Acta Cryst. (1996). D**52**, 1174–1180

Phasing Proteins at Low Resolution

KLAS M. ANDERSSON* AND SVEN HOVMÖLLER

Structural Chemistry, Arrhenius Laboratory, University of Stockholm, S-106 91 Stockholm, Sweden. E-mail: klasa@struc.su.se

(*Received 19 April 1996; accepted 29 July 1996*)

Abstract

A method for obtaining phases of low-order reflections is presented. It is based on four observations: (1) the electron density inside proteins is smooth and uniform at low resolution. (2) Since all proteins have almost the same density, the total volume of the protein is known if the molecular weight is known. (3) The overall shape of many proteins is fairly spherical. (4) The total scattering from a sphere of uniform density is in phase with a point scatterer at its centre of gravity, up to a well defined cross-over. After the first cross-over the total protein molecule scatters out of phase with its centre. If the centre of the protein can be found, the phases of typically the ten lowest resolution reflections can be very accurately determined. The method works, provided low-order reflections can be measured accurately and the centre of gravity can be well positioned from these data. The correctly phased low-resolution reflections may be used as a starting set for phase extension. By combining the measured amplitudes with these phases we believe that the size and low-resolution shape of an unknown protein, *i.e.* the envelope of the molecule, can be obtained.

1. Introduction

The phase problem is still a rate-limiting step in many macromolecular structure determinations by X-ray crystallography. The most widespread phasing techniques used in protein crystallography today are multiple isomorphous replacement (MIR) (Blow & Crick, 1959) and multiwavelength anomalous dispersion (MAD) (Hendrickson, 1991; Smith, 1991). These methods often suffer from large phase errors, particularly at high resolution and are also time consuming. Density modification is a method that reduces these phase errors and extends the phases to higher resolution (Podjarny, Bhat & Zwick, 1987). Sjölin & Svensson (1993) have used solvent flattening (Wang, 1985) in combination with maximum entropy in order to improve the performance of density modification. Main and co-workers used a combination of histogram matching, Sayre's equation (Sayre, 1952) and density modification (Main, 1990; Cowtan & Main, 1993; Zhang & Main, 1990*a*,*b*) to obtain better phases. A number of interesting developments in direct methods and maximum-entropy calculations have shown promising results when applied to small proteins at atomic resolution; see for example Gilmore (1992); Bricogne (1993) and the 'shake and bake' method (Weeks, Hauptman, Smith & Blessing, 1995). Despite recent advances in direct methods and maximum-entropy calculations, these techniques do not yet present a general solution of the phase problem and the challenging task of solving the phase problem for macromolecules still remains.

In most structure determinations of macromolecules not much attention is being paid to the very low resolution reflections. Indeed, quite often data before 15 or 20 Å is not even measured. However, during the past few years the interest in low-resolution data for *ab initio* phasing has grown (Lunin *et al.*, 1995; Urzhumstev & Podjarny, 1995). The phases of the lowest order reflections are dominated by the molecule as a whole and thus carry the information of the shape and size of the molecule (Kraut, 1958; Teeter & Hendrickson, 1979).

Subbiah (1991) has tried to find the shape of the solvent, using a method where a simulated gas of hard point scatterers are moved around in the unit cell until they condense in a cluster under the constraint of the solvent fraction and the restraint of the observed Fourier amplitude data, $|F_o|$. The condensed cluster defines the solvent and thus inversely also defines the envelope of the molecule. A similar method has been developed by Harris (1995).

Several of the previous methods exploit the fact that the solvent regions outside the protein molecule have a uniform density. We now try to make use of the fact that the density inside the protein is also quite uniform at low resolution. If it is justified to approximate the protein shape as a sphere, we only need to determine three unknowns, the x, y, z coordinates of the centre of gravity of the protein, from the ten or so lowest reflections, to be able to estimate the phases at low resolution. This problem should be over-determined.

The method relies on the assumption that the centre of gravity of a protein sphere will scatter in phase with the whole protein to a certain d value (see Fig. 1).

