Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Characterization of gadolinium complexes for SAD phasing in macromolecular crystallography: application to CbpF

Rafael Molina,[a] Meike Stelter,[b]‡ Richard Kahn[b] and Juan A. Hermoso[a]*

[a]Grupo de Cristalografía Macromolecular y Biología Estructural, Instituto Química Física Rocasolano, CSIC, Serrano 119, 28006 Madrid, Spain, and [b]Institut de Biologie Structurale J. P. Ebel, UMR 5075 CEA CNRS UJF, 41 Rue Jules Horowitz, 38027 Grenoble CEDEX 1, France

‡ Present address: European Synchrotron Radiation Facility, BP 220, 38043 Grenoble CEDEX 9, France.

Correspondence e-mail: xjuan@iqfr.csic.es

Received 13 March 2009
Accepted 12 May 2009

Seven Gd complexes were used in the preparation of heavy-atom derivatives for solving the structure of choline-binding protein F (CbpF), a 36 kDa surface protein from *Streptococcus pneumoniae*, by the SAD method. CbpF was used as a model system to analyse the phasing capability of each of the derivatives. Three different aspects have been systematically characterized: the efficacy of cocrystallization *versus* soaking in the binding of the different Gd complexes, their mode of interaction and a comparative study of SAD phasing using synchrotron radiation and using a rotating-anode generator. This study reveals the striking potential of these complexes for SAD phasing using a laboratory source and further reinforces their relevance for high-throughput macromolecular crystallography.

1. Introduction

The single-wavelength anomalous diffraction (SAD) method is now a well established and highly effective approach in structural biology. Recent advances in instrumentation and data-analysis software (notably in phasing and density-modification programs) combined with an increased awareness of radiation damage at third-generation sources have led to a marked increase in the popularity and efficiency of the SAD method. SAD experiments generally involve a single data set collected at or above the absorption edge of an anomalous scatterer incorporated in the sample. SAD experiments generally require fewer data than MAD experiments, depending on the type of heavy atom, its binding to the protein and the strength of its anomalous signal; because of this reduction in the amount of data required for phasing, they are faster and the potential for radiation damage is limited. Moreover, SAD phasing has been reported at remote wavelengths above the absorption edge (Leonard *et al.*, 2005) and even with laboratory X-ray sources (Dauter *et al.*, 2002).

One obstacle associated with heavy-atom-based phasing is the process of obtaining a useful heavy-atom derivative. Nowadays, the majority of anomalous diffraction experiments are carried out using proteins in which methionine residues are substituted by selenomethionine, a method that is readily applicable to recombinant proteins expressed in bacterial or insect-cell systems (Doublié, 1997; Hendrickson *et al.*, 1990).

Alternative methods to obtain derivatives for proteins that cannot be recombinantly produced or for which the seleno-

methionine-substituted material is either insoluble or fails to crystallize have been proposed, including the use of halides (Dauter *et al.*, 1999; Evans & Bricogne, 2002) and of noble gases introduced either into the native protein (Vitali *et al.*, 1991) or into binding sites created by direct mutagenesis (Quillin & Matthews, 2002). The anomalous signal from the intrinsic S atoms in proteins has increasingly been used to solve the structure using the SAD method (Hendrickson & Teeter, 1981; Micossi *et al.*, 2002). However, owing to the weak anomalous signal from sulfur at the wavelengths accessible to most modern protein crystallography beamlines, the method is really only suitable for well diffracting crystals. The traditional heavy-atom soaking or cocrystallization protocols have been revisited in recent years, with the current trend being towards the application of shorter soaking times (Sun & Radaev, 2002; Sun *et al.*, 2002). Despite attempts to develop rational approaches for heavy-atom derivatization using techniques such as native gel electrophoresis (Boggon & Shapiro, 2000) or mass spectrometry (Cohen *et al.*, 2000; Sun & Hammer, 2000), the procedure remains largely empirical.

A new class of Gd complexes have been demonstrated to be a successful tool for obtaining heavy-atom derivative crystals that are suitable for experimental phasing, exploiting the strong anomalous signal at the Gd L_{III} absorption edge ($f'' = 28$ e$^-$, $\lambda = 1.711$ Å) and for Cu $K\alpha$ radiation ($f'' = 12$ e$^-$; Girard, Stelter, Anelli *et al.*, 2003; Girard *et al.*, 2004). Noteworthily, the latter value is higher than the anomalous signal obtained from Se at the K absorption edge ($f'' = 6$–10 e$^-$).

Using Gd complexes, we have previously solved the structures of the pneumococcal surface proteins phosphocholinesterase (Pce; 69.4 kDa; (Hermoso *et al.*, 2005) and choline-binding protein F (CbpF; 39.6 kDa; Molina *et al.*, 2009). Here, we further explore the potential of Gd complexes for SAD phasing by contrasting the phasing power of Gd derivatives obtained by cocrystallization or by soaking methods and by comparing the diffraction data obtained using synchrotron radiation or X-rays from a rotating-anode laboratory source.

