Acta Crystallographica Section F
Structural Biology and Crystallization Communications
ISSN 1744-3091

In-house SAD phasing with surface-bound cerium ions

K. Natesan Vennila and Devadasan Velmurugan*

CAS in Crystallography and Biophysics, University of Madras, Maraimalai (Guindy) Campus, Chennai, Tamilnadu 600 025, India

Correspondence e-mail: shirai2011@gmail.com

Received 10 June 2011
Accepted 2 September 2011

The anomalous signal of cerium(III) ions present in a derivative of hen egg-white lysozyme (HEWL) crystals obtained by the addition of 0.025 *M* cerium chloride to the crystallization medium was used for phasing. X-ray intensity data were collected to 2 Å resolution using an in-house Cu $K\alpha$ radiation data-collection facility. Phasing of a single-wavelength data set purely based on its f'' led to a clearly interpretable electron-density map. Automated substructure solution by *AutoSol* in *PHENIX* resulted in four highest peaks corresponding to cerium(III) ions with data limited to 3 Å resolution, and about 90% of the residues were built automatically by *AutoBuild* in *PHENIX*. Cerium(III) ions bound on the surface of the enzyme are found to interact mainly with the main-chain and side-chain carbonyl groups of Asn, Glu, Tyr and Asp and with water molecules. Ce^{3+} ions were used as potential anomalous scatterers for the in-house single-wavelength anomalous scattering technique, and this is proposed as a tool for macromolecular phasing and for the study of the interactions of trivalent metal ions with proteins and other macromolecules.

1. Introduction

The automation of protein crystallography and its application to difficult, large and complex biological systems require a deeper understanding of various aspects of crystallography. The non-availability of synchrotron X-ray sources, expense and difficulties in dealing with the precious protein crystals has led to advances in home-source X-ray diffraction experiments and phasing. From the perspective of in-house SAD phasing, sulfur SAD and halide soaking have been successfully utilized for various proteins (Dauter & Dauter, 1999; Dauter *et al.*, 2000). Other heavy atoms used for derivatization and for SAD phasing include Hg, Ba (Moiseeva & Allaire, 2007), Sm (Appleby *et al.*, 2005), Rb (Korolev *et al.*, 2001), Cs, Gd (Nagem *et al.*, 2001), Cd (Yogavel *et al.*, 2010), Ho (Ismaya *et al.*, 2011; Jakoncic *et al.*, 2006) and Co (Gunčar *et al.*, 2007). In addition to monovalent, divalent and trivalent cations and anions, compounds that contain the heavy atom as one of their constituents, such as I3C (Beck *et al.*, 2008) and the so-called 'magic seven' compounds including $KAuCl_4$, K_2PtCl_6 and tetramethyllead acetate (Sun *et al.*, 2002), have also been reported to be successful in SAD phasing. In the case of in-house S-SAD experiments, high redundancy and accuracy of data collection are important (Dauter *et al.*, 1999; Weiss *et al.*, 2001; Ramagopal *et al.*, 2003) and are difficult to achieve in many cases. The quick halide-soaking experiments also have some demerits such as lattice instability at higher salt concentrations and at long soaking times of greater than 1 min (Sun *et al.*, 2002). In this direction, the use of a new source of anomalous scattering, Ce(III) ions, which can also be effectively used for automated SAD phasing, is described in the present paper.

Cerium, the second element in the lanthanide series, has its L_I absorption edge at 1.8932 Å and its anomalous scattering coefficients are $f' = -1.88$ e and $f'' = 9.76$ e (http://skuld.bmsc.washington.edu/scatter/AS_periodic.html) for the Cu $K\alpha$ wavelength. Cerium can exist as a trivalent ion (ionic radius 1.15 Å) in solution. The present work attempts to determine whether cerium can be incorporated into the protein crystal while maintaining the crystal isomorphism and also to find out whether sufficient cerium ions occupy ordered posi-

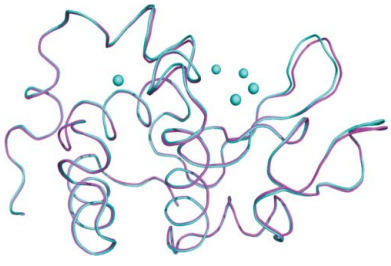

Table 1
Data-processing, substructure-solution and phasing statistics.