We believe that this method can provide correct phase information for low-resolution reflections of

unknown protein structures provided these reflection intensities, which are normally discarded in protein crystallographic work, are accurately measured.

2. Methodology

Structure factors and phases were calculated in two ways; all the non-H atoms were given (*a*) their respective scattering factor of C, N, O atoms *etc.* and (*b*) an equal scattering factor, namely that of carbon. The amplitude difference between these calculations was less than 5%. In fact, the bigger the globular proteins are, the more uniformly distributed are the atoms which results in a smaller difference in amplitudes, $R = \sum |F_c^a - F_c^b| / \sum F_c^a$, where the superscripts a and b are the protein and the carbon-approximated protein, respectively.

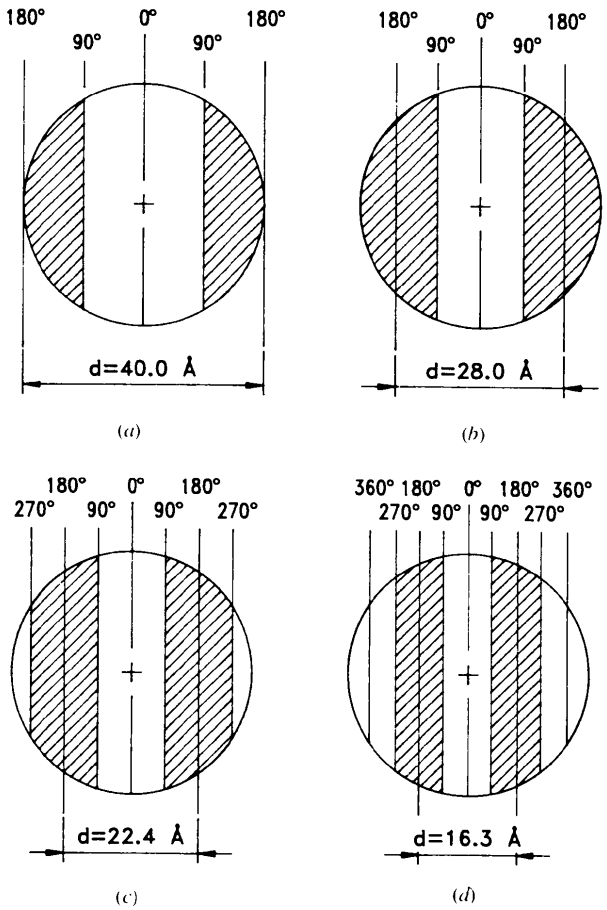

Fig. 1. The total molecular scattering from a sphere at different d values. Regions of atoms scattering in phase with the center of gravity of the protein (+) are shown in white. Regions of atoms scattering 180° out of phase with the centre of gravity are hatched. Here a molecule with a radius of 20 Å is used. (*a*) The total scattering from all the atoms is in phase with the point at the cross, $d = 40.0$ Å. (*b*) The first cross-over of the G function occurs when the total scattering from the atoms near the centre (white segment) is equal to the total scattering from the atoms out of phase (within the hatched segments), resulting in zero scattering, d value of 28.0 Å. (*c*) At $d = 22.4$ Å the total scattering from all atoms will be 180° out of phase with the centre of gravity since the sum of hatched segments is larger than the sum of the white segments. (*d*) The total scattering of the atoms in the hatched segments again equals the white segments and the total scattering will be zero. This situation represents the second cross-over, at $d = 16.3$ Å.

The following five proteins have been used to test the method; chymotrypsinogen (Freer, Kraut, Robertus, Wright & Xuong, 1970), carbonic anhydrase (Eriksson, Jones & Liljas, 1988), cytochrome P-450 (Raag & Poulos, 1991), glucose oxidase (Hecht, Kalisz, Hendle, Schmid & Schomburg, 1993) and haemoglobin (T state) (Waller & Liddington, 1990). Coordinates, and other crystallographic data have been taken from the Protein Data Bank (PDB) (Bernstein *et al.*, 1977).

