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

Room-temperature ultrahigh-resolution time-of-flight neutron and X-ray diffraction studies of H/D-exchanged crambin

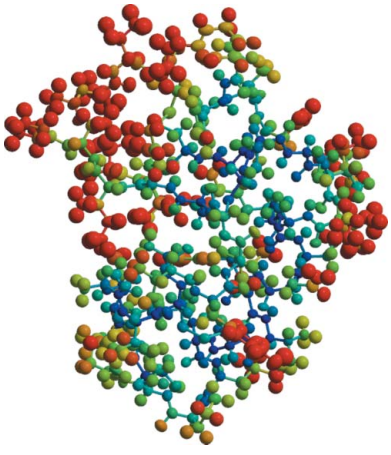
Julian C.-H. Chen,[a]* Zoë Fisher,[b] Andrey Y. Kovalevsky,[b] Marat Mustyakimov,[b] B. Leif Hanson,[c] Vladimir V. Zhurov[c] and Paul Langan[c,d]

[a]Institute of Biophysical Chemistry, Goethe University Frankfurt, Max-von-Laue-Strasse 9, 60438 Frankfurt, Germany, [b]Bioscience Division, Los Alamos National Laboratory, Los Alamos, NM 87545, USA, [c]Department of Chemistry, University of Toledo, Toledo, OH 53606, USA, and [d]Biology and Soft Matter Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA

Correspondence e-mail: chen@chemie.uni-frankfurt.de

Received 14 October 2011
Accepted 30 November 2011

PDB Reference: H/D-exchanged crambin, 3u7t.



The room-temperature (RT) X-ray structure of H/D-exchanged crambin is reported at 0.85 Å resolution. As one of the very few proteins refined with anisotropic atomic displacement parameters at two temperatures, the dynamics of atoms in the RT and 100 K structures are compared. Neutron diffraction data from an H/D-exchanged crambin crystal collected at the Protein Crystallography Station (PCS) showed diffraction beyond 1.1 Å resolution. This is the highest resolution neutron diffraction reported to date for a protein crystal and will reveal important details of the anisotropic motions of H and D atoms in protein structures.

1. Introduction

H atoms make up approximately one-half of all atoms in a protein molecule. H atoms are involved in enzyme mechanisms and stabilizing hydrogen-bond interactions and are key indicators of the microenvironment within a protein structure. However, they are very difficult to locate within a protein structure using X-ray diffraction techniques owing to their weak scattering power and their proximity to heavier atoms.

