Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

A new class of gadolinium complexes employed to obtain high-phasing-power heavy-atom derivatives: results from SAD experiments with hen egg-white lysozyme and urate oxidase from *Aspergillus flavus*

É. Girard,[a] M. Stelter,[a] P. L. Anelli,[b] J. Vicat[a] and R. Kahn[a]*

[a]Laboratoire de Cristallographie Macromoléculaire, Institut de Biologie Structurale J.-P. Ebel CEA–CNRS–UJF, 41 Rue Jules Horowitz, 38027 Grenoble CEDEX 1, France, and [b]Bracco Imaging spa, Via E. Folli 50, 20134 Milan, Italy

Correspondence e-mail: kahn@ibs.fr

Received 26 July 2002
Accepted 5 November 2002

Seven gadolinium complexes are shown to be excellent compounds for the preparation of heavy-atom derivatives for macromolecular crystallography projects. *De novo* phasing has been carried out using single-wavelength anomalous diffraction (SAD) on a series of gadolinium-derivative crystals of two proteins: hen egg-white lysozyme and urate oxidase from *Aspergillus flavus*. Lysozyme derivative crystals were obtained by co-crystallizing the protein with the corresponding gadolinium complex at a concentration of 100 m*M*. Diffraction data were collected to a resolution of 1.7 Å using Cu $K\alpha$ radiation from a rotating-anode generator, making use of the high anomalous signal of gadolinium at this wavelength. Urate oxidase derivative crystals were obtained by soaking native crystals in 100 m*M* gadolinium complex solutions. Diffraction data were collected to a resolution close to 3 Å using X-rays at the Gd $L_{\rm III}$ absorption edge, taking advantage of the sharp white line on that edge. For all urate oxidase derivative crystals and three of the lysozyme crystals, SAD phasing led to electron-density maps of very high quality, allowing unambiguous chain tracing. From this study, the binding effectiveness of the gadolinium complexes seems to be related to the nature of the precipitant used for crystallization. These gadolinium complexes represent a new class of high-phasing-power heavy-atom derivatives that may be used for high-throughput structure-determination projects.

1. Introduction

The current emphasis on high-throughput biological structure determination has led to the development of methods to prepare heavy-atom derivatives that are more reliable and less disruptive than traditional heavy-atom soaking. The multiple-wavelength anomalous diffraction (MAD) method at the Se *K* absorption edge combined with the incorporation of Se in proteins by replacing methionine with selenomethionine, especially using *Escherichia coli* (Doublié, 1997; Hendrickson *et al.*, 1990), has revolutionized macromolecular crystallography (Hendrickson & Ogata, 1997). Although selenomethionine incorporation can also be achieved in insect (Bellizzi *et al.*, 1999; Chen & Bahl, 1991; McWhirter *et al.*, 1999) and mammalian cells (Lustbader *et al.*, 1995), having to optimize these eukaryotic systems does not facilitate selenomethionine incorporation. Moreover, many eukaryotic proteins cannot be overexpressed and folded properly in prokaryotic systems owing to necessary post-translational modifications such as glycosylation or disulfide-bond formation. Even when the selenomethionine incorporation is successful, the selenomethionine-containing protein may not crystallize as readily as the native protein.

Alternative methods to obtain derivatives for anomalous diffraction experiments have therefore been proposed, including the use of halides (Dauter & Dauter, 1999, 2001; Dauter *et al.*, 2000, 2001; Evans & Bricogne, 2002) and of noble gases in native binding sites (Vitali *et al.*, 1991; Schiltz *et al.*, 1994; Cohen *et al.*, 2001) or in binding sites created by direct mutagenesis (Quillin & Matthews, 2002). The anomalous signal from the intrinsic S atoms has increasingly been used to solve the structure by the SAD method (Hendrickson & Teeter, 1981; Wang, 1985; Dauter *et al.*, 1999; Liu *et al.*, 2000; Bond *et al.*, 2001; Gordon *et al.*, 2001; Micossi *et al.*, 2002). For a general review on the potential of the SAD method, see Dauter *et al.* (2002). Recently, a labelling method based on a thiol-reactive lanthanide chelate has been proposed by Purdy *et al.* (2002) and the traditional heavy-atom preparation has been revisited using short soaking (Sun & Radaev, 2002; Sun *et al.*, 2002). Techniques such as native-gel electrophoresis (Boggon & Shapiro, 2000) or mass spectrometry (Cohen *et al.*, 2000; Sun & Hammer, 2000) have been used to screen rapidly for derivatives.