We have used the value of 20 Å^3 for the volume of each non-H atom (Andersson & Hovmöller, 1996; Gellatly & Finney, 1982). Once the number of atoms in the protein is known, we can directly calculate the radius r of the protein, assuming a mathematically perfect sphere. We obtain a radius of 19.7 Å for chymotrypsinogen using the number of atoms of the protein and the atomic volume of 20 Å^3, which can be compared to the work by Kraut (1958) who predicted a radius of 19.3 Å using hydrodynamic properties such as viscosity, diffusion constants and molecular weight. The relevant data for all the tested proteins are listed in Table 1.

The molecular scattering of a protein is illustrated in Fig. 1. The atoms lying on the 0° line will all have the same phase as the point at the centre of gravity of the protein. The centre of gravity of the protein is depicted with a cross in Fig. 1. The further an atom is from the 0° line the larger is the phase difference. At a certain distance from the 0° line, the scattering from the atoms will start to contribute negatively to the total scattering. For the about ten reflections with the largest d values, the total scattering of the sphere will be in phase with a point at the centre of gravity, Fig. 1(*a*). For any spherical protein the scattering, in phase and out of phase with the centre, will exactly cancel at a d value of $1.40 \times r$. Thus, the total scattering will be zero at this point, Fig. 1(*b*). This d value is the first cross-over, $d_{c\text{-}o1}$.

For reflections with a d value smaller than $d_{c\text{-}o1}$ the whole protein will scatter 180° out of phase with an atom at its centre as illustrated in Fig. 1(*c*).

A second cross-over, $d_{c\text{-}o2}$, is reached at a d value of $0.815 \times r$, Fig. 1(*d*). The scattering continues to oscillate in this way, but we do not expect a good predictive value from our spherical model much beyond the first cross-over.

This oscillating dependence is described by the average intensity distribution, $i(hr)$. For a compact sphere (Guiner & Fournet, 1955) the function describes how the scattering intensity varies with radius.

Table 1. *Crystal information of the five proteins used as test structures*

	Carbonic anhydrase	Chymotrypsinogen	Cytochrome P-450	Glucose oxidase	Haemoglobin (T state)
PDB code	4CAC	1CHG	8CPP	1GAL	1THB
No. of atoms	2040	1643	3209	4622	4560
Volume of molecule* (Å^3)	40800	32860	64180	92440	91200
Radius of sphere* (Å)	21.4	19.7	24.8	28.0	27.9
Molecular dimension ratios	10:11:12	10:12:10	11:12:7	6:10:10	12:14:14
d value at 1st cross-over (Å)	29.9	27.6	34.7	39.2	39.1
d value at 2nd cross-over (Å)	17.4	16.1	20.2	22.8	22.7
Space group	$P2_1$	$P2_12_12_1$	$P2_12_12_1$	$P3_121$	$P2_12_12$
Z	2	4	4	6	4
a (Å)	42.7	52.0	108.7	66.5	95.8
b (Å)	41.7	63.9	103.9	66.5	97.8
c (Å)	73.0	77.1	36.4	214.5	65.5
α (°)	90.0	90.0	90.0	90.0	90.0
β (°)	104.6	90.0	90.0	90.0	90.0
γ (°)	90.0	90.0	90.0	120.0	90.0
Unit-cell volume (Å^3)	125786	256188	410760	821488	613685
Protein content in unit cell (%)	64.9	51.3	62.5	67.5	59.4

* Calculated using a volume of 20 Å^3 per non-H atom.

$$\frac{\overline{F^2(hr)}}{V^2\rho^2} = i(hr) = \left[3\frac{\sin(hr) - hr\cos(hr)}{h^3r^3}\right]^2. \quad (1)$$

