Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Using surface-bound rubidium ions for protein phasing

S. Korolev,[a] I. Dementieva,[a] R. Sanishvili,[a] W. Minor,[b] Z. Otwinowski[c] and A. Joachimiak[a]*

[a]Biosciences Division and Structural Biology Center, Argonne National Laboratory, 9700 South Cass Avenue, Building 202, Argonne, IL 60439, USA, [b]Department of Molecular Physiology and Biological Physics, University of Virginia, Charlottesville, VA 22903, USA, and [c]Department of Biochemistry, UT Southwestern Medical Center at Dallas, Dallas, TX 75235, USA

Correspondence e-mail: andrzejj@anl.gov

Received 15 February 2001
Accepted 2 May 2001

Rubidium is a monovalent metal that can be used as a counterion in protein solutions. X-ray anomalous scattering from rubidium ions bound to the protein surface was used for phasing of the crystal structure of the hsp60 apical domain from *Thermus thermophilus*. Multiple-wavelength anomalous dispersion (MAD) data were collected from a crystal obtained from a solution containing 0.2 *M* rubidium salt. One molecule of protein (147 amino acids) binds one well ordered and one poorly ordered Rb atom. Phases calculated with the program *SHARP* were sufficient for automatic tracing and side-chain assignment using the program *ARP*/*wARP*. The data show that bound rubidium ions can be used to determine protein structures and to study the interaction of monovalent metal ions with proteins and other macromolecules.

1. Introduction

Multi(single)-wavelength anomalous dispersion (MAD/SAD) is the most powerful method for solving novel macromolecular crystal structures (Hendrickson *et al.*, 1985; Hendrickson, 1991; Wang, 1985). It requires a single crystal of the macromolecule with an incorporated anomalous scatterer having an appropriate X-ray absorption edge and a tunable X-ray radiation source available at synchrotron facilities (Walsh, Evans *et al.*, 1999; Hendrickson & Ogata, 1997; Hendrickson, 1999).

Heavy atoms that are traditionally used in the multiple isomorphous replacement method (MIR) with appropriate absorption edges can be used for a MAD experiment (Hendrickson & Ogata, 1997). However, since heavy-atom compounds tend to disturb macromolecule structures and usually decrease the crystals' diffraction power, the search for suitable elements is a very time- and labor-intensive procedure.

In the MAD approach, the most common and successfully used element today is selenium (Hendrickson *et al.*, 1990; Doublie, 1997), incorporated into protein structures by substitution of the naturally occurring amino acid methionine with selenomethionine *in vivo*. Although the anomalous contribution from the Se *K* edge is weak compared with the scattering of the heavy atoms traditionally used in MIR, a well ordered selenium in a crystal structure nevertheless provides a good anomalous diffraction signal which is sufficient for the phasing of a large protein structure (Deacon & Ealick, 1999).

Bromine chemically introduced into a nucleic acid (Gao *et al.*, 1999) or naturally present in a protein, metals such as iron (Hendrickson *et al.*, 1988), zinc (Hosfield *et al.*, 1999), cobalt and copper (Guss *et al.*, 1988; Williams *et al.*, 1996; Walter *et al.*, 1996) are also good candidates for MAD phasing.

The noble gases Xe or Kr can be introduced into some protein crystals by pressurizing them in an atmosphere of the gas (Prange *et al.*, 1998; Schiltz, Shepard *et al.*, 1997). Currently, only Kr has been used for MAD experiments (Cohen *et al.*, 2001), while Xe has been utilized primarily for single-wavelength experiments (Schiltz, Kvick *et al.*, 1997; Malashkevich *et al.*, 1996), as there are major experimental difficulties working at the absorption edges of Xe (Cohen *et al.*, 2001).

