Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

A dipicolinate lanthanide complex for solving protein structures using anomalous diffraction

Guillaume Pompidor,[a]*‡ Olivier Maury,[b] Jean Vicat[a] and Richard Kahn[a]*

[a]Institut de Biologie Structurale J.-P. Ébel, UMR 5075, 41 Rue Jules Horowitz, 38027 Grenoble CEDEX 1, France, and [b]Université de Lyon, École Normale Supérieure de Lyon, Laboratoire de Chimie, UMR 5182, 46 Allée d'Italie, 69364 Lyon CEDEX 07, France

‡ Present address: Swiss Light Source at the Paul Scherrer Institut, 5232 Villigen PSI, Switzerland.

Correspondence e-mail: guillaume.pompidor@psi.ch, richard.kahn@ibs.fr

Received 27 January 2010
Accepted 23 March 2010

PDB References: dipicolinate lanthanide complexes of thaumatin, 2pe7; urate oxidase, 2pes; lysozyme, 2pc2; xylanase, 3lgr.

Tris-dipicolinate lanthanide complexes were used to prepare derivative crystals of six proteins: hen egg-white lysozyme, turkey egg-white lysozyme, thaumatin from *Thaumatococcus daniellii*, urate oxidase from *Aspergillus flavus*, porcine pancreatic elastase and xylanase from *Trichoderma reesei*. Diffraction data were collected using either synchrotron radiation or X-rays from a laboratory source. In all cases, the complex turned out to be bound to the protein and the phases determined using the anomalous scattering of the lanthanide led to high-quality electron-density maps. The binding mode of the complex was characterized from the refined structures. The lanthanide tris-dipicolinate was found to bind through interactions between carboxylate groups of the dipicolinate ligands and hydrogen-bond donor groups of the protein. In each binding site, one enantiomeric form of the complex is selected from the racemic solution according to the specific site topology. For hen egg-white lysozyme and xylanase, derivative crystals obtained by cocrystallization belonged to a new monoclinic *C*2 crystal form that diffracted to high resolution.

1. Introduction

Lanthanide ions, which all exhibit strong white lines at their $L_{\rm III}$ absorption edge with anomalous contributions (f'' values) of almost 30 e^-, are excellent candidates for structure determination using anomalous diffraction. Rare-earth ions can substitute for Ca^{2+} in calcium-binding proteins (Reuben, 1975). This property was used in early studies that exploited the MAD method to solve the structures of parvalbumin (Kahn *et al.*, 1985) and mannose-binding protein (Weis *et al.*, 1991). Derivative crystals can be obtained by incorporating lanthanide salts in the cryoprotectant solution (Nagem *et al.*, 2001), but soaking in solutions containing lanthanide salts often damages the crystals. To avoid crystal degradation, lanthanides can be introduced into protein crystals using Gd complexes initially designed as MRI contrast agents (Girard *et al.*, 2003) or by a covalent linkage using thio-reactive functionalities (Purdy *et al.*, 2002). Recently, the structure of ubiquitin has been solved using the anomalous signal of Tb^{3+} bound to a double lanthanide-binding tag (Silvaggi *et al.*, 2007). Two rare-earth ions are particularly interesting for anomalous diffraction experiments: (i) Lu^{3+}, which is the lanthanide with the $L_{\rm III}$ absorption edge at the shortest wavelength (λ = 1.3412 Å) and allows the collection of data at the anomalous peak with the highest possible resolution, and (ii) Sm^{3+}, which has the highest f'' value (13.3 e^-) with Cu $K\alpha$ radiation and is the most effective lanthanide for in-house phasing. The latter property has been exploited to solve the structure of uridine-cytidine kinase 2 by the SIRAS method

Table 1
Crystallization conditions for HEWL, TEWL, thaumatin, urate oxidase, xylanase and PPE.

	Protein concentration (mg ml^{-1})	Precipitating agent	Buffer	Derivatization method
HEWL	30	0.3–0.8 *M* NaCl	MES pH 5.1	Cocrystallization, 100 m*M* $Na_3[Eu(DPA)_3]$
TEWL	30	1.2–2.0 *M* NaCl	Sodium acetate pH 4.6	Cocrystallization, 50–100 m*M* $Na_3[Eu(DPA)_3]$
Thaumatin	50	0.6–1.2 *M* Na/K tartrate	BES or bis-tris-propane pH 6.5	Cocrystallization and soaking, 50–100 m*M* $Na_3[Eu(DPA)_3]$
Urate oxidase	11	5% PEG 8000	Tris–HCl pH 8.5	Soaking, 100–150 m*M* $Na_3[Eu(DPA)_3]$
Xylanase	18	35% PEG 3350	HEPES pH 7.2	Cocrystallization, 100 m*M* $Na_3[Eu(DPA)_3]$
PPE	40	10–35% PEG 3350	Sodium acetate pH 5.2	Soaking, 120 m*M* $Na_3[Eu(DPA)_3]$

using X-rays from a rotating-anode generator (Appleby *et al.*, 2005).

Furthermore, among the lanthanide ions, Eu^{3+} and Tb^{3+} are well known to exhibit intense long-lived luminescence in the visible spectral range; these emission properties are widely used for bio-imaging applications (Pandia *et al.*, 2006). It is worth noting that aqueous solvated ions present weak emission owing to both the low absorption cross-section of the ions and the luminescence quenching induced by coordinated water molecules (OH vibrators). Importantly, this emission can be enhanced by several orders of magnitude upon coordination of the central metal ion with chromophore-containing ligands, which act both as as a shield to protect the metal from water and as antennae to efficiently sensitize the lanthanide luminescence.

