Acta Cryst. (1995). D**51**, 140–144

A Probe Molecule Composed of Seventeen Percent of Total Diffracting Matter Gives Correct Solutions in Molecular Replacement

BY BYUNG-HA OH*

Department of Macromolecular Science, SmithKline Beecham Pharmaceuticals, 709 Swedeland Road, King of Prussia, PA 19406, USA

(*Received* 1 *November* 1993; *accepted* 1 *September* 1994)

Abstract

It is often found in the crystallization of enzyme–inhibitor complexes that an inhibitor causes crystal packing which is different to that of native protein. This is the case for crystals of human non-pancreatic secreted phospholipase A_2 (124 residues) containing six molecules in the asymmetric unit when the protein is complexed with a potential acylamino analogue of a phospholid. The hexameric structure was determined by molecular replacement using the structure of monomeric native protein as a probe. As an extension to the experiment, it was tested whether a backbone polypeptide composed of 17% of a known monomeric structure could find its correct position on a target molecule in molecular replacement. A probe model composed of the backbone atoms of the N-terminal 77 residues of lysine-, arginine-, ornithine-binding protein (LAO, a total of 238 residues) liganded with lysine correctly finds its position on LAO liganded with histidine which crystallizes as a monomer in the asymmetric unit. The results indicate that as little as 17% of total diffracting matter can be used in molecular replacement to solve crystal structures or to obtain phase information which can be combined with phases obtained by the isomorphous-replacement method.

Molecular replacement is a primary technique for the determination of crystal structures of enzyme–inhibitor complexes when the three-dimensional structure of native protein is available and the crystal packing of a complex is different from that of native enzyme. In molecular replacement, the smaller the size of the probe molecule which is used as compared with that of the molecule(s) in the asymmetric unit, the harder it is to obtain correct solutions. We have crystallized human non-pancreatic secreted phospholidpase A_2 (hnps-PLA2) complexed with 1-octadecyl-2-acetamido-2-deoxy-*sn*-*glycero*-3-phosphoethylmethyl sulfide in space group $P2_1$ with cell parameters $a = 64.51$, $b = 114.72$, $c = 64.43$ Å, $\alpha = \gamma = 90.0$, $\beta = 120.13°$. The crystal lattice volume indicated that there were six molecules of the hnps-PLA2–inhibitor complex in the asymmetric unit, corresponding to a crystal volume per protein mass (V_m) of 2.36 Å^3 Da^{-1} and a solvent content of 47.9% by volume (Matthews, 1968). A self-rotation search clearly indicated the presence of a non-crystallographic twofold symmetry. It was possible to determine the hexameric structure by the molecular-replacement method using the previously determined monomeric native hnps-PLA2 structure (Scott *et al.*, 1991) as a search model. The probe molecule constitutes less than 16% of a total diffracting matter considering the size of the bound inhibitor.

This represents the first example demonstrating that the structure of a multimer in a crystal containing as many as six molecules in the asymmetric unit can be determined by molecular replacement using a single molecule as a probe. The closest other example is a cyclophilin–cyclosporin structure of a pentamer in the asymmetric unit solved by molecular replacement (Pflugl *et al.*, 1993). Molecular-replacement experiments with small probe molecules can be found in the literature. Nordman (1972) demonstrated that an intragroup vector set characteristic of a four- or five-turn α-helix can be used to find the axial direction of a major α-helices in myoglobin by a vector-space search method, and suggested that the method could be applied to detect and determine the orientation of assumed relatively large substructures; another group reported test experiments showing that a model composed of C^α atoms of an Fab subdomain, comprising only 6% of the asymmetric content, yielded the correct orientation of the fragment in rotation searches (Cygler & Anderson, 1988*a*), and that a correctly oriented Fab subdomain with as little as 24% of the asymmetric unit content yielded a correct position in translation searches (Cygler & Anderson, 1988*b*). Since phase information cannot be obtained from an orientation solution alone (except for $P1$ symmetry), the size of a good incomplete model used for the molecular replacement should be large enough to yield correct solutions both in the rotation and translation search. This paper describes the structure determination of the hnps-PLA2–inhibitor complex along with test experiments which show that 16–17% of a total diffracting matter can be

* Current address: Department of Life Science, Pohang University of Science and Technology, PO Box 125, Pohang, S. Korea.

