

A new class of lanthanide complexes to obtain high-phasing-power heavy-atom derivatives for macromolecular crystallography

Éric Girard,† Meike Stelter, Jean Vicat and Richard Kahn*

Institut de Biologie Structurale J.-P. Ébel
CEA-CNRS-UJF, 41 Rue Jules Horowitz,
38027 Grenoble CEDEX 1, France

† Present address: Synchrotron Soleil, l'Orme
des Merisiers, Saint-Aubin BP 48,
91192 Gif-sur-Yvette CEDEX, France.

Correspondence e-mail: richard.kahn@ibs.fr

Received 27 May 2003

Accepted 16 September 2003

Because of their intense white lines and large f'' values, lanthanide atoms are of great interest for solving structures of biological macromolecules using single-wavelength anomalous diffraction (SAD) or multiple-wavelength anomalous diffraction (MAD) methods. In this work, a series of seven gadolinium complexes are described which provide excellent derivatives for anomalous diffraction experiments in biological systems. These highly soluble lanthanide complexes can easily be introduced into protein crystals either by soaking or by co-crystallization, without significantly affecting the crystallization conditions, by employing highly concentrated complex solutions (≈ 100 mM). *De novo* phasing by the SAD method was carried out with several proteins of known as well as previously unknown structures by employing this new class of heavy-atom compounds. Diffraction data were collected either with a laboratory source, making use of the high anomalous signal ($f'' = 12 e^-$) of gadolinium with Cu $K\alpha$ radiation, or with synchrotron radiation at the peak of the gadolinium L_{III} absorption edge, which exhibits a strong white line ($\lambda = 1.711 \text{ \AA}$, $f'' = 28 e^-$). Using one of these gadolinium complexes, Gd-HPDO3A, the structure of a bacterial chimeric ornithine carbamoyl transferase, OTC_{Case3630}, a dodecameric protein of 450 kDa, was determined. Employed with the SAD method, these seven complexes could be of particular interest for high-throughput macromolecular crystallography.

1. Introduction

The multiple-wavelength anomalous solvent-contrast (MASC) method (Fourme *et al.*, 1995; Ramin *et al.*, 1999; Shepard *et al.*, 2000) uses the anomalous scattering of a chemical species dispersed at high concentration in the solvent of a crystal of a biological macromolecule in order to determine the molecular envelope of the protein as well as the fluctuations of the electron density at low resolution. The contrast variation is obtained by changing the wavelength of the incident X-rays around one of the absorption edges of the anomalous scatterer.

In contrast to the multiple-wavelength anomalous diffraction (MAD) method, for which the anomalous scatterers need to be well localized, the anomalous partial structure in the MASC method is supposed to correspond to a uniform electron density extending over the whole solvent domain, where the solvent molecules are virtually free. As a consequence, in an ideal MASC experiment the contrast agent should not bind significantly to the protein even at high concentration ($\approx 1 M$). From this point of view, ionic compounds are not well suited as MASC agent candidates because they usually bind quite strongly to the protein (Ramin *et al.*, 1999). Indeed,

the binding of ionic compounds has been used in cryosoaking methods of preparing heavy-atom derivatives (Dauter *et al.*, 2000). Sound specifications for an ideal MASC contrast agent are therefore (i) it should contain a strong anomalous scatterer, preferably with a white line in the selected absorption edge, (ii) it should be electrically neutral, (iii) it should be water-soluble at high concentration ($\approx 1\text{ M}$) and (iv) it may be introduced at high concentration into the crystal either by soaking or by co-crystallization without damaging the crystal or disturbing the crystallization conditions.

The electrically neutral gadolinium complex Gd-HPDO3A, which is employed as a contrast agent for magnetic resonance imaging, fulfilled the first three characteristics. It was necessary to verify whether it was possible to introduce it at high concentration into biological macromolecular crystals without significant binding of the complex to the macromolecule. In fact, Gd-HPDO3A turned out to bind strongly to the first tested protein, which was hen egg-white lysozyme (HEWL). The structure of the Gd-HPDO3A derivative of HEWL was easily solved applying the single-wavelength anomalous diffraction (SAD) method to diffraction data collected with a conventional X-ray source (Girard *et al.*, 2002). Consequently, Gd-HPDO3A looked to be of particular interest as a tool for solving the phase problem in biological macromolecular crystallography. Subsequently, seven gadolinium complexes, including Gd-HPDO3A, have been tested on several proteins and have been found to provide excellent heavy-atom derivatives for anomalous diffraction experiments in macromolecular crystallography (Girard *et al.*, 2003).

2. A series of gadolinium complexes

The seven gadolinium complexes are composed of ligand molecules (Fig. 1) that chelate a single Gd^{3+} ion. The core of HPDO3A, DO3A, DOTMA, HPSA-DO3A and Gd-DOTA is a tetraazacyclododecane macrocycle. DTPA and DTPA-BMA are acyclic ligands. When the ligand molecules chelate the Gd^{3+} ion, their carboxylic groups are deprotonated. Therefore, the resulting 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. They are all highly soluble in water.

The sources of the seven gadolinium complexes are given by Girard *et al.* (2003).