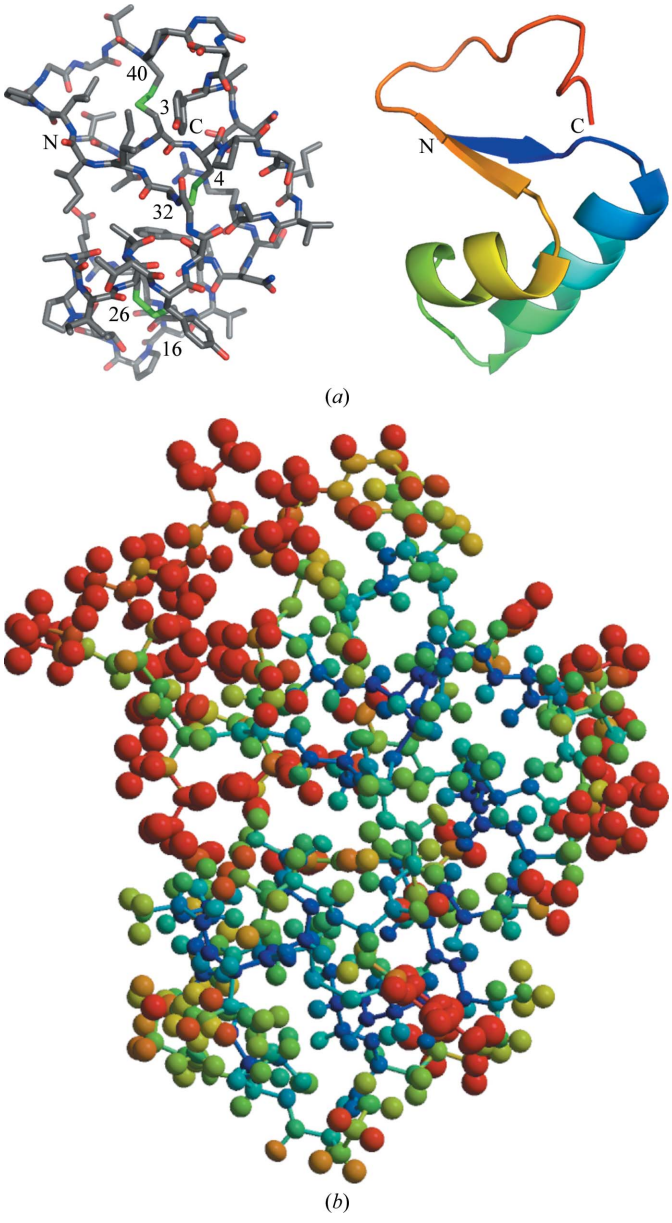
Ultrahigh-resolution X-ray data have the potential to yield a significant amount of structural information about a protein. At resolutions beyond ∼0.70 Å, interatomic scattering from bonding electrons and even lone electron pairs can be resolved. Charge-density refinement has been applied to a number of macromolecules at this resolution (Jelsch *et al.*, 2000; Guillot *et al.*, 2008). H atoms may also be visible at ultrahigh resolution, but normally only well ordered H atoms, typically at the core of proteins. However, most of the mobile and labile H atoms of critical mechanistic interest remain difficult to identify. Even with the best-diffracting protein crystals, only around two-thirds of H atoms can be resolved (Gardberg *et al.*, 2010). Furthermore, the vibrational characteristics of H atoms in proteins remain poorly understood, with little experimental information on their properties.

Neutron crystallography is the only method that can be used to directly and precisely determine the locations of H and D atoms in macromolecular structures, even at a medium resolution of ∼1.8–2.5 Å (Blakeley, 2009). Neutrons are scattered by atomic nuclei, in contrast to X-rays, which are scattered by electrons. The scattering power of X-rays is dependent on the atomic number; in contrast, the neutron scattering properties of atomic nuclei are relatively evenly spread across the periodic table. Notably, the heavy elements in proteins have similar neutron scattering lengths. More importantly, the neutron scattering length of deuterium (+6.671 fm) is comparable to those of oxygen (+5.803 fm), carbon (+6.646 fm) and nitrogen (+9.36 fm), while hydrogen has a negative scattering length (−3.7406 fm). This property has been utilized to accurately locate H atoms of critical interest in enzyme mechanisms (Kossiakoff & Spencer, 1981; Bennett *et al.*, 2006; Blakeley *et al.*, 2008; Kovalevsky *et al.*, 2008, 2010) and also to obtain information on the dynamics of proteins through their backbone amide-exchange characteristics (Kossiakoff, 1982). Furthermore, neutron crystallography has been successful in resolving water species such as D_2O, OD^- and D_3O^+, orienting hydrogen-bond donors and acceptors, and for orienting

metal-coordinated solvent molecules (Blum *et al.*, 2009; Chen *et al.*, 2010; Fisher *et al.*, 2010; Kovalevsky *et al.*, 2011).

