Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Rapid visualization of hydrogen positions in protein neutron crystallographic structures

Parthapratim Munshi,[a,b] Shang-Lin Chung,[a] Matthew P. Blakeley,[c] Kevin L. Weiss,[a] Dean A. A. Myles[a] and Flora Meilleur[a,d]*

[a]Neutron Sciences Directorate, Oak Ridge National Laboratory, PO Box 2008, Oak Ridge, TN 37831, USA, [b]Department of Chemistry, Middle Tennessee State University, TN 37132, USA, [c]Institut Laue–Langevin, BP 156, Grenoble 38042, France, and [d]Department of Molecular and Structural Biochemistry, North Carolina State University, Raleigh, NC 27695, USA

Correspondence e-mail: meilleurf@ornl.gov

Received 9 August 2011
Accepted 15 November 2011

PDB References: perdeuterated rubredoxin, 3rz6; 3ss2; 3ryg; 3rzt.

Neutron crystallography is a powerful technique for experimental visualization of the positions of light atoms, including hydrogen and its isotope deuterium. In recent years, structural biologists have shown increasing interest in the technique as it uniquely complements X-ray crystallographic data by revealing the positions of D atoms in macromolecules. With this regained interest, access to macromolecular neutron crystallography beamlines is becoming a limiting step. In this report, it is shown that a rapid data-collection strategy can be a valuable alternative to longer data-collection times in appropriate cases. Comparison of perdeuterated rubredoxin structures refined against neutron data sets collected over hours and up to 5 d shows that rapid neutron data collection in just 14 h is sufficient to provide the positions of 269 D atoms without ambiguity.

1. Introduction

X-ray crystallography has thrived as the most powerful technique for rapid protein structure determination at the atomic level. Structure determination from small crystals in minutes gives X-ray crystallography at third-generation synchrotrons an undisputable lead. However, because of the weak interaction of X-rays with light elements, H atoms can only be seen (if at all) in X-ray crystal structures if they are well ordered and if data are available at atomic resolution (<1.0 Å). Whilst the positions of many H atoms can be reliably inferred from the chemical groups to which they are bound, the positions of other more labile – and perhaps more interesting – atoms cannot. The need for experimental hydrogen localization in macromolecules has encouraged structural biologists to brave the technical hurdles of neutron crystallography (Myles, 2006; Meilleur, Myles *et al.*, 2006; Blakeley, 2009) to tackle essential questions in enzymatic mechanisms, protein–ligand binding interactions, DNA structures and hydration. The advantage of neutron crystallography over X-ray crystallography comes from its unique ability to visualize hydrogen, or its isotope deuterium, at 'moderate' resolutions (1.5–2.5 Å), while hydrogen remains difficult to localize in subatomic resolution X-ray structures (Blakeley *et al.*, 2006; Gardberg *et al.*, 2010). Recent technical advances in neutron instrumentation for protein crystallography have provided orders-of-magnitude gains in efficiency compared with conventional neutron diffractometers. Significant instrument developments include the use of high-performance neutron image plates mounted on cylindrical detectors (Tanaka *et al.*, 2002) and the development of fast Laue and quasi-Laue diffractometers at reactor facilities (Myles *et al.*, 1997; Cole *et al.*, 2001; Wilkinson *et al.*, 2009)

© 2012 International Union of Crystallography
Printed in Singapore – all rights reserved

and of fast electronic detectors for time-of-flight Laue diffraction instruments at spallation neutron sources (Langan *et al.*, 2004).

Similarly, the ability to exchange hydrogen with deuterium in protein structures delivers an order-of-magnitude gain in the signal-to-noise ratio of the neutron diffraction data and enhances the interpretation of the resulting neutron maps (Meilleur *et al.*, 2005, 2009). Together, these technical advances have made feasible studies of larger biological complexes and smaller crystals than was previously possible (Blakeley, Ruiz *et al.*, 2008*b*; Howard *et al.*, 2011) and have underpinned the resurgence of interest in neutron crystallography in structural biology (Blakeley *et al.*, 2004; Meilleur, Snell *et al.*, 2006; Budayova-Spano *et al.*, 2006; Blakeley, Ruiz *et al.*, 2008; Adachi *et al.*, 2009; Yamaguchi *et al.*, 2009; Fisher *et al.*, 2010; Kovalevsky, Hanson *et al.*, 2010; Howard *et al.*, 2011; Tomanicek *et al.*, 2010, 2011).

