Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Selling candles in a post-Edison world: phasing with noble gases bound within engineered sites

Michael L. Quillin and Brian W. Matthews*

Institute of Molecular Biology, Howard Hughes Medical Institute and Department of Physics, 1229 University of Oregon, Eugene, OR 97403-1229, USA

Correspondence e-mail: brian@uoxray.uoregon.edu

Received 24 February 2003
Accepted 12 August 2003

The utility of noble gases for phase determination has been limited by the lack of naturally occurring binding sites in proteins. Wild-type T4 lysozyme contains one such binding site. By mutating large hydrophobic residues to alanine, additional noble-gas binding sites have been successfully introduced into this protein. Using data from xenon derivatives of the wild type, two single mutants and the corresponding double mutant, experimental phases for T4 lysozyme have been determined using standard multiple isomorphous replacement (MIR) techniques. These phases, which were obtained from room-temperature data collected on a rotating-anode source, are comparable in quality with phases calculated using selenomethionine-based multiwavelength anomalous dispersion (MAD) methods on frozen crystals at a synchrotron. In addition, this method of introducing noble-gas binding sites near specific residues should provide useful information for determining the register of amino acids within electron-density maps and the positions of molecules within the unit cell.

1. Introduction

Recent improvements in crystallographic methods have revolutionized structural biology. Advances in MAD phasing (Hendrickson & Ogata, 1997; Ogata, 1998), selenomethionine incorporation (Hendrickson *et al.*, 1990; Doublié, 1997), cryocrystallography (Hope, 1988; Rodgers, 1994) and synchrotron sources (Helliwell, 1998; Hendrickson, 2000) have led to structural genomics initiatives inconceivable 20 years ago (Kim, 1998; Burley, 2000). In spite of its usefulness, however, MAD phasing with selenomethionine is not a panacea for protein crystallographers.

To begin with, incorporation of selenomethionine gives decreased levels of expression in prokaryotes owing to growth conditions that call for minimal media (Doublié, 1997). In eukaryotes, although considerable progress has been made, incorporation of selenomethionine remains non-trivial (Chen & Bahl, 1991; Lustbader *et al.*, 1995; Bellizzi *et al.*, 1999; McWhirter *et al.*, 1999). Once expressed, selenomethionine-labelled protein is subject to oxidation, which can affect both protein folding and crystallization. Finally, although experimental limits are constantly being extended (Deacon & Ealick, 1999; Merckel *et al.*, 2002), proteins with a deficiency or excess of methionine remain challenging candidates for MAD phasing.

As is often the case in science, new technologies in protein crystallography have spawned new problems. The freezing process, which minimizes radiation damage during a selenomethionine MAD experiment, can itself be quite destructive. In addition, the need for a tunable source requires that crystals

© 2003 International Union of Crystallography
Printed in Denmark – all rights reserved

be transported to a synchrotron and data be collected in an unfamiliar environment. While on an individual basis none of these problems may be insurmountable, the aim of structural genomics to solve representatives from all protein families requires that even recalcitrant structures be determined in a brief amount of time with as little intervention as possible.

2. Alternatives to selenomethionine-based phasing

What recourse do protein crystallographers have when selenomethionine-based phasing fails? Although the traditional method of heavy-metal soaks is still feasible, finding heavy-metal compounds that bind well enough to perturb native intensities without reducing diffraction quality can be quite time-consuming. Despite recent progress in rapid-soaking techniques (Sun *et al.*, 2002), this method may still be a poor choice for high-throughput phasing.

New techniques for phasing proteins are also being developed. Rapid soaks in high concentrations of negatively (Dauter & Dauter, 1999, 2001; Dauter *et al.*, 2001) or positively (Korolev *et al.*, 2001; Nagem *et al.*, 2001) charged ions that exhibit anomalous scattering have recently been introduced as a generally applicable means of phasing native crystals. While such soaks have a high probability of producing useful sites, the concentrations of salt solutions required for binding may preclude the use of this technique in crystals grown under conditions of low ionic strength.