Generally, proteins need counterions for efficient folding (Katsumata *et al.*, 1996). Binding of metal cations to proteins effects a large heat-capacity change (Guinto & Di Cera, 1996). Both divalent and monovalent metal cations are often required for protein stability (Henzl *et al.*, 2000) and function (Machius *et al.*, 1998). Binding of monovalent anions and cations to protein surfaces has been demonstrated in a number of high-resolution crystal structures (Lim *et al.*, 1998; Dauter *et al.*, 1999; Berman *et al.*, 2000). This has recently been utilized for MAD phasing by Z. Dauter and coworkers (Dauter & Dauter, 1999; Dauter *et al.*, 2000). Diffusion of Br^- ions into protein crystals provides a quick and easy method of introducing anomalous scatterers into protein crystals. However, it requires a high ($\sim$1 *M*) salt concentration for crystal soaking.

In this paper, we demonstrate that Rb ions can be introduced into protein crystals at a relatively low concentration and used for MAD phasing. In our study, we used the peptide-binding domain of a thermophilic chaperonin as a model system (Walsh, Dementieva *et al.*, 1999; PDB code 1srv). Prior to our experiments there were no data on binding of monovalent metal ions to this protein. Protein crystals obtained from or soaked in solutions containing 0.2 *M* rubidium salt provided an anomalous signal sufficient to phase the protein structure. Therefore, this approach provides a simple method of introducing anomalous scattering atoms into protein crystals.

2. Materials and methods

The expression and purification of the hsp60 apical domain from *T. thermophilus* were similar to those described previously (Walsh, Dementieva *et al.*, 1999). Crystals with typical dimensions 0.8 $\times$ 0.4 $\times$ 0.15 mm were grown at room temperature in 8% PEG 8K, 5% MPD, 0.1 *M* HEPES pH 7.5, 0.2 *M* RbBr in hanging drops using a streak-seeding technique. Crystals were gradually transformed in a cryoprotectant solution consisting of 20% PEG 8000, 20% MPD, 0.2 *M* RbBr, 0.1 *M* HEPES pH 7.5 and flash-frozen in liquid nitrogen. This monoclinic crystal form of the apical domain crystallizes in space group $P2_1$, with unit-cell parameters a = 44.6, b = 73.9, c = 45.3 Å, β = 91.37° and two protein molecules per asymmetric unit, and diffracts to a maximum resolution of 1.1 Å using synchrotron radiation.

MAD data were collected to 1.7 Å resolution at three different wavelengths (peak, inflection point and high-energy remote) near the Rb absorption edge and three wavelengths near the Br *K* absorption edge on the 19ID beamline of the Structural Biology Center at the Advanced Photon Source, Argonne National Laboratory. Absorption edges for each element were experimentally detected with a fluorescent scan from the crystal as described previously (Walsh, Evans *et al.*, 1999). Data were processed with *HKL*2000 (Otwinowski & Minor, 1997) and data statistics are shown in Table 1. There were no detectable changes in unit-cell parameters and crystal mosaicity between all six data sets and there were no indications of crystal X-ray radiation damage. The lower resolution of data near the bromine edge is explained by the experimental setup (longer wavelength with a constant crystal-to-detector distance).

The program *SOLVE* (Terwilliger & Berendzen, 1999) was used to locate anomalous Rb/Br sites. Additional sites were found by analyzing Fourier difference maps and were subsequently confirmed using phases from a refined model (PDB code 1srv) by calculating an anomalous difference Fourier map using the *CCP*4 program suite (Collaborative Computational Project, Number 4, 1994). For phase calculations, we used the program *SHARP* (de La Fortelle & Bricogne, 1997) followed by density modification (Abrahams & Leslie, 1996) as implemented in *SHARP*.

The final phases were used in the program *ARP*/*wARP* (Perrakis *et al.*, 1999; Lamzin & Perrakis, 2000) for automated building of protein structures. Atomic models and electron-density maps were visualized using the program *O* (Jones, 1985; Jones *et al.*, 1991). Refinement was carried out using the program *CNS* (Brunger *et al.*, 1998).

