Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

A preliminary time-of-flight neutron diffraction study of *Streptomyces rubiginosus* D-xylose isomerase

B. Leif Hanson,[a] Paul Langan,[b] Amy K. Katz,[c] Xinmin Li,[b] Joel M. Harp,[a] Jenny P. Glusker,[c] Benno P. Schoenborn[b] and Gerard J. Bunick[d]*

[a]University of Tennessee/Oak Ridge National Laboratory Graduate School of Genome Science and Technology, Oak Ridge, TN 37831, USA, [b]Bioscience Division, Los Alamos National Laboratory, Los Alamos, NM 87545, USA, [c]Fox Chase Cancer Center, Philadelphia, PA 19111, USA, and [d]Life Sciences Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA

Correspondence e-mail: gjbunick@utk.edu

Received 14 August 2003
Accepted 10 November 2003

The metalloenzyme D-xylose isomerase forms well ordered crystals that diffract X-rays to ultrahigh resolution (<1 Å). However, structural analysis using X-ray diffraction data has as yet been unable to differentiate between several postulated mechanisms that describe the catalytic activity of this enzyme. Neutrons, with their greater scattering sensitivity to H atoms, could help to resolve this by determining the protonation states within the active site of the enzyme. As the first step in the process of investigating the mechanism of action of D-xylose isomerase from *Streptomyces rubiginosus* using neutron diffraction, data to better than 2.0 Å were measured from the unliganded protein at the Los Alamos Neutron Science Center Protein Crystallography Station. Measurement of these neutron diffraction data represents several milestones: this is one of the largest biological molecules (a tetramer, MW $\simeq$ 160 000 Da, with unit-cell lengths around 100 Å) ever studied at high resolution using neutron diffraction. It is also one of the first proteins to be studied using time-of-flight techniques. The success of the initial diffraction experiments with D-xylose isomerase demonstrate the power of spallation neutrons for protein crystallography and should provide further impetus for neutron diffraction studies of biologically active and significant proteins. Further data will be measured from the enzyme with bound substrates and inhibitors in order to provide the specific information needed to clarify the catalytic mechanism of this enzyme.

1. Introduction

Many enzyme mechanisms in biological systems use hydrogen ions, often derived from water molecules, as part of their catalytic activity. Although X-ray diffraction can often provide atomic resolution structural data on these enzymes, our inability to place H atoms unambiguously in these structures has caused problems in the elucidation of information on enzyme mechanisms using this method. Neutron diffraction, however, is unparalleled for imaging the H-atom locations in a protein, providing significantly less ambiguity than is found with X-ray diffraction data. It can locate H-atom positions at 2.0 Å resolution, compared with the 0.9 Å resolution or better necessary for X-ray structure determinations. Neutron diffraction has been useful in providing support for mechanistic studies of enzyme action in, for example, the serine proteases (Kossiakoff & Spencer, 1980), lysozyme (Mason *et al.*, 1984; Niimura *et al.*, 1997; Bon *et al.*, 1999), ribonuclease (Wlodawer *et al.*, 1983; Borah *et al.*, 1985), and more recently endothiapepsin, an aspartic proteinase (Coates *et al.*, 2001). Crystals of the enzyme D-xylose isomerase (XI; EC 5.3.1.5), for which the structure of the *Streptomyces rubiginosus*

enzyme was first determined in our laboratory in Philadelphia (Carrell *et al.*, 1984, 1989), grow very large, a precondition for successful neutron diffraction experiments. It was thought that neutron diffraction might provide some of the information needed to elucidate the enzymatic mechanism of XI, despite the fact that this enzyme has a large molecular weight and unit cell, greater than that studied by neutron diffraction to date.

The enzyme XI catalyzes the interconversion of D-xylose to D-xyulose and D-glucose to D-fructose by transferring an H atom between C1 and C2 of the substrate. It binds two metal ions, approximately 4.9 Å apart. Magnesium is the metal ion normally found in this enzyme, but divalent manganese or cobalt can also bind to give a catalytically active enzyme. One metal-binding site (M1) contains four carboxylate groups (Glu181, Glu217, Asp245 and Asp287) and two water molecules which bind magnesium; this is the site to which the substrate binds, displacing the two metal-bound water molecules. The other metal-binding site (M2) binds magnesium, manganese or cobalt; this site contains three carboxylate groups (Glu217, bidentate Asp255 and Asp257), one water molecule and a histidine residue (His220) bound to the metal ion. The carboxylate group of Glu217 is shared by both metal ions. At the other end of the active site lies His54, which is held in a rigid orientation by three hydrogen bonds to Asp57. This histidine acts as a base that is presumed to assist in opening the sugar ring system (Smart *et al.*, 1992).

The overall action of the enzyme involves transfer of an H atom from one C atom of the substrate to an adjacent C atom. Three mechanisms have been proposed for this hydrogen transfer (Fig. 1). One involves a base-catalyzed proton transfer with a *cis*-enediol intermediate (Wohl & Neuberg, 1900; Rose *et al.*, 1969). The other two involve a simple hydride-shift mechanism (Ramachander & Feather, 1977; Farber *et al.*, 1989; Collyer *et al.*, 1990; Collyer & Blow, 1990) or a hydride shift mediated by a metal ion (Whitlow *et al.*, 1991). The metal-ion-assisted hydride-shift mechanism implies some movement of the metal ion M2 and, indeed, a second site for this metal ion has been observed in some crystal structure determinations (Lavie *et al.*, 1994).

Crystals of this protein from *S. rubiginosus* diffract well: 1.6 Å X-ray diffraction data have been measured at room temperature and the resulting crystal structure has been reported (Carrell *et al.*, 1989). Several other crystal structures of the enzyme from this and other sources have since been reported (Henrick *et al.*, 1989; Farber *et al.*, 1987; Rey *et al.*, 1988; Whitlow & Howard, 1989), including structures at 0.99 Å (PDB code 1mnz; E. Nowak, S. Panjikar & P. A. Tucker, unpublished work) and 0.86 Å resolution (PDB code 1muw; T. D. Fenn, D. Ringe & G. Petsko, unpublished work). Additionally, a liquid-nitrogen-cooled data set was collected by us at IMCA-CAT at APS to 0.94 Å resolution from the *S. rubiginosus* enzyme (Hanson *et al.*, 2002). However, ultrahigh-resolution X-ray data have not yet provided the information sought for an understanding of the catalytic mechanism. The binding of substrates and inhibitors to XI has also been studied and several different binding modes have been identified (Carrell *et al.*, 1994).

Until now, flux limitations and detector arrangements have hindered the use of neutron diffraction in protein crystallography. Experiments have required large samples ($\sim$5 mm^3) and extended data-collection periods (weeks to months). The use of Laue-diffraction geometry can be used to overcome these limitations (Wilkinson & Lehmann, 1991). Laue-diffraction geometry gives a rapid and efficient survey of reciprocal space since it maximizes the flux at the sample. It also generates large numbers of reflections over all incident wavelengths, enabling the measurement of neutron data from proteins with higher molecular weights and/or smaller crystals (Arzt *et al.*, 1999). Some disadvantages of Laue diffraction at steady-state reactor neutron sources are the overlap of reflections on the detector and a large build-up of incoherent background scattering which can severely reduce the signal-to-noise ratio. In order to minimize these problems, $\delta\lambda/\lambda$ is typically limited to 20% or less. This quasi-Laue technique (Schoenborn, 1992) has been used effectively at ILL with the LADI detector (Myles *et al.*, 1997). Laue-diffraction geometry, however, will find its most powerful expression at spallation neutron sources, where the time of flight of the particle permits a determination of the wavelength. With the wavelength-resolved Laue technique it is found that reflections and background scattering are resolved in wavelength, so that all available neutron wavelengths can be used without reflection overlap and background accumulation.