Using such lanthanide complexes to prepare derivative crystals permits easy verification of the efficiency of the derivatization by checking the luminescence of the derivative crystals under UV radiation. This fast and simple method presents the significant advantages that the binding of anomalous scatterers can be assessed prior to X-ray data collection and that the derivatization protocol can thus be modified in a rational manner to ensure success.

Here, we report the use of multifunctional lanthanide tris-dipicolinate complexes, which combine luminescence and anomalous scattering properties, for the preparation of derivative crystals in order to determine experimental phases from anomalous diffraction experiments.

2. Materials and methods

2.1. $[Ln(DPA)_3]^{3-}$ synthesis

$Na_3[Ln(DPA)_3].6H_2O$ complexes, where DPA and Ln represent pyridine-2,6-dicarboxylate (dipicolinate) and lanthanide, respectively, were prepared using the procedure described by Tancrez *et al.* (2005). The water solubilites of the complexes are greater than 300 m*M* for all the lanthanide ions.

The structure of one enantiomeric form of the complex is shown in Fig. 1. The three DPA ligands complete the first coordination shell of the Ln^{3+} ion with nine coordinating atoms; each dipicolinate ion provides one N and two O atoms.

The study was carried out using either the Lu^{3+} or the Eu^{3+} complex. The Lu^{3+} ion was selected for data collection using synchrotron radiation because of the wavelength of its L_{III} absorption edge. Eu^{3+} was chosen because it exhibits luminescence properties in the visible domain and a high value of the anomalous contribution at the Cu $K\alpha$ radiation wavelength ($f'' = 11$ e$^-$).

2.2. Crystallization

Hen egg-white lysozyme (HEWL) was purchased from Boehringer. Thaumatin from *Thaumatococcus daniellii* and turkey egg-white lysozyme (TEWL) were purchased from Sigma–Aldrich. Xylanase from *Trichoderma reesei* and porcine pancreatic elastase (PPE) were purchased from Hampton Research and Serva, respectively. All proteins were used without any further purification. Urate oxidase from *Aspergillus flavus* was a generous gift from Sanofi–Aventis.

Derivative crystals were prepared by the vapour-diffusion technique using the hanging-drop method. $Na_3[Ln(DPA)_3]$ was dissolved in water to obtain solutions of the complexes at the required concentrations. Crystallization conditions are summarized in Table 1. Derivative crystals were obtained by cocrystallization for HEWL, TEWL and xylanase, by the soaking method for urate oxidase and PPE and by both methods for thaumatin. HEWL crystallization conditions were adapted from standard conditions that give rise to tetragonal crystals (Ducruix & Giegé, 1999; Vaney *et al.*, 1996). Thaumatin crystallization conditions were modified from those of Ko *et al.* (1994). Native crystals of urate oxidase were obtained according to the conditions described by Vivarès & Bonneté (2002) and Retailleau *et al.* (2004). TEWL crystals were prepared using similar conditions to those used by Howell (1995) and Margiolaki *et al.* (2005). PPE was crystallized using

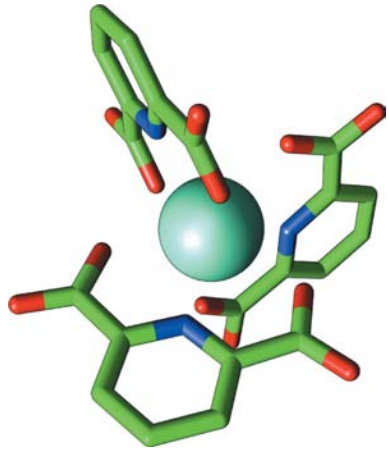

Figure 1
Tris-dipicolinate lanthanide complex.

Table 2
Summary of data-collection parameters and processing statistics for HEWL, TEWL, thaumatin, urate oxidase, xylanase and PPE.

Values in parentheses are for the highest resolution shell.

Data set (wavelength)	HEWL (λ = 1.542 Å)	TEWL (λ = 1.542 Å)	Thaumatin, Eu peak (λ = 1.776 Å)	Thaumatin, remote (λ = 0.920 Å)	Urate oxidase, Lu peak (λ = 1.341 Å)	Urate oxidase, remote (λ = 0.980 Å)	Xylanase (λ = 1.542 Å)	PPE (λ = 1.542 Å)
Space group	*C*2	$P6_122$	$P4_12_12$	$P4_12_12$	*I*222	*I*222	*C*2	$P2_12_12_1$
Unit-cell parameters								
a (Å)	50.40	70.21	57.25	57.55	79.21	79.55	110.89	50.41
b (Å)	33.88	70.21	57.25	57.55	94.57	95.32	38.37	58.15
c (Å)	69.86	85.09	148.70	149.40	103.65	104.33	47.59	74.89
β (°)	108.18						99.83	
Resolution range (Å)	17.05–1.54 (1.63–1.54)	18.55–1.73 (1.83–1.73)	45.36–2.80 (2.95–2.80)	45.60–1.46 (1.53–1.46)	47.30–2.07 (2.18–2.07)	47.67–1.60 (1.69–1.60)	17.05–1.63 (1.72–1.63)	19.68–1.63 (1.72–1.63)
No. of observed reflections	51088	272663	81920	543542	148449	368272	106992	216171
No. of unique reflections	16375	32881	6582	44672	23324	51801	23239	32517
$I/\sigma(I)$	15.3 (12.1)	36.6 (7.7)	17.9 (11.3)	30.4 (6.8)	31.1 (10.7)	21.5 (5.4)	20.2 (8.8)	39.1 (18.7)
Multiplicity	3.1 (2.5)	19.8 (17.9)	12.4 (8.0)	12.2 (7.6)	6.4 (3.1)	7.1 (6.7)	4.6 (4.3)	6.6 (6.0)
Completeness (%)	98.1 (98.1)	99.6 (97.9)	99.9 (93.3)	99.1 (95.6)	97.1 (87.4)	98.7 (99.9)	94.0 (89.1)	98.9 (98.8)
$R_{\rm merge}$ (%)	4.6 (5.7)	4.3 (40.8)	4.4 (6.7)	5.6 (36.9)	4.1 (30.6)	5.2 (41.7)	4.5 (16.1)	3.0 (8.3)
$R_{\rm ano}$ (%)	10.6 (16.2)	4.6 (11.5)	4.2 (4.4)	2.3 (11.5)	7.1 (22.4)	3.1 (14.8)	5.6 (16.3)	2.4 (4.9)

