Acta Cryst. (1998). D**54**, 615–621

Direct Phase Determination in Protein Electron Crystallography: Aquaporin Channel-Forming Integral Membrane Protein

Douglas L. Dorset[a]* and Bing K. Jap[b]

[a]*Electron Diffraction Department, Hauptman-Woodward Medical Research Institute, 73 High Street, Buffalo, NY 14203, USA, and* [b]*Life Sciences Division, Lawrence Berkeley Laboratory, University of California, Berkeley, CA 94720, USA. E-mail: dorset@hwi.buffalo.edu*

(*Received 27 August 1997; accepted 9 December 1997*)

Abstract

The location of helix sites in the projected structure of the aquaporin channel-forming integral membrane protein from bovine red blood cells was determined by multisolution direct methods to a mean accuracy of ±1.9 Å, based on *hk*0 electron diffraction data extending to 6 Å. The structure was assumed to be composed of pseudo-atoms, corresponding to the helix cross sections, and after re-scaling, normalized structure factors were used to order $\sum_n$ triples according to the *A* values. Initial phases were found by symbolic addition with algebraic unknowns. Probable solutions could be isolated by an overall Luzzati test for density flatness and restrictions on local density extremes. The best solution was identified by matching Patterson functions, generated from the trial map density sites, to the one calculated from observed intensities.

1. Introduction

As shown recently by the NATO Advanced Studies Institute (Fortier, 1998), direct methods for determination of crystallographic phases is at last being taken quite seriously in protein crystallography, despite numerous reservations given earlier (Karle, 1989). With new algorithms that circumvent the well known limitations of traditional probabilistic methods applied to large molecules, much of the work has concentrated on atomic resolution data (Weeks *et al.*, 1995). Currently, it appears that 1000-atom proteins might be solved by such techniques without the need for heavy-atom derivatives or utilization of an anomalous scattering signal.

Direct methods are not limited to atomic resolution data, however. One of the significant problems in any protein structure determination is the definition of the molecular envelope, a process requiring data collected to rather modest resolutions (Podjarny *et al.*, 1997). There have been cases where reasonably accurate crystallographic phases have been found but the resulting electron density could not be interpreted because the boundary between protein and solvent could not be defined. In X-ray crystallography, pioneering studies of direct phase determination with low-resolution diffraction data attempted to solve this problem (Podjarny *et al.*, 1981; Schevitz *et al.*, 1981) and this significant effort is still in progress (Lunin *et al.*, 1995; Urzhumtzev & Podjarny, 1995). The reason why traditional methods are useful with such data is because most of the scattering power is in this region of reciprocal space, hence the most probable phase invariants should be nearly valid, even if there are fewer of them than for a small-molecule data set (Fan *et al.*, 1991).

One further problem that limits the application of direct methods to low-resolution X-ray data is that a complete set of intensities is only rarely measured. For example, geometry of the diffraction experiment may occlude very low angle reflections by a beam stop. On the other hand, electron-diffraction patterns from thin microcrystals almost always include all of the available data in zones that can be accessed by tilting the specimen (Dorset, 1995*a*). For this reason, there has been some interest recently in exploring the potential of direct methods with such intensity sets (Gilmore *et al.*, 1993, 1996; Dorset *et al.*, 1995; Dorset, 1995*b*, 1996, 1997*a*), even though electron microscopy is acknowledged to be a powerful source of crystallographic phases not available to the X-ray crystallographer.

Initial applications of direct methods, initially maximum entropy and likelihood (Gilmore *et al.*, 1993), but later convolutional techniques (Dorset *et al.*, 1995) considered the accuracy of phase extension from a very low resolution phase set (*e.g.* derived from the Fourier transform of an averaged electron micrograph) to the limit of an electron-diffraction pattern. Useful results were reported for a variety of proteins. Later the concept of *ab initio* phase determination was considered, first assuming (somewhat incorrectly) that density flatness would be a rigorous figure of merit (Dorset, 1995*b*, 1996), but later relying on likelihood predictions (Gilmore *et al.*, 1996). The old concept of pseudo-atom 'globs' (Harker, 1953) as a model for the protein mass distribution was also considered. For proteins with density distributions (*e.g.* projections of helices) that could be treated as pseudo-atomic scatterers, there was shown to be an improved prospect for direct phase

determination, say to 6 Å resolution (Dorset, 1997*a*,*b*). However, for the three examples considered successfully, *i.e.* halorhodopsin, and two trigonal forms of bacteriorhodopsin, it could be said that the same structural motif was being considered repeatedly. Indeed, recently, the orthorhombic bacteriorhodopsin structure has also been analyzed successfully (Dorset, 1998).

