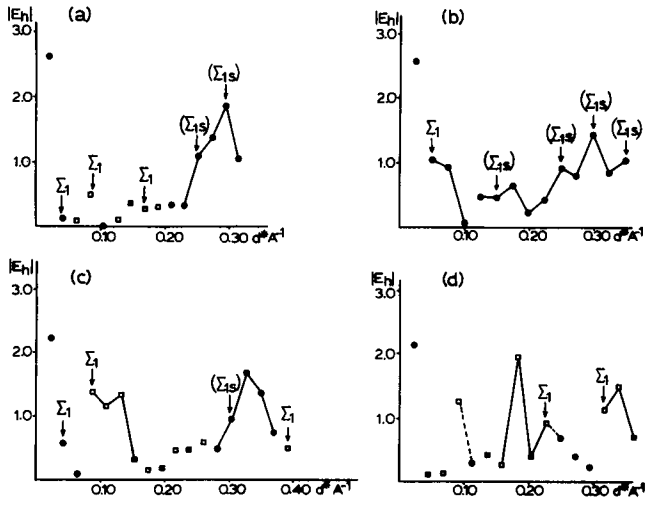

FIGURE 3 Normalized structure factors $|E_h|$ vs. $d^*_{00\ell}$ for difficultly analyzed bilayer structure. Symbols are as in Fig. 1 except that "Σ_1S" denotes and signifies a phase assignment where the Σ_1-triple value needs to be shifted by π. (*a*) DLPE · Ac, (*b*) DMPA, (*c*) CER, (*d*) MCER (for this latter example, two incorrect Σ_2-triple relationships are indicated by a dashed line [---]).

as indicated in Fig. 3 *d*. The two most negative Σ_1-triples ($A_1 \leq -0.08$) are correct. Using the origin definition $\psi_{001} = 0$ and the connectivity indicated, phase values can be given for 9 of 16 reflections, one of which (ψ_{0011}) is incorrect. The initial map is shown in Fig. 4 *d*.

Phase refinement

Within the constraints discussed above, total phase assignments are given to two structures (DMPA and GDLT) by direct methods and no further discussion is necessary for these. A third structure, DMPC, has only one very weak reflection which is undefined, and assignment of either phase value to it does not change the appearance of the electron density map in Fig. 2 *a*.

As suggested in numerous studies, including some recent work on protein crystal structures (Zhang and Main, 1990), at least two approaches can be taken for phase refinement when the structure is not resolved to atomic positions. One of these is to use the Sayre (1952) equation, a convolution of phased structure factors

$$\sum_{\vec{p}} F(\vec{p})F(\vec{h} - \vec{p}) \simeq w(\vec{h})F(\vec{h}),$$

where the predominance of known values, hopefully, will define the few that remain unknown. (Here $F(\vec{n}) \equiv F_{\vec{n}}$, and $\vec{n} = \vec{p}, \vec{h}, \vec{h} - \vec{p}$ are Miller indices used in the summation. The $w(\vec{h})$ value is a scale factor.) This relationship was tested for all the partial determinations

