Lawrence Berkeley National Laboratory

Recent Work

Title

Solvent flows, conformation changes and lattice reordering in a cold protein crystal.

Permalink

<https://escholarship.org/uc/item/7214v1z7>

Journal

Acta crystallographica. Section D, Structural biology, 75(Pt 11)

ISSN

2059-7983

Authors

Moreau, David W
Atakisi, Hakan
Thorne, Robert E

Publication Date

2019-11-01

DOI

10.1107/s2059798319013822

Peer reviewed

Solvent flows, conformation changes and lattice reordering in a cold protein crystal

David W. Moreau, Hakan Atakisi and Robert E. Thorne*

Physics Department, Cornell University, Ithaca, NY 14853, USA. *Correspondence e-mail: ret6@cornell.edu

Received 6 March 2019

Accepted 10 October 2019

Edited by E. F. Garman, University of Oxford, England

Keywords: protein crystallography; solvent behaviour; cryocrystallography; thermal expansion; cold denaturation.

Supporting information: this article has supporting information at journals.iucr.org/d

When protein crystals are abruptly cooled, the unit-cell, protein and solvent-cavity volumes all contract, but the volume of bulk-like internal solvent may expand. Outflow of this solvent from the unit cell and its accumulation in defective interior crystal regions has been suggested as one cause of the large increase in crystal mosaicity on cooling. It is shown that when apoferritin crystals are abruptly cooled to temperatures between 220 and 260 K, the unit cell contracts, solvent is pushed out and the mosaicity grows. On temperature-dependent timescales of 10 to 200 s, the unit-cell and solvent-cavity volume then expand, solvent flows back in, and the mosaicity and *B* factor both drop. Expansion and reordering at fixed low temperature are associated with small-amplitude but large-scale changes in the conformation and packing of apoferritin. These results demonstrate that increases in mosaicity on cooling arise due to solvent flows out of or into the unit cell and to incomplete, arrested relaxation of protein conformation. They indicate a critical role for time in variable-temperature crystallographic studies, and the feasibility of probing interactions and cooperative conformational changes that underlie cold denaturation in the presence of liquid solvent at temperatures down to ~200 K.

1. Introduction

Protein crystallography remains the primary tool for determining atomic-resolution information about protein structure and function. Protein crystals are damaged by X-rays, and the rate of damage with X-ray dose drops dramatically on cooling from ~300 to 200 K, with a smaller additional drop on cooling to 100 K (Warkentin & Thorne, 2010*a*; Warkentin *et al.*, 2013). The ability to collect much more useful diffraction information per crystal and a convenient experimental infrastructure for handling cryocooled crystals have led to the near-total dominance of *T* = 100 K data collection in the last 20 years. However, increasing recognition of what is lost by cryocooling (Fraser *et al.*, 2011; Keedy *et al.*, 2015) and increasing competition from single-particle cryoelectron microscopy for the determination of initial structures (Vinothkumar & Henderson, 2016) will drive a large expansion of both room-temperature and variable/multi-temperature X-ray crystallography in the coming decade.

Although cryocooling reduces the amplitude of thermal atomic motions and often increases the achievable diffraction resolution, it increases nonthermal crystal disorder. Crystal mosaicities, a measure of lattice orientational order, increase from as little as a few thousands of a degree to tenths of a degree or more, the spread in unit-cell parameters measured in reciprocal-space mapping (Kriminski *et al.*, 2002) increases, and crystallographic *B* factors may also increase. This disorder

reduces diffraction peak-to-background ratios and limits the data-quality improvements that can be achieved by reducing the crystal rotation per frame (for example in fine φ slicing). It is particularly limiting in the study of biomolecular complexes and viruses, the crystals of which have large unit cells and large solvent cavities. These are targets for which single-particle cryoelectron microscopy has made strong inroads.

The same factors that lead to the improved resolution of cryocooled crystals also perturb the conformational ensemble of the protein away from its biologically relevant form. Minority side-chain conformations that may be important in molecular recognition or allosteric interactions are depopulated (Lang *et al.*, 2010; Fraser *et al.*, 2011; Fenwick *et al.*, 2014). New conformations may appear (Fraser *et al.*, 2011), especially for side chains involved in crystal contacts but also in non-contact solvent-exposed and buried regions (Atakisi *et al.*, 2018).

Variable/multi-temperature crystallography (Tilton *et al.*, 1992; Teeter *et al.*, 2001; Warkentin & Thorne, 2009, 2010b; Warkentin *et al.*, 2012) performed at temperatures near and above the protein–solvent glass transition (~ 200 K) has great promise as a tool for mapping the conformational ensemble and energy landscape of a protein and for identifying and understanding key interactions (Keedy *et al.*, 2015, 2018). The key obstacles to such studies have been the formation of diffraction-destroying crystalline ice and the need to use large, crystal-perturbing cryoprotectant concentrations to prevent it. We have recently shown that ice formation inside protein crystals is dramatically suppressed by nanoconfinement (Moreau *et al.*, 2019). Following abrupt cooling, the internal solvent in cryoprotectant-free protein crystals with solvent

cavities as large as ~ 70 Å can be maintained in a (super-cooled) liquid state at temperatures down to 200 K for a time sufficient to collect a complete data set.