To extend the direct approach to a completely different helical array, the *ab initio* phase analysis of the red blood cell aquaporin channel-forming integral membrane protein (AQP-CHIP) was attempted. The results are described in this communication.

2. Materials and methods

2.1. *Crystallization and electron diffraction*

The purification of bovine red blood cell AQP-CHIP and its reconstitution in dimyristoyl phosphatidylcholine bilayers has been described in detail in an earlier publication (Jap & Li, 1995). Electron diffraction patterns from untilted crystals extend to about 3.0 Å from preparations embedded in vitreous ice. The two-dimensional unit-cell constants are a square lattice with $a = b = 96.4$ (2) Å in the centrosymmetric plane group $p4mg$, a projection of the $P42_12$ space-group symmetry along [001]. Strongest reflections from 16 averaged diffraction patterns were observed to about 10 Å resolution or lower, even though the diffraction resolution extended to 3 Å. The phase determination described in this paper used all 106 reflections to the 6 Å limit, thus excluding those in the 6–3 Å range. With the phases provided from averaged low-dose electron micrographs the projected potential map at 6 Å resolution is depicted in Fig. 1.

2.2. *Intensity normalization*

In earlier studies of proteins composed mostly of α-helices (Dorset, 1997*a*,*b*), it was assumed that the dimensionality of the structure analysis could be reduced by a factor of 10. For example, the typical center-to-center distance for two helical cylinders is about 10–15 Å (Tatarinova & Vainshtein, 1962; Parsons & Martius, 1964), compared with the 1.54 Å length of a C—C single bond. If the helix profile were regarded as a Gaussian glob, then its Fourier transform, or scattering factor, might be well modeled by a carbon scattering factor (that has really more of a Lorentzian shape) if the unit cell were reduced to $a = b = 9.64$ Å. In simulations for halorhodopsin and bacteriorhodopsin, this was not a bad approximation (Dorset, 1997*a*,*b*, 1998).

Using the carbon electron scattering factor (Doyle & Turner, 1968) to model all projected helix transforms, a Wilson (1942) plot indicated that $B = -3.5$ Å^2. [Note, however, that this temperature factor is meaningless in its traditional sense. Here it is only used to optimize the fit of an approximate scattering factor envelope to the observed scattering fall-off. There is also nothing especially significant about the choice of scattering factor, except that it should be approximately Gaussian and convenient to model. Tests of other Gaussian or Lorentzian functions, also explored in earlier work (Dorset, 1997*b*), could marginally improve the R factor to the experimental structure-factor amplitudes but did not appreciably influence the accuracy of the phase model. This procedure, therefore, appears to be the best one for normalization of intensities when the glob approximation is made. See below for more discussion.] Normalized structure factors $|E_h|$ (Hauptman, 1972) were then calculated from I_{obs} for 106 unique data. To evaluate the glob model, a structure-factor calculation was carried out, based on helix positions identified from Fig. 1. With the carbon scattering factor model, the calculated phases in a structure-factor calculation deviated from the image-derived phases by a mean value of 49°, with $R = 0.63$ when $B = 0.0$ Å^2 was used. However, for the 25 most intense reflections, the mean phase deviation was only 14.4°. Hence, the model is better for predicting phases than amplitudes, an observation consistent with earlier glob simulations of rhodopsins (Dorset, 1997*a*,*b*).

2.3. *Phase determination*

Note (Fig. 1) that the unit-cell origin, adapted from the image analysis of Jap & Li (1995), is specified on the twofold axis of the plane group rather than the fourfold specified for the plane group. This implies the symmetry and origin definition for the actual space group $P42_12$ [also assumed in the earlier direct phase determination of halorhodopsin (Dorset, 1995*b*, 1997*a*)]. From simple

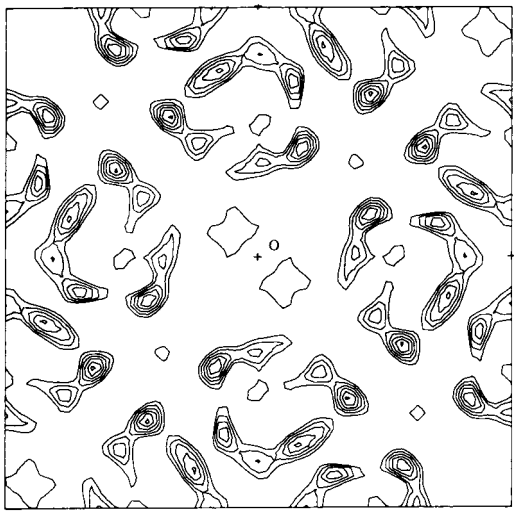

Fig. 1. Potential map for AQP-CHIP at 6 Å, calculated from electron diffraction $|F_h|$ and image-derived φ_h (Jap & Li, 1995).

