Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

High-phasing-power lanthanide derivatives: taking advantage of ytterbium and lutetium for optimized anomalous diffraction experiments using synchrotron radiation

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

[a]Laboratoire de Cristallographie Macromoléculaire, Institut de Biologie Structurale J. P. Ébel CEA CNRS UJF, 41 Rue Jules Horowitz, 38027 Grenoble CEDEX 01, France, and [b]Bracco Imaging spa, Via E. Folli 50, 20134 Milan, Italy

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

Correspondence e-mail: kahn@ibs.fr

Received 14 March 2003
Accepted 30 July 2003

Ytterbium and lutetium are well suited for optimized anomalous diffraction experiments using synchrotron radiation. Therefore, two lanthanide complexes Yb-HPDO3A and Lu-HPDO3A have been produced that are similar to the Gd-HPDO3A complex already known to give good derivative crystals. Derivative crystals of hen egg-white lysozyme were obtained by co-crystallization using 100 m*M* solutions of each lanthanide complex. *De novo* phasing has been carried out using single-wavelength anomalous diffraction on data sets collected on each derivative crystal at the L_{III} absorption edge of the corresponding lanthanide ($f'' = 28\ e^-$). A third data set was collected on a Lu-HPDO3A derivative crystal at the Se *K* absorption edge with $f''_{Lu} = 10\ e^-$. The structures were refined and compared with the known structure of the Gd-HPDO3A lysozyme derivative. The quality of the experimental electron-density maps allows easy model building. With L_{III} absorption edges at shorter wavelengths than the gadolinium absorption edge, lutetium and ytterbium, when chelated by a ligand such as HPDO3A, form lanthanide complexes that are especially interesting for synchrotron-radiation experiments in structural biology.

1. Introduction

The use of anomalous signals has improved protein-structure solution in single/multiple isomorphous replacement (SIRAS/MIRAS method; North, 1965; Matthews, 1966). The multiple-wavelength anomalous diffraction (MAD) method (Hendrickson *et al.*, 1990; Hendrickson & Ogata, 1997) using genetic engineering to replace methionines by selenomethionines (Doublié, 1997) has revolutionized macromolecular crystallography. More recently, the development of the single-wavelength anomalous diffraction (SAD) method (for a review, see Dauter *et al.*, 2002) has reduced the time required for data collection.

In a previous paper (Girard *et al.*, 2003), seven gadolinium complexes were discussed for obtaining derivative crystals that make use of the very high anomalous signal of gadolinium. This paper presents the advantage of using ytterbium or lutetium in similar complexes for synchrotron-radiation experiments. For this purpose, we have substituted gadolinium in Gd-HPDO3A, the gadolinium complex of 10-(2-hydroxypropyl)-1,4,7,10-tetraazacyclododecane-1,4,7-triacetic acid, with either ytterbium or lutetium. SAD experiments were performed on Yb-HPDO3A and Lu-HPDO3A derivatives of hen egg-white lysozyme (HEWL), allowing a direct comparison with results obtained using Gd-HPDO3A (Girard *et al.*, 2002).

2. Methods

2.1. Crystallization

HEWL was purchased from Boehringer. Yb-HPDO3A and Lu-HPDO3A were kindly provided by Bracco Imaging spa, Milan, Italy. Tetragonal crystals were obtained by co-crystallization using 100 m*M* of the lanthanide complex according to the previously described protocol (Girard *et al.*, 2002).

2.2. Data collection and data processing

Two optimized SAD data sets were collected using synchrotron radiation from beamline BM30A at the ESRF with Lu-HPDO3A and Yb-HPDO3A HEWL-derivative crystals. The wavelength was chosen at the L_{III} absorption edge of the lanthanide and was set to the maximum of f'' ($\sim 28\ e^-$).

At wavelengths around 1 Å, far away from the Lu L_{III} absorption edge, the anomalous signal from Lu remains large ($f'' \simeq 10\ e^-$). To evaluate the potential use of this anomalous signal to solve the structure of a small- or medium-size protein, a SAD data set was collected on a Lu-HPDO3A HEWL-derivative crystal at the Se *K* absorption edge, a standard method for collecting native data using synchrotron radiation.