Crambin (46 amino acids, 4.7 kDa) is a small protein derived from the seed of the Abyssinian cabbage (*Crambe abyssinica*). Crambin contains both α-helical and β-strand secondary structures and is stabilized by three disulfide bridges (Fig. 1*a*). Its overall fold is similar to that of a family of plant toxins, although it has no known biological function. The protein is hydrophobic, is soluble in polar organic solvents such as acetone and ethanol, and is insoluble in water and alkanes. Crambin forms the best-diffracting protein crystals known; synchrotron radiation has yielded data to a usable resolution of 0.48 Å, with observed reflections to 0.40 Å (Teeter & Hendrickson, 1979; Jelsch *et al.*, 2000; Schmidt *et al.*, 2011). With the possibility of obtaining ultrahigh-resolution neutron diffraction data, it is expected that it should be possible to resolve not only the locations of H and D atoms but also to refine their anisotropic vibrational characteristics. As such, crambin is an excellent target for neutron crystallography and the aim of this study is to obtain neutron diffraction to ultrahigh resolution as a first step towards understanding the motions of H and D atoms in proteins in the context of their heavier neighbors. These ultrahigh-resolution data may also help in improving the parameterization of neutron structure refinement techniques.

The ultrahigh-resolution RT X-ray structure of H/D-exchanged crambin is reported at 0.85 Å, together with preliminary 1.1 Å neutron diffraction data collected at the PCS at the Los Alamos Neutron Science Center, demonstrating the exceptional order of crambin crystals. Future prospects for the use of ultrahigh-resolution neutron and X-ray data are discussed.



(*a*)

(*b*)

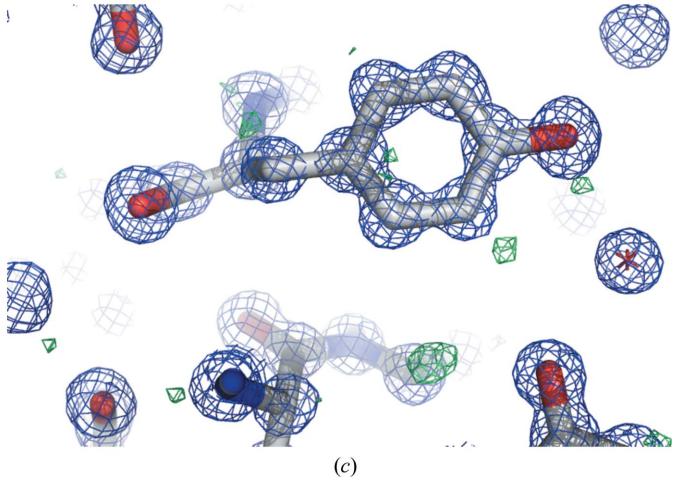
2. Experimental

2.1. Protein preparation and crystallization

Crambin was purified according to the procedures of Van Etten *et al.* (1965). Crystals were grown by vapor diffusion of an ∼20 mg ml^{-1} solution of crambin in 80% ethanol against a well solution of 55% ethanol at RT. For the ultrahigh-resolution X-ray data collection, a single prismatic crystal measuring approximately 1.6 × 0.4 × 0.7 mm was mounted in a quartz capillary and was exchanged for two weeks against mother liquor containing 55% perdeuterated ethanol (Cambridge Isotopes) in D_2O. For neutron data collection, a larger crystal measuring 3.0 × 2.0 × 0.7 mm (4.2 mm^3) was mounted in a quartz capillary and was exchanged against perdeuterated mother liquor by vapor diffusion for six weeks prior to data collection.

2.2. X-ray data collection and refinement

Data collection was performed with a Rigaku diffractometer equipped with an UltraX-18 Mo rotating-anode generator (50 kV, 300 mA), a graphite monochromator and a RAPID cylindrical image-plate detector. Data were processed and scaled in *d*TREK* (Pflugrath, 1999) and the structure was solved by molecular replacement using the crambin structure 1crn (Teeter, 1984) as a search model. The structure was refined to 0.85 Å resolution in *PHENIX* using positional and anisotropic *B*-factor refinement. Riding H atoms were



(*c*)

Figure 1
(*a*) Stick and secondary-structure representations of crambin. Disulfide bridges are labeled in green. (*b*) X-ray structure of H/D-exchanged crambin with anisotropic ellipsoids, colored by *B* factor from low (blue) to high (red). The orientation of the molecule is the same in (*a*) and (*b*). (*c*) Representative electron density for residue Tyr44 of H/D-exchanged crambin. $2F_o - F_c$ electron density contoured at 1.5σ is shown in blue, with $F_o - F_c$ density at 3.0σ in green. Difference density for the hydrogen of the tyrosine hydroxyl group allows clear orientation of the H atom.