3. Tests on proteins of known structure

As already mentioned, Gd-HPDO3A, the first tested compound, bound strongly to HEWL (Girard *et al.*, 2002). Tetragonal gadolinium-derivative crystals of HEWL were obtained by co-crystallization using NaCl as precipitant and different concentrations of the complex. Diffraction data from three derivative crystals (100, 50 and 10 mM Gd-HPDO3A concentrations) were collected to a resolution of 1.7 Å using Cu $K\alpha$ radiation from a laboratory X-ray generator. In both the 100 and the 50 mM derivatives, two highly occupied

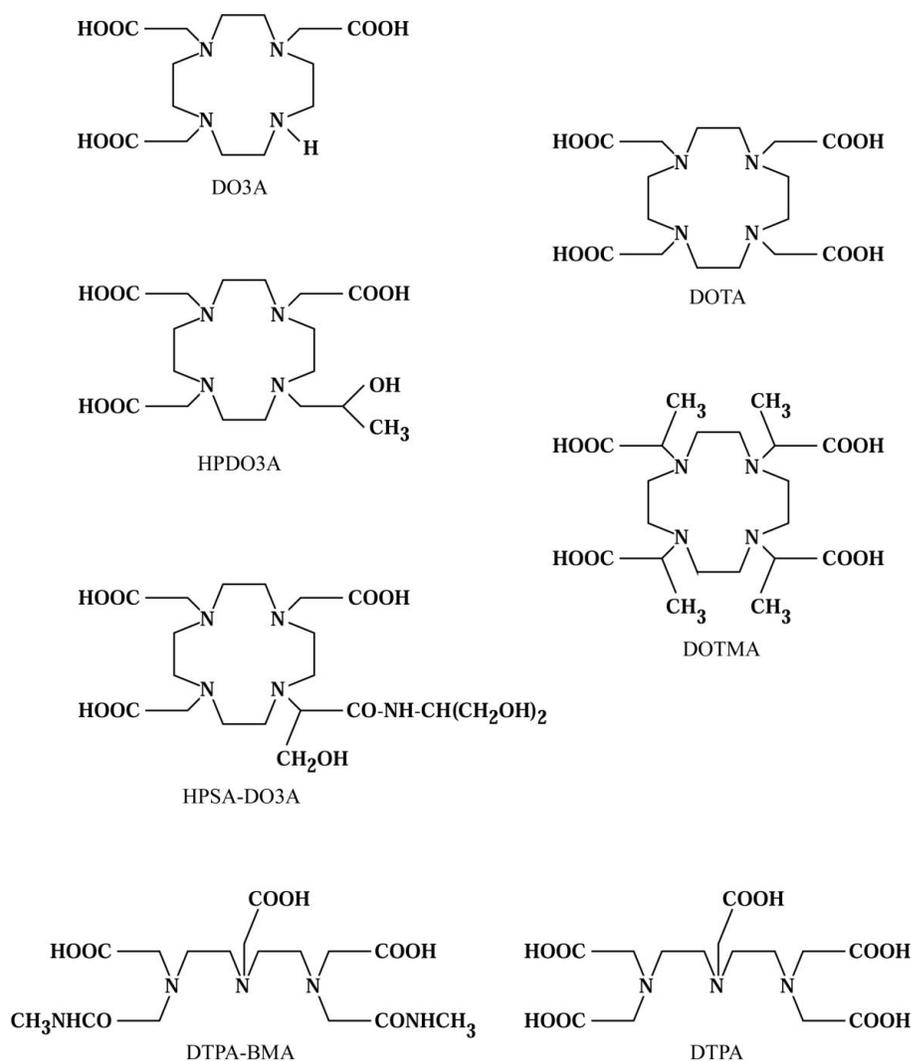


Figure 1
The different ligands in the seven Gd complexes presented in this study. DO3A, 1,4,7,10-tetraazacyclododecane-1,4,7-triacetic acid; DOTA, 1,4,7,10-tetraazacyclododecane-1,4,7,10-tetraacetic acid; HPDO3A, 10-(2-hydroxypropyl)-1,4,7,10-tetraazacyclododecane-1,4,7-triacetic acid; DOTMA, *a,a',a'',a'''*-tetramethyl-1,4,7,10-tetraazacyclododecane-1,4,7,10-tetraacetic 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. Ligands leading to non-charged Gd complexes are on the left-hand side of the figure.

Table 1

Occupancies and *B* factors for the two gadolinium sites of lysozyme-derivative crystals obtained by co-crystallization with 100 mM Gd-HPDO3A.

The first crystal was back-soaked for 8 h in a 10 mM solution of Gd-HPDO3A. The second crystal was used without back-soaking.

	Back-soaked crystal	Crystal without back-soaking
Gd 1		
Occupancy	0.77	0.81
<i>B</i> factor (Å ²)	16.4	15.3
Gd 2		
Occupancy	0.68	0.73
<i>B</i> factor (Å ²)	18.6	15.9

binding sites of the gadolinium complex were localized. In the 10 mM derivative, only one site was occupied. With both the 100 and the 50 mM derivatives, SAD phasing using the program *OASIS* (Hao *et al.*, 2000) followed by density modification with the program *DM* (Cowtan & Main, 1996) from the CCP4 suite of programs (Collaborative Computational Project, Number 4, 1994) led to electron-density maps of high quality. The structure of the 100 mM Gd-HPDO3A derivative (PDB code 1h87) was refined using the program *CNS* (Brünger *et al.*, 1998). The final refinement residuals *R* and *R*_{free} (Brünger, 1992) were 18.0 and 20.0%, respectively. The Gd-HPDO3A binding mode was analyzed on the basis of this structure. It combines hydrophobic interactions and hydrogen bonds. The hydrophobic interactions involve tryptophan residues and the tetraazacyclododecane macrocycle of the ligand molecule. On the opposite side of the Gd-HPDO3A complex, pointing toward the solvent, a well structured hydrogen-bonding network involves water molecules (including a water molecule directly bound to the gadolinium), the carboxylic functions of the HPDO3A ligand and protein residues. These interactions, which are rather weak, may explain why high concentrations of the gadolinium complex can be used without significantly modifying the crystallization conditions. It may also explain why concentrated solutions of the gadolinium complex must be used to obtain sufficiently occupied gadolinium sites. Indeed, the site occupancies increase with increasing concentrations of the gadolinium complex (Girard *et al.*, 2002). Nevertheless, a lysozyme derivative co-crystallized with 100 mM Gd-HPDO3A that was back-soaked for 8 h in a 10 mM solution of Gd-HPDO3A gave Gd-site occupancies that were very similar to those of a crystal also prepared with 100 mM Gd-HPDO3A that was not back-soaked (Table 1). This suggests that once the lysozyme-derivative crystal has grown, the Gd-HPDO3A complex cannot leave the binding site and diffuse outside the crystal. Conversely, we never succeeded in preparing lysozyme-derivative crystals by soaking native crystals in Gd-HPDO3A solutions. This behaviour may be quite specific to lysozyme, which is characterized by an unusually low solvent content of the native crystals (27%).