Here, we report the changes in unit-cell and solvent-cavity volume, protein conformation and packing, and crystal order that occur at constant temperature *after* apoferritin crystals are cooled to temperatures between 220 and 260 K. These results illuminate the roles of solvent transport and of protein relaxations on different timescales in generating cooling-induced disorder, show how this disorder can be reduced and suggest the feasibility of studying cooperative conformational changes, including those associated with cold denaturation at temperatures down to ~ 200 K.

2. Materials and methods

2.1. Structure of cubic apoferritin crystals

This study focused on cubic crystals of apoferritin, which were chosen because of their very large solvent cavities and large bulk-like solvent fraction. The atomic structure of cubic apoferritin has been extensively studied at both room and cryogenic temperatures. Ferritin monomers (Fig. 1*a*) are composed of five α -helices labelled A–E (Hempstead *et al.*, 1997) that are roughly aligned within a volume of $45 \times 20 \times 20$ Å. These monomers readily form dimers (Fig. 1*b*), with contacts formed between the A and B helices and along the BC loop. The interface between dimers is highly diverse and is stabilized by hydrophobic interactions, hydrogen bonding between the monomers, and bound solvent. 12 dimers, having a total molecular weight of 476 kDa, assemble into a nearly

Figure 1

(*a*) A ferritin monomer, comprised of five helices and a long flexible loop. (*b*) A ferritin dimer. 12 of these dimers assemble into a spherical shell with an internal cavity of diameter 68 Å. In the apo form this cavity is filled with solvent. (*c*) In the crystal, the interface between spherical shells involves contacts between the BC loops of adjacent dimers, mediated through Cd atoms. (*d*) Coordinate reference directions and (*e*) the distances described in the text and used to characterize structural changes during cooling and cold unit-cell expansion.

spherical shell having an internal cavity of 68 Å in diameter for the storage of iron. In the apo form this cavity is filled with solvent. In cubic apoferritin crystals these solvent cavities have a face-centred cubic arrangement. These cavities are larger than those found in roughly 98% of PDB-deposited protein crystal structures (Moreau *et al.*, 2019). A second set of large cavities lies between the spherical shells, having a characteristic size of 65 Å (the distance between the second nearest-neighbour shells) and together containing roughly 65% of the total solvent volume. In addition to these large cavities inside and between the spherical shells, eight hydrophilic channels of 6 Å in length and 3.4 Å in diameter (providing conduits for water, ions and glycerol when the crystals are soaked or cooled) and six hydrophobic channels, each

12.5 Å in length and 2 Å in diameter, pass through each spherical protein shell.

The spherical shells contact each other at the dimer interface along the BC loops (Fig. 1c). Residues Asp80 and Gln82 protrude from the BC loop and connect to a Cd atom between the shells. Asp80 has well defined electron density and is always involved in the contact. Gln82 is found in either an 'in' or 'out' conformation in a given crystal (Supplementary Fig. S1); the specific conformation may depend on the pH of the mother liquor or soaking solution (de Val *et al.*, 2012). In the 'in' conformation Gln82 is connected to the Cd atom by well defined electron density. In the 'out' conformation Gln82 protrudes outward into the solvent cavity and an additional water molecule coordinates to the Cd atom. The distance between the spherical shells was slightly different in crystals with each of these conformations. To quantify the effects of temperature and glycerol concentration, crystals exhibiting the 'in' and 'out' conformations were separately analysed.

2.2. Protein crystallization and harvesting

Cubic crystals of equine spleen apoferritin (Sigma, catalog No. A-3641) were grown by the hanging-drop vapour-diffusion method. Drops comprising of 2 µl protein solution, containing 10 mg ml⁻¹ apoferritin in 0.1 M sodium acetate buffer pH 6.5, and 2 µl of a well solution consisting of 2% (w/v) cadmium sulfate and 15% (w/v) (~1.1 M) ammonium sulfate in the same buffer were equilibrated against well solution in 24-well plates. Cubic crystals in space group *F*432 with room-temperature unit-cell dimensions of $a = b = c = 183.5$ Å and sizes of 300–500 µm were obtained within a week. The crystals had a solvent content of 63% (v/v) as derived from the Matthews coefficient.

Crystals were used as grown, or else were soaked in solutions containing 10, 20 or 40% (v/v) glycerol for at least 5 min. The crystals were then transferred to a separate drop of NVH oil (Cargille) and manipulated until all external solvent was removed from their surface, as indicated by the near-disappearance of the crystal owing to the close match between its refractive index and that of the oil (Warkentin & Thorne, 2010b). Crystals were mounted on MicroGrippers (MiTeGen) in a spherical blob of NVH oil to prevent dehydration during data collection; the fingers of the MicroGrippers helped to immobilize both the oil and the crystal. Mounted crystals were stored in MicroRT tubes (MiTeGen) containing mother liquor or cryoprotectant solution for ~1 h prior to data collection. Although the apoferritin crystallization solutions and the solvent present in the crystal cavities contain significant salt concentrations, these only reduce the freezing temperature of the solution by 4°C (Clegg, 1995). Ice formation is dominated by the effects of nanoconfinement by the protein lattice (Moreau *et al.*, 2019) and, at the concentrations of 10% (v/v) and larger used here, by glycerol.