Table 1
Data-collection and refinement statistics.

Values in parentheses are for the highest resolution shell.

	X-ray	Neutron
Source	Rotating anode	PCS, LANSCE
Wavelength (Å)	0.71073 [Mo $K\alpha$]	0.6–6.0
Temperature (K)	293	
Unit-cell parameters		
a (Å)	22.79	
b (Å)	18.83	
c (Å)	41.04	
β (°)	90.89	
Space group	$P2_1$	$P2_1$
Resolution (Å)	30–0.85 (0.88–0.85)	20–1.10 (1.16–1.10)
Completeness (%)	99.8 (99.9)	77.4 (64.6)
Unique reflections	30963 (3067)	10930 (1302)
Multiplicity	3.41 (3.21)	2.6 (1.9)
R_{merge} (%)	5.0 (34.2)	22.1 (29.6)
$R_{p.i.m.}$ (%)	N/A	13.9 (21.8)
$\langle I/\sigma(I)\rangle$	8.2 (1.9)	5.6 (1.9)
Refinement		
No. of reflections	30659	
R_{free}/R_{cryst} (%)	13.60/11.94	
R.m.s.d. bonds (Å)	0.008	
R.m.s.d. angles (°)	1.380	
$\langle B\rangle$ protein/solvent (Å^2)	6.6/18.1	
Ramachandran outliers	0	

added in the later rounds of refinement and model building was performed in *Coot*. Solvent molecules were refined as O atoms only with anisotropic atomic displacement parameters (ADPs). Data-collection and refinement statistics are shown in Table 1.

2.3. Neutron data collection

Neutron time-of-flight Laue data were collected at room temperature on the Protein Crystallography Station at the Los Alamos Neutron Science Center (LANSCE; Schoenborn & Langan, 2004). The detector is a position-sensitive ^{3}He-filled detector and the sample was mounted on a kappa-circle goniometer from Huber. Each exposure was 12 h and to date 44 crystal settings have been collected to a maximum usable resolution of 1.1 Å. With the detector positioned with $2\theta = 0^\circ$ the diffraction was observed to extend to the edge of the detector. One image was collected at $2\theta = 70^\circ$ for 20 h to assess the limit of diffraction (Fig. 4*c*). In order to collect a complete high-resolution data set the detector was set to $2\theta = 30^\circ$. The images were processed with an in-house-modified version of *d*TREK* and wavelength-normalized using *LAUENORM* (Langan & Greene, 2004; Helliwell *et al.*, 1989). Data were merged with *SCALA* from the *CCP*4 suite (Winn *et al.*, 2011). Data-collection statistics are summarized in Table 1.

3. Results and discussion

3.1. RT X-ray structure of H/D-exchanged crambin

The RT X-ray structure of H/D-exchanged crambin was solved at 0.85 Å resolution. Labile H atoms were exchanged for D atoms through vapor exchange against perdeuterated mother liquor prior to data collection. H atoms in the structure were refined in their riding positions with isotropic ADPs, and a number of them could be identified as $F_o - F_c$ difference peaks throughout the structure (Fig. 1*c*). Future refinement of a complete neutron data set will clarify which atoms have exchanged for deuteriums and the extent of the exchange.

