Journal of
Synchrotron Radiation

ISSN 0909-0495

Received 31 July 2007
Accepted 11 January 2008

Protein structures by spallation neutron crystallography

Paul Langan,[a]* Zoë Fisher,[a] Andrii Kovalevsky,[a] Marat Mustyakimov,[a] Amanda Sutcliffe Valone,[a] Cliff Unkefer,[a] Mary Jo Waltman,[a] Leighton Coates,[a]‡ Paul D. Adams,[b] Pavel V. Afonine,[b] Brad Bennett,[c] Chris Dealwis[c] and Benno P. Schoenborn[a]

[a]Bioscience Division, Los Alamos National Laboratory, NM 87545, USA, [b]Lawrence Berkeley National Laboratory, CA 94720, USA, and [c]Department of Pharmacology, School of Medicine, Case Western Reserve University, 10900 Euclid Avenue, Cleveland, OH 44106-4965, USA. E-mail: langan_paul@lanl.gov

The Protein Crystallography Station at Los Alamos Neutron Science Center is a high-performance beamline that forms the core of a capability for neutron macromolecular structure and function determination. This capability also includes the Macromolecular Neutron Crystallography (MNC) consortium between Los Alamos (LANL) and Lawrence Berkeley National Laboratories for developing computational tools for neutron protein crystallography, a biological deuteration laboratory, the National Stable Isotope Production Facility, and an MNC drug design consortium between LANL and Case Western Reserve University.

Keywords: neutrons; proteins; macromolecular crystallography; deuteration; enzyme mechanisms; drug binding; hydration; joint XN structure refinement.

1. Introduction

Neutron crystallography (NC) is a powerful technique for locating H atoms and protons, and can be used to complement X-ray crystallography (XC) with unique information about how biological macromolecules function and interact with each other and smaller molecules (Niimura *et al.*, 2006). NC can also provide detailed information on the hydration of biomacromolecules, including the exact orientation and coordination of water molecules. In addition to providing structural information, NC can be used to identify H atoms that can be exchanged with deuterium and the extent of this replacement, thus providing a tool for the study of protein dynamics, complementary to NMR techniques.

The Protein Crystallography Station (PCS) is the first NC beamline to be built at a spallation source (Langan *et al.*, 2004). Neutrons are produced from a coupled high-flux water moderator (Schoenborn *et al.*, 1999), in pulses at a rate of 20 Hz, and then travel 28 m down a vacuum pipe with collimation inserts tapering the beam to a final divergence, at the sample position, of 0.12°. The time and energy structure of the beam, and also its relatively small divergence, allow data to be collected efficiently and with enhanced signal-to-noise with time-of-flight (wavelength-resolved) Laue techniques (Langan & Greene, 2004). The data are recorded as three-dimensional time-of-flight Laue patterns on a large position-sensitive electronic detector with a spatial resolution of less than a millimetre (Mahler *et al.*, 1998). The PCS is run as a user facility with open access to beam time, protein deuteration, expression and robotic crystallization facilities, the synthesis of substrates with stable isotopes, structure refinement software, and support for data reduction and structure analysis. In this paper we illustrate the capabilities of the PCS for determining protein structures by spallation neutron crystallography, and we describe the methodological and technological advances that are emerging from the Macromolecular Neutron Crystallography (MNC) consortium.

2. Recent highlights from the PCS user program

Diisopropyl fluorophosphatase (DFPase) is a calcium-dependent phosphotriesterase with potential as a candidate for enzymatic decontamination of nerve-agent stocks. Proposed reaction mechanisms include esterase hydrolysis by nucleophilic attack on the P atom of the bound substrate by an activated water molecule. The water is activated by proton abstraction by either a basic histidine or the catalytic calcium. These ideas have recently been challenged. In particular, the work of Blum *et al.* (2006) suggests that the substrate forms a covalent intermediate with the carboxylate group of an aspartic acid and that it is the carboxylate C atom that is then attacked by water. In order to obtain more evidence for this new mechanism, Blum *et al.* (2007) have determined the holoenzyme structure through joint X-ray (1.8 Å) and neutron (2.2 Å) (XN) refinement. The catalytic calcium coordinates a key solvent molecule (W33) that is identified as a water molecule, as required for the new proposal, and not as a hydroxy ion, as required in previous proposals. In addition, the key aspartic acid (Asp229) is clearly deprotonated, as required for the new proposal, and not protonated, as expected in previous proposals (Fig. 1).