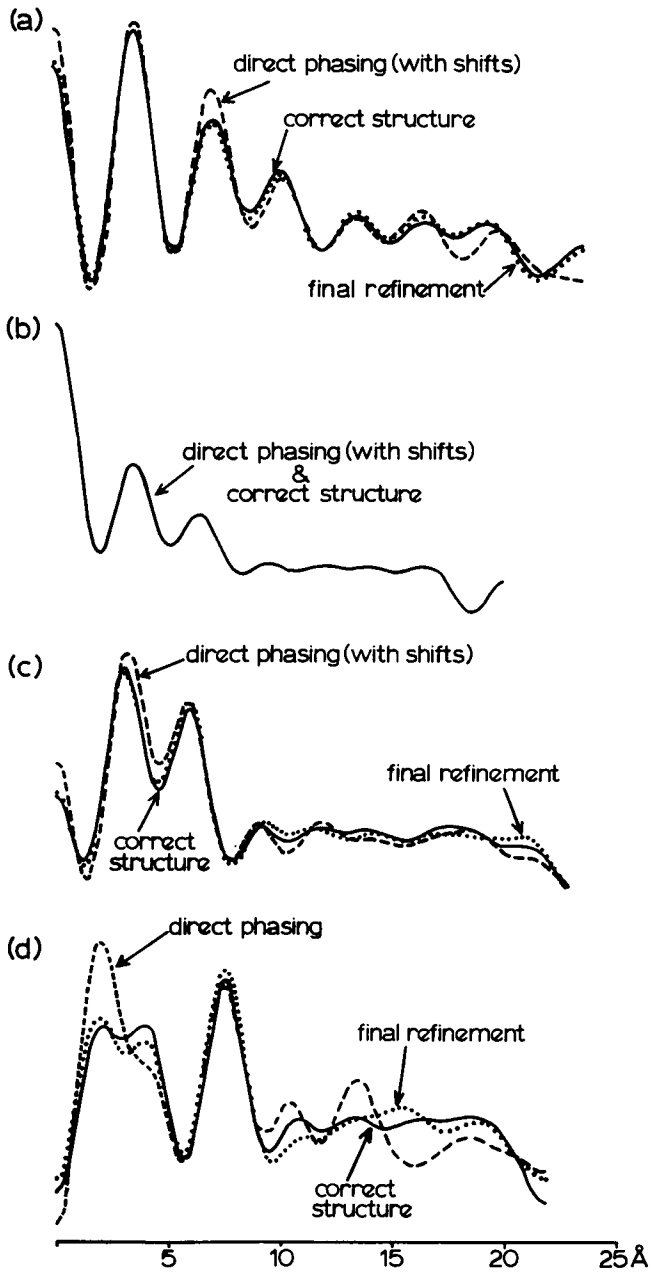

FIGURE 4 Electron density maps for lipids analyzed as in Fig. 3. Notations are made as in Fig. 2. (*a*) DLPE · Ac, (*b*) DMPA, (*c*) CER, (*d*) MCER.

indicated in Fig. 1 and 3, assuming that the starting phase set found by direct phasing can be accepted. In general, although some phase values are correctly determined, it is not a satisfactory procedure, undoubtedly due to the low resolution of these data. This presumption can be justified. For example, in the refinement of the $DLPEM_2$ structure, at 3.1 Å resolution, the Sayre equation predicts $\psi_{005} = \psi_{006} = \psi_{007} = \pi$ and attempts to change ψ_{004} to 0. Only one of these values is correct. If one extends the resolution of the original data set to 1.4 Å, then the determined sequence $\psi_{005} = 0$, $\psi_{006} = 0$, $\psi_{007} = \pi$ is correct.

The alternative approach is to use density modification, known as "solvent flattening" in protein crystallography (Wang, 1985). The region of these amphiphile structures, which can be most reliably modified in this way, is the hydrocarbon chain segments, e.g., as implied in the work of Worthington and Khare (1978) among others (e.g., Nagle and Wiener, 1989). Thus, in the analysis of these incompletely determined structures, the features of the headgroup are accepted as being correct but the rippling of the hydrocarbon portion is flattened to some mean value with some fall-off allowed near the chain ends (although without too much attention paid to the exact functional form). The Fourier transform is calculated as in Eq. 1 above, except that the usual scattering factor f_j is replaced by $\mathcal{F}(\delta(r) \cdot \rho(r))$ where $\rho(r)$ is the observed map density at locus r. When shifted to an origin to simulate a point scatterer the Fourier transform is uniform over all reciprocal space (see Champeney [1973]). The phase values are found in the signs of the resultant structure factors. These are then used to calculate a new map by

$$\rho(r) = \frac{1}{c} \sum_h F_{\vec{h}} \cos 2\pi(\vec{h} \cdot \vec{r}_j),$$

and the process is continued until the phase assignments are stable.