The samples were all N_2-cryocooled at 100 K. The Yb-HPDO3A data were collected using a MAR345 detector. Both Lu-HPDO3A data sets were collected using a MAR CCD

detector of 165 mm diameter. The resolution of all data sets collected at the L_{III} absorption edge was not limited by the intrinsic diffraction quality of the derivative crystal, but was limited by the experimental setup.

Measurements were integrated with the program *DENZO* (Otwinowski & Minor, 1997). Integrated intensities were scaled using *SCALA* from the *CCP*4 suite (Collaborative Computational Project, Number 4, 1994). A summary of data-collection parameters and processing statistics is given in Table 1.

2.3. *De novo* SAD phasing

To compare the results with those observed for the Gd-HPDO3A HEWL derivative (Girard *et al.*, 2002), *de novo* phasing was performed in a similar way on each derivative data set.

The lanthanide coordinates were determined with the program *SnB* (Weeks & Miller, 1999). As evidenced by the anomalous difference Patterson maps (Fig. 1), the site positions in both the ytterbium and lutetium derivatives were identical to the two Gd-site positions found in the Gd-HPDO3A derivative.

Phases were determined with the direct-methods-based program *OASIS* (Hao *et al.*, 2000) from the *CCP*4 suite (Collaborative Computational Project, Number 4, 1994). f'' was fixed at 28.0 e^- for data collected at the L_{III} absorption edges and at 10.4 e^- for the Lu-HPDO3A-derivative data measured at the Se *K* absorption edge (Henke *et al.*, 1993).

Phases were improved by solvent flattening and histogram matching with the program *DM* (Cowtan & Main, 1996). Results of the phasing are summarized in Table 2. Phasing statistics for the Gd-HPDO3A derivative of HEWL (Girard *et al.*, 2002) are also indicated.

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

Values in parentheses refer to the highest resolution shell. R_{sym} and R_{ano} were computed using the program *SCALA* (Collaborative Computational Project, Number 4, 1994).

	Yb-HPDO3A, λ = 1.3859 Å, Yb L_{III} edge	Lu-HPDO3A, λ = 1.3406 Å, Lu L_{III} edge	Lu-HPDO3A, λ = 0.9798 Å, Se *K* edge
Total angular range (°)	161	180	180
Unit-cell parameters (Å)	*a* = 77.045, *c* = 38.462	*a* = 77.192, *c* = 38.681	*a* = 77.281, *c* = 38.648
Resolution range (Å)	25–2.15 (2.27–2.15)	25–2.37 (2.53–2.37)	25–1.82 (1.92–1.82)
No. of observed reflections	76714	59288	145459
No. of unique reflections	6598	4915	10993
Acentrics	5300	3831	9087
Centrics	1298	1084	1906
Signal-to-noise ratio $I/\sigma(I)$	10.4 (8.9)	11.2 (9.4)	14.9 (9.4)
Completeness (%)	99.4 (99.4)	99.8 (100.0)	100.0 (100.0)
Multiplicity	11.6 (11.4)	12.1 (10.0)	13.2 (13.6)
R_{sym}† (%)	5.1 (6.8)	4.8 (6.3)	3.4 (7.2)
R_{ano}‡ (%)	10.8 (13.0)	11.0 (13.3)	5.0 (6.7)

† $R_{sym} = \sum_{\mathbf{h}} \sum_i |I(\mathbf{h})_i - \langle I(\mathbf{h})\rangle| / \sum_{\mathbf{h}} \sum_i I(\mathbf{h})_i$, where $\langle I(\mathbf{h})\rangle$ indicates the mean intensity of symmetry-related reflections $(\mathbf{h})_i$ of the same Bijvoet mate $\mathbf{h}$. ‡ $R_{ano} = \sum_{\mathbf{h}} |\langle I(\mathbf{h})\rangle - \langle I(-\mathbf{h})\rangle| / \sum_{\mathbf{h}} [\langle I(\mathbf{h})\rangle + \langle I(-\mathbf{h})\rangle]$, where $\langle I(\mathbf{h})\rangle$ indicates the mean intensity of symmetry-related reflections of the same Bijvoet mate $\mathbf{h}$.