Values in parentheses are for the highest resolution shell.

Crystal data	45°	65°	90°	180°	360°
Unit-cell parameters (Å, °)	a = 77.44, b = 77.44, c = 37.09, α = 90.00, β = 90.00, γ = 90.00				
Resolution limits (Å)	30.0–2.01 (2.08–2.01)				
R_{sym}† (%)	2.08 (2.6)	2.25 (3.9)	2.56 (4.7)	3.42 (5.3)	5.74 (8.7)
Mosaicity (°)	0.32	0.32	0.34	0.34	0.34
R_{merge}† (%)	2.2	2.4	2.7	3.2	5.7
Completeness (%)	75.7 (78.5)	92.6 (92.7)	99.5 (97.7)	99.6 (97.7)	99.5 (96.4)
Multiplicity	4.20 (3.0)	5.03 (4.7)	7.16 (6.8)	13.1 (12.2)	25.1 (23.8)
$\langle I/\sigma(I)\rangle$	22.4 (11.4)	21.7 (11.0)	22.7 (11.2)	20.8 (11.2)	15.6 (8.1)
R_{anom}† (%)	3.29 (4.5)	3.28 (4.2)	3.12 (4.1)	3.02 (5.64)	3.00 (3.8)
$R_{r.i.m.}$† (%)	2.69 (5.0)	2.85 (5.2)	3.15 (5.7)	3.85 (5.86)	5.96 (9.1)
$R_{p.i.m.}$† (%)	1.56 (3.2)	1.58 (3.1)	1.56 (2.9)	1.38 (2.26)	1.59 (2.5)
Unique reflections	6053	7402	7954	7961	7969
Bijvoet pairs	4874	5993	6422	6437	6456
Lone Bijvoet mates	115	196	31	19	13
Anomalous signal	0.046	0.045	0.043	0.042	0.042
$R_{anom}/R_{p.i.m.}$	2.11	2.08	2.00	2.79	1.89
AutoSol					
Heavy-atom sites	22	22	22	22	22
Bayesian CC‡	17.76 ± 14.31	25.37 ± 14.52	33.2 ± 26.2	38.04 ± 11.46	32.7 ± 26.6
FOM	0.39	0.43	0.46	0.48	0.47
Model–map correlation‡	0.23	0.32	0.48	0.50	0.52
Residues built	31	48	88	89	87
Side chains	0	0	27	30	40
AutoBuild					
R	0.52	0.40	0.37	0.27	0.25
R_{free}	0.55	0.45	0.39	0.32	0.30
Residues built	75	115	115	115	115
Side chains	5	92	115	115	115
Model–map correlation	0.43	0.58	0.61	0.70	0.71

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

tions which can be used for phasing. Hen egg-white lysozyme was used for this study.

2. Materials and methods

Hen egg-white lysozyme derivative crystals were prepared as per previously described reports (Evans & Bricogne, 2002, 2003) but using cerium chloride as the heavy-atom solution in the protein drop. The cerium chloride used was of analytical grade (Sigma–Aldrich, USA). Concentrations of cerium chloride in the range 25–100 m*M* were used and well diffracting crystals were obtained in this range. The crystal was cryoprotected with 25% ethylene glycol and flash-cooled in a stream of liquid nitrogen at 100 K.

A crystal from a well in which 25 m*M* cerium chloride was used in the protein drop was used for data collection using a Bruker Microstar rotating-anode generator equipped with a MAR345dtb image-plate system. The crystal was mounted arbitrarily and diffraction images were collected over a total angular range of 360° with an oscillation angle of 1°. The crystal was exposed to X-rays for 60 s per frame.

3. Results and discussion

3.1. Anomalous signal and substructure-solution analysis

The data were processed with *automar* (MAR Research GmbH) by separately merging the anomalous pairs as I^+ and I^-. The redundancy of the data collected to 2 Å resolution at 100 K for the complete data set was 25.1 overall and 23.8 in the high-resolution bin. The intensity data were analysed for the presence of anomalous signal using *phenix.xtriage* in *PHENIX* (Adams *et al.*, 2010) and *HKL2MAP* (Pape & Schneider, 2004). The full resolution range of the data was found to have anomalous signal measurability greater than 10%. The expected anomalous signal with four cerium ions calculated using the modified Hendrickson formula (Hendrickson & Teeter, 1981; Dauter *et al.*, 2002) was 0.012 and the intensity data collected had an observed anomalous signal of 0.04, which was greater than the expected value.