3. Results and discussion

3.1. Determination of crystal structures using Rb anomalous signal

In hsp60 apical domain crystals obtained in the presence of 0.2 *M* RbBr salt, two Rb sites were found by the program *SOLVE*, one for each protein molecule in the asymmetric unit. These sites were geometrically identical and were located on the surface. Difference Fourier maps revealed two additional sites. They were also located at identical places on the surfaces of two molecules in an asymmetric unit, although the peak corresponding to one of these sites had a very low signal level (Fig. 1). *SOLVE* could not find any peaks for data collected near the Br edge and only two weak sites were found in an anomalous difference Fourier map with phases calculated from the refined model. The top ten peaks found by the program *PEAKSEARCH* (Collaborative Computational Project, Number 4, 1994) using the anomalous signal and phases calculated from the final model are shown in Fig. 1. The first two sites were those found by *SOLVE*. The data collected at the Br *K* edge did not provide sufficient phasing information to build a model.

Using three data sets collected near the Rb absorption edge, phases were calculated with the program *SHARP* using four rubidium sites and were subjected to solvent modification. The refined occupancies for the four Rb sites were 0.9, 0.6, 0.5 and 0.1 and the *B* factors were 11, 10, 18 and 12 Å^2, respectively. The different values of the occupancies and temperature

Table 1
Data-collection statistics.

Values for the highest resolution shell are shown in parentheses.

	Rb (30–1.7 Å)			Br (30–1.9 Å)		
Wavelength (Å)	0.8147	0.8151	0.7948	0.9195	0.9197	0.9393
R_{merge} (%)	7.3 (12.7)	7.6 (13.2)	7.5 (14.1)	7.6 (11.9)	7.7 (11.6)	8.6 (14.3)
Completeness (%)	98 (96)	98 (97)	98 (96)	98 (96)	98 (96)	97 (85)
$I/\sigma(I)$	15.1 (4.8)	16.6 (5.4)	16.0 (5.0)	14.1 (5.9)	14.1 (6.0)	18.2 (8.4)
Redundancy	2.5	3	3.2	2.5	2.5	3.6

factors for geometrically identical sites are most likely to result from the different packing environments of each monomer in the asymmetric unit. The centric and acentric figures of merit before solvent flattening were 0.34 and 0.47, respectively. An example of the resulting electron-density map is shown in Fig. 2. Compared with phases obtained from selenomethionine crystals of the same protein (Walsh, Dementieva *et al.*, 1999; data not shown), the anomalous signal of rubidium is roughly 2.5 times weaker, with two Rb sites *versus* two selenium sites per protein molecule. Nevertheless, the phases obtained were sufficient for automatic tracing of the protein structure with the program *ARP/wARP*. We applied a simple protocol provided as an example in the *ARP/wARP* manual. 144 residues out of 147 for both protein molecules in the asymmetric units were traced in 30 cycles (three macrocycles) with a connectivity index of 0.99. There were only two missing residues at the N-terminus and one at the C-terminus. All side chains were built by *ARP/wARP* with 100% confidence.

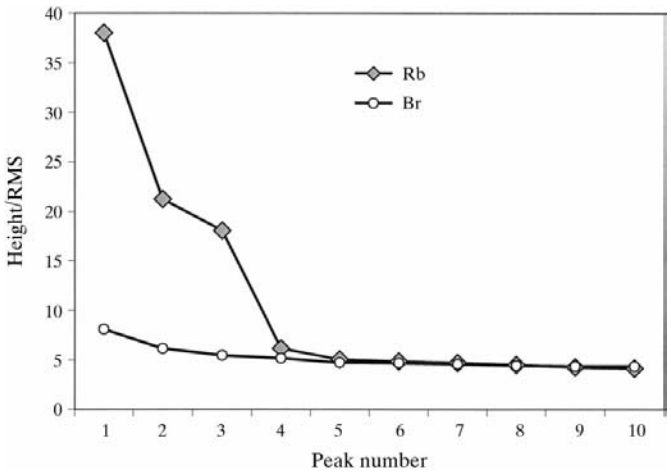


Figure 1
The top ten peaks found by the *PEAKMAX* program in an anomalous difference Fourier map using the anomalous signal from data collected at the Rb f'' peak (shaded diamonds) and at the Br f'' peak (open circles) and phases calculated using the refined atomic model of the apical domain (PDB code 1srv).