The overall structure is identical to the 1.50 Å resolution RT structure 1crn, with an average 0.07 Å r.m.s.d. at the C$^\alpha$ positions, and closely resembles the 0.54 Å resolution (PDB entry 1ejg; Jelsch *et al.*, 2000) and 0.48 Å resolution (PDB entry 3nir; Schmidt *et al.*, 2011) structures solved at 100 K, with average r.m.s.d.s of 0.116 and 0.128 Å at C$^\alpha$ positions, respectively. A comparison of the hydration patterns showed little difference between the structures. Therefore, the H/D-exchange process did not affect the overall structure of the protein or its hydration. The X-ray structures of perdeuterated proteins, in which all H atoms have been replaced by deuteriums, also showed nearly identical structures to their H/D-exchanged and hydrogenated counterparts (Blum *et al.*, 2010). The crystal is notable in having a low solvent content of 34.4%. Much of the solvent region between symmetry-related protein molecules is ordered, with relatively little 'bulk' solvent. Such an extensive network of hydrogen-bonded contacts may in part explain the very high degree of order in these crystals.

3.2. Anisotropic motions of crambin at RT *versus* 100 K

This is one of the few ultrahigh-resolution protein structures reported at ambient temperature and provides a view of anisotropic atomic motions of a macromolecule at RT (Fig. 1*b*). Furthermore, the RT structure and the anisotropic vibrational characteristics of the atoms can be compared with the set of structures solved at 100 K. Anisotropic atomic displacement parameters (ADPs) are expressed as a tensor consisting of six values relating to the atomic motions in three dimensions.

Data for 1ejg and 3nir were obtained at 100 K and the structures reported at 0.54 and 0.48 Å, respectively. As expected for a RT structure, the average B factor for the protein atoms in our current structure is 6.6 Å^2, higher than 1ejg and 3nir, and the distribution of B factors is similar in all three structures (Fig. 2).

The RT structure is considerably more isotropic in nature, with a mean anisotropy of 0.578 for the protein atoms, compared with 0.404 for 1ejg, 0.420 for 3nir and an average of 0.45 for ∼1000 structures in the PDB (http://skuld.bmsc.washington.edu/parvati/parvati_survey.html). The anisotropy value is defined as the square of the ratio of the smallest and largest of the eigenvalues of the 3 × 3 matrix of the

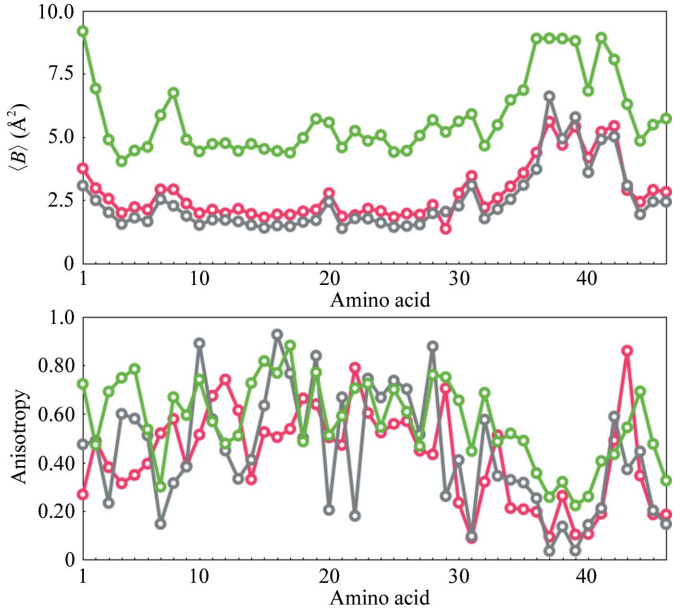


Figure 2
Distribution of C$^\alpha$ B factors (top) and anisotropy (bottom) in RT H/D-exchanged crambin (green) and the 100 K 1ejg (gray) and 3nir (magenta) structures.