trigonometric considerations, starting from the symmetry-reduced structure-factor expression, it can be shown that an effective origin-shift by $(x + \frac{1}{2})$ does not affect the phases of invariant reflections with indices $gg0$ or $uu0$ but will change the phases of those with index $ug0$ or $gu0$, the origin-sensitive reflections. A closely analogous situation of an origin shift from a fourfold to a twofold axis is treated formally by Hauptman (1972).

For the phase determination, it was assumed that there were seven 'atoms' in the asymmetric unit. From the 106 calculated $|E_h|$ values 859 $\sum_2$ triple invariants were calculated to $A_{\min} = (2/N^{1/2})|E_{h_1}E_{h_2}E_{h_3}| = 0.5$, as well as a small number of $\sum_1$ invariants. A convergence procedure (Germain *et al.*, 1970) established that, if $\varphi_{580} = 0$ defined the origin, then 85 phases could be accessed if (150) and (670) reflections were also assigned algebraic phase values. Symbolic addition (Karle & Karle, 1966) was employed for all phase determinations.

3. Results

Initially, it was found that the value $\varphi_{10,0,0} = 0$ also could be predicted accurately from a highly probable $\sum_1$ invariant. With the total basis set of four reflections, 21 clustered $\sum_2$-triple relationships, each contributing to a new phase term, were taken in the sequence established by the convergence procedure. However, unlike most small-molecule determinations, where an average is made over all contributors, weighted according to a figure of merit α_{est} (DeTitta *et al.*, 1975), only the first triple in the sequence (*i.e.* the most probable) was accepted for the phase assignment. During this determination the algebraic values, $\varphi_{150} = a$; $\varphi_{670} = b$ were