Unfortunately, neutron crystallography still comes with a number of obstacles, including the requirement for large (0.1–1.0 mm^3) fully or partially deuterated crystals, long data-collection times compared with X-ray data collection and, as an emerging limiting step, scarce access to single-crystal neutron diffractometers dedicated to macromolecular crystallography. Indeed, while crystallographers now have access to more than 100 X-ray crystallography beamlines at synchrotrons worldwide and to thousands of X-ray generators, at present only four neutron facilities operate single-crystal diffractometers that are dedicated to macromolecular crystallography: LADI-III (Blakeley *et al.*, 2010) and the upgraded D19 (Teixeira *et al.*, 2008) at the Institut Laue–Langevin (ILL); PCS (Kovalevsky, Fisher *et al.*, 2010) at the Los Alamos Neutron Science Center (LANSCE); iBIX (Tanaka *et al.*, 2010) at the Japan Proton Accelerator Research Complex (J-PARC); and BIX3 (Kurihara *et al.*, 2004) and BIX4 (Tanaka *et al.*, 2002) at the Japan Atomic Energy Research Institute (JAERI). With such limited access and increasing demand, the lead time for neutron protein structure determination can be dramatically long. Oak Ridge National Laboratory (ORNL) is bringing on-line three new instruments that will accept international user proposals for neutron macromolecular crystallographic studies: the Laue time-of-flight diffractometers TOPAZ (Zikovsky *et al.*, 2011) and MaNDi (Schultz *et al.*, 2005; Coates *et al.*, 2010) at the Spallation Neutron Source (SNS) and the quasi-Laue diffractometer IMAGINE (Meilleur, personal communication) at the High Flux Isotope Reactor (HFIR). The IMAGINE beamline is designed and optimized for supra-macromolecular assemblies and macromolecular Laue and quasi-Laue crystallography. Whilst the IMAGINE diffractometer itself is modelled upon LADI-III (Blakeley *et al.*, 2010), a novel neutron optics system has been designed to maximize the flux delivered to the sample (Robertson & Stoica, unpublished work). With the increasing number of proposals submitted to a limited number of available beamlines, the question arises as to what the best experimental strategy may be to provide meaningful structure determination and analysis. Therefore, we are undertaking a series of analyses to plan and design experiments that will ensure best use of the valuable, but limited, neutron beam time on Laue and quasi-Laue neutron diffractometers.

Previous studies have focused on the increased visibility of hydrogen and deuterium positions in neutron-density maps compared with electron-density maps (Habash *et al.*, 2000; Blakeley *et al.*, 2006; Gardberg *et al.*, 2010). Here, we investigate data-collection schemes to determine what the minimal data-collection time may be to visualize hydrogen positions. We report the neutron structure of perdeuterated *Pyrococcus furiosus* rubredoxin (PfRd) protein determined from data sets collected over 14, 40, 48 and 128 h. We compare the information content, specifically the visibility of D atoms, in the '14 h structure' with structures refined against data sets collected over longer time scales and with published neutron structures. The results suggest approaches that may make it possible to considerably speed up and improve neutron data-collection strategies and structure analyses in favorable cases.

2. Materials and methods

2.1. Perdeuteration

The expression and purification of perdeuterated PfRd (D-rubredoxin) have been described elsewhere (Weiss *et al.*, 2008; Meilleur *et al.*, 2009; Gardberg *et al.*, 2010). Briefly, perdeuterated rubredoxin was purified using a combination of heat treatment, anion-exchange chromatography and gel filtration (Jenney & Adams, 2001). The yield of purified perdeuterated PfRd was ∼45 mg protein per litre of flask culture.

2.2. Crystal growth

D-Rubredoxin was crystallized by vapor diffusion at room temperature following the protocol previously reported for H/D-rubredoxin crystallization (Weiss *et al.*, 2008). Orthorhombic crystals were grown from two crystallization macroseeding experiments. The 14 h data set was collected from a 3.2 mm^3 crystal. All other data sets were collected consecutively using another crystal of 3.9 mm^3 in volume.

2.3. Neutron data collection and reduction

Crystals were transported to the ILL in microcentrifuge tubes containing 3.4 *M* sodium/potassium phosphate and were mounted in quartz capillaries upon arrival. Neutron quasi-Laue data for both crystals were collected at room temperature (295 K) on the LADI-III instrument using a narrow-bandpass filter for wavelength selection (λ = 3.5 Å, $d\lambda/\lambda$ = 20%; Blakeley, Langan *et al.*, 2008; Blakeley *et al.*, 2010). Each quasi-Laue image corresponds to an approximate 6–7° rotation of the crystal. Crystals were aligned to enable a unique quadrant of reciprocal space to be explored in a minimal number of Laue images. For the 3.9 mm^3 D-rubredoxin crystal three sub-data sets were collected (Gardberg *et al.*, 2010). The first sub-data set consisted of ten images collected with a φ separation of 10° (starting at φ = 0° and ending at φ = 90°) and an exposure time of 4 h per frame. A second interleaving sub-data set of ten images was collected with a φ separation of 10°, with the first image offset by 5° in φ

Table 1
Neutron data-collection, processing and refinement statistics.

Values in parentheses are for the highest resolution shell.

