Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Gd-HPDO3A, a complex to obtain high-phasing-power heavy-atom derivatives for SAD and MAD experiments: results with tetragonal hen egg-white lysozyme

Éric Girard, Laurent Chantalat, Jean Vicat and Richard Kahn*

Laboratoire de Cristallographie Macromoléculaire, Institut de Biologie Structurale J.-P. Ebel CEA–CNRS–UJF, 41 Rue Jules Horowitz, 38027 Grenoble CEDEX 01, France

Correspondence e-mail: kahn@ibs.fr

Received 18 July 2001
Accepted 3 October 2001

PDB Reference: Gd derivative of lysozyme, 1h87.

A neutral gadolinium complex, Gd-HPDO3A, is shown to be a good candidate to use to obtain heavy-atom derivatives and solve macromolecular structures using anomalous dispersion. Tetragonal crystals of a gadolinium derivative of hen egg-white lysozyme were obtained by co-crystallization using different concentrations of the complex. Diffraction data from three derivative crystals (100, 50 and 10 m*M*) were collected to a resolution of 1.7 Å using Cu $K\alpha$ radiation from a rotating anode. Two strong binding sites of the gadolinium complex to the protein were located from the gadolinium anomalous signal in both the 100 and 50 m*M* derivatives. A single site is occupied in the 10 m*M* derivative. Phasing using the anomalous signal at a single wavelength (SAD method) leads to an electron-density map of high quality. The structure of the 100 m*M* derivative has been refined. Two molecules of the gadolinium complex are close together. Both molecules are located close to tryptophan residues. Four chloride ions were found. The exceptional quality of the SAD electron-density map, only enhanced by solvent flattening, suggests that single-wavelength anomalous scattering with the Gd-HPDO3A complex may be sufficient to solve protein structures of high molecular weight by synchrotron-radiation experiments, if not by laboratory experiments.

1. Introduction

Since obtaining good isomorphous derivatives is often the bottleneck in the multiple isomorphous replacement method, phasing methods taking advantage of anomalous dispersion are now widely used for *de novo* structure determination in macromolecular crystallography. In this context, the single-wavelength anomalous dispersion (SAD) phasing method is particularly interesting, as it can be applied to data collected using conventional X-ray sources.

Successful SAD experiments have used the anomalous signal from atoms either present in the native protein, such as sulfur (Hendrickson & Teeter, 1981; Dauter *et al.*, 1999), copper (Zheng *et al.*, 1996; Harvey *et al.*, 1998), iron (Royer *et al.*, 1989; Wu *et al.*, 2001) and cadmium (Furey *et al.*, 1986), giving the advantage of fully occupied sites, or incorporated in the crystal, such as caesium (Wallace *et al.*, 1990; Burkhart *et al.*, 1998), iodine (Chen *et al.*, 1991), lead (Biou *et al.*, 1995), selenium (Sha *et al.*, 1995), mercury (Fan *et al.*, 1990) and platinum (Sabini *et al.*, 2000). Potential use of bromide ions for SAD experiments has been noted by Dauter & Dauter (1999). Different halides were used to phase proteins from derivatives obtained by short cryosoaking (Dauter *et al.*, 2000; Nagem *et*

al., 2001; Dauter & Dauter, 2001). Such a method has been used to solve the structures of human acyl protein thioesterase (Devedjiev *et al.*, 2000) and of pepstatin-insensitive carboxyl proteinase (Dauter *et al.*, 2001).

SAD phasing is more likely to succeed when using anomalous scatterers that exhibit high f'' values. For Cu $K\alpha$ radiation, lanthanides, especially samarium ($f'' = 13.3\ e^-$) and gadolinium ($f'' = 12.0\ e^-$) (*International Tables for X-ray Crystallography*, 1985, Vol. III, pp. 214–215), are of particular interest. Surprisingly, no SAD experiments using lanthanides, except for the special case of Ca substitution by holmium (Brodersen *et al.*, 2000), have been reported up to now.

This paper reports the use of a neutral highly soluble gadolinium complex called Gd-HPDO3A to obtain derivatives with a strong anomalous signal. Gd-HPDO3A is a gadolinium complex of 10-(2-hydroxypropyl)-1,4,7,10-tetraazacyclododecane-1,4,7-triacetic acid (Fig. 1), which is used as a contrast agent for magnetic resonance imaging. The complex is stable over a fairly large pH range from 4 to 9. Our initial intention was to use this complex for the MASC method (Ramin *et al.*, 1999). To observe a pure MASC effect, the anomalous scatterers, dispersed at high concentration in the crystal solvent channels, should not bind to the macromolecule. In fact, we found that Gd-HPDO3A binds tightly to hen egg-white lysozyme (HEWL). This property was used to solve the phase problem using the gadolinium anomalous signal collected at a single wavelength from a copper rotating-anode X-ray source away from the absorption edge. Based on these results, phasing using the anomalous signal from Gd-HPDO3A should provide a broadly applicable method for the determination of new macromolecular structures.

2. Methods

2.1. Crystallization

A single batch of hen egg-white lysozyme (Boehringer No. 13032022-90) was used without further purification. Protein solutions were prepared by dissolving 4 mg of lyophilized HEWL in 100 µl of Gd-HPDO3A solution at the required complex concentration. Tetragonal crystals were grown by the vapour-diffusion technique using the hanging-drop method, with 0.7–1.0 *M* sodium chloride and 50 m*M* sodium acetate buffer at pH 4.5. Crystals were obtained using concentrations of Gd-HPDO3A in the range 10–500 m*M* without further changes in the crystallization conditions for the native crystals. Crystals prepared with Gd-HPDO3A concentrations of 10, 50 and 100 m*M* were used for data collection. The derivative crystals are similar in size and shape to the native ones, except at high complex concentrations (greater than 100 m*M*), where the crystals are flattened along the *c* cell-parameter direction. All crystals belong to the space group $P4_32_12$.