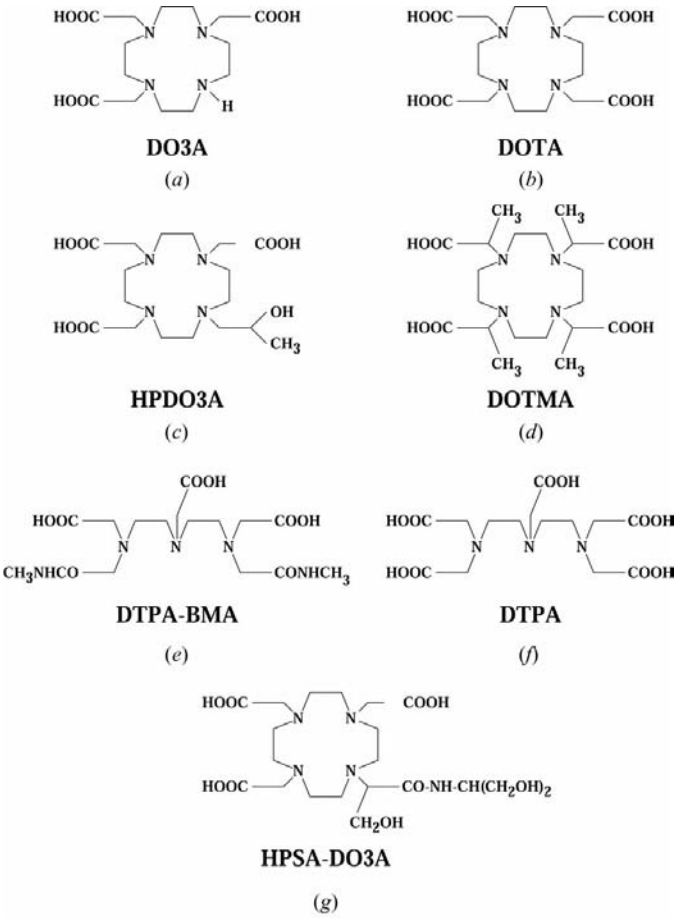

Figure 1
The different ligands in the seven Gd complexes presented in this study. DO3A, 1,4,7,10-tetraazacyclododecane-1,4,7-triacetic acid; DOTA, 1,4,7,10-tetraazacyclododecane-1,4,7,10-tetraacetic acid; HPDO3A, 10-(2-hydroxypropyl)-1,4,7,10-tetraazacyclododecane-1,4,7-triacetic acid; DOTMA, *a*,*a*′,*a*″,*a*‴-tetramethyl-1,4,7,10-tetraazacyclododecane-1,4,7,10-tetraacetic acid; HPSA-DO3A, 10-(2-{[2-hydroxy-1-(hydroxymethyl)ethyl]amino}-1-(hydroxymethyl)-2-oxoethyl)-1,4,7,10-tetraazacyclododecane-1,4,7-triacetic acid; DTPA-BMA, diethylenetriaminepentaacetic acid bismethylamide; DTPA, diethylenetriaminepentaacetic acid.

We propose a new class of Gd derivatives that exploit the high anomalous signal of gadolinium with Cu $K\alpha$ radiation ($f'' = 12\ e^-$) or, when using synchrotron radiation, its very strong anomalous signal at its L_{III} absorption edge ($f'' = 28\ e^-$). The derivatives are obtained using highly soluble lanthanide complexes that can be easily introduced into protein crystals at high concentration by either soaking or co-crystallization. This ability has already been demonstrated for the Gd–HPDO3A derivative of hen egg-white lysozyme described in Girard *et al.* (2002).

This paper describes the use of derivative crystals obtained with these complexes to generate phases from SAD experiments performed on two proteins: hen egg-white lysozyme (HEWL) and urate oxidase from *Aspergillus flavus*. For HEWL derivative crystals the SAD data were collected using Cu $K\alpha$ radiation from a conventional X-ray source, whereas for urate oxidase derivative crystals they were collected with synchrotron radiation at the Gd L_{III} absorption edge.

2. Methods

2.1. Gadolinium complexes

Seven gadolinium complexes were employed in this study, each of them being composed of a ligand that chelates a single Gd^{3+} ion. The different ligands are represented in Fig. 1. These ligands are based on a tetraazacyclododecane macrocycle, except for DTPA and DTPA-BMA, which are acyclic.

The gadolinium complexes of HPDO3A, DO3A, DOTMA and HPSA-DO3A were kindly provided as powders by Bracco Imaging spa, Milan, Italy. For the three other complexes, Gd–DTPA, Gd–DTPA-BMA and Gd–DOTA, the magnetic resonance imaging commercial solutions (MAGNEVIST™, Schering, Germany, OMNISCAN™, Nycomed Inc., Norway and DOTAREM, Guerbet, France, respectively) were used.

The complexes are soluble in water at concentrations higher than 1 *M*, except for the Gd–HPSA-DO3A complex, whose solubility in water is ∼0.3 *M*. The commercial solutions of Gd–DTPA-BMA, Gd–DOTA and Gd–DTPA that were used were provided at a concentration of 0.5 *M*.

The complexes are neutral except for Gd–DOTA and Gd–DOTMA, which carry a single negative global charge, and Gd–DTPA, which bears a double negative charge.

2.2. Crystallization

Hen egg-white lysozyme (HEWL) was purchased from Boehringer and was used without any further purification. Protein solutions were prepared by dissolving 4 mg of lyophilized HEWL in 100 µl of the given gadolinium complex solution at a concentration of 100 m*M*. 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 pH 4.5. Co-crystallized derivative crystals were obtained without further changes to the

Table 1
Data-collection parameters and processing statistics for urate oxidase derivative crystals obtained with the seven Gd complexes.

The diffraction data were collected with synchrotron radiation on the Gd L_{III} absorption edge at the wavelength corresponding to the maximum of f''. Values in parentheses refer to the highest resolution shell.

	Gd–DTPA	Gd–DTPA-BMA	Gd–DO3A	Gd–HPDO3A	Gd–HPSA-DO3A	Gd–DOTA	Gd–DOTMA
Soaking time	20 min	45 min	3.5 d	48 h	30 min	240 min	360 min
Unit-cell parameters							
a (Å)	78.900	76.621	79.064	77.462	77.883	78.664	79.291
b (Å)	95.162	95.215	95.211	95.524	95.099	95.399	95.043
c (Å)	104.087	104.544	104.138	104.539	104.239	104.244	104.009
Resolution range (Å)	25.0–3.14 (3.31–3.14)	25.0–3.14 (3.31–3.14)	20.0–2.90 (3.06–2.90)	25.0–2.92 (3.08–2.92)	25.0–3.00 (3.16–3.00)	25.0–3.00 (3.16–3.00)	25.0–3.00 (3.16–3.00)
No. of observed reflections	44807	43249	59555	53214	48103	49052	49526
No. of unique reflections	7088	6854	8960	8522	7800	8025	7968
Acentric	6090	5911	7801	7430	6782	6954	6909
Centric	998	943	1159	1092	1018	1071	1059
Signal-to-noise ratio $I/\sigma(I)$	17.1 (12.8)	15.2 (12.1)	10.1 (8.5)	17.1 (5.6)	20.1 (16.8)	13.9 (12.8)	18.0 (13.3)
Completeness (%)	99.8 (99.0)	99.8 (99.2)	99.6 (99.4)	99.4 (97.2)	99.6 (95.4)	99.6 (98.6)	99.2 (95.6)
Multiplicity	6.3 (5.4)	6.3 (5.4)	6.6 (6.6)	6.2 (4.5)	6.1 (4.1)	6.1 (4.3)	6.2 (4.1)
R_{sym} (%)	3.3 (4.7)	3.5 (4.8)	5.6 (7.1)	3.2 (8.9)	2.8 (3.6)	4.0 (4.7)	3.1 (4.3)
R_{ano} (%)	3.1 (3.8)	9.6 (9.2)	6.3 (6.5)	4.0 (7.0)	3.4 (3.6)	3.9 (4.6)	5.2 (5.9)

standard crystallization conditions. All HEWL derivative crystals belong to space group $P4_32_12$.