In most cases, hydrocarbon chain flattening is highly successful, as reviewed in Table 3. (Note that phases not listed are identical to those in Figs. 1 and 3.) Even when one or two phases remain with an incorrect phase value, the maps are very close to the ones calculated from the true phases (see Figs. 2 and 4), mainly because the associated structure factor magnitudes are weak. The most difficult analysis experienced so far was for MCER. After five refinement cycles based on the procedure outlined above, two phase errors remain, but the map is very close to that of the true structure (Fig. 4 *d*). Again, only the density of the chain packing was altered in this refinement.

As suggested by Luzzati et al. (1972), the progress of the phase refinement can be evaluated by use of a figure of merit which tests the flatness of the density distribution. It is said to be related to a maximum entropy value when this function is minimized (Luzzati et al., 1988). If the data are normalized (e.g., by setting $F_{000} = 0$) so that the mean density of the map $\bar{\rho} = 0$ and $\Delta\rho = \rho - \bar{\rho}$, then the value of $\langle \Delta\rho^4 \rangle$ serves this purpose. Progress of the structure refinements shown in Figs. 2 and 4 are assessed in Table 4 for the structures that are not completely

TABLE 3 **Results of phase refinement by density modification of hydrocarbon chains (only missing reflections shown)**

ℓ	refined	correct
DLPE · Ac		
3	0*	π
5	π*	0
6	π	π
7	π	π
10	0	0
DMPG		
4	π	π
5	0	0
6	0	0
7	0	0
$DLPEM_2$		
5	0*	0
6	π*	0
7	π	π
13	π*	0
GDP		
3	0	0
5	π*	π
7	π	π
9	π	π
GDL		
3	0	0
4	π	π
5	π	π
CER		
3	π‡	0
8	π	π
9	0‡	π
10	π	π
11	π	π
12	π	π
MCER		
2	π	π
3	π	π
4	π	π
5	0	0
6	0	π
11§	0	0
12	π	0
13	π	0

*Weak reflection. ‡Relatively weak. §Incorrect value in direct phase set.

phased by direct methods. The refinement always improves the map calculated from the initial incomplete phase set and, in general, even if the correct (model) structure is not reached, this should always have the lowest $\langle \Delta\rho^4 \rangle$. There is only one slight deviation from this rule in Table 4, i.e., a refined value is found for CER which is lower than the value calculated for the actual structure.

An alternative figure of merit can also be suggested.

TABLE 4 **Progress of structure analysis evaluated by the value of $\langle \Delta\rho^4 \rangle$ (values × 10^2)**

Compound	Direct phases	Real space refinement	Actual model structure
DMPG	2.40	1.61	1.61
$DLPEM_2$	3.12	2.34	2.22
DLPE · Ac	2.90	1.90	1.80
GDL	2.47	1.49	1.49
GDP	1.58	1.34	1.34
CER	1.49	0.94	0.99
MCER	1.90	0.81	0.72

Instead of testing the flatness of the density function, one could just as well evaluate its smoothness, assuming that the relatively most smooth profile should be the most likely. For this, one can calculate $s = \langle |\partial\rho/\partial r| \rangle$, where an average is taken over local differences in the magnitude of the density function. As shown in Table 5, this is just as successful as the method proposed by Luzzati. However, in the structure of MCER the refined phase set leads to a calculated value of s which is slightly smaller than the phase set of the actual structure.

DISCUSSION

In the initial evaluation of direct phase determinations of lamellar diffraction data from phospholipids (Dorset, 1990*b*), the reliable use of Σ_2 and Σ_1 triples was discussed. Combining these findings with the analyses carried out above, it is possible to make the following generalizations.