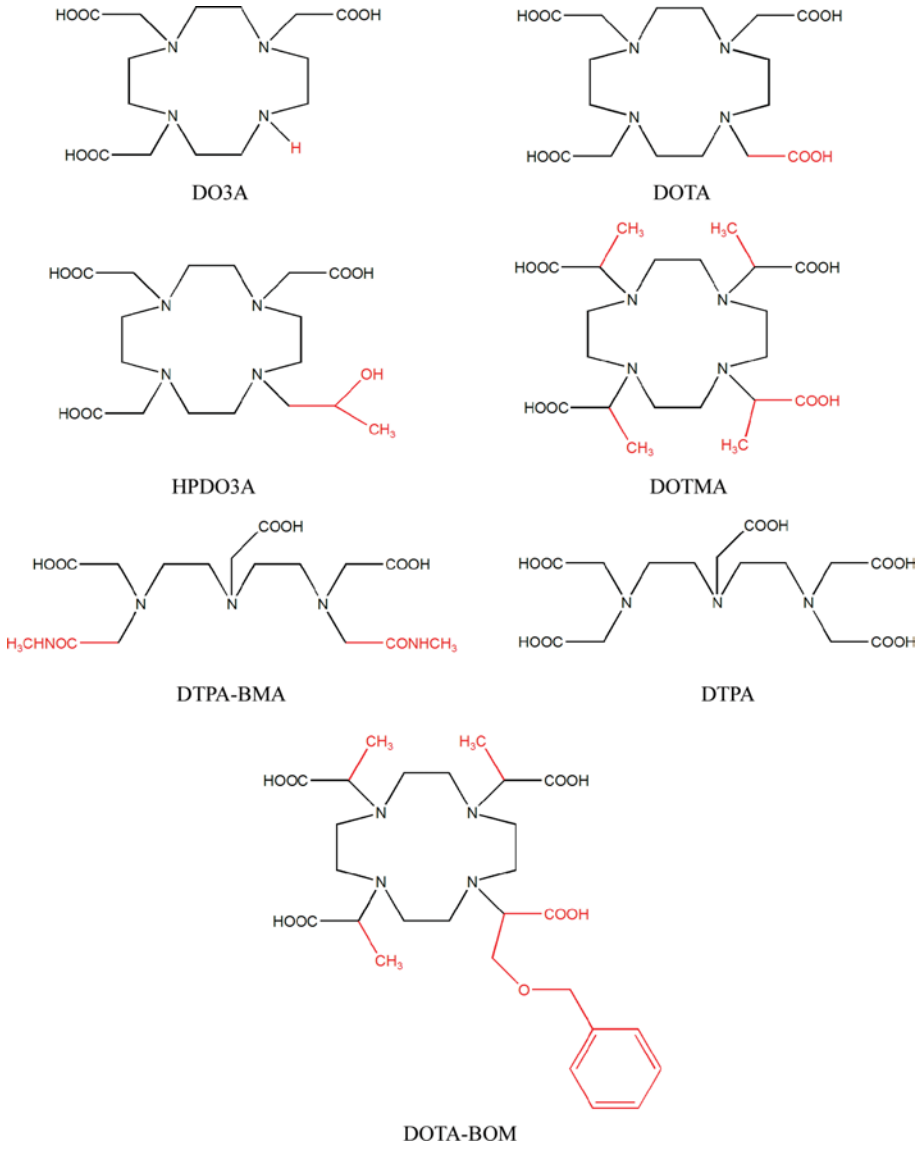

Figure 1
The different ligands constituting the seven Gd complexes presented in this study. DO3A, 1,4,7,10-tetraazacyclododecane-1,4,7-triacetic acid; DOTA, 1,4,7,10-tetraazacyclododecane-1,4,7,10-tetraacetic acid; HPDO3A, 10-(2-hydroxypropyl)-1,4,7,10-tetraazacyclododecane-1,4,7-triacetic acid; DOTMA, *a*,*a*′,*a*″,*a*‴-tetramethyl-1,4,7,10-tetraazacyclododecane-1,4,7,10-tetraacetic acid; DTPA-BMA, diethylenetriaminepentaacetic acid bismethylamide; DTPA, diethylenetriaminepentaacetic acid; DOTA-BOM, (phenylmethoxy)methyl-1,4,7,10-tetraazacyclododecane-1,4,7,10-tetraacetic acid.

2. Methods

2.1. Gadolinium complexes

The seven gadolinium complexes used in this study are composed of ligand molecules (Fig. 1) that chelate a single Gd^{3+} ion and are commonly used in medical imaging (Port *et al.*, 2008). The core of the HPDO3A, DO3A, DOTMA and DOTA ligands is a tetraazacyclododecane macrocycle, while DTPA and DTPA-BMA are linear molecules. When chelating the Gd^{3+} ion, the carboxylic groups of the ligand molecules are deprotonated. Therefore, the resulting complexes are neutral apart from Gd-DOTA and Gd-DOTMA, which both carry a single negative global charge, and Gd-DTPA, which bears a double negative charge. Gd-HPDO3A, Gd-DO3A and Gd-DOTMA are soluble in water at concentrations higher than 1 *M*; Gd-DOTA, Gd-DOTA-BOM, Gd-DTPA and Gd-

Table 1
Data-collection, refinement and phasing statistics for the seven gadolinium derivatives obtained by cocrystallization.

Diffraction data were collected using synchrotron radiation at the Gd L_{III} absorption edge from CbpF derivative crystals obtained by soaking and by cocrystallization with the seven Gd complexes. Values in parentheses are for the highest resolution shell. The resolution in phasing statistics was limited to 2.98 Å for all data sets.

	Gd-HPDO3A	Gd-DOTMA	Gd-DTPA	Gd-DO3A	Gd-DTPA-BMA	Gd-DOTA-BOM	Gd-DOTA
Data collection							
Unit-cell parameters							
a (Å)	49.35	49.18	48.42	48.98	48.71	49.03	49.19
b (Å)	114.52	115.00	113.49	114.92	114.00	114.37	114.24
c (Å)	76.23	75.80	75.38	76.25	76.29	76.07	75.66
Resolution range (Å)	100–2.23 (2.36–2.23)	100–2.23 (2.36–2.23)	100–2.23 (2.35–2.23)	100–2.29 (2.42–2.29)	100–2.98 (3.14–2.98)	100–2.23 (2.35–2.23)	100–2.30 (2.42–2.30)
No. of observed reflections	143495	144217	144177	77081	60403	118066	134747
No. of unique reflections	21579	21372	20726	19934	9076	20962	19830
Acentric	18835	18571	18108	17349	7528	18338	17202
Centric	2744	2801	2618	2585	1548	2624	2628
Signal-to-noise ratio $I/\sigma(I)$	10.4 (7.7)	10.5 (4.6)	3.7 (3.6)	11.3 (5.4)	8.3 (3.1)	7.5 (1.3)	9.2 (7.4)
Redundancy	6.6 (5.5)	6.7 (5.8)	7.0 (7.0)	3.9 (3.7)	6.7 (5.1)	5.6 (3.6)	6.8 (6.7)
Completeness (%)	98.9 (95.0)	99.7 (98.4)	98.7 (95.5)	98.6 (93.9)	98.9 (92.8)	97.2 (96.1)	99.4 (97.7)
R_{merge} (%)	4.7 (8.9)	5.4 (15.0)	10.9 (17.1)	4.8 (12.4)	7.1 (19.7)	4.9 (13.7)	5.4 (10.0)
R_{ano} (%)	10.6 (5.9)	4.5 (5.4)	9.2 (11.7)	10.1 (12.5)	4.3 (7.3)	5.6 (9.0)	5.4 (6.6)
$\Delta F/\sigma$†	4.249	1.541	1.760	2.788	1.307	2.164	1.852
Phasing statistics							
No. of sites found by *SHELXD*	5	1	2	5	4	2	2
No. of sites finally modelled‡	4	1	1	3	2	1	1
Site occupancies§	0.81, 0.68, 0.47, 0.38	0.41	0.84	0.74, 0.64, 0.61	0.34, 0.18	0.82	0.68
$\langle\Delta F\rangle/F$	0.100	0.037	0.069	0.095	0.032	0.067	0.056
$\Delta F/\sigma$†	5.496	1.997	1.865	3.629	1.307	2.711	2.246
FOM before *DM*§	0.588	0.371	0.456	0.537	0.392	0.416	0.391
R_{cullis} for acentric reflections§	0.298	0.752	0.549	0.436	0.632	0.695	0.737
Phasing power for acentric reflections§	4.905	1.371	2.340	3.231	1.949	1.854	1.441
FOM after *DM*§	0.860	0.765	0.821	0.844	0.802	0.812	0.778
Free *R* factor in real space§	0.093	0.104	0.088	0.100	0.101	0.104	0.107
Map correlation with *OVERLAPMAP*	0.724	0.584	0.672	0.711	0.626	0.654	0.605

† Data from *XDS*. ‡ Data from *O*. § Data from *SHARP*.