Dihydrofolate reductase (DHFR) is an enzyme conserved across species that is essential for most biosynthetic pathways involving one-C-atom transfer reactions, and has become an important chemo-

‡ Present address: Oak Ridge National Laboratory, TN 3781, USA.

therapeutic and antimicrobial drug target. Bennett *et al.* (2006) have determined the 2.2 Å neutron structure of the binary complex of *Escherichia coli* DHFR with the anticancer drug MTX bound, in order to understand the preferential binding of MTX over DHF. The NC structure directly reveals that a ring N atom (N1) of MTX undergoes induced protonation on binding, imparting it with a favorable positive charge for ionic interaction with an unprotonated aspartic acid (Asp27). This result agrees with nearly all available biochemical and complementary structural data but disagrees with the theoretical calculations of Cannon *et al.* (1997). In addition, a detailed analysis of solvent molecules related their interaction with DHFR to differences in H/D exchange. Two molecules in the asymmetric unit bind the drug with the regulatory Met-20 loop (residues 9–24) in two different conformations, occluded and closed. The occluded conformation undergoes H/D exchange more readily than the closed, and this difference is explained by greater water intrusion and solvent accessibility.

D-Xylose isomerase (XI) is one of a large class of aldose–ketose isomerases that require two divalent metal ions for function. A number of possible reaction mechanisms have been proposed. Katz *et al.* (2006) have determined the 1.8 Å neutron structure in order to determine the catalytic role of a metal-bound solvent molecule, located in the active site, which is thought to be involved in the isomerization step in which D-xylose is converted to D-xylulose or D-glucose to D-fructose. Under the conditions of structure determination (pH 8.0) the solvent molecule is shown to be a water molecule (rather than a hydroxy ion) with one of its lone pair electrons pointing towards the C—C substrate bond. This lone pair of electrons could act as a proton-abstracting group. Alternatively, at more alkali conditions, the water molecule may be activated by deprotonation by a metal-coordinated aspartic acid residue (Asp257). Interestingly, a lysine residue in the active site (Lys289) is observed to have an $-NH_2$ terminal group (rather than NH_3^+) and to be disordered between two positions, suggesting a possible role in shuffling an abstracted proton away from Asp257 during the reaction. The ionization state of each histidine residue was also determined. Of particular importance, His54 is doubly protonated, providing an H atom to a water molecule in the absence of substrate and implying its role in sugar-ring opening as promoting a hydrogen-bond interaction with the ring O atom of glucose (rather than its previously proposed role as a proton-abstracting base) (Fig. 2).

Photoactive yellow protein (PYP) from *Halorhodospira halophila* absorbs light *via* its *para*-coumaric acid chromophore (pCA), covalently attached to Cys69, and is thought to be involved in the negative phototactic response of the organism to blue light. Hydrogen bonding in the chromophore binding pocket is thought to be crucial to photocycle kinetics. Fisher *et al.* (2007) have determined the ground-state structure by joint X-ray (1.1 Å) and neutron (2.5 Å) refinement in order to determine the nature of this hydrogen bonding. The structure reveals that the phenolate O atom of pCA accepts two short hydrogen bonds (from Glu46 $O^{\varepsilon 2}$ and Tyr42 O^{η}) and that Thr50 $O^{\gamma 1}$ stabilizes this arrangement by donating a hydrogen bond to Tyr42 O^{η}. However, the deuteron position between pCA and Tyr42 is only partially occupied, indicating that this atom may also interact with Thr50, possibly in a resonance between the two bonds, with important implications for ground-state kinetics.

3. Advances in deuteration

For data collection on the PCS, crystals have to be 0.3 mm^3 or greater in volume (Blum *et al.*, 2007). However, this requirement can be reduced if the protein is deuterated (H replaced by D). Water and labile H atoms in proteins can be replaced simply by soaking crystals in D_2O mother liquor. To replace the remaining H atoms (perdeuteration) requires gene expression in a deuterated growth medium. The effectiveness of perdeuteration for improving NC was first demonstrated by Shu *et al.* (2000). While straightforward in principle, protein expression systems can be sensitive both to the replacement of H_2O by D_2O and to the exact amount and types of deuterated nutrients used, and can also require lengthy adaptation periods before substantial growth is achieved.