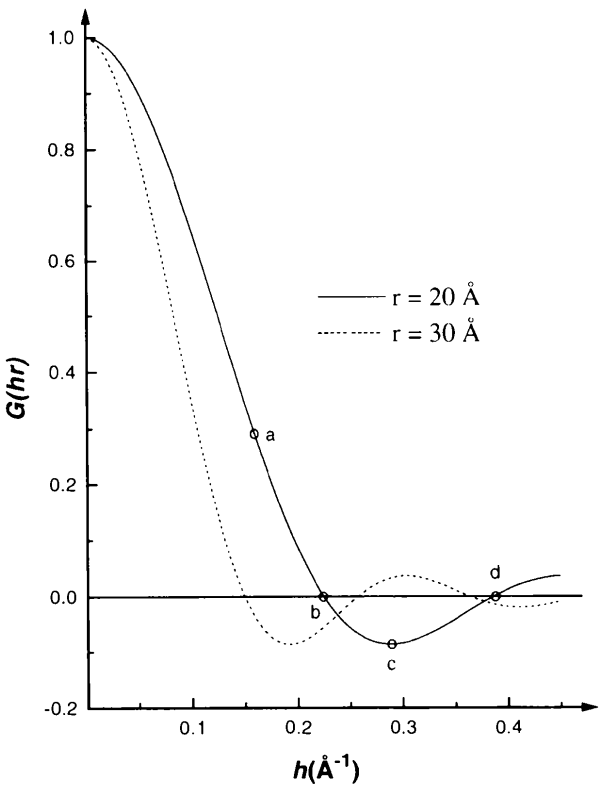

Fig. 2. The molecular scattering from perfect spheres of radii 20 and 30 Å. The total scattering from all atoms is in phase with a point at the centre of the protein out to the point marked *b*. The total scattering from all atoms will be 180° out of phase with the point at the centre of the protein between positions marked *b* and *d* in the graph. The four positions marked *a*, *b*, *c* and *d* correspond to Figs. 1(*a*), (*b*), (*c*) and (*d*), respectively.

where $F^2(hr)$ is the average of the square of the structure factor, $V^2\rho^2$ = squared total number of electrons in the molecule, $h = 2\pi/d$, r = radius of sphere.

The *d* values of the first and second cross-over in this work have been calculated using the *G* function (Rossmann & Blow, 1962) or the interference function,

$$G(hr) = 3\left[\frac{\sin(hr) - hr\cos(hr)}{h^3r^3}\right], \quad (2)$$

h being a point from the origin in reciprocal space defined above.

The interference, scattering or *G* function does not give any information about the specific value of the phase angles relative to the phase origin of the unit cell. It only tells us for which reflections the entire protein will scatter in phase or 180° out of phase with a point at the centre of gravity of the protein. The scattering dependence for different *d* values and radii are shown in Fig. 2.

We have calculated the centre of gravity, $\langle x\rangle$, $\langle y\rangle$, $\langle z\rangle$, of the proteins using,

$$(\langle x\rangle, \langle y\rangle, \langle z\rangle) = \left(\frac{\sum_{i=0}^{N} x_i}{N}, \frac{\sum_{i=0}^{N} y_i}{N}, \frac{\sum_{i=0}^{N} z_i}{N}\right), \quad (3)$$

treating all non-H atoms as C atoms. The centres of gravity of the tested proteins are presented in Table 2. Calculations of structure factors and phases for the centre of gravity of the protein were carried out using *MathCad*® and *SHELXL*93 (Sheldrick, 1993).

The atomic scattering profile in the low-angle region (0–2°) is not affected much by the displacement factor (*B* factor) and we have set the displacement factors to 20 Å^2 for the calculations of structure factors of the centre of gravity of the protein using,

Table 2. *Calculated R factors and number of correctly determined phases for the five tested proteins*

	Carbonic anhydrase	Chymotrypsinogen	Cytochrome P-450	Glucose oxidase	Haemoglobin (T state)
Centre of gravity (x, y, z)	(−0.1310, −0.0409, 0.2267)	(0.3145, 0.5036, 0.2964)	(0.4784, 0.4150, 0.3426)	(0.6359, 0.1137, 0.2468)	(0.0830, 0.0143, 0.2342)
No. of reflections (total/correct) before 1st cross-over	7/7	10/10	6/4	10/10	10/7
$\Delta\varphi_w$* before 1st cross-over (°)	0	3.9	16.4	0	11.2
No. of reflections (total/correct) between 1st and 2nd cross-over	26/11	29/18	29/15	22/12	30/17
$\Delta\varphi_w$ between 1st and 2nd cross-over (°)	58.3	57.5	93.3	74.4	36.2
R value before 1st cross-over	22.2	34.9	33.1	47.0	51.2
R value between 1st and 2nd cross-over	31.1	62.9	67.1	52.5	62.4
CC† before 1st cross-over	83.9	89.2	88.6	43.9	50.1
CC between 1st and 2nd cross-over	19.9	36.7	−13.1	14.1	46.6