the conditions given by Würtele *et al.* (2000) and Mueller-Dieckmann *et al.* (2004). Xylanase was cocrystallized using similar conditions to those reported by McPherson & Cudney (2006).

Owing to the global charge of the complex, addition to the mother liquor of $Na_3[Ln(DPA)_3]$ at concentrations in the range 50–150 m*M* strongly increases the ionic strength of the solution used in cocrystallization and soaking essays. Furthermore, the $[Ln(DPA)_3]^{3-}$ complex tends to precipitate in the presence of divalent cations such as Ca^{2+} and Mg^{2+}, which substitute Na^+ as the counterion. The use of lanthanide tris-dipicolinate complexes may thus be incompatible with crystallization conditions that use salts at high concentrations or contain divalent cations.

2.3. Data collection

Diffraction data were collected using synchrotron radiation for thaumatin and urate oxidase and X-rays from a laboratory source for HEWL, TEWL, xylanase and PPE. In order to obtain highly redundant data, a wedge of 180° was collected for each data set with an oscillation range of 0.5° per image. Data-collection parameters and processing statistics are summarized in Table 2.

2.3.1. HEWL. Diffraction data were collected from a single crystal at the Cu *Kα* wavelength using a Nonius FR 591 rotating-anode X-ray generator equipped with a multilayer optical system from Osmic. Data were recorded on a MAR 345 image-plate detector. Cryoprotection was performed with a solution corresponding to the mother liquor supplemented with 25% glycerol. The crystal was soaked for 20 s in this solution and then cryocooled using N_2 at 100 K (Oxford Cryosystems).

2.3.2. TEWL. Data were collected from a single crystal using the same procedure as for HEWL except that Paratone-N was used for cryoprotection.

2.3.3. Thaumatin. Cryoprotection was performed with a solution corresponding to the mother liquor supplemented with 25% ethylene glycol. Diffraction data were collected using synchrotron radiation on ESRF beamline BM30A. Two data sets were collected at different wavelengths from the same derivative crystal in order to determine phases by the MAD method. One data set was collected at the Eu $L_{\rm III}$ absorption edge (λ = 1.776 Å) to maximize the anomalous signal and the other was collected at a shorter wavelength (λ = 0.920 Å) to extend the resolution to 1.46 Å.

2.3.4. Urate oxidase. Three data sets were collected on BM30A at two different wavelengths: two at the Lu $L_{\rm III}$ absorption edge and one at a shorter wavelength. Cryoprotection was performed using a solution corresponding to the mother liquor supplemented with 25% PEG 400.

2.3.5. Xylanase and PPE. Data were collected from a single crystal at the Cu *Kα* wavelength using a Rotaflex RU-200 rotating-anode X-ray generator from Rigaku equipped with multi-layer optics from Xenocs. Cryoprotection was performed with Paratone-N.

For all the data sets, the intensities were integrated with the program *XDS* (Kabsch, 2010). The integrated intensities were scaled with the program *SCALA* from the *CCP*4 suite (Collaborative Computational Project, Number 4, 1994).

3. Results

The derivatized crystals of HEWL and xylanase obtained by cocrystallization belonged to the monoclinic space group *C*2, which has never previously been observed for these proteins.

3.1. Phasing and automatic model building

For all derivative crystals, anomalous difference Patterson maps were calculated with the program *FFT* (Collaborative Computational Project, Number 4, 1994). For xylanase and thaumatin, one binding site of the complex located on the twofold axis was found manually on the basis on anomalous difference Patterson maps. In the case of the HEWL derivative crystal five Eu^{3+} sites were found with the program *Shake-*

Table 3
Summary of phase determination, solvent flattening and automatic model building.

$\Delta\varphi$ and MCC correspond to the mean absolute phase difference and map correlation coefficient, respectively.

	Solvent content (%)	Phasing power, centric	Phasing power, acentric	*SHARP*/*SOLOMON*	*DM*	Automatic building with *ARP*/*wARP*
HEWL	28		Ano = 3.56	$\Delta\varphi$ = 44.00°, MCC = 0.76	$\Delta\varphi$ = 39.46°, MCC = 0.76	113/129 (88%)
TEWL	30		Ano = 3.19			117/129 (91%)
PPE	30		Ano = 1.184			207/240 (86%)
Xylanase	40		Ano = 1.08	$\Delta\varphi$ = 40.08°, MCC = 0.83	$\Delta\varphi$ = 40.35°, MCC = 0.83	180/190 (95%)
Thaumatin (peak)	55	Iso = 0.639	Ano = 1.64, Iso = 0.62	$\Delta\varphi$ = 32.77°, MCC = 0.89	$\Delta\varphi$ = 36.74°, MCC = 0.88	198/207 (96%)
Thaumatin (remote)			Ano = 0.40			
Urate oxidase (peak)	50	Iso = 0.367	Ano = 2.85, Iso = 0.42	$\Delta\varphi$ = 29.93° MCC = 0.91	$\Delta\phi$ = 28.60°, MCC = 0.91	282/295 (96%)
Urate oxidase (remote)			Ano = 1.41			

and-Bake (Blessing & Smith, 1999), one of which was located on the twofold crystallographic axis. For the urate oxidase, PPE and TEWL derivative crystals, binding sites were determined using the program *SHELXD* (Sheldrick, 2008).