The first dedicated protein crystallography station (PCS) at a spallation neutron source, at the Los Alamos Neutron

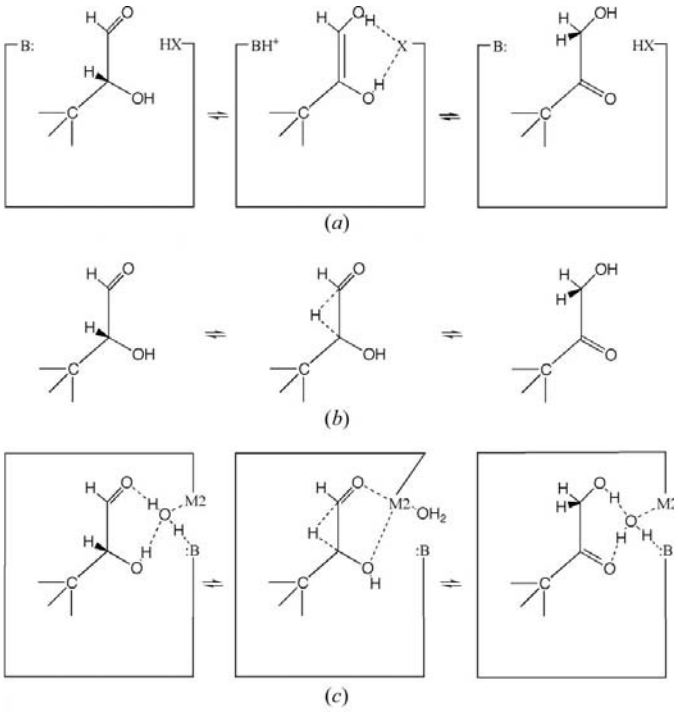


Figure 1
Diagrammatic representations of the proposed mechanisms described in the text for D-xylose isomerase catalytic activity. (*a*) Base-catalyzed proton transfer *cis*-enediol intermediate. (*b*) Hydride-shift mechanism. (*c*) Hydride shift mediated by metal ions. B = base, X = hydrogen-donating group, M2 = a specific metal ion, that which binds His220.

Scattering Center (LANSCE), provides a pulse of neutrons from the impact of protons on a tungsten target. These neutrons have a broad wavelength range and the pulses occur 20 times a second. Use of a partially coupled water moderator and a large curved detector, as well as improved neutron flux from the LANSCE target, means that the PCS permits faster data collection from smaller crystals with larger unit cells. This recently commissioned PCS provides us with a means for testing our hypothesis that Laue-diffraction geometry with time-of-flight neutrons can be used to study larger macromolecules than previously possible. Crystals of XI, a bimetallic enzyme, provide an excellent choice for such a study.

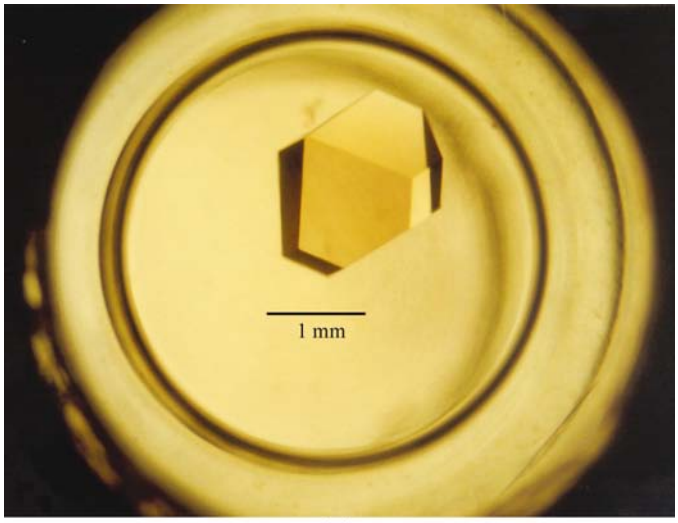
2. Methods

The enzyme XI used for crystallization was purified from a stock solution of commercial food-grade XI (Gensweet SGI) kindly provided by Genencor International. 4 ml of stock material was mixed with an equal volume of saturated ammonium sulfate solution and allowed to sit overnight at 277 K. The sample was then centrifuged at 8000*g* for 30 min. The supernatant was discarded and the pellet was resuspended in 4 ml of 25% saturated $(NH_4)_2SO_4$. After incubating at 277 K for 4 h, the sample was centrifuged at 8000*g* for 30 min. After centrifugation, the supernatant was again discarded and the pellet was resuspended in 4 ml of 25% saturated $(NH_4)_2SO_4$ and allowed to incubate overnight at 277 K. The sample was centrifuged a final time at 8000*g* for 30 min and the supernatant was discarded.

Since the pH optimum for this enzyme is 7.5–9.0 (Callens *et al.*, 1986), the samples for X-ray studies were maintained in this range. The pelleted sample was redissolved in a buffer consisting of 50 m*M* Tris pH 8.0, 10 m*M* EDTA, placed in dialysis tubing with an 8000 Da cutoff (Spectrum Industries) and dialyzed against 2 l of 50 m*M* Tris pH 8.0 containing 10 m*M* EDTA at 277 K. After three changes of buffer over a 2 d period, the sample was dialyzed against 100 m*M* Tris pH 8.0 and 2 m*M* $MnCl_2$ at 277 K with three changes of buffer over a 2 d period. The protein was then concentrated to 125 mg ml^{-1} using a 3000 MW CentriPrep (Millipore) and the final volume of protein spun at 50 000*g* in an airfuge for 10 min.

The growth of large protein crystals was enhanced by the use of the Diffusion Controlled Apparatus for Microgravity (DCAM) and the Counter-Diffusion Cell (CDC; Fig. 2*b*; Carter *et al.*, 1999). This crystallization apparatus also facilitated the exchange of deuterated reagents after the crystals had formed. 45 µl of protein was placed in the intrinsic button of the cap for the small chamber and then covered with a single sheet of dialysis membrane (MW cutoff = 8000 Da). The membrane was secured in place with a silicon rubber O-ring. The small chamber of the CDC (2.0 ml) was filled with 5% saturated $(NH_4)_2SO_4$ in stock buffer (100 m*M* Tris pH 8.0, 2 m*M* $MgCl_2$). The large chamber (10.5 ml) was filled with 45% saturated $(NH_4)_2SO_4$ in stock buffer, resulting in an equilibrium concentration of 38.6% $(NH_4)_2SO_4$. An agarose plug positioned between the two chambers of the CDC prevents free mixing of the two solutions and controls the rate of exchange between the two chambers.