Many photosynthetic algae are capable of autotrophic growth in an inorganic environment using CO_2 as their sole carbon source. These organisms can be fully deuterated with D_2O as the only deuterium-containing compound required to support their growth (Katz &

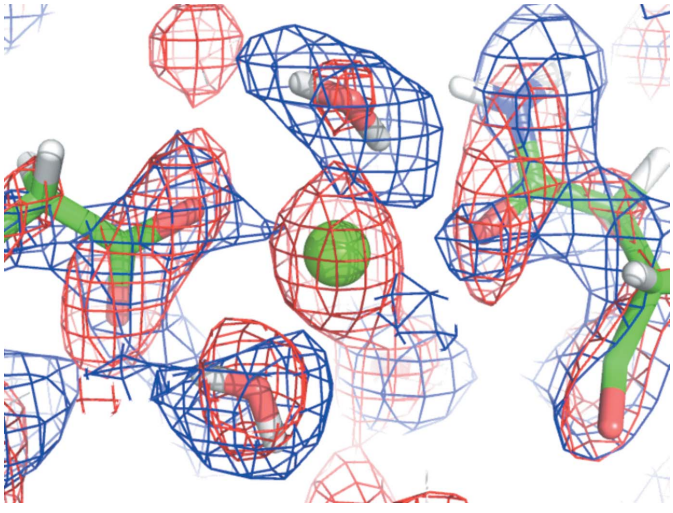

Figure 1
The active site in the XN structure of DFPase showing the coordination of the catalytic calcium ion with E21, N175 and key solvent molecules (Blum *et al.*, 2007). Neutron and X-ray $2F_o - F_c$ scattering density maps are represented in blue and red, respectively.

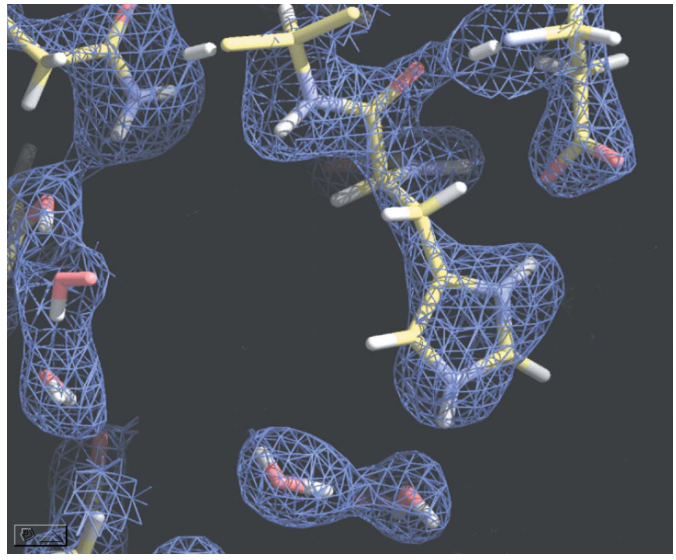

Figure 2
The interaction of His 54 with water molecules in the neutron structure of D-xylose isomerase in the absence of substrate with $2F_o - F_c$ neutron scattering density represented in blue (Katz *et al.*, 2006).

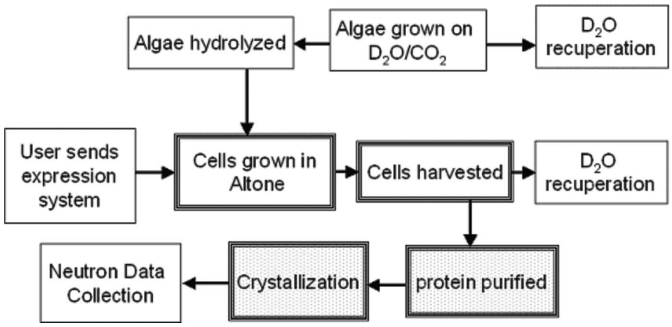

Figure 3
A schematic representation of the process used for perdeuterating proteins at the BDL at Los Alamos National Laboratory (LANL). A bacterial culture medium, which we designate Altone, is made from the hydrolyzate of algae (*e.g. Scenedesmus obliquus*) grown in D_2O and used for protein expression in *E. coli*. Steps represented by framed boxes can be performed robotically at LANL. Steps represented by shaded boxes can be performed either at LANL or the user's home laboratory.

Crespi, 1966). We use mass production of deuterated algae as an inexpensive source of deuterated metabolites. This approach, described in more detail in its application to studying the catalytic mechanism of haloalkane dehalogenase from *Xanthobacter autrophicus* by Liu *et al.* (2007) and schematically illustrated in Fig. 3, is exploited at the biological deuteration laboratory (BDL) in order to provide a facility for expressing large quantities of perdeuterated proteins.