(*a*) Sigma 2 triples are more reliable for phase determination than are the Σ_1-triples. In 9 of 10 of the examples considered above, it was possible to set a threshold value for A_2, above which the prediction of all the Σ_2-triple phases was correct. For the remaining example, a nearly correct structure could be reached, despite two errors in the Σ_2-triples used for phase determination. Such behavior is consistent with the use

TABLE 5 **Progress of structure analysis evaluated by the value of $\langle |\partial\rho/\partial x| \rangle$ (values × 10^2)**

Compound	Direct phases	Real space refinement	Actual model structure
DMPG	3.39	3.26	3.26
$DLPEM_2$	4.15	3.80	3.75
DLPE · Ac	6.04	5.54	5.44
GDL	3.93	3.33	3.33
GDP	4.44	3.90	3.90
CER	3.97	3.50	3.46
MCER	5.09	3.81	3.83

of direct methods to determine molecular organic crystal structures where diffraction data are measured to better than atomic resolution.

(*b*) The use of sigma 1 triples, which must be done more cautiously than the evaluation of the Σ_2 phase sums, seems to be governed, in part, by the envelope of normalized $|E_h|$ in Figs. 1 and 3, reminiscent of the search for "intensity zeroes" when using bilayer swelling techniques to phase x-ray diffraction data (Burge and Fiddy, 1981). For example, within the first low angle intensity region, it is often possible to assign a phase $\psi_{002} = 0$, as indicated by a large positive A_1 value. Negative Σ_1-triples for higher angle reflections are often reliable, especially when the reflection indices can be demonstrated to be in an intensity region separated from the low angle reflections by a "node." In some cases, e.g., DLPE · Ac, CER, evidence can be found for three distinct intensity regions (Fig. 3) so that an empirical phase shift might be justified for negative Σ_1-triples in the highest angle region. In other cases, e.g., DMPA, this distinction is not so clear.

Within these constraints, the use of phase invariants as a technique for ab initio determination of unknown bilayer structures seems to be reliable. Obviously, there was no problem with six of the ten structures considered in this paper. This is encouraging, because the headgroup structures and/or conformations were not the more "typical" ones considered in our initial work and, hence, indicates that direct phasing may not be restricted to the solution of just one type of structural motif. From the electron density maps in Fig. 2, it will be noted that all of these structures have minimal density at the unit cell origin.

For the four structures that were more difficult to solve, three have density at the unit cell origin due to solvation or a counterion. As described, a somewhat subjective constraint could be placed on the structure analysis to accept structures where such features appear, e.g., in a multisolution approach where clusters of phases within intensity domains are given uniform phase signs. Without atomic resolution, however, the correctness of a model can be difficult to prove (Karle, 1989), since the use of a crystallographic R-factor as a figure of merit has been found to be a poor determinant in this case (Hamilton, 1964). An illustration of this point can be made with the DMPA structure. If an intensity zero is expected near (004) so that all lower angle reflections have a phase value 0 and higher angle reflection phase π (as would be expected from the Σ_1 assignments in our determination), then the electron density map obtained is found to be "reasonable" (Fig. 5 *b*) especially in the headgroup profile. The crystallographic residual for the experimental structure factor magnitudes used to calculate the map and those obtained from the reverse Fourier transform of the map is 4.5%. However, if all the phases are 0, as known from the crystal structure, then the map (Fig. 5 *a*) may not look "reasonable" (i.e., consistent with a preconceived notion of density distribution), even if it is correct! The reverse Fourier transform gives an R-factor of 7.5% with the model structure factors, i.e., the agreement is "worse" for the correct structure. Similarly, while the value of $\langle \Delta\rho^4 \rangle$ used above is sufficient for evaluating the progress of a refinement, this moment calculation is not a good figure of merit to distinguish between two near homometric structures. Thus, for DMPA, the correct structure yields $\langle \Delta\rho^4 \rangle = 3.10 \times 10^{-2}$, whereas a value 1.77×10^{-2} is found for the incorrect model. An evaluation of the density smoothness is much more successful here. For the correct model, one obtains a value $s = 2.74 \times 10^{-2}$ while $s = 2.77 \times 10^{-2}$ for the incorrect structure. Although these values are not greatly different, their similarity should indicate that the solvated structure might be considered as an alternative.