2.2. Data collection and data processing

Data collections were performed at the Cu $K\alpha$ wavelength (1.5418 Å) using a Rigaku RU-200 rotating-anode X-ray generator equipped with a Supper 7600 double-mirror system. The X-ray generator was operated at 40 kV and 100 mA with a focal size of 0.3 × 3 mm (effective focal size of 0.3 × 0.3 mm). Data were recorded on a MAR300 imaging-plate detector. In order to remove any remaining Gd-HPDO3A in the mother liquor surrounding the sample, crystals were soaked for 30 s in reservoir solution containing 25 or 30% PEG 400 as a cryo-

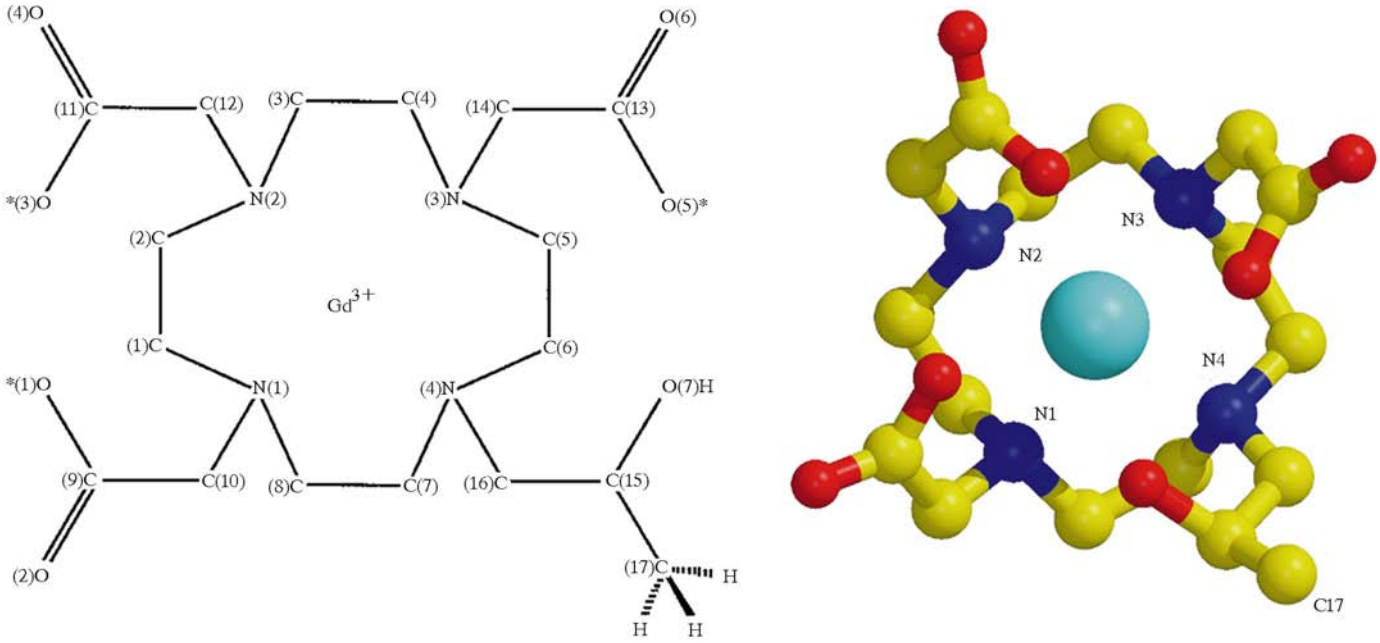

Figure 1
The Gd-HPDO3A molecule: planar formula (left) and *MOLSCRIPT* view (right) (Kraulis, 1991). The *MOLSCRIPT* view was rendered with *Raster3D* (Merritt & Bacon, 1997).

protectant (Rodgers, 1994). The samples were then N_2-cryo-cooled at 100 K (Oxford Cryosystems). No attempt was made to orient the crystals along any particular crystallographic axis. For each data collection the crystal-to-detector distance was set to 110 mm. The oscillation range per image was 1° for a total oscillation range of 180° in order to obtain high multiplicity. The total exposure time per crystal was 30 h. Under these experimental conditions, data for the three derivative crystals were collected to a resolution of 1.7 Å.

Measurements were integrated using the program *XDS* (Kabsch, 1988). The integrated intensities were scaled and merged using *SCALA* from the *CCP*4 suite (Collaborative Computational Project, Number 4, 1994).

A summary of data-collection parameters and processing statistics is given in Table 1. The respective numbers of observed and unique reflections, the signal-to-noise ratios, the completeness and multiplicity as well as the $R_{\rm sym}$ factors are similar for the three data sets. In contrast, as evidenced by the $R_{\rm ano}$ factor, the anomalous signal increases with increasing concentrations of Gd-HPDO3A from 10 to 100 m*M* in the derivative solutions.

Table 1
Summary of data-collection parameters and processing statistics.

Values in parentheses refer to the highest resolution shell.

Gd-HPDO3A	10 m*M*	50 m*M*	100 m*M*
Crystal size (mm)	0.2 × 0.2 × 0.3	0.2 × 0.2 × 0.35	0.3 × 0.3 × 0.2
Unit-cell parameters (Å)	*a* = 77.59, *c* = 37.80	*a* = 77.11, *c* = 38.17	*a* = 77.25, *c* = 38.66
Resolution range (Å)	34.7–1.71 (1.80–1.71)	34.2–1.71 (1.80–1.71)	17.2–1.72 (1.81–1.72)
No. of observed reflections	170432	169410	166568
No. of unique reflections	13025	12958	12876
Acentric	10904	10855210	10760
Centric	2121	2103	2116
Signal-to-noise ratio $I/\sigma(I)$	7.6 (2.6)	6.6 (5.6)	8.1 (4.0)
Completeness (%)	100.0 (100.0)	99.8 (99.9)	99.5 (96.8)
Multiplicity	12.7 (11.4)	12.9 (11.0)	12.5 (8.9)
$R_{\rm sym}$ (%)	6.5 (28.4)	6.5 (12.5)	5.9 (17.3)
$R_{\rm ano}$ (%)	3.5 (9.2)	5.9 (10.5)	7.1 (14.5)

2.3. Determination of the gadolinium positions

Anomalous difference Patterson maps were calculated with the program *FFT* (Collaborative Computational Project, Number 4) using data in the 10.0–1.72 Å resolution range. Large anomalous differences, greater than four times the root-mean-square deviation (r.m.s.) of the difference distribution, were rejected. Related $w = 1/4$ and $w = 1/2$ Harker sections of the anomalous difference Patterson maps of the three derivatives are shown in Fig. 2: they exhibit noticeable differences in peak intensities as well as in the number of peaks. The anomalous scatterers were located from these maps using the program *RSPS* (Knight, 2000). For both the 100 and 50 m*M* Gd-HPDO3A derivatives, two sites were found and were assigned to two Gd atoms. Only one of these two sites was