anisotropic displacement parameters. A higher anisotropy value represents a more isotropic atom. The distribution of ADPs and their anisotropy are similar in the three crambin structures discussed here, with a region of higher ADP values between residues 34 and 42 which correlates with lower anisotropy values (more anisotropic, less isotropic; Fig. 2). In the case of the two low-temperature structures 1ejg and 3nir, several discrepancies are seen in the anisotropy values which may be a consequence of incomplete modeling of alternate conformations. These figures are likely to be accurate, although they may in part reflect differences in the refinement programs used. The structure reported in this work was refined in *PHENIX*, while 1ejg and 3nir were refined using *SHELX*. Merritt compiled information about the anisotropy of structures refined using *SHELX*, *REFMAC* and *PHENIX* and noted a slight but statistically significant difference between the two programs (http://skuld.bmsc.washington.edu/parvati/parvati_survey.html). Nevertheless, the anisotropy of the 100 K structures is comparable to other protein structures in the PDB, while that of the RT structure is higher. At this time, it is difficult to make a comparison of the anisotropy of other RT structures in the PDB, as it is difficult to accurately assess the data-collection temperature in the PDB entries. Other possible reasons for this discrepancy may be the separation of static and dynamic disorder at lower temperatures.

3.3. Neutron diffraction of a H/D-exchanged crambin crystal

To assess the diffraction limit currently possible for crystals of crambin using current neutron sources, 44 diffraction images have been collected from a single crystal (Fig. 3) at the PCS in Los Alamos, with a total coverage of 77% of reflections to 1.1 Å resolution and 65% of reflections in the outermost shell (1.16–1.10 Å resolution; Table 1, Figs. 4*a* and 4*b*).

A full data set is currently being collected. We expect complete data to ∼1.1 Å resolution using the current data-collection protocol, although it is clear that the crystal diffracts to beyond 1.05 Å from inspection of a diffraction image collected at a 2θ of 70° (Figs. 4*c* and 4*d*). A significant amount of beam time is necessary to collect data beyond 1.1 Å resolution and is currently not practical given the constraints on beam time and the geometry of the detector. The preliminary data show strong diffraction even in the highest resolution shell, with a $\langle I/\sigma(I)\rangle$ of 1.9 in the outermost shell, with very good merging statistics. The data are thus of excellent quality. As the data collection currently stands, there is sufficient completeness to 1.8 Å resolution (∼85%), which is a high enough resolution to determine the orientation of solvent molecules and amino-acid side chains and to determine the extent of H/D exchange in the backbone. A joint X-ray and neutron refinement of the structure to 1.8 Å resolution is under way.

We expect that when the full data set is collected the data-to-parameter ratio for the neutron data alone will be sufficient to justify anisotropic refinement of all atoms, including H and D atoms, and solvent molecules. We can also exploit the advantages of joint refinement protocols, which in the case of crambin will combine ultrahigh-resolution X-ray data and atomic resolution neutron data, possibly yielding the most detailed and accurate protein structure to date.

Figure 3
Photograph of the crambin crystal used for neutron diffraction data collection, measuring 3 × 2 × 0.7 mm (4.2 mm^3).

3.4. Perspective

This work may open up a number of new possibilities in macromolecular neutron crystallography. Of great benefit to our understanding of protein structure and dynamics is the possibility of refining anisotropic ADPs for H and D atoms. Although refinement of anisotropic ADPs is routinely performed for the heavier atoms in proteins (C, N, O and S) using atomic resolution X-ray data (typically to better than 1.2 Å resolution), such treatment is not possible for H atoms. Neutron crystallography, on the other hand, offers insight into H-atom positions owing to their scattering properties, and now with ultrahigh-resolution neutron diffraction data, both positions and dynamics can be determined with high accuracy.

To date, 11 structures in the PDB have been resolved to beyond 0.70 Å resolution, a number that will undoubtedly increase, especially