Data set	40 h (data set 1)	48 h (data set 3)	128 h (data sets 1 + 2 + 3)	14 h†
PDB code	3rz6	3ss2	3ryg	3rzt
Crystal volume (mm^3)	3.9	3.9	3.9	3.2
Source	LADI-III, ILL	LADI-III, ILL	LADI-III, ILL	LADI-III, ILL
Wavelength range (Å)	3.3–4.2	3.3–4.2	3.3–4.2	3.2–4.2
Settings	10	12	32	14
Setting spacing (°)	10	7	7, 10	10
Exposure time/frame (h)	4	4	4	1
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Unit-cell parameters (Å)	a = 33.92, b = 34.93, c = 43.53	a = 33.92, b = 34.93, c = 43.53	a = 34.37, b = 35.34, c = 44.11	a = 33.92, b = 34.87, c = 43.50
Resolution (Å)	21.77–1.75 (1.84–1.75)	24.33–1.75 (1.84–1.75)	27.11–1.75 (1.84–1.75)	27.21–1.75 (1.84–1.75)
Unique reflections	4416 (442)	4283 (408)	5125 (605)	4187 (421)
Redundancy	3.0 (1.9)	3.5 (2.5)	6.8 (3.5)	4.6 (2.6)
Completeness (%)	79.6 (58.3)	77.2 (53.4)	92.2 (76.0)	75.6 (56.2)
R_{merge} (%)	8.6 (12.3)	7.4 (11.8)	7.8 (9.9)	15.3 (18.9)
$R_{p.i.m.}$ (%)	5.4 (9.9)	4.3 (7.6)	2.8 (5.9)	7.2 (11.5)
Mean $I/\sigma(I)$	11.0 (6.1)	14.5 (7.7)	22.0 (10.7)	8.4 (3.6)
R_{work} (%)	20.78	21.05	18.08	20.22
R_{free} (%)	23.81	24.24	19.96	24.90
R.m.s.d. bonds (Å)	0.016	0.011	0.009	0.013
R.m.s.d. angles (°)	1.234	1.183	1.203	1.257

† These data were collected from a different crystal.

(starting at $\varphi = 5^\circ$ and ending at $\varphi = 95^\circ$) and an exposure time of 4 h per frame. This protocol covers reciprocal space rapidly, although sparsely, in the first pass and adds to data completeness and redundancy in the second pass. For the third sub-data set, the crystal was reoriented by >45° using an extended kappa arc goniometer and 12 images were collected with a φ separation of 7° and an exposure time of 4 h per frame (starting at $\varphi^* = 0^\circ$ and ending at $\varphi^* = 77^\circ$). In a routine data-collection scheme, this third pass is designed to complete the blind region of the first orientation. The diffraction limit for each sub-data set was 1.65 Å resolution. Finally, a 14 h data set consisting of 14 images collected with a φ separation of 10° (starting at $\varphi' = 0^\circ$ and ending at $\varphi' = 130^\circ$) and exposure time of just 1 h per frame was collected from a second crystal of 3.2 mm^3 in volume. The diffraction limit for this data set was 1.75 Å resolution.

All four data sets were processed using the modified version of the Daresbury Laboratory software *LAUEGEN* (Helliwell *et al.*, 1989; Campbell *et al.*, 1998). The program *LSCALE* (Arzt *et al.*, 1999) was used to derive the wavelength-normalization curve. The data for each set were then scaled and merged using *SCALA* (Winn *et al.*, 2011). The first, second and third sub-data sets collected from the 3.9 mm^3 crystal were combined, scaled and merged to provide a 128 h data set. Data-collection and refinement statistics are summarized in Table 1. For this comparative study, the diffraction limit was set to 1.75 Å resolution for all data sets.