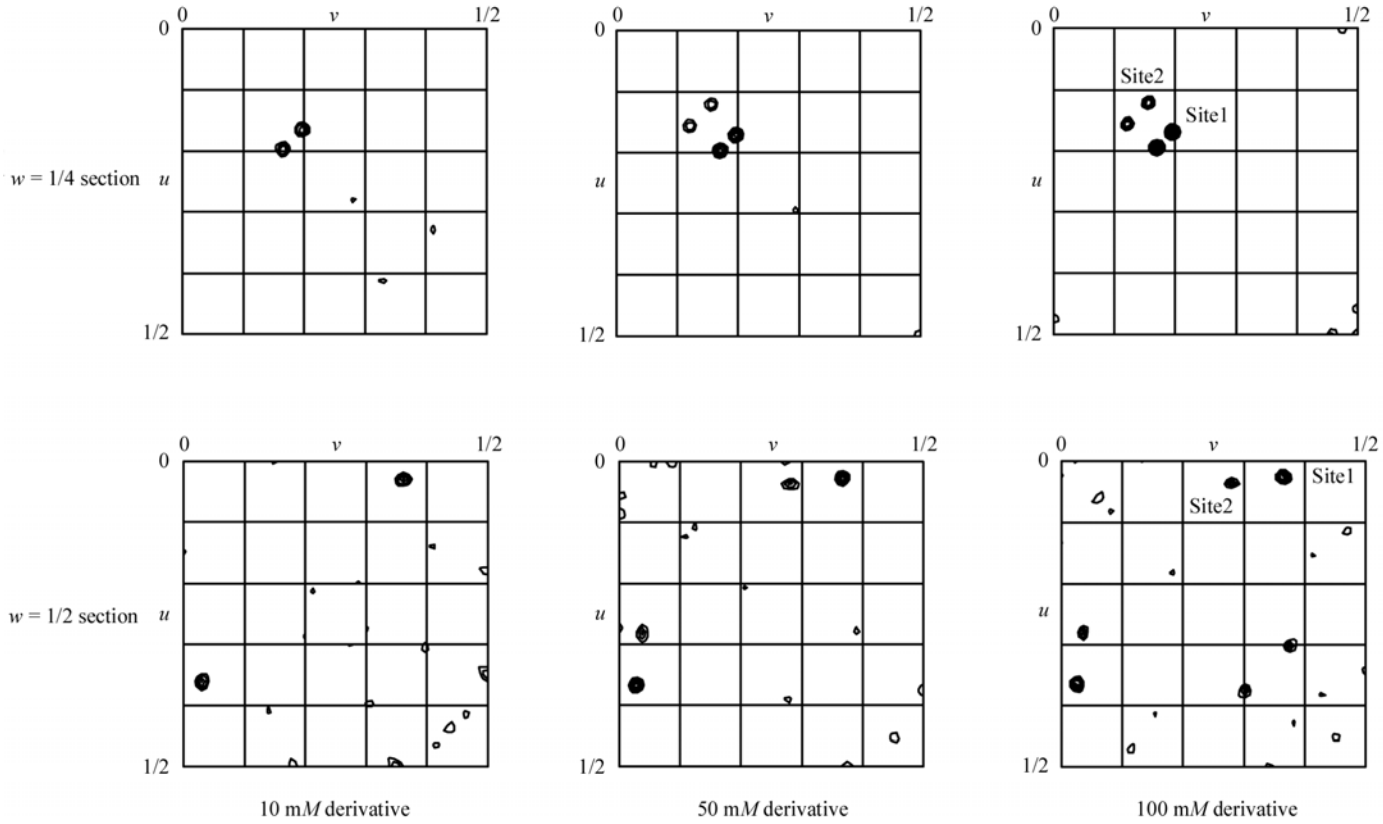

Figure 2
Harker sections of anomalous difference Patterson maps for the 10, 50 and 100 m*M* derivatives. Levels are contoured in 1σ steps starting at 2σ.

found for the 10 m*M* derivative. The coordinates of these sites were used without further refinement to solve the structure.

2.4. Phase determination

Several successful procedures have been proposed for breaking the SAD phase ambiguity.

In the context of direct methods, Hauptman (1982) and Giacovazzo (1983) have derived relations based on three-phase structure invariants to break the SAD phase ambiguity. Fan *et al.* (1984) have designed direct phasing procedures, also based on three-phase structure invariants, for the case of known positions of the scatterers. These procedures developed by Fan & Gu (1985) have been successfully tested on avian pancreatic polypeptide (Fan *et al.*, 1990), on core streptavidin (Sha *et al.*, 1995) and azurin II (Zheng *et al.*, 1996). They were implemented in the program *OASIS* (Hao *et al.*, 2000), which has been used to solve the structure of the copper protein rusticyanin (Harvey *et al.*, 1998). *OASIS* is now part of the *CCP*4 suite. In the present study, since the positions of the Gd atoms were known from the anomalous difference Patterson map, this program was perfectly adapted to phase the anomalous data.

The cell content was calculated using the atomic content of eight protein molecules and assuming two Gd-HPDO3A molecules per asymmetric unit for the 100 and 50 m*M* derivatives and only one Gd-HPDO3A molecule per asymmetric unit for the 10 m*M* derivative. The gadolinium f'' value was set to 12.0 e^-.

Both hands for anomalous scatterer positions were tested using one cycle of *OASIS* phase determination for each hand of each derivative. Results on the 100 m*M* derivative are indicated in Table 2. As anticipated, the correct solution cannot be deduced from the figures of merit obtained after *OASIS* only (FOM = 0.700 for both hands), but can be stated from the comparison of both FOM and free *R* values after density modification using the program *DM* (Cowtan & Main, 1996).

As shown in Figs. 3(*a*) and 3(*c*), the correct solution already leads to interpretable maps for both the 50 and 100 m*M* derivatives, although the connectivity in the 100 m*M* *OASIS* map is slightly better than in the 50 m*M* map (the displayed region, which corresponds to Tyr53, is representative of the map quality but is not the best region). The *OASIS* map obtained for the 10 m*M* derivative shows peaks of electron density that could be assigned to protein, even though the connectivity between these peaks is rather poor.