The structure was visually inspected using the program *O*. Two N-terminal residues were built manually. Refinement was carried out with the program *CNS*, with a final *R* factor of 19% for all data in the resolution range 20–1.7 Å and R_{free} = 22% (for 5% of data). R.m.s. deviation for bonds was 0.02 Å and for angles was 1.9°, with 96.5% of residues in the core and 3.5% in the allowed regions of the Ramachandran plot. The average *B* factor for all atoms was 12.5 Å^2, with r.m.s. deviations of 1.1 Å^2 for bonded main-chain atoms and 2.7 Å^2 for side-chain atoms.

Structure-determination experiments were also performed with an orthorhombic form of the hsp60 apical domain ($C222_1$) described previously (Walsh, Dementieva *et al.*, 1999), with one protein molecule per asymmetric unit. These crystals diffract X-rays to 1.7 Å using synchrotron radiation. A small crystal grown in 0.25 *M* RbCl was slowly equilibrated against a solution of 0.5 *M* RbCl and a 2.0 Å MAD data set was obtained (data not shown). Essentially the same Rb sites (one strong and one weak) were found in this crystal form, in spite of the 2.5-fold higher concentration of rubidium. Although a considerably weaker diffractor, this crystal provided MAD phases that were sufficient for autotracing 131 of the 147 residues using the program *ARP/wARP* (data not shown).

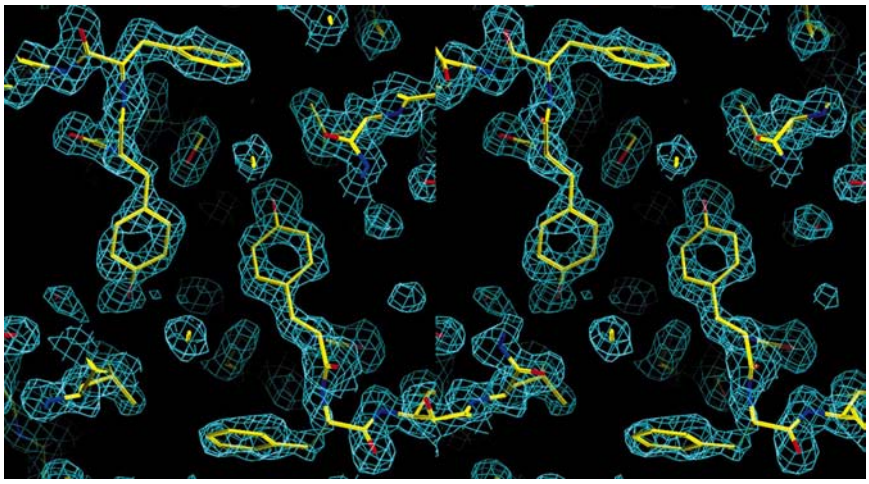


Figure 2
Stereoview of a representative part of the electron-density map at the 1σ level obtained with experimental phases calculated with the program *SHARP* subjected to the solvent-flattening procedure as implemented in *SHARP*.

In addition to cocrystallization of protein with Rb salts, we also tested short (30 s) soaking of hsp60 apical domain crystals obtained from solutions containing NaCl in cryoprotectants with RbCl or RbBr salts (data not shown). Crystals soaked in solutions containing up to 0.5 *M* salt did not show considerable changes in diffraction power. The Rb sites were similar to those described above. The binding of Br^- ions

was still very weak. Short (30 s) crystal soaking in a solution containing 1.0 *M* RbBr or NaBr substantially lowered the quality of the crystals.

3.2. Binding of rubidium ions to the apical domain

As shown in Fig. 3, Rb^+ ions bound to protein surfaces are coordinated mainly by main-chain carbonyl groups. In the major site (Fig. 3*a*), a Rb^+ ion is pentacoordinated and interacts with the main-chain carbonyl groups of Gly297, Ser315 and Leu317 and two water molecules, with an average distance of 3.0 Å. Gly297 is the last residue of α-helix, while residues 314–317 adopt a 3_{10}-helix turn conformation. The second site (Fig. 3*b*) is located close to the first one. Here, the Rb^+ ion is hexacoordinated and interacts with the main-chain carbonyl O atoms of Gly297 and Gly298, the OG1 atom of Thr299 and with a main-chain carbonyl O atom of Gly306 from a second chain (crystal packing contact). The site is completed by two water molecules located within a 3.5 Å distance of rubidium.