2.4. Structure refinement

Neutron structure refinement of D-rubredoxin was performed using *phenix.refine* (Adams *et al.*, 2002; Afonine *et al.*, 2010). Neutron-density map examination and model building was performed in *Coot* (Emsley *et al.*, 2010). For each set of data the room-temperature X-ray structure of selectively H-labeled deuterated rubredoxin (PDB entry 3kyw; Weiss *et al.*, 2008; Gardberg *et al.*, 2010) was used as the starting model. D atoms, water molecules and the iron ion were stripped out from the initial model. The program *phenix.ready_set* was used to generate D atoms at all positions. Rigid-body refinement followed by cycles of individual positional refinement was performed. D-atom positions were refined individually (*i.e.* D-atom positions were not refined as a riding model). Group atomic displacement parameters (B factors) were then refined. The iron ion and water-molecule O atoms were built manually into the model by inspection of $F_o - F_c$ maps contoured at +3σ. *Phenix.ready_set* was used to add both water-molecule D atoms to the corresponding O atoms. Water molecules were manually oriented in *Coot* and *phenix.refine* was used to refine their B factors only. Water molecules showing $2F_o - F_c$ neutron density at 1σ were kept as D_2O molecules. Water molecules for which no significant $2F_o - F_c$ neutron density was observed or which presented negative $F_o - F_c$ neutron density were modeled as OD or O. Inspection of the $F_o - F_c$ negative peaks also revealed that up to six backbone amide deuterium positions (Trp3, Val4, Tyr10, Ile40 in the 14 h structure only, Phe48 and Glu49) were at least partially occupied by H atoms. The $F_o - F_c$ negative feature arises from the negative coherent neutron scattering length of hydrogen, whereas that of deuterium is positive (Meilleur, Myles *et al.*, 2006). These positions, which are involved in strong hydrogen-bond networks, exchanged to hydrogen during the heat-treatment step of the purification but did not exchange back to deuterium during the final deuterium-exchange steps or the crystallization process. This has been observed in previously determined rubredoxin neutron structures (Chatake *et al.*, 2004; Kurihara *et al.*, 2004). *Phenix.refine* was used to refine H and D occupancies at these positions. Rounds of solvent building, individual positional refinement of all protein atoms and group B-factor refinement of protein and solvent atoms were performed until the refinement converged.

2.5. Computation of D-OMIT maps and peak-height analyses

The concept of a composite D-OMIT map and its use has been described elsewhere (Gardberg *et al.*, 2010). D-OMIT maps were calculated for each residue by setting D-atom occupancies to zero one residue at a time. Neutron density peak heights of omitted D atoms that presented positive

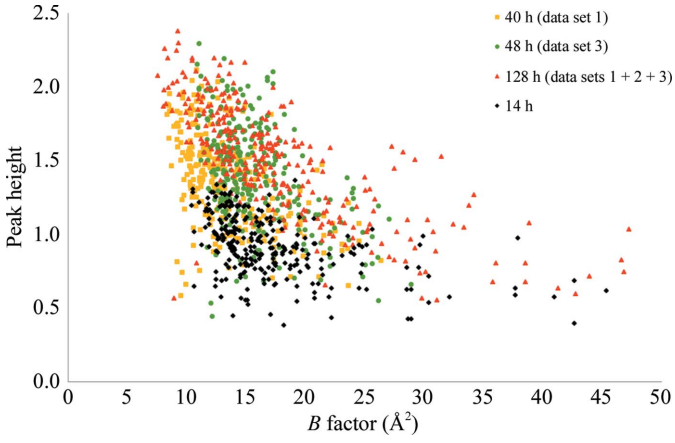

Figure 1
Distribution of peak heights at each D-atom position *versus* their *B* factors. The color codes and symbols for the four data sets are shown in the inset.

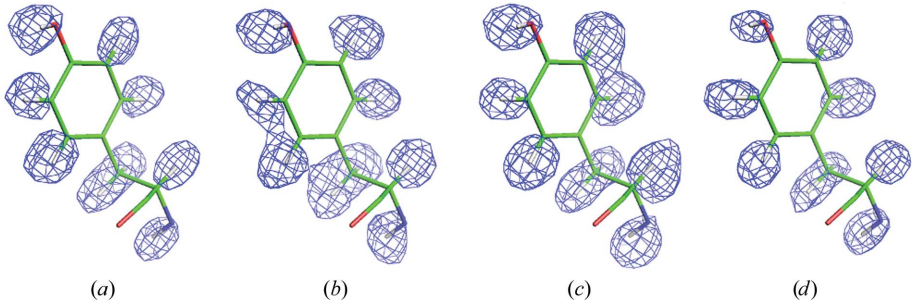

Figure 2
Neutron-density maps ($F_o - F_c$) at Tyr12. D-OMIT maps at 3σ for D-rubredoxin computed from (*a*) 40 h (data set 1), (*b*) 48 h (data set 3), (*c*) 128 h (data sets 1 + 2 + 3) and (*d*) 14 h data.

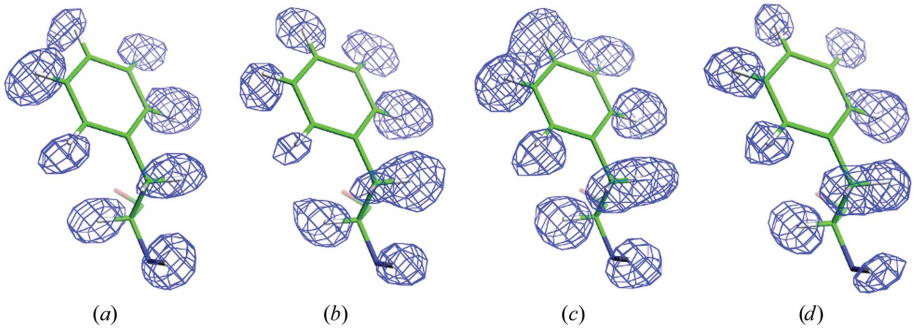

Figure 3
Neutron-density maps ($F_o - F_c$) at Phe29. D-OMIT maps at 3σ for D-rubredoxin computed from (*a*) 40 h (data set 1), (*b*) 48 h (data set 3), (*c*) 128 h (data sets 1 + 2 + 3) and (*d*) 14 h data.