DTPA-BMA were provided at a concentration of 0.5 *M*. Gd-HPDO3A, Gd-DO3A, Gd-DOTMA and Gd-DOTA-BOM were kindly provided by Bracco Imaging SpA, Milan, Italy. Commercial solutions of Gd-DOTA, Gd-DTPA and Gd-DTPA-BMA were kindly provided by Professor J.-F. Le Bas, CHU-Hôpital Nord, Grenoble, France; Eu and Yb complexes of HPDO3A, DO3A, DOTA and DTPA-BMA are now distributed by the company NatX-ray, Grenoble, France (contact@natx-ray.com).

2.2. Crystallization

The crystallization strategy of *Streptococcus pneumoniae* CbpF has been described elsewhere (Molina *et al.*, 2007). In brief, crystals of CbpF were grown by the hanging-drop vapour-diffusion technique using a protein solution containing 140 m*M* choline chloride, 20 m*M* Tris–HCl buffer pH 8.0 and 3.9 mg ml^{-1} protein. The best crystals were obtained in 0.01 *M* $NiCl_2$, 20% PEG MME 2K, 0.1 *M* Tris pH 8.5 in drops containing 3 µl protein solution and 1 µl precipitant solution. Derivative crystals were prepared by cocrystallization or by soaking using the native crystallization conditions (see above) plus 50 m*M* of each Gd complex. The best crystals were obtained with drops containing 3 µl protein solution and 1 µl heavy-atom solution. For all soaked crystals the optimized soaking time was around 10 s (a common soaking time which preserved crystal quality with all complexes), which is significantly shorter than previously reported soaking times (Girard, Stelter, Anelli *et al.*, 2003). All derivative crystals were isomorphous to the $P2_12_12$ native crystals.

2.3. Data collection and processing

All crystals were back-soaked for several seconds in the crystallization solution complemented with 17% glycerol as a cryoprotectant to remove any unbound gadolinium complex from the crystals. All the crystals were cryocooled directly in a nitrogen-gas stream at 100 K (Oxford Cryosystems).

Diffraction data were collected at the ESRF on beamline ID29 for all cocrystallized derivative crystals and on beamline BM30A for soaked derivative crystals (Gd-HPDO3A, Gd-DOTMA and Gd-DTPA derivatives). Data collections were carried out at the wavelength corresponding to the white line (λ = 1.711 Å) of the L_{III} absorption edge of Gd as determined from X-ray fluorescence spectra. The crystal-to-detector distance was 110 mm on both beamlines. The crystals typically diffracted to 1.9 Å resolution, but the high-resolution limit was limited by the detector size to 2.23 Å on ID29 and to 2.68 Å on BM30A. Data were recorded over an angular range of 180° for

Table 2
Data-collection, refinement and phasing statistics comparing soaking and cocrystallization methods.

Diffraction data were collected using synchrotron radiation at the Gd L_{III} absorption edge from CbpF derivative crystals obtained by soaking and by cocrystallization with three Gd complexes. Values in parentheses are for the highest resolution shell. The resolution in phasing statistics was limited to 2.70 Å for all data sets.

	Soaking			Cocrystallization		
	Gd-HPDO3A	Gd-DOTMA	Gd-DTPA	Gd-HPDO3A	Gd-DOTMA	Gd-DTPA
Data collection						
Unit-cell parameters						
a (Å)	48.97	48.89	48.87	49.35	49.18	48.42
b (Å)	114.93	114.86	114.06	114.52	115.00	113.49
c (Å)	76.17	76.16	75.76	76.23	75.80	75.38
Resolution range (Å)	100–2.70 (2.85–2.70)	100–2.70 (2.85–2.70)	100–2.70 (2.85–2.70)	100–2.23 (2.36–2.23)	100–2.23 (2.36–2.23)	100–2.23 (2.35–2.23)
No. of observed reflections	78896	76577	78780	143495	144217	144177
No. of unique reflections	11791	11598	11879	21579	21372	20726
Acentric	9999	9761	9995	18835	18571	18108
Centric	1792	1837	1884	2744	2801	2618
Signal-to-noise ratio $I/\sigma(I)$	18.6 (3.0)	10.3 (1.5)	8.6 (1.6)	10.4 (7.7)	10.5 (4.6)	3.7 (3.6)
Redundancy	6.7 (4.8)	6.6 (4.5)	6.6 (4.7)	6.6 (5.5)	6.7 (5.8)	7.0 (7.0)
Completeness (%)	99.3 (98.4)	96.3 (96.1)	100 (99.9)	98.9 (95.0)	99.7 (98.4)	98.7 (95.5)
R_{merge} (%)	3.0 (14.0)	4.5 (21.9)	3.8 (21.0)	4.7 (8.9)	5.4 (15.0)	10.9 (17.1)
R_{ano} (%)	5.3 (10.3)	6.9 (18.7)	4.3 (15.8)	10.6 (5.9)	4.5 (5.4)	9.2 (11.7)
$\Delta F/\sigma$†	3.614	2.951	2.370	4.249	1.541	1.760
Phasing statistics						
No. of sites found by *SHELXD*	4	4	4	5	1	2
No. of sites finally modelled‡	4	2	3	4	1	1
Site occupancies	0.41, 0.24, 0.22, 0.21	0.51, 0.41	0.42, 0.19, 0.12	0.55, 0.53, 0.40, 0.31	0.39	0.74
$\langle\Delta F\rangle/F$	0.046	0.054	0.038	0.075	0.032	0.061
$\Delta F/\sigma$†	4.081	3.331	2.539	5.136	1.855	1.833
FOM after ML refinement§	0.518	0.509	0.491	0.567	0.346	0.441
R_{cullis} for acentric reflections§	0.424	0.455	0.552	0.313	0.773	0.586
Phasing power for acentric reflections§	3.256	2.990	2.405	4.745	1.257	2.194
FOM after *DM*§	0.861	0.837	0.802	0.845	0.741	0.801
Free *R* factor in real space§	0.113	0.104	0.088	0.101	0.107	0.096
Map correlation with *OVERLAPMAP*	0.701	0.686	0.667	0.718	0.566	0.641

† Data from *XDS*. ‡ Data from *SHARP*. § Data from *O*.

all derivatives in order to obtain highly redundant data and thereby improve the precision of the data measurements.