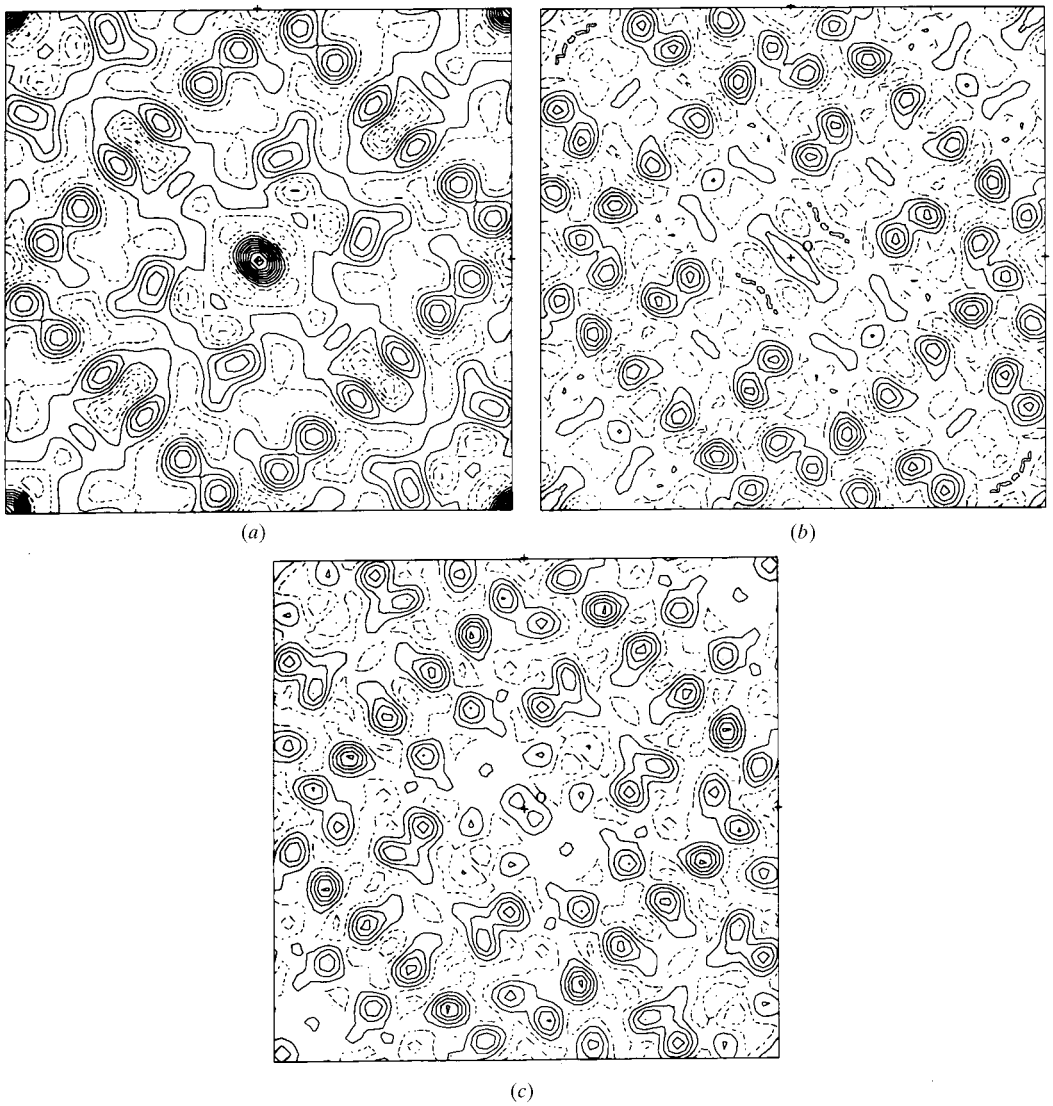

Fig. 2. Direct phasing and refinement of AQP-CHIP *via* convergence procedure. (*a*) Initial potential map; (*b*) map after first Fourier refinement cycle; (*c*) map after second Fourier refinement cycle.