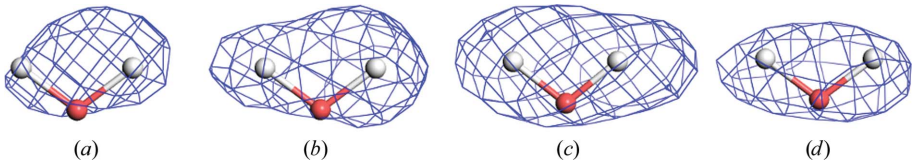

Figure 4
Representative neutron-density maps ($F_o - F_c$) at D_2O. D-OMIT maps at 3σ computed from (*a*) 40 h (data set 1), (*b*) 48 h (data set 3), (*c*) 128 h (data sets 1 + 2 + 3) and (*d*) 14 h data.

$F_o - F_c$ density were measured in *Coot* and plotted as a function of the associated *B* factors (Fig. 1).

3. Results

3.1. Data and structure quality assessment

Data-processing and structure-refinement statistics for all neutron data sets are presented in Table 1. As expected, the standard measures of diffraction quality [R_{merge} and mean $I/\sigma(I)$] are slightly improved with longer data-collection times. However, we note that the extended angular range and greater completeness and redundancy of the 14 h data set help to compensate in part for the shorter exposure times and that the precision-indicating $R_{p.i.m.}$ values that take account of redundancy are comparable across the 14, 40, 48 and 128 h data sets. Critically, we note that the standard quality metrics (R_{work} and R_{free}) for structure refinement compare well for the neutron structures refined against the 14 h data set and the 128 h data set.

3.2. Visibility of D atoms

For each data set, the D-OMIT maps were analysed and the neutron density present at each D-atom position was measured. Neutron-density maps ($F_o - F_c$ contoured at 3σ) from D-OMIT maps for Tyr12 and Phe29 and water molecules (as D_2O) are shown in Figs. 2, 3 and 4, respectively. Remarkably, the OMIT maps computed from all data sets appear similar when D atoms are visible.

To produce a complete analysis, we measured and plotted (Fig. 1) the peak heights of all visible D atoms for all four data sets using *Coot*. For each data set, we calculated the percentage of visible D atoms. The total number of H atoms (as D atoms in the case of perdeuterated PfRd) in PfRd is 371 (residues 52 and 53 were omitted from the analysis as they were disordered in our structures). The visibility statistics are presented in Table 2. We found that 269 D atoms (73%) were visible in the '14 h structure' D-OMIT maps, while the 128 h data set only revealed an additional 25 D atoms (6.7%). The fraction of D atoms visible in the 14 h data set is similar to that seen for the 40 and 48 h data sets. The occupancies of the D atoms that did not show positive $F_o - F_c$ peaks in the D-OMIT maps were set to 0.00 in the deposited PDB files, even when $2F_o - F_c$ positive neutron density could be observed. For comparison, we

Table 2
D-atom visibility summary, excluding D_2O molecules.

Data set	40 h (data set 1)	48 h (data set 3)	128 h (data sets 1 + 2 + 3)	14 h
Visible D	281 (76%)	286 (77%)	294 (79%)	269 (73%)
Average absolute value	1.4	1.4	1.5	0.9
Average r.m.s. (σ)	5.0	5.1	5.8	5.0
Average σ/absolute value	3.7	3.6	4.0	5.5
Average B (Å^2)	13.3	15.7	17.8	17.1

Table 3
Summary of water molecules.

Data set	Total No. of waters	D_2O	OD	O
40 h (data set 1)	35	13	2	20
48 h (data set 3)	33	12	10	11
128 h (data sets 1 + 2 + 3)	36	21	2	13
14 h	35	11	15	9

calculated D/H-OMIT maps for the deposited PfRD neutron structure (PDB entry 1vcx) refined to 1.5 Å resolution against monochromatic neutron data (Kurihara *et al.*, 2004). This data set was collected over 840 h (35 d) of neutron beam time. This structure revealed 313 D/H atoms, a marginal increase on the number of D atoms observed in the structures reported here.

The plotted peak heights show that the longer the data-collection time (*i.e.* the better the data statistics), the clearer the visibility of the D atoms (higher peak heights). This is expected, as are the reduced peak heights with increasing B factors (Table 2, Fig. 1). While a similar number of water molecules can be modeled in the 14 and 128 h data sets (Table 3), the 128 h data set does allow a significantly higher number of water molecules to be refined as full D_2O molecules (21 *versus* 11).