2.3. X-ray data collection

As was discussed in the supporting information to Moreau *et al.* (2019), time-dependent X-ray diffraction data were

collected on the F1 beamline at the Cornell High-Energy Synchrotron Source (CHESS) using the experimental configuration shown in Supplementary Fig. S2. Samples were illuminated using a Gaussian beam with full width at half maximum (FWHM) values in the horizontal and vertical directions of 63 and 69 µm, respectively, the tails of which were clipped using a 100 µm aperture placed 2 cm upstream of the sample. The beam had a divergence of 0.015°, a photon energy of 12.7 keV and a flux of 2.2×10^9 photons s⁻¹. Dose rates were ~2 kGy s⁻¹. Diffraction patterns were recorded using a PILATUS 6M detector with a frame rate of 5 Hz.

The sample temperature was controlled using a nitrogen-gas stream generated by an Oxford Cryosystems Cryostream 700 nitrogen-gas cryocooler. The gas flow rate was 5 l min⁻¹, corresponding to an ~1 m s⁻¹ velocity through a 1 cm diameter aperture. The temperature of the gas stream was varied between 180 and 260 K using the internal heater of the cryocooler and was monitored using both the internal temperature sensor of the cryocooler and a thermocouple that was periodically placed at the sample position.

Before X-ray data collection, with the cryostream retracted and blocked, a sample in its MicroRT tube was manually placed on the beamline goniometer stage, the crystal was allowed to settle in the oil and the crystal was then translated into the X-ray beam path. The MicroRT tube was then removed and automated data collection was initiated using the *ADX* software at the beamline. In a typical experiment, an initial set of ten frames (0.5° sample rotation and 0.2 s exposure time per frame, 5° rotation and 2 s total per exposure) was collected at room temperature. The crystal was then returned to its initial orientation, the cryostream head was extended and unblocked and a single set of 200 frames (0.5° sample rotation and 0.1 or 0.2 s exposure time per frame, 100° rotation and 20–40 s total per exposure) was acquired. With a dose rate of ~2 kGy s⁻¹, the total doses ranged from ~40 to ~100 kGy, which was much less than the half dose at all temperatures studied, so changes in the diffraction properties with time were dominated by effects other than radiation damage. Sample-cooling rates, as monitored using the unit cell of the crystal, varied with sample (crystal + surrounding oil) volume, with maximum cooling rates of ~300 K s⁻¹ and cooling times to the final sample temperature in the range 0.5–2 s. As illustrated in Fig. 2, the time required for each crystal to cool and the stability of its final temperature were verified by monitoring the positions of the strong, broad peak near 5.7 Å owing to scattering from the surrounding NVH oil. Temperature versus time at the sample position (with the sample removed) was also measured using a 250 µm thermocouple (Supplementary Fig. S3; Moreau *et al.*, 2019). All three measurements gave consistent results.

A total of 142 crystals of apoferritin were examined at CHESS on eight different dates between November 2015 and April 2018. The measurement of a large number of samples was essential to draw robust conclusions because of variations in the extent to which external solvent was removed, the stochastic nature of ice formation within supercooled internal crystal solvent, and because the sample response depended to

Figure 2

(a) Images of glycerol-free apoferritin crystals in NVH oil at (left) 220 K and (right) 240 K, for which unit-cell, mosaicity and Wilson B -factor data are shown in Fig. 5. The images were taken at the end of data collection. The crystals are ice-free and appear clear. When internal ice forms, the crystals become darker and opaque. (b) Diffraction patterns recorded from the apoferritin crystals in (a) at room temperature and at the indicated times after the start of cooling to 220 K (left) and 240 K (right). Azimuthally averaged diffuse diffraction intensity from the patterns in (b) after Bragg peaks from the protein lattice have been deleted, leaving diffuse scatter primarily generated from the NVH oil surrounding the crystal and also from internal solvent. (c) Azimuthally averaged diffuse diffraction intensity from the patterns in (b). Peaks from NVH oil and from internal solvent undergo only very small shifts in position during cooling, indicating small changes in density. No evidence for ice or other major changes in solvent structure (for example as might occur in a transition between low-density and high-density amorphous ice) is observed. (d) Position of the diffraction peak from NVH oil versus time during and following cooling. The peak position monotonically increases from its room-temperature position (indicated by the dashed green line) to its final position at the final temperature. The direction and size of the shifts on the left and right are consistent with an increase in the density of the oil with decreasing temperature. The timescale of this increase gives the timescale for crystal cooling. The relatively weak diffraction from internal solvent made the precise tracking of its peak position with time/temperature difficult.

some extent on the cooling time, which in turn depended on the sample size.

2.4. Processing of protein lattice diffraction

The diffraction frames were indexed, integrated and scaled using *XDS* (Kabsch, 2010), with an input file acquired from the MacCHESS website and modified for use with the PILATUS 6M detector. The resolution and *XDS*-determined mosaicity at room temperature, averaged over all crystals, were 2.2 ± 0.3 Å and $0.06 \pm 0.02^\circ$, respectively. The Bragg peaks were thus strong and well separated, so data processing using a single initial input file generally proceeded with few issues.