Table 1. *Parameters used in the rotation search using hnps-PLA2 (monomer) as a probe*

*P*1 cell for the probe model (Å)	110 × 110 × 110
Resoluton (Å)	10–4
Patterson vector length (Å)	< 45, 3σ above the mean
No. of Patterson vectors	3000
Sampling interval for θ_2* (°)	2.5

* For definition, see Rossman & Blow (1962).

generally used in molecular replacement for crystal structure determination.

Molecular replacement was carried out using the *X-PLOR* program package (Brünger, 1991). Various parameters used for the cross-rotation search are shown in Table 1. The top 110 rotation-function peaks were further refined by the Patterson correlation (PC) refinement (Brünger, 1991) which consisted of 15 steps of rigid-body conjugate-gradient minimization (Powell, 1977) against reflections at 10–4.0 Å resolution. The PC refinement resulted in three different refined rotation solutions distinguishable from the noise level (Fig. 1). At this point, it was confirmed that any pairs of the three solutions were not related by the non-crystallographic twofold symmetry defined earlier.

The two-dimensional (*xz* plane) translation searches were carried out for the three refined rotation solutions. The reflections at 10–3.5 Å resolution were used, and the search with 1 Å sampling interval was restricted to the half of *x* and *z* axis of the unit cell. The first, second and third orientation yielded a top translation solution only 1.25σ (molecule *A*), 1.34σ (molecule *B*) and 1.12σ (molecule *C*) higher than the next highest peak, respectively. At this point, it was decided to use the local twofold symmetry information to check whether the position of molecule *A* is actually the correct translation solution. After generating a molecule (molecule *D*) by applying the non-crystallographic twofold to molecule *A*, two translation searches (along *x* and *z*, and along *y*) were carried out to determine the position of molecule *D* with respect to molecule *A*. The second translational search (along *y*), in which molecule *D* was translated with fixed position of molecule *A*, resulted in a top translation peak 3.4σ higher than the next highest peak. The two molecules showed favorable contacts with each other (figure not shown), and there was no overlap between symmetry-related molecules. This result confirmed that the position of molecule *A* was indeed the correct translation solution and that the correct position of molecule *D* was determined. The same procedure could have been repeated for molecules *B* and *C*. However, since the packing of molecules *A* and *D* might be a repeating unit for the packing of the hexamer in the unit cell, the two molecules were used as a probe for subsequent molecular replacement to test the idea and, if true, to find the positions of the other four molecules. In using molecules *A* and *D* as a probe, a rotation search followed by PC refinement resulted in the three different strong rotation-function peaks (Fig. 1*b*) representing