Large single crystals of XI were obtained in the CDC after several months. Their size range was 1.5–3.5 mm on a side and they were suitable for neutron diffraction (Fig. 2*a*). The exchange with deuterated reagents was initiated by first exchanging the large chamber with 100 m*M* D_4Tris pH 7.5 (pD 7.9), 2 m*M* $MnCl_2$, 2 m*M* $CoCl_2$, 40% saturated $(ND_4)_2SO_4$. After 10 d, the contents of the large chamber were removed and reserved. A fresh aliquot of deuterated crystallization buffer was then added to the chamber. The solution in the small chamber was then extracted and reserved buffer was added to that chamber. The swapping and replacement of



(*a*)

(*b*)

Figure 2
The growth of single large crystals for neutron diffraction made use of hardware developed, manufactured and marketed by Crystal Systems Inc. under an exclusive license by NASA. (*a*). A low-power microscopic view of a single XI crystal is shown in the protein receptacle of a CDC. (*b*) The 45 µl protein receptacle is molded into the end cap visible in the right foreground of the CDC components. See Carter *et al.* (1999) for more detail on this device. (Counter-Diffusion Cell is a trademark of Crystal Systems Inc.)

Table 1
Data-collection parameters for D-xylose isomerase.

Space group	*I*222
Unit-cell parameters† (Å, °)	$a = 93.78$, $b = 99.53$, $c = 102.90$, $\alpha = \beta = \gamma = 90$
Wavelength range (Å)	0.6–5
Station	PCS
Temperature (K)	295
No. of crystal settings	23
Average time per setting (h)	20
Average No. of T_0 values per setting	1440000
No. observed reflections	181797
No. observed reflections to 1.8 Å resolution	146996
No. unique reflections to 1.8 Å resolution	38029

† Unit-cell parameters are from X-ray diffraction.

buffers occurred four times in this fashion over a 4–5 week period. The crystals were regularly monitored for any signs of stress (cracks, dissolution *etc*.). Diffusion between the chambers of the CDC is controlled by the cross-section and length of the agarose plug. Studies monitoring pH or colorimetric changes show that it can take weeks to months to attain equilibrium between the chambers. The exchange procedure was crafted to minimize potential stresses (osmotic, pH, variations in precipitant concentration) on the crystals and to complete the process prior to the scheduled start date at LANSCE.

The crystals were then extracted from the CDC and placed in quartz capillaries for data collection. Special precautions were taken to minimize back-exchange of hydrogen during crystal mounting. A plastic shroud was constructed around the microscope used for mounting crystals. The atmosphere inside the shroud was nitrogen that had been bubbled through a side-arm flask holding D_2O. Because of the size of the crystal, it did not fully descend into the capillary, but became lodged before the taper had reached a diameter of 2 mm. The capillary, with the crystal in it, was then mounted in modeling clay atop a goniometer head and centered for data collection.

3. Data collection

Neutron diffraction data were measured on the PCS (beamline 15) at LANSCE, as summarized in Table 1. This PCS was originally described by Schoenborn & Pitcher (1996) and recently in more detail by Langan *et al.* (2004). A κ-circle goniometer was used to step the crystal through 23 different orientations, measuring a wavelength-resolved Laue pattern at each setting over the useful wavelength range of 0.7–7 Å. The strategy used was to collect data from an octant of reciprocal space by scanning the crystal around its spindle axis in four steps of 30° in φ. The detector was positioned in the straight-through orientation so that it subtended ±60° in its horizontal curved 2θ plane and ±8° in its vertical flat direction. The detector was not moved between crystal settings.

The sample-to-detector distance of 70 cm corresponds to the cylindrical radius of the detector. At this position, the minimum *d*-spacings ($d_{\min}$, where *d* is the resolution) that can be collected on the detector match the corresponding wavelengths (*i.e.* at $\lambda = 0.7$ Å, $d_{\min} = 0.7$ Å). However, the possible resolution of the data is restricted by the time-resolution of the PCS. Using the formulation of Jauch & Dachs (1994) and the measured time-dispersion of the PCS neutron beam reported by Langan *et al.* (2004), the reflections for this particular unit cell remain resolved in time at $\lambda = 0.7$ Å only out to *d*-spacings of about 0.9 Å. Because of the limited coverage of the detector in the vertical direction, the crystal was reoriented five times using the κ and ω circles of the goniometer, with φ scans being collected at each orientation.

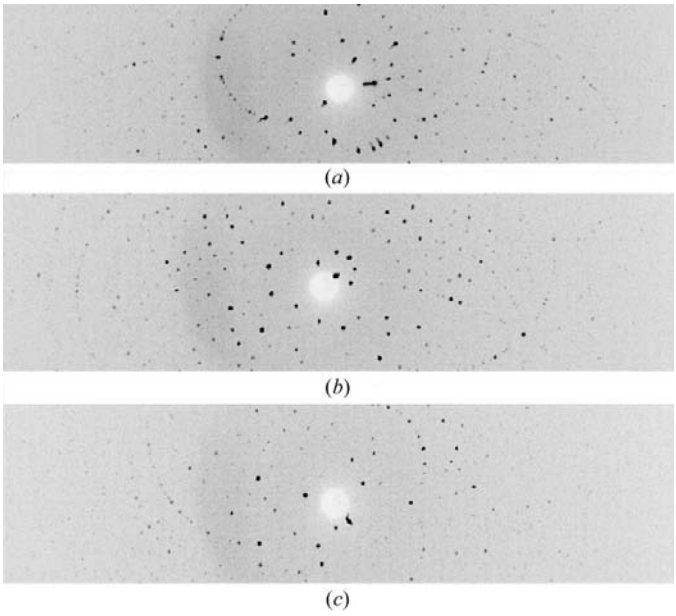


Figure 3
Neutron diffraction data from three different crystal settings for the native XI crystal. Each image is an overlay of all time intervals for a setting; as a consequence, weaker reflections from longer wavelength diffraction can be lost in the background from the short-wavelength time slices. However, since the data from each time slice is analyzed separately, these reflections can be measured against the background within their time slice. This reduces the time needed to measure the signal for each reflection and improves the signal-to-noise ratio for the weak data (Langan *et al.*, 2004).

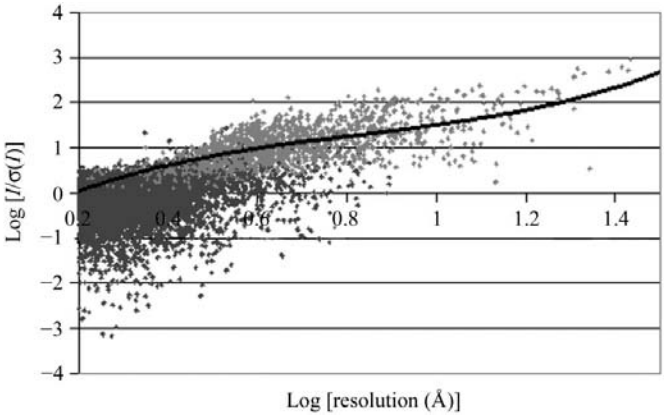


Figure 4
A scatter plot of $\log[I/\sigma(I)]$ verses log(resolution) for strong (gray points) and weak (black points) reflection peaks, where the resolution corresponds to the *d*-spacings of the reflections. A trend line (polynomial of order 3) has been fitted to the strong reflection points as a visual guide.

