**Balakrishnan Vijayakumar and
Devadasan Velmurugan***Centre of Advanced Study in Crystallography
and Biophysics, University of Madras,
Maraimalai (Guindy) Campus,
Chennai 600 025, India

Correspondence e-mail: shirai2011@gmail.com

Received 13 July 2012
Accepted 19 November 2012**PDB Reference:** lysozyme, europium derivative,
4h1p

Use of europium ions for SAD phasing of lysozyme at the Cu $K\alpha$ wavelength

Europium is shown to be a good anomalous scatterer in SAD phasing for solving the structure of biological macromolecules. The large value of the anomalous contribution of europium, $f'' = 11.17\text{ e}^-$, at the Cu $K\alpha$ wavelength is an advantage in *de novo* phasing and automated model building. Tetragonal crystals of hen egg-white lysozyme (HEWL) incorporating europium(III) chloride (50 mM) were obtained which diffracted to a resolution of 2.3 Å at a wavelength of 1.54 Å (Cu $K\alpha$). The master data set (360° frames) was split and analyzed for anomalous signal-to-noise ratio, multiplicity, completeness, SAD phasing and automated building. The structure solution and model building of the split data sets were carried out using *phenix.autosol* and *phenix.autobuild*. The contributions of the Eu ions to SAD phasing using in-house data collection are discussed. This study revealed successful lysozyme phasing by SAD using laboratory-source data involving Eu ions, which are mainly coordinated by the side chains of Asn46, Asp52 and Asp101 together with some water molecules.

1. Introduction

The single-wavelength anomalous diffraction (SAD) method has been effectively established in structural biology in terms of data-collection techniques, accurate detectors and software and is used for data processing, phasing, density modification and model building for the determination of new macromolecular structures. Recent developments in the preparation of heavy-atom derivatives are more accurate and less disruptive than traditional heavy-atom soaking in high-throughput structure determination using X-ray crystallography (Moiseeva & Allaire, 2007). The incorporation of heavy atoms into protein crystals for phasing was pioneered by Green *et al.* (1954). Successful SAD phasing has been carried out using the anomalous signal from native protein atoms, including sulfur (Hendrickson & Teeter, 1981; Dauter *et al.*, 1999), iron (Royer *et al.*, 1989; Wu *et al.*, 2001) and copper (Zheng *et al.*, 1996; Harvey *et al.*, 1998), or from incorporated heavy atoms such as caesium (Wallace *et al.*, 1990; Burkhart *et al.*, 1998), iodine (Chen *et al.*, 1991), lead (Biou *et al.*, 1995), mercury (Fan *et al.*, 1990), selenium (Sha *et al.*, 1995), platinum (Sabini *et al.*, 2000), cadmium (Yogavel *et al.*, 2010), cerium (Vennila & Velmurugan, 2011), gadolinium (Nagem *et al.*, 2001), the $[\text{Eu}(\text{DPA})_3]^{3-}$ complex (Pompidor *et al.*, 2010) and EuCl_3 (Ruggiero *et al.*, 2011). In this study, diffraction data were collected to 2.3 Å resolution at the Cu $K\alpha$ wavelength; the master data set (360°) was split and the anomalous pairs were separately merged as I^+ and I^- in order to analyze the anomalous signal-to-noise ratio, multiplicity and completeness. Europium(III) ions were used to determine the structure by Eu-SAD phasing, and the use of the $R_{\text{anom}}/R_{\text{p.i.m.}}$ ratio for estimating the significance of the anomalous signal is discussed.

Europium, with atomic number 63, is the seventh element of the lanthanide series. It has an L_{III} absorption edge at 1.77 Å. This wavelength is near to Cu $K\alpha$ and is an easily accessible wavelength range on synchrotron beamlines. The anomalous scattering coefficients of europium are $f' = -7.838\text{ e}^-$ and $f'' = 11.17\text{ e}^-$ at the Cu $K\alpha$ wavelength (http://skuld.bmsc.washington.edu/scatter/AS_periodic.html). The present work, SAD phasing at the Cu $K\alpha$ wavelength, was successful owing to the high f'' value of the anomalous scatterer.