Another source of phase information, which has been largely overlooked in the past owing to its weak signal, is anomalous scattering from the sulfur atoms that occur naturally in proteins. While this method may be helpful for improving phases from another source (Yang & Pflugrath, 2001), it requires very accurate data and may not be effective for *de novo* phasing unless high-resolution data (Hendrickson & Teeter, 1981; Dauter *et al.*, 1999; Gordon *et al.*, 2001) or soft X-ray sources (Lehmann *et al.*, 1993; Stuhrmann *et al.*, 1997) are available. A recent report on the incorporation of selenotryptophan into dihydrofolate reductase may provide yet another avenue for obtaining phases (Boles *et al.*, 2002).

A technique that holds great promise for phase determination is the use of noble gases (Vitali *et al.*, 1991). One of the advantages of this method is that the conditions required to generate derivatives are extremely gentle owing to the noncovalent character of the protein–ligand interactions. The apparent underutilization of noble gases may in part be because of user inexperience and technical impediments associated with collecting data or freezing crystals under pressure; however, several devices have been developed to improve the ease and reliability of these processes (Schiltz *et al.*, 1994; Stowell *et al.*, 1996; Sauer *et al.*, 1997; Soltis *et al.*, 1997; Djinovic Carugo *et al.*, 1998; Machius *et al.*, 1999). Furthermore, some synchrotrons currently provide equipment for noble-gas pressurization and support experiments at the absorption edges of krypton (Schiltz, Shepard *et al.*, 1997; Cohen *et al.*, 2001) and xenon (Schiltz, Kvick *et al.*, 1997).

Another possible explanation for the small numbers of structures solved using noble-gas derivatives is the lack of suitable binding sites. The likelihood that a native protein will contain a noble-gas binding cavity has been estimated to be as high as 0.5 (Stowell *et al.*, 1996). In practice, however, most structures that have been solved using noble-gas derivatives either bind gases (Vitali *et al.*, 1991; Schiltz *et al.*, 1994; Cohen *et al.*, 2001) or apolar ligands (Bourguet *et al.*, 1995; Weston *et al.*, 1998; Cianci *et al.*, 2001) as part of their biological function or contain fortuitous packing defects at subunit interfaces (Malashkevich *et al.*, 1996; Wang *et al.*, 2001; Mittl *et al.*, 2002). Serine proteases comprise an interesting exception to these classes: in contrast to the apolar environments typically associated with binding sites, noble gases bind tightly within the highly polar catalytic sites of many of these enzymes (Schiltz *et al.*, 1994). Although there are examples of structures phased with noble gases that do not fall into these categories (including, ironically, the subject of this paper), noble-gas phasing of native proteins, while still a valid approach, may not be as generally applicable as initially anticipated.

3. Removing the limitations of noble-gas phasing using protein engineering

Protein engineering has been applied with great success to many problems in crystallography (Price & Nagai, 1995). The familiar bottleneck of obtaining derivatives suitable for phase determination is no exception. To generate novel mercury derivatives for isomorphous replacement, for instance, cysteine residues have been introduced into proteins using site-directed mutagenesis (Sun *et al.*, 1987; Hatfull *et al.*, 1989; Martinez *et al.*, 1993; Nagai *et al.*, 1990). Similarly, constellations of methionine residues found in native proteins have been altered to simplify selenomethionine-based phasing (Leahy *et al.*, 1994; Gassner & Matthews, 1999). In both cases, the resulting derivatives contain heavy atoms that are *covalently* bound to the protein.

In contrast to these methods, the formation of noble-gas derivatives relies upon much weaker *noncovalent* forces. Results from our laboratory have shown that noncovalent interactions with exogenous small molecules can complement deficiencies in protein structure that either occur naturally or have been introduced *via* site-directed mutagenesis. Although the association of polar molecules with sites of mutation on the protein surface has been the subject of some experiments (Baldwin *et al.*, 1998), much greater effort has been directed at the binding of apolar ligands within hydrophobic cavities created by mutations within the protein core (Eriksson, Baase, Wozniak *et al.*, 1992; Morton & Matthews, 1995). These cavity-creating mutations typically involve the replacement of large apolar amino acids (specifically, phenylalanine, methionine, leucine, isoleucine and valine) with alanine.