Table 2
R_{merge}, completeness (%Comp), $I/\sigma(I)$ and $\langle I/s_A \rangle$ *versus* resolution for D-xylose isomerase.

The cumulative R_{merge} and completeness are also given (CumR_{merge} and Cum%Comp, respectively). Only reflections with $I/\sigma(I) > 1.5$ are included in the above terms. The completeness for all reflections without an $I/\sigma(I)$ threshold, %AllComp, is also given. s_A corresponds to the standard deviation in merging reflections as opposed to $\sigma(I)$, the error involved in measuring individual intensities. Laue data tend to have lower completeness than those measured with monochromatic radiation (Yang *et al.*, 1998).

Resolution	No. reflections	%Comp	Cum%Comp	All%Comp	R_{merge}	CumR_{merge}	$\langle I/s_A \rangle$	$I/\sigma(I)$
∞–5.69	9263	97.2	97.2	99.6	0.109	0.109	16.4	3.4
5.69–4.02	16480	93.7	94.9	99.1	0.181	0.150	7.9	3.4
4.02–3.29	17800	87.0	91.3	96.9	0.213	0.172	4.8	2.9
3.29–2.85	17665	72.7	84.8	94.0	0.256	0.185	2.5	2.5
2.85–2.55	17013	55.8	76.5	91.3	0.285	0.192	1.6	1.4
2.55–2.32	15998	40.8	68.0	88.1	0.272	0.195	1.3	2.2
2.32–2.15	15082	30.0	60.1	84.6	0.255	0.197	1.2	2.5
2.15–2.01	13884	22.1	53.2	80.8	0.237	0.198	1.1	2.6
2.01–1.90	12520	16.2	47.2	76.1	0.221	0.198	1.0	2.9
1.90–1.80	11291	12.4	42.1	70.8	0.223	0.198	1.0	3.1

Table 3
Wavelength-bin-resolved R_{merge}, $I/\sigma(I)$ and I/σ_A values determined from D-xylose isomerase TOF neutron diffraction data.

Detector geometry and the relative number of neutrons at each wavelength produced by the LANSCE source contribute to the number and intensity of the observed reflections for the three wavelength ranges. The width of these wavelength bins is similar to the wavelength selection used for data measurement with LADI at ILL.

	λ = 1.0–1.99 Å		λ = 2.0–2.99 Å		λ = 3.0–3.99 Å	
Resolution	R_{merge} (No. reflections)	$I/\sigma(I)$ (I/s_A)	R_{merge} (No. reflections)	$I/\sigma(I)$ (I/s_A)	R_{merge} (No. reflections)	$I/\sigma(I)$ (I/s_A)
∞–5.69	0.092 (2641)	5.1 (13.9)	0.074 (1461)	6.8 (12.8)	0.094 (933)	6.2 (7.8)
5.69–4.02	0.151 (3522)	3.2 (6.7)	0.125 (2462)	5.1 (6.8)	0.150 (1588)	4.8 (4.4)
4.02–3.29	0.183 (3120)	3.2 (4.6)	0.162 (2857)	4.4 (4.7)	0.169 (865)	4.4 (2.9)
3.29–2.85	0.227 (1676)	2.6 (2.2)	0.208 (2148)	3.5 (2.7)	0.175 (26)	4.6 (1.8)
2.85–2.55	0.250 (919)	2.6 (2.2)	0.246 (949)	2.8 (1.9)	(0)	
2.55–2.32	0.259 (636)	2.3 (2.0)	0.230 (361)	3.2 (1.6)	(0)	
2.32–2.15	0.258 (524)	2.4 (1.9)	0.211 (86)	3.3 (1.4)	(0)	
2.15–2.01	0.241 (354)	2.6 (1.8)	(0)		(0)	
2.01–1.90	0.246 (208)	2.5 (1.7)	(0)		(0)	
1.90–1.80	0.241 (150)	2.7 (1.7)	(0)		(0)	
Cumulative	0.176 (13750)	3.1 (4.2)	0.169 (10324)	4.3 (4.5)	0.149 (3412)	5.0 (4.6)

In addition to the corresponding 20 wavelength-resolved Laue patterns, three patterns were collected at supplementary settings at the end of data collection. The exposure time at each crystal setting was typically 12 h. However, a few settings were collected at up to 48 h and down to 6 h, depending on beam time availability. The quality of the data, as indicated by the number of strongly measured reflections, was proportional to the exposure time. Data-collection strategies on the PCS are described in greater detail by Langan & Greene (2004).

Neutrons at LANSCE are produced by high-energy protons hitting a metal target in microsecond pulses at a rate of 20 Hz. The exact time the protons strike the target is designated T_0. Neutrons in pulses are time-stamped and travel down the beamline from their time of production, T_0, as a function of their energy, so that neutrons of different energies or wavelengths are detected at different times. The position and the time of flight (TOF) of diffracted neutrons are recorded on a position-sensitive detector and associated electronics. The pixel resolution of the detector is 0.77 mm and the neutrons are binned into 96 time channels, each of 500 µs in duration. Since the neutron beam is generally interrupted for maintenance several times a day, a more meaningful measure of the exposure of the crystal to neutrons is the number of elapsed T_0 values. The number of elapsed T_0 values for each crystal setting ranged from about 40 000 to about 3 000 000. Data collected at three of the settings are represented in Fig. 3. This figure was constructed by overlaying the data collected at each time channel for each crystal setting in order to produce a projection similar to a conventional Laue diffraction pattern.