The site shown in Fig. 3(*a*) was reproducible in all our experiments, including short soaking of crystals grown in the presence of NaCl, and seems quite specific. Both sites are located on the protein surface, with nested interactions with main-chain carbonyl groups, water molecules and nearby side chains.

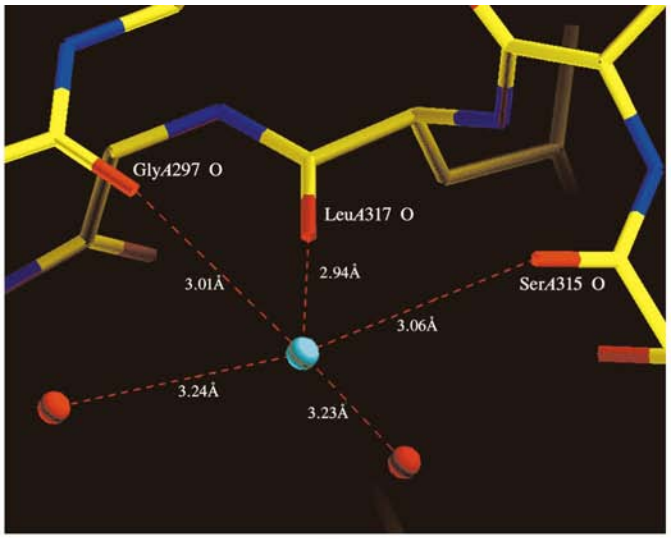
3.3. How general can the use of rubidium for protein phasing be?

Rubidium is a small (ionic radius = 1.49 Å; Sharpe, 1992), stable monovalent ion with an absorption edge at 15 199 eV (0.8157 Å), easily accessible at many synchrotron beamlines. Rubidium anomalous scattering factors are similar to those of selenium and bromine ($f' = -10$, $f'' = 4$ electrons; Cromer, 1983). Therefore, rubidium could potentially serve as a useful source of anomalous signal in the phasing of protein structures. The questions are (i) whether these ions can be introduced under gentle conditions into protein crystals and (ii) whether there will be a sufficient number of ordered rubidium sites that can be used for phasing.

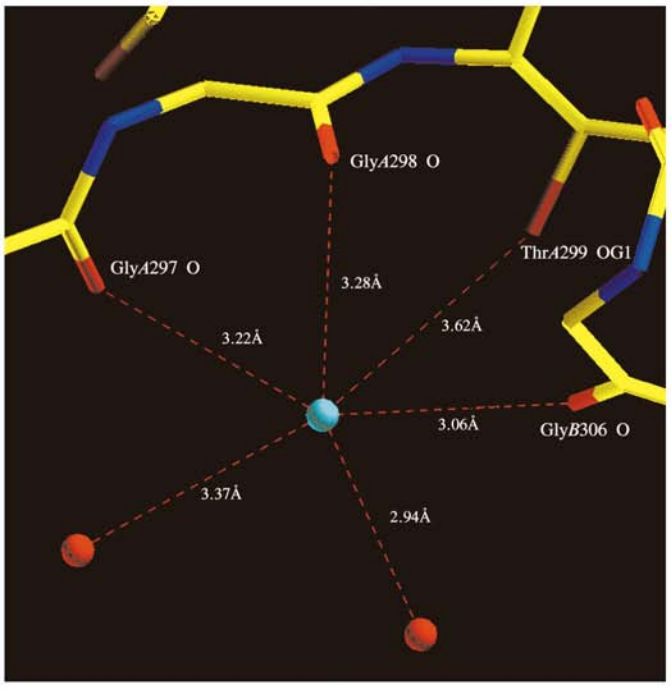
In this work, we demonstrated that the binding of Rb^+ ions to a protein surface can be utilized for the phasing of crystal structures in a similar way to Br^- ions. In our study, the number of Rb-binding sites was independent of ion concentration between 0.2 and 0.5 *M* Rb salt and the ions were well coordinated on the protein surface. Therefore, Rb^+ ions may need appropriate ordered sites on a protein surface to achieve specific coordination geometry, in contrast to the non-specific binding of Br^- ions, which can substitute for water in solvent sites around a protein surface (Dauter & Dauter, 1999).