Catalase from *Proteus mirabilis* (PMC) was used as a second test protein with the Gd-HPDO3A complex. PMC is a 55 kDa enzyme of known structure (Gouet *et al.*, 1995;

Andrèoletti *et al.*, 2003) that catalyzes the decomposition of hydrogen peroxide into oxygen. Native crystals were obtained using ammonium sulfate as precipitant at a concentration in the range 2.3–2.7 *M*. The crystals were then conserved in a solution containing 3.7 *M* ammonium sulfate. Soaking conditions had to take into account this particularly high salt content. Thanks to the high solubility of the Gd-HPDO3A complex and its availability as a powder, native crystals could be soaked in a solution containing 3.7 *M* ammonium sulfate and 50 or 100 mM Gd-HPDO3A. A SAD data set was collected at the maximum of *f*^{''} in the white line of the Gd *L*_{III} absorption edge ($\lambda = 1.711 \text{ \AA}$, $f'' = 28 e^-$) on a 100 mM Gd-HPDO3A derivative of PMC. Initial site positions were found using the program *SHELXS97* (Sheldrick, 1990). Subsequent cycles of heavy-atom site refinement using the program *MLPHARE* (Otwinowski, 1991) from the CCP4 suite of programs led to a first set of phases. The initial phases were improved through density modification using the program *DM*. New sites were detected in the anomalous difference Fourier maps and were introduced into the phasing process. At the end of this procedure, 13 Gd-HPDO3A binding sites were located. Very good phases were obtained with a final figure of merit of 0.8. The structure of the Gd-HPDO3A derivative of PMC was refined using the program *CNS*. Although the crystals were soaked for no longer than 10 min, the occupancies of the two main sites were found to be higher than 0.7. A full report will be published elsewhere (É. Girard, V. Stojanoff, H. Jouve, J. Vicat & R. Kahn, in preparation), including a comparison between the SAD phasing of data collected at the Gd *L*_{III} absorption edge and the SAD phasing of data collected with a conventional X-ray source on a Gd-HPDO3A derivative of PMC. To our knowledge, these experiments made PMC the largest protein structure solved directly by the SAD method with Cu *K* α radiation on a derivative crystal.

In order to compare the seven gadolinium complexes for SAD structure determination, studies were carried out on two proteins of known structure. Results on the corresponding derivatives of HEWL and of urate oxidase from *Aspergillus flavus*, a medium-size protein of 295 residues, have been described in a previous paper (Girard *et al.*, 2003).

Lysozyme-derivative crystals were obtained by co-crystallizing the protein with the seven gadolinium complexes at a concentration of 100 mM using the usual crystallization conditions with NaCl as precipitant. Diffraction data were collected to a resolution of 1.7 Å using Cu *K* α radiation from a laboratory X-ray generator. For three lysozyme-derivative crystals containing the complexes Gd-HPDO3A, Gd-DO3A and Gd-DOTA, respectively, SAD phasing led to electron-density maps of good quality, allowing chain tracing.

Urate oxidase native crystals grew with PEG 8000 as precipitating agent. Derivative crystals were obtained by soaking native crystals in 100 mM solutions of the seven gadolinium complexes. Diffraction data were collected at the maximum of *f*^{''} in the white line of the Gd *L*_{III} absorption edge. Thanks to the large anomalous signal of gadolinium at this wavelength and despite the resolution of the data being

limited to 3 Å by the finite size of the detector coupled with the smallest crystal-to-detector distance available on the beamline, SAD phasing led to electron-density maps of very high quality for all urate oxidase derivative crystals. Three of the tested derivatives (the Gd-DTPA-BMA, the Gd-DO3A and the Gd-DOTMA derivatives) led to particularly good phases as shown by the experimental electron-density maps as well as the high figures of merit ($FOM > 0.4$) after phasing using the program *MLPHARE* and before any density modification. Comparison of the results obtained with soaked and co-crystallized derivative crystals of urate oxidase is in progress. A preliminary analysis of the results indicates no significant difference in complex binding for the derivative crystals obtained with the two methods.

Table 2

Data-collection parameters and data processing and phasing statistics for three gadolinium-derivative crystals of zinc β -lactamase VIM-2.

Values in parentheses correspond to the highest resolution bin.

Gadolinium complex	Gd-DTPA-BMA	Gd-DO3A	Gd-DOTMA
Crystal dimensions (mm)	0.20 × 0.10 × 0.04	0.16 × 0.07 × 0.03	0.15 × 0.06 × 0.03
Soaking time (min)	420	1020	300
Unit-cell parameters (Å)	$a = 67.80, b = 78.17,$ $c = 80.20$	$a = 67.65, b = 78.18,$ $c = 80.28$	$a = 67.87, b = 77.82,$ $c = 80.09$
Resolution range (Å)	25.0–3.14 (3.31–3.14)	25.0–3.14 (3.31–3.14)	25.0–3.14 (3.31–3.14)
Observed reflections	23940	23580	22956
Unique reflections	3888	3886	3880
Acentric	3245	3236	3229
Centric	643	650	651
$I/\sigma(I)$	14.8 (11.4)	16.2 (10.0)	14.9 (10.4)
Completeness (%)	99.6 (99.9)	99.6 (99.9)	99.6 (99.9)
Multiplicity	6.2 (5.3)	6.1 (5.3)	5.9 (5.1)
R_{sym} (%)	4.0 (5.6)	3.9 (6.6)	3.9 (6.3)
R_{ano}^{\dagger} (%)	8.6 (9.8)	7.8 (8.4)	4.3 (5.2)
No. of sites determined by SnB	1	2	3
Final No. of sites	4	5	3
FOM after <i>MLPHARE</i>	0.538	0.476	0.329
Anomalous R_{cull}	0.46 (0.56)	0.58 (0.68)	0.71 (0.84)
No. of cycles of <i>MLPHARE</i>	24	18	6
FOM after <i>DM</i>	0.738	0.698	0.627
Direct-space R_{free}	0.137	0.146	0.205
Map correlation	0.590	0.499	0.468

$\dagger R_{\text{ano}}$, computed by the program *SCALA* (Collaborative Computational Project, Number 4, 1994), is defined by $R_{\text{ano}} = \sum_{\mathbf{h}} |I(\mathbf{h}) - I(-\mathbf{h})| / \sum_{\mathbf{h}} [I(\mathbf{h}) + I(-\mathbf{h})]$, where $\langle I(\mathbf{h}) \rangle$ indicates the mean intensity of symmetry-related reflections of the same Bijvoet mate \mathbf{h} .