To assess the anomalous phasing potential of these compounds when using conventional laboratory X-ray sources, an in-house data collection was carried out with a cocrystallized Gd-HPDO3A-derivative crystal. In-house data were collected on a MAR345 image-plate detector using Cu $K\alpha$ radiation from an FR571 Enraf–Nonius rotating-anode generator.

All data sets collected at ESRF were processed using *XDS* (Kabsch, 1988) and scaled with *SCALA* (Evans, 1993) from the *CCP*4 package (Collaborative Computational Project, Number 4, 1994). Data sets collected with Cu $K\alpha$ radiation were integrated using *MOSFLM* (Leslie, 1987) and scaled with *SCALA*. A summary of data-collection parameters and processing statistics is given in Tables 1, 2 and 3.

2.4. SAD phasing and refinement

Seven different isomorphous Gd derivatives were produced by cocrystallization using the native crystallization conditions supplemented with the corresponding complex at a concentration of 50 m*M* (concentrations of 20, 50 and 100 m*M* were tested, with 50 m*M* being the most successful). All data sets were collected on the same beamline (ID29 at ESRF). The heavy-atom positions were determined with *SHELXD* (Sheldrick, 2008) for each derivative and their coordinates were used in the phasing protocol implemented in *SHARP* (de La Fortelle & Bricogne, 1997). Phases were subsequently improved by solvent flattening and histogram matching using the program *DM* (Cowtan & Main, 1996), assuming 58% solvent content as estimated by the *CCP*4 program *TRUNCATE*. The number of Gd sites as well as the phasing statistics for the derivative crystals obtained with the seven different Gd complexes are given in Table 1. In order to study the protein–ligand interactions, binding sites were visualized with the graphics program *O* (Jones *et al.*, 1991). Refinement of the ligand positions was carried out with *CNS* (Brünger *et al.*, 1998).

For three of the seven complexes (Gd-HPDO3A, Gd-DOTMA and Gd-DTPA), derivative crystals were obtained by both cocrystallization and soaking using the same crystallization conditions and the same complex concentration (50 m*M*). The procedure for identifying heavy-atom positions and phase calculation and modification was the same as described above. The final number of Gd sites and the phasing statistics for CbpF derivatives obtained by cocrystallization and by soaking are given in Table 2.

Phasing statistics for the diffraction data of the Gd-HPDO3A derivative crystal collected using synchrotron radiation and using Cu $K\alpha$ radiation are given in Table 3.

Table 3
Data-collection, refinement and phasing statistics comparing synchrotron and rotating-anode data.

Data from CbpF derivative crystals obtained by cocrystallization with the Gd-HPDO3A complex and collected using in-house rotating-anode Cu $K\alpha$ radiation and using synchrotron radiation. Values in parentheses are for the highest resolution shell. The resolution in phasing statistics was limited to 2.39 Å for both data sets.

	Rotating anode	Synchrotron (ESRF)
Data collection		
Unit-cell parameters		
a (Å)	49.04	49.35
b (Å)	114.86	114.52
c (Å)	76.05	76.23
Resolution range (Å)	100–2.39 (2.51–2.39)	100–2.23 (2.36–2.23)
No. of observed reflections	115647	143495
No. of unique reflections	16575	21579
Acentric	14195	18835
Centric	2380	2744
Signal-to-noise ratio $I/\sigma(I)$	13.7 (2.3)	10.4 (7.7)
Redundancy	7.0 (6.9)	6.6 (5.5)
Completeness (%)	99.6 (97.1)	98.9 (95.0)
$R_{\rm merge}$ (%)	4.9 (33.9)	4.7 (8.9)
$R_{\rm ano}$ (%)	4.5 (15.6)	10.6 (5.9)
$\Delta F/\sigma$†	1.661	4.249
Phasing statistics		
No. of sites found by *SHELXD*	4	5
No. of sites finally modelled‡	4	4
Site occupancies	0.51, 0.28, 0.20, 0.19	0.64, 0.63, 0.49, 0.38
$\langle\Delta F\rangle/F$	0.023	0.090
$\Delta F/\sigma$†	1.661	4.654
FOM after ML refinement§	0.426	0.521
$R_{\rm cullis}$ for acentric reflections§	0.719	0.352
Phasing power for acentric reflections§	1.581	4.546
FOM after *DM*	0.802	0.823
Free *R* factor in real space§	0.096	0.103
Map correlation with *OVERLAPMAP*	0.589	0.682

† Data from *XDS*. ‡ Data from *O*. § Data from *SHARP*.

The CbpF structure was modelled manually with *O* (Jones *et al.*, 1991) and refined with *CNS* (Brünger *et al.*, 1998). The correlation coefficients between experimental electron-density maps after density modification and maps calculated from the final refined models were determined using the program *OVERLAPMAP*. Statistics for the refinement of all Gd derivatives obtained are summarized in Supplementary Tables S1, S2 and S3.[1]

3. Results and discussion

3.1. Phasing capability of Gd derivatives

The derivatization of CbpF by either cocrystallization or soaking protocols was straightforward and did not require significant changes to the original crystallization conditions. The high solubility of the complexes permitted the use of concentrations considerably higher than those generally used in the preparation of heavy-atom derivatives. Despite this high concentration, derivative crystals with good diffraction characteristics were routinely obtained and in some cases they diffracted to higher resolution than the native crystals. For instance, while the best native CbpF crystal diffracted to 1.67 Å resolution (Molina *et al.*, 2007), a Gd-DOTA derivative crystal provided good-quality data to 1.40 Å resolution. In these cases, new interactions mediated by Gd complexes that further stabilized the crystal packing were observed.

The success of Gd phasing depends on the number of binding sites and their occupancies. The strong anomalous scattering of the Gd atoms, combined with the high binding-site occupancies, yielded large anomalous signals as demonstrated directly by the high values of $\Delta F/\sigma$ (*XDS*) and $R_{\rm ano}$ (*SCALA*), by the peak heights in anomalous Patterson maps and by the phasing parameters ($R_{\rm cullis}$, phasing power and figure of merit). For all derivative crystals, the $\Delta F/\sigma$ and $R_{\rm ano}$ values were in the ranges 1.30–4.25 and 4.3–12.1%, respectively (Table 1). The lowest values corresponded to the cocrystallized Gd-DTPA-BMA derivative. In all cases phasing was successful, demonstrating that all the Gd complexes bound to CbpF, which was confirmed from the anomalous Patterson maps (Fig. 2).