The advantage of the more specific binding is the lower salt concentration required for crystal soaking. While for Br the salt concentration should be around 1 *M*, in our case 0.2 *M* Rb salt was enough to obtain a high occupancy of bound Rb^+ ions sufficient for phasing. Consequently, soaking protein crystals in solutions with a relatively low Rb salt concentration provides sufficient ion binding for phasing of a protein structure. Lower salt concentration also moderates the background scattering from solvent.

At present, it is not clear how many proteins may have appropriate sites for Rb binding. In the hsp60 apical domain, a protein for which no monovalent metal binding was reported, we found two Rb-binding sites per 147 amino acids. From our experience, we can conclude that one well coordinated Rb^+ ion may provide an anomalous signal strong enough to phase a medium-size protein structure (100–200 amino acids) for crystals diffracting to 2.0 Å resolution.



(*a*)



(*b*)

Figure 3
Binding of Rb^+ ions to the hsp60 apical domain. Rb^+ ions are shown as cyan spheres and water molecules as red spheres. (*a*) The Rb site that corresponds to the stronger anomalous peak. (*b*) The Rb site that corresponds to the weaker anomalous peak.

There have been very few attempts to search the Protein Data Bank (Berman *et al.*, 2000) for potential monovalent metal-binding sites (Nayal & Di Cera, 1996). There are about 300 structures of more than 60 different proteins in the PDB where sodium ions were detected bound to proteins and about 70 structures containing a potassium ion (Kleywegt & Jones, 1998; Berman *et al.*, 2000). Although this is a small percentage of all known protein structures, it does not necessarily mean that monovalent cations are generally not bound to proteins. Indeed, it is feasible that many sodium ions in deposited structures are identified as water molecules, as the electron densities of this cation and water are comparable. They can be distinguished based on coordination geometry only in limited cases of protein structures solved at near-atomic resolution.

Using Rb in protein solutions can provide valuable knowledge about ion–protein interactions. In fact, Rb was used for Na substitution in thrombin crystals (Di Cera *et al.*, 1995) to prove the presence of Na^+ cations in specific Na-binding sites. Collecting data at the Rb absorption edge provides an excellent opportunity to directly locate metal cation-binding sites in proteins and other macromolecular crystals.

We wish to thank all members of the Structural Biology Center at Argonne National Laboratory for their help in conducting experiments and for fruitful discussions. This research was supported by the US Department of Energy, Office of Biological and Environmental Research under contract W-31-109-Eng-38.