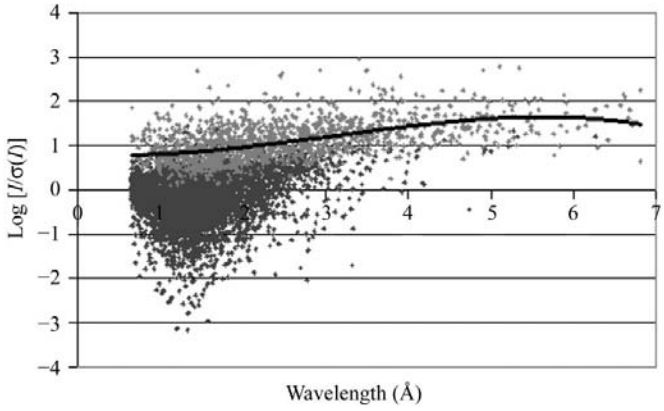
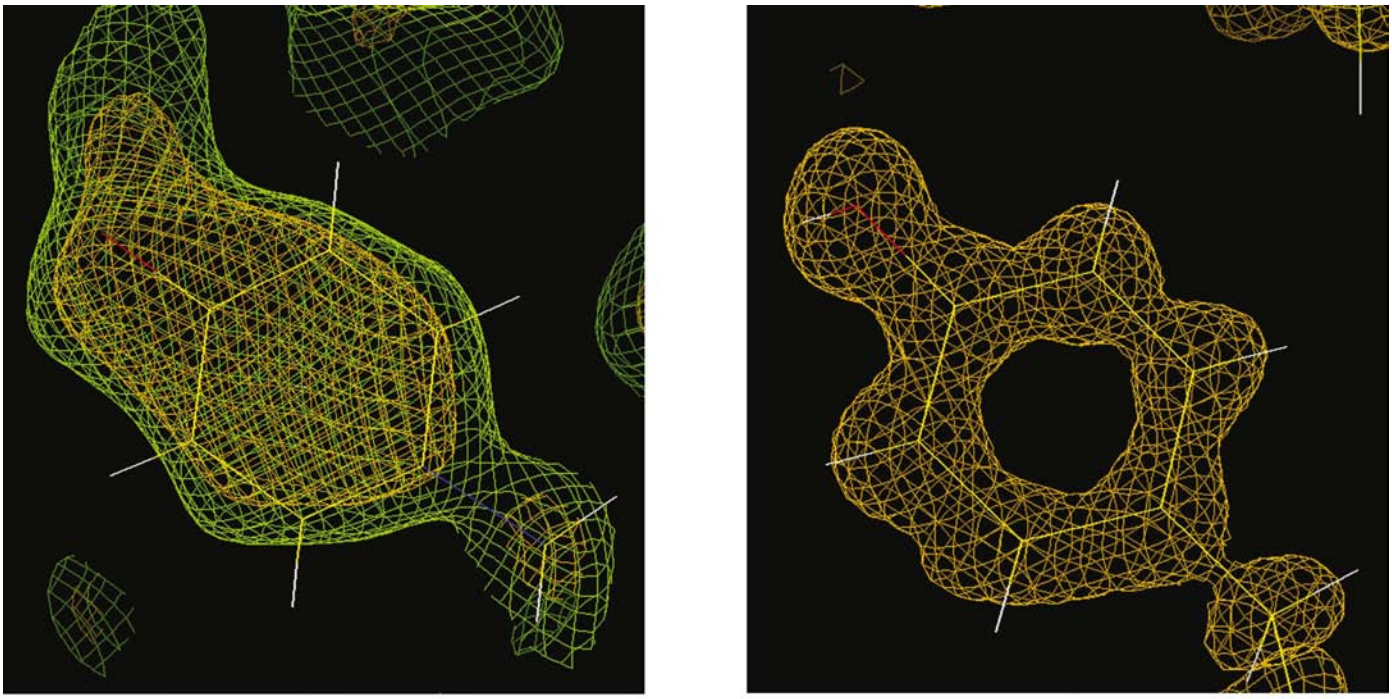


Figure 5
A scatter plot of $\log[I/\sigma(I)]$ verses wavelength for strong (gray points) and weak (black points) reflection peaks. A trend line (polynomial of order 3) has been fitted to the strong reflection points as a visual guide.

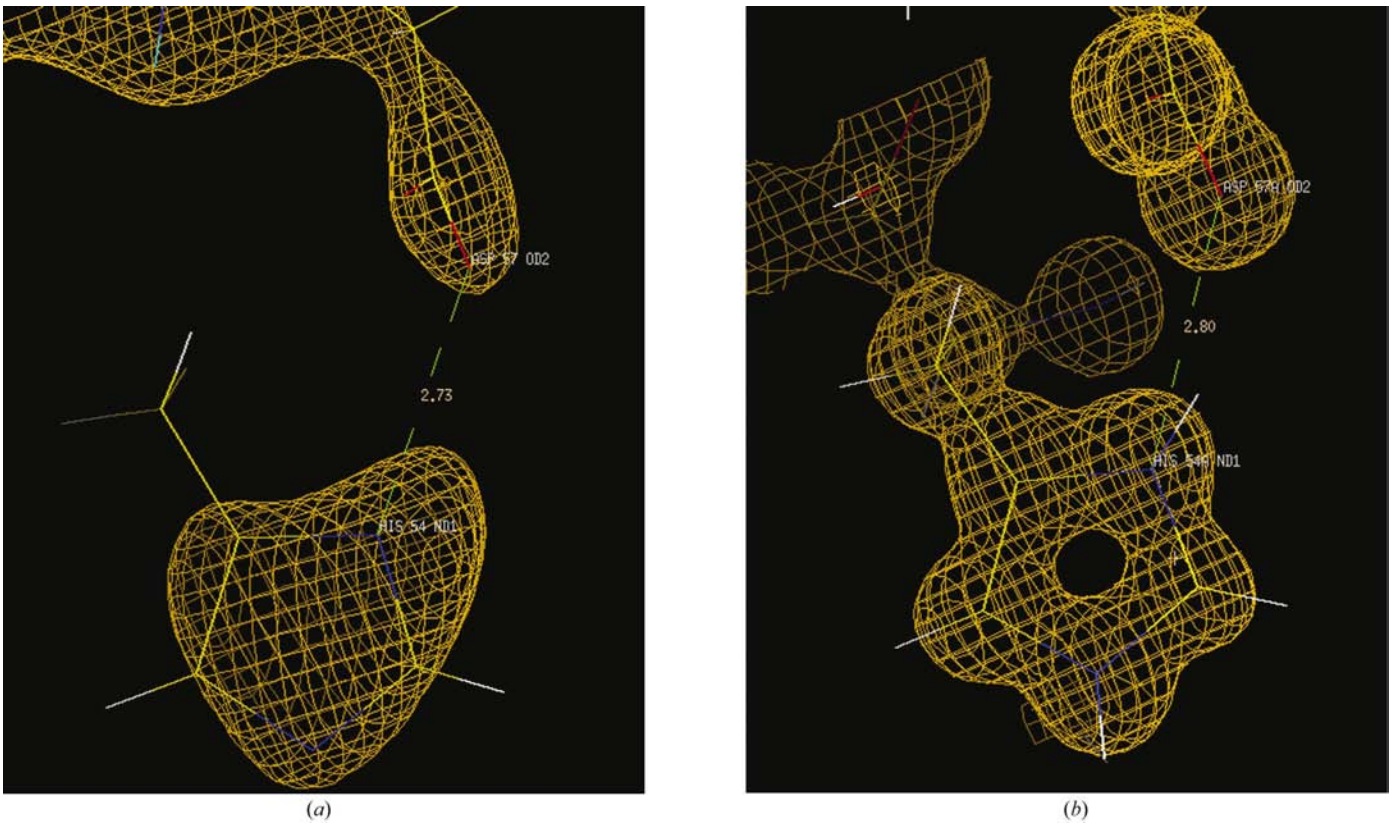
Data indexing and integration were performed using a version of the computer program *d*TREK* (Pflugrath, 1999) that had been modified for wavelength-resolved Laue neutron protein crystallography and that is described in detail by Langan & Greene (2004). An automatic peak-finding routine located between 452 and 1263 strong peaks ($d > 1.8$ Å) at each crystal setting. These peaks were used to determine the crystal orientation matrix. After refining the matrix, together with a number of instrument parameters, the r.m.s. deviation of the predicted positions of reflections on the detector from the observed positions varied between 0.6 and 1.3 mm. The average size of a reflection on the detector was approximately 3 mm full-width at half-height (FWHH). The total number of integrated reflections was 181 797, of which 38 029 were unique with $d > 1.8$ Å.