Data sets were processed in segments of five frames. The refined outputs for a segment, consisting of unit cell, beam centre and sample-to-detector distance, were used as inputs in the *XDS*.INP file when processing subsequent segments. As a check on these results, the data were also processed using *XDS* in segments of two and ten frames, and using *HKL-2000*. The use of five-frame segments balanced time resolution with variance in parameter estimates.

Unit-cell values were taken directly from the *XDS* CORRECT.LP output file. Wilson *B* factors were estimated as half the slope of a linear fit to the natural logarithm of the Bragg peak intensities (obtained using *phenix.merging_stats*; Adams *et al.*, 2010) versus $(\sin \theta / \lambda)^2$. Only Bragg peaks with resolutions numerically smaller than 4 Å and worse than a cutoff resolution where $I/\sigma > 2$, typically ~ 2 Å and no larger than 3 Å, were used to determine *B* factors.

The refined crystal mosaicities, calculated using *XDS* as the standard deviation of a Gaussian, had a 'floor' of 0.045° . This was imposed by *XDS* and did not reflect the actual crystal mosaicities. Diffraction frames were also indexed, integrated and scaled using *HKL-2000*, which had a mosaicity 'floor' close to the refined beam divergence, which was estimated by *XDS* to be 0.015° . The mosaicity values reported here were calculated using values from the *HKL-2000* *.x integration files, converted from FWHM to standard deviation to allow comparison with *XDS* results.

2.5. Protein structure modelling and refinement

The data sets used for structure determination and analysis all had resolutions of ≤ 2.1 Å and an average resolution of 1.96 Å, where the resolution cutoff was set so that the highest resolution shell had at least 98% completeness and $I/\sigma > 2$. Molecular replacement and model refinement were performed in *Phenix* (Adams *et al.*, 2010) using PDB entry 3f32 (Vedula *et al.*, 2009) as a starting model for all crystals/temperatures/glycerol concentrations. An initial refinement cycle using *phenix.refine* was performed with simulated-annealing, rigid-body, real-space, xyz coordinates and individual *B* factor options. In *Coot* (Emsley *et al.*, 2010), these models were checked for large peaks in the difference maps, Ramachandran outliers, rotamer outliers and regions of poor geometry. Cd atoms located at the interface of the spherical shells and residues in a partially disordered loop, Gly155 and Ser156,

were added to the model. A second round of refinement was performed to identify and add ordered solvent. These models were checked residue by residue for errors and alternative conformations were added. A third round of refinement including occupancies and *B* factors for alternative conformers and target weights was performed. The placement of the water molecules at the interface of the spherical shells were checked for accuracy and a fourth round of refinement was performed with the same parameters as the third round. Final model validation was performed using *MolProbity* in the *Phenix* software package (Chen *et al.*, 2010).

At room temperature, only 5° of data were collected from each crystal, so no single crystal gave enough data for structure determination. Data from multiple room-temperature crystals, collected on the same day, were merged using *XSCALE* to create four, two, two and one structures at 0, 10, 20 and 40% (v/v) glycerol, respectively.

2.6. Protein and solvent volume calculations

The volume of protein within the unit cell, $v_{\text{protein}}(T)$, was calculated as the volume enclosed by the solvent-excluded surface (SES) using a custom 'ball-rolling' algorithm very similar to that implemented by the program *3vee* (Voss & Gerstein, 2010). In *3vee*, the SES is calculated by 'rolling' a fixed-diameter probe over the surface of a voxelized 3D model of the protein, with atoms modelled as spheres having radii matching their van der Waals radii (Li & Nussinov, 1998). Our program was designed to also find the volume for multiple copies of the protein extended to fill the unit cell, accounting for crystal contacts and periodicity. Additionally, it uses probe radii of 1.4 and 1.7 Å for polar and apolar atoms, respectively (Li & Nussinov, 1998). The program divided the unit cell into voxels of dimension 0.15 Å and recorded all regions in which the probe clashes with an atom as part of the volume of the protein.

Solvent-cavity volumes $v_{\text{cavity}}(T)$ were determined by subtracting the protein volumes $v_{\text{protein}}(T)$ from the unit-cell volumes $v_{\text{cell}}(T)$. At room temperature, all crystals were highly ordered with very small mosaicities. Consequently, the solvent volume within a crystal could be assumed to be fully contained within the solvent cavities of the crystallographically determined unit cell, and the solvent-volume fraction of the crystal was equal to the volume fraction of the unit cell occupied by cavities,

$$f_{\text{solvent}}(300 \text{ K}) = \frac{v_{\text{cavity}}(300 \text{ K})}{v_{\text{cell}}(300 \text{ K})}. \quad (1)$$

We also separately calculated the solvent-cavity volumes located inside and outside the spherical protein shell. Using the voxel grid created with our ball-rolling method, we calculated the fraction of voxels belonging to solvent enclosed within thin spherical shells concentric with the apoferritin shell as a function of the shell radius, starting at radii much smaller than that of the apoferritin shell. For the radius at which this solvent-volume fraction reached a minimum, the enclosed solvent volume was taken as the solvent cavity

volume located inside the spherical protein shell. The volume outside the spherical protein shell was then obtained by subtracting this from the total solvent-cavity volume in the unit cell.