* $\Delta\varphi_w$ is the amplitude-weighted phase error as defined in the text. † CC = correlation coefficient in percent.

$$F_{\mathbf{h}}^{\text{sphere}} = |G_{\mathbf{h}}| \sum_{j=1}^{N} f_j \exp[i2\pi(hx_j + ky_j + lz_j)], \quad (4)$$

in which $F_{\mathbf{h}}^{\text{sphere}}$ is the structure factor for the centre of gravity of the protein multiplied with the corresponding d value in the scattering function, $G_{\mathbf{h}}$. The structure factors of the protein have been calculated using B factors for the individual atoms omitting $|G_{\mathbf{h}}|$ in (4). The phases and amplitudes for our spherical model were then compared with the calculated amplitudes and phases of the real protein.

All the presented results are based on calculated structure-factor amplitudes and phases using the coordinates from the PDB.

3. Results and discussion

3.1. *Phases*

All the phases for the reflections before the first cross-over were correctly predicted by our spherical model for three of the five tested proteins, Table 2 and Table 3. We considered a phase as correctly predicted if the phase error was less than 45°. For the other two proteins, cytochrome P-450 and haemoglobin four out of six and seven out of ten reflections, respectively, were correctly predicted. This implies that the phases of the low-order reflections of these proteins indeed can be predicted by the spherical scattering function. All the five proteins we used are fairly spherical, except glucose oxidase and cytochrome P-450 which are ellipsoid and disc-shaped, respectively.

For haemoglobin and cytochrome P-450 not all reflections before the first cross-over were correct. For cytochrome P-450 this may be due to the fact that the protein has a disc-like shape with dimensions about 11:12:7 (see Table 1). One might expect even poorer prediction power for our spherical model for such a non-spherical protein, but when viewing the molecule in the xy plane the molecule has almost a circular shape and before the first cross-over all reflections are ($hk0$). This explains why the number of correct phases is still good. The number of reflections before the first cross-over for cytochrome P-450 is only six, because accidentally four reflections have almost the same d value as the point at the first cross-over.

The difference in shape between cytochrome P-450 and the other three globular proteins (haemoglobin, carbonic anhydrase and chymotrypsinogen) is much larger than the difference between glucose oxidase and these three proteins. This explains why there are more correctly phased reflections before the first cross-over for glucose oxidase than for cytochrome P-450.

For haemoglobin only seven out of ten reflections were correctly predicted. Haemoglobin is an $\alpha_2\beta_2$ tetramer. A cylindrical hole with a diameter of approximately 7 Å passes through the centre of the tetramer. The tetramer is spherical, similar to chymotrypsinogen and carbonic anhydrase, and consists of more than twice the number of atoms compared with carbonic anhydrase. The hole in the middle of the molecule does not affect the spherical scattering dependence seriously, since the number of atoms which could fill the hole would be only about 2% of the total number of atoms in the molecule. When we look closely at the three incorrectly phased reflections we see that one of the reflections is very close to the first cross-over and, therefore, must be considered as uncertain, one has a phase error of 55° which is acceptable and one reflection is very weak compared to the other reflections.