Urate oxidase from *A. flavus* (295 residues) was kindly provided by Sanofi-Synthélabo, France and was used without any further purification. The protein solution at a concentration of 11 mg ml^{-1} contained 1 mg ml^{-1} EDTA and an inhibitor, 8-azaxanthine, at a concentration of 0.3 mg ml^{-1}. Crystallization conditions for the production of native crystals were adapted from Bonneté *et al.* (2001). Native crystals were grown at 288 K using the hanging-drop method with 4–9% PEG 8000 and 50–100 m*M* Tris base buffer pH 8.5. Derivative crystals were prepared by soaking the native crystals in a solution containing 9% PEG 8000, 50–100 m*M* Tris base buffer and 100 m*M* of the given gadolinium complex. As indicated in Table 1, soaking times vary from 20 min to a few days. This is mainly owing to contextual reasons, *i.e.* synchrotron beam-time access. All urate oxidase derivative crystals belong to the same space group, *I*222, as the native crystals.

2.3. Data collection and data processing

In order to remove any remaining gadolinium complex from the liquor surrounding the derivative crystals, the samples were soaked for 30 s in the native solution containing 25–30% PEG 400 as a cryoprotectant. The crystals were cooled at 100 K (Oxford Cryosystems). For Gd–DTPA-BMA, Gd-–DOTA, Gd–DO3A, Gd–DOTMA and Gd–HPSA-DO3A, the quality of the cryocooled HEWL derivative crystals was increased using paraffin oil (Riboldi-Tunnicliffe & Hilgenfeld, 1999). No attempt was made to orient the crystals along any particular crystallographic axis or to measure Bijvoet mates simultaneously.

Data from HEWL derivative crystals were collected on a MAR300 imaging-plate detector using Cu $K\alpha$ radiation from a Rigaku RU-200 rotating-anode X-ray generator equipped with a Supper 7600 double-mirror system. The Gd–DTPA HEWL derivative data were recorded on a MAR345 imaging-plate detector using a Nonius rotating-anode generator equipped with an OSMIC focusing system. In order to obtain high multiplicity, 180° of data were recorded for each derivative crystal. A rotation range limited to 136° was recorded for the Gd–DOTMA derivative.

Data from urate oxidase derivative crystals were collected at the ESRF on the BM30A beamline. The wavelength was chosen at the L_{III} gadolinium absorption edge, at which a large white line is observed (Fig. 2). The wavelength was set to 1.711 Å, which corresponds to the maximum value of f'' ($\sim$28 e$^-$). The data were collected on a MARCCD detector of 165 mm in diameter, with the exception of the Gd–DO3A derivative data, which were collected on a MAR345 imaging-plate detector. Since the shortest crystal-to-detector distance for data sets collected with the MARCCD detector was 132 mm, the resolution of these data sets was limited to approximately 3 Å.

HEWL data were integrated using the program *XDS* (Kabsch, 1988), whilst urate oxidase data were integrated using the program *DENZO* (Otwinowski & Minor, 1997). Integrated intensities for HEWL derivatives as well as for urate oxidase derivatives were scaled using *SCALA* from the *CCP*4 suite of programs (Collaborative Computational Project, Number 4, 1994). A summary of data-collection parameters and processing statistics is given in Table 1 for Gd-derivative crystals of urate oxidase and in Table 2 for Gd-derivative crystals of HEWL.

2.4. SAD phasing

In order to evaluate the ability of the different Gd derivatives to solve the two structures, *de novo* phasing was undertaken on the whole series of derivatives for both proteins. The coordinates of the Gd atoms were found using the program *SHELX* (Sheldrick, 1986, 1998) for HEWL derivative data and using the program *SnB* (Weeks & Miller, 1999) for urate oxidase derivative data. In order to be able to compare the electron-density maps from different derivative crystals of the

Table 2
Data-collection parameters and processing statistics for diffraction data collected with Cu $K\alpha$ radiation from lysozyme derivative crystals obtained with the seven Gd complexes.

Values in parentheses refer to the highest resolution shell.

	Gd–DTPA	Gd–DTPA-BMA	Gd–DO3A	Gd–HPDO3A	Gd–HPSA-DO3A	Gd–DOTA	Gd–DOTMA
Unit-cell parameters							
a (Å)	78.276	78.624	78.715	77.250	78.734	77.321	78.095
c (Å)	37.231	37.356	37.249	38.660	37.214	38.077	37.504
Resolution range (Å)	20.0–1.71 (1.80–1.71)	20.0–1.71 (1.80–1.71)	20.0–1.70 (1.80–1.70)	17.2–1.72 (1.81–1.72)	30.0–1.71 (1.81–1.71)	31.0–1.71 (1.80–1.71)	34.0–1.70 (1.80–1.70)
No. of observed reflections	165609	167433	166661	166568	164169	170962	129961
No. of unique reflections	12920	13108	13184	12876	13026	13019	13074
Acentric	10877	11035	11080	10760	10923	10908	10937
Centric	2043	2073	2104	2116	2103	2111	2137
Signal-to-noise ratio $I/\sigma(I)$	11.0 (4.9)	9.7 (8.1)	7.5 (5.8)	8.1 (4.0)	6.8 (5.6)	5.2 (3.2)	5.2 (1.9)
Completeness (%)	99.4 (99.5)	99.4 (98.6)	99.7 (98.5)	99.5 (96.8)	99.9 (99.9)	99.7 (98.6)	99.5 (96.9)
Multiplicity	12.8 (10.7)	12.7 (10.5)	12.6 (10.0)	12.5 (8.9)	12.6 (10.1)	13.1 (11.6)	9.9 (8.9)
$R_{\rm sym}$ (%)	3.8 (15.6)	6.0 (8.3)	7.6 (11.9)	5.9 (17.3)	6.2 (12.2)	7.4 (22.5)	7.9 (36.0)
$R_{\rm ano}$ (%)	2.6 (6.5)	2.8 (3.8)	5.2 (6.7)	7.1 (14.5)	2.7 (4.9)	4.2 (8.5)	3.6 (12.0)