Crystals with Gln82 in both the 'in' and 'out' conformations as discussed in Section 2.1 were observed in our experiments, with all crystals measured on a given date having the same conformation. Because the two conformations gave slightly different unit cells and subtle differences in backbone, we only compared structures with the same conformation.

2.7. Protein structure analysis

To explore the structural changes to apoferritin during the initial unit-cell contraction on cooling and during subsequent cold expansion, models of the ferritin monomers were constructed using data sets from 29 glycerol-free apoferritin crystals at temperatures between 180 and 300 K, including four initial room-temperature data sets, the 16 data sets between 180 and 260 K where the unit cell contracted on cooling but did not expand appreciably during the data-acquisition time, and nine data sets collected from crystals at temperatures between 220 and 260 K that were deemed to have fully expanded by a plateauing of their unit cell in time and a drop in the *HKL*-2000 mosaicity to near the beam-divergence limit. One crystal measured at $T = 220$ K, where the time for expansion was longer, allowed separate structures to be determined after initial cooling while the unit cell remained close to its minimum cold value (having expanded by less than 19% of the total expansion) and then after expansion by at least 74%. The modelled structures were overlaid and compared residue by residue for differences within the protein monomers, in the monomer packing and in the interaction between the apoferritin shells. No significant differences, either in the backbone structure or side-chain conformations with corresponding differences in observable electron density, were observed between any structure, other than the previously discussed variation to Gln82.

To characterize more subtle shifts in protein structure and packing during cooling and cold expansion, several quantities shown in Figs. 1(*d*) and 1(*e*) were evaluated from the positions of the backbone C, N and C α atoms of the refined models. Before this analysis, residues 1–3 and 155–158 were removed from the models owing to their ill-defined electron density. The refined models were reduced to a single conformer and superposed to the same location in the unit cell by symmetry operations.

In order to characterize the changes in protein structure and packing during cooling and cold expansion, several quantities shown in Figs. 1(*d*) and 1(*e*) were evaluated from the positions of the backbone C, N and C α atoms in the refined models. The centre of volume of the protein shell is a known location in the unit cell. The centre of volume of the monomer was estimated as the average position of its atoms, $\langle \mathbf{r} \rangle$. The radial distance between the shell centre and a monomer centre was defined as Δ_{radial} , and the unit axis connecting these points as $\mathbf{r}_{\text{radial}}$. A second protein monomer was added in the dimer symmetry

location. The distance between the centres of volume of the monomers was defined as Δ_{dimer} and the unit axis connecting these points as $\mathbf{r}_{\text{dimer}}$. This axis is roughly tangential to the surface of the protein shell. A third axis is defined perpendicular to these axes, $\mathbf{r}_\alpha = \mathbf{r}_{\text{radial}} \times \mathbf{r}_{\text{dimer}}$, and is roughly oriented along the long α -helices of the monomer and tangential to the shell. These axes are shown relative to the apoferritin dimer in Fig. 1(*d*). A second dimer belonging to the adjacent protein shell and in contact with the first dimer was added at the symmetry location. The centre of volume of each dimer was estimated, and the distance between their centres Δ_{tetramer} was interpreted as the distance between protein shells.

The radius of gyration of a monomer was estimated from the distances of its backbone atoms from the centre of volume of the monomer,

$$R_g = \left[\frac{1}{n} \sum_{i=1}^n (\mathbf{r}_i - \langle \mathbf{r} \rangle)^2 \right]^{1/2}, \quad (2)$$

where $\mathbf{r}_i$ is the position of the i th backbone atom. To evaluate how the shape of the protein changed, a radius of gyration along each of the axes defined above was estimated as

$$R_{g,\text{axis}} = \left\{ \frac{1}{n} \sum_{i=1}^n [(\mathbf{r}_i - \langle \mathbf{r} \rangle) \cdot \mathbf{r}_{\text{axis}}]^2 \right\}^{1/2}. \quad (3)$$

To image the structural changes, two structures at room temperature, three structures at 220 K that did not show post-cooling expansion and four structures at 220 K that showed post-cooling expansion were converted to a set of average structures, one for each condition, by averaging the backbone-atom coordinates of the structures determined at that condition,

$$\{\mathbf{r}_i\}^{\text{condition}} = \frac{1}{m} \sum_{j=1}^m \mathbf{r}_i^j. \quad (4)$$

The index j represents structures obtained from different data sets acquired under the same conditions. The scalar displacement $\varepsilon_i^{a,b}$ of backbone atom i between averaged structures at conditions a and b was calculated as

$$\varepsilon_i^{a,b} = [(r_{i,x}^a - r_{i,x}^b)^2 + (r_{i,y}^a - r_{i,y}^b)^2 + (r_{i,z}^a - r_{i,z}^b)^2]^{1/2}. \quad (5)$$