For centric reflections there are only two possible values for the phase and thus the phase difference between the predicted and true phase is either 0 or 180°. For acentric reflections, on the other hand, the phases can take any value between 0 and 360°. We have defined a correct reflection rather strictly as having a phase error $\Delta\varphi = |\varphi_{\text{protein}} - \varphi_{\text{sphere}}| \leq 45°$. This is not too serious an error considering phases from isomor-

Table 3. *Calculated amplitudes and phases for carbonic anhydrase*

The phase error or phase difference $\Delta\varphi = |\varphi_{protein} - \varphi_{sphere}|$. Maximum phase error is 180°. $\varphi_{protein}$ = phase calculated using the non-H atoms from the coordinates of PDB. φ_{sphere} = phase calculated assuming a perfect sphere with the same volume as the protein. The amplitude weighted phase error $\Delta\varphi_w = \sum(180 - \Delta\varphi)|F_h|_{protein}/\sum|F_h|_{protein}$, is 39.5° for the first nine reflections after the first cross-over, 63.3° for the following nine reflections and 76.4° for the last eight reflections. $\Delta\varphi|F_h|_{protein}$ is the product of the phase error and the amplitude.

	Amplitudes		Phases				
hkl	$\|F_h\|_{protein}$	$\|F_h\|_{sphere}$	$\varphi_{protein}$	φ_{sphere}	$\Delta\varphi$	$\Delta\varphi \times \|F_h\|_{protein}$	*d*
001	2378	2375	0	0	0	0	70.64
100	4075	4424	0	0	0	0	41.32
−101	5488	3759	180	180	0	0	40.38
011	3760	3421	75	75	0	0	35.91
002	1390	2978	180	180	0	0	35.32
101	1086	1104	0	0	0	0	32.30
−102	678	513	180	180	0	0	30.98
First cross-over						$(180 - \Delta\varphi)\times$ $\|F_h\|$	
110	133	193	295	255	40	18620	29.35
−111	652	335	341	75	94	56072	29.01
012	1828	378	305	75	130	91400	26.95
111	961	988	300	75	135	43245	25.53
−112	2546	954	49	255	154	66196	24.87
102	1107	879	0	180	180	0	24.02
003	982	862	0	180	180	0	23.55
−103	644	760	0	0	0	115920	23.12
−201	1063	1775	0	180	180	0	21.33
020	804	1640	260	330	70	88440	20.85
112	1980	1465	266	75	169	21780	20.82
200	482	119	0	180	180	0	20.66
013	162	1377	25	255	130	8100	20.50
−113	1055	1312	55	255	160	21100	20.22
−202	96	303	180	180	0	17280	20.19
021	989	193	154	330	176	3956	20.00
201	861	642	0	0	0	154980	18.61
120	1251	448	286	330	44	170136	18.61
103	1447	597	0	180	180	0	18.55
−121	1603	386	330	150	180	0	18.53
210	874	604	126	255	129	44574	18.51
−104	120	302	0	0	0	21600	17.96
022	894	298	20	150	130	44700	17.96
−203	1616	276	0	0	0	290880	17.93
004	1053	126	0	0	0	189540	17.66
121	662	63	211	330	119	40382	17.52
Second cross-over							

phous replacement often have errors of 50–60° before refinement is started.

It is more important to obtain good phases for the strong reflections than for the weak ones. We have calculated an amplitude-weighted phase error $\Delta\varphi_w$ as a figure of merit of the phase error as,

$$\Delta\varphi_w = \frac{\sum \Delta\varphi|F_h|}{\sum|F_h|}. \qquad (5)$$

The expectation value of the weighted phase error assuming random phases will be 90°. For all the five proteins (see Table 2) $\Delta\varphi_w$ is <20° for the reflections before the first cross-over.

As expected, the phases between the first and second cross-overs are not predicted as well by our spherical model as the phases before the first cross-over. At higher resolution both the deviations from a sphere and the internal structural features (α-helices *etc.*) will become increasingly important. However, for the more or less spherical protein haemoglobin the phases were predicted quite well also between the first and second cross-over, with $\Delta\varphi_w = 36°$.