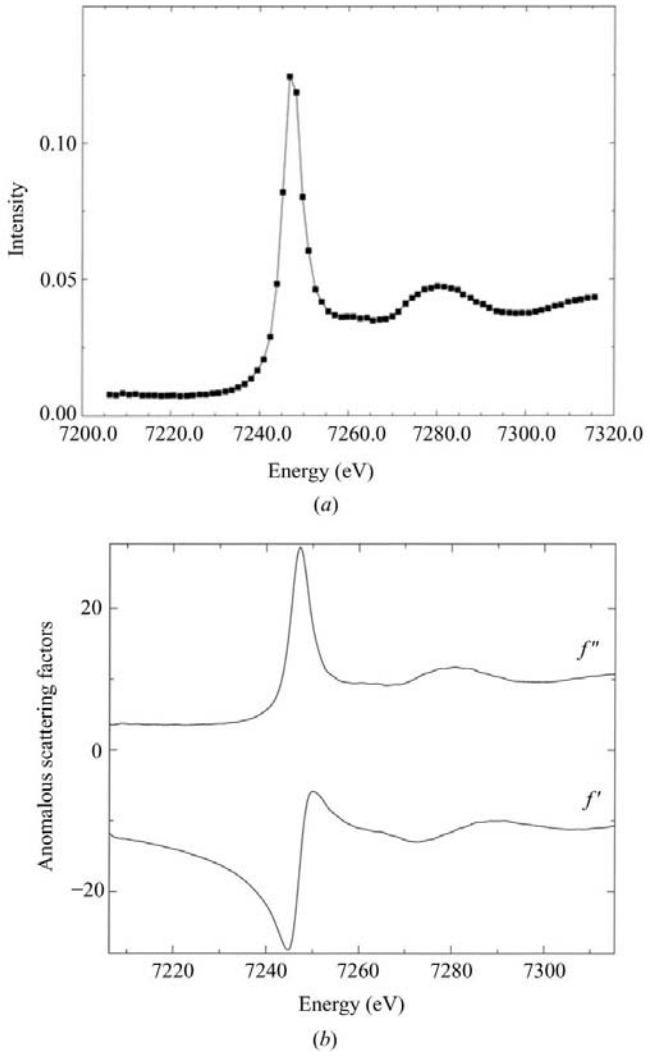

Figure 2
(*a*) Fluorescence spectrum from a 100 m*M* Gd–HPDO3A solution at the Gd $L_{\rm III}$ absorption edge. (*b*) f' and f'' values derived from the fluorescence spectrum using the program *CHOOCH* (Evans & Pettifer, 2001). Fluorescence data were collected using beamline BM30A at the ESRF.

same protein, a common origin should be found for all these derivatives. Therefore, using allowed origin translations for the crystal space group, the atomic positions initially found were translated into the same asymmetric unit as the Gd positions found from anomalous difference Fourier syntheses calculated for each derivative crystal using phases derived from the same atomic model of the protein. This model was used only for fixing a common origin and thus this procedure did not in any way affect the *de novo* nature of the phasing.

2.4.1. Procedure for SAD phasing. The resulting coordinates were directly used for phasing with the program *MLPHARE* (Otwinowski, 1991). Based on this program, we established a procedure that describes the heavy-atom structure more precisely and allows the estimation of phases of centric reflections during the SAD phasing process. As usual in procedures that phase SAD data using *MLPHARE*, the mean structure factor is utilized to describe both the native structure and the derivative structure. The anomalous partial structure is refined from the anomalous differences. A command script has been written that allows the linking together of different successive cycles of *MLPHARE*. As soon as the coordinates of an anomalous scatterer are known, they are introduced into the refinement procedure. In a first cycle, these coordinates, as well as the corresponding anomalous occupancies, are refined, whilst the isotropic B factors are fixed to 25 Å^2. During the next two cycles, for each heavy atom, in addition to the atomic coordinates and to the anomalous occupancy, the thermal B factor is refined, first as an isotropic scalar and then as an anisotropic tensor. The value of the real occupancy, which is never refined by the program, is fixed to the value of the refined anomalous occupancy. This requires that in the atomic structure-factor library, atomsf.lib, f' and f'' be set to their actual values, either by being tabulated or derived from the experiment. The values for both occupancies converge generally after one or two subsequent refinement cycles.

2.4.2. Phasing the HEWL derivative data. Results obtained with the Gd–HPDO3A complex have been described previously (Girard *et al.*, 2002). As in the six other gadolinium

Table 3
Number of Gd sites and phasing statistics for lysozyme derivative crystals obtained with the seven Gd complexes.

Values in parentheses refer to the highest resolution shell.

	Gd–DTPA	Gd–DTPA-BMA	Gd–DO3A	Gd–HPDO3A	Gd–HPSA-DO3A	Gd–DOTA	Gd–DOTMA
No. of sites determined by *SHELX*	4	2	3	2†	1	2	1
FOM after *MLPHARE*	0.244	0.147	0.411	0.592	0.103	0.330	0.068
Anomalous $R_{\rm cull}$	0.83 (0.98)	0.90 (0.99)	0.73 (0.87)	0.57 (0.79)	0.95 (1.00)	0.77 (0.92)	0.96 (1.00)
FOM after *DM*	0.560	0.544	0.671	0.729	0.390	0.659	0.424
Free *R* factor in real space	0.188	0.233	0.202	0.133	0.300	0.199	0.317
No. of cycles	6	8	6	8	7	6	6

† Sites determined using *RSPS* (Knight, 2000). See Girard *et al.* (2002).

Table 4
Number of Gd sites and phasing statistics for urate oxidase derivative crystals obtained with the seven Gd complexes.

Values in parentheses refer to the highest resolution shell.

	Gd–DTPA	Gd–DTPA-BMA	Gd–DO3A	Gd–HPDO3A	Gd–HPSA-DO3A	Gd–DOTA	Gd–DOTMA
No. of sites determined by *SnB*	3	3	2	3	3	3	1
No. of sites finally determined	4	4	11	6	5	6	3
FOM after *MLPHARE*	0.340	0.582	0.402	0.272	0.375	0.299	0.432
Anomalous $R_{\rm cull}$	0.71 (0.79)	0.37 (0.45)	0.62 (0.69)	0.79 (0.95)	0.66 (0.77)	0.77 (0.83)	0.54 (0.63)
No. of cycles	8	8	50	10	12	19	12
FOM after *DM*	0.767	0.809	0.792	0.755	0.773	0.767	0.793
Free *R* factor in real space	0.117	0.124	0.110	0.134	0.129	0.119	0.116