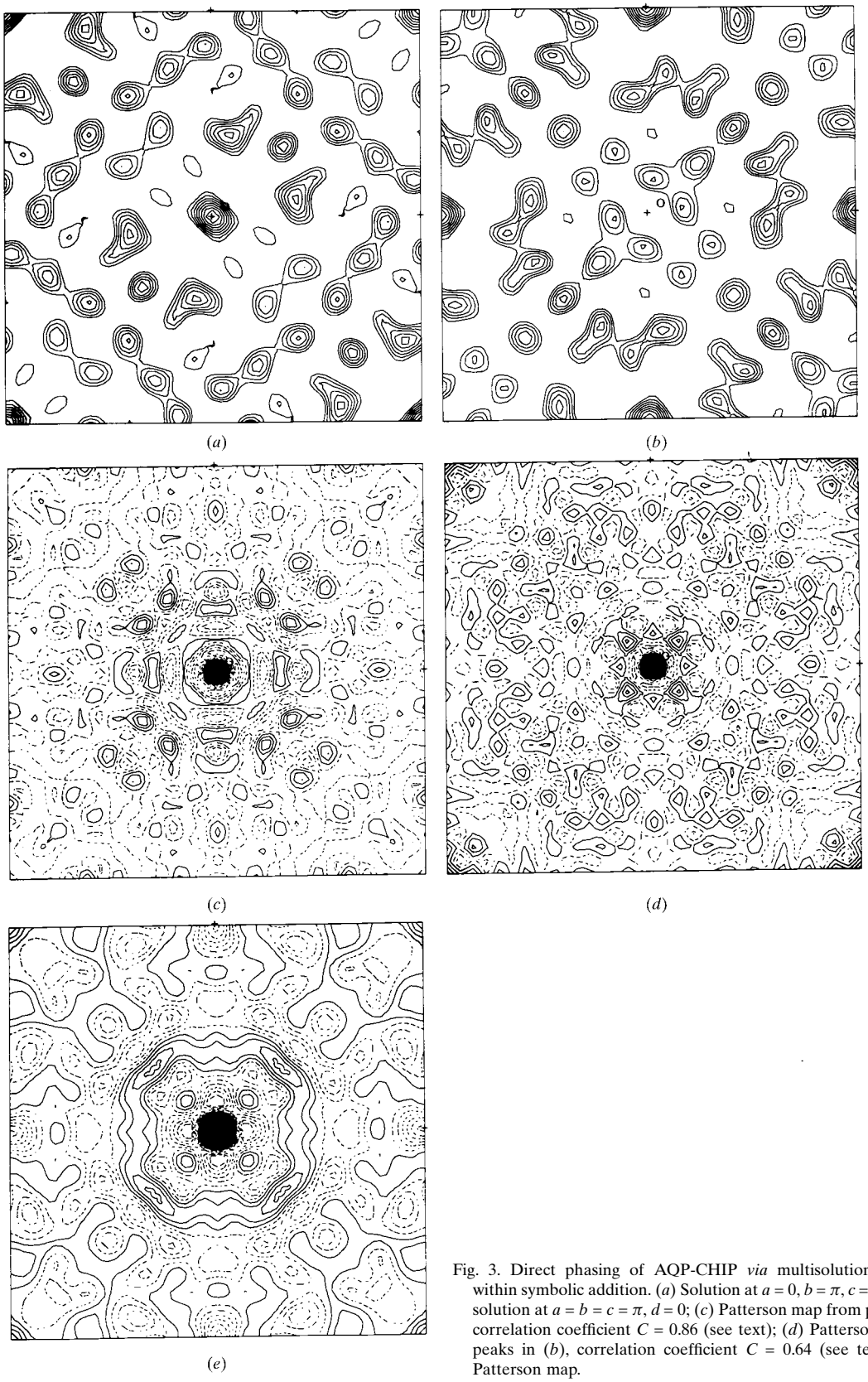

Fig. 3. Direct phasing of AQP-CHIP *via* multisolution procedure within symbolic addition. (*a*) Solution at $a = 0, b = \pi, c = 0, d = \pi$; (*b*) solution at $a = b = c = \pi, d = 0$; (*c*) Patterson map from peaks in (*b*), correlation coefficient $C = 0.86$ (see text); (*d*) Patterson map from peaks in (*b*), correlation coefficient $C = 0.64$ (see text); (*e*) I_{obs} Patterson map.

Table 1. *Final helical centers for AQP-CHIP*

Direct methods and refinement		Jap & Li (1995)	
x/a	*y/b*	*x/a*	*y/b*
0.272	−0.172	0.280	−0.175
0.357	−0.095	0.374	−0.076
0.413	−0.016	0.413	0.000
0.337	0.043	0.358	0.068
0.233	0.077	0.222	0.085
0.193	0.038	0.191	0.007
0.226	−0.108	0.228	−0.119

found to have the values 0 and π, respectively. In the list of 25 phases, obtained by direct methods, there were seven errors.

A potential map, calculated from these 25 phase terms (*i.e.* combined with the corresponding $|F_{obs}|$ amplitudes), is shown in Fig. 2(*a*). Some interpretation could be made, despite the fact that protein structures at this resolution have no conservative 'bonding pattern' of subunits expected for molecular crystals. For example, the large peak at the unit-cell origin, the twofold axis, is highly unlikely (Podjarny *et al.*, 1981) and can be rejected. Six likely peak locations were chosen from the other major density sites on the map and these were then used to calculate a first complete phase set (structure factors with pseudo-atom model). In the second map (Fig. 2*b*), the central peak is weakened and the chosen sites reinforced. At least two other density sites could be chosen as another helix positions in the asymmetric unit. However, it was assumed that actual protein density would not lie directly on a (projected) mirror plane so all peaks on this plane were also rejected. When seven helix sites are chosen for a structure-factor calculation, the peaks were reinforced in the ensuing map (Fig. 2*c*). From this model, the mean phase error for the top 25 reflections (to 6 Å resolution) was 50.4°.

If the strict analogy to small-molecule crystallography were to apply, the approach used above to arrive at an arrangement of helical sites can be viewed almost as an *a posteriori* contrivance, even though the result is somewhat like the true protein structure. Without knowing the structure beforehand, there is no real reason to select only the first triple contributor to any phase in the convergence list.