The quality of the data varies with the resolution. Fig. 4 shows a scatter plot of $\log(I/\sigma)$ *versus* $\log(d)$ (d is the reso-



(*a*) (*b*)

Figure 6
(*a*) Neutron density for Tyr137 contoured at 1.5σ (green) and 2.0σ (goldenrod). The orientation of the deuteron on the O atom of the tyrosine is clearly indicated by the positive neutron density. (*b*) The protonation state of Tyr254 cannot be determined from the 0.94 Å resolution nitrogen-cooled X-ray structure. The program *XtalView* was used to generate the neutron and electron-density images (McRee, 1999).



(*a*) (*b*)

Figure 7
(*a*) Neutron density for His54 contoured at 2σ as described in the text. The neutron density indicates the protonation state of the two ring N atoms. ND1 is protonated as shown by the positive neutron density. (*b*) The protonation state of His54 cannot be determined from the 0.94 Å resolution nitrogen-cooled X-ray structure.

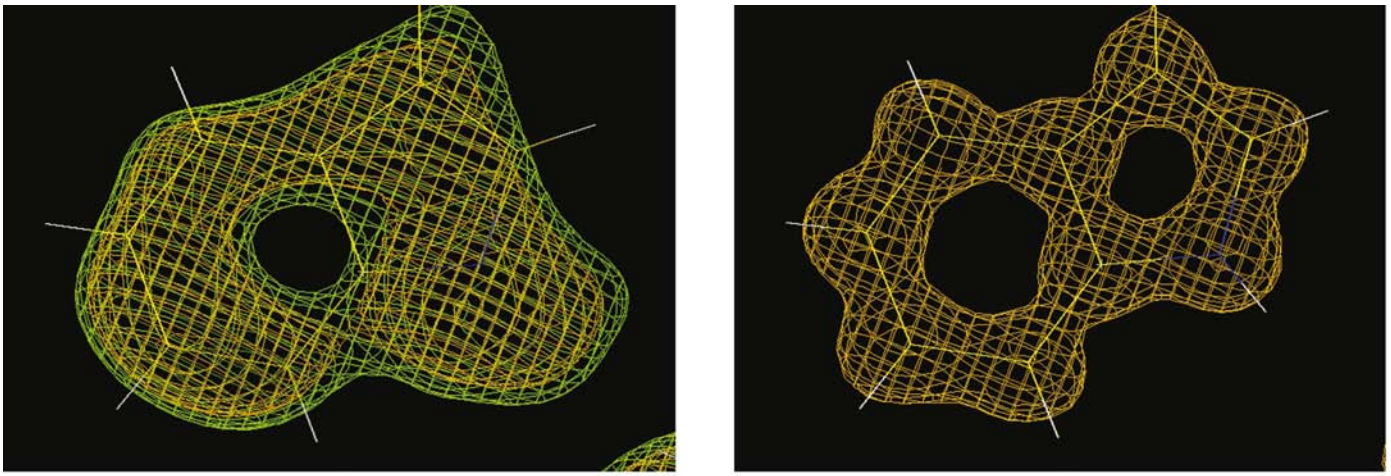


Figure 8
(*a*) Neutron density for Trp137 contoured at 1.5 and 2.0σ (green and goldenrod, respectively) and (*b*) electron density contoured at 2.0σ. The protonation state of the ring N atom cannot be ascertained in the electron-density map from the 0.94 Å resolution low-temperature X-ray data. The neutron density map calculated at 2.0 Å resolution from the room-temperature neutron data shows that the ring N atom is deuterated.

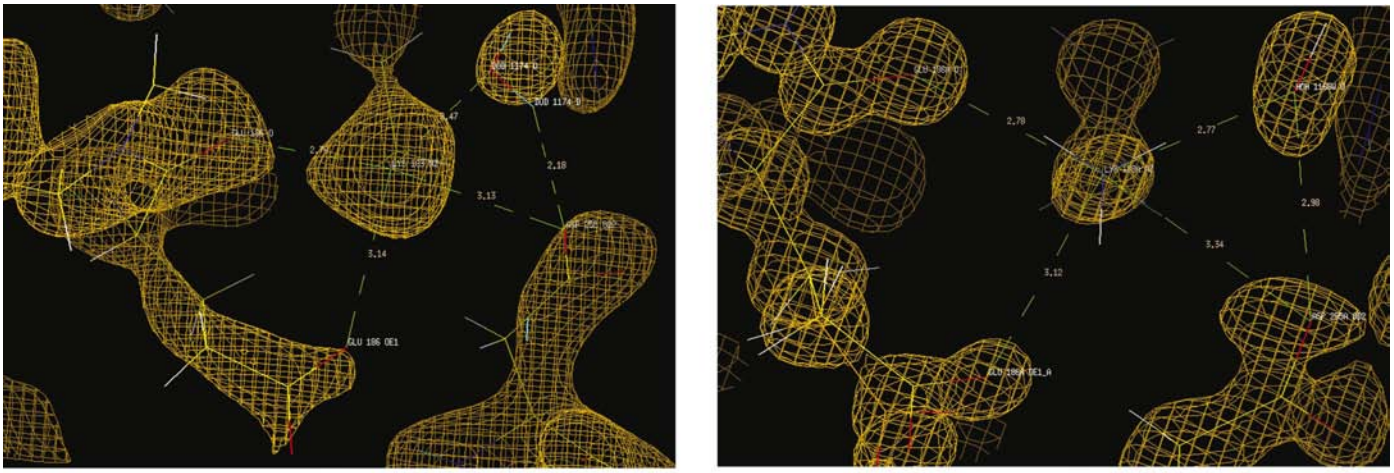


Figure 9
(*a*) The terminal N atom of Lys183 is triply deuterated in the neutron density map contoured at 2σ. Hydrogen bonds are made from the N atom to the main-chain O atom of Glu186, a carboxylate O atom of the same Glu186 and a D_2O molecule. The D_2O molecule makes a hydrogen bond to Asp255, which in turn may share the third deuteron on the terminal N atom with Glu186 OE1. Lys183 may function as a hydrogen donor during catalysis. The protonation state of the terminal N atom cannot be determined from the 0.94 Å resolution electron-density map shown in (*b*).

lution) for all reflections measured out to *d*-spacings of 1.5 Å in one of the wavelength-resolved Laue patterns collected for 48 h. In this figure, I is the measured intensity and σ is the standard error associated with its measurement. 1440 strong reflection peaks were integrated (gray points) and used to construct empirical peak shapes for eight different regions of the detector. The peak shapes were then used to integrate intensities at the positions of 8400 other possible reflections (black points). Strong reflections extend to 1.5 Å resolution and have values of $I/\sigma(I)$ that vary from 1.7 to 1000 ($\log^{-1}0.2$ and $\log^{-1}3$). Beyond 2.5 Å resolution ($\log^{-1}0.4$), the weak reflections are more numerous. Although the weak reflections have values of $I/\sigma(I)$ as large as 10 or more, the mean value for $I/\sigma(I)$ drops to 0.3 ($\log^{-1}-0.5$) at 1.5 Å resolution.

The quality of the data varies with wavelength. Fig. 5 shows a scatter plot of $\log[I/\sigma(I)]$ *versus* wavelength (0.7–7 Å) for intense (gray points) and weak (black points) reflection peaks. The number of measured reflections increases with decreasing wavelength because of the constant solid angle covered by the detector. There is a slight decrease in the number of measured reflections at the smallest wavelengths because some reflections are rejected from integration owing to time-dispersion overlap. The number of both strong and weak peaks is greatest between 1 and 2 Å and this is also where the neutron flux is at a maximum. $I/\sigma(I)$ increases with wavelength, probably as a result of increased sample reflectivity.