In all experiments, phases were determined with the program *SHARP* (de La Fortelle & Bricogne, 1997) using either the SAD or MAD method, depending on the data set. Phases were improved using the programs *SOLOMON* (Abrahams & Leslie, 1996) and *DM* (Cowtan & Main, 1996) using the solvent content given by the program *TRUNCATE*. Phasing statistics are given in Table 3. Using the phases obtained after solvent flattening, automatic chain tracing and model building were attempted with the program *ARP*/*wARP* (Morris & Perrakis, 2003), which was able to build about 90% of the model for each of the six structures. As shown by the phasing statistics and the high proportion of the model docked in the experimental electron-density maps, the use of $[Ln(DPA)_3]^{3-}$ led to experimental phases of excellent quality with both the SAD and MAD methods.

3.2. Refinement

The six PDB files and the free reflection sets given by *ARP*/*wARP* were used as a starting point for refinement. Refinement of the HEWL and xylanase structures was performed without introducing any external structural information. As the thaumatin and urate oxidase structures have already been solved in the same space groups and deposited in the PDB, the native structures with PDB codes 1lxz (Charron *et al.*, 2002) and 1xt4 (Retailleau *et al.*, 2005) were placed in the same orientation as the automatically built model and then used for refinement. The TEWL and PPE structures were not further refined because the complex was not sufficiently bound to identify the ligands in the electron density.

The four structures were refined using *REFMAC* (Murshudov *et al.*, 1999). A description library for the dipicolinate ligand created using the programs *PRODRG* (Schüttelkopf & van Aalten, 2004) and *Monomer Library Sketcher* from the *CCP*4 suite was used to model each DPA ligand of both enantiomeric forms of the lanthanide complex. No other restraints were added. Complex occupancies were estimated by careful examination of both the difference electron-density maps and the *B* factors of the ligands.

Table 4
Summary of refinement statistics.

	HEWL	Thaumatin	Urate oxidase	Xylanase
PDB code	2pc2	2pe7	2pes	3lgr
R_{work}	0.138	0.152	0.179	0.136
R_{free}	0.167	0.176	0.195	0.165
Standard deviations				
Bond lengths (Å)	0.009	0.012	0.011	0.010
Bond angles (°)	1.520	1.416	1.358	1.370
Ramachandran plot: residues in (%)				
Favoured region	98.4	99.0	98.3	97.9
Allowed region	1.6	1.0	1.7	2.1
Disallowed region	0.0	0.0	0.0	0.0
Average temperature (*B*) factors (Å^2)				
Main-chain atoms	10.0	10.1	19.4	9.3
Side-chain atoms	11.9	11.6	22.5	10.5
Water molecules	25.9	26.5	32.6	25.7
Lanthanide atoms	14.1	11.5	30.9	12.2
Dipicolinate	11.1	14.5	28.1	8.6

In each case, anisotropic *B* factors were introduced and refined for the lanthanide atoms. Anisotropic *B* factors were refined for all atoms of the thaumatin structure. Refinement statistics are summarized in Table 4.

3.3. $[Ln(DPA)_3]^{3-}$ binding mode

The binding mode of the complex was elucidated in each of the four refined structures. Since the first coordination sphere of the lanthanide is complete, $[Ln(DPA)_3]^{3-}$ is bound to the protein through the three DPA ions; more precisely, by hydrogen bonds between carboxylate groups of the DPA and hydrogen-donor groups of the protein.

3.3.1. HEWL. The negatively charged $[Eu(DPA)_3]^{3-}$ complex is localized in positively charged regions. In each binding site, Arg residues are involved in the binding of the complex.

Five $[Eu(DPA)_3]^{3-}$ sites were found per HEWL molecule, two of which were highly occupied. In the four most occupied sites the binding of the tris-dipicolinate lanthanide complex turned out to be enantioselective. The weak electron-density around the less occupied site did not allow us to locate the DPA ligands.

The lanthanide complexes bind to the protein through hydrogen bonds between the O atoms of the carboxylate groups of the dipicolinate ligands and a hydrogen-bond donor of the protein.

The first site, which is almost fully occupied (occupancy of 0.88), is depicted in Fig. 2. The $[Eu(DPA)_3]^{3-}$ complex is bound to three lysozyme molecules. One DPA ligand is hydrogen-bonded to the N^{α} atoms of Arg5, Cys6 and Glu7 of the HEWL molecule at (x, y, z), to the O^{η} atom of Tyr20 of the HEWL molecule at $(x + 1/2, y - 1/2, z)$ and to three water molecules. A second DPA ligand is bound to the protein *via* hydrogen bonds between an O atom from one carboxylate group and the N^{α} atom of Arg128 of the protein molecule at (x, y, z) and between the other carboxylate group and the $N^{\eta1}$ and $N^{\eta2}$ atoms of Arg112 from the lysozyme molecule at $(x, y - 1, z)$. Two water molecules complete the hydrogen-bond network. The two carboxylate groups of the third DPA are hydrogen-bonded to the $N^{\eta1}$ atom of Arg128 of the molecule at (x, y, z) and to the N^{ζ} atom of Lys97 of the molecule at $(x + 1/2, y - 1/2, z)$.