The test of density smoothness as a figure of merit thus appears to be another viable way of evaluating the phases in these diffraction patterns. As suggested above, clusters of phases related by the Σ_2 formula can be assigned an algebraic value $a = 0, \pi$. For example, in the cerebroside (CER) structure (see Fig. 3 *c*), the highest resolution data, which are falsely assigned a phase sign $\psi = \pi$ by Σ_1-triples, can be evaluated by two separate density maps calculated with an algebraic ambiguity.

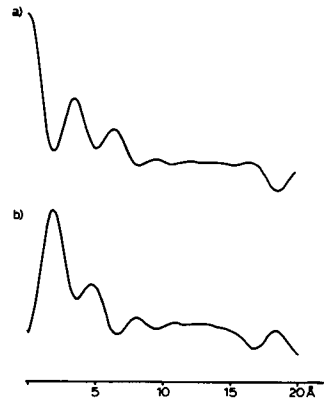

FIGURE 5 Comparison of correct vs. "reasonable" choice of phase values based on a presumed zero near (004) for the lipid DMPA (see Fig. 3 *b*). (*a*) The actual structure has significant density near the origin and is calculated from cell $\psi_h = 0$. (*b*) The incorrect structure, calculated from $\psi_h = 0$ for $h = 001, 002, 003, 004$, and $\psi_h = \pi$ for all other indices of 00ℓ, is very "reasonable" in appearance because the density profile resembles that of, e.g., GDP (Fig. 2*f*) after direct phase determination. Because both structures produce the same Patterson function, they are termed "homometric" (Hosemann and Bagachi, 1962).

After phase refinement, the correct solution is indicated by $s = 3.54 \times 10^{-2}$ ($a = 0$) versus a value $s = 3.95 \times 10^{-2}$ ($a = \pi$). The correct solution is also indicated in the initial phase set for DLPE · Ac when an ambiguity is used for the highest resolution phase cluster, i.e., $s = 6.04 \times 10^{-2}$ when $a = 0$ for the cluster of highest resolution data interrelated by Σ_2-triples whereas $s = 6.08 \times 10^{-2}$ when $a = \pi$. Because these values are very close, the alternative solution must be considered as another possibility. The appearance of density at the unit cell origin could be used then as an expected chemical feature of the structure to enable a choice to be made.

Problems encountered in the analysis of the methyl cerebroside (MCER) lamellar packing has nothing to do with the evaluation of Σ_1-triplet phase sums. MCER is merely a difficult structure to solve, as also encountered sometimes in atomic-resolution small molecule crystallography. The analysis summarized above seems to indicate a procedure that could be useful in such cases, i.e., to set a reasonably high threshold for A_2 values for acceptance of Σ_2-triples. Even with one or two phase errors, there may be enough of a start toward finding the correct structure in the initial phase set that a subsequent phase refinement will produce satisfactory results. However, unlike the previous examples where, generally, a single cycle is needed to find a phase set that will remain stable for repeated cycles, the MCER structure will only find a stable solution after five refinement cycles.

To summarize, it is clear that traditional direct phasing methods can be very useful for solving lipid bilayer structures. As demonstrated in this work, a variety of amphiphilic materials can be analyzed and, for the phospholipids, a number of headgroup classes or conformations can be successfully studied. While some modifications in the methods must be made in some cases, e.g., for structures where significant density will be found at the unit cell origin, it is also evident that high resolution electron microscopy could enable the structure analysis to be carried out without any subjective bias about which result should appear to be most "reasonable."

Received for publication 6 May 1991 and in final form 22 August 1991.