Table 2
Statistics of phase determination by *OASIS* and of phase improvement by *DM* as well as refined occupancies and thermal *B* factors of the lanthanide atoms in the three lysozyme-derivative crystals.

The last row reports the correlation coefficients determined with the program *OVERLAPMAP* (Collaborative Computational Project, Number 4, 1994) between the experimental maps for the three derivatives and the respective refined model maps. The correlation coefficient between two maps is defined as $\langle xy\rangle - \langle x\rangle\langle y\rangle / [(\langle x^2\rangle - \langle x\rangle^2)(\langle y^2\rangle - \langle y\rangle^2)]^{1/2}$, where x represents the density values from one map and y the values from the other and where $\langle\rangle$ represents the mean value of the quantities inside the angle brackets. Results for the Gd-HPDO3A lysozyme derivative (Girard *et al.*, 2002) are indicated for comparison.

	100 m*M* Yb-HPDO3A, λ = 1.3859 Å	100 m*M* Lu-HPDO3A, λ = 1.3406 Å	100 m*M* Lu-HPDO3A, λ = 0.9798 Å	100 m*M* Gd-HPDO3A, λ = 1.5418 Å
Estimated f'' value (electrons)	28.0	28.0	10.4	12.0
Maximum resolution (Å)	2.15	2.37	1.82	1.72
FOM after *OASIS*	0.733	0.721	0.644	0.700
FOM after *DM*	0.771	0.761	0.772	0.836†
Real-space free *R* factor	0.164	0.201	0.152	0.127
R factor (%)	17.0	17.0	17.9	18.0
R_{free} factor (%)	21.8	22.9	20.9	21.0
Site 1				
Occupancy	0.79	0.84	0.78	0.81
B factor (Å^2)	21.9	23.3	16.6	15.3
Site 2				
Occupancy	0.70	0.74	0.69	0.73
B factor (Å^2)	30.9	37.8	24.4	15.9
Correlation coefficient	0.685	0.653	0.773	0.775

† After 12 cycles of solvent flattening and histogram matching.

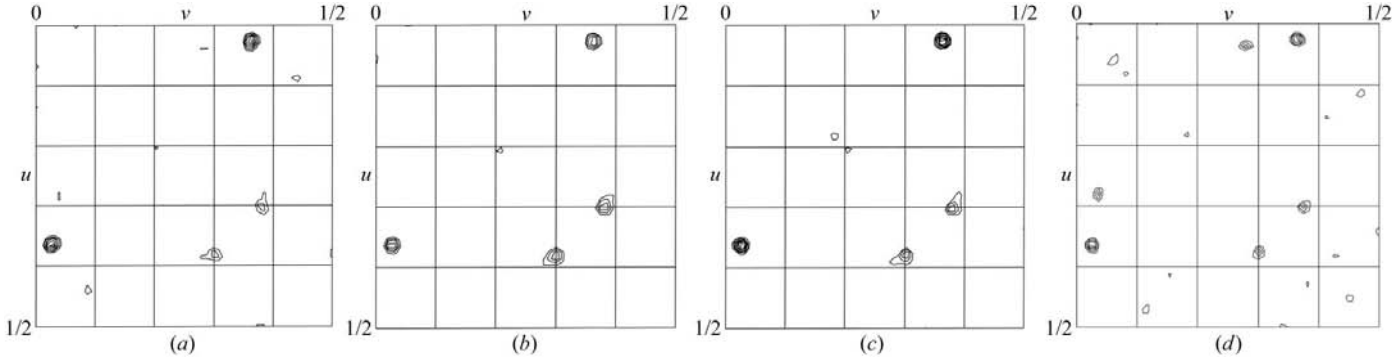

Figure 1
Harker section w = 1/2 of the anomalous Patterson map for each lysozyme derivative: (*a*) 100 m*M* Yb-HPDO3A derivative (λ = 1.3859 Å), (*b*) 100 m*M* Lu-HPDO3A derivative (λ = 1.3406 Å), (*c*) 100 m*M* Lu-HPDO3A derivative (λ = 0.9797 Å), (*d*) 100 m*M* Gd-HPDO3A derivative (λ = 1.5418 Å) (Girard *et al.*, 2002), shown for comparison. Levels are contoured at increments of 1σ, starting at 2σ.