All derivatives provided phases of sufficient quality to facilitate model building (Table 1). Gd-HPDO3A led to the highest figure of merit (0.588) and the lowest $R_{\rm cullis}$ factor (0.298) after phasing with *SHARP*. After density modification, the figure of merit for this derivative was 0.860. The quality of the phases obtained with this derivative crystal allowed straightforward model tracing using the experimental map at 2.23 Å resolution. Gd-DO3A and Gd-DTPA derivatives also provided excellent phasing statistics, with figures of merit of greater than 0.45 before density modification. Even for Gd-DTPA-BMA-derivative and Gd-DOTMA-derivative crystals, the figures of merit were 0.37 and 0.39, respectively, before *DM* and were greater than 0.75 after density modification, allowing easy model building.

3.2. Cocrystallization *versus* soaking

Three of the seven gadolinium complexes were used to prepare derivatized crystals by both soaking and cocrystallization. High $\Delta F/\sigma$ values (greater than 1.5 in all cases) were obtained from *XDS*, indicating a significant anomalous signal in each case. The phasing statistics corresponding to the soaked crystals (Table 2) confirmed the effective binding of all three complexes. For the three derivatives analysed, the figures of merit before density modification were greater than 0.49 and the free *R* factors in real space were lower than 0.12, reflecting the quality of the initial phases. Moreover, high values of phasing power (greater than 2.4), low values of $R_{\rm cullis}$ (lower than 0.55) and high figures of merit after density modification (greater than 0.77) were obtained. Hence, both cocrystallization and soaking protocols facilitated the derivatization of the crystals and led to successful phasing of the structure.

The best derivative crystals, obtained with Gd-HPDO3A, were characterized by several binding sites with high occu-

[1] Supplementary material has been deposited in the IUCr electronic archive (Reference: DZ5159). Services for accessing this material are described at the back of the journal.

pancies (Table 2) using either protocol. Four Gd-HPDO3A molecules were bound to the protein, with higher occupancies for the cocrystallized derivative (0.81, 0.68, 0.47 and 0.38) than for the soaked crystal (0.41, 0.24, 0.22 and 0.21), which explains the better phasing results obtained with cocrystallized derivative crystals.

In contrast, for the remaining two Gd complexes, Gd-DOTMA and Gd-DTPA, the phasing statistics were better for the soaked derivatives than for the cocrystallized samples. For the cocrystallized Gd-DOTMA derivative only one molecule was bound to the protein, with a binding-site occupancy of 0.41, while in the soaked derivative two molecules were bound to the protein with occupancies of 0.51 and 0.41, respectively. Similarly, for the Gd-DTPA derivatives the cocrystallized derivative crystal presented one highly occupied site (0.74) while the soaked derivative revealed two binding sites with lower occupancies.

With each of the complexes analysed, the number of Gd-binding sites was equal or higher for the soaked crystals compared with the cocrystallized derivatives. Conversely, the binding-site occupancies of the cocrystallized derivative crystals were generally higher than for the soaked crystals, which is further illustrated by the intensities of the peaks in the anomalous Patterson maps (data not shown).

It is worth noting that although the soaking times were very short (no more than 10 s), they resulted in effective binding of the complex, generating good phasing. The small differences observed between the soaking and cocrystallization methods could be related to the high accessibility of the binding sites in this case. In agreement with this, it has been reported that

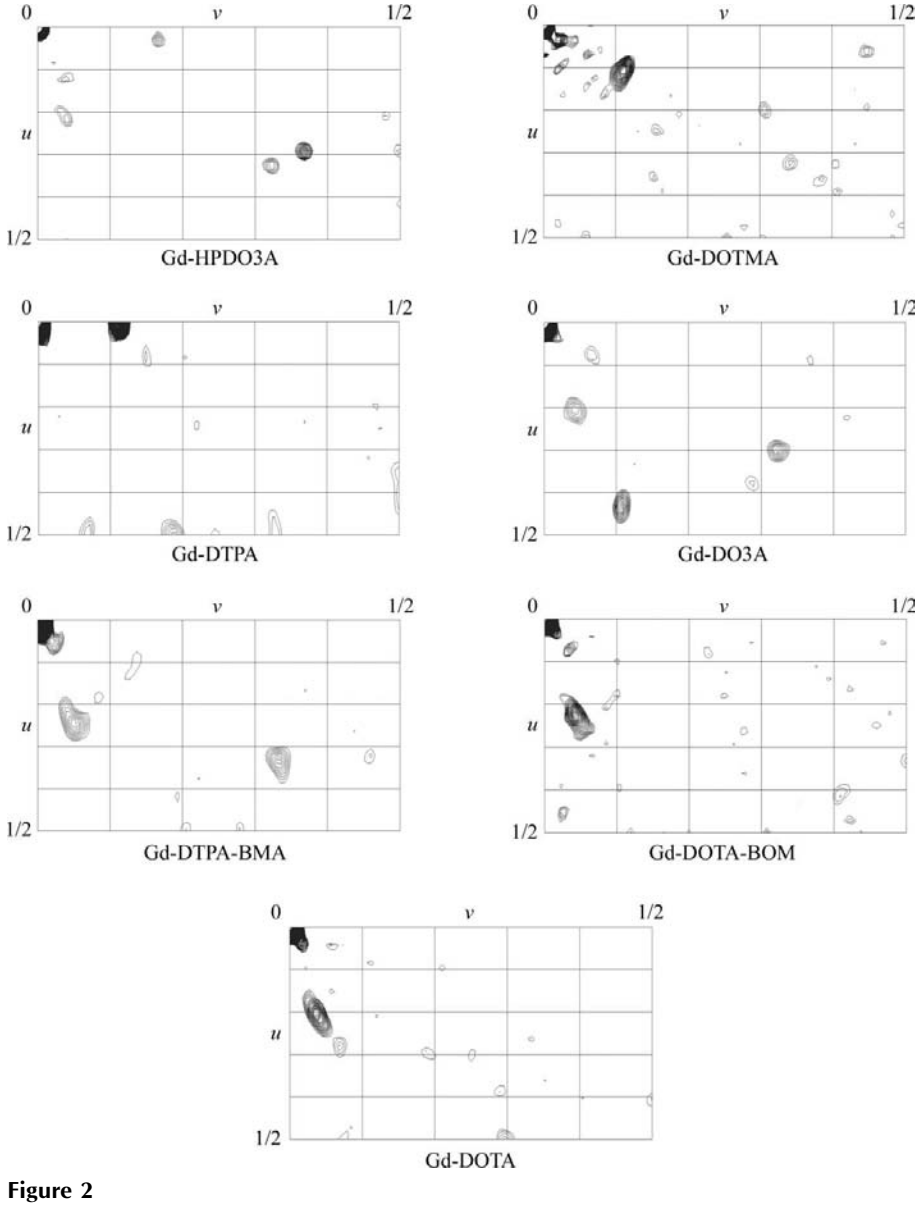

Figure 2
$w = 0$ Harker sections of anomalous Patterson maps for the CbpF derivative crystals obtained with the seven Gd complexes. Levels are contoured in steps of 0.5σ starting from 2σ.