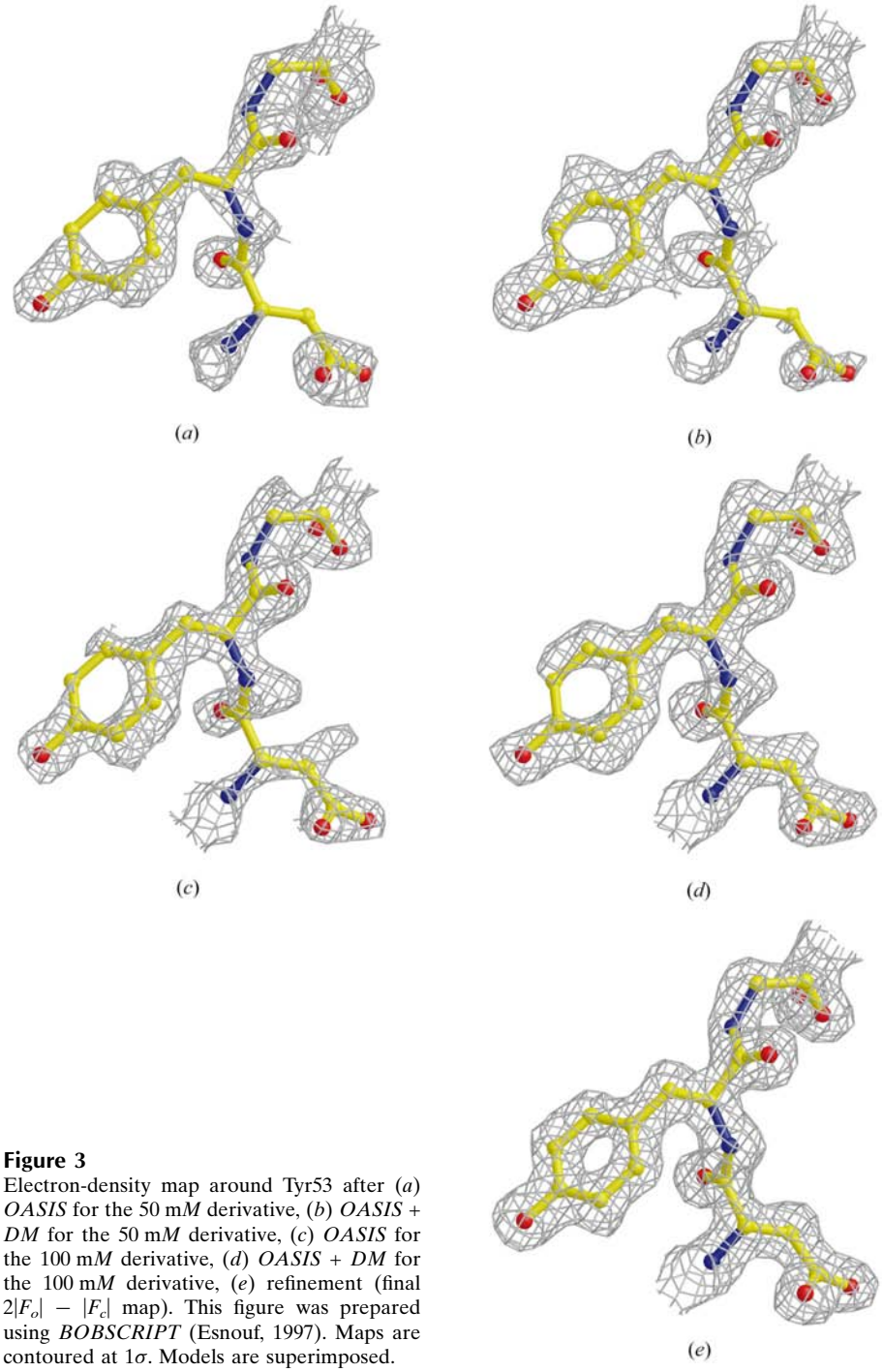

Figure 3
Electron-density map around Tyr53 after (*a*) *OASIS* for the 50 m*M* derivative, (*b*) *OASIS* + *DM* for the 50 m*M* derivative, (*c*) *OASIS* for the 100 m*M* derivative, (*d*) *OASIS* + *DM* for the 100 m*M* derivative, (*e*) refinement (final $2|F_o| - |F_c|$ map). This figure was prepared using *BOBSCRIPT* (Esnouf, 1997). Maps are contoured at 1σ. Models are superimposed.

2.5. Phase improvement

The phases obtained from the *OASIS* program were improved using the program *DM* for density modification over the whole resolution range of each data set. The solvent content was estimated as 0.27 by the program *TRUN-*

CATE. Table 3 lists the number of cycles of histogram matching and solvent flattening as determined automatically by *DM*, the final mean figures of merit as well as the final free *R* values. The electron-density maps were examined using the program *O* (Jones *et al.*, 1991). Only the correct hand leads to an interpretable electron-density map.

The FOM and the free *R* values obtained after *DM* for the 50 m*M* derivative are similar to those obtained for the 100 m*M* derivative, whereas they are clearly not as good for the 10 m*M* derivative. This was confirmed after electron-density map inspection. Whereas the maps from the 10 m*M* derivative could not be interpreted, the quality of both the maps from the 100 m*M* derivative (Fig. 3*d*) as well as from the 50 m*M* derivative (Fig. 3*b*) allows direct model building.

As pointed out by Terwilliger (1999), the effectiveness of solvent flattening decreases considerably at low solvent content, which may explain why when starting from phases with a rather low FOM for the 10 m*M* derivative, *DM* was unable to lead finally to an interpretable electron-density map.

2.6. Refinement

The structure was refined using the 100 m*M* Gd-HPDO3A data set. The maximum-likelihood target function implemented in the program *CNS* (Brunger *et al.*, 1998) was utilized. The scripts were set up to automatically compute a cross-validated σ_A estimate. For all cycles of slow-cooling refinement and Powell minimization, the weight between the X-ray refinement target and the geometric energy function was set to 1.0. Data from the whole 30–1.7 Å resolution range were used without any σ cutoff. The anomalous signal was taken into account. Corrections for bulk solvent and for data anisotropy were applied. In order to monitor the free residual (Brünger, 1992), 10% of the reflections were excluded from the refinement.

The atomic coordinates from the Protein Data Bank corresponding to the tetragonal form at 298 K (entry 6lyt; Young *et al.*, 1994) were used as a starting point. A first rigid-body refinement of this model was followed by a second refinement where both Gd atoms were inserted as ions with full occupancy. The WATER_PICK script from *CNS* was used to automatically locate water molecules. These molecules were retained if they formed stereochemically reasonable hydrogen bonds (distances to closest atoms in the range 2.6–3.5 Å), if their densities in the $2|F_o| - |F_c|$ map were above 1.0 r.m.s. deviation and if they corresponded to peaks greater than 3.0 r.m.s. in the difference density map $\big||F_o| - |F_c|\big|$. About 150 water molecules were located after two steps of this procedure. The occupancies of the gadolinium sites were then refined using the QGROUP script from *CNS*.

The insertion of the ligand HPDO3A around the Gd atoms has been the main difficulty in this refinement. The model used for Gd-HPDO3A was obtained from the refined atomic coordinates of the structure of the complex (Kumar *et al.*, 1994). The electron density corresponding to the ligand was clearly visible, but as HPDO3A is a fairly symmetrical molecule (Fig. 1) with just one arm differing slightly from the others, it was quite difficult to determine its orientation. The two HPDO3A molecules were inserted manually in the electron density and water molecules previously assigned to this