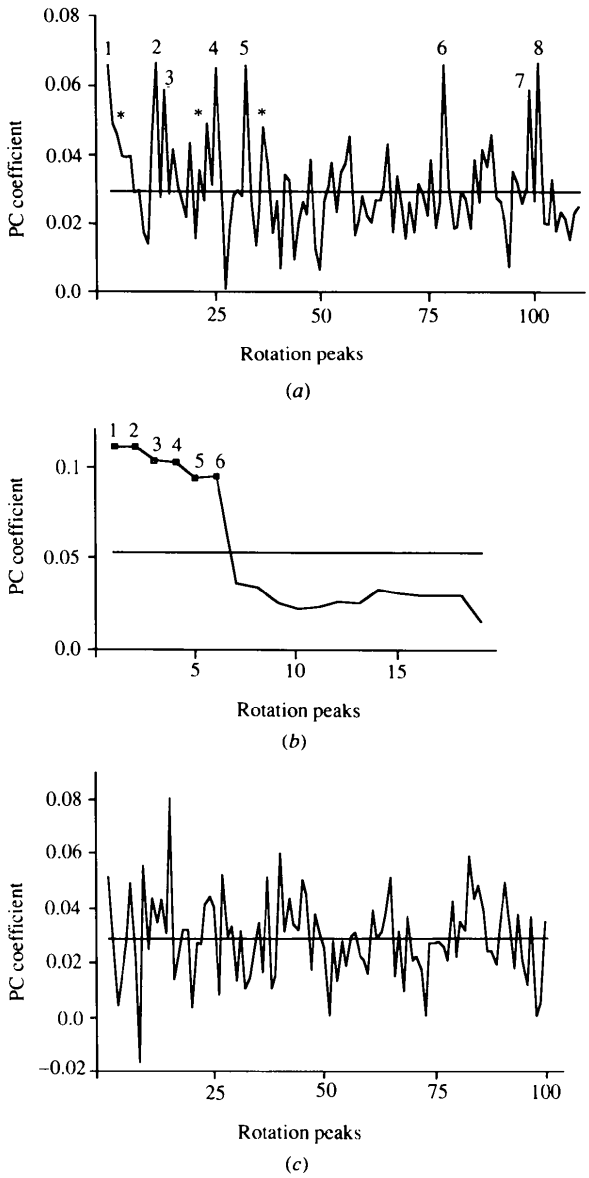

Fig. 1. Patterson correlation refinement of rotation peaks. (*a*) Peaks from the rotation search using monomeric hnps-HPLA2 as a probe model. The rotation peaks (1,5,6), (2,4,8) and (3,7) were converged to the same refined peak, respectively. The asterisks indicate a correct rotation peak which can be readily distinguishable from the noise level in the PC refinement at 10–3.5 Å resolution. (*b*) Peaks from the second rotation search using two molecules of hnps-HPLA2 as a probe model (see text). The rotation peaks (1,2), (3,4) and (5,6) were converged to the same solution, respectively. The peaks (1,2) are non-rotated self peaks, and the peaks (3,4) and (5,6) represent rotational positions of the other two pairs of two molecules in the asymmetric unit. (*c*) Peaks from the rotation search using the backbone atoms of the N-teminal 77 residues of LAO as a probe model against reflection daa of LAO–histidine complex. The variable parameters used in the rotation search were the same as in Table 1 except for a maximum Patterson vector length of 30 Å and for the cube size of 80 Å for the *P*1 cell. The horizontal lines in the figures indicate the mean values of PC coefficients. In all three rotation searches the search space was correctly restricted according to Rao, Jih & Hartsuck (1980).

rotational positions of all six molecules in the asymmetric unit. At this point, all subsequent translation searches to find the positions of other two 'dimers' were straightforward, since the probe molecule constituted of about 33% of the total diffracting matter. Subsequent rigid-body refinement (20 steps of conjugate-gradient minimization against reflections at 8–3.0 Å resolution) of the six molecules lowered the R factor to 39.1%. In the $2F_o - F_c$ map, the strong electron density for bound inhibitor unambiguously guided how it should be fitted (Fig. 2). After some manual refitting of side chains, subsequent atoms position and B-factor refinement including the inhibitor lowered the R factor to 20.5% for reflections at 8–2.8 Å resolution. The details of the structure will be presented elsewhere.*

After the positions of the six molecules were determined, it was confirmed that one-dimensional translation searches along y found correct positions of molecules B and C with respect to molecule A, showing that the positions of molecules B and C were also the correct translation solutions. This indicates that the positions of the other two molecules could have been detemined alternatively; two molecules can be generated by applying the non-crystallographic twofold symmetry to molecules B and C, and their positions with respect to molecule A can be determined by translation searches along y. The same procedure used for determining the position of molecule D.