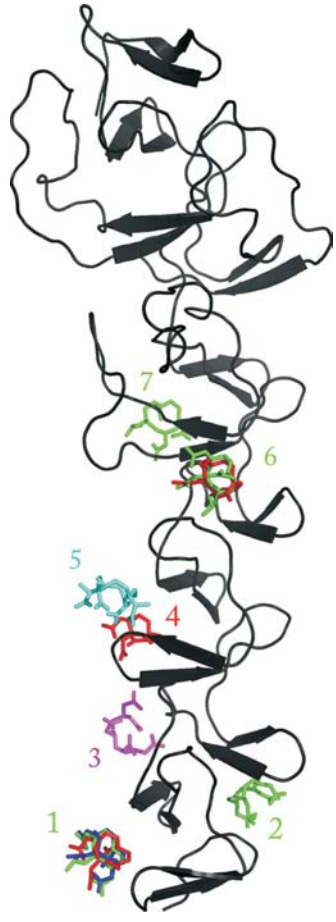

Figure 3
Location of the different binding sites, numbered 1–7. The ligands of the complexes are coloured green (HPDO3A), red (DO3A), cyan (DOTMA), magenta (DTPA) and blue (DOTA). The DOTA-BOM and DTPA-BMA ligands were not well defined in the electron-density map, but the Gd atom was localized in site 1 in both cases.

cocrystallization was the only way to prepare derivatized crystals when binding sites were inaccessible to the Gd complexes in the crystal (Girard, Stelter, Vicat *et al.*, 2003).

3.3. Binding mode of the different Gd complexes

The number of ligand-binding sites and their location was analysed for the seven complexes (Fig. 3). All the binding sites were located in the C-terminal module of CbpF. Differences in stereochemical and net global charge between the complexes resulted in binding sites that differed in number and in location. The carboxylate extensions of the ligand molecule (Fig. 1) bend towards the gadolinium ion in all complexes, forming a cage around the charged ion (Fig. 4). Three main interaction modes were observed: (i) CH–π hydrophobic interactions (Brandl *et al.*, 2001) were typically observed for electrically neutral complexes (Gd-HPDO3A, Gd-DO3A and Gd-DTPA-BMA) and for two complexes bearing a single negative global charge (Gd-DOTA and Gd-DOTA-BOM), (ii) polar interactions played a role in the binding of the Gd-DOTMA ligand and (iii) both electrostatic and polar interactions were implicated in the binding of the Gd-DTPA ligand.

CH–π interactions play a major role in the binding of five of the seven Gd complexes and generate the highest number of ligand-binding sites (ranging from one to four depending on the ligand). This kind of interaction was observed at several different sites, the most common of which (observed for all the five ligands built around a tetraazacyclododecane macrocycle) corresponded to a large hydrophobic pocket (site 1) formed by the side chains of four aromatic residues (three Trp residues and a Tyr residue; Fig. 4*a*). In this case the macrocycle is fully encapsulated within an approximately square cavity (around 8 Å in size). The remaining hydrophobic binding sites, sites 2, 4, 6 and 7 (Figs. 4*b*, 4*d*, 4*f* and 4*g*), stabilize the ligand through interactions with the three aromatic residues (two structurally conserved Trp residues and one Tyr residue) that are implicated in choline binding (Molina *et al.*, 2009). In these smaller aromatic pockets the macrocycle is oriented towards one of the Trp residues. Binding of the ligand is mainly achieved through CH–π interactions, but polar interactions with an Asn residue, sometimes mediated by one water molecule, are also observed