In order to find out the redundancy limit for solving the structure using a cerium derivative, the master data were split up into five data sets with oscillation angles of 45, 65, 90, 180 and 360°. The data-processing statistics for all of these data sets are shown in Table 1. Substructure solution was carried out for all of the data sets using *AutoSol* in *PHENIX*. *HySS* located two peaks corresponding to cerium. *Phaser* was used to refine the positions, occupancies and *B*

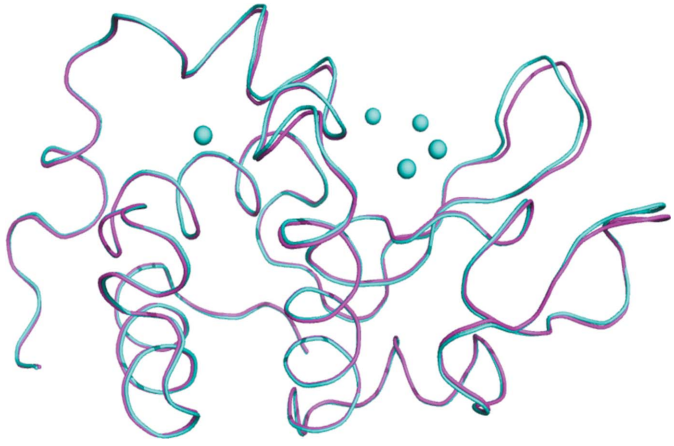

Figure 1
Superposition of the atomic resolution structure of HEWL (PDB entry 193l; magenta; Vaney *et al.*, 1996) with the 2.0 Å resolution structure of HEWL cocrystallized with $CeCl_3$ (cyan). The r.m.s.d. value was found to be 0.18 Å.

factors of the cerium ions and to derive the estimates of the protein phases, and resulted in 22 refined sites including those of sulfur and chlorine; the highest four peaks corresponded to cerium. The occupancies of the four cerium ions refined to 0.45, 0.22, 0.22 and 0.30, respectively; the second and third of these were interpreted as weakly bound cerium ions. The temperature factors of the four cerium ions refined to 15.7, 29.0, 31.3 and 13.2 Å^2, respectively. The phases calculated by *Phaser* were further improved by density modification using *RESOLVE* in *AutoBuild*. Automated model building was carried out by *AutoBuild* in *PHENIX* and a total of 115 residues were built, all with side chains, for the master data set. For the data set with redundancy 5.03, automated solution and model building by *PHENIX* resulted in the building of 115 amino-acid residues, of which 92 were built with side chains, indicating that the data set with reflections collected up to 65° was sufficient to solve the structure. The substructure solution and refinement statistics are shown in Table 1. The overall figure of merit (FOM) ranged from 0.39 to 0.47 for the five data sets.

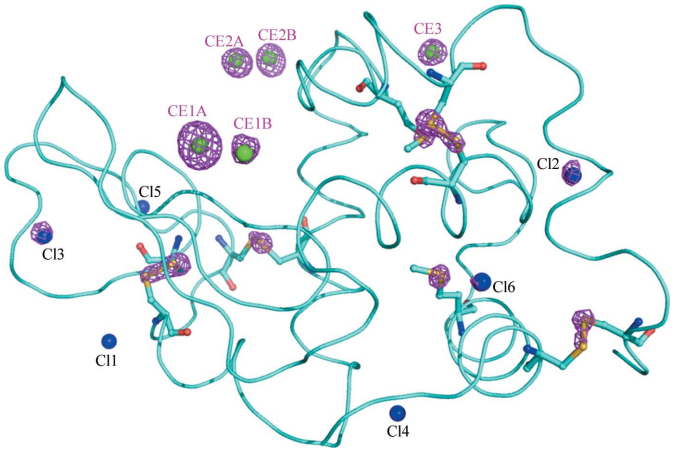

Figure 2
Anomalous map at the 5σ level showing the peaks of four cerium ions (magenta), two chloride ions (green) and sulfurs (yellow)