Among the small molecules that have been analyzed for binding within engineered cavities in T4 lysozyme are the noble gases argon, krypton and xenon (Quillin *et al.*, 2000). A single noble-gas binding site is present in the wild-type protein. The introduction of individual 'large-to-small' mutations typically results in the creation of one or two additional

binding sites. The only known exception is the L121A mutant, which contains a single site located midway between the wild-type site and the site of mutation. The locations of binding sites are highly conserved among the noble gases. In most cases the extent of binding of each noble gas is more dependent upon polarizability than size, with the strongest binding observed for xenon (although a third binding site in the F153A mutant appears to bind krypton better than xenon, suggesting that ligand size can be limiting in smaller cavities).

In a practical extension of this work, data obtained from xenon complexes of the wild type, two single cavity-creating mutants (L99A and F153A) and the corresponding double mutant (L99A/F153A) have been successfully used to determine structure-factor phases for T4 lysozyme (see Quillin & Matthews, 2002 for experimental details). Considering that these data were collected from unfrozen crystals on a rotating-anode source, the quality of the resulting phases is remarkably high compared with the quality of phases derived from a selenomethionine MAD data set collected from a single frozen crystal at a synchrotron. Figures of merit calculated to 1.9 Å before density modification are 0.6 and 0.8 for xenon- and selenomethionine-derived phases, respectively. After solvent flattening, the figures of merit for both sets of phases increased to 0.9. To gauge the accuracy of these phases, experimentally phased maps were compared with maps phased using refined coordinates (Fig. 1). Real-space correlation coefficients calculated to 1.9 Å are 0.53 for xenon and 0.61 for selenomethionine before solvent flattening, and 0.80 for xenon and 0.74 for selenomethionine after solvent flattening, again suggesting that xenon MIRAS is a viable alternative to selenomethionine MAD. Given the success of xenon MIRAS, MAD experiments using krypton bound within engineered cavities may also prove to be a useful substitute for selenomethionine MAD phasing, although synchrotron data