Later, different resolution shells or different lengths of Patterson search vectors in the rotation search using monomeric probe were tested for the possibiity of finding the orientations of all six molecules in the asymmetric unit. A rotation search using the same parameters as in Table 1, except for employing shorter Patterson search vectors (<15 Å), yielded only two of the three correct rotation solutions described earlier. This is probably due to the fact that shorter interatomic vectors are much more isotropically distributed than longer interatomic vectors in the probe molecule as shown in Fig. 3. The interatomic vectors in the distance ranges of 10–15 and 40–45 Å of the probe molecule are calculated and displayed as a function of angles φ and θ, respectively, which are defined in the Fig. 3. The 40–45 Å distance vectors are quite non-isotropic, whereas 10–15 Å distance vectors are fairly uniformly distributed. Thus, selection of the longer Patterson vectors in rotation searches could yield correct rotation solutions more easily. In fact, the first and eighth highest rotation-function peaks (the two correct solutions, Fig. 1*a*) from the rotation search using longer Patterson vectors (<45 Å) were only the 11th and 74th highest rotation-function peaks, respectively, when shorter vectors (<15 Å) were used (figure not shown). Subsequently, test experiments varying the resolution limit were carried out all with longer Patterson search vectors (maximum length of 45 Å). When reflections of 10–3.5 Å resolution were used, an additional correct rotation solution was readily distinguishable from the noise level (figure not shown) in the PC refinement along with the other three rotation solutions described earlier. It was noticed that the additional solution is indeed present in Fig. 1 (labeled with asterisks), but is hardly distinguishable from the noise level. Apparently, the PC coefficient (for definition, see, Brüger, 1991) of this rotation-function peak becomes relatively larger than those of faulty rotation-function peaks at higher resolution. The rotational position of any one of the two other molecules in the asymmetric unit was not found in the rotation searches and PC refinements using other resolution shells, 15–4.0, 15–3.5 or 15–3 Å. In carrying out each rotation search and PC refinement, the same resolution shell was used, and the top 100 rotation-function peaks from each rotation search were refined by the PC refinement.

The results presented here show clearly that a monomeric probe can be used to determine the orientation of several, though not all, molecules of a hexamer in the asymmetric unit by conventional molecular replacement plus PC refinement. However,

* Atomic coordinates and structure factors have been deposited with the Protein Data Bank, Brookhaven National Laboratory (Reference: 1AYP, R1AYPSF). Free copies may be obtained through The Managing Editor, International Union of Crystallography, 5 Abbey Square, Chester CH1 2HU, England (References: GR0370). At the request of the authors the atomic coordinates will remain priviliged until 31 July 1996. A list of deposited data is given at the end of this issue.

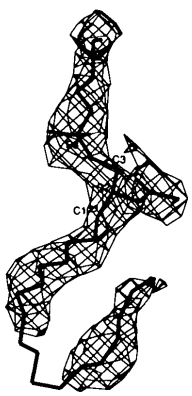

Fig. 2. $2F_o - F_c$ electron density for the bound inhibitor in one of the six molecules of hnps-PLA2–inhibitor complex in the asymmetric unit. Electron density was calculated for the rigid-body refined coordinates (R factor = 39.1%) with reflections at 10–2.8 Å. The bound inhibitor was not included in the calculation. The contour level is 1.0σ. The bound inhibitor is shown by a thick line. The three C atoms of the glycerol backbone and the phosphate group of the inhibitor are labeled. The other five molecules in the asymmetric unit have similar strong electron density for the bound inhibitor. The electron density revealed that the long 18-C-atom chain foids back onto itself in the hydrophobic channel of the hnps-PLA2. The portion of the aliphatic chain, which does not have electron density, is exposed to the solvent. Diffraction data were collected on a Siemens multiwire area detector and were reduced using the program *XENGEN* (Howard *et al.*, 1987). The internal R_{merge} on I_{hkl} is 5.79% for a total of 24374 reflections (16515 unique reflections). The cumulative completeness of the data is 91.0 and 76.3% (> 1σ) up to 3.5 and 2.8 Å, respectively.

since a hexameric packing of molecules is most likely to have a non-crystallographic symmetry, this combined with the molecular replacement can be generally used to determine the positions of all six molecules.