2. Materials and methods

2.1. Crystallization and data collection

Crystallization-grade hen egg-white lysozyme (HEWL) powder from Hampton Research was used for this study. The reservoir solution (1 ml) was made up of 1 M NaCl in 50 mM sodium acetate pH 4.5. The protein was cocrystallized with europium chloride in the protein-solution droplet (5 μ l), which consisted of 2.5 μ l protein solution (30 mg ml⁻¹ protein in 50 mM sodium acetate pH 4.5), 1.5 μ l reservoir solution and 1 μ l 50 mM europium chloride, using the hanging-drop vapour-diffusion method at 293 K. Tetragonal crystals were obtained within 2 d. A crystal was fished out using a copper magnetic loop (Hampton Research) and was immersed into 25% ethylene glycol for about 10–15 s. The crystal was quickly cooled in a liquid-nitrogen stream (Oxford Cryosystems) at 100 K. Data were collected at the Cu K α wavelength (1.54 Å) using the home source at the GNR X-ray facility: a MICROSTAR rotating high-brilliance anode (HBA) generator (Bruker AXS) operating at 45 kV and 60 mA equipped with Helios focal optics. Diffraction data were collected from the crystal using 3 min exposure and 1° oscillation per frame (missetting angles 17.254°, 29.779°, -19.352°) and 360° images were collected to 2.3 Å resolution on a MAR345 dtb image-plate detector (MAR Research Inc.) at a distance of 200 mm from the crystal.

2.2. Software packages used

The master images (360°) were split into five subsets of data, 45°, 90°, 180°, 270° and 360°, to analyze the effectiveness of the anomalous scattering of the europium in macromolecular phasing and automated model building. These data sets were integrated and scaled by the *automar* software (MAR Research GmbH, Germany). The anomalous pairs were merged separately as I^+ and I^- . The relevant data-collection statistics are summarized in Table 1. The standard SAD protocol was used in the *PHENIX* software (Adams *et al.*, 2010) for substructure solution, SAD phasing, automatic model building and refinement. The *phenix.xtriage* (Zwart *et al.*, 2005) program was used to analyze the anomalous measurability from the scaled intensity data sets. *HySS* (Hybrid Substructure Search) was used to find substructure sites, *Phaser* was used to calculate experimental phases and *RESOLVE* (Terwilliger, 2000) was used for density modification and partial model building in *phenix.autosol* (Terwilliger *et al.*, 2009). These partial models were fed into *phenix.autobuild* (Terwilliger *et al.*, 2008), a combination of density modification and chain tracing was performed using *RESOLVE*, and *phenix.refine* (Afonine *et al.*, 2005) was used to generate a quality model. *Coot* (Emsley & Cowtan, 2004) was used to build missing residues and side chains from the autobuilt model. *REFMAC5* (Murshudov *et al.*, 2011) and *phenix.refine* were used to refine the manually built model. Figures were generated using *Chimera* (Pettersen *et al.*, 2004) and *PyMOL* (DeLano, 2002). The quality of all atomic models was assessed with *PROCHECK* (Laskowski *et al.*, 1993) and *MolProbity* (Chen *et al.*, 2010).

3. Results and discussion

3.1. Analysis of the data quality and anomalous signal

Mean multiplicities of 3.31, 6.31, 12.4, 18.6 and 24.7 were observed in the 45°, 90°, 180°, 270° and 360° data sets, respectively. The $\langle I/\sigma(I) \rangle$ values for the 27.4–2.3 Å resolution range were 20.7, 15.8, 28.5, 25.0 and 25.2 for the respective data sets. The overall completeness was 88% for the 45° data set and 99% for remaining data sets. The R_{merge}

Table 1

Summary of data-collection statistics, substructure solution, SAD phasing, model building and refinement.