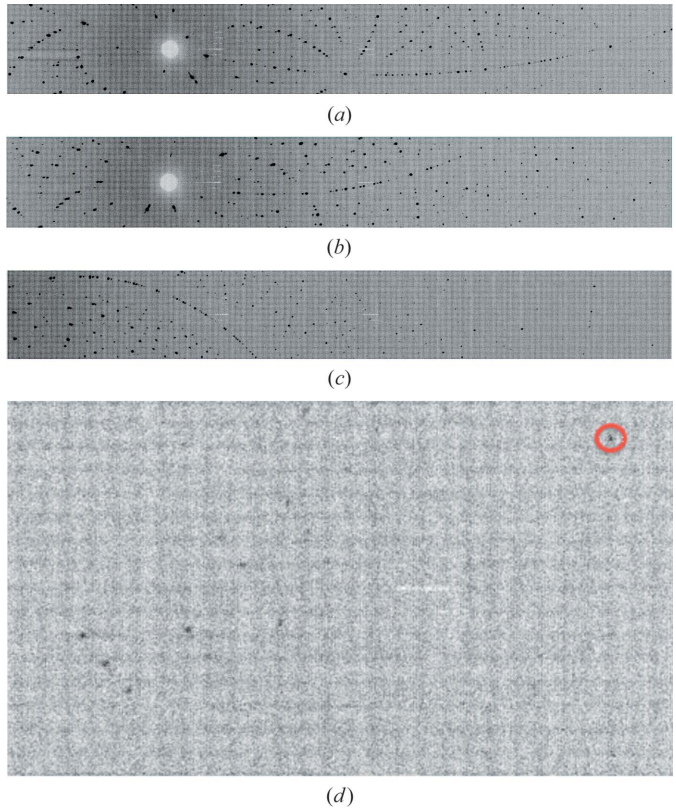


Figure 4
(*a*, *b*) Laue time-of-flight neutron diffraction images from two different settings of H/D-exchanged crambin collected at the PCS. The exposure time was 12 h, with the 2θ arm set at 30°. (*c*) For this image the 2θ arm was moved to 70° and the exposure time was 20 h. The acquired data are displayed as an overlay of all time bins, generating the Laue diffraction pattern in the image. (*d*) Diffraction slice at a wavelength of 1.35 Å between 57° and 82° in 2θ. A reflection at 1.05 Å resolution is circled.

with the development of new synchrotron sources and free electron lasers combined with new and more accurate detectors and hardware. Therefore, more accurate parameterization for crystallographic refinement that includes H atoms will be needed in order to provide higher quality models at high resolutions.

At a software level, this is an ideal test case for joint XN refinement protocols. Joint XN refinement, described over 30 years ago by Wlodawer and Hendrickson, has become a practical tool for refining neutron structures and has been implemented in the commonly used refinement programs *CNS* (as *nCNS*) and *PHENIX* (Adams *et al.*, 2009).

Synchrotron data have also been collected on crambin crystals at DESY to a nominal resolution of 0.48 Å, with observed diffraction spots out to 0.40 Å resolution (Schmidt *et al.*, 2011). With a combination of synchrotron and neutron diffraction data, it should be possible to separate the static and dynamic contributions to disorder. These methods will afford the most detailed view of a macromolecule to date and may provide experimental information about bonding and electrostatics that can be tested against theoretical calculations.

The coordinates and structure factors for the RT X-ray structure of H/D-exchanged crambin have been deposited in the PDB under accession code 3u7t.

We thank Dr Blaine Schatz for providing crambe seed for this study, Dr Pavel Afonine for advice on *PHENIX* refinement and Mary Jo Waltman, Dr A. Alan Pinkerton and Dr Kristin Kirschbaum for technical assistance. JC-HC thanks Dr Justin and Cynthia Chen for their support. The PCS is funded by the US Department of Energy Office of Biological and Environmental Research (DOE-OBER).