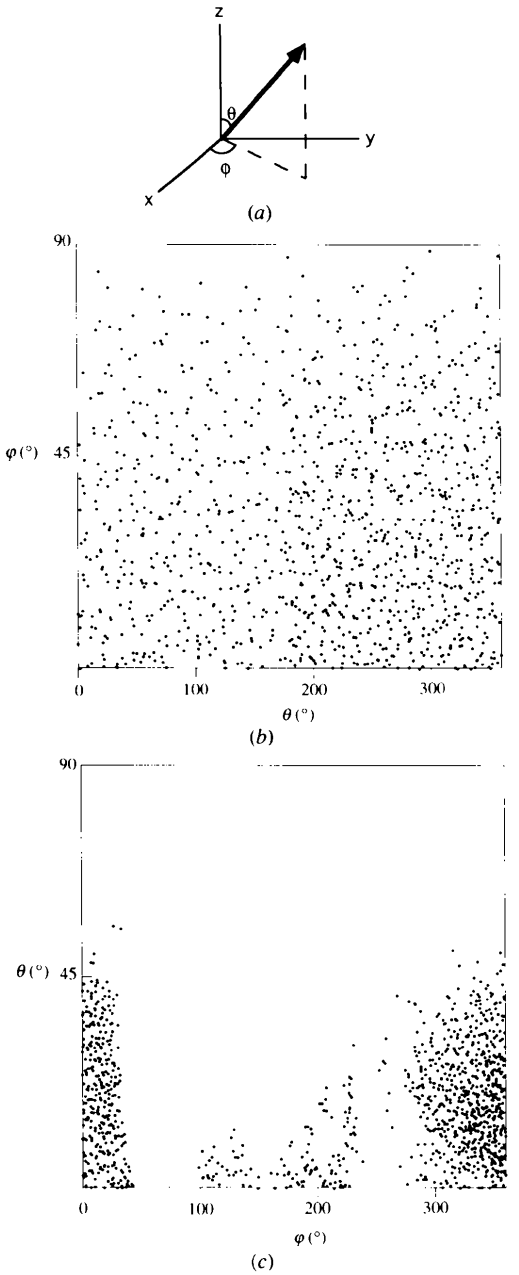

Fig. 3 Two-dimensional display of interatomic vectors of the native hnps-PLA2 coordinates as a function of angles φ and θ. Because any two atomic coordinates can yield two vectors with opposite directionality, only one vector with θ greater than 0° was selected. (*a*) Definition of φ and θ angle. (*b*) Vectors with magnitude of 10–15 Å. (*c*) Vectors with magnitude of 40–45 Å. Ensemble of the longer distance vectors (*a*) is obviously more non-isotropically distributed than that of the shorter distance vectors (*b*). In both cases, vectors closer than 2.0 Å each other were clustered to one vector.

Table 2. *Top ten solutions of the translation search carried out with the top rotation solution using the backbone atoms of* N*-terminal* 77 *residues of LAO–lysine complex against diffraction data* (10–4 Å) *of the LAO–histidine complex*

(*x*,*y*,*z*) translation (Å)	Translation function*
(0.375 26.185 57.648)	0.1904
(0.000 26.185 54.119)	0.1880
(0.000 26.185 57.648)	0.1880
(0.375 26.185 45.884)	0.1846
(0.000 26.185 45.884)	0.1843
(0.375 26.767 45.884)	0.1820
(0.375 26.185 56.472)	0.1804
(0.000 26.767 45.884)	0.1784
(0.000 55.862 45.884)	0.1784
(0.375 26.185 54.119)	0.1764

*σ of a total of 1000 translation functions is 0.022.

Finally, it was considered whether a partial model composed of 17% of a whole structure could find the correct position in the known structure of a monomer in the asymmetric unit in molecular replacement. The partial model was constructed as backbone atoms of N-terminal 77 residues of highly refined coordinates of lysine-, arginine-, ornithine-binding protein (LAO, 238 residues) liganded with lysine (Oh *et al.*, 1993). The N-terminal 77 residues was chosen arbitrarily, and does not represent unique structural motif of the protein as shown in Fig. 4. This probe was searched against the diffraction data of LAO liganded with histidine. The crystal of LAO–histidine contains one molecule in the asymmetric unit and the crystal packing of LAO–histidine is completely different from that of LAO–lysine (Oh, Ames & Kim, 1994). The r.m.s. deviation between backbone atoms of LAO–lysine and LAO–histidine is 0.28 Å. Even in this test case, the top refined rotation peak (Fig. 1*c*), and top translation peak (Table 2) were