3.2. *Amplitudes*

The *R* values of the structure-factor amplitudes, shown in Table 2, were calculated using,

$$R = \frac{\sum \left||F_h^{protein}| - sc|F_h^{sphere}|\right|}{\sum|F_h^{protein}|}, \qquad (6)$$

where sc is the scaling factor. The correlation coefficient, CC has been calculated as a complement to the structure factor *R* value defined as,

$$CC = [\sum|F_c^a||F_c^b| - (\sum|F_c^a|\sum|F_c^b|)/N] \div [\sum(|F_c^a|)^2 - (|F_c^a|)^2/N]^{1/2} \times [\sum(|F_c^b|)^2 - \sum(|F_c^b|)^2/N]^{1/2}.$$

The *R* value of the structure-factor amplitudes is an important parameter since it will be the restraining parameter for finding the centre of gravity of the protein.

The *R* values differ very much between the five tested proteins. For carbonic anhydrase the *R* value before the first cross-over is only 22%, see Table 2. Chymotrypsinogen and cytochrome P-450 have somewhat higher *R* values (34.9 and 33.1%). Cytochrome P-450 has also fewer reflections before the first cross-over compared to chymotrypsinogen. Finally, glucose oxidase and haemoglobin have quite high *R* values (47.0 and 51.2%).

The low *R* value for cytochrome P-450 can be explained in the same way as for the phases, namely that only the (*hk*0) reflections are present in this region and thus, the amplitudes comply with the two-dimensional projection of a sphere.

The shape of glucose oxidase is more or less elliptical, but circular like cytochrome P-450 in the *xy* plane. However, the *R* value of glucose oxidase is significantly higher than the *R* value of cytochrome P-450. A plausible explanation for this is that the packing of atoms in the molecule is not as uniform as in cytochrome P-450 and that glucose oxidase have some residues sticking out which will slightly perturb the distribution of atoms.

Almost the same arguments can explain the higher *R* value of haemoglobin. It is plausible that the amplitudes

are affected to a larger extent than the phases by the hole in the middle of the molecule. Another contributing effect might be that the interfacial geometry between the subunits will be somewhat different compared with a monomeric protein and that the centre of haemoglobin does not see the same distribution of atoms in all directions.

The R value for the reflections between the first and second cross-over are random or very close to random which for non-centrosymmetric space groups is 59%. Errors in our simplified model affect the amplitudes more than the phases and thus the geometrical considerations probably play a more important role for the amplitudes than for the phases. Although the R value between the first and second cross-over for haemoglobin and cytochrome P-450 are nearly the same, the correlation coefficient indicates that the amplitude data for cytochrome P-450 is much worse, see Table 2.

However, for three of the five proteins tested the reflections before the first cross-over have both correct phases and a low R value and for the other two proteins the phases have been predicted to an acceptable accuracy. We believe that we can find the centre of gravity of the protein from a Patterson search or some other search algorithm in the asymmetric unit using the experimentally measured low-order reflections.

We believe it will be necessary to switch from a spherical model to an ellipsoid in order to phase the reflections in the region between the first and second cross-over of the G function. This means that we need to find six further numbers in addition to the three (x, y, z) coordinates of the centre of the molecule. These are the lengths of the three ellipsoid axes and the orientation of the ellipsoid. At this stage we will need to estimate a total of nine unknowns from 40 or so reflections out to the second cross-over. This might be possible since the over-determination is still about four.

4. Conclusions

We have presented a method which produces phases for low-order reflections to a satisfactory accuracy for globular proteins consisting of more than 1500 atoms. A feature of the method is that the interference function contained in the rotation function (Rossmann & Blow, 1962) indeed works for our purposes and allows us to get useful phase information also of molecules deviating significantly from a globular shape. By using intensities of the ten lowest order reflections we believe that we can find the position of the protein in the unit cell using Patterson or some other search algorithm for proteins with a shape not deviating more from the spherical shape than those proteins presented here. The error in the position of the protein should not be more than 3–4% of the cell parameters. The technique of collecting good-quality data in the low-angle region and the correction of the bulk solvent to the observed amplitudes are some of the things which are under development and that have to be improved to facilitate the search of the centre of gravity of the protein and thereby a starting envelope. The search program for finding the centre of gravity of the protein is under construction and will be presented in the near future.