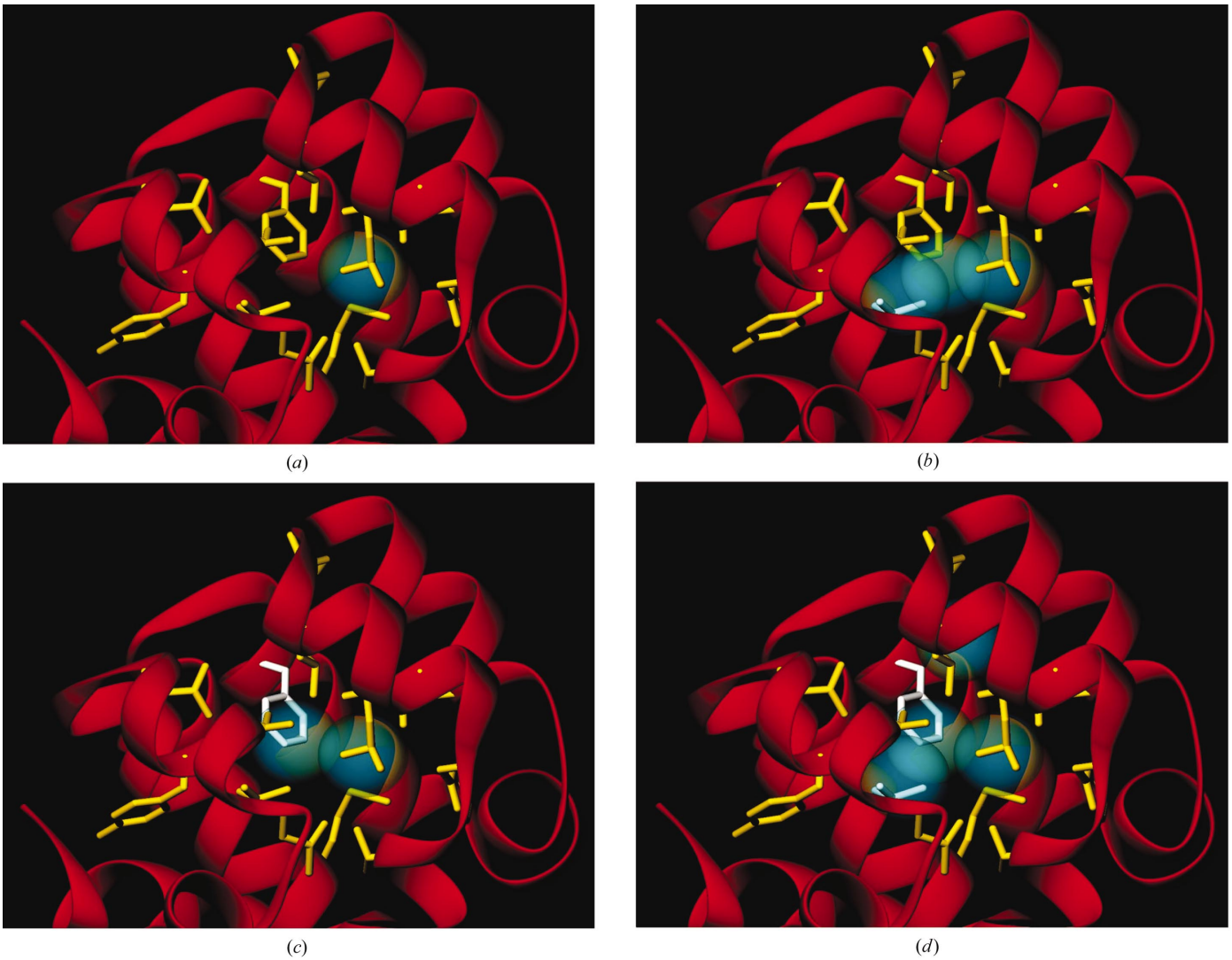

Figure 1
Noble-gas binding sites present in (*a*) wild-type, (*b*) L99A, (*c*) F153A and (*d*) L99A/F153A lysozyme. Refined positions of bound xenon atoms are shown as transparent cyan spheres. Wild-type coordinates for residues surrounding the binding sites are also displayed, with mutated residues shown in white and non-mutated residues in yellow. This figure was generated using *RIBBONS* (Carson, 1991) and *POV-Ray* (http://www.povray.org).

Table 1
Summary of 'large-to-small' mutations in T4 lysozyme (adapted from Xu *et al.*, 1998).

			No. of mutations that		
Amino acid	Total No.	No. buried	Crystallize isomorphously	Form cavities or crevices	Form apolar cavities
Val	9	5	4	4	3
Ile	10	8	6	5	5
Leu	16	10	6	6	6
Met	5	2	2	2	1
Phe	5	3	3	3	3
Total	45	28	21	20	18

collection on frozen crystals would again become necessary. In addition, the gas pressure required for full krypton occupancy is significantly higher than for xenon (Schiltz, Shepard *et al.*, 1997; Quillin *et al.*, 2000; Cohen *et al.*, 2001).

4. Potential pitfalls of engineered noble-gas binding sites

One of the caveats of this method is that in the absence of a known structure it is unclear which residues to mutate in order to generate a noble-gas binding site. Although large hydrophobic residues tend to be buried (Rose *et al.*, 1985), our experience with T4 lysozyme suggests that not all 'large-to-small' mutations generate cavities suitable for noble-gas binding (Table 1). Although the sample size is not sufficient to justify strong conclusions, the percentage of buried hydrophobic residues in T4 lysozyme that form apolar cavities when mutated to alanine range from 50% for methionine to 100% for phenylalanine. Since our studies have concentrated on sites that are known to be buried, we must extrapolate these ratios to account for the total number of residues. Making the conservative assumption that only completely buried residues will produce noble-gas binding sites gives overall success rates of from 20% for methionine to 60% for phenylalanine, with an average value of 30%. The utility of this method for high-throughput phasing is enhanced by the realisation that 86% of mutant proteins that crystallize isomorphously with wild-type lysozyme contain apolar cavities, facilitating the screening process.