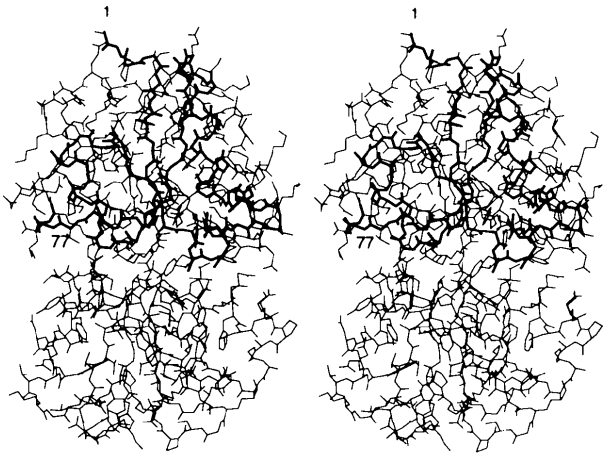

Fig. 4. The crystal structure of LAO–lysine complex. The thick lines show the location of the backbone atoms of the N-terminal 77 residues used as a probe model in the test molecular replacement (see text). The numbers (1 and 77) indicate the first and last residues of the probe model.

the correct solutions. The top translation solution showed an r.m.s. deviation of 0.65 Å with LAO–histidine coordinates. The translation function of the top peak is barely larger than those of the following peaks. However, it was noted that the top eight peaks show nearly the same values for the x and y coordinates as shown in Table 2. With backbone atoms of N-terminal 74 residues (16% of total atoms), the correct orientation solution was found unambiguously, but a correct translation solution was not found. Probably, 16–17% of total diffracting matter is the minimum number of atoms that can provide correct molecular-replacement solutions.

I thank Dr J. Hadju for providing the inhibitor. I also thank Drs S.-H. Kim, D. Egglestone and S. Abdel-Meguid for helpful discussions, and Dr S. Dixon for providing computing resources during the research.

References

BRÜNGER, A. T. (1991). *X-PLOR*, Version 3.0. Yale Univ., New Haven, Connecticut, USA.

CYGLER, M. & ANDERSON, W. F. (1988*a*). *Acta Cryst.* A**44**, 38–45.

CYGLER, M. & ANDERSON, W. F. (1988*b*). *Acta Cryst.* A**44**, 300–308.

HOWARD, A. J., GILLIAND, G. L., FINZEL, B. C., POULOS, T. L., OHLENDORF, D. H. & SALEMME, F. R. (1987). *J. Appl. Cryst.* **20**, 383–403.

MATTHEWS, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.

NORDMAN, C. E. (1972). *Acta Cryst.* A**28**, 134–143.

OH, B.-H., AMES, G. F.-L. & KIM, S.-H. (1994). *J. Biol. Chem.* **269**. In the press.

OH, B.-H., PANDIT, J., KANG, C. H., NIKAIDO, K., GIOKCEN, K., AMES, G. F.-L. & KIM, S.-H. (1993). *J. Biol. Chem.* **268**, 11348–11355.

PFLUGL, G., KALLEN, J., SCHIRMER, T., JANSONIUS, J. N., ZURINI, M. G. & WALKINSHAW, M. D. (1993). *Nature (London)*, **361**, 91–94.

POWELL, M. J. D. (1977). *Math. Programming*, **12**, 273–278.

RAO, S. N., JIH, J.-H. & HARTSUCK, J. A. (1980). *Acta Cryst.* A**36**, 878–884.

ROSSMAN, M. G. & BLOW, D. M. (1962). *Acta Cryst.* **15**, 24–31.

SCOTT, D. L., WHITE, S. P., BROWNING, J. L., ROSA, J. J., GELB, M. H. & SIGLER, P. B. (1991). *Science*, **254**, 1007–1010.