This was normalized by subtracting the average displacement of all backbone atoms,

$$\langle \varepsilon^{a,b} \rangle = \frac{1}{n} \sum_{i=1}^n \varepsilon_i^{a,b}, \quad (6)$$

and dividing by the square root of the sum of variances σ_i^2 for each condition,

$$\sigma_i^2 = \frac{1}{m} \sum_{j=1}^m (r_{i,x}^j - \langle r_{i,x} \rangle)^2 + (r_{i,y}^j - \langle r_{i,y} \rangle)^2 + (r_{i,z}^j - \langle r_{i,z} \rangle)^2, \quad (7)$$

to obtain

$$\widehat{\varepsilon}_i^{a,b} = \frac{\varepsilon_i^{a,b} - \langle \varepsilon^{a,b} \rangle}{[(\sigma_i^a)^2 + (\sigma_i^b)^2]^{1/2}}. \quad (8)$$

Normalizing in this manner pulls out the more significant displacements by downweighting atoms that might show large variances owing to imprecise placement from weak electron density.

3. Results and analysis

3.1. Unit-cell, protein and solvent-cavity contraction on cooling

Fig. 3 replots our previously reported data (Moreau *et al.*, 2019) on how the unit-cell volume, protein volume and solvent-cavity volume of apoferritin crystals, determined immediately after cooling has completed, vary with temperature and glycerol concentration. All three contract on cooling. The protein volume shows a fractional contraction that is 2–3 times smaller than that of the unit cell, so the fractional contraction of the solvent-cavity volume is nearly double that of the unit cell. Addition of glycerol increases the unit-cell and solvent-cavity volumes at room temperature, and has a small effect on the amount that these contract on cooling from room

temperature to 220 K. In contrast, glycerol has a large effect on the contraction of bulk aqueous solutions. Between 300 and 77 K, pure water expands by 6% (Loerting *et al.*, 2011) and a 40%(v/v) glycerol solution contracts by $\sim 6\%$ (Shen *et al.*, 2016) as they cool into amorphous ice. Between 300 and 240 K liquid water expands by 2% (Hare & Sorensen, 1987) and a 40%(v/v) glycerol solution contracts by 2% (Glycerine Producers Association, 1963). This suggests that solvent must be expelled from or flow into the unit cell to account for the difference between the solvent-cavity and solvent contractions (Juers & Matthews, 2001, 2004; Kriminski *et al.*, 2002; Tyree *et al.*, 2018; Juers *et al.*, 2018). As shown in Supplementary Fig. S3, for glycerol-free crystals and crystals soaked in 40%(v/v) glycerol, 1.7% and 0.9% of the solvent must exit the unit cell on cooling to 240 K. These values assume that solvent in the first hydration shell has the same contraction as the protein and that the remaining solvent has the solvent contraction of the bulk solution. Other reasonable assumptions about the contraction of solvent in the first hydration shell yield similar results.

Fig. 4 shows how crystal mosaicity (Fig. 4a) and Wilson *B* factor (Fig. 4b), determined immediately after cooling has completed, vary with final temperature and with glycerol concentration. For glycerol-free crystals, the mosaicity determined by HKL-2000 remains near the ‘floor’ value of $\sim 0.03^\circ$, which is set by the incident beam divergence, on cooling to temperatures as low as 240 K. On cooling to temperatures of

Figure 3 Absolute values and percentage changes from room temperature, respectively, of unit-cell volume (a, b), protein volume (c, d) and solvent-cavity volume (e, f) as a function of temperature and glycerol concentration for cubic apoferritin crystals. For $220 \leq T \leq 260$ K, where unit-cell volumes increased on long timescales after cooling, minimum unit-cell volumes measured a few seconds after cooling had completed are plotted. Protein volume within the unit cell was deduced from full structural models by evaluating the solvent-excluded surface of the protein, and the total solvent-cavity volume was obtained by subtracting this volume from the unit-cell volume. Error bars indicate cell variations between crystals prepared and measured under nominally identical conditions. The number of crystals examined for each condition is given in Supplementary Table S1. Adapted from Fig. 5 of Moreau *et al.* (2019).

Figure 4 (a) Mosaicity and (b) Wilson *B* factor, determined using data acquired 3–5 s after cooling, versus temperature and glycerol concentration for ice-free cubic apoferritin crystals. Mosaicities have a ‘floor’ determined by the incident X-ray beam divergence of $\sim 0.03^\circ$. Actual room-temperature crystal mosaicities are likely to be $\sim 0.01^\circ$ or less. The significant scatter in values arises from crystal-size-related variations in illuminated volume and background scatter. Adapted from Supplementary Fig. S11 in the supporting information to Moreau *et al.* (2019).

180 and 200 K, the mosaicities increase to $\sim 0.15\text{--}0.2^\circ$. Wilson B factors show at most a small decrease with decreasing temperature down to 180 K. Both mosaicities and Wilson B factors tend to be somewhat lower for glycerol-free crystals than for those soaked in 20 and 40%(v/v) glycerol at temperatures above ~ 200 K.