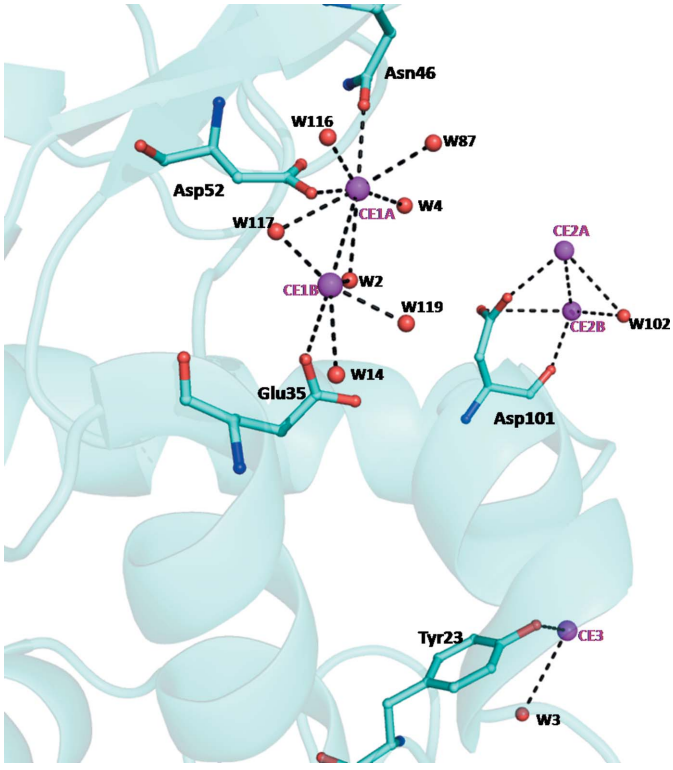

Figure 3
Cerium(III) ion binding sites.

Clearly interpretable maps were obtained after automated model building. Manual rebuilding was performed using *Coot* (Emsley & Cowtan, 2004) and the model was finally refined to R and $R_{\rm free}$ values of 16.1% and 21.4%, respectively, by *phenix.refine* (Afonine *et al.*, 2005). The two cerium ions (CE1 and CE2) were found to occupy two positions each with occupancies of 0.45/0.10 and 0.22/0.23, respectively; the corresponding thermal factors are 15.7/32.3 and 31.5/35.1 Å^2. The third cerium ion also bound on the surface with 35% occupancy and was more stable, with a thermal factor of 16.8 Å^2.

4. Ce^{3+} ion binding sites

The molecular structure remains the same after cerium ion binding and is shown in Fig. 1. Inspection of the cerium ions in the structure reveals that the cerium ions are bound on the surface of the protein, replacing water molecules in the crystal lattice. The positions of the cerium ions were validated by the anomalous map shown in Fig. 2. The cerium ions are mainly coordinated to water molecules and the side-chain and main-chain carbonyl groups of glutamic acid, tyrosine, aspartic acid and asparagine. The other peaks in the anomalous map were assigned as chloride ions and sulfurs. Chloride ions were distinguished from cerium ions based on their binding patterns as described in previously reported structures (Dauter & Dauter, 1999; Evans & Bricogne, 2002).

Fig. 3 shows that the first cerium ion occupies two closely related positions 3.6 Å apart in an ∼45%/10% ratio in the structure. It is highly coordinated by five water molecules and by the side-chain carbonyl groups of Asn46, Glu35 and Asp52. The coordination distances to the side-chain carbonyl groups of Asn46, Glu35 and Asp52 are 2.22, 3.4 and 2.6 Å, respectively. HEWL has been found to have the same binding site for Ho^{3+} ion (Jakoncic *et al.*, 2006). The cerium ion binding site is also similar to the gadolinium binding site reported by Nagem *et al.* (2001). This type of disorder was also observed for the second cerium ion, which occupies two close positions 3.4 Å apart in a ∼22%/23% ratio. The second cerium ion shown in Fig. 3 interacts with the main-chain and side-chain carbonyl groups of Asp101 with coordination distances of 2.6 and 2.7 Å, respectively. The third cerium ion is at a coordination distance of 3 Å from Tyr23. These common types of interactions suggest that the cerium ions may bind to other proteins in a similar way.