In an alternative, more objective, approach, if only the ten $\sum_2$ triples with highest A values were evaluated with six basis set phases, $\varphi_{10,0,0} = 0(\sum_1)$; $\varphi_{580} = 0$ (origin); $\varphi_{150} = a$; $\varphi_{670} = b$; $\varphi_{350} = c$; $\varphi_{550} = d$, then 15 phases could be found containing only two errors (with the correct values for the algebraic unknowns). Permutation of the unknowns gave $2^4 = 16$ solutions, from which an optimal set had to be chosen. Earlier, the Luzzati test for density flatness (Luzzati *et al.*, 1986) in the corresponding potential maps, *i.e.* $\langle \Delta\rho^4 \rangle$ is a minimum when $\bar{\rho} = 0$, was shown not to be a rigorous criterion for choosing the best solution (Dorset, 1996). Nevertheless, this FOM was employed to select a few most likely solutions. Within this subset, we were also looking for a solution that had the most 'atomistic' character so the two solutions with the greatest range of ρ_{max} *versus* ρ_{min} were chosen as likely candidates. For all 16 possible solutions, $\langle \Delta\rho^4 \rangle$ encompasses values from 0.3048 to 0.6064. (These values are not absolute ones, however, since the $|F_{obs}|$ are only relative.) Within the solutions where $\langle \Delta\rho^4 \rangle \leq 0.40$, assumed to be the most favorable case, the value of $(\rho_{max} - \rho_{min})$ lies between 0.37 and 0.49; two solutions with the latter value being tested. The corresponding potential maps are depicted in Figs. 3(*a*) and 3(*b*). From the peak positions (again ignoring those appearing on rotation axes or projected mirror planes), Patterson functions were calculated as depicted in Figs. 3(*c*) and 3(*d*). While neither Patterson map perfectly matched the one calculated from the complete set of I_{obs} (Fig. 3*e*), only the solution $0,\pi,0,\pi$ gave a satisfactory match to numerous peak positions. A linear correlation coefficient, calculated from $C = \sum_h m_o m_c / \sum_h |m_o||m_c|$, where $m_o = |F_o|^2 - \langle |F_o|^2 \rangle$ and $m_c = |F_c|^2 - \langle |F_c|^2 \rangle$ was also effective for selecting the correct solution. (Only the values for the 25 most intense observed reflections were compared in this calculation.) This FOM is slightly different to the one given by Drenth (1994). Again ignoring peaks on mirrors or rotation axes, seven positions could be found that corresponded closely to the helix sites identified by Jap & Li (1995) (Fig. 3*a*). After three cycles of Fourier refinement, these sites were reinforced (Fig. 4), and a model could be proposed for helix locations that were, on average, only 1.9 Å away from the model based on image phases (Table 1). For the top 25 reflections the mean phase difference was only 29° (Table 2).

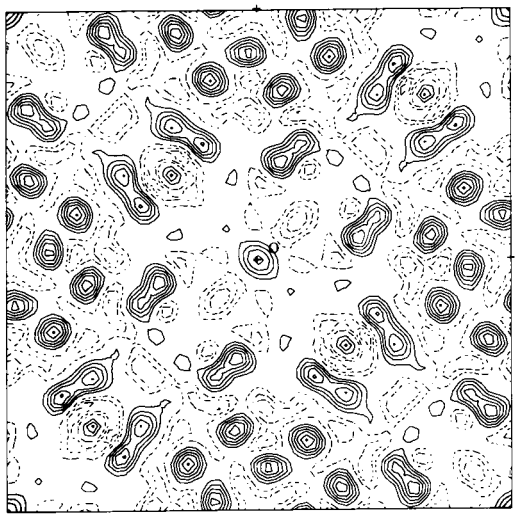

Fig. 4. Solution in Fig. 3(*a*) after three cycles of Fourier refinement.

Table 2. *Refined phase values or aquaporin after direct determination compared with values determined from the Fourier transform of electron micrographs*

$h\ k\ 0$	$\|F_{obs}\|$	φ (direct methods)	φ (image transform)
0 6 0	55.63	π	π
0 10 0	40.65	π	0
1 5 0	99.63	0	0
1 6 0	52.73	π	π
2 5 0	45.48	0	0
2 11 0	38.33	π	π
3 4 0	72.18	0	0
3 5 0	70.23	0	0
3 6 0	40.90	0	0
3 9 0	57.94	0	π
4 4 0	47.09	π	π
4 5 0	40.49	π	π
4 6 0	37.65	0	0
4 8 0	36.08	π	π
4 10 0	46.23	0	0
5 5 0	38.30	π	π
5 7 0	77.99	0	0
5 8 0	77.93	0	0
5 10 0	40.01	0	π
6 6 6	43.63	π	π
6 7 0	68.43	π	π
6 8 0	62.39	π	π
7 7 0	64.58	0	0
8 9 0	35.65	0	π
9 9 0	51.55	0	0