2.4. Lanthanide-site occupancies

Subsequent refinements were performed with the program *CNS* (Brünger *et al.*, 1998) to evaluate the occupancies of the two lanthanide sites in both derivatives and compare them with the refined occupancies of the gadolinium sites in the Gd-HPDO3A HEWL derivative. Data from the whole resolution range were used without any σ-cutoff. The anomalous signal was taken into account.

The atomic coordinates from the PDB corresponding to the 100 m*M* Gd-HPDO3A lysozyme derivative (PDB code 1h87; Girard *et al.*, 2002) were used as the starting model for each refinement. Systematic bias between $R_{\rm free}$ (Brünger, 1992) values for the different derivatives was avoided by selecting for the free-reflection set of a given derivative all reflections of the Gd-HPDO3A free set present in the corresponding data set. Values for f' and f'' were modified to take into account the different experimental conditions. According to the values estimated from the fluorescence spectra by the program *CHOOCH* (Evans & Pettifer, 2001), they were fixed at -15.0 and 28.0 e^-, respectively, for the lanthanide when data were collected at the $L_{\rm III}$ absorption edge. For the Lu-HPDO3A derivative data collected at the Se *K* absorption edge, f' and f'' were set to -2.6 and 10.4 e^-, respectively (Henke *et al.*, 1993).

The complete Gd-HPDO3A HEWL structure was used as the starting model for subsequent refinement of each derivative structure, consisting of a rigid-body refinement followed by slow-cooling and individual *B*-factor refinement cycles. The occupancies of the lanthanide atoms and of the HPDO3A ligands were then refined using the *CNS* script QGROUP. The final occupancies for each derivative crystal are listed in Table 2.

3. Results and discussion

The same concentrations of gadolinium, ytterbium and lutetium HPDO3A complex were used to prepare the three lysozyme-derivative crystals. The resulting binding-site positions as well as the occupancies of the different complexes are identical.

In the early stages of data analysis, the binding effectiveness can easily be detected from the high values of $R_{\rm ano}$ (Table 1), which is about 11% for data collected at the lanthanide $L_{\rm III}$ absorption edge. $R_{\rm ano}$ remains large (5%) even for data collected on the Lu-HPDO3A derivative at the Se *K* edge.

Figures of merit (FOM) resulting from SAD phasing with *OASIS* before phase improvement are in the range 0.64–0.73. The higher the f'', the larger the FOMs. Although the solvent content of the HEWL tetragonal crystal is rather low, density modification is effective. The FOM values are greater than 0.75 after density modification with *DM*, indicating effective SAD phasing. The higher the resolution, the larger the FOMs and the lower the real-space free *R* factors.

The effectiveness of the phasing is confirmed by the quality of the electron-density maps calculated from the experimental phases. With values greater than 65%, correlation coefficients between these experimental density maps and the maps calculated from each corresponding model that were used to evaluate the lanthanide occupancies (Table 2) also confirmed the quality of the phases.

Subsequent refinements of each derivative structure led to comparable results (Table 2) for *R* and $R_{\rm free}$ values as well as for lanthanide occupancies. Nevertheless, results depend on the resolution of the data. The site occupancies are nearly identical: they are in the range 0.78–0.84 for the first binding site and in the range 0.69–0.74 for the second site. The *B* factors, which are more sensitive to resolution, are larger than those of the Gd atoms in the Gd-HPDO3A derivative. This is especially true for data collected at the $L_{\rm III}$ absorption edge. Similarly, the higher the resolution, the higher the map correlation coefficients.

The resolution of data collected at the Se *K* absorption edge on the Lu-HPDO3A derivative is comparable to that of the Gd-HPDO3A derivative. Correlation coefficients, lanthanide occupancies and *B*-factor values are similar (Table 2). This confirms that binding of the complex depends mainly on the HPDO3A ligand as shown by Girard *et al.* (2002).