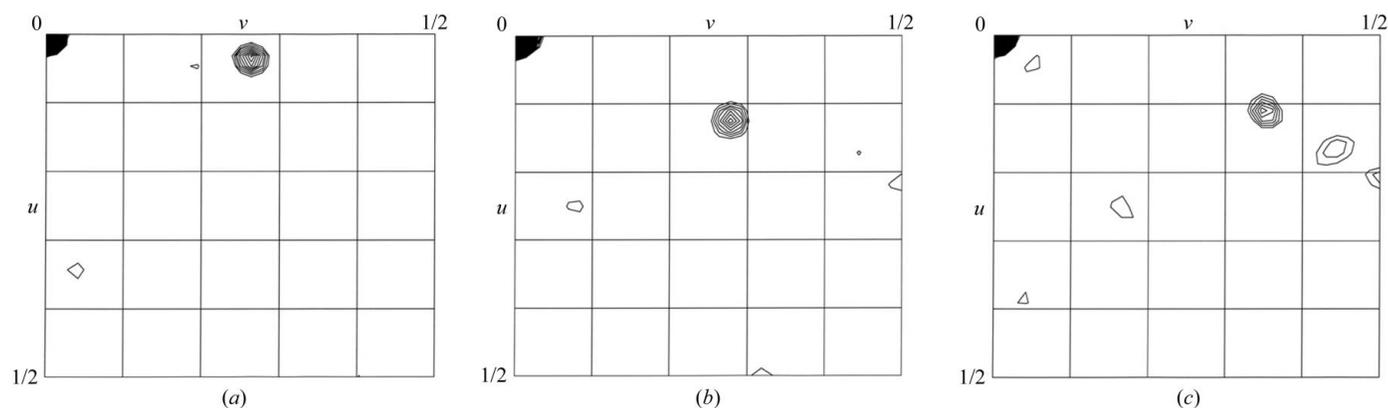


Figure 2

Harker section $w = 0$ of the anomalous Patterson maps for the three derivative crystals of zinc β -lactamase VIM-2. (a) Gd-DTPA-BMA, (b) Gd-DO3A, (c) Gd-DOTMA. Levels are contoured in steps of 1σ starting from 2σ .

The binding modes of the complexes to the proteins in the best derivative crystals of HEWL and of urate oxidase are very different. Whereas Gd-HPDO3A binds to tryptophan residues of HEWL, the binding-site environment of Gd-DTPA-BMA involves charged residues of urate oxidase (Asp141 and Asp133 in the most occupied site, Glu131 and Lys143 in the second site).

The three gadolinium complexes giving the best results with urate oxidase, Gd-DTPA-BMA, Gd-DO3A and Gd-DOTMA, were also tested on VIM-2, a zinc β -lactamase of class B from a *Pseudomonas aeruginosa* clinical isolate (Lauret *et al.*, 1999). The structure of this 240-residue protein had already been solved in our laboratory using molecular replacement (I. Garcia-Saez, J.-D. Docquier, G. M. Rossolini, M. Galleni, J.-M.

Frère & O. Dideberg, in preparation). Crystals of the native protein were obtained with 30% PEG 8000 and 200 mM sodium acetate as precipitating agent. Derivative crystals were obtained by soaking the native crystals in 100 mM solutions of the three complexes.

Diffraction data were collected at 3.14 Å resolution at the Gd L_{III} absorption edge using beamline BM30A at the ESRF (Grenoble, France). Likewise for urate oxidase, the resolution of the diffraction data was limited by the 165 mm diameter of the MARCCD detector located at 132 mm from the sample. Data were processed using the program *DENZO* (Otwinowski & Minor, 1997) and scaled together using the program *SCALA* (Collaborative Computational Project, Number 4, 1994). A summary of data-collection parameters and processing statistics is given in Table 2. The binding of all the complexes to the protein was clearly evidenced by the strong anom-

alous signal leading to large R_{ano} values, as well as by high peaks in the anomalous Patterson maps (Fig. 2).

The procedure used to phase the SAD diffraction data from VIM-2 was very similar to the procedure used to phase the urate oxidase data (Girard *et al.*, 2003). Initial site positions were found using the program *SnB* (Miller *et al.*, 1994; Weeks & Miller, 1999). Phasing was carried out by performing subsequent cycles of heavy-atom site refinement using the program *MLPHARE*, density modification using the program *DM* and detection of new sites in the anomalous difference Fourier maps. The resulting experimental electron-density maps were easily interpretable (Fig. 3). The best results were obtained with the Gd-DTPA-BMA derivative (Table 2) as shown by the large FOM (0.538) obtained before any density modification, the high correlation coefficient between the experimental electron-density map and the map calculated from the known structure (0.590) and by the overall quality of the experimental electron-density map. The comparison of the results obtained with derivative crystals prepared with the seven gadolinium complexes is in progress.

4. Solving an unknown structure

We present here a short description of the structure determination of a chimeric ornithine carbamoyl transferase, OTCase 3630, using Gd-HPDO3A for SAD phasing. A full report of this structure determination will be published elsewhere (É. Girard, G. Sainz, P. L. Anelli, J. Vicat & R. Kahn, in preparation).