A second shortcoming that must be taken into account when using this technique is the energetic cost of cavity formation. Leucine-to-alanine substitutions in T4 lysozyme result in decreases in stability ranging from 11.3 to 20.9 kJ mol^{-1} (Eriksson, Baase, Zhang *et al.*, 1992; Xu *et al.*, 1998). While this destabilization is not insignificant, it may be less detrimental in proteins with larger hydrophobic cores and may have less of an effect on expression levels than the requirement for minimal media during selenomethionine incorporation. While this hypothesis has not been tested directly, it may be advantageous to store protein and grow crystals in the presence of xenon to reclaim some of the stability lost to cavity formation.

5. Other applications of engineered noble-gas sites

In addition to facilitating initial phase determination, the ability to introduce cavities that bind noble gases at predefined locations may prove invaluable during later stages of structure determination. Noble-gas binding sites created by 'large-to-small' substitutions correspond closely to the positions of the mutated side chains in the wild-type protein. In the L99A protein, xenon-binding sites are located 1.2 and 1.5 Å away from the positions of the terminal methyl groups of the Leu99 side chain in wild-type T4 lysozyme. Similarly, in the F153A protein, a xenon-binding site is located 0.6 Å away from the *para*-carbon of Phe153 in the wild-type protein. These binding sites may serve as useful guides during model building when the sequence register is ambiguous or the fold of the protein is unclear.

In cases of molecular replacement in which there are many copies per unit cell, it may prove beneficial to create engineered cavities at homologous sites in both crystal forms. Comparison of the positions of bound noble gases may sufficiently restrict the search so that a solution can be found. A similar strategy may be applied to difficult cases of non-crystallographic symmetry.

6. Concluding remarks

The current explosion in structural genomics initiatives can be largely attributed to the development of selenomethionine MAD phasing. When this method cannot be used, however, the probability of automated structure determination decreases substantially. Using protein engineering, we have extended the utility of noble-gas phasing to proteins that do not contain a noble-gas binding site. The addition of engineered noble-gas binding sites to the repertoire of methods for *de novo* phase determination should dramatically improve the prospects of structural genomics.

We thank the organizers of the CCP4 Study Weekend 2003 for the opportunity to present this work. We also thank Thomas Earnest and the staff of the Macromolecular Crystallography Facility at the Advanced Light Source for help with synchrotron data collection. MLQ acknowledges the support of a fellowship from the Helen Hay Whitney foundation. This work was also supported in part by NIH grant GM20066 to BWM.

References

Baldwin, E., Baase, W. A., Zhang, X., Feher, V. & Matthews, B. W. (1998). *J. Mol. Biol.* **277**, 467–485.
Bellizzi, J. J., Widom, J., Kemp, C. W. & Clardy, J. (1999). *Structure*, **7**, R263–R267.
Boles, J. O., Henderson, J., Hatch, D. & Silks, L. A. (2002). *Biochem. Biophys. Res. Commun.* **298**, 257–261.
Bourguet, W., Ruff, M., Chambon, P., Gronemeyer, H. & Moras, D. (1995). *Nature (London)*, **375**, 377–382.
Burley, S. K. (2000). *Nature Struct. Biol.* **7**, 932–934.
Carson, M. (1991). *J. Appl. Cryst.* **24**, 958–961.
Chen, W. & Bahl, O. P. (1991). *J. Biol. Chem.* **266**, 8192–8197.