complexes, Gd-atom positions were found using the program *SHELX*. After six cycles of the phasing procedure described previously, phases were improved by solvent flattening and histogram matching using the program *DM* (Cowtan & Main, 1996), assuming 27% solvent content as estimated by the *CCP*4 program *TRUNCATE*. The number of phase-improvement cycles was determined automatically by *DM* on the basis of the real-space free residual, free *R* (Abrahams, 1997). In all cases, no extra sites were detected in anomalous difference Fourier synthesis using the phases improved with *DM*. Only the anomalous difference Fourier synthesis of the Gd–DOTMA derivative of HEWL showed one low-level extra peak in the electron density that could be attributed to a Gd atom with a low occupancy. Interestingly, this potential site is on a special position along a twofold axis, which is consistent with the high symmetry of the DOTMA–gadolinium complex. Nevertheless, this extra site was not taken into account for phasing. To allow a direct comparison with the other derivatives, the same protocol was applied to the Gd–HPDO3A derivative, for which the structure had been previously determined *de novo* using direct methods on SAD data (Girard *et al.*, 2002). Results of the phasing procedure applied to the whole series of gadolinium derivatives are summarized in Table 3. The Gd site with the highest peak in the anomalous difference Fourier synthesis is common to all derivatives. When there are two sites, the gadolinium positions are close to those found in the Gd–HPDO3A derivative.

2.4.3. Phasing the urate oxidase derivative data. For the seven gadolinium derivatives of urate oxidase, the heavy-atom positions were determined with the program *SnB*. The bimodal histograms of the minimum function $R_{\rm min}$ observed at the end of all *SnB* runs gave a clear indication that the Gd-atom partial structures had been determined successfully (Miller *et al.*, 1994; Howell *et al.*, 2000). As evidenced by the anomalous Patterson maps (Fig. 3), the Gd sites are different in all of the derivatives. The numbers of sites found are indicated in Table 4. The coordinates found by *SnB* were directly used for the phasing protocol previously described in which, based on the fluorescence spectrum (Fig. 2), f' and f'' values for Gd were set to -15 and $28\,e^-$, respectively, in the atomsf.lib file.

Phases were subsequently improved by solvent flattening and histogram matching using the program *DM*, assuming 50% solvent content as estimated by the *CCP*4 program *TRUNCATE*. In order to be able to compare results after phase improvement for different derivative crystals, a fixed number of ten cycles of solvent flattening and histogram matching was imposed on each derivative crystal data set. New sites were found from anomalous difference Fourier syntheses using phases improved with *DM*. These new sites were introduced in subsequent cycles of the whole procedure. The final number of sites as well as phasing statistics with the urate oxidase derivatives obtained with the seven different Gd complexes are given in Table 4.

2.5. Correlation between maps

To evaluate the quality of the different gadolinium derivatives for phasing SAD data from the same protein, a correlation coefficient between the electron-density map obtained from the experimental phases and the electron-density map calculated from an atomic model was determined with the *CCP*4 program *OVERLAPMAP*. The atomic model was first determined by molecular replacement using the program *AMoRe* (Navaza, 1994) and was then positioned in the same asymmetric unit as the original model using allowed origin translations for the crystal space group.

Table 5
Correlation between the electron-density maps obtained from experimental phases and the calculated maps from an atomic model determined by molecular replacement for the lysozyme derivatives as well as for the urate oxidase derivatives obtained with the seven Gd complexes.

	Gd–DTPA	Gd–DTPA-BMA	Gd–DO3A	Gd–HPDO3A	Gd–HPSA-DO3A	Gd–DOTA	Gd–DOTMA
Correlation for HEWL derivatives	0.446	0.275	0.563	0.704	0.240	0.534	0.204
Correlation obtained for urate oxidase derivatives	0.548	0.661	0.657	0.498	0.654	0.557	0.668

The HEWL atomic model was the model determined by Dauter *et al.* (1999) using the anomalous signal of S and Cl atoms (PDB code 1lz8). All residues, water molecules and ions were kept for molecular replacement and map correlation. The starting model for urate oxidase was the protein part of the structure (PDB code 1uox) determined by Colloc'h *et al.* (1997).

The correlation coefficients between the different maps are indicated in Table 5. One should note that these correlation coefficients were calculated without taking into account the Gd atoms, which were not included in the atomic model.

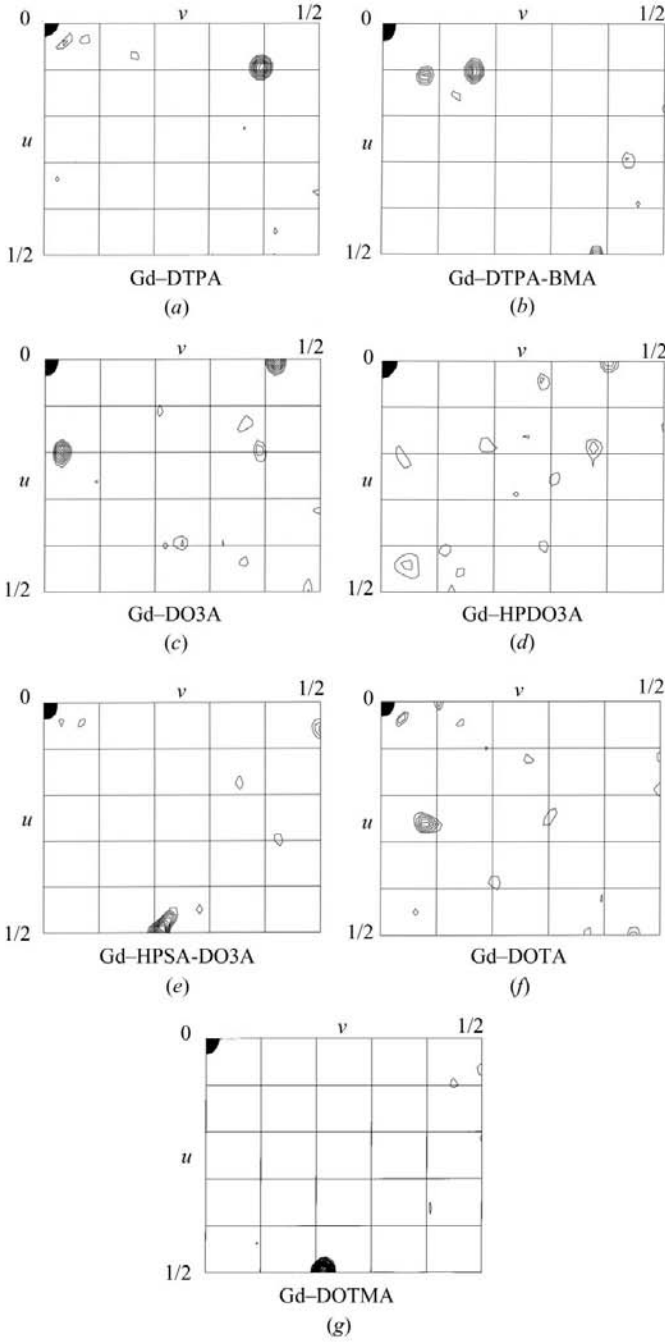

Figure 3
Harker section $w = 0$ of the anomalous Patterson maps for the urate oxidase derivative crystals obtained with the seven Gd complexes. Levels are contoured in steps of 1σ starting from 2σ.