Chimeric OTCase3630 was obtained by intragenic fusion of the gene coding for the catabolic OTCase from *P. aeruginosa* and of the gene coding for the anabolic OTCase from *Escherichia coli* (Nguyen *et al.*, 1996). It is formed by the first 100 amino acids of *E. coli* OTCase, which is a non-allosteric enzyme, followed by the last 234 residues of the *P. aeruginosa* OTCase, which presents allosteric properties. OTCase3630 is

still allosteric. Previous trials to solve the structure of OTCase3630 by molecular replacement did not give satisfactory results.

Crystals were obtained using ammonium sulfate as precipitating agent. They belong to space group $P2_1$. The asymmetric unit contains one dodecamer of 450 kDa.

Preliminary trials using native crystals soaked in 100 mM Gd-HPDO3A solutions allowed the detection of clear binding of the gadolinium complex. Derivative crystals for the final data collection were obtained by co-crystallization with 100 mM Gd-HPDO3A complex added to the native crystallization conditions. The structure was solved using 3 Å resolution SAD data collected on co-crystallized derivative crystals at the Gd L_{III} absorption edge using beamline BM14 at the ESRF (Grenoble, France).

The coordinates of 20 sites were determined using the program *SnB*. Refinement of coordinates and occupancies of these sites and phasing of the SAD data were performed with the program *SHARP* (de La Fortelle & Bricogne, 1997). New gadolinium sites were identified from the residual maps, leading to a final total of 67 detected binding sites and a figure of merit of 0.43 before density modification. Solvent flattening and histogram matching using the program *DM* led to a figure of merit of 0.76 and to a readily interpretable electron-density map.

Phases were extended to higher resolution with diffraction data collected on native crystals to 2.5 Å resolution at $\lambda = 0.9797$ Å using beamline ID14-EH4 at the ESRF (Grenoble, France). An excellent electron-density map with a figure of merit of 0.81 was obtained after phase extension using density modification and application of the non-crystallographic symmetry of the dodecamer with the program *DM*. An atomic model has been built from this electron-density map using the programs *QUANTA* (Accelrys Inc.) and *O* (Jones *et al.*, 1991). Refinement of the structure is in progress.

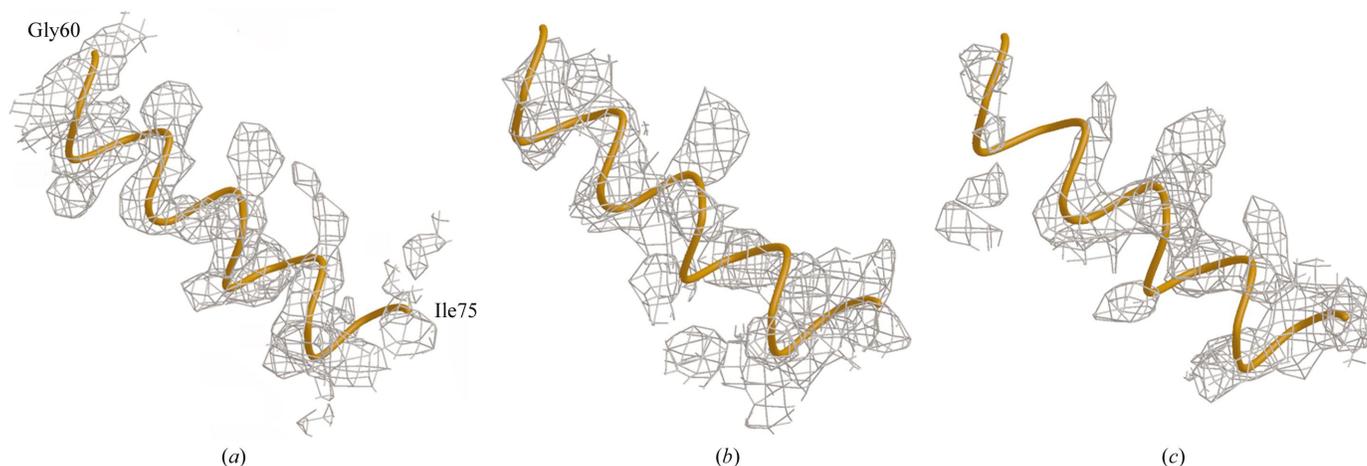


Figure 3

Experimental electron-density maps obtained after *DM* for the zinc β -lactamase derivative crystals of the three studied Gd complexes. (a) Gd-DTPA-BMA, (b) Gd-DO3A, (c) Gd-DOTMA. The represented protein region corresponds to helix H1 (residues 60–75). Maps are contoured at 1σ . This figure was drawn with the program *BOBSCRIPT* (Esnouf, 1997).

As can be seen in Fig. 4, the good quality of the experimental electron-density map obtained directly after density modification of the SAD data using neither the non-crystallographic symmetry information nor data at higher resolution confirms the high phasing power of the Gd-complex derivative combined with the SAD method. Therefore, such a combination allows the phasing of diffraction data for a protein assembly of 450 kDa.

5. Discussion

When the anomalous scatterers are chemically identical and their sites are fully occupied, the strength of the anomalous signal can be estimated by the Bijvoet ratio (Hendrickson & Teeter, 1981),

$$\frac{\Delta F^{\pm}}{F} = \left(\frac{2N_A}{N_P} \right)^{1/2} \frac{f_A''}{Z_{\text{eff}}},$$

where N_A is the number of anomalous scatterers with anomalous scattering factor f_A'' and where N_P is the number of atoms of the protein of mean scattering factor Z_{eff} . More generally, when the crystal contains different anomalous scatterers with different site occupancies this relationship becomes (see Appendix A)

$$\frac{\Delta F^{\pm}}{F} = \frac{1}{Z_{\text{eff}}} \left(\frac{2 \sum q_j^2 f_j''^2}{N_P} \right)^{1/2},$$

where q_j and f_j'' are the site occupancy and the imaginary part of the atomic scattering factor of the anomalous scatterer j , respectively.