4. Discussion

As discussed in a recent paper (Dorset, 1997*b*) where similar techniques were used to determine the structures of two bacteriorhodopsin crystal forms, the approximation of projected globular density by a pseudoatom model is more successful in some cases than in others. Other phenomenological scattering-factor models (Dorset, 1997*b*), *e.g.* pure Gaussian or Lorentzian functions, could produce a better fit to the intensity transform, even if the phase accuracy was not so much affected by the exact form of the scattering model. In a test of various models at idealized helix sites, it was found that the Gaussian function, $\exp(-2.5\,s^2)$, was better than the carbon scattering factor used above (R = 0.54). The overall phase accuracy was only 44°, not greatly different from the 49° value obtained before. Choice of a slightly better phenomenological scattering factor, however, does not necessarily improve the direct phase determination. If the Gaussian scattering factor was used to calculate $|E_h|$ and these used to order $\sum_2$ triples according to the decreasing order of A values, then the the phase determination was degraded by five errors in the ten top triples (instead of just two, following the procedure given above), including that with the highest A value. Using the convergence routine (Germain *et al.*, 1970) to order the sequence of phases to be picked up by direct methods, as before, 16 phase values contained six errors but the resulting potential map was, in fact, not quite so easily interpreted as the ones shown above, even though density is found to lie near the optimal helix positions. A Fourier refinement was not attempted from this starting point.

We found, therefore, that the most accurate determination utilized the carbon electron scattering factor as the model for the globular Fourier transform. For example, the mean value of the $\sum_2$ invariants with the ten largest A values was $36 \pm 76°$ (where 0° is predicted), an accuracy similar to the one found for a $p3$ form of bacteriorhodopsin (Dorset, 1997*b*). Thus, even though variations of the structure-factor model may have little effect on simulating the protein phases in a structure-factor calculation, these variations can affect the normalization of intensities for ranking of phase invariants.

Could the unknown aquaporin structure have been solved directly from the observed 6 Å resolution diffraction data? Obviously, the solution would not have been found as easily as in the earlier determinations of the two centrosymmetric rhodopsin projections (Dorset, 1997*a*, 1998). In these earlier studies, where the projected helix densities were nearly isotropic, a strict analogy to small-molecule determinations was observed, after the data were normalized by pseudo-atom scattering factors.

From the recently published three-dimensional structure of AQP-CHIP (Li *et al.*, 1997), it is evident that tilt of at least half of the α-helices in the molecule account for the strongly elliptical shapes of the peaks in Fig. 1. Comparing the results of Fig. 4 to the map in Fig. 1 (also see Table 1), an anisotropic density distribution will not be found by direct methods when an isotropic model is used for the phenomenological scattering factor. Nevertheless, even though the anisotropy of globular density affects the accuracy of the phase invariants when an isotropic scattering envelope is assumed, a multisolution approach can still find a phase set for the most intense reflections that is reasonably accurate. Again the top ten $\sum_2$ invariants contain only two false estimates. The Harker (1953) pseudo-atom approach to determining the structures of appropriate membrane protein projections has considerable merit, therefore, even for such examples. The remaining problem in the direct analysis is finding an optimal figure of merit to select the best solution among many, as experienced earlier in the analysis of noncentrosymmetric data sets (Dorset, 1997*b*). If this problem could be solved decisively, then the answer to the above question would be affirmative.

We are only at a threshold of understanding how to find a correct structure from a multisolution set when data are limited to low resolution. Obviously, the same constraints used in interpretation of trial potential maps in small-molecule crystallography cannot be made in protein applications. There is no conservative entity to helix (or other globular density) interaction analogous to the well known bonding constraints of covalent molecules so that the chemical knowledge actually used

to solve many small molecules by direct methods does not apply. Nevertheless, there are other criteria that might be effective. If the structure of analogous proteins is known, histogram matching (Zhang & Main, 1990) might be useful although it has not been considered so far in our work. This option is mentioned since glob-generation in real space has referred to this FOM for accepting possible solutions (Lunin *et al.*, 1990). Maximum likelihood (Bricogne & Gilmore, 1990) is another constraint to consider, *e.g.* already shown by Gilmore & Nicholson (1997) to be effective for producing a reasonable replica of this structure at 10 Å resolution. In keeping with the atomistic idea, we have preferred to impose the reasonable restrictions that density cannot occur or dyad axes or on projected symmetry planes (Podjarny *et al.*, 1981), due to the chiral nature of the protein itself. In addition, the matching of observed asnd calculated Patterson functions, also employed in earlier work (Dorset, 1997*b*), may be an effective constraint for selecting among a small range of possible initial phase choices, especially when correlation coefficients are calculated.

Given the success of real-space 'glob' models for solving two-dimensional protein crystal structures at low resolution (Podjarny *et al.*, 1997), it would be interesting to evaluate, in future studies, if these reciprocal-space methods can be extended to three dimensions. Also, suitable methods for refining the initial structures found by this approximation must be be determined in future work.

Research was supported by a grants from the National Institute of General Medical Sciences, GM-46733 to DLD and GM-51487 to BKJ, which are gratefully acknowledged.