The second binding site is located on a twofold crystallographic axis, onto which a twofold axis of $[Eu(DPA)_3]^{3-}$ complex is superimposed (Fig. 3). The complex is bound to two symmetry-related protein molecules at (x, y, z) and $(1 - x, y, -z)$. The site occupancy is also 0.88. One O atom of the DPA crossed by the axis is bound to the N^{α} atom of Gly126 of the molecule at (x, y, z) and to the $N^{\eta2}$ atom of Arg125 of the molecule at $(1 - x, y, -z)$. The symmetry-related O atom is bound to the corresponding symmetry-related atoms of both protein molecules. The two other DPA ligands are hydrogen-bonded to the N^{ε} and $N^{\eta2}$ atoms of Arg125 of the molecules at (x, y, z) and $(1 - x, y, -z)$.

The third $[Eu(DPA)_3]^{3-}$-binding site, which has a lower occupancy (0.62), is surrounded by three HEWL molecules (Fig. 4). One DPA ligand is bound to the protein by hydrogen bonds between the O atoms of one carboxylate group and the N^{ε} atom of Trp62, the N^{ε} and the $N^{\eta2}$ atoms of Arg61 and the $N^{\eta2}$ atom of Arg73 of the molecule at (x, y, z), while the other carboxylate group is hydrogen-bonded to the N^{α} and O^{γ} atoms of Ser86 of the protein molecule at $(x, y + 1, z)$ and a water molecule. A second DPA ligand is connected to the protein *via* hydrogen bonds to the N^{ε} atom of Arg68 of the molecule at $(1/2 - x, y + 1/2, 1 - z)$, the O^{γ} atom of Ser86 at $(x, y + 1, z)$ and two water molecules. The third DPA is hydrogen-bonded to water molecules and makes a π-stacking interaction with Trp62 of the molecule at (x, y, z).

The fourth site, with an occupancy of 0.39, is located between three HEWL molecules at (x, y, z), $(x - 1/2, y - 1/2, z)$ and $(x, y - 1, z)$ (Fig. 5). One DPA is bound to the N^{α} atoms of Gly102, Asn103 and Gly104 of the molecule at $(x, y - 1, z)$ and to the $N^{\eta2}$ atom of Arg21 of the same molecule as well as to two water molecules. A second ligand is hydrogen-bonded to the N^{ε} and $N^{\eta2}$ atoms of Arg14 of the

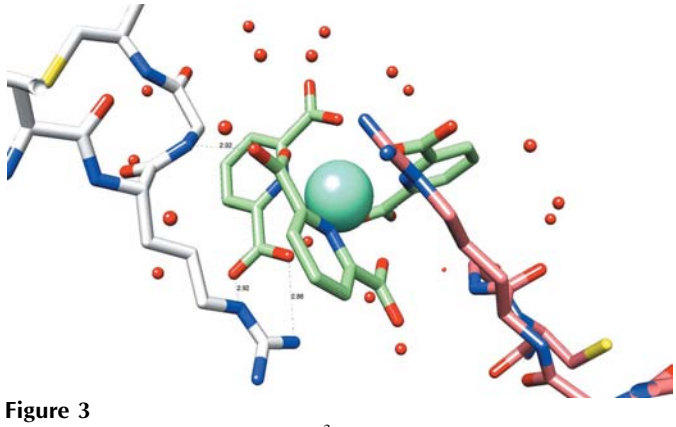

Figure 3
Binding site No. 2 of $[Eu(DPA)_3]^{3-}$ located on the twofold axis in the HEWL derivative crystal. DPA ligands are depicted in green. HEWL molecules located at (x, y, z) and $(1 - x, y, -z)$ are depicted in white and salmon, respectively.

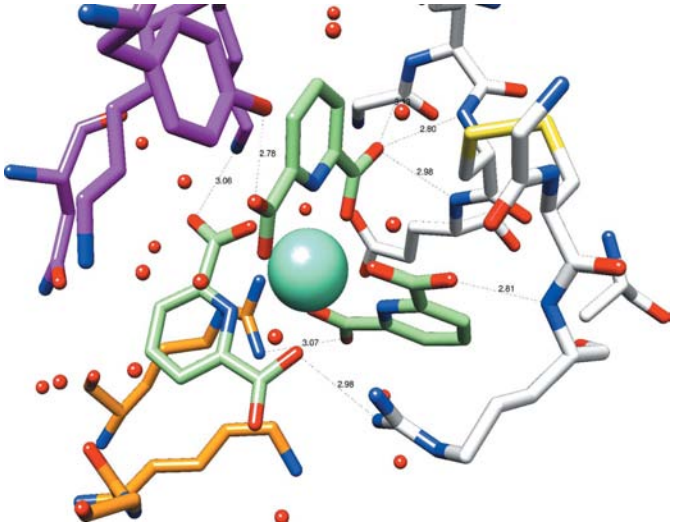

Figure 2
Binding site No. 1 of $[Eu(DPA)_3]^{3-}$ in the HEWL derivative crystal. DPA ligands are depicted in green. HEWL molecules located at (x, y, z), $(x + 1/2, y - 1/2, z)$ and $(x, y - 1, z)$ are depicted in white, purple and orange, respectively.