3. Results

Incorporation of the seven complexes in macromolecular crystals was straightforward and did not significantly change the crystallization conditions. Incorporation could be achieved either by co-crystallization, as for the lysozyme derivative crystals, or by soaking, as for the urate oxidase derivative crystals, using heavy-atom concentrations that are considerably higher than those generally used to prepare heavy-atom derivatives.

With respect to the anomalous signal, Gd presents two great advantages. The f'' value for X-rays emitted by a conventional X-ray source is 12 e^-, greater than that for Se with radiation at the Se K absorption edge. For data collected at the Gd L_{III} absorption edge, f'' can reach 30 e^- (Fig. 2). As already observed, combined with high Gd-site occupancies this leads to large anomalous effects.

Thanks to this anomalous signal, effective binding of some complexes can be directly deduced from the high value of the corresponding anomalous factor R_{ano}. For HEWL, the R_{ano} values for Gd–HPDO3A, Gd–DO3A and Gd–DOTA derivatives are 7.1, 5.2 and 4.2%, respectively (Table 2), providing strong evidence that these three gadolinium complexes are bound to lysozyme. In a similar manner, effective binding of Gd–DTPA-BMA, Gd–DO3A and Gd–DOTMA to urate oxidase can directly be deduced from the corresponding R_{ano} values of 9.6, 6.7 and 5.2%, respectively.

In fact, all the gadolinium complexes bind to both proteins. The binding sites were confirmed in anomalous Fourier syntheses calculated using phases from an atomic model. As evidenced by the anomalous Patterson map shown in Fig. 3, all the gadolinium complexes also bound to urate oxidase. Since all of the tested molecules bound to both the test proteins, phasing with the SAD method was attempted on all the derivatives.

For HEWL, *de novo* phasing confirmed that Gd–HPDO3A, Gd–DO3A and Gd–DOTA were bound to the protein sufficiently strongly to produce very good phases, as evidenced by high values of the figures of merit after *MLPHARE* as well as