Data set (°)	45	90	180	270	360
Unit-cell parameters (Å, °)	$a = b = 77.48$, $c = 36.99$, $\alpha = \beta = \gamma = 90$				
Space group	$P4_32_12$				
Resolution (Å)	30.0–2.3				
Mosaicity (°)	0.073	0.050	0.050	0.050	0.050
$R_{\text{merge}}^\dagger$ (%)	2.64 (4.12)	3.60 (6.53)	3.39 (6.73)	4.17 (7.31)	4.46 (7.46)
Completeness (%)	88.0 (91.6)	99.7 (99.8)	99.9 (99.3)	99.9 (100)	99.9 (100)
Multiplicity	3.31 (3.20)	6.31 (6.48)	12.4 (12.6)	18.6 (19.0)	24.7 (24.9)
$\langle I/\sigma(I) \rangle$	20.7 (14.0)	15.8 (7.1)	28.5 (13.4)	25.0 (14.3)	25.2 (11.2)
$R_{\text{anom}}^\dagger$ (%)	5.41 (6.63)	5.15 (6.21)	4.97 (6.23)	4.80 (5.81)	4.77 (5.73)
$R_{\text{r.i.m.}}^\dagger$ (%)	3.44 (5.73)	4.24 (7.72)	4.31 (7.28)	4.38 (7.70)	4.63 (7.75)
$R_{\text{p.i.m.}}^\dagger$ (%)	2.17 (3.76)	2.22 (4.08)	1.62 (2.78)	1.35 (2.39)	1.23 (2.10)
Unique reflections	4838	5351	5354	5353	5353
Bijvoet pairs	3131	4211	4211	4211	4211
Lone Bijvoet pairs	836	0	0	0	0
Anomalous signal	0.079	0.076	0.068	0.066	0.066
<i>phenix.autosol</i>					
FOM	0.39	0.46	0.47	0.49	0.49
Refined heavy atoms	3	4	4	6	6
Bayesian score ‡	18.7	39.1	44.0	46.0	45.0
Model-map CC	0.33	0.46	0.48	0.59	0.52
Residues built	47	62	81	97	84
<i>phenix.autobuild</i>					
No. of residues/side chains built	81/0	112/112	119/113	113/113	119/119
Model-map CC	0.46	0.72	0.72	0.73	0.76
Waters	0	56	49	50	51
R_{work}	0.48	0.29	0.29	0.29	0.28
R_{free}	0.53	0.35	0.34	0.33	0.33
Refinement					
Resolution (Å)					26.75–2.30
$R_{\text{work}}/R_{\text{free}}$					0.17/0.23
No. of residues/waters					128/125
Eu ions					2
Overall mean B factor (Å ²)					21.59
Stereochemistry					
Bond lengths (Å)					0.007
Bond angles (°)					1.057
Ramachandran plot, residues in (%)					
Favoured regions					96.2
Allowed regions					3.8
Disallowed regions					0

† Weiss (2001). ‡ Terwilliger *et al.* (2009).

and R_{anom} values were below 4.46% and 5.51%, respectively, in all rotation ranges (Table 1). The expected anomalous signal ($\langle \Delta F \rangle / \langle F \rangle = 0.10$) was estimated using the modified Hendrickson formula (Hendrickson & Teeter, 1981; Dauter *et al.*, 2002), $\langle \Delta F \rangle / \langle F \rangle_{\text{exp}} = 2^{1/2} [(f''_A N_A^{1/2})] / [f_{\text{eff}}(\theta) N_P^{1/2}]$, where f''_A are the imaginary scattering contributions of europium ions and N_A and N_P are the number of Eu atoms and the total number of non-H atoms in the molecule, respectively. Here, N_A is 2 and N_P is 1001. $f_{\text{eff}}(\theta)$, which is the average number of electrons for protein atoms, is 6.7. The ratio of R_{anom} to $R_{\text{p.i.m.}}$ has been used as a practical indicator of the level of anomalous signal (Mueller-Dieckmann *et al.*, 2004, 2005; Weiss, 2001; Weiss *et al.*, 2004, 2005; Weiss, Sicker, Djinoovic-Carugo *et al.*, 2001; Weiss, Sicker & Hilgenfeld, 2001), which can provide information on whether or not the data are sufficient for structure determination. As a rule of thumb, an $R_{\text{anom}}/R_{\text{p.i.m.}}$ ratio of ≥ 1.5 is sufficient to determine the structure (Mueller-Dieckmann *et al.*, 2005). Analysis of the $R_{\text{anom}}/R_{\text{p.i.m.}}$ ratio for the europium data revealed that sufficient anomalous signal was exhibited in each data subset (Table 2).