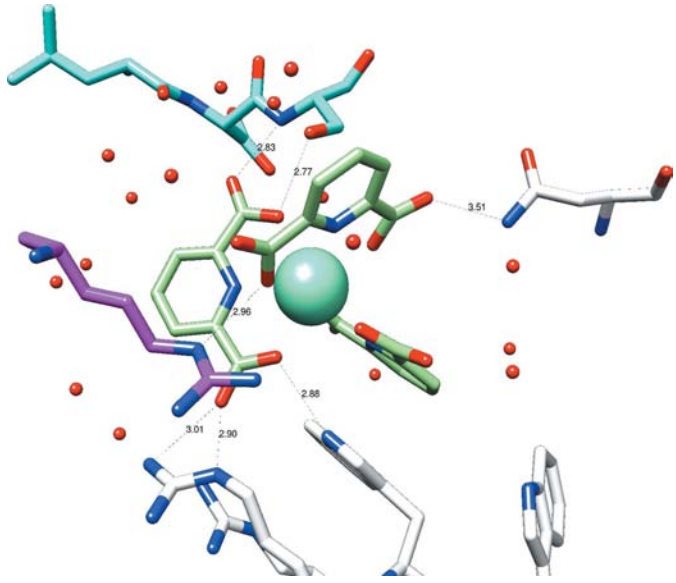

Figure 4
Binding site No. 3 of $[Eu(DPA)_3]^{3-}$ in the HEWL derivative crystal. DPA ligands are depicted in green. HEWL molecules located at (x, y, z), $(x, y + 1, z)$ and $(1/2 - x, y + 1/2, 1 - z)$ are depicted in white, cyan and purple, respectively.

lysozyme molecule at (x, y, z), to the $N^{\eta 1}$ atom of Arg125 of the molecule at ($x - 1/2$, $y - 1/2$, z), to the $N^{\eta 2}$ atom of Arg21 of the molecule at (x, $y - 1$, z) and to two water molecules. The third DPA ligand is bound to water molecules.

3.3.2. Thaumatin. The $[Eu(DPA)_3]^{3-}$ complex, which is located on the twofold axis, is bound to two symmetry-related protein molecules. The complex is bound to each of them by two arginine residues. The site occupancy was refined to a value of 0.26, which corresponds to an effective occupancy of 0.52 when taking the special position of the complex into account. Double conformations were introduced to describe Arg79 and Arg82 involved in the binding of $[Eu(DPA)_3]^{3-}$: one for the position when $[Eu(DPA)_3]^{3-}$ is bound and one for the position when $[Eu(DPA)_3]^{3-}$ is not bound.

3.3.3. Urate oxidase. Two of the three sites were sufficiently occupied to identify the DPA ligands in the electron-density maps. In contrast to HEWL and thaumatin, each $[Lu(DPA)_3]^{3-}$ complex is bound to one protein molecule only. The two most highly occupied sites are located in the largest solvent channel between the protein molecules at (x, y, z) and ($-x + 1/2$, $y + 1/2$, $-z + 1/2$). Hydrophobic π-stacking interactions between two DPA rings permit the proximity of two complexes in the structure, with a distance between lanthanide ions of 9.2 Å.

In the most occupied site, two O atoms from two of the DPA ligands are hydrogen-bonded to the N^{ζ} atom of Lys203 and the N^{ε} atom of Gln242. The second complex is bound to the protein through interactions with water molecules.

3.3.4. Xylanase. The $[Eu(DPA)_3]^{3-}$-binding site located on the crystallographic twofold axis was fully occupied (Fig. 6). The DPA ligand bisected by the twofold axis is bound to the O^{η} group of Tyr17 of two protein molecules. The other DPA ligands are hydrogen-bonded to the $N^{\delta 2}$ atom of Asn19 and to the N^{α} atoms of Gly23 and Gly24 of the two molecules. The occupancies of the two other sites were too low to locate DPA ligands in the electron-density maps.

4. Discussion

The four refined structures allowed the binding modes of the complex to be established in principle. The binding of the complex is enantioselective. The main interaction between the protein and the $[Ln(DPA)_3]^{3-}$ complex occurs through hydrogen bonds between the O atoms of the carboxylate groups of the DPA ligands and hydrogen-bond donor residues, more specifically positively charged residues. Hydrophobic π-stacking interactions were also observed between DPA rings and aromatic residues.

At the molecular level two different types of interactions can be distinguished. On one hand the $[Ln(DPA)_3]^{3-}$ ion can act as a classical heavy-atom compound that can be used to prepare derivative crystals, while on the other it can greatly affect the crystallization process as a cross-linking agent.

For urate oxidase, PPE, TEWL and thaumatin the binding rate was low and the $[Ln(DPA)_3]^{3-}$ complexes were bound to the protein in accessible regions of the solvent channel. Nevertheless, owing to the large anomalous contributions of the lanthanide ions either using synchrotron radiation or Cu $K\alpha$ radiation, the weak binding observed in the structures of urate oxidase, PPE, TEWL and thaumatin was sufficient to determine experimental phases of high quality, as assessed by the high proportion of the whole model which was docked into the initial experimental electron-density maps. In all of these cases the $[Ln(DPA)_3]^{3-}$ complexes were not involved in the crystal packing. Moreover, the thaumatin derivative crystals prepared by cocrystallization and by soaking were isomorphous and the binding sites were identical.

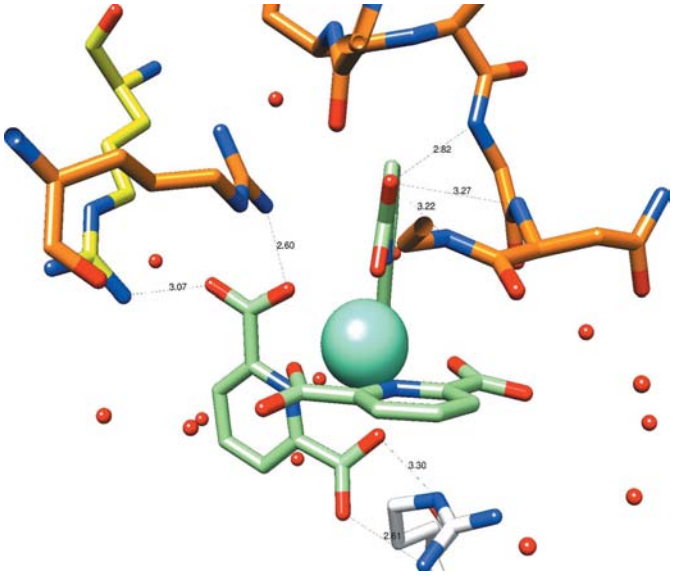

Figure 5
Binding site No. 4 of $[Eu(DPA)_3]^{3-}$ in the HEWL derivative crystal. DPA ligands are depicted in green. HEWL molecules located at (x, y, z), ($x - 1/2$, $y - 1/2$, z) and (x, $y - 1$, z) are depicted in white, orange and yellow, respectively.