References

Adams, P. D., Mustyakimov, M., Afonine, P. V. & Langan, P. (2009). *Acta Cryst.* D**65**, 567–573.
Bennett, B., Langan, P., Coates, L., Mustyakimov, M., Schoenborn, B., Howell, E. E. & Dealwis, C. (2006). *Proc. Natl Acad. Sci. USA*, **103**, 18493–18498.
Blakeley, M. P. (2009). *Crystallogr. Rev.* **15**, 157–218.
Blakeley, M. P., Ruiz, F., Cachau, R., Hazemann, I., Meilleur, F., Mitschler, A., Ginell, S., Afonine, P., Ventura, O. N., Cousido-Siah, A., Haertlein, M., Joachimiak, A., Myles, D. & Podjarny, A. (2008). *Proc. Natl Acad. Sci. USA*, **105**, 1844–1848.
Blum, M.-M., Mustyakimov, M., Rüterjans, H., Kehe, K., Schoenborn, B. P., Langan, P. & Chen, J. C.-H. (2009). *Proc. Natl Acad. Sci. USA*, **106**, 713–718.
Blum, M.-M., Tomanicek, S. J., John, H., Hanson, B. L., Rüterjans, H., Schoenborn, B. P., Langan, P. & Chen, J. C.-H. (2010). *Acta Cryst.* F**66**, 379–385.
Chen, J. C.-H., Mustyakimov, M., Schoenborn, B. P., Langan, P. & Blum, M.-M. (2010). *Acta Cryst.* D**66**, 1131–1138.
Fisher, S. Z., Kovalevsky, A. Y., Domsic, J. F., Mustyakimov, M., McKenna, R., Silverman, D. N. & Langan, P. A. (2010). *Biochemistry*, **49**, 415–421.
Gardberg, A. S., Del Castillo, A. R., Weiss, K. L., Meilleur, F., Blakeley, M. P. & Myles, D. A. A. (2010). *Acta Cryst.* D**66**, 558–567.
Guillot, B., Jelsch, C., Podjarny, A. & Lecomte, C. (2008). *Acta Cryst.* D**64**, 567–588.
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.
Jelsch, C., Teeter, M. M., Lamzin, V., Pichon-Pesme, V., Blessing, R. H. & Lecomte, C. (2000). *Proc. Natl Acad. Sci. USA*, **97**, 3171–3176.
Kossiakoff, A. A. (1982). *Nature (London)*, **296**, 713–721.
Kossiakoff, A. A. & Spencer, S. A. (1981). *Biochemistry*, **20**, 6462–6474.
Kovalevsky, A. Y., Hanson, L., Fisher, S. Z., Mustyakimov, M., Mason, S. A., Forsyth, V. T., Blakeley, M. P., Keen, D. A., Wagner, T., Carrell, H. L., Katz, A. K., Glusker, J. P. & Langan, P. (2010). *Structure*, **18**, 688–699.
Kovalevsky, A. Y., Hanson, B. L., Mason, S. A., Yoshida, T., Fisher, S. Z., Mustyakimov, M., Forsyth, V. T., Blakeley, M. P., Keen, D. A. & Langan, P. (2011). *Angew. Chem. Int. Ed. Engl.* **50**, 7520–7523.
Kovalevsky, A. Y., Katz, A. K., Carrell, H. L., Hanson, L., Mustyakimov, M., Fisher, S. Z., Coates, L., Schoenborn, B. P., Bunick, G. J., Glusker, J. P. & Langan, P. (2008). *Biochemistry*, **47**, 7595–7597.
Langan, P. & Greene, G. (2004). *J. Appl. Cryst.* **37**, 253–257.
Pflugrath, J. W. (1999). *Acta Cryst.* D**55**, 1718–1725.
Schmidt, A., Teeter, M., Weckert, E. & Lamzin, V. S. (2011). *Acta Cryst.* F**67**, 424–428.
Schoenborn, B. P. & Langan, P. (2004). *J. Synchrotron Rad.* **11**, 80–82.
Teeter, M. M. (1984). *Proc. Natl Acad. Sci. USA*, **81**, 6014–6018.
Teeter, M. M. & Hendrickson, W. A. (1979). *J. Mol. Biol.* **127**, 219–223.
Van Etten, C. H., Nielsen, H. C. & Peters, J. E. (1965). *Phytochemistry*, **4**, 467–473.
Winn, M. D. *et al.* (2011). *Acta Cryst.* D**67**, 235–242.