Perdeuteration at the BDL is free to PCS users and many proteins have been perdeuterated to date. This inexpensive and efficient approach is being combined with robotic overexpression to produce a series of apo, binary and ternary complexes in mechanistic and drug design studies of *E. coli* and *Bacillus anthracis* DHFR (Langan & Dealwis, 2006). In addition, at the National Stable Isotope Production Facility (NSIPF), we have developed methods for the synthesis of specifically deuterated amino acids based on the Oppolzer chiral glycine template.

4. Advances in computational tools

Most biologists who use the PCS are not experienced in NC and for structure refinement rely on adapting programs used in XC. This approach can be time-consuming and complicated and is never completely satisfactory because there are fundamental differences between the way neutrons and X-rays are scattered. With the increasing number of NC structures and their increasing size and complexity this situation is untenable, and we have therefore established the MNC consortium to develop computational tools that are fully accessible to the growing community of occasional and non-expert NC users (these tools are free and can be downloaded from the MNC website at http://mnc.lanl.gov). A first goal is to make software available quickly, to meet the immediate needs of the MNC community. A patch, designated *nCNS*, for the existing structure solution program called *CNS* (Brünger *et al.*, 1998) has been developed. This allows joint X-ray and neutron refinement with the use of cross-validated maximum-likelihood refinement targets and simulated annealing optimization (Langan & Mustyakimov, 2007). *nCNS* has already been used to determine the joint XN structures of PYP (Fisher *et al.*, 2007) and DFPase (Blum *et al.*, 2007).

In addition, we have added NC capabilities to *PHENIX* (Adams *et al.*, 2004). The software *phenix.refine* (Afonine *et al.*, 2005*a*) offers a broad variety of efficient and fully automated tools for structure refinement using XC, NC or both, such as restrained refinement at low resolution, simulated annealing and TLS modelling using maximum-likelihood targets, automatic detection and use of NCS, automatic detection and use of twinning information, sophisticated bulk-solvent correction and anisotropic scaling protocols (Afonine *et al.*, 2005*b*), efficient handling of H atoms, and proper refinement at ultrahigh resolution. *phenix.refine* has been used to determine the joint XN structure of aldose reductase, an NADPH-dependent enzyme that reduces a wide range of substrates, such as aldehydes, aldoses and corticosteroids (Afonine *et al.*, 2007).

5. Conclusions

The studies of DFPase (Blum *et al.*, 2007), DHFR (Bennett *et al.*, 2006), XI (Katz *et al.*, 2006) and PYP (Fisher *et al.*, 2007) highlighted here, as well as additional studies of amicyanin (Sukumar *et al.*, 2005), nucleic acids (Langan *et al.*, 2006), endothiapepsin cocrystallized with a *gem*-diol inhibitor (Tuan *et al.*, 2007), and human deoxyhaemaglobin (Kovalevskyi *et al.*, 2008) reported elsewhere, demonstrate how spallation neutrons are being used to greatly advance our understanding of biomacromolecular structure and function. The proteins being studied on the PCS are becoming larger and more complex and the required sample sizes are becoming smaller. A number of technological and methodological advances are being made by MNC consortia centered at the PCS, particularly in the development of deuteration and computation tools. A key feature of the computational tools being developed is the ability to refine joint XN structures, an approach originally developed for proteins by Wlodawer & Hendrickson (1982).

The PCS is funded by the Office of Biological and Environmental Research of the Department of Energy. MM, PA, PDA and PL were supported by an NIH–NIGMS-funded consortium (1R01GM071939-01) between Los Alamos and Lawrence Berkeley National Laboratories to develop computational tools for neutron protein crystallography. AK, ASV, BB, CD and PL were supported by a Los Alamos National Laboratory LDRD grant (20070131ER).