Cianci, M., Rizkallah, P. J., Olczak, A., Raftery, J., Chayen, N. E., Zagalsky, P. F. & Helliwell, J. R. (2001). *Acta Cryst.* D**57**, 1219–1229.
Cohen, A., Ellis, P., Kresge, N. & Soltis, S. M. (2001). *Acta Cryst.* D**57**, 233–238.
Dauter, Z. & Dauter, M. (1999). *J. Mol. Biol.* **28**, 93–101.
Dauter, Z. & Dauter, M. (2001). *Structure*, **9**, R21–R26.
Dauter, Z., Dauter, M., de La Fortelle, E., Bricogne, G. & Sheldrick, G. M. (1999). *J. Mol Biol.* **289**, 83–92.
Dauter, Z., Li, M. & Wlodawer, A. (2001). *Acta Cryst.* D**57**, 239–249.
Deacon, A. M. & Ealick, S. E. (1999). *Structure*, **7**, R161–R166.
Djinovic Carugo, K., Everitt, P. & Tucker, P. A. (1998). *J. Appl. Cryst.* **31**, 812–814.
Doublié, S. (1997). *Methods Enzymol.* **276**, 523–530.
Eriksson, A. E., Baase, W. A., Wozniak, J. A. & Matthews, B. W. (1992). *Nature (London)*, **355**, 371–373.
Eriksson, A. E., Baase, W. A., Zhang, X. J., Heinz, D. W., Blaber, M., Baldwin, E. P. & Matthews, B. W. (1992). *Science*, **255**, 178–183.
Gassner, N. C. & Matthews, B. W. (1999). *Acta Cryst.* D**55**, 1967–1970.
Gordon, E. J., Leonard, G. A., McSweeney, S. & Zagalsky, P. F. (2001). *Acta Cryst.* D**57**, 1230–1237.
Hatfull, G. F., Sanderson, M. R., Freemont, P. S., Raccuia, P. R., Grindley, N. D. & Steitz, T. A. (1989). *J. Mol. Biol.* **208**, 661–667.
Helliwell, J. R. (1998). *Nature Struct. Biol.* **5**, 614–617.
Hendrickson, W. A. (2000). *Trends Biochem. Sci.* **25**, 637–643.
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. & Teeter, M. M. (1981). *Nature (London)*, **290**, 107–113.
Hope, H. (1988). *Acta Cryst.* B**44**, 22–26.
Kim, S.-H. (1998). *Nature Struct. Biol.* **5**, 643–645.
Korolev, S., Dementieva, I., Sanishvili, R., Minor, W., Otwinowski, Z. & Joachimiak, A. (2001). *Acta Cryst.* D**57**, 1008–1012.
Leahy, D. J., Erickson, H. P., Aukhil, I., Joshi, P. & Hendrickson, W. A. (1994). *Proteins*, **19**, 48–54.
Lehmann, M. S., Muller, H.-H. & Stuhrmann, H. B. (1993). *Acta Cryst.* D**49**, 308–310.
Lustbader, J. W., Wu, H., Birken, S., Pollak, S., Gawinowicz Kolks, M. A., Pound, A. M., Austen, D., Hendrickson, W. A. & Canfield, R. E. (1995). *Endocrinology*, **136**, 640–650.
Machius, M., Henry, L., Palnitkar, M. & Deisenhofer, J. (1999). *Proc. Natl Acad. Sci. USA*, **96**, 11717–11722.
McWhirter, S. M., Pullen, S. S., Holton, J. M., Crute, J. J., Kehry, M. R. & Alber, T. (1999). *Proc. Natl Acad. Sci. USA*, **96**, 8408–8413.
Malashkevich, V. N., Kammerer, R. A., Efimov, V. P., Schulthess, T. & Engel, J. (1996). *Science*, **274**, 761–765.
Martinez, C., de Geus, P., Stanssens, P., Lauwereys, M. & Cambillau, C. (1993). *Protein Eng.* **6**, 157–165.