3.2. Substructure determination, phasing and model building

The anomalous substructures were located in *phenix.autosol* using a 3.0 Å resolution cutoff. Between three and six heavy-atom sites (including S, Na and Cl atoms) were refined for phasing for the

various φ rotation ranges and the highest two peaks corresponded to europium ions. The initial phases obtained from europium were sufficient for automatic tracing of the protein structure in preliminary model building. Figures of merit (FOMs) of 0.39–0.49 and correlation coefficients (CCs) of 0.33–0.52 were observed for the 45–360° data sets. In an asymmetric unit, preliminary models consisting of 47, 62, 81, 97 and 84 residues were built in the data sets. These partial models were fed into *phenix.autobuild* for iterative model building with refinement. The overall CC was 0.76 and 92% of the residues were built for the 360° data. The other data sets also showed good CCs with good automated model building (Figs. 1*a–e*). The 360° data set alone

was used for manual building and refinement in *Coot* and *phenix.refine*, respectively. This model was finally refined to R_{work} and R_{free} values of 17.3% and 22.9%, respectively. The occupancy values of the bound Eu ions were refined using *phenix.refine*.

3.3. Europium-binding sites

Two europium ions are bound to the surface of the crystal structure without affecting the crystal lattice. The molecular structure remains the same after binding and is shown in Fig. 1(*f*). The ion-binding sites were validated by the anomalous difference Fourier map up to the

Table 2

Statistics of the $R_{\text{anom}}/R_{\text{p.i.m.}}$ ratio for different φ rotation ranges.

Total φ rotation (°)	Resolution (Å)									
	30.0–4.96	4.96–3.93	3.93–3.44	3.44–3.12	3.12–2.90	2.90–2.73	2.73–2.59	2.59–2.48	2.48–2.38	2.38–2.30
45	3.65	2.33	2.45	2.09	2.67	2.61	2.34	2.43	2.11	1.76
90	4.16	2.32	2.08	1.95	2.47	2.17	2.08	2.02	1.81	1.52
180	5.51	2.94	2.71	2.57	2.92	3.08	2.74	2.88	2.48	2.24
270	6.71	3.51	3.08	2.94	3.55	3.36	3.11	3.18	2.67	2.43
360	6.78	3.67	3.33	3.31	3.92	3.62	3.51	3.61	2.98	2.73

Figure 1

Cartoon representations of the automatically built lysozyme-molecule model in the asymmetric unit: (*a*) 45°, (*b*) 90°, (*c*) 180°, (*d*) 270° and (*e*) 360° data sets. (*f*) Europium ions and sodium ion in the final refined model (360°) are shown in cyan and magenta, respectively.

Figure 2

(a) Anomalous difference Fourier map of binding sites of europium ions (red) at the 15σ level. (b) The Eu-1 ion and (c) the Eu-2 ion (cyan), showing interactions with amino-acid residues and water molecules.

15σ level (Fig. 2a). The europium ions are mainly coordinated by water molecules and by aspartic acid and asparagine side chains. Chloride and sodium ions were assigned based on the difference peak and binding environment of the protein. A distance of 16.7 Å is observed between the two europium ions. Eu-1 is coordinated by five water molecules and the Asp101 OD1 and Asp101 OD2 atoms at the solvent boundary. The occupancy of the ion is 0.57. Eu-2 is bound by six water molecules and two residues, Asn46 OD1 and Asp52 OD2, in the catalytic cleft; its occupancy is 0.69 (Figs. 2b and 2c). Both europium-binding sites are identical to those found for Ho^{3+} (Jakoncic *et al.*, 2006) and Ce^{3+} (Vennila & Velmurugan, 2011) ions in HEWL.