4. Discussion

We have analyzed the structure of perdeuterated rubredoxin against four neutron data sets collected in 14, 40, 48 and 128 h in order to assess how the completeness and statistical quality of the data impacts the qualitative and quantitative interpretation of the respective neutron maps. Our comparison of these structures is striking and important in showing that 269 D atoms can be located unambiguously in the protein using a 1.75 Å neutron data set collected in just 14 h. Furthermore, we show that increasing the data-collection time by almost an order of magnitude helps to visualize only an additional 25 D atoms in the 128 h neutron protein structure. Moreover, comparison with the 1.5 Å resolution PfRd neutron structure (PDB entry 1vcx) of Kurihara *et al.* (2004), which required a total of 840 h (35 d) of neutron beam time to complete, revealed 313 D/H atoms, a marginal increase on the number of D atoms observed in the structures reported here. For comparison, the 0.95 Å resolution X-ray structure of the same protein revealed only half that number of H atoms (Bau *et al.*, 1998).

As in all single-crystal diffraction, the choices made in neutron data collection balance the desired precision of the results against the diffracting power of the crystal, the brightness of the source, the efficiency of the instrumentation and the time that is required and available to complete the experiment. This study aimed to compare the information content of data sets collected over different reciprocal volumes for different data-collection times and to assess the impact of statistical quality, completeness and redundancy on the results. The 14 h data set provided a large 130° survey of reciprocal space with a minimal exposure time of just 1 h per frame. The two interleaving 40 h data sets and the third offset 48 h data sets aimed to provide more complete and accurate data sets that were restricted to the unique asymmetric unit of reciprocal space. The latter three data sets were then combined and merged to provide a composite 128 h data set that is typical of the data-collection protocols and times at the LADI-III instrument. As Table 1 shows and as can be expected, the standard data-quality statistics $R_{\rm merge}$, $R_{\rm p.i.m.}$ and mean $I/\sigma(I)$ improved overall as data-collection times increased. Again as expected, increasing the quality, completeness and redundancy of the data did improve the interpretability and quality of the neutron maps, but only marginally so: hence, whilst 269 D atoms and 35 waters were indentified in the 14 h structure, increasing the data-collection time by near-tenfold led to the identification of 294 D atoms and 36 waters. However, 21 of these waters could then be modeled as complete D_2O molecules, while only 11 could from the 14 h data set. Breaks in neutron-density maps are an acknowledged issue in structure refinement against neutron data at moderate resolutions (1.5–2.5 Å). Here, we noted six breaks in density at $C^{\alpha}-C$ bonds in the 14 h structure and just one at a $C^{\alpha}-C$ bond in the 128 h structure. The density breaks, which were limited to $C^{\alpha}-C$ bonds, did not prevent interpretation of the maps. The 14 h structure provides a wealth of information on the hydrogen-bonding interactions in the protein and the positions of all backbone amide D atoms are clearly revealed. Labile D atoms are clearly visible, allowing hydroxyl groups to be oriented experimentally, as illustrated in Fig. 2 for Tyr12. Critically, we note that refinement statistics are closely similar for the neutron structures refined against both the 14 and 128 h data sets.

More generally, our study emphasizes that data requirements for neutron refinement match those for X-ray atomic model refinement, which primarily requires data to the highest resolution that a crystal may reach rather than perfect data completeness and accuracy (Cruickshank, 1999; Blow, 2002; Dauter, 2010). While no single optimal data-collection strategy can be designed from this one comparative study, our analysis does suggest that the total time required for neutron data collection might generally be reduced by (i) choosing the shortest exposure time per frame that provides the highest acceptable resolution, rather than attempting to maximize counting statistics and data accuracy in an extended data-collection run, (ii) collecting data frames in large φ-steps that rapidly survey reciprocal space and quickly build data completeness and redundancy and (iii) refining and analysing the structure as data are being collected and monitoring the impact of additional frames of data to decide when and where

to stop data collection. It should be emphasized that refining a structure against neutron data is far less demanding in time than refining against ultrahigh-resolution X-ray data and can be performed to a great extent at the facility while accumulating data. In order to do so, a perdeuterated sample might be desirable as (i) perdeuterated samples deliver a near-tenfold improvement in the signal-to-noise ratio of the data, (ii) D atoms become visible at more modest resolutions (2.0–2.5 Å) than hydrogen (1.5–2.0 Å) (Shu *et al.*, 2000) and (iii) the resulting neutron maps do not suffer from cancellation effects or partial H/D exchange, both of which can complicate the interpretation of neutron maps (Shu *et al.*, 2000).