4. Conclusion

Owing to their strong anomalous signal, lanthanide atoms are of special interest in determining new protein structures. We have already demonstrated (Girard *et al.*, 2002, 2003) that it is possible to solve the phase problem using the SAD method with Cu $K\alpha$ radiation, since gadolinium exhibits a rather large anomalous signal with $f'' = 12$ e^- at this wavelength. When data are collected at the $L_{\rm III}$ absorption edge, f'' reaches 28 e^-. However, not all beamlines devoted to protein crystallography using anomalous scattering are well suited to access the Gd $L_{\rm III}$ absorption edge, which is at a wavelength of 1.71 Å. Moreover, in this wavelength range, absorption becomes a serious issue.

Ytterbium and lutetium are more interesting for optimized anomalous diffraction experiments. As for all lanthanide atoms, f'' reaches 28 e^- at the $L_{\rm III}$ edge. These edges, at 1.39 and 1.34 Å respectively, are easily accessed with most synchrotron-radiation sources. Lutetium is the most advantageous since f'' is still about 10 e^- at the Se *K* absorption edge.

When data resolution is limited by the detector size, it may be advantageous to work at a shorter wavelength. Since anomalous effects are more important at high resolution, this gain facilitates the phasing step. From this point of view, lutetium is also the most interesting.

In a recent paper (Girard *et al.*, 2003), we presented a new class of seven gadolinium complexes, including the Gd-HPDO3A complex, that can be used to prepare heavy-atom derivatives. These complexes are composed of ligands that chelate a single Gd atom. They can be easily incorporated into protein crystals either by soaking or co-crystallization. Their binding effectiveness to the protein seems to mainly be related to the ligand. All these ligands are good chelating agents for the entire lanthanide series. In order to optimize anomalous diffraction experiments using synchrotron radiation, gadolinium should be substituted by lutetium in the seven complexes. To optimize the anomalous signal for in-house SAD experiments with Cu $K\alpha$ radiation (*International Tables for X-ray Crystallography*, 1985, Vol. III, pp. 214–215), it is better to substitute Gd ($f'' = 12.0$ e^-) with Sm ($f'' = 13.3$ e^-) rather than with Dy ($f'' = 8$ e^-) as proposed by Purdy *et al.* (2002). Nevertheless, since most of the Gd complexes are commercially available, they actually remain a good choice for laboratory experiments.

By using gadolinium complexes for in-house experiments and lutetium complexes for synchrotron data collection, a large set of lanthanide complexes is thus available for protein crystallography, especially for high-throughput genomics projects.

The authors would like to thank Daniela Palano at Bracco Imaging spa, Milan, Italy for the synthesis of the Yb-HPDO3A and Lu-HPDO3A samples.

References

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. & Dodson, E. (2002). *Acta Cryst.* D**58**, 494–506.
Doublié, S. (1997). *Methods Enzymol.* **276**, 523–530.
Evans, G. & Pettifer, R. F. (2001). *J. Appl. Cryst.* **34**, 82–86.
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.
Hao, Q., Gu, Y. X., Zheng, C. D. & Fan, H. F. (2000). *J. Appl. Cryst.* **33**, 980–981.
Hendrickson, W. A., Horton, J. R. & LeMaster, D. M. (1990). *EMBO J.* **9**, 1665–1672.
Hendrickson, W. A. & Ogata, C. M. (1997). *Methods Enzymol.* **276**, 494–523.
Henke, B. L., Gullikson, E. M. & Davis, J. C. (1993). *At. Data Nucl. Data Tables*, **54**, 181–342.
Matthews, B. W. (1966). *Acta Cryst.* **20**, 82–86.
North, A. C. T. (1965). *Acta Cryst.* **18**, 212–216.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Purdy, M. D., Ge, P., Chen, J., Selvin, P. R. & Wiener, M. C. (2002). *Acta Cryst.* D**58**, 1111–1117.
Weeks, C. M. & Miller, R. (1999). *J. Appl. Cryst.* **32**, 120–124.