4. Conclusion

In this study, the results using the five subsets of data show that europium ions can effectively be used as an anomalous scatterer for SAD phasing and model building at the Cu $K\alpha$ wavelength. The 45° data set with a multiplicity of 3.31 was sufficient for SAD phasing and structure determination of the 14.4 kDa protein. The protein cocrystallized with a lower concentration of europium (50 mM) produced sufficient phasing power to allow automatic structure determination. The absorption edge of europium (1.77 Å) is near the Cu $K\alpha$ wavelength. As a consequence, the europium ion has a large f'' value which is sufficient for phasing and model building. Without altering the structure, the europium ions are bound to the surface of the protein and interact with the side chains of Asn and Asp residues and with water molecules. Eu^{3+} , Ho^{3+} and Ce^{3+} ions interact with the same binding-site residues in HEWL. The results show that Eu can provide sufficient anomalous signal for phasing and model building to determine new macromolecular structures using the SAD technique.

BV thanks the University Grants Commission (UGC), Government of India, New Delhi for a Meritorious Fellowship. The authors

acknowledge the Department of Biotechnology, Government of India for financial support of the in-house macromolecular data-collection facility

References

- Adams, P. D. *et al.* (2010). *Acta Cryst.* **D66**, 213–221.
- Afonine, P. V., Grosse-Kunstleve, R. W. & Adams, P. D. (2005). *CCP4 Newsl. Protein Crystallogr.* **42**, contribution 8.
- Biou, V., Shu, F. & Ramakrishnan, V. (1995). *EMBO J.* **14**, 4056–4064.
- Burkhardt, B. M., Li, N., Lango, D. A., Pangborn, W. A. & Duax, W. L. (1998). *Proc. Natl Acad. Sci. USA*, **95**, 12950–12955.
- Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. & Richardson, D. C. (2010). *Acta Cryst.* **D66**, 12–21.
- Chen, L., Rose, J. P., Breslow, E., Yang, D., Chang, W. R., Furey, W. F., Sax, M. & Wang, B.-C. (1991). *Proc. Natl Acad. Sci. USA*, **88**, 4240–4244.
- Dauter, Z., Dauter, M., de La Fortelle, E., Bricogne, G. & Sheldrick, G. M. (1999). *J. Mol. Biol.* **289**, 83–92.
- Dauter, Z., Dauter, M. & Dodson, E. J. (2002). *Acta Cryst.* **D58**, 494–506.
- DeLano, W. L. (2002). *PyMOL*. <http://www.pymol.org>.
- Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* **D60**, 2126–2132.
- Fan, H.-F., Hao, Q., Gu, Y.-X., Qian, J.-Z., Zheng, C.-D. & Ke, H.-M. (1990). *Acta Cryst.* **A46**, 935–939.
- Green, D. W., Ingram, V. M. & Perutz, M. F. (1954). *Proc. R. Soc. London Ser. A*, **225**, 87–307.
- Harvey, I., Hao, Q., Duke, E. M. H., Ingledew, W. J. & Hasnain, S. S. (1998). *Acta Cryst.* **D54**, 629–635.
- Hendrickson, W. C. & Teeter, M. M. (1981). *Nature (London)*, **290**, 107–113.
- Jakoncic, J., Di Michiel, M., Zhong, Z., Honkimaki, V., Jouanneau, Y. & Stojanoff, V. (2006). *J. Appl. Cryst.* **39**, 831–841.
- Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.
- Moiseeva, N. & Allaire, M. (2007). *Acta Cryst.* **D63**, 1025–1028.
- Mueller-Dieckmann, C., Panjikar, S., Tucker, P. A. & Weiss, M. S. (2005). *Acta Cryst.* **D61**, 1263–1272.
- Mueller-Dieckmann, C., Polentarutti, M., Djinnovic Carugo, K., Panjikar, S., Tucker, P. A. & Weiss, M. S. (2004). *Acta Cryst.* **D60**, 28–38.
- Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* **D67**, 355–367.