Identical Bijvoet ratios are therefore obtained for a protein that is four times larger either when f_j'' is doubled for each anomalous scatterer and the site occupancies are fixed or when the site occupancies are doubled and the value of each f_j'' is fixed. In contrast, to obtain identical Bijvoet ratios for a protein that is four times larger with q_j and f_j'' fixed, the number of anomalous scatterers must be multiplied by 4.

Therefore, with their large f'' values, lanthanide atoms are of special interest for solving macromolecular structures using anomalous scattering. The gadolinium complexes presented in this paper can be introduced into crystals at high concentration without affecting the crystal quality and may bind to the macromolecule giving sites with high occupancies thus providing excellent derivatives for SAD or MAD experiments.

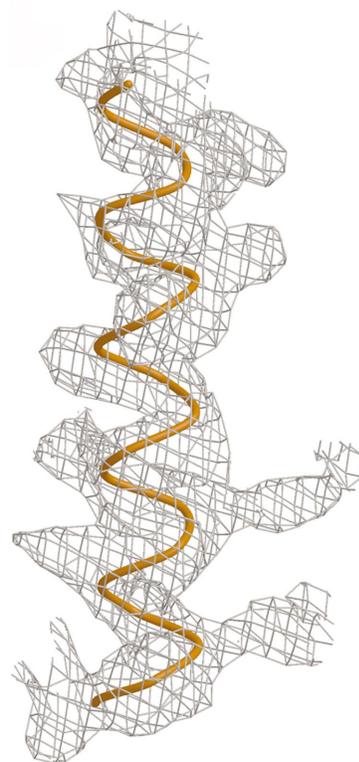


Figure 4 Experimental electron-density map of OTCase3630 obtained with phases determined after *SHARP* and *DM*, without application of NCS for the 3 Å resolution SAD data set. The depicted protein region corresponds to helix H1 (residues 18–37). The map is contoured at 1σ .

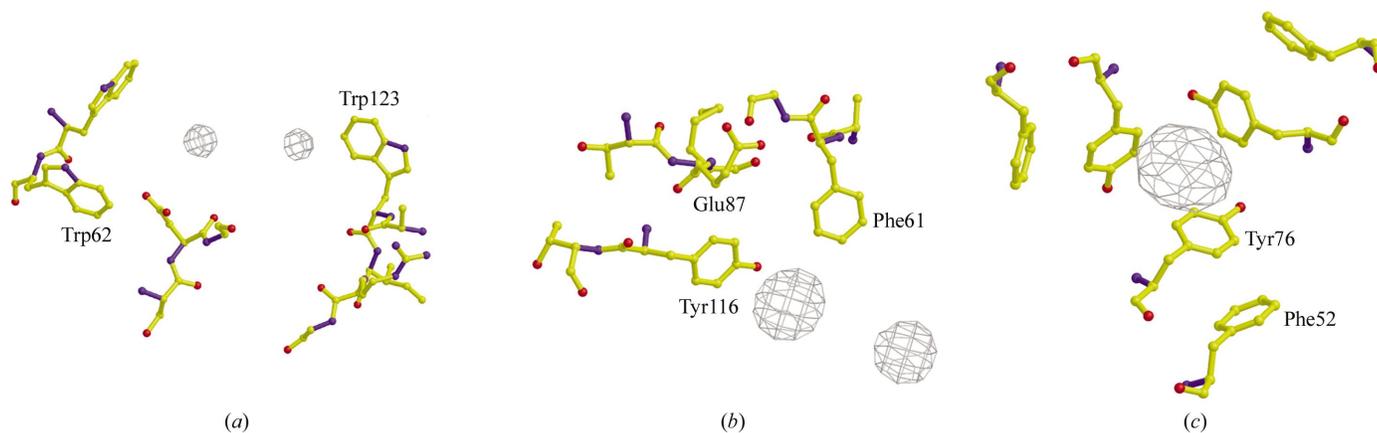


Figure 5

Environment of the major binding sites of the complex Gd-HPDO3A for (a) lysozyme, (b) catalase and (c) OTCase3630 (view along a non-crystallographic threefold-symmetry axis of the dodecamer). The electron density corresponds to the anomalous difference Fourier maps (level 15σ).

Table 3

Solvent content of the native crystals and number of Gd sites in the best derivative for each studied protein.

Protein	HEWL	Catalase	Urate oxidase	VIM-2	OTCase3630
Best derivative	Gd-HPDO3A	Gd-HPDO3A	Gd-DTPA-BMA	Gd-DTPA-BMA	Gd-HPDO3A
Solvent content† (%)	27	65	50	32	55
No. of Gd sites	2	13	4	4	67

† Solvent content was calculated using the program *TRUNCATE* (Collaborative Computational Project, Number 4, 1994).

Since the anomalous scattering of gadolinium is still high with Cu $K\alpha$ radiation ($f'' = 12 e^-$), actual binding of different complexes to the macromolecule in the crystal can be tested using a conventional X-ray source. In case of sufficiently strong binding, the structure can be solved from a SAD data set collected at this wavelength. High data redundancy may then be necessary to compensate for the weakness of the anomalous signal, which is weaker than would be obtained at the peak in the L_{III} absorption edge ($f'' = 28 e^-$). From a purely statistical point of view, neglecting other sources of benefit mainly arising from the geometrical quality of the source when using synchrotron radiation and taking only into account that the precision of a measurement varies with the square root of the number of measurements N , a redundancy multiplied by

$$\frac{N_{Cu\ K\alpha}}{N_{Gd\ L_{III}}} = \left(\frac{\Delta F_{Gd\ L_{III}}^{\pm}}{\Delta F_{Cu\ K\alpha}^{\pm}} \right)^2 = \left(\frac{f''_{Gd\ L_{III}}}{f''_{Cu\ K\alpha}} \right)^2 = \left(\frac{28}{12} \right)^2 = 5.4$$

should be necessary to compensate for the weaker anomalous signal.