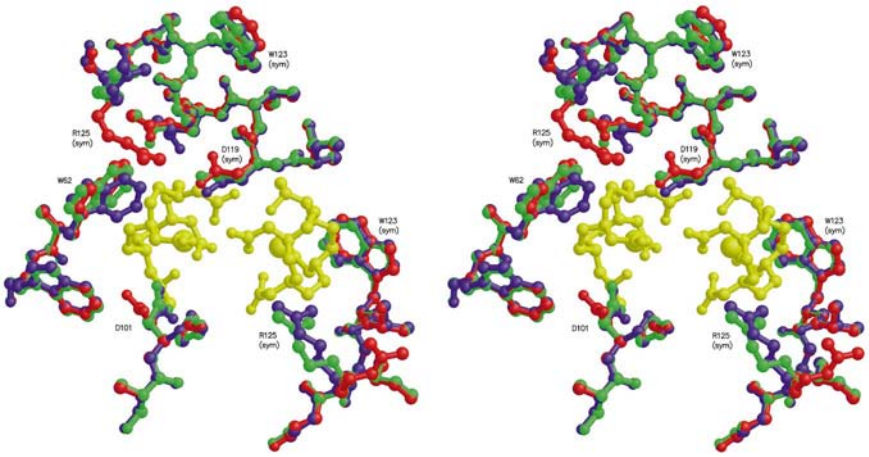

Figure 4
Stereoview of the two Gd-HPDO3A sites (in yellow) in the hen egg-white lysozyme structure determined in this study (in red). For comparison the native HEWL structure observed by Vaney *et al.* (PDB code 193l) is shown in blue and the HEWL structure with a substrate analogue determined by Maenaka *et al.* (PDB code 1lzb) is shown in green. Displacements of Trp62, Asp101 and Arg125 can be observed.

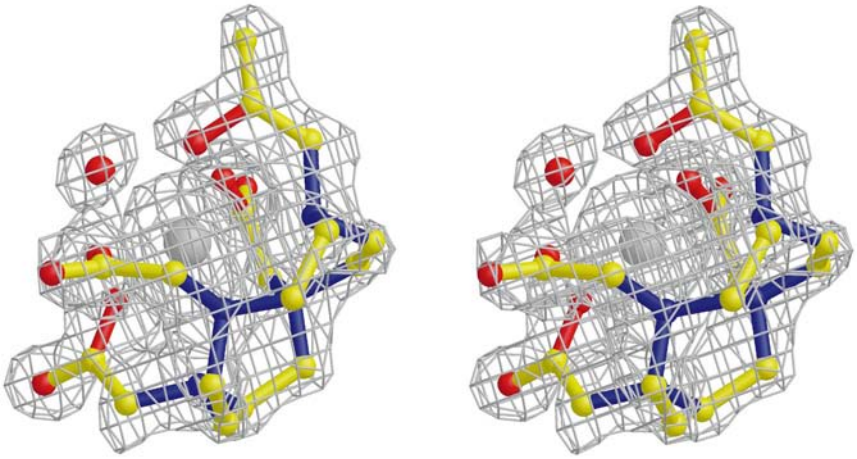

Figure 5
Stereoview of the final $2|F_o| - |F_c|$ electron-density map around Gd-HPDO3A (site Gd 1) and its ninth ligand, water molecule Wat152. The map is contoured at the 1σ level.

Table 2
Statistics of phase determination by *OASIS* and phase improvement by *DM* using the 100 m*M* Gd-HPDO3A lysozyme derivative data set for each possible set of positions of the anomalous scatterers.

Space group	Anomalous scatterer positions	FOM after *OASIS*	No. of *DM* cycles	FOM after *DM*	*DM* free *R* value
$P4_32_12$	$(x_1, y_1, z_1), (x_2, y_2, z_2)$	0.700	12	0.836	0.367
$P4_12_12$	$(\overline{x}_1, \overline{y}_1, \overline{z}_1), (\overline{x}_2.\overline{y}_2, \overline{z}_2)$	0.700	18	0.750	0.497

Table 3
Statistics of phase determination by *OASIS* and phase improvement by *DM* for the three Gd-HPDO3A lysozyme derivative data sets.

[Gd-HPDO3A] (m*M*)	FOM after *OASIS*	No. of *DM* cycles	FOM after *DM*	*DM* free *R* value
100	0.700	12	0.836	0.367
50	0.719	11	0.838	0.374
10	0.596	10	0.772	0.394

density were removed. At this stage, the Gd-HPDO3A complex was considered to be formed of two distinct parts, the ligand and the gadolinium ion, whose occupancies were forced to be identical. After three cycles of Powell minimization, the electron-density map around the Gd atoms was still not well defined.

A few protein side chains were adjusted manually and new water molecules were added. Two chloride ions were identified. The two ligand molecules were removed from the model and water molecules were placed manually around the Gd atoms in order to mimic the ligand molecule in the electron density. Peaks higher than 2.0 r.m.s. in the $2|F_o| - |F_c|$ map and 3.0 r.m.s. in the $|F_o| - |F_c|$ map were assigned to either O or N atoms of the HPDO3A ligand. In order to assign C atoms, these levels were decreased to 1.5 and 2.3 r.m.s., respectively. The refinement of the ligand atoms was achieved using different parameter values in the water-parameter files. During the refinement of O and N atoms of the ligand molecules, the van der Waals radii for oxygen and for hydrogen were set to 1.7 Å and 1.1 Å, respectively, in the water-parameter file. When C atoms were refined, these values were set to 1.5 and 1.0 Å, respectively. Each addition of new atoms was followed by a Powell minimization and was checked by inspecting the resulting electron-density map. Finally, the orientation of each HPDO3A ligand molecule appeared clearly, allowing the two ligand molecules to be positioned unambiguously. In order to allow four O atoms [labelled O(1), O(3), O(5) and O(7) in Fig. 1] to approach the Gd atom, the gadolinium van der Waals radius was set to zero in the ion-parameter file. Two additional chloride ions were found. The four chloride sites finally found correspond to those observed by Lim *et al.* (1998) and Dauter *et al.* (1999). Refinement was ended with several cycles of Powell minimization including individual thermal *B*-factor refinement. The final *R* factor and $R_{\rm free}$ factor are 18.0 and 21.0%, respectively. Final refinement parameters are given in Table 4.

Table 4
Final statistics for the 100 m*M* Gd-HPDO3A lysozyme derivative refined structure.

Weighted r.m.s. deviations from ideality	
Bond lengths (Å)	0.009
Bond angles (°)	1.5
Improper angles (°)	0.83
Dihedral angles (°)	23.2
Average thermal *B* factors (Å^2)	
Main-chain protein atoms	12.8
Side-chain protein atoms	13.8
Overall protein atoms	13.3
Water molecules	24.0
HPDO3A	18.1
Gd atoms	12.0
Cl^- ions	20.3

3. Results and discussion

3.1. General description of the refined lysozyme structure

A total of 1017 non-H protein atoms, 154 water molecules, four chloride ions, two Gd atoms and two HPDO3A ligand molecules have been identified in the structure. The quality of the protein structure has been checked using *PROCHECK* (Laskowski *et al.*, 1993) and PDBSUBMISSION from *CNS*. All non-glycine residues are localized in the favoured and allowed regions of the Ramachandran diagram (Ramachandran & Sasisekharan, 1968) of the main-chain torsion angles φ and ψ. The standard deviation on coordinates estimated from the Luzzati plot (Luzzati, 1952) is 0.17 Å.