- Nagem, R. A. P., Dauter, Z. & Polikarpov, I. (2001). *Acta Cryst.* **D57**, 996–1002.
- Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C. & Ferrin, T. E. (2004). *J. Comput. Chem.* **25**, 1605–1612.
- Pompidor, G., Maury, O., Vicat, J. & Kahn, R. (2010). *Acta Cryst.* **D66**, 762–769.
- Royer, W. E., Hendrickson, W. A. & Chiancone, E. (1989). *J. Biol. Chem.* **264**, 21052–21061.
- Ruggiero, A., Squeglia, F., Marasco, D., Marchetti, R., Molinaro, A. & Berisio, R. (2011). *Biochem. J.* **435**, 33–41.
- Sabini, E., Schubert, H., Murshudov, G., Wilson, K. S., Siika-Aho, M. & Penttilä, M. (2000). *Acta Cryst.* **D56**, 3–13.
- Sha, B.-D., Liu, S.-P., Gu, Y.-X., Fan, H.-F., Ke, H., Yao, J.-X. & Woolfson, M. M. (1995). *Acta Cryst.* **D51**, 342–346.
- Terwilliger, T. C. (2000). *Acta Cryst.* **D56**, 965–972.
- Terwilliger, T. C., Adams, P. D., Read, R. J., McCoy, A. J., Moriarty, N. W., Grosse-Kunstleve, R. W., Afonine, P. V., Zwart, P. H. & Hung, L.-W. (2009). *Acta Cryst.* **D65**, 582–601.
- Terwilliger, T. C., Grosse-Kunstleve, R. W., Afonine, P. V., Moriarty, N. W., Zwart, P. H., Hung, L.-W., Read, R. J. & Adams, P. D. (2008). *Acta Cryst.* **D64**, 61–69.
- Vennila, K. N. & Velmurugan, D. (2011). *Acta Cryst.* **F67**, 1662–1665.
- Wallace, B. A., Hendrickson, W. A. & Ravikumar, K. (1990). *Acta Cryst.* **B46**, 440–446.
- Weiss, M. S. (2001). *J. Appl. Cryst.* **34**, 130–135.
- Weiss, M. S., Mander, G., Hedderich, R., Diederichs, K., Ermler, U. & Warkentin, E. (2004). *Acta Cryst.* **D60**, 686–695.
- Weiss, M. S., Panjikar, S., Mueller-Dieckmann, C. & Tucker, P. A. (2005). *J. Synchrotron Rad.* **12**, 304–309.
- Weiss, M. S., Sicker, T., Djinojic-Carugo, K. & Hilgenfeld, R. (2001). *Acta Cryst.* **D57**, 689–695.
- Weiss, M. S., Sicker, T. & Hilgenfeld, R. (2001). *Structure*, **9**, 771–777.
- Wu, C.-K., Dailey, H. A., Rose, J. P., Burden, A., Sellers, V. M. & Wang, B.-C. (2001). *Nature Struct. Biol.* **8**, 156–160.
- Yogavel, M., Nithya, N., Suzuki, A., Sugiyama, Y., Yamane, T., Velmurugan, D. & Sharma, A. (2010). *Acta Cryst.* **D66**, 1323–1333.
- Zheng, X.-F., Fan, H.-F., Hao, Q., Dodd, F. E. & Hasnain, S. S. (1996). *Acta Cryst.* **D52**, 937–941.
- Zwart, P. H., Grosse-Kunstleve, R. W. & Adams, P. D. (2005). *CCP4 Newsl. Protein Crystallogr.* **43**, contribution 7.