References

Abrahams, J. P. & Leslie, A. G. W. (1996). *Acta Cryst.* D**52**, 30–42.
Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N. & Bourne, P. E. (2000). *Nucleic Acids Res.* **28**, 235–242.
Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* D**54**, 905–921.
Cohen, A., Ellis, P., Kresge, N. & Soltis, S. M. (2001). *Acta Cryst.* D**57**, 233–238.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Cromer, D. T. (1983). *J. Appl. Cryst.* **16**, 437–438.
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.* **289**, 83–92.
Dauter, Z., Dauter, M. & Rajashankar, K. R. (2000). *Acta Cryst.* D**56**, 232–237.
Deacon, A. M. & Ealick, S. E. (1999). *Structure Fold. Des.* **7**, R161–R166.
Di Cera, E., Guinto, E. R., Vindigni, A., Dang, Q. D., Ayala, Y. M., Wuyi, M. & Tulinsky, A. (1995). *J. Biol. Chem.* **270**, 22089–22092.
Doublie, S. (1997). *Methods Enzymol.* **276**, 523–530.
Gao, Y. G., Robinson, H., Sanishvili, R., Joachimiak, A. & Wang, A. H. (1999). *Biochemistry*, **38**, 16452–16460.
Guinto, E. R. & Di Cera, E. (1996). *Biochemistry*, **35**, 8800–8804.
Guss, J. M., Merritt, E. A., Phizackerley, R. P., Hedman, B., Murata, M., Hodgson, K. O. & Freeman, H. C. (1988). *Science*, **241**, 806–811.
Hendrickson, W. A. (1991). *Science*, **254**, 51–58.
Hendrickson, W. A. (1999). *J. Synchrotron Rad.* **6**, 845–851.
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.
Hendrickson, W. A., Smith, J. L., Phizackerley, R. P. & Merritt, E. A. (1988). *Proteins*, **4**, 77–88.
Hendrickson, W. A., Smith, J. L. & Sheriff, S. (1985). *Methods Enzymol.* **115**, 41–55.
Henzl, M. T., Larson, J. D. & Agah, S. (2000). *Biochemistry*, **39**, 5859–5867.
Hosfield, D. J., Guan, Y., Haas, B. J., Cunningham, R. P. & Tainer, J. A. (1999). *Cell*, **98**, 397–408.
Jones, T. A. (1985). *Methods Enzymol.* **115**, 157–171.
Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, G. J. (1991). *Acta Cryst.* A**47**, 110–119.
Katsumata, K., Okazaki, A., Tsurupa, G. P. & Kuwajima, K. (1996). *J. Mol. Biol.* **264**, 643–649.
Kleywegt, G. J. & Jones, T. A. (1998). *Acta Cryst.* D**54**, 1119–1131.
La Fortelle, E. de & Bricogne, G. (1997). *Methods Enzymol.* **276**, 472–494.
Lamzin, V. S. & Perrakis, A. (2000). *Nature Struct. Biol.* **7** (Suppl.), 978–981.
Lim, K., Nadarajah, A., Forsythe, E. L. & Pusey, M. L. (1998). *Acta Cryst.* D**54**, 899–904.
Machius, M., Declerck, N., Huber, R. & Wiegand, G. (1998). *Structure*, **6**, 281–292.
Malashkevich, V. N., Kammerer, R. A., Efimov, V. P., Schulthess, T. & Engel, J. (1996). *Science*, **274**, 761–765.
Nayal, M. & Di Cera, E. (1996). *J. Mol. Biol.* **256**, 228–234.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Perrakis, A., Morris, R. & Lamzin, V. S. (1999). *Nature Struct. Biol.* **6**, 458–63.
Prange, T., Schiltz, M., Pernot, L., Colloc'h, N., Longhi, S., Bourguet, W. & Fourme, R. (1998). *Proteins*, **30**, 61–73.
Sharpe, A. G. (1992). *Inorganic Chemistry.* New York: John Wiley.
Schiltz, M., Kvick, A., Svensson, O., Shepard, W., de La Fortelle, E., Prange, T., Kahn, R., Bricogne, G. & Fourme, R. (1997). *J. Synchrotron Rad.* **4**, 287–297.
Schiltz, M., Shepard, W., Fourme, R., Prange, T., de La Fortelle, E. & Bricogne, G. (1997). *Acta Cryst.* D**53**, 78–92.
Terwilliger, T. C. & Berendzen, J. (1999). *Acta Cryst.* D**55**, 849–861.
Walsh, M. A., Dementieva, I., Evans, G., Sanishvili, R. & Joachimiak, A. (1999). *Acta Cryst.* D**55**, 1168–1173.
Walsh, M. A., Evans, G., Sanishvili, R., Dementieva, I. & Joachimiak, A. (1999). *Acta Cryst.* D**55**, 1726–1732.
Walter, R. L., Ealick, S. E., Friedman, A. M., Blake, R. C. N., Proctor, P. & Shoham, M. (1996). *J. Mol. Biol.* **263**, 730–751.
Wang, B. C. (1985). *Methods Enzymol.* **115**, 90–111.
Williams, P. A., Blackburn, N. J., Sanders, D., Bellamy, H., Stura, E. A., Fee, J. A. & McRee, D. E. (1996). *Nature Struct. Biol.* **6**, 509–516.