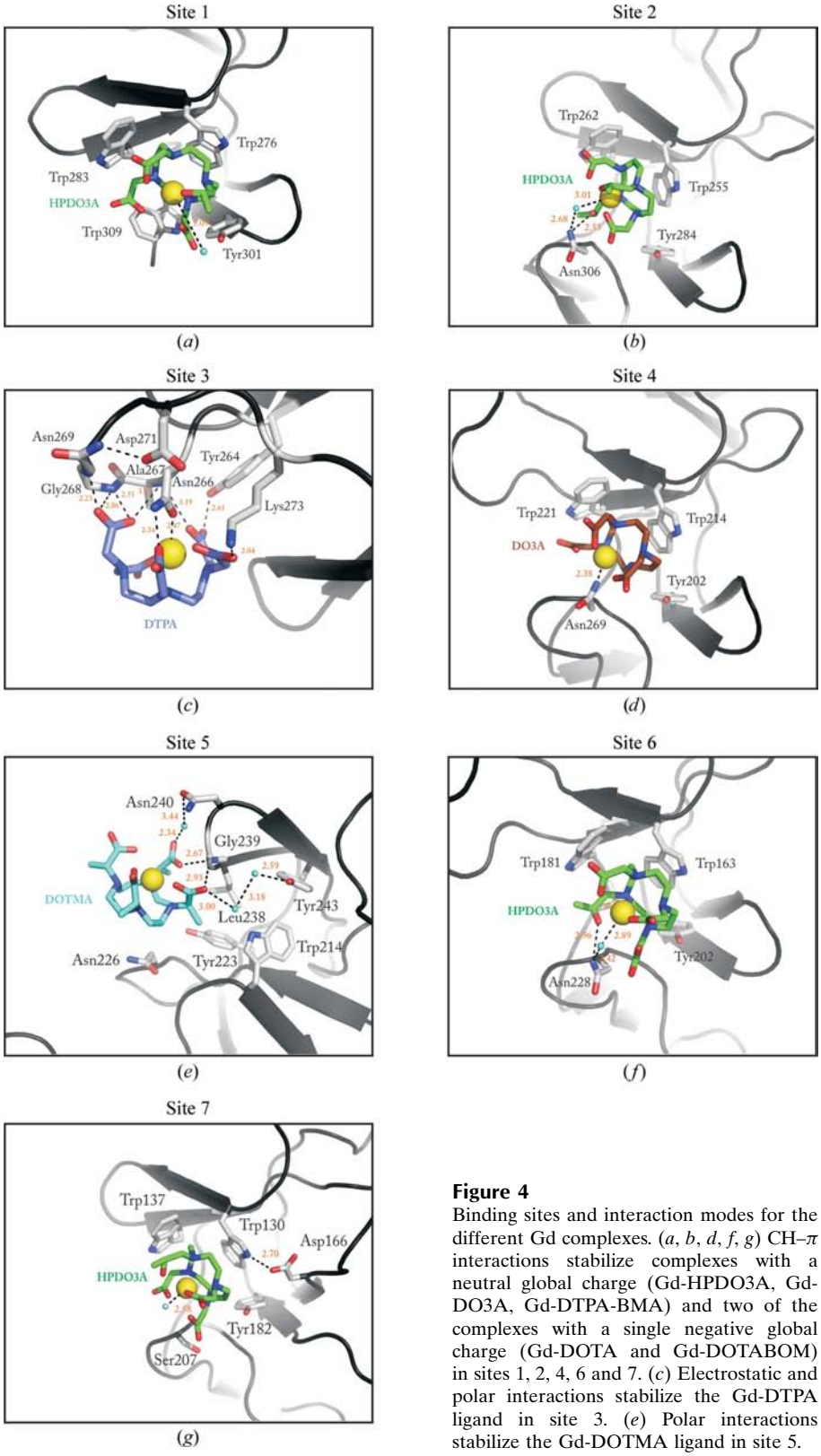

Figure 4
Binding sites and interaction modes for the different Gd complexes. (*a*, *b*, *d*, *f*, *g*) CH–π interactions stabilize complexes with a neutral global charge (Gd-HPDO3A, Gd-DO3A, Gd-DTPA-BMA) and two of the complexes with a single negative global charge (Gd-DOTA and Gd-DOTABOM) in sites 1, 2, 4, 6 and 7. (*c*) Electrostatic and polar interactions stabilize the Gd-DTPA ligand in site 3. (*e*) Polar interactions stabilize the Gd-DOTMA ligand in site 5.

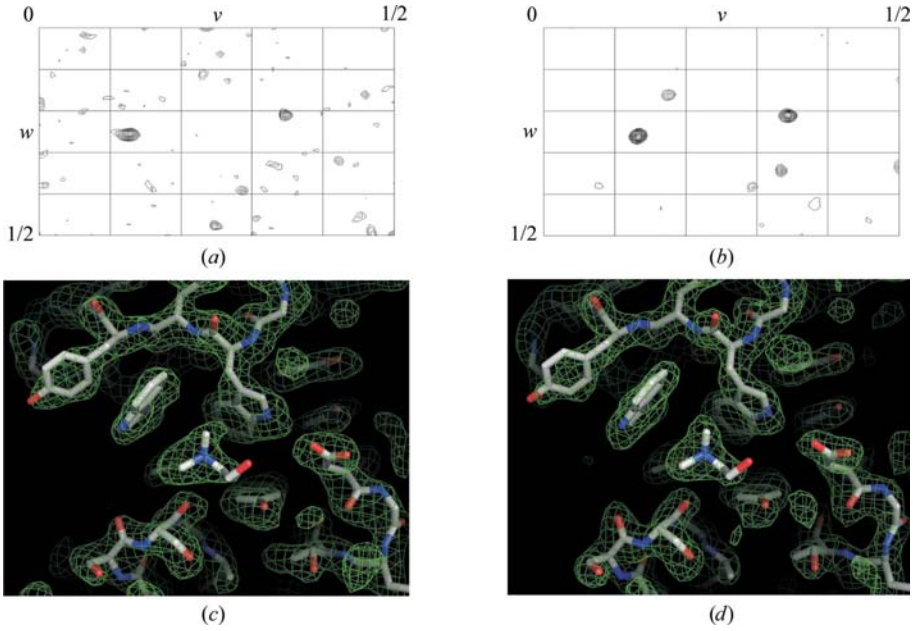

Figure 5
Comparison of rotating-anode and synchrotron data. $u = \frac{1}{2}$ Harker sections from the anomalous Patterson map for a CbpF derivative crystal obtained with the Gd-HPDO3A complex computed from synchrotron data (*b*) and from data collected using an in-house rotating anode (*a*). Experimental electron-density maps from the Gd-HPDO3A-derivatized crystals collected with in-house Cu $K\alpha$ radiation (*c*) and with synchrotron radiation at the Gd L_{III} absorption edge (*d*). The images depict the experimental electron-density maps in the region of a choline-binding site.

(Figs. 4*b*, 4*d* and 4*f*). Interestingly, the linear Gd-DTPA-BMA complex adopts a similar conformation to the complexes based on the macrocycle and interacts with the protein in a similar fashion. Smaller Gd complexes such as Gd-DO3A and Gd-HPDO3A manifest the highest number of sites (three and four, respectively), presumably because they can be more easily accommodated into hydrophobic pockets of variable size than larger complexes. However, the highest occupancy (0.82) among the ligands interacting through CH–π interactions is attained by Gd-DOTA-BOM (Table 1). This result presumably arises from the presence of an extra phenyl ring, allowing Gd-DOTA-BOM to interact with aromatic residues through both the macrocycle and the phenyl ring.

Another type of binding mode was observed for Gd-DOTMA (site 5). This complex, which presents a single negative global charge, binds to the protein through hydrogen bonds to the main-chain amides of the backbone (*e.g.* Gly239) *via* the two carboxylate chains of the ligand and through two other hydrogen bonds mediated by water molecules (Fig. 4*e*). Two methyl groups of DOTMA also contact Leu238, further stabilizing the binding of the complex. No interaction with the macrocycle was observed for this ligand and only one site was occupied (with an occupancy of 0.41), indicating a weaker binding mode.

The binding of the Gd-DTPA complex to CbpF involves coordination of the Gd^{3+} ion by the amide O atom of Asn266 and the formation of a strong salt bridge (2.04 Å) between Lys273 and one of the carboxylic groups of the ligand (Fig. 4*c*). This interaction is reinforced by a network of seven hydrogen bonds between the O atoms of the carboxylic groups of the ligand and the side chains of Asn266, Tyr264 and two main-chain amides (Ala267 and Gly268; Fig. 4*e*). In contrast to the scenario observed for other complexes, the carboxylic groups of Gd-DTPA are oriented towards the protein not the solvent. This mode of interaction is also observed for the only ligand presenting two negative charges. The linear structure of DTPA probably induces a high plasticity of the ligand, thus maximizing the number of potential interactions with the protein. Its high occupancy (0.84), the highest observed for all seven complexes, reflects the strength of this interaction. For the CbpF–Gd-DTPA complex the electron density was good enough to allow us to build and refine the structure of the Gd complex itself, the structure of which has not been previously reported. In the refined structures the coordination sphere of the Gd ion in the different complexes is usually completed by a water molecule, although amino-acid side chains have also been implicated in this role.

In summary, three principal modes of binding have been observed for the seven complexes with the same protein. While the CH–π interaction is the most frequent, electrostatic interactions also led to high binding-site occupancies. This versatility augments the potential for the exploitation of these ligands for preparing heavy-atom derivatives with different proteins.