Exploiting this series of seven gadolinium complexes, it is possible to obtain a variety of different derivatives. As previously mentioned for the Gd-HPDO3A derivative of HEWL, this gadolinium complex binds to the protein through hydrophobic contacts. It can be seen in Fig. 5 that a similar situation is observed for all tested proteins that gave high Gd-HPDO3A site occupancies, namely HEWL, catalase and OTCase3630. For these three derivatives the most highly occupied binding sites have aromatic residues (tryptophan, tyrosine, phenylalanine) in their direct environment.

Except for HEWL, for which the close crystal packing seems to restrict the location of the possible binding sites, different complexes tend to bind in different locations and consequently in different ways to the protein.

Even if all the complexes are rather large, there is enough room in the solvent channels to allow diffusion of the complex in all the investigated crystals, leading to a number of Gd-binding sites (Table 3).

This study shows that excellent derivatives can be obtained under various crystallization conditions and that effective binding of the complexes can be achieved even with high concentrations of precipitating agents (PEGs or salts). For a few tested proteins we could not obtain any useful derivative with the presented compounds. Nevertheless, in most cases excellent derivatives were obtained using this series of gado-

linium complexes, which makes it a valuable tool for macromolecular crystallography.

APPENDIX A Derivation of Bijvoet ratio for a crystal containing anomalous scatterers of different types with different site occupancies

The structure factor of reflection \mathbf{h} for a crystal containing atoms at rest with fully occupied sites is

$$F(\mathbf{h}) = \sum_j f_j \exp(i2\pi\mathbf{h} \cdot \mathbf{x}_j),$$

where f_j and \mathbf{x}_j are the atomic scattering factor and the vector of the fractional coordinates of atom j , respectively. Starting from this expression, it is easy to derive the structure factor for a crystal containing atoms with different occupancies by just exchanging $q_j f_j$ for f_j , where q_j is the site occupancy of atom j . Thus, we will derive the expression of the Bijvoet ratio for a crystal containing atoms of different species with fully occupied sites and derive the expression of the Bijvoet ratio for a crystal containing atoms with different site occupancies just by a final substitution of f_j by $q_j f_j$.

In the case of significant anomalous scattering from atom j , f_j should be considered as a complex number that may be written

$$f_j = {}^0f_j + f'_j + if''_j = |f_j| \exp(i\delta_j).$$

The intensity of reflection \mathbf{h} is given by

$$\begin{aligned} I(\mathbf{h}) &= |F(\mathbf{h})|^2 = \sum_j \sum_{j'} f_j \bar{f}_{j'} \exp[i2\pi\mathbf{h} \cdot (\mathbf{x}_j - \mathbf{x}_{j'})] \\ &= \sum_j |f_j|^2 + 2 \sum_j \sum_{j' > j} |f_j| |f_{j'}| \cos[2\pi\mathbf{h} \cdot (\mathbf{x}_j - \mathbf{x}_{j'}) + (\delta_j - \delta_{j'})]. \end{aligned}$$

The double sum adds a large number of cosine terms with quasi-randomly distributed phases. As a consequence, this sum vanishes with respect to the single sum that just adds positive terms. Therefore,

$$\begin{aligned} |F(\mathbf{h})| &= I(\mathbf{h})^{1/2} \\ &\cong \left(\sum_j |f_j|^2 \right)^{1/2} + \frac{\sum_j \sum_{j' > j} |f_j| |f_{j'}| \cos[2\pi\mathbf{h} \cdot (\mathbf{x}_j - \mathbf{x}_{j'}) + (\delta_j - \delta_{j'})]}{\left(\sum_j |f_j|^2 \right)^{1/2}}. \end{aligned}$$

One can then derive the first-order approximation of the Bijvoet difference

$$\begin{aligned}\Delta F^\pm &= |F(\mathbf{h})| - |F(-\mathbf{h})| \\ &= -2 \sum_j \sum_{j>j'} |f_j| |f_{j'}| \sin(\delta_j - \delta_{j'}) \sin[2\pi\mathbf{h} \cdot (\mathbf{x}_j - \mathbf{x}_{j'})] \\ &\cong \frac{-2 \sum_j \sum_{j>j'} |f_j| |f_{j'}| \sin(\delta_j - \delta_{j'}) \sin[2\pi\mathbf{h} \cdot (\mathbf{x}_j - \mathbf{x}_{j'})]}{\left(\sum_j |f_j|^2 \right)^{1/2}}.\end{aligned}$$

The developed expression of the squared Bijvoet difference $(\Delta F^\pm)^2$ contains products of sine functions of the form $\sin[2\pi\mathbf{h} \cdot (\mathbf{x}_j - \mathbf{x}_{j'})] \sin[2\pi\mathbf{h} \cdot (\mathbf{x}_{j''} - \mathbf{x}_{j'''})]$. These products can be transformed into sums of cosine terms, $\cos[2\pi\mathbf{h} \cdot (\mathbf{x}_j - \mathbf{x}_{j'} - \mathbf{x}_{j''} + \mathbf{x}_{j'''})]$ and $\cos[2\pi\mathbf{h} \cdot (\mathbf{x}_j - \mathbf{x}_{j'} + \mathbf{x}_{j''} + \mathbf{x}_{j'''})]$. The phases of these cosine terms taken as a whole are distributed quasi-randomly, except for terms for which the phase is zero ($\mathbf{x}_{j''} - \mathbf{x}_{j'''} = \mathbf{x}_j - \mathbf{x}_{j'}$). The contribution of terms with quasi-random phase can be neglected with respect to the contribution of terms whose phases are zero. Therefore,

$$(\Delta F^\pm)^2 \cong \frac{2 \sum_j \sum_{j>j'} |f_j|^2 |f_{j'}|^2 \sin(\delta_j - \delta_{j'})}{\sum_j |f_j|^2}.$$