References

Adams, P. D., Gopal, K., Grosse-Kunstleve, R. W., Hung, L.-W., Ioerger, T. R., McCoy, A. J., Moriarty, N. W., Pai, R. K., Read, R. J., Romo, T. D., Sacchettini, J. C., Sauter, N. K., Storoni, L. C. & Terwilliger, T. C. (2004). *J. Synchrotron Rad.* **11**, 53–55.
Afonine, P. V. *et al.* (2007). ACA Meeting, Salt Lake City, Utah, USA, Vol. 34, p. 44.
Afonine, P. V., Grosse-Kunstleve, R. W. & Adams, P. D. (2005*a*). *CCP4 Newsl.* **42**, contribution 8.
Afonine, P. V., Grosse-Kunstleve, R. W. & Adams, P. D. (2005*b*). *Acta Cryst.* D**61**, 850–855.
Bennett, B., Langan, P., Coates, L., Mustyakimov, M., Schoenborn, B. P., Howell, E. E. & Dealwis, C. (2006). *Proc. Natl Acad. Sci. USA*, **103**, 18493–18498.
Blum, M.-M., Koglin, A., Rüterjans, H., Schoenborn, B., Langan, P. & Chen, J. C.-H. (2007). *Acta Cryst.* F**63**, 42–45.
Blum, M.-M., Loehr, F., Richardt, A., Rueterjans, H. & Chen, J. C.-H. (2006). *J. Am. Chem. Soc.* **128**, 12750–12757.
Brünger, A. T. *et al.* (1998). *Acta Cryst.* D**54**, 905–921.
Cannon, W. R., Garrison, B. J. & Benkovic, S. J. (1997). *J. Mol. Biol.* **271**, 656–668.
Fisher, S. Z., Anderson, S., Henning, R., Moffat, K., Langan, P., Thiyagarajan, P. & Schultz, A. J. (2007). *Acta Cryst.* D**63**, 1178–1184.
Katz, A. K., Li, X., Carrell, H. L., Hanson, B. L., Langan, P., Coates, L., Schoenborn, B. P., Glusker, J. P. & Bunick, G. J. (2006). *Proc. Natl Acad. Sci. USA*, **103**, 8342–8347.
Katz, J. J. & Crespi, H. L. (1966). *Science*, **151**, 1187–1194.

Kovalevsky, A. Y., Chatake, T., Shibayama, N., Park, S.-Y., Ishikawa, T., Mustyakimov, M., Fisher, S. Z., Langan, P. & Morimoto, Y. (2008). *Acta Cryst.* F**64**. Submitted.

Langan, P & Dealwis, C. (2006). *A Los Alamos National Laboratory-funded consortium (LDRD 20070131ER) to develop understand the mechanism of E. coli DHFR and develop drugs against B. anthracis DHFR.* Los Alamos National Laboratory, NM, USA, and Case Western Reserve University, OH, USA.

Langan, P. & Greene, G. (2004). *J. Appl. Cryst.* **37**, 253–257.

Langan, P., Greene, G. & Schoenborn, B. P. (2004). *J. Appl. Cryst.* **37**, 24–31.

Langan, P., Li, X., Hanson, B. L., Coates, L. & Mustyakimov, M. (2006). *Acta Cryst.* F**62**, 453–456.

Langan, P. & Mustyakimov, M. (2007). *C-06, 104 Patch for CNS: nCNS. An open source distribution patch for CNS for macromolecular structure refinement.* Los Alamos National Security, LLC (download from http://mnc.lanl.gov).

Liu, X., Hanson, B. L., Langan, P. & Viola, R. E. (2007). *Acta Cryst.* D**63**, 1000–1008.

Mahler, G. J., Radeka, V., Schaknowski, N. A., Smith, G. C., Yu, B. & Zojceski, Z. (1998). *IEEE Nuclear Science Symposium*, Vol. 46, No. 6, Part 2, pp. 1916–1919.

Niimura, N., Arai, S., Kurihara, K., Chatake, T., Tanaka, I. & Bau, R. (2006). *Cell. Mol. Life Sci.* **62**, 285–300.

Schoenborn, B. P., Court, J. D., Larson, A. C. & Ferguson, P. J. (1999). *Neutron Res.* **7**, 89–106.

Shu, F., Ramakrishnan, V. & Schoenborn, B. P. (2000). *Proc. Natl Acad. Sci. USA*, **97**, 3872–3877.

Sukumar, N., Langan, P., Mathews, F. S., Jones, L. H., Thiyagarajan, P., Schoenborn, B. P. & Davidson, V. L. (2005). *Acta Cryst.* D**61**, 640–642.

Tuan, H.-F., Erskine, P., Langan, P., Cooper, J. & Coates, L. (2007). *Acta Cryst.* F**63**, 1080–1083.

Wlodawer, A. & Hendrickson, W. A. (1982). *Acta Cryst.* A**38**, 239–247.