Other shapes, such as ellipsoids, must be used in order to use reflections out to higher resolution. The orientation of the ellipsoid may be determined using the molecular replacement method. Phase extension methods, such as direct methods or maximum entropy, might be used to obtain phases for even higher resolution reflections, which are needed in order to discern the internal features of a protein.

The authors would like to thank Drs Pär Nordlund, Salam al Karadghi, Eila Cedergren-Zeppezauer and Professors Per-Erik Werner and Anders Liljas for fruitful and encouraging discussions. We would also like to thank Mr Karl-Erik Johansson for drawing some of the figures.

References

Andersson, K. M. & Hovmöller, S. (1996). In preparation.
Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. F. Jr, Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T. & Tasumi, M. (1977). *J. Mol. Biol.* **112**, 535–542.
Blow, D. M. & Crick, F. H. C. (1959). *Acta Cryst.* **12**, 794–802.
Bricogne, G. (1993). *Acta Cryst.* D**49**, 37–60.
Cowtan, K. D. & Main, P. (1993). *Acta Cryst.* D**49**, 148–157.
Eriksson, A. E., Jones, A. T. & Liljas, A. (1988). *Proteins*, **4**, 274–282.
Freer, S. T., Kraut, J., Robertus, J. D., Wright, H. D. & Xuong, N. H. (1970). *Biochemistry*, **9**, 1997–2009.
Gellatly, B. J. & Finney, J. L. (1982). *J. Mol. Biol.* **161**, 305–322.
Gilmore, C. J. (1992). *Curr. Opin. Struct. Biol.* **2**, 806–810.
Guiner, A. & Fournet, G. (1955). *Small-angle Scattering of X-rays*. New York: John Wiley.
Harris, G. W. (1995). *Acta Cryst.* D**51**, 695–702.
Hecht, H. J., Kalisz, H. M., Hendle, J., Schmid, R. D. & Schomburg, D. (1993). *J. Mol. Biol.* **229**, 153–172.
Hendrickson, W. A. (1991). *Science*, **254**, 51–58.
Kraut, J. (1958). *Biochim. Biophys. Acta*, **30**, 265–270.
Lunin, V. Y., Lunina, N. L., Petrova, T. E., Vernoslova, E. A., Urzhumtsev, A. G. & Podjarny, A. D. (1995). *Acta Cryst.* D**51**, 896–903.
Main, P. (1990). *Acta Cryst.* D**46**, 372–377
Podjarny, A. D., Bhat, T. N. & Zwick, M. (1987). *Ann. Rev. Biophys. Chem.* **16**, 551–573.
Raag, R. & Poulos, T. L. (1991). *Biochemistry*, **30**, 2674–2684.
Rossmann, M. G. & Blow, D. M. (1962). *Acta Cryst.* **15**, 24–31.
Sayre, D. (1952). *Acta Cryst.* **5**, 60–65.

Sheldrick, G. M. (1993). *SHELXL93. Program for the Refinement of Crystal Structures.* University of Göttingen, Germany.
Sjölin, L. & Svensson, A. L. (1993). *Acta Cryst.* D**49**, 66–74.
Smith, J. L. (1991). *Curr. Opin. Struct. Biol.* **1**, 1002–1011.
Subbiah, S. (1991). *Science*, **252**, 128–133.
Teeter, M. M. & Hendrickson, W. A. (1979). *J. Mol. Biol.* **127**, 219–223.
Urzhumtsev, A. & Podjarny, A. (1995). *Acta Cryst.* D**51**, 888–895.
Waller, D. A. & Liddington, R. C. (1990). *Acta Cryst.* B**46**, 409–418.
Wang, B. (1985). *Methods Enzymol.* **115**, 90–112.
Weeks, C. M., Hauptman, H. A., Smith, G. D. & Blessing, R. H. (1995). *Acta Cryst.* D**51**, 33–38.
Zhang, K. Y. J. & Main, P. (1990*a*). *Acta Cryst.* A**46**, 41–46.
Zhang, K. Y. J. & Main, P. (1990*b*). *Acta Cryst.* A**46**, 377–381.