3.2. Time-dependent unit cells, mosaicity and Wilson B factors in cold crystals

When crystals are cooled to ~ 140 K or below, vitrified internal solvent immobilizes the protein atoms and the crystal diffraction properties are (in the absence of radiation damage) independent of time. As crystals are warmed from ~ 140 to ~ 190 K, solvent atoms develop increasing diffusive mobility and ice crystals grow (Weik, Kryger *et al.*, 2001). Ice formed within the solvent cavities of the unit cell can cause unit-cell expansion. Ice formed at crystal surfaces draws water from the rest of the crystal for its growth, and can cause unit cells to shrink (Juers *et al.*, 2018), just as ice formed in intercellular spaces leads to cellular dehydration in biological cryopreservation (Fahy & Wowk, 2015).

Figure 5
Unit-cell volume, mosaicity and Wilson B factor during and following cooling from room temperature to (a) 220 K and (b) 240 K, measured for the two glycerol-free apoferritin crystals shown in Fig. 2 that remained ice-free throughout. The unit-cell volume decreases sharply during the initial cooling transient and then expands on a timescale that increases with decreasing temperature. Similarly, the mosaicity increases and then decreases. Dashed blue, green and red lines indicate the room-temperature unit-cell volume, mosaicity and B factor, respectively. The dashed vertical lines indicate the time at which the NVH oil peak position reached 95% of its final value; at larger times, the sample temperature was constant to within a few degrees.

How do crystal diffraction and unit-cell volume evolve at *constant* temperature following abrupt cooling under conditions where the internal solvent remains liquid and ice does not form? On cooling, the unit-cell volume initially drops and the mosaicity grows, on the timescale of the cooling transient ($\sim 0.2\text{--}2$ s). As illustrated by the examples shown in Fig. 5, for roughly 70% of ice-free apoferritin crystals at temperatures between 220 and 260 K, after the final temperature was reached the unit cells then expanded, by an average of $\sim 0.8\%$ at 220 K for glycerol-free crystals and $\sim 2.4\%$ at 240 K for crystals soaked in 40%(v/v) glycerol. As shown in Fig. 6, the final unit-cell volumes often exceeded their room-temperature values. As shown in Fig. 7, the unit-cell expansion timescale, obtained from an exponential fit, tended to increase with decreasing temperature and increasing glycerol concentration. For glycerol concentrations of 0, 10 and 20%(v/v) but not 40%(v/v), cell expansion was often accompanied by major improvements in crystal order, as reflected in large drops in crystal mosaicity (Figs. 5 and 8) and sometimes also substantial

Figure 6
Change in unit-cell volume relative to room temperature for apoferritin crystals that were slowly cooled at $\sim 0.1\text{ K s}^{-1}$ (solid lines and small symbols) and that were abruptly cooled (in <1 s) to each temperature (large symbols) versus temperature. Closed and open symbols indicate the unit-cell volumes measured just after cooling had completed and after the completion of unit-cell expansion, respectively. For all glycerol concentrations, the expanded unit-cell volumes approach the slow-cooled volumes, which between 220 and 250 K were generally larger than the room-temperature volumes. The slow-cooled data end at the lowest temperature at which ice diffraction was not observed.

decreases in Wilson B factors (Fig. 5). As discussed later, these diffraction-quality improvements in oil-encased crystals held at fixed low temperature in dry nitrogen-gas cryostreams are unrelated to those seen in 'cryoannealing', where crystals are briefly warmed (typically to room temperature) so that the external solvent melts before recooling. No appreciable cell expansion or diffraction-quality changes were observed at 180 or 200 K on timescales of 200 s, mostly like because of diverging protein relaxation timescales as the protein–solvent glass transition was approached (Ringe & Petsko, 2003; Doster, 2010; Weik, Ravelli *et al.*, 2001; Warkentin & Thorne, 2010a). In contrast to these results for apoferritin, similar measurements on thaumatin crystals (Supplementary Fig. S4) showed no low-temperature unit-cell expansion and/or crystal re-ordering on <200 s timescales at any temperature between 180 and 260 K.

3.3. Unit cells and mosaicities obtained during slow cooling

To gain insight into the possible causes of low-temperature unit-cell expansion in apoferritin, diffraction data were acquired as apoferritin crystals were slowly cooled at 0.1 K s^{-1} to 200 K (requiring ~ 1000 s) by a programmed ramp of the nitrogen-gas stream temperature. This very slow cooling allowed a more complete approach towards equilibrium at each temperature. Over the temperature range in which each crystal remained free of ice diffraction, the crystal mosaicities determined by *HKL-2000* remained at their room-temperature (beam-divergence limited) values. The solid lines in Fig. 6 show the change in the unit-cell volume of apoferritin relative to its room-temperature value over this temperature range. For all glycerol concentrations, the initial unit cells obtained after abrupt cooling (closed symbols) are smaller than those

Figure 7

Timescale for unit-cell expansion at constant temperature following abrupt cooling versus temperature and glycerol concentration. Error bars indicate the standard deviation of values obtained from fits to crystals at each temperature/concentration. The expansion timescale τ was obtained from a fit to the unit-cell data of the form $V_{\text{cell}}(t) = [V_{\text{cell}}(t_{\text{final}}) - V_{\text{cell}}(t_{\text{initial}})]\{1 - \exp[-(t - t_{\text{initial}})/\tau]\} + V_{\text{cell}}(t_{\text{initial}})$, where t_{initial} is the time at which the expansion began.

obtained by slowly cooling, whereas the unit cells measured after expansion at fixed temperature (open symbols) closely match those obtained by slowly cooling, for temperatures between 230 and 260 K. This suggests that the expanded state more nearly approximates the equilibrium state of the protein and crystal at each temperature.