Five residues (Lys1, Asn77, Ser85, Ser86 and Val109) display double conformations in the electron-density map. The occupancy of each conformation has been set to 50%.

Disordered conformations are observed for four side chains at the surface of the protein, leading to poorly defined electron density around a few atoms in residues Arg21, Arg45, Arg61 and Arg73.

Some 3σ residual peaks, either positive or negative, in the $|F_o| - |F_c|$ map, with corresponding peaks in the $2|F_o| - |F_c|$ electron-density map, were observed for some residues (Lys13, Thr47, Asn59, Arg68, Gly71, Asp87, Asn93, Asn103, Arg125, Arg128 and Leu129). A positive 3σ $|F_o| - |F_c|$ residual peak was also observed for one disulfide bridge (Cys6–Cys127).

Some extra density in the $|F_o| - |F_c|$ map close to the Gd atoms which were flanked by two opposite negative peaks as well as two opposite positive peaks was interpreted as arising from anisotropic displacements of the Gd atoms. This anisotropy could not be modelled with the current version of *CNS*, which cannot refine anisotropic *B* factors.

No attempt was made to refine the occupancies of the four identified chloride sites, which correspond to the first four chloride sites found by Dauter *et al.* (1999). As described by these authors, an anomalous Fourier synthesis was computed in order to find the four remaining chloride sites. No peak could be detected at the expected positions. This may be a consequence of the strong anomalous signal from the two Gd atoms, which introduces noise in the anomalous Fourier

Table 5
Environment and refined *B* factors for Cl atoms of the 100 m*M* Gd-HPDO3A lysozyme derivative.

Cl atoms	1	2	3	4
Neighbouring residues	Tyr53, Asn113, Val109, Ala110, Arg114	Gly26, Ser24, Val120, Gln121	Gly67, Thr69	Asp87, Ile88
B factor (Å^2)	13.8	14.4	23.3	29.6

synthesis which is higher than the weak anomalous signal from the chloride ions.

The refined *B* factors of the four chloride ions are listed in Table 5. Their hierarchy is the same as found by Dauter *et al.* (1999), but their values are slightly different: they are lower for Cl atoms 1 and 2 and higher for Cl atoms 3 and 4. The neighbouring residues of each Cl atom are also indicated in Table 5.

3.2. Gd-HPDO3A structure and binding mode

From the known structure of the Gd-HPDO3A complex (Kumar *et al.*, 1994), it can be seen that this complex exhibits a double character: the HPDO3A macrocycle is rather hydrophobic, whereas, on the opposite side, the gadolinium environment is very hydrophilic with a water molecule as a ninth ligand.

In the 100 m*M* Gd-HPDO3A HEWL derivative crystal the two gadolinium complex molecules are located in the lysozyme active site with a distance of 6 Å between Gd atoms. An overall view of the binding sites of Gd-HPDO3A to lysozyme in the crystal structure is shown in Fig. 4. Both HPDO3A macrocycles are located in hydrophobic environments (close to Trp62 for site Gd1 and to Trp123 for site Gd2). On the hydrophilic side of both Gd-HPDO3A complexes, water molecules play the role of ninth ligand to the gadolinium. The Gd-HPDO3A complex in site Gd1 with water molecule Wat152 as the ninth ligand can be seen in Fig. 5.

Fig. 4 shows the HEWL native structure obtained by Vaney *et al.* (1996), the HEWL structure with a substrate analogue (Maenaka *et al.*, 1995) and the Gd-HPDO3A derivative structure in the active-site region where both Gd-HPDO3A complexes are located. In the Gd-HPDO3A derivative structure, Trp62 has moved to stack against the HPDO3A macrocycle corresponding to gadolinium site Gd1. Trp62 is a flexible residue that was found to be disordered by NMR (Smith *et al.*, 1993) as well as X-ray (Vaney *et al.*, 1996) studies of the native structure. When the sugar inhibitor is bound to lysozyme, Trp62 is displaced by hydrophobic interactions and is less disordered (Kurachi *et al.*, 1976; Perkins *et al.*, 1978; Strynadka & James, 1991; Turner & Howell, 1995). A co-operative movement of Trp62 and Asp101 has been observed for the recognition of both the substrate and the reaction product (Maenaka *et al.*, 1998). In the Gd-HPDO3A derivative structure the $2|F_o| - |F_c|$ electron-density map contoured at 1σ is well defined for all atoms of Trp62, which is in agreement with the hydrophobic interaction. Instead of making a hydrogen bond, Asp101 is pushed away.

A similar situation is observed for the second Gd-HPDO3A molecule. Even if Trp123 does not directly stack the HPDO3A macrocycle, the environment is mainly hydrophobic. The side chain of Arg125 is flipped like Asp101 for site Gd1, disallowing hydrogen bonding.

As can be shown in Fig. 6, on the hydrophilic side of the two HPDO3A ligand molecules a hydrogen-bonding network takes place. This network involves the HPDO3A carboxyl groups, six water molecules (including the two water molecules acting as ninth gadolinium ligand), Val120 N, Lys33 NH, Asn103 ND2 and the carboxyl group of Asp119. Arg125 NH2 makes a bifurcated hydrogen bond with Gln121 OE1 and Asp119 OD2.

The combination of the hydrophobic stacking and hydrogen bonding is consistent with the high values obtained for the two Gd-HPDO3A molecule occupancies of the 100 m*M* derivative.

Table 6
Refined occupancies and thermal *B* factors of the Gd atoms in the three HPDO3A lysozyme derivatives.

Correlation coefficients between the experimental maps for the three derivatives and the refined 100 m*M* model map.

	10 m*M*	50 m*M*	100 m*M*
Gd1			
Occupancy	0.30	0.64	0.81
B factor (Å^2)	15.6	15.6	15.3
Gd2			
Occupancy	—	0.55	0.73
B factor (Å^2)	—	16.1	15.9
Map correlation coefficients			
After *OASIS*	0.277	0.518	0.624
After *DM*	0.294	0.629	0.775

3.3. Gadolinium-site occupancies and phasing power of the different derivatives

The structures of both the 50 and the 10 m*M* derivatives were refined using similar conditions to those used for the refinement of the 100 m*M* derivative structure. Bias in the comparison between $R_{\rm free}$ was avoided by selecting for the free reflection set of a given derivative all reflections of the 100 m*M* free set present in the corresponding data set. The 100 m*M* derivative final structure, replaced close to the origin to reduce shifts introduced by differences in unit-cell parameters, was used as a starting model for subsequent refinement of each derivative structure: a rigid-body refinement followed by a step of slow cooling and individual *B*-factor refinements. Thermal *B* factors for the Gd atoms and for the HPDO3A ligand were arbitrarily set to the final values obtained for the corresponding atoms in the 100 m*M* model. The gadolinium and the HPDO3A-ligand occupancies were refined using the *CNS* script QGROUP. The final Gd-site occupancies for the three derivatives are listed in Table 6.