References

Bricogne, G. & Gilmore, C. J. (1990). *Acta Cryst.* A**46**, 284–297.
DeTitta, G. T., Edmonds, J. W., Langs, D. A. & Hauptman, H. (1975). *Acta Cryst.* A**31**, 472–479.
Dorset, D. L. (1995*a*). *Structural Electron Crystallography*. New York: Plenum Press.
Dorset, D. L. (1995*b*). *Proc. Natl Acad. Sci. USA*, **92**, 10074–10078.
Dorset, D. L. (1996). *Acta Cryst.* A**52**, 480–489.
Dorset, D. L. (1997*a*). *Proc. Natl Acad. Sci. USA*, **94**, 1791–1794.
Dorset, D. L. (1997*b*). *Acta Cryst.* A**53**, 445–455.
Dorset, D. L. (1998). *Acta Cryst.* A**54**. In the press.
Dorset, D. L., Kopp, S., Fryer, J. R. & Tivol, W. F. (1995). *Ultramicroscopy*, **57**, 59–89.
Doyle, P. A. & Turner, P. S. (1968). *Acta Cryst.* **24**, 390–397.
Drenth, J. (1994). *Principles of Protein X-ray Crystallography*, p. 229. New York: Springer-Verlag.
Fan, H. F., Hao, Q. & Woolfson, M. M. (1991). *Z. Kristallogr.* **197**, 196–208.
Fortier, S. (1998). Editor. *Direct Methods for Solving Macromolecular Structures*. Dordrecht: Kluwer Academic Publishers.
Germain, G., Main, P. & Woolfson, M. M. (1970). *Acta Cryst.* B**26**, 274–285.
Gilmore, C. J. & Nicholson, W. V. (1997). *Direct Methods for Solving Macromolecular Structures*, edited by S. Fortier, Dordrecht: Kluwer Academic Publishers.
Gilmore, C. J., Nicholson, W. V. & Dorset, D. L. (1996). *Acta Cryst.* A**52**, 937–946.
Gilmore, C. J., Shankland, K & Fryer, J. R. (1993). *Ultramicroscopy*, **49**, 132–146.
Harker, D. (1953). *Acta Cryst.* **6**, 731–736.
Hauptman, H. A. (1972). *Crystal Structure Determination. The Role of the Cosine Seminvariants.* New York: Plenum Press.
Jap, B. K. & Li, H. L. (1995). *J. Mol. Biol.* **251**, 413–420.
Karle, J. (1989). *Acta Cryst.* A**45**, 765–781.
Karle, J. & Karle, I. L. (1966). *Acta Cryst.* **21**, 849–859.
Li, H, L., Lee, S. & Jap, B. K. (1997). *Nature Struct. Biol.* **4**, 263–265.
Lunin, V. Yu., Lunina, N. L., Petrova, T. E., Vernosolva, E. A., Urzhumtsev, A. G. & Podjarny, A. D. (1995). *Acta Cryst.* D**51**, 896–903.
Lunin, V. Yu., Urzhumtzev, A. G. & Skovoroda, T. P. (1990). *Acta Cryst.* A**46**, 540–544.
Luzzati, V., Mariani, P. & Delacroix, H. (1986). *Makromol. Chem. Macromol. Symp.* **15**, 1–17.
Parsons, D. F. & Martius, U. (1964). *J. Mol. Biol.* **10**, 530–533.
Podjarny, A. D., Schevitz, R. W. & Sigler, P. B. (1981). *Acta Cryst.* A**37**, 662-668.
Podjarny, A., Urzhumtsev, A. & Lunin, V. (1997). *Direct Methods for Solving Macromolecular Structures*, edited by S. Fortier. Dordrecht: Kluwer Academic Publishers.
Schevitz, R. W., Podjarny, A. D., Zwick, M., Hughes, J. J. & Sigler, P. B. (1981). *Acta Cryst.* A**37**, 669–677.
Tatarinova, L. I. & Vainshtein, B. K. (1962). *Vysokomolek. Soed.* **4**, 261–269.
Urzhumtsev, A. & Podjarny, A. (1995). *Acta Cryst.* D**51**, 888–895.
Weeks, C. M., Hauptman, H. A., Smith, G. D., Blessing, R. H., Teeter, M. M. & Miller, R. (1995). *Acta Cryst.* D**51**, 33–38.
Wilson, A. J. C. (1942). *Nature (London)*, **150**, 151–152.
Zhang, K. Y. J. & Main, P. (1990). *Acta Cryst.* A**46**, 41–46.