3.4. Synchrotron radiation *versus* rotating anode

Diffraction data from a Gd-HPDO3A-derivative crystal collected on an in-house diffractometer were compared with data collected using synchrotron radiation in order to assay the phasing potential of these compounds for use with conventional X-ray sources (Table 3). High values of R_{ano} and $\Delta F/\sigma$ (4.5% and 1.66, respectively) indicated a substantial anomalous signal. Harker sections of the anomalous Patterson maps clearly revealed the presence of anomalous scatterers (Fig. 5). Four Gd sites were identified which corresponded to the sites found using the synchrotron data (phasing statistics are shown in Table 3). The figure of merit was around 0.45 before density modification. A phasing power of 1.74, an R_{cullis} of 0.64 and a figure of merit of 0.84 were obtained following density modification. A well defined experimental electron-density map allowed us to easily build the model. These results confirmed the suitability of Gd complexes for SAD phasing using X-rays from a laboratory source.

Not surprisingly, the phasing results for the data collected using synchrotron radiation at the Gd L_{III} absorption edge were better than those obtained using the in-house rotating-

anode generator. However, the anomalous signal and phases obtained using an in-house X-ray source led to well defined experimental electron-density maps of sufficient quality to allow straightforward model building. From a Gd-HPDO3A-derivative crystal obtained by cocrystallization, data were collected with synchrotron radiation at the Gd L_{III} absorption edge, providing an excellent anomalous signal to 2.23 Å resolution which, when combined with phase extension, allowed us to obtain an initial experimental map at 1.91 Å resolution (see Fig. 5*d*). The high quality of the map permitted the easy tracing and enabled model building of most side chains. Similarly, data collected from a Gd-HPDO3A-derivative crystal obtained by cocrystallization using a conventional X-ray source allowed the determination of phases at 2.93 Å resolution, leading to a well defined experimental electron-density map at 1.86 Å resolution after solvent flattening and density modification (Fig. 5*c*). Although the map obtained using the in-house source showed less detailed information than that obtained using synchrotron radiation, the difference between both maps was insignificant, as can be observed for a choline-binding site.

4. Conclusions

The seven Gd complexes presented in this study are excellent candidates for experimental phasing methods based on anomalous scattering. They were used at high concentrations to obtain derivative crystals by soaking or cocrystallization with our test protein and SAD experiments led to electron-density maps that allowed unambiguous building of the protein model in all cases.

Soaking and cocrystallization methods were found to be equally suited to providing good-quality derivative crystals. Whereas the number of ligand-binding sites was equal or higher for the soaked crystals, the binding-site occupancies were generally higher for the cocrystallized crystals.

These Gd complexes also offer the possibility of obtaining good phases using diffraction data from an in-house Cu $K\alpha$ X-ray source by making use of their f'' of 12 e^- at this wavelength. Introduction of the complexes into the crystals did not impair their diffraction quality, unlike as often happens with conventional heavy-atom compounds.

Three different types of binding interactions have been observed for our test protein depending on the chemical nature of the ligand. Since for a given protein binding sites may be different depending on the ligand, these complexes could be used as a mixture to take advantage of their different binding modes. The versatility of the tetraazacyclododecane gadolinium complexes, together with their high phasing power even using conventional X-ray sources, could make them a useful tool in macromolecular crystallography.

The authors thank the ID29 and BM30A beamlines at ESRF for data collection and Gavin Connor Fox for critically reading the manuscript. This work was supported by grants from the Spanish Ministry of Science and Technology (BFU2008-01711/BMC, CONSOLIDER-INGENIO 2010 CSD2006-00015). RM holds a fellowship from the Spanish Ministry of Science and Technology.

References

Boggon, T. J. & Shapiro, L. (2000). *Structure*, **8**, R143–R149.
Brandl, M., Weiss, M. S., Jabs, A., Sühnel, J. & Hilgenfeld, R. (2001). *J. Mol. Biol.* **307**, 357–377.
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.
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.
Cowtan, K. D. & Main, P. (1996). *Acta Cryst.* D**52**, 43–48.
Dauter, Z., Dauter, M., de La Fortelle, É., Bricogne, G. & Sheldrick, G. M. (1999). *J. Mol. Biol.* **289**, 83–92.
Dauter, Z., Dauter, M. & Dodson, E. J. (2002). *Acta Cryst.* D**58**, 494–506.
Doublié, S. (1997). *Methods Enzymol.* **276**, 523–530.
Evans, G. & Bricogne, G. (2002). *Acta Cryst.* D**58**, 976–991.
Evans, P. R. (1993). *Proceedings of the CCP4 Study Weekend. Data Collection and Processing*, edited by L. Sawyer, N. Isaacs & S. Bailey, pp. 114–123. Warrington: Daresbury Laboratory.
Girard, É., Pebay-Peyroula, E., Vicat, J. & Kahn, R. (2004). *Acta Cryst.* D**60**, 1506–1508.
Girard, É, Stelter, M., Anelli, P. L., Vicat, J. & Kahn, R. (2003). *Acta Cryst.* D**59**, 118–126.
Girard, É., Stelter, M., Vicat, J. & Kahn, R. (2003). *Acta Cryst.* D**59**, 1914–1922.
Hendrickson, W. A., Horton, J. R. & LeMaster, D. M. (1990). *EMBO J.* **9**, 1665–1672.
Hendrickson, W. A. & Teeter, M. M. (1981). *Nature (London)*, **290**, 107–113.
Hermoso, J. A., Lagartera, L., González, A., Stelter, M., García, P., Martinez-Ripoll, M., García, J. L. & Menéndez, M. (2005). *Nature Struct. Mol. Biol.* **12**, 533–538.
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.
La Fortelle, É. de & Bricogne, G. (1997). *Methods Enzymol.* **276**, 472–494.
Leonard, G. A., Sainz, G., de Backer, M. M. E. & McSweeney, S. (2005). *Acta Cryst.* D**61**, 388–396.
Leslie, A. G. W. (1987). *Proceedings of the CCP4 Study Weekend. Computational Aspects of Protein Crystal Data Analysis*, edited by J. R. Helliwell, P. A. Machin & M. Z. Papiz, pp. 39–50. Warrington: Daresbury Laboratory.
Micossi, E., Hunter, W. N. & Leonard, G. A. (2002). *Acta Cryst.* D**58**, 21–28.
Molina, R., González, A., Moscoso, M., García, P., Stelter, M., Kahn, R. & Hermoso, J. A. (2007). *Acta Cryst.* F**63**, 742–745.
Molina, R., González, A., Stelter, M., Pérez-Dorado, I., Kahn, R., Morales, M., Campuzano, S., Campillo, N. E., Mobashery, S., García, J. L., García, P. & Hermoso, J. A. (2009). *EMBO Rep.* **10**, 246–251.
Port, M., Idée, J.-M., Medina, C., Robic, C., Sabatou, M. & Corot, C. (2008). *Biometals*, **21**, 469–490.
Quillin, M. L. & Matthews, B. W. (2002). *Acta Cryst.* D**58**, 97–103.
Sheldrick, G. M. (2008). *Acta Cryst.* A**64**, 112–122.
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.