5. Conclusion

These results confirm cerium chloride as a potential source of automated phasing using anomalous scattering data collected using an in-house Cu $K\alpha$ source. Owing to its high anomalous scattering coefficient ($f'' = 9.74$) for home-source Cu $K\alpha$, cerium proves to be an excellent heavy-atom marker for automated phasing. The addition of three cerium ions led to sufficient anomalous signal for the phasing of 129 amino-acid residues with a redundancy of 5.03. The high solubility of cerium chloride in water and the binding of the cerium(III) ions to the surface of the enzyme are added advantages for phasing.

The authors acknowledge the Department of Biotechnology, Government of India for financial support of the in-house macro-

molecular data-collection facility. VN thanks the CSIR, New Delhi for providing a Senior Research Fellowship.

References

Adams, P. D. *et al.* (2010). *Acta Cryst.* D**66**, 213–221.
Afonine, P. V., Grosse-Kunstleve, R. W. & Adams, P. D. (2005). *CCP4 Newsl. Protein Crystallogr*. **42**, contribution 8.
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.
Beck, T., Krasauskas, A., Gruene, T. & Sheldrick, G. M. (2008). *Acta Cryst.* D**64**, 1179–1182.
Dauter, Z. & Dauter, M. (1999). *J. Mol. Biol.* **289**, 93–101.
Dauter, Z., Dauter, M., de La Fortelle, E., Bricogne, G. & Sheldrick, G. M. (1999). *J. Mol. Biol.* **28**, 83–92.
Dauter, Z., Dauter, M. & Dodson, E. J. (2002). *Acta Cryst.* D**58**, 494–506.
Dauter, Z., Dauter, M. & Rajashankar, K. R. (2000). *Acta Cryst.* D**56**, 232–237.
Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.
Evans, G. & Bricogne, G. (2002). *Acta Cryst.* D**58**, 976–991.
Evans, G. & Bricogne, G. (2003). *Acta Cryst.* D**59**, 1923–1929.
Gunčar, G., Wang, C.-I. A., Forwood, J. K., Teh, T., Catanzariti, A.-M., Ellis, J. G., Dodds, P. N. & Kobe, B. (2007). *Acta Cryst.* F**63**, 209–213.
Hendrickson, W. A. & Teeter, M. M. (1981). *Nature (London)*, **290**, 107–113.
Ismaya, W. T., Rozeboom, H. J., Weijn, A., Mes, J. J., Fusetti, F., Wichers, H. J. & Dijkstra, B. W. (2011). *Biochemistry*, **50**, 5477–5486.
Jakoncic, J., Di Michiel, M., Zhong, Z., Honkimaki, V., Jouanneau, Y. & Stojanoff, V. (2006). *J. Appl. Cryst.* **39**, 831–841.
Korolev, S., Dementieva, I., Sanishvili, R., Minor, W., Otwinowski, Z. & Joachimiak, A. (2001). *Acta Cryst.* D**57**, 1008–1012.
Moiseeva, N. & Allaire, M. (2007). *Acta Cryst.* D**63**, 1025–1028.
Nagem, R. A. P., Dauter, Z. & Polikarpov, I. (2001). *Acta Cryst.* D**57**, 996–1002.
Pape, T. & Schneider, T. R. (2004). *J. Appl. Cryst.* **37**, 843–844.
Ramagopal, U. A., Dauter, M. & Dauter, Z. (2003). *Acta Cryst.* D**59**, 1020–1027.
Sun, P. D., Radaev, S. & Kattah, M. (2002). *Acta Cryst.* D**58**, 1092–1098.
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.* D**65**, 582–601.
Vaney, M. C., Maignan, S., Riès-Kautt, M. & Ducruix, A. (1996). *Acta Cryst.* D**52**, 505–517.
Weiss, M. S. (2001). *J. Appl. Cryst.* **34**, 130–135.
Weiss, M. S., Sicker, T., Djinovic-Carugo, K. & Hilgenfeld, R. (2001). *Acta Cryst.* D**57**, 689–695.
Yogavel, M., Nithya, N., Suzuki, A., Sugiyama, Y., Yamane, T., Velmurugan, D. & Sharma, A. (2010). *Acta Cryst.* D**66**, 1323–1333.