REFERENCES

Burge, R. E., and M. A. Fiddy. 1981. Phase assignment to diffraction patterns based on the analytic properties of scattered fields applied to the structure of nerve myelin. *J. Appl. Cryst.* 14:455–461.

Champeney, D. C. 1973. Fourier Transforms and Their Physical Applications. Academic Press, New York. p. 31.

Cowley, J. M., A. L. G. Rees, and J. A. Spink. 1951. Secondary scattering in electron diffraction. *Proc. Phys. Soc. (Lond.).* A64:609–619.

Dorset, D. L. 1987*a*. Molecular packing of a crystalline ether-linked phosphatidylcholine: an electron diffraction study. *Biochim. Biophys. Acta.* 898:121-128.

Dorset, D. L. 1988*a*. Two untilted lamellar packings for an ether-linked phosphatidyl-*N*-methylethanolamine. An electron crystallographic study. *Biochim. Biophys. Acta.* 938:279–292.

Dorset, D. L. 1988*b*. How different are the crystal structures of chiral and racemic diacylphosphatidylethanolamines? *Z. Naturforsch.* 43c: 319–327.

Dorset, D. L. 1990*a*. Concentration domain for the epitaxial orientation of phospholipids on naphthalene. *J. Phys. Chem.* 94:6854–6858.

Dorset, D. L. 1990*b*. Direct determination of crystallographic phases for diffraction data from phospholipid multilamellar arrays. *Biophys. J.* 58:1077–1087.

Dorset, D. L., and W. A. Pangborn. 1988. Polymorphic forms of 1,2-dipalmitoyl-*sn*-glycerol: a combined x-ray and electron diffraction study. *Chem. Phys. Lipids.* 48:19–28.

Dorset, D. L., and W. P. Zhang. 1990. Lamellar packing of a chiral N,N-dimethylphosphatidylethanolamine: electron diffraction. Evidence for a lecithin-type headgroup conformation. *Biochim. Biophys. Acta.* 1028:299–303.

Dorset, D. L., E. Beckmann, and F. Zemlin. 1990. Direct determination of phospholipid lamellar structure at 0.34nm resolution. *Proc. Natl. Acad. Sci. USA.* 87:7570–7573.

Dorset, D. L., A. K. Massalski, and J. R. Fryer. 1987. Interpretation of lamellar electron diffraction data from phospholipids. *Z. Naturforsch.* 42a:381–391.

Dorset, D. L., W. A. Pangborn, and A. J. Hancock. 1988. Epitaxial crystallization of alkane chain lipids for electron diffraction analysis. *J. Biochem. Biophys. Meth.* 8:29–40.

Doyle, P. A., and P. S. Turner. 1968. Relativistic Hartree-Fock x-ray and electron scattering factors. *Acta Crystallogr.* A24:390–397.

Elder, M., P. Hitchcock, R. Mason, and G. G. Shipley. 1977. A refinement analysis of the crystallography of the phospholipid 1,2-dilauroyl-*DL*-phosphatidylethanolamine and some remarks on lipid and lipid-protein interaction. *Proc. R. Soc. Lond. B. Biol. Sci.* 354:157–170.

Fryer, J. R., and D. L. Dorset. 1987. Direct imaging of paracrystalline phospholipid structure in the electron microscope. *J. Microsc. (Oxford).* 145:61–68.

Hamilton, W. C. 1964. Statistics in Physical Science. Estimation, Hypothesis Testing and Least Squares. Ronald Press Co., New York. 157–162.

Harlos, K., H. Eibl, I. Pascher, and S. Sundell. 1984. Conformation and packing properties of phosphatides and the crystal structure of monosodium dimyristoyl phosphatide. *Chem. Phys. Lipids.* 34:116–126.

Hauptman, H. A. 1972. Crystal Structure Determination. The Role of the Cosine Seminvariants. Plenum Publishing Corp., New York.