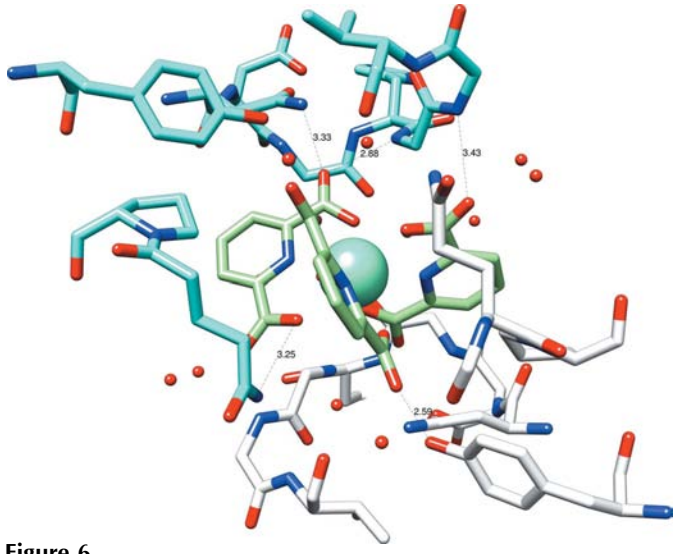

Figure 6
Binding site of $[Eu(DPA)_3]^{3-}$ located on the twofold axis in the xylanase derivative crystal. DPA ligands are depicted in green. The two xylanase molecules related by the twofold-symmetry axis are depicted in white and cyan.

In the xylanase and HEWL derivative structures, the $[Ln(DPA)_3]^{3-}$ complexes participate in the crystal packing by modifying protein–protein contacts in the lattice, explaining the change of space group that was observed for the derivative crystals. Native crystals of xylanase are smaller than derivative crystals obtained under similar conditions but in the presence of the complex. The native crystals belonged to space group $P222_1$ and diffracted to a resolution of 1.80 Å using synchrotron radiation, while data were collected to 1.63 Å resolution from derivative crystals using Cu $K\alpha$ radiation from a laboratory source. In the case of HEWL the monoclinic derivative crystals exhibited higher diffraction quality than the native crystals, which belonged to space group $P4_32_12$. A high-resolution data set was collected at 0.81 Å resolution using synchrotron radiation at a short wavelength distant from the lanthanide L_{III} absorption edge.

As suggested by McPherson & Cudney (2006), the incorporation of small molecules in crystallization drops can promote nucleation and crystal growth by establishing numerous weak interactions between protein molecules. The presence of such cross-linking agents in the crystallization medium can be advantageous for obtaining crystals and for improving diffraction quality. By establishing hydrogen-bond networks between protein molecules in the crystal, $[Ln(DPA)_3]^{3-}$ complexes can promote lattice formation as well as providing derivative crystals with high phasing power. In the case of HEWL the hydrogen-bond network is strengthened by electrostatic forces between the negatively charged $[Ln(DPA)_3]^{3-}$ complexes and the positively charged arginine residues. The high affinity of the complex for arginine residues leads to a supramolecular organization on the crystal scale (Pompidor *et al.*, 2008). Through its binding, the complex improves the protein–protein interactions involved in crystal packing. This is illustrated by the values of the *B* factors for arginine residues located on the surface of the HEWL molecule and that are involved in the binding of $[Ln(DPA)_3]^{3-}$, which are lower than in the native structure.

In addition to these strong supramolecular interactions, the unique luminescence properties of Eu^{3+} and Tb^{3+} complexes offer an easy method to verify complex binding in derivative crystals obtained by cocrystallization (D'Aléo *et al.*, 2007). As the lanthanide complexes are 'trapped' between protein molecules in the derivative crystals of HEWL and xylanase, simple backsoaking removes the free complexes in the bulk solvent, removing the luminescence from unbound complex in the mother liquor. After backsoaking, simple irradiation with a classical UV lamp permits the detection of luminescence from Eu^{3+} and Tb^{3+} complexes bound to the protein. This simple procedure is much faster than collecting and processing diffraction data in order to verify the binding of anomalous scatterers in the prepared crystals.

5. Conclusion

We have demonstrated the potential of $[Ln(DPA)_3]^{3-}$ complexes in the preparation of derivative crystals for *de novo* structure determination by anomalous diffraction methods using either synchrotron or Cu $K\alpha$ radiation. Derivative crystals can be obtained by either soaking or cocrystallization procedures. Owing to the large anomalous contribution of the lanthanide ions, the $[Ln(DPA)_3]^{3-}$ complexes facilitate the preparation of derivative crystals with high phasing power even in cases where the binding rate is not high, as in the thaumatin, PPE, TEWL and urate oxidase examples.

For some proteins, the presence of the complex can promote crystallization or improve the quality of the resulting crystals. In these cases, use of the complex offers the opportunity to prepare crystals of better quality with very high phasing power.