After integration, the data were converted into a format for input into the program *LAUENORM* (Helliwell *et al.*, 1989).

Using the data from all 23 crystal settings, six iterations of a combined five cycles of wavelength-normalization and four cycles of intercycle scaling were performed. The wavelength-normalization curve is modeled using nine Chebyshev polynomials (Arzt *et al.*, 1999). In order to obtain reasonable values for $R_{\rm merge}$, the wavelength range was restricted to 0.9–6.5 Å and only reflections with $I > 3\sigma$ were used in determining the wavelength-normalization polynomial; $R_{\rm merge}$ is 0.258 for all measurements of a reflection, 0.235 for all measurements of a reflection of the same sign and 0.173 for all measurements of a reflection of the same sign and with a wavelength range of 0.1 Å. The *LAUENORM* data were output in unmerged form so that *SCALA* could be used for statistical analysis (Collaborative Computational Project, Number 4, 1994; Diederichs & Karplus, 1997; Weiss & Hilgenfeld, 1997; Weiss, 2001). Only data with $I/\sigma(I)$ greater than 1.5 were used in data scaling.

4. Results

The quality of the neutron diffraction data were then assessed. The completeness and other statistics of the merged data are presented in Table 2. The values of $R_{\rm merge}$ for the XI data for most resolution bins listed in this table are larger than values reported to date in neutron studies of other proteins. The cumulative $R_{\rm merge}$ value of 0.198 compares with values of 0.17 for concanavalin A (Habash *et al.*, 1997), 0.19 for a concanavalin–sugar complex (Kalb *et al.*, 2001), 0.13 for lysozyme (Myles *et al.*, 1997), 0.14 for endothiapepsin (Cooper & Myles, 2000) and 0.11 for *trp* repressor (Daniels *et al.*, 2003). The two main differences between this study and previous ones are the size of the system under investigation and the wavelength range of neutrons used. With a unit-cell volume of approximately 1×10^6 Å^3, this study represents by far the largest system studied to high resolution with neutron diffraction.

Previous neutron data sets on other proteins have been collected at steady-state neutron sources using either monochromatic neutron beams or quasi-Laue neutron beams with a wavelength band of approximately 1 Å, whereas the neutron data set reported here has been collected over a large wavelength range of 0.7–7 Å (restricted to 0.9–6.5 Å during data merging). In order to investigate the effects of using a large wavelength band on the data-merging statistics, the data were normalized and then merged over small wavelength bands of 1 Å width (Table 3). The values of $R_{\rm merge}$ within each wavelength band (cumulative values are 0.176 for λ = 1–2 Å, 0.156 for λ = 2–3 Å and 0.139 for 3–4 Å) are comparable to or better than the values reported for neutron data collected from large proteins using equivalent wavelength bands (Habash *et al.*, 1997; Kalb *et al.*, 2001). Over equivalent wavelength ranges, the accuracy of the data is at least as good as previous data collected on other neutron instruments.

The cumulative values for $R_{\rm merge}$ and $I/\sigma(I)$ given in Table 3 improve as the wavelength increases. However, this is partly due to the number of weaker high-resolution reflections decreasing as the neutron wavelength increases and therefore $d_{\rm min}$ increases (see Fig. 4). Given the results from Table 3, the large values of $R_{\rm merge}$ reported in Table 2 would appear to result from the use of a large range of wavelengths. Variations in the detector efficiency or the incident neutron spectrum with wavelength cannot explain this observation, because these effects are accounted for during wavelength normalization with *LAUENORM*. Variations in incoherent scattering from hydrogen result in an effective absorption factor that is a function of both wavelength and crystal shape. The scaled intensities of equivalent reflections collected at the same wavelength will differ because of the different associated path lengths through the crystal and the size of this difference will change with wavelength. The effective linear absorption coefficient for the crystal used in this study has approximate values of 0.5 and 2 cm^{-1} at 0.7 and 7 Å, respectively. The corresponding transmissions in the short (1 mm) and long (2 mm) dimensions of the crystal are 95 and 90% at 0.7 Å, and 82 and 67% at 7 Å. The magnitude of this effect may explain the large $R_{\rm merge}$ values. The unusually high data redundancy in the data collected on the PCS (the average multiplicity of reflections reported here is 4) will compensate for this effect to some degree and result in accurate averaged intensities. However, methods are also being developed to properly correct for this effect so that the accuracy of the data can be further improved.

The ultimate test of the quality of the data set is its ability to locate single D atoms. Figs. 6, 7, 8 and 9 show details from a $2F_o - F_c\ \sigma_A$ map calculated using phases from the room-temperature structure determined from a 1.55 Å X-ray data set (Carrell *et al.*, 1994; PDB code 1xib), with all of the water-molecule atoms and labile H atoms omitted and using neutron amplitudes in the resolution range 10–2 Å. There are clear indications of the protonation states of residues around the catalytic site as illustrated in Figs. 7, 8 and 9. Direct observation of these protonation states is of crucial importance in determining the details of the catalytic mechanism. Fig. 9 shows density at the end of a lysine residue (Lys183) that defines the orientation of the terminal ND_3 group. This particular orientation allows the individual D atoms in the ND_3 group to hydrogen bond to a main-chain O atom from Glu186, an O atom of a water molecule (D_2O) and the third D atom to hydrogen bond to either a carboxylate O atom of the same Glu186 or a carboxylate O atom from an active-site Asp255.

5. Discussion

The neutron diffraction data measured from the unliganded XI represents an excellent first step in an ongoing research program. Additional neutron data collection will focus on information needed to elucidate the mechanisms of XI. The three mechanisms of action of this enzyme proposed to date, shown in Fig. 1, involve a *cis*-enediol intermediate, a hydride shift and a metal-assisted hydride shift, respectively. The first requires a neighboring base in the active site plus a group that would bind the two O atoms of the enediol intermediate. The second does not have any neighboring-group requirements,

except for those binding the two O atoms of the substrate. The third implies a movement of the metal ion (Garcia-Viloca *et al.*, 2002), for which the ultrahigh-resolution X-ray data may provide evidence. Presumably, this mechanism requires a metal-bound water molecule that is appropriately oriented. Each of these possible mechanisms involve differences in H-atom orientations in the active site and these can be investigated in our detailed structural studies.

In Figs. 6, 7, 8 and 9, the locations of many exchanged D atoms are clearly revealed, even in these initial maps of neutron density. Further molecular refinement and interpretation of the protonation state of the active-site amino-acid residues may help indicate the suitability of each of the proposed XI mechanisms. It is next planned to investigate an enzyme–ligand complex of D-xylose isomerase by neutron diffraction using deuterated xylose as the enzymatic substrate.

We acknowledge NIH grants CA-10925 (to JPG and AK) and CA-06927 (to FCCC). BLH and GJB acknowledge support by the Oak Ridge National Laboratory Laboratory Director's Research and Development seed money, grants NAG8-1568 and NAG8-1826 from the National Aeronautics and Space Administration and GM-29818 from the National Institutes of Health. The PCS is funded by the Office of Science and the Office of Biological and Environmental Research of the US Department of Energy.