Hitchcock, P. B., R. Mason, and G. G. Shipley. 1975. Phospholipid arrangements in multilayers and artificial membranes: quantitative analysis of the x-ray diffraction data from a multilayer of 1,2-dimyristoyl-DL-phosphatidylethanolamine. *J. Mol. Biol.* 94:297–299.

Hosemann, R., and S. N. Bagachi. 1962. Direct Analysis of Diffraction by Matter. North-Holland Publishing Co., Amsterdam. p. 57.

Ibers, J. A., and W. C. Hamilton, editors 1974. International Tables for X-ray Crystallography, Vol. IV. Kynoch Press, Birmingham, England. pp. 72–98; 155–173.

Karle, J. 1989. Direct methods in protein crystallography. *Acta Cryst.* A45:765–781.

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.

Luzzati, V., P. Mariani, and H. Delacroix. 1988. X-ray crystallography at macromolecular resolution. A solution to the phase problem. *Makromol. Chem. Macromol. Symp.* 15:1–17.

Nagle, J. F., and M. C. Wiener. 1989. Relations for lipid bilayers. Connection of electron density profiles to other structural quantities. *Biophys. J.* 55:309–313.

Nyholm, P.-G., I. Pascher, and S. Sundell. 1990. The effect of hydrogen bonds on the conformation of glycosphingolipids. Methylated and unmethylated cerebroside studied by x-ray single crystal analysis and model calculation. *Chem. Phys. Lipids.* 52:1–10.

Pangborn, W. A. 1973. The crystal and molecular structures of the anhydride of 11-bromoundecanoic acid and racemic glycerol 1,2(di-11-bromoundecanoate)-3-(-p-toluenesulfonate). Ph.D. Thesis, University of Maryland, College Park.

Pascher, I., and S. Sundell. 1977. Molecular arrangement in sphingolipids. The crystal structure of cerebroside. *Chem. Phys. Lipids.* 20:175–191.

Pascher, I., and S. Sundell. 1986. Membrane lipids: preferred conformational states and their interplay. The crystal structure of dilauroyl phosphatidyl-N,N-dimethylethanolamine. *Biochim. Biophys. Acta.* 855:68–78.

Pascher, I., S. Sundell, and H. Hauser. 1981. Glycerol conformation and the molecular packing of membrane lipids. The crystal structure of 2,3-dilauroyl-D-glycerol. *J. Mol. Biol.* 153:791–806.

Pascher, I., S. Sundell, K. Harlos, and H. Eibl. 1987. Conformation and packing properties of membrane lipids. The crystal structure of sodium dimyristoyl phosphatidylglycerol. *Biochim. Biophys. Acta.* 896:77–88.

Pearson, R. H., and I. Pascher. 1979. The molecular structure of lecithin dihydrate. *Nature (Lond.).* 281:499–501.

Sayre, D. 1952. The squaring method: a new method for phase determination. *Acta Crystallogr.* 5:60–65.

Torbet, J., and M. H. F. Wilkins. 1976. X-ray diffraction studies of lecithin bilayers. *J. Theor. Biol.* 62:447–458.

Wang, B.-C. 1985. Resolution of phase ambiguity in macromolecular crystallography. *In* Methods of Enzymology. S. P. Colowick and N. O. Kaplan, editors. Academic Press, New York. 90–112.

Watts, P. H., Jr., W. A. Pangborn, and A. Hybl. 1972. X-ray structure of racemic glycerol 1,2-(di-11-bromoundecanoate)-3-(-p-toluenesulfonate). *Science (Wash. DC).* 175:60–61.

Worthington, C. K., and R. S. Khare. 1978. Structure determination of lipid bilayers. *Biophys. J.* 23:407–425.

Zhang, K. Y. J., and P. Main. 1990. The use of Sayre's equation with solvent flattening and histogram matching for phase extension and refinement of protein structures. *Acta Crystallogr.* A46:377–381.