The main advantages of $[Ln(DPA)_3]^{3-}$ complexes, (i) high anomalous contribution, (ii) cross-linking properties and (iii) luminescence emission in the visible domain for Eu^{3+} and Tb^{3+} complexes, can be used to obtain derivative crystals of high quality with a fast and simple method to check the binding of the complex prior to data collection and processing.

These lanthanide complexes could also be used as an additive in crystallization conditions. It would be beneficial to include the $[Ln(DPA)_3]^{3-}$ complex in the screening of crystallization conditions which are compatible with the use of the complex. Taken together with the fact that complex binding in the resulting crystals can be easily checked by luminescence, the advantages of using $[Ln(DPA)_3]^{3-}$ complexes as a tool for high-throughput structure determination are clear.

References

Abrahams, J. P. & Leslie, A. G. W. (1996). *Acta Cryst.* D**52**, 30–42.
Appleby, T. C., Larson, G., Cheney, I. W., Walker, H., Wu, J. Z., Zhong, W., Hong, Z. & Yao, N. (2005). *Acta Cryst.* D**61**, 278–284.
Blessing, R. H. & Smith, G. D. (1999). *J. Appl. Cryst.* **32**, 664–670.
Charron, C., Kadri, A., Robert, M.-C., Giegé, R. & Lorber, B. (2002). *Acta Cryst.* D**58**, 2060–2065.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Cowtan, K. D. & Main, P. (1996). *Acta Cryst.* D**52**, 43–48.
D'Aléo, A., Pompidor, G., Elena, B., Vicat, J., Baldeck, P. L., Toupet, L., Kahn, R., Andraud, C. & Maury, O. (2007). *ChemPhysChem*, **8**, 2125–2132.
Ducruix, A. & Giegé, R. (1999). *Crystallization of Nucleic Acids and Proteins. A Practical Approach*, p. 144. Oxford University Press.
Girard, É., Stelter, M., Anelli, P. L., Vicat, J. & Kahn, R. (2003). *Acta Cryst.* D**59**, 118–126.
Howell, P. L. (1995). *Acta Cryst.* D**51**, 654–662.
Kabsch, W. (2010). *Acta Cryst.* D**66**, 125–132.
Kahn, R., Fourme, R., Bosshard, R., Chiadmi, M., Risler, J. L., Dideberg, O. & Wery, J. P. (1985). *FEBS Lett.* **179**, 133–137.
Ko, T.-P., Day, J., Greenwood, A. & McPherson, A. (1994). *Acta Cryst.* D**50**, 813–825.
La Fortelle, E. de & Bricogne, G. (1997). *Methods Enzymol.* **276**, 472–494.
Margiolaki, I., Wright, J. P., Fitch, A. N., Fox, G. C. & Von Dreele, R. B. (2005). *Acta Cryst.* D**61**, 423–432.
McPherson, A. & Cudney, B. (2006). *J. Struct. Biol.* **156**, 387–406.
Morris, R. J., Perrakis, A. & Lamzin, V. S. (2003). *Methods Enzymol.* **374**, 229–244.
Mueller-Dieckmann, C., Polentarutti, M., Djinovic Carugo, K., Panjikar, S., Tucker, P. A. & Weiss, M. S. (2004). *Acta Cryst.* D**60**, 28–38.

Murshudov, G. N., Vagin, A. A., Lebedev, A., Wilson, K. S. & Dodson, E. J. (1999). *Acta Cryst.* D**55**, 247–255.
Nagem, R. A. P., Dauter, Z. & Polikarpov, I. (2001). *Acta Cryst.* D**57**, 996–1002.
Pandia, S., Yu, J. & Parker, D. (2006). *Dalton Trans.* **23**, 2757–2766.
Pompidor, G., D'Aléo, A., Vicat, J., Toupet, L., Giraud, N., Kahn, R. & Maury, O. (2008). *Angew. Chem. Int. Ed.* **48**, 3388–3391.
Purdy, M. D., Ge, P., Chen, J., Selvin, P. R. & Wiener, M. C. (2002). *Acta Cryst.* D**58**, 1111–1117.
Retailleau, P., Colloc'h, N., Vivarès, D., Bonneté, F., Castro, B., El Hajji, M., Mornon, J.-P., Monard, G. & Prangé, T. (2004). *Acta Cryst.* D**60**, 453–462.
Retailleau, P., Colloc'h, N., Vivarès, D., Bonneté, F., Castro, B., El Hajji, M. & Prangé, T. (2005). *Acta Cryst.* D**61**, 218–229.
Reuben, J. (1975). *Naturwissenschaften*, **62**, 172–178.
Schüttelkopf, A. W. & van Aalten, D. M. F. (2004). *Acta Cryst.* D**60**, 1355–1363.
Sheldrick, G. M. (2008). *Acta Cryst.* A**64**, 112–122.
Silvaggi, N. R., Martin, L. J., Schwalbe, H., Imperiali, B. & Allen, K. N. (2007). *J. Am. Chem. Soc.* **129**, 7114–7120.
Tancrez, N., Feuvrie, C., Ledoux, I., Zyss, J., Toupet, L., Le Bozec, H. & Maury, O. (2005). *J. Am. Chem. Soc.* **127**, 13474–13475.
Vaney, M. C., Maignan, S., Riès-Kautt, M. & Ducruix, A. (1996). *Acta Cryst.* D**52**, 505–517.
Vivarès, D. & Bonneté, F. (2002). *Acta Cryst.* D**58**, 472–479.
Weis, W. I., Kahn, R., Fourme, R., Drickamer, K. & Hendrickson, W. A. (1991). *Science*, **254**, 1608–1611.
Würtele, M., Hahn, M., Hilpert, K. & Höhne, W. (2000). *Acta Cryst.* D**56**, 520–523.