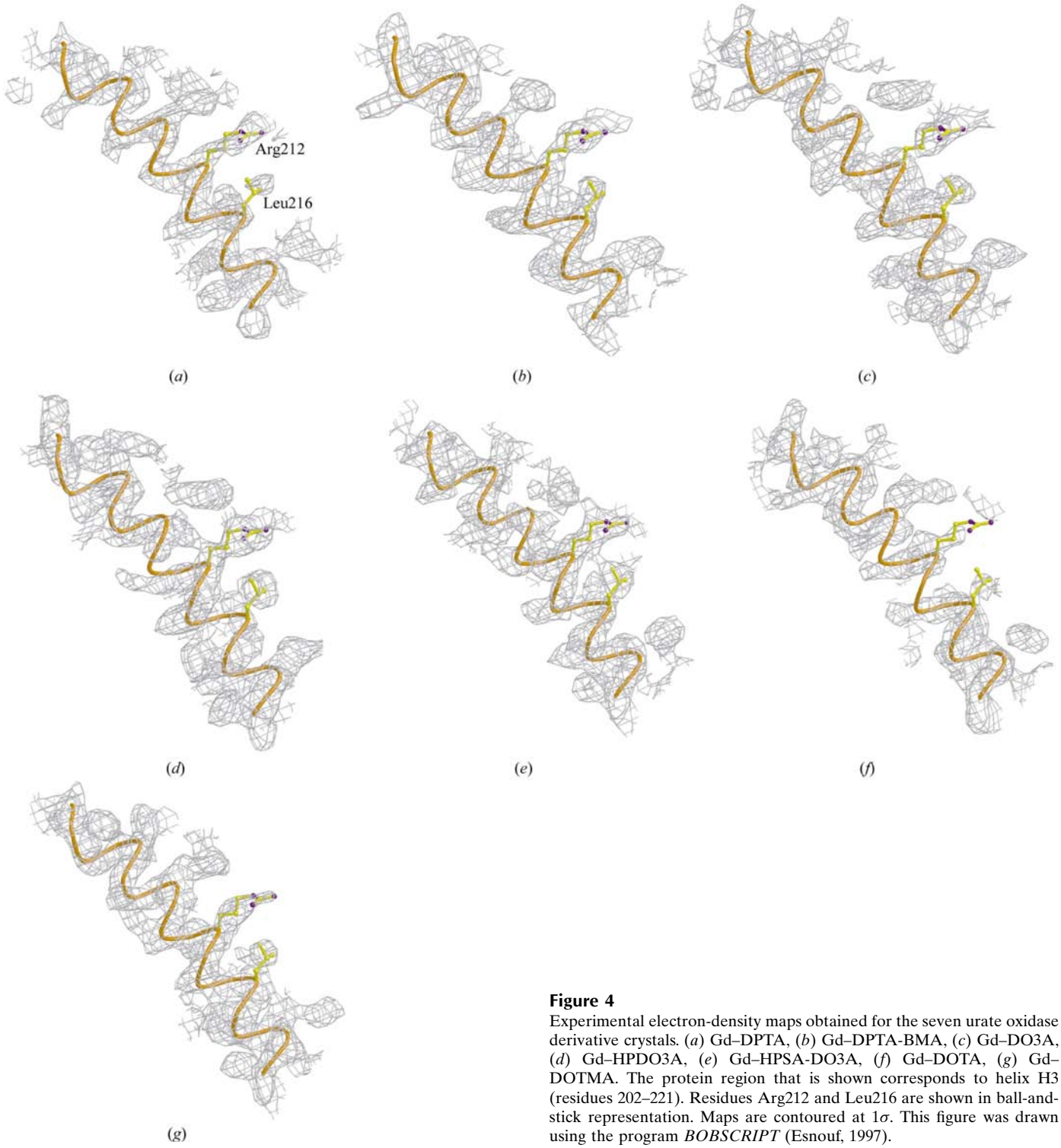

Figure 4
Experimental electron-density maps obtained for the seven urate oxidase derivative crystals. (*a*) Gd–DPTA, (*b*) Gd–DPTA-BMA, (*c*) Gd–DO3A, (*d*) Gd–HPDO3A, (*e*) Gd–HPSA-DO3A, (*f*) Gd–DOTA, (*g*) Gd–DOTMA. The protein region that is shown corresponds to helix H3 (residues 202–221). Residues Arg212 and Leu216 are shown in ball-and-stick representation. Maps are contoured at 1σ. This figure was drawn using the program *BOBSCRIPT* (Esnouf, 1997).

by low values of the $R_{\rm cull}$ factor (Table 3). All of the urate oxidase derivatives led to experimental SAD phases of high quality, allowing unequivocal model building. Gd–DTPA-BMA gives the best derivative, with the highest FOM (0.582) and the lowest $R_{\rm cull}$ factor (0.37) after *MLPHARE* (Table 4). For the Gd–DTPA-BMA, Gd–DO3A and Gd–DOTMA derivatives the phases were excellent, as the FOM values just after *MLPHARE*, without any phase improvement, were greater than 0.40. For all of the derivatives the FOM values after density modification are greater than 0.75 and the free R factors in real space are very low, about 0.12 (Table 4), which reflects the good quality of the final phases.

The previous observation is also supported by the correlation between maps calculated from the experimental phases and maps calculated from an atomic model (Table 5). When a given gadolinium complex was actually bound to the protein,

the experimental electron-density map was good, with correlation factors between experimental and calculated maps of better than 0.50. This was confirmed in the case of urate oxidase by inspecting the different experimental electron-density maps in *O* (Jones *et al.*, 1991) as shown in Fig. 4.

4. Discussion

The strength of the anomalous signal can be estimated by the Bijvoet ratio (Hendrickson & Teeter, 1981),

$$\frac{\langle \Delta F^{\pm} \rangle}{\langle F \rangle} = \left(\frac{2N_A}{N_P} \right)^{1/2} \frac{f''_A}{Z_{\rm eff}},$$

where N_A is the number of anomalous scatterers with an anomalous factor f''_A and N_P is the number of atoms of the protein of mean scattering factor $Z_{\rm eff}$. This relationship assumes that the anomalous scatterers are chemically identical and that their sites are fully occupied. When different anomalous scatterers with different site occupancies are present in the crystal, this relationship becomes

$$\frac{\langle \Delta F^{\pm} \rangle}{\langle F \rangle} = \frac{1}{Z_{\rm eff}} \left(\frac{2 \sum q_i^2 f''^2_i}{N_P} \right)^{1/2},$$

where q_i is the site occupancy and f''_i the imaginary part of the atomic scattering factor of the anomalous scatterer *i*.

This expression shows that, assuming identical site occupancies for the anomalous scatterers, identical Bijvoet ratios are obtained for a protein which is four times larger when f'' is simply doubled. Similarly, assuming given values for f''_i, identical Bijvoet ratios are obtained for a protein four times larger when only the site occupancies are doubled. Notably, if the number of anomalous scatterers is doubled, the protein size can only be doubled in order for the same result to be achieved.

Thus, with their large anomalous scattering effects, lanthanide atoms are of special interest in protein crystallography. However, they have been used mainly in the case of proteins with a specific calcium-binding site, which can be correlated with the special coordination of lanthanides. Indeed, this coordination requires approximately nine neighbouring atoms, which cannot be fully achieved in most proteins.

The complexes presented in this paper overcome this problem since the ligand provides the major part of the coordination needed. Compared with lanthanide salts, this could explain why these Gd complexes are not disruptive to protein crystals and why they can be introduced easily into such crystals at high concentration by either soaking or co-crystallization.

The results from lysozyme and urate oxidase show that derivative crystals can be obtained over a large pH range, from 4.5 to 8.5, with precipitant agents commonly used in macromolecular crystallization such as salts or PEGs. Whilst the Gd–HPDO3A complex provides the best lysozyme derivative, with NaCl as the precipitant agent, the Gd–DTPA-BMA complex leads to the best urate oxidase derivative, for which PEG 8000 was used as the precipitant agent. This observation could indicate a specific binding of the Gd complexes depending on the nature of the precipitant agent and/or pH. The Gd–DO3A complex allowed the preparation of good heavy-atom derivatives in both cases. Evans & Bricogne (2002) have suggested that the use of iodide clusters in combination with PEG may be hindered by the affinity of PEG for iodide. Our results on urate oxidase clearly show that PEG does not prevent the fixation of Gd complexes on the protein. The possible precipitant/pH dependency should be further analyzed by refining each derivative structure and by studying the binding mode of each Gd complex. This work is in progress. Comparison between the fixation of complexes in a soaked crystal and fixation in a co-crystallized crystal of the same protein is foreseen.

As observed for Gd–HPDO3A (Girard *et al.*, 2002), the binding mode of the Gd complexes seems to mainly employ the ligand molecule and not the lanthanide atom itself. This binding mode also explains the easy incorporation of these complexes into crystals and above all the high occupancies of the binding sites.

Such strong binding allowed us to solve the phase problem for three lysozyme derivatives for which data were collected at high resolution with a medium anomalous signal (f'' = 12 e$^-$) using a conventional X-ray source. For urate oxidase, data could only be collected at 3 Å resolution whilst the derivative crystal diffracts to 2 Å . Thanks to the high anomalous signal of the Gd atom at the $L_{\rm III}$ absorption edge (f'' = 28 e$^-$) good phases could be determined even with single-wavelength data collected at a rather low resolution.

5. Conclusions

The seven Gd complexes presented in this study are excellent candidates for derivatizing agents in macromolecular X-ray crystallography. We have shown that the Gd complexes can be introduced at high concentration in both lyzozyme and urate oxidase crystals by co-crystallization and soaking, respectively. For a large number of derivative crystals for which binding is sufficient, the good quality of the phases derived from SAD experiments allows unambiguous model building.