Whilst we recognize that deuterated rubredoxin is an unusually good candidate for neutron structure analysis, and that the results of our 14 h analysis are perhaps exceptional, we suggest that this approach should be considered and scaled appropriately (to 2 or 3 d of beam time) for the neutron analysis of other systems. Clearly, if the H or D atoms of interest to the investigator do not appear in such an analysis, or do not appear with the required precision, then this rapid data-collection approach can, as a minimum, serve to guide decisions on the viability and design of more extended data-collection strategies that would warrant the allocation of substantially more beam time at the facility.

5. Conclusion

Our results show that a rapid neutron data-collection strategy was able to unambiguously locate 73% of the D atoms in the crystal structure of perdeuterated PfRd (D-rubredoxin) after just 14 h of accumulated beam time. Increasing the total data-collection time by almost an order of magnitude revealed only 25 more D atoms in the protein structure. These results suggest that in favorable cases neutron analysis may be accelerated by data-collection protocols that exploit rapid high-resolution surveys of reciprocal space. Since the aim of most neutron protein diffraction analyses is to determine the hydration and protonation states of a system and is most frequently focused on just a few key residues at an active site or along a reaction pathway, then in the best case the information derived from a rapid survey may already answer in part the question being asked of the system. In the worst case, in which H or D atoms of interest cannot be visualized with the required precision in the neutron maps, then the results can help to guide decisions on the design, value and feasibility of additional experiments and of the investment of beam time that would be required. Finally, and more generally, this study emphasizes the value for all such experiments of refining and analysing structures as data are being collected at the facility, which enables data acquisition to be adjusted, optimized and then stopped when the required precision in structural information is met.

PM and the IMAGINE project are supported by an award from the National Science Foundation (Award 0922719) to FM. Oak Ridge National Laboratory is managed by UT-Battelle LLC for the US Department of Energy under contract No. DO-AC05-00OR22725. The work at ORNL was supported by the Division of Scientific User Facilities, DOE Basic Energy Sciences. This manuscript has been authored by UT-Battelle LLC under Contract No. DE-AC05-00OR22725 with the US Department of Energy.

References

Adachi, M. *et al.* (2009). *Proc. Natl Acad. Sci. USA*, **106**, 4641–4646.
Adams, P. D., Grosse-Kunstleve, R. W., Hung, L.-W., Ioerger, T. R., McCoy, A. J., Moriarty, N. W., Read, R. J., Sacchettini, J. C., Sauter, N. K. & Terwilliger, T. C. (2002). *Acta Cryst.* D**58**, 1948–1954.
Afonine, P. V., Mustyakimov, M., Grosse-Kunstleve, R. W., Moriarty, N. W., Langan, P. & Adams, P. D. (2010). *Acta Cryst.* D**66**, 1153–1163.
Arzt, S., Campbell, J. W., Harding, M. M., Hao, Q. & Helliwell, J. R. (1999). *J. Appl. Cryst.* **32**, 554–562.
Bau, R., Rees, D. C., Kurtz, D. M., Scott, R. A., Huang, H. S., Adams, M. W. W. & Eidsness, M. K. (1998). *J. Biol. Inorg. Chem.* **3**, 484–493.
Blakeley, M. P. (2009). *Crystallogr. Rev.* **66**, 1198–1205.
Blakeley, M. P., Kalb, A. J., Helliwell, J. R. & Myles, D. A. A. (2004). *Proc. Natl Acad. Sci. USA*, **101**, 16405–16410.
Blakeley, M. P., Langan, P., Niimura, N. & Podjarny, A. (2008). *Curr. Opin. Struct. Biol.* **18**, 593–600.
Blakeley, M. P., Mitschler, A., Hazemann, I., Meilleur, F., Myles, D. A. A. & Podjarny, A. (2006). *Eur. Biophys. J.* **35**, 577–583.
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.
Blakeley, M. P., Teixeira, S. C. M., Petit-Haertlein, I., Hazemann, I., Mitschler, A., Haertlein, M., Howard, E. & Podjarny, A. D. (2010). *Acta Cryst.* D**66**, 1198–1205.
Blow, D. M. (2002). *Acta Cryst.* D**58**, 792–797.
Budayova-Spano, M., Bonneté, F., Ferté, N., El Hajji, M., Meilleur, F., Blakeley, M. P. & Castro, B. (2006). *Acta Cryst.* F**62**, 306–309.
Campbell, J. W., Hao, Q., Harding, M. M., Nguti, N. D. & Wilkinson, C. (1998). *J. Appl. Cryst.* **31**, 496–502.
Chatake, T., Kurihara, K., Tanaka, I., Tsyba, I., Bau, R., Jenney, F. E., Adams, M. W. W. & Niimura, N. (2004). *Acta Cryst.* D**60**, 1364–1373.
Coates, L., Stoica, A. D., Hoffmann, C., Richards, J. & Cooper, R. (2010). *J. Appl. Cryst.* **43**, 570–577.
Cole, J. M., McIntyre, G. J., Lehmann, M. S., Myles, D. A. A., Wilkinson, C. & Howard, J. A. K. (2001). *Acta Cryst.* A**57**, 429–434.
Cruickshank, D. W. J. (1999). *Acta Cryst.* D**55**, 583–601.
Dauter, Z. (2010). *Acta Cryst.* D**66**, 389–392.
Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* D**66**, 486–501.
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.
Habash, J., Raftery, J., Nuttall, R., Price, H. J., Wilkinson, C., Kalb (Gilboa), A. J. & Helliwell, J. R. (2000). *Acta Cryst.* D**56**, 541–550.
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.
Howard, E. I., Blakeley, M. P., Haertlein, M., Petit-Haertlein, I., Mitschler, A., Fisher, S. J., Cousido-Siah, A., Salvay, A. G., Popov, A., Muller-Dieckmann, C., Petrova, T. & Podjarny, A. (2011). *J. Mol. Recognit.* **24**, 724–732.
Jenney, F. E. Jr & Adams, M. W. (2001). *Methods Enzymol.* **334**, 45–55.
Kovalevsky, A., Fisher, Z., Johnson, H., Mustyakimov, M., Waltman, M. J. & Langan, P. (2010). *Acta Cryst.* D**66**, 1206–1212.