Merckel, M. C., Kajander, T., Deacon, A. M., Thompson, A., Grossman, J. G., Kalkkinen, N. & Goldman, A. (2002). *Acta Cryst.* D**58**, 727–734.
Mittl, P. R. E., Fritz, G., Sargent, D. F., Richmond, T. J., Heizmann, C. W & Grütter, M. G. (2002). *Acta Cryst.* D**58**, 1255–1261.
Morton, A. & Matthews, B. W. (1995). *Biochemistry*, **34**, 8576–8588.
Nagai, K., Oubridge, C., Jessen, T. H., Li, J. & Evans, P. R. (1990). *Nature (London)*, **348**, 515–520.
Nagem, R. A. P., Dauter, Z. & Polikarpov, I. (2001). *Acta Cryst.* D**57**, 996–1002.
Ogata, C. M. (1998). *Nature Struct. Biol.* **5**, 638–640.
Price, S. R. & Nagai, K. (1995). *Curr. Opin. Biotechnol.* **6**, 425–430.
Quillin, M. L., Breyer,W. A., Griswold, I. J. & Matthews, B. W. (2000). *J. Mol. Biol.* **302**, 955–977.
Quillin, M. L. & Matthews, B. W. (2002). *Acta Cryst.* D**58**, 97–103.
Rodgers, D. W. (1994). *Structure*, **2**, 1135–1140.
Rose, G. D., Geselowitz, A. R., Lesser, G. J., Lee, R. H. & Zehfus, M. H. (1985). *Science*, **229**, 834–838.
Sauer, O., Schmidt, A. & Kratky, C. (1997). *J. Appl. Cryst.* **30**, 476–486.
Schiltz, M., Kvick, A., Svensson, O. S., Shepard,W., de La Fortelle, E., Prangé, T., Kahn, R., Bricogne, G. & Fourme, R. (1997). *J. Synchrotron Rad.* **4**, 287–297.
Schiltz, M., Prangé, T. & Fourme, R. (1994). *J. Appl. Cryst.* **27**, 950–960.
Schiltz, M., Shepard, W., Fourme, R., Prangé, T., de La Fortelle, E. & Bricogne, G. (1997). *Acta Cryst.* D**53**, 78–92.
Soltis, S. M., Stowell, M. H. B., Wiener, M. C., Phillips, G. N. Jr & Rees, D. C. (1997). *J. Appl. Cryst.* **30**, 190–194.
Stowell, M. H. B., Soltis, S. M., Kisker, C., Peters, J. W., Schindelin, H., Rees, D. C., Cascio, D., Beamer, L., Hart, P. J., Wiener, M. C. & Whitby, F. G. (1996). *J. Appl. Cryst.* **29**, 608–613.
Stuhrmann, S., Bartels, K. S., Braunwarth, W., Doose, R., Dauvergne, F., Gabriel, A., Knochel, A., Marmotti, M., Stuhrmann, H. B., Trame, C. & Lehmann, M. S. (1997). *J. Synchrotron Rad.* **4**, 298–310.
Sun, D. P., Alber, T., Bell, J. A., Weaver, L. H. & Matthews, B. W. (1987). *Protein Eng.* **1**, 115–123.
Sun, P. D., Radaev, S. & Kattah, M. (2002). *Acta Cryst.* D**58**, 1092–1098.
Vitali, J., Robbins, A. H., Almo, S. C. & Tilton, R. F. (1991). *J. Appl.Cryst.* **24**, 931–935.
Wang, W. C., Hsu, W. H., Chien, F. T. & Chen, C. Y. (2001). *J. Mol. Biol.* **306**, 251–261.
Weston, S. A., Camble, R., Colls, J., Rosenbrock, G., Taylor, I., Egerton, M., Tucker, A. D., Tunnicliffe, A., Mistry, A., Mancia, F., de La Fortelle, E., Irwin, J., Bricogne, G. & Paupit, R. A. (1998). *Nature Struct. Biol.* **5**, 213–221.
Xu, J., Baase, W. A., Baldwin, E. & Matthews, B. W. (1998). *Protein Sci.* **7**, 158–177.
Yang, C. & Pflugrath, J. W. (2001). *Acta Cryst.* D**57**, 1480–1490.