The site occupancies for both the 50 and the 100 m*M* derivatives are higher than those reported for a 50 m*M* gadolinium acetate lysozyme derivative (Perkins *et al.*, 1979), where two different sites have been found with occupancies of 0.38 and 0.24, respectively. Nagem *et al.* (2001) report three major sites in a lysozyme derivative obtained by short cryo-soaking with 250 m*M* gadolinium chloride, but no refined Gd-site occupancy is given.

As can be anticipated from the binding mode of Gd-HPDO3A in the crystal, the higher the Gd-HPDO3A concentration in the crystallization solution, the higher the site occupancy in the derivative crystal. This leads to fairly high occupancies and consequently to a high phasing power for both the 50 and 100 m*M* derivatives.

The correlation coefficients between experimental maps for the three derivatives and the refined 100 m*M* model map have been calculated by *OVERLAPMAP* (Collaborative Computational Project, Number 4, 1994) and are reported in Table 6. The high values for the map correlation coefficients obtained after *OASIS* and after *DM* show the high phasing power of both derivatives, particularly the 100 m*M* derivative.

4. Conclusions

The development of the selenomethionine (SeMet) based multiwavelength anomalous dispersion (MAD) method in protein phasing has overcome the cumbersome process of heavy-atom screening (Hendrickson, 1991). The primary advantage of SeMet phasing over the conventional heavy-atom screening method derives from the ability of the SeMet method to provide heavy-atom derivatives avoiding uncertain and lengthy screening processes. This method provides full site occupancy for the Se anomalous atoms. Nevertheless, it is not fully general and when it fails other derivatives are needed. As shown in this work, Gd derivatives are easily obtained using Gd-HPDO3A either by co-crystallization or by soaking. The crystallization conditions for the derivative are unmodified with respect to the native crystals even with very high concentrations (up to 500 m*M*) of the Gd complex.

With such derivatives, the phasing method that combines accurately measured SAD data with density modification is particularly interesting for the structure determination of biological molecules. This approach is applicable to many types of anomalous scatterers. Using this method, Dauter and co-workers (Dauter *et al.*, 1999, 2000; Nagem *et al.*, 2001), mimicking Cu $K\alpha$ radiation at a synchrotron source, have shown that the anomalous signal produced by S and Cl atoms found in lysozyme crystal or halide ions introduced into the crystal allows the phase problem to be solved from data collected away from the absorption edge. This work demonstrates that it is possible to determine the Gd-HPDO3A lysozyme derivative structure from a SAD experiment using Cu $K\alpha$ radiation from a conventional X-ray rotating anode, by making use of the fairly strong anomalous signal of Gd atoms ($f'' = 12.0$) at this wavelength. High-quality electron-density maps have been derived from SAD data collected in the laboratory. Owing to the moderate average signal-to-noise ratio (∼8; see Table 1) of this data set, it is likely that the phasing power of data collected with synchrotron radiation at the same wavelength with a higher signal-to-noise ratio would be increased. For SAD data collected with synchrotron radiation at the Gd L_{III} absorption edge (1.71 Å), where f'' can reach 28–30 e^-, the phasing power should be enhanced again. Finally, the MAD method at the Gd L_{III} absorption edge could be used for crystals with lower Gd-site occupancies.

Thanks to its charge neutrality, Gd-HPDO3A can be used at high concentrations in either crystallizing or soaking solutions leading to high site occupancies. Large anomalous effects can thus be obtained, especially if data are collected at the Gd L_{III} absorption edge. To take full advantage of this method, the absorption and the fluorescence should be reduced as far as possible. This is obtained by washing the crystal for a short time in a Gd-free cryosolution while keeping high Gd-site occupancies in the crystal.

Further studies on derivatives obtained with Gd-HPDO3A and other proteins, crystallizing either with different salts or with PEGs, are in progress to generalize the use of

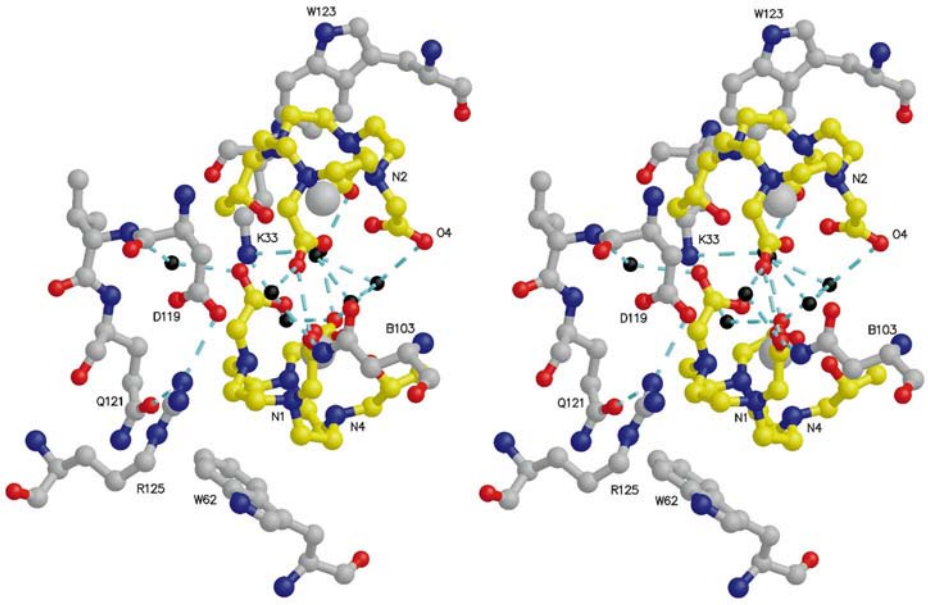

Figure 6
Stereoview of the environment of the two Gd-HPDO3A molecules. Distances between Trp62 and the first HPDO3A macrocycle and between Trp123 and the second HPDO3A macrocycle are about 4 and 3.5 Å, respectively. Hydrogen-bond distances less than 3.3 Å are represented. Water molecules are drawn in black for clarity.

such derivatives to determine the structure of biological macromolecules by SAD or MAD experiments.

We thank Bracco s.p.a., Milan, Italy for kindly providing a sample of Gd-HPDO3A.