References

Arzt, S., Campbell, J. W., Harding, M. M., Hao, Q. & Helliwell, J. R. (1999). *J. Appl. Cryst.* **32**, 554–562.
Bon, C., Lehmann, M. S. & Wilkinson, C. (1999). *Acta Cryst.* D**55**, 978–987.
Borah, B., Chen, C.-W., Egan, W., Miller, M., Wlodawer, A. & Cohen, J. S. (1985). *Biochemistry*, **24**, 2058–2067.
Callens, M., Kersters-Hilderson, H., Van Opetal, O. & De Bruyne, C. K. (1986). *Enzyme Microb. Technol.* **8**, 696–700.
Carrell, H. L., Glusker, J. P., Burger, V., Manfre, F., Tritsch, D. & Biellmann, J.-F. (1989). *Proc. Natl Acad. Sci. USA*, **86**, 4440–4444.
Carrell, H. L., Hoier, H. & Glusker, J. P. (1994). *Acta Cryst.* D**50**, 113–123.
Carrell, H. L., Rubin, B. H., Hurley, T. & Glusker, J. P. (1984). *J. Biol. Chem.* **259**, 3230–3236.
Carter, D. C., Wright, B., Miller, T., Chapman, J., Twigg, P., Keeling, K., Moody, K., White, M., Click, J. Ruble, J., Ho, J. X., Adcock-Downey, L., Bunick, G. J. & Harp, J. M. (1999). *J. Cryst. Growth*, **196**, 602–609.
Coates, L., Erskine, P. T., Wood, S. P., Myles, D. A. & Cooper, J. B. (2001). *Biochemistry*, **40**, 13149–13157.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Collyer, C. A. & Blow, D. A. (1990). *Proc. Natl Acad. Sci. USA*, **87**, 1362–1366.
Collyer, C. A., Henrick, K. & Blow, D. M. (1990). *J. Mol. Biol.* **212**, 211–235.
Cooper, J. B. & Myles, D. A. A. (2000). *Acta Cryst.* D**56**, 246–248.
Daniels, B. V., Myles, D. A. A., Forsyth, V. T. & Lawson, C. L. (2003). *Acta Cryst.* D**59**, 136–138.
Diederichs, K. & Karplus, P. A. (1997). *Nature Struct. Biol.* **4**, 269–275.
Farber, G. K., Glasfeld, A., Tiraby, G., Ringe, D. & Petsko, G. A. (1989). *Biochemistry*, **28**, 7289–7297.
Farber, G. K., Petsko, G. A. & Ringe, D. (1987). *Protein Eng.* **1**, 459–466.
Garcia-Viloca, M., Alhanbra, C., Truhlar, D. G. & Gao, J. (2002). *J. Comput. Chem.* **24**, 177–190.
Habash, J., Raftery, J., Weisgerber, S., Cassetta, A., Lehmann, M. S., Wilkinson, C., Campbell, J. W. & Helliwell, J. R. (1997). *J. Chem. Soc. Faraday Trans.* **93**, 4313–4317.
Hanson, B. L., Harp, J. M., Kirschbaum, K., Schall, C. A., DeWitt, K., Howard, A., Pinkerton, A. A. & Bunick, G. J. (2002). *J. Synchrotron. Rad.* **9**, 375–381.
Helliwell, J. R., Habash, J., Cruickshank, D. W. J., Harding, M. M., Greenhough, T. J., Campbell, J. W., Clifton, I. J., Elder, M., Machin, P. A., Papiz, M. Z. & Zurek, S. (1989). *J. Appl. Cryst.* **22**, 483–497.
Henrick, K., Collyer, C. A. & Blow, D. W. (1989). *J. Mol. Biol.* **208**, 129–157.
Jauch, W. & Dachs, H. (1994). *Proceedings of the Workshop on Neutron Scattering Instrumentation for SNQ*, edited by R. Scherm & H. H. Stiller. Report Jül-1954, FZ Jülich, Jülich, Germany.
Kalb, A. J. K., Myles, D. A. A., Habash, J., Raftery, J. & Helliwell, J. R. (2001). *J. Appl. Cryst.* **34**, 454–457.
Kossiakoff, A. A. & Spencer, S. A. (1980). *Nature (London)*, **288**, 414–416.
Langan, P. & Greene, G. (2004). Submitted.
Langan, P., Greene, G. & Schoenborn, B. P. (2004). *J. Appl. Cryst.* **37**, 24–31.
Lavie, A., Allen, K. N., Petsko, G. A. & Ringe, D. (1994). *Biochemistry*, **33**, 5469–5480.
McRee, D. E. (1999). *J. Struct. Biol.* **125**, 156–165.
Mason, S. A., Bentley, G. A. & McIntyre, G. J. (1984). *Basic Life Sci.* **27**, 323–334.
Myles, D. A. A., Bon, C., Langan, P., Cipriani, F., Castagna, J. C., Lehmann, M. S. & Wilkinson, C. (1997). *Physica B*, **241**, 1122–1130.
Niimura, N., Minezaki, Y., Nonaka, T., Castgna, J.-C., Ciprianin, F., Høghøj, P., Lehmann, M. S. & Wilkinson, C. (1997). *Nature Struct. Biol.* **4**, 909–914.
Pflugrath, J. W. (1999). *Acta Cryst.* D**55**, 1718–1725.
Ramachander, S. & Feather, M. S. (1977). *Arch. Biochem. Biophys.* **178**, 576–580.
Rey, F., Jenkins, J., Janin, J., Lasters, I., Alard, P., Claessens, M., Matthyssens, G. & Wodak, S. (1988). *Proteins Struct. Funct. Genet.* **4**, 165–173.
Rose, I. A., O'Connell, E. L. & Mortlock, R. P. (1969). *Biochem. Biophys. Acta*, **178**, 376–379.
Schoenborn, B. P. (1992). *Proc. SPIE*, **1738**, 192–199.
Schoenborn, B. P. & Pitcher, E. (1996). *Neutrons in Biology*, edited by B. P. Schoenborn & R. B. Knott. New York: Plenum Press.
Smart, O. S., Akins, J. & Blow, D. M. (1992). *Proteins*, **13**, 100–111.
Weiss, M. (2001). *J. Appl. Cryst.* **34**, 130–135.
Weiss, M. & Hilgenfeld, R. (1997). *J. Appl. Cryst.* **30**, 203–205.
Whitlow, M. & Howard, A. J. (1989). *Trans. Am. Crystallogr. Assoc.* **25**, 105–126.
Whitlow, M., Howard, A. J., Finzel, B. C., Poulos, T. L., Winborne, E. & Gilliland, G. L. (1991). *Proteins Struct. Funct. Genet.* **9**, 153–173.
Wilkinson, C. & Lehmann, M. S. (1991). *Nucl. Instrum. Methods Phys. Res. A*, **310**, 411–415.
Wlodawer, A., Miller, M. & Sjolin, L. (1983). *Proc. Natl Acad. Sci. USA*, **80**, 3628–3631.
Wohl, A. & Neuberg, C. (1900). *Berichte*, **33**, 3095–3110.
Yang, X., Ren, Z. & Moffat, K. (1998). *Acta Cryst.* D**54**, 367–377.