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.
Kurihara, K., Tanaka, I., Chatake, T., Adams, M. W. W., Jenney, F. E., Moiseeva, N., Bau, R. & Niimura, N. (2004). *Proc. Natl Acad. Sci. USA*, **101**, 11215–11220.
Langan, P., Greene, G. & Schoenborn, B. P. (2004). *J. Appl. Cryst.* **37**, 24–31.
Meilleur, F., Dauvergne, M.-T., Schlichting, I. & Myles, D. A. A. (2005). *Acta Cryst.* D**61**, 539–544.
Meilleur, F., Myles, D. A. A. & Blakeley, M. P. (2006). *Eur. Biophys. J.* **35**, 611–620.
Meilleur, F., Snell, E. H., van der Woerd, M. J., Judge, R. A. & Myles, D. A. A. (2006). *Eur. Biophys. J.* **35**, 601–609.
Meilleur, F., Weiss, K. L. & Myles, D. A. A. (2009). *Methods Mol. Biol.* **544**, 281–292.
Myles, D. A. A. (2006). *Curr. Opin. Struct. Biol.* **16**, 630–637.
Myles, D. A. A., Bon, C., Langan, P., Cipriani, F., Castagna, J. C., Lehmann, M. S. & Wilkinson, C. (1997). *Physica B*, **241**, 1122–1130.
Schultz, A. J., Thiyagarajan, P., Hodges, J. P., Rehm, C., Myles, D. A. A., Langan, P. & Mesecar, A. D. (2005). *J. Appl. Cryst.* **38**, 964–974.
Shu, F., Ramakrishnan, V. & Schoenborn, B. P. (2000). *Proc. Natl Acad. Sci. USA*, **97**, 3872–3877.
Tanaka, I., Kurihara, K., Chatake, T. & Niimura, N. (2002). *J. Appl. Cryst.* **35**, 34–40.
Tanaka, I., Kusaka, K., Hosoya, T., Niimura, N., Ohhara, T., Kurihara, K., Yamada, T., Ohnishi, Y., Tomoyori, K. & Yokoyama, T. (2010). *Acta Cryst.* D**66**, 1194–1197.
Teixeira, S. C. *et al.* (2008). *Chem. Phys.* **345**, 133–151.
Tomanicek, S. J., Blakeley, M. P., Cooper, J., Chen, Y., Afonine, P. V. & Coates, L. (2010). *J. Mol. Biol.* **396**, 1070–1080.
Tomanicek, S. J., Wang, K. K., Weiss, K. L., Blakeley, M. P., Cooper, J., Chen, Y. & Coates, L. (2011). *FEBS Lett.* **585**, 364–368.
Wilkinson, C., Lehmann, M. S., Meilleur, F., Blakeley, M. P., Myles, D. A. A., Vogelmeier, S., Thoms, M., Walsh, M. & McIntyre, G. J. (2009). *J. Appl. Cryst.* **42**, 749–757.
Weiss, K. L., Meilleur, F., Blakeley, M. P. & Myles, D. A. A. (2008). *Acta Cryst.* F**64**, 537–540.
Winn, M. D. *et al.* (2011). *Acta Cryst.* D**67**, 235–242.
Yamaguchi, S., Kamikubo, H., Kurihara, K., Kuroki, R., Niimura, N., Shimizu, N., Yamazaki, Y. & Kataoka, M. (2009). *Proc. Natl Acad. Sci. USA*, **106**, 440–444.
Zikovsky, J., Peterson, P. F., Wang, X. P., Frost, M. & Hoffmann, C. (2011). *J. Appl. Cryst.* **44**, 418–423.