Since $\delta_j = 0$ for the normal scattering atoms ($j = n$), which represent the major part of the structure, all cross-terms between normal scattering atoms vanish in the numerator. This is also the case for anomalous scatterers ($j = a$) of the same type, which have identical phases. Thus,

$$\begin{aligned}\sum_j \sum_{j>j'} |f_j|^2 |f_{j'}|^2 \sin^2(\delta_j - \delta_{j'}) \\ &= \sum_n |f_n|^2 \sum_a |f_a|^2 \sin^2(\delta_a) + \sum_a \sum_{a'>a} |f_a|^2 |f_{a'}|^2 \sin^2(\delta_{a'} - \delta_a) \\ &\cong \sum_n |f_n|^2 \sum_a |f_a|^2 \sin^2(\delta_a) = \sum_n |f_n|^2 \sum_a f_a'^2.\end{aligned}$$

With

$$\sum_j |f_j|^2 = \sum_n |f_n|^2 + \sum_a |f_a|^2 \cong \sum_n |f_n|^2,$$

we finally obtain

$$(\Delta F^\pm)^2 \cong 2 \sum_a f_a'^2 = 2 \sum_j f_j'^2.$$

For a given reflection \mathbf{h} , the squared Bijvoet difference may deviate from this value, but the mean value of the squared Bijvoet difference, $\langle (\Delta F^\pm)^2 \rangle$, taken over a large number of reflections is very close to this value.

Introducing the mean structure-factor modulus

$$\langle |F| \rangle_\pm = \frac{|F(\mathbf{h})| + |F(-\mathbf{h})|}{2}$$

and taking into account the relationship

$$\langle (|F|)_\pm^2 \rangle \cong \sum_j |f_j|^2,$$

one derives the Bijvoet ratio

$$\frac{\Delta F^\pm}{F} = \left[\frac{\langle (\Delta F^\pm)^2 \rangle}{\langle (|F|)_\pm^2 \rangle} \right]^{1/2} = \left(\frac{2 \sum_j f_j'^2}{\sum_j |f_j|^2} \right)^{1/2}.$$

For a protein composed of N_p atoms excluding H atoms, the definition of Z_{eff} corresponds to the relationship

$$\sum_j |f_j|^2 = N_p Z_{\text{eff}}^2.$$

Therefore,

$$\frac{\Delta F^\pm}{F} = \frac{1}{Z_{\text{eff}}} \left(\frac{2 \sum_j f_j'^2}{N_p} \right)^{1/2}.$$

Finally, for a crystal containing anomalous scatterers with different site occupancies q_j , we obtain

$$\frac{\Delta F^\pm}{F} = \frac{1}{Z_{\text{eff}}} \left(\frac{2 \sum_j q_j^2 f_j'^2}{N_p} \right)^{1/2}.$$

At zero diffraction angle, $Z_{\text{eff}} \cong 6.7 e^-$ for proteins.

The authors would like to thank Dr P. L. Anelli of Bracco Imaging spa, Milan, Italy for kindly providing samples of Gd-HPDO3A, Gd-DO3A, Gd-DOTMA and Gd-HPSA-DO3A; Professor 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 of Sanofi-Synthelabo, Aramon, France for kindly providing urate oxidase; Dr F. Bonneté and Dr D. Vivarès, Université Paris 6-Paris 7, France for discussions on the crystallization of urate oxidase; Dr H. Jouve and Dr P. Andréoletti, IBS, Grenoble, France for kindly providing catalase crystals and Dr V. Stojanoff, BNL, Brookhaven, USA for help in collecting diffraction data from catalase crystals on beamline BM14 at the ESRF; Dr I. Garcia-Saez, IBS, Grenoble, France for kindly providing crystals of β -lactamase from *P. aeruginosa*; Dr G. Sainz, ESRF, Grenoble, France for her strong collaboration on the OTCase project; the BM30A staff for help on the beamline.

References

- Andréoletti, P., Sainz, G., Jaquinod, M., Gagnon, J. & Jouve, H.-M. (2003). *Proteins*, **50**, 261–271.
- Brünger, A. T. (1992). *Nature (London)*, **355**, 472–475.
- Brünger, 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.
- 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. & Rajashankar, K. R. (2000). *Acta Cryst. D* **56**, 232–237.
- Esnouf, R. M. (1997). *J. Mol. Graph. Model.* **15**, 132–134.
- Fourme, R., Shepard, W., Kahn, R., l'Hermite, G. & Li de La Sierra, I. (1995). *J. Synchrotron Rad.* **2**, 36–48.
- Girard, É., Chantalat, L., Vicat, J. & Kahn, R. (2002). *Acta Cryst. D* **58**, 1–9.
- Girard, É., Stelter, M., Anelli, P. L., Vicat, J. & Kahn, R. (2003). *Acta Cryst. D* **59**, 118–126.

- Gouet, P., Jouve, H.-M. & Dideberg, O. (1995). *J. Mol. Biol.* **249**, 933–954.
- Hao *et al.* (2000). *J. Appl. Cryst.* **33**, 980–981.
- Hendrickson, W. A. & Teeter, M. M. (1981). *Nature (London)*, **290**, 107–113.
- Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* **A47**, 110–119.
- La Fortelle, E. de & Bricogne, G. (1997). *Methods Enzymol.* **276**, 472–494.
- Lauretti, L., Riccio, M. L., Mazzariol, A., Cornaglia, G., Amicosante, G., Fontana, R. & Rossolini, G. M. (1999). *Antimicrob. Agents Chemother.* **43**, 1584–1590.
- Miller, R., Gallo, S. M., Khalak, H. G. & Weeks, C. M. (1994). *J. Appl. Cryst.* **27**, 613–621.
- Nguyen, V. T., Baker, D. P., Tricot, C., Baur, H., Villeret, V., Dideberg, O., Gigot, D., Stalon, V. & Haas, D. (1996). *Eur. J. Biochem.* **236**, 283–293.
- 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.
- Ramin, M., Shepard, W., Fourme, R. & Kahn, R. (1999). *Acta Cryst.* **D55**, 157–167.
- Sheldrick, G. M. (1990). *Acta Cryst.* **A46**, 467–473.
- Shepard, W., Kahn, R., Ramin, M. & Fourme, R. (2000). *Acta Cryst.* **D56**, 1288–1303.
- Weeks, C. M. & Miller, R. (1999). *J. Appl. Cryst.* **32**, 120–124.