Since the binding sites may be different with different complexes for a given protein, these molecules could possibly be used as mixtures, taking advantage of their different binding modes.

These complexes form a new set of heavy atoms for protein crystallography and, combined with the SAD method, are of particular interest for high-throughput projects.

The authors would like to thank Bracco Imaging spa, Milan, Italy for kindly providing samples of Gd–HPDO3A, Gd–DO3A, Gd–DOTMA and Gd–HPSA-DO3A, Pr J.-F. Le Bas, CHU–Hôpital Nord, Grenoble, France for kindly providing commercial solutions of Gd–DOTA, Gd–DTPA and Gd–DTPA-BMA, Dr M. El Hajji from Sanofi-Synthélabo, Aramon, France for kindly providing urate oxidase, Drs F. Bonneté and D. Vivarès, Université Paris 6–Paris 7, France for

discussions on the crystallization of urate oxidase and the BM30A staff for help on the beamline.

References

Abrahams, J. P. (1997). *Acta Cryst.* D**53**, 371–376.
Bellizzi, J. J., Widom, J., Kemp, C. W. & Clardy, J. (1999). *Structure Fold. Des.* **7**, R263–R267.
Boggon, T. J. & Shapiro, L. (2000). *Structure Fold. Des.* **8**, R143–R149.
Bond, C. S., Shaw, M. P., Alphey, M. S. & Hunter, W. N. (2001). *Acta Cryst.* D**57**, 755–758.
Bonneté, F., Vivarès, D., Robert, Ch. & Colloc'h, N. (2001). *J. Cryst. Growth*, **232**, 330–339.
Chen, W. & Bahl, O. P. (1991). *J. Biol. Chem.* **266**, 8192–8197.
Cohen, A., Ellis, P., Kresge, N. & Soltis, S. M. (2001). *Acta Cryst.* D**57**, 233–238.
Cohen, S. L., Padovan, J. C. & Chait, B. T. (2000). *Anal. Chem.* **72**, 574–579.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Colloc'h, N., el Hajji, M., Bachet, B., L'Hermite, G., Schiltz, M., Prangé, T., Castro, B. & Mornon, J. P. (1997). *Nature Struct. Biol.* **4**, 947–952.
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. & Dodson, E. (2002). *Acta Cryst.* D**58**, 494–506.
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.
Doublié, S. (1997). *Methods Enzymol.* **276**, 523–530.
Esnouf, R. M. (1997). *J. Mol. Graph. Model.* **15**, 132–134.
Evans, G. & Bricogne, G. (2002). *Acta Cryst.* D**58**, 976–991.
Evans, G. & Pettifer, R. F. (2001) *J. Appl. Cryst.* **34**, 82–86.
Girard, E., Chantalat, L., Vicat, J. & Kahn, R. (2002). *Acta Cryst.* D**58**, 1–9.
Gordon, E. J., Leonard, G. A., McSweeney, S. & Zagalsky, P. F. (2001). *Acta Cryst.* D**57**, 1230–1237.
Hendrickson, W. A., Horton, J. R. & LeMaster, D. M. (1990). *EMBO J.* **9**, 1665–1672.
Hendrickson, W. A. & Ogata, C. M. (1997). *Methods Enzymol.* **276**, 494–523.
Hendrickson, W. A. & Teeter, M. M. (1981). *Nature (London)*, **290**, 107–113.
Howell, P. L., Blessing, R. H., Smith, G. D. & Weeks, C. M. (2000). *Acta Cryst.* D**56**, 604–617.
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.
Liu, Z. J., Vysotski, E. S., Chen, C. J., Rose, J. P., Lee, J. & Wang, B. C. (2000). *Protein Sci.* **9**, 2085–2093.
Lustbader, J. W., Wu, H., Birken, S., Pollak, S., Gawinowicz Kolks, M. A., Pound, A. M., Austen, D., Hendrickson, W. A. & Canfield, R. E. (1995). *Endocrinology*, **136**, 640–650.
McWhirter, S. M., Pullen, S. S., Holton, J. M., Crute, J. J., Kehry, M. R. & Alber, T. (1999). *Proc. Natl Acad. Sci. USA*, **96**, 8408–8413.
Micossi, E., Hunter, W. N. & Leonard, G. A. (2002). *Acta Cryst.* D**58**, 21–28.
Miller, R., Gallo, S. M., Khalak, H. G. & Weeks, C. M. (1994). *J. Appl. Cryst.* **27**, 613–621.
Navaza, J. (1994). *Acta Cryst.* A**50**, 157–163.
Otwinowski, Z. (1991). *Proceedings of the CCP4 Study Weekend. Isomorphous Replacement and Anomalous Scattering*, edited by W. Wolf, P. R. Evans & A. G. W. Leslie, pp. 80–86. Warrington: Daresbury Laboratory.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Purdy, M. D., Ge, P., Chen, J., Selvin, P. R. & Wiener, M. C. (2002). *Acta Cryst.* D**58**, 1111–1117.
Quillin, M. L. & Matthews, B. W. (2002). *Acta Cryst.* D**58**, 97–103.
Riboldi-Tunnicliffe, A. & Hilgenfeld, R. (1999). *J. Appl. Cryst.* **32**, 1003–1005.
Schiltz, M., Prangé, T. & Fourme, R. (1994). *J. Appl. Cryst.* **27**, 950–960.
Sheldrick, G. M. (1986). *SHELX*86. *Program for Crystal Structure Solution.* University of Göttingen, Germany.
Sheldrick, G. M. (1998). *Direct Methods for Solving Macromolecular Structures*, edited by S. Fortier, pp. 401–411. Dordrecht: Kluwer Academic Publishers.
Sun, P. D. & Hammer, C. H. (2000). *Acta Cryst.* D**56**, 161–168.
Sun, P. D. & Radaev, S. (2002). *Acta Cryst.* D**58**, 1099–1103.
Sun, P. D., Radaev, S. & Kattah, M. (2002). *Acta Cryst.* D**58**, 1092–1098.
Vitali, J., Robbins, A. H., Almo, S. C. & Tilton, R. F. (1991). *J. Appl. Cryst.* **24**, 931–935.
Wang, B.-C. (1985). *Methods Enzymol.* **115**, 90–111.
Weeks, C. M. & Miller, R. (1999). *J. Appl. Cryst.* **32**, 120–124.