References

Biou, V., Shu, F. & Ramakrishnan, V. (1995). *EMBO J.* **14**, 4056–4064.
Brodersen, D. E., de La Fortelle, E., Vonrhein, C., Bricogne, G., Nyborg, J. & Kjeldgaard, M. (2000). *Acta Cryst.* D**56**, 431–441.
Brünger, A. T. (1992). *Nature (London)*, **355**, 472–475.
Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* D**54**, 905–921.
Burkhart, B. M., Li, N., Langs, D. A., Pangborn, W. A. & Duax, W. L. (1998). *Proc. Natl. Acad. Sci. USA*, **95**, 12950–12955.
Chen, L. Q., Rose, J. P., Breslow, E., Yang, D., Chang, W. R., Furey, W. F. Jr, Sax, M. & Wang, B. C. (1991). *Proc. Natl Acad. Sci. USA*, **88**, 4240–4244.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Cowtan, K. D. & Main, P. (1996). *Acta Cryst.* D**52**, 43–48.
Dauter, Z. & Dauter, M. (1999). *J. Mol. Biol.* **289**, 93–101.
Dauter, Z. & Dauter, M. (2001). *Structure*, **9**, R21–R26.
Dauter, Z., Dauter, M., de La Fortelle, E., Bricogne, G. & Sheldrick, G. M. (1999). *J. Mol. Biol.* **289**, 83–92.
Dauter, Z., Dauter, M. & Rajashankar, K. R. (2000). *Acta Cryst.* D**56**, 232–237.
Dauter, Z., Li, M. & Wlodawer, A. (2001). *Acta Cryst.* D**57**, 239–249.
Devedjiev, Y., Dauter, Z., Kuznetsov, S. R., Jones, T. L. & Derewenda, Z. S. (2000). *Structure Fold. Des.* **8**, 1137–1146.
Esnouf, R. M. (1997). *J. Mol. Graph.* **15**, 133–138.
Fan, H. F. & Gu, Y. X. (1985). *Acta Cryst.* A**41**, 280–284.
Fan, H. F., Han, F. S., Qian, J. Z. & Yao, J. X. (1984). *Acta Cryst.* A**40**, 489–495.
Fan, H. F., Hao, Q., Gu, Y. X., Qian, J. Z., Zheng, C. D. & Hengming, K. (1990). *Acta Cryst.* A**46**, 935–939.
Furey, W. F., Robbins, A. H., Clancy, L. L., Winge, D. R., Wang, B. C. & Stout, C. D. (1986). *Science*, **231**, 704–710.
Giacovazzo, C. (1983). *Acta Cryst.* A**39**, 585–592.
Hao, Q., Gu, Y. X., Zheng, C. D. & Fan, H. F. (2000). *J. Appl. Cryst.* **33**, 980–981.
Harvey, I., Hao, Q., Duke, E. M., Ingledew, W. J. & Hasnain, S. S. (1998). *Acta Cryst.* D**54**, 629–635.
Hauptman, H. (1982). *Acta Cryst.* A**38**, 632–641.
Hendrickson, W. A. (1991). *Science*, **254**, 51–58.
Hendrickson, W. C. & Teeter, M. M. (1981). *Nature (London)*, **290**, 107–113.
Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* A**47**, 110–119.
Kabsch, W. (1988). *J. Appl. Cryst.* **21**, 916–924.
Knight, S. D. (2000). *Acta Cryst.* D**56**, 42–47.
Kraulis, P. (1991). *J. Appl. Cryst.* **24**, 946–950.
Kumar, K., Chang, C. A., Francesconi, L. C., Dischino, D. D., Malley, M. F., Gougoutas, J. Z. & Tweedle, M. F. (1994). *Inorg. Chem.* **33**, 3567–3575.
Kurachi, K., Sieker, L. C. & Jensen, L. H. (1976). *J. Mol. Biol.* **101**, 11–24.
Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.
Lim, K., Nadarajah, A., Forsythe, E. L. & Pusey, M. L. (1998). *Acta Cryst.* D**54**, 899–904.
Luzzati, V. (1952). *Acta Cryst.* **5**, 802–810.
Maenaka, K., Matsushima, M., Kawai, G., Kidera, A., Watanabe, K., Kuroki, R. & Kumagai, I. (1998). *Biochem. J.* **333**, 71–76.
Maenaka, K., Matsushima, M., Song, H., Sunada, F., Watanabe, K. & Kumagai, I. (1995). *J. Mol. Biol.* **247**, 281–293.
Merritt, E. A. & Bacon, D. J. (1997). *Methods Enzymol.* **277**, 505–524.
Nagem, R. A., Dauter, Z. & Polikarpov, I. (2001). *Acta Cryst.* D**57**, 996–1002.
Perkins, S. J., Johnson, L. N., Machin, P. A. & Phillips, D. C. (1978). *Biochem. J.* **173**, 607–616.
Perkins, S. J., Johnson, L. N., Machin, P. A. & Phillips, D. C. (1979). *Biochem. J.* **181**, 21–36.
Ramachandran, G. N. & Sasisekharan, V. (1968). *Adv. Protein Chem.* **23**, 283–438.
Ramin, M., Shepard, W., Fourme, R. & Kahn, R. (1999). *Acta Cryst.* D**55**, 157–167.
Rodgers, D. W. (1994). *Structure*, **2**, 1135–1140.
Royer, W. E. Jr, Hendrickson, W. A. & Chiancone, E. (1989). *J. Biol. Chem.* **264**, 21052–21061.
Sabini, E., Schubert, H., Murshudov, G., Wilson, K. S., Siika-Aho, M. & Penttila, M. (2000). *Acta Cryst.* D**56**, 3–13.
Sha, B. D., Liu, S. P., Gu, Y. X., Fan, H. F., Hengming, K., Yao, J. X. & Woolfson, M. M. (1995). *Acta Cryst.* D**51**, 342–346.
Smith, L. J., Sutcliffe, M. J., Redfield, C. & Dobson, C. M. (1993). *J. Mol. Biol.* **229**, 930–944.
Strynadka, N. C. & James, M. N. (1991). *J. Mol. Biol.* **220**, 401–424.
Terwilliger, T. C. (1999). *Acta Cryst.* D**55**, 1863–1871.
Turner, M. A. & Howell, P. L. (1995). *Protein Sci.* **4**, 442–449.
Vaney, M. C., Maignan, S., Ries-Kautt, M. & Ducruix, A. (1996). *Acta Cryst.* D**52**, 505–517.
Wallace, B. A., Hendrickson, W. A. & Ravikumar, K. (1990). *Acta Cryst.* B**46**, 440–446.
Wu, C. K., Dailey, H. A., Rose, J. P., Burden, A., Sellers, V. M. & Wang, B. C. (2001). *Nature Struct. Biol.* **8**, 156–160.
Young, A. C., Tilton, R. F. & Dewan, J. C. (1994). *J. Mol. Biol.* **235**, 302–317.
Zheng, X. F., Fan, H. F., Hao, Q., Dodd, F. E. & Hasnain, S. S. (1996). *Acta Cryst.* D**52**, 937–941.