For all glycerol concentrations, the expanded and slow-cooled unit cells both show little variation with temperature between 220–235 and 300 K, even though there is a strong decrease over this temperature range in the initial unit cell after abrupt cooling.

3.4. Analysis of cooling and post-cooling structural changes

The structural changes to apoferritin that occurred as the unit cell contracted during cooling and expanded after cooling were subtle, and were somewhat challenging to confidently interpret given the resolution of the data sets analysed. We focused on structures obtained using glycerol-free crystals, as these tended to have the highest resolutions and the lowest

Figure 8

Mosaicity before (closed symbols) and after (open symbols) unit-cell expansion at constant temperature versus temperature. For glycerol-free crystals and crystals soaked in solutions containing 10 and 20% (v/v) glycerol, the mosaicity drops dramatically during expansion toward the 0.03° floor (horizontal dashed line) set by the incident beam divergence. For crystals soaked in 40% (v/v) glycerol, the mosaicity tended to become worse during expansion. The small symbols and solid lines indicate the mosaicity measured when the crystal was slowly cooled at 0.1 K s^{-1} .

mosaicities. We separately analysed crystals in which the 'in' and 'out' conformations of Gln82 dominated, as this conformational difference substantially affected some of the results. Supplementary Table S4 lists the number of crystals used for the structural analysis at each condition, and a supplementary spreadsheet provides refinement statistics for all analysed crystal structures (see the supporting information). Figs. 9(a) and 9(b) show the solvent-cavity volume measured following abrupt cooling and following cold unit-cell expansion, determined by subtracting the protein volume from the unit-cell volume. Figs. 9(c) and 9(d) show that the volume outside the protein shell of apoferritin contracts more during cooling and expands more during the cold expansion than does the solvent volume inside the shell. This suggests that the largest contribution to the unit-cell changes comes from interactions between apoferritin shells rather than from changes within them.

Fig. 10 shows the fractional changes from room temperature of the linear dimension of the unit cell and of Δ_{radial} , Δ_{tetramer} and Δ_{dimer} as defined in Fig. 1(e). While each distance metric decreased on cooling and increased on expansion, Δ_{tetramer} showed by far the largest fractional changes, further implicating the interface between protein shells as the dominant

location for changes to the unit cell. As shown in Supplementary Fig. S5, fractional changes in R_g of each monomer and its components along $\mathbf{r}_{\text{radial}}$, $\mathbf{r}_{\text{dimer}}$ and $\mathbf{r}_\alpha$ are much smaller.

Figs. 11(a) and 11(b) show representations of the dimers in the 'out' conformation, created in *PyMOL*. The colours from blue to red indicate increasing values of $\widehat{\epsilon}_i^{a,b}$ obtained by comparing room-temperature and initial cold structures (Fig. 11a) and initial cold structures with final cold structures after expansion (Fig. 11b). The arrow locations indicate positions where the normalized atomic displacements are at least one standard deviation larger than the mean, the arrow directions indicate the direction of the displacement and the arrow lengths are 50 times larger than the actual displacement. During cold expansion (Fig. 11b) the atomic changes are somewhat localized to the BC loop at the tetramer interface, consistent with post-cooling expansion occurring primarily between the spherical shells. During cooling (Fig. 11a) the structural changes are more uniformly distributed throughout the protein shell.

4. Discussion

4.1. Solvent flows, cooling-induced disordering and cold reordering

The present results provide the clearest evidence to date for the role of solvent flows in causing cooling-induced disorder in protein crystals. During cooling, the unit-cell volume generally contracts more than the protein volume, so that the solvent-cavity volume shows the largest fractional decrease (Juers & Matthews, 2001, 2004; Juers *et al.*, 2018). Unless the solvent internal to the crystal contains large cryoprotectant concentrations [$\sim 30\%(\nu/\nu)$ or more], its volume will contract less than the solvent cavities. If cooling times are short, the exiting solvent cannot make its way to the crystal surface and the lattice of protein molecules must undergo elastic deformations and plastic failure to accommodate it.

Here, we observe a substantial expansion, often to values larger than at room temperature, of the crystal lattice and a concurrent crystal reordering, as reflected in a drop in mosaicity and also sometimes in *B* factor, at *constant* temperatures between 220 and 260 K. All external solvent was removed from the crystals (which was essential to allow cooling below 260 K without ice formation), and an oil coating prevented all solvent transport between the crystal and the surrounding dry nitrogen-gas cryostream. Consequently, the solvent required to fill the expanding solvent cavities must come from inside the crystal and nowhere else. The solvent itself could expand if, for example, there was a change in water structure toward the open structure of ice, but this would be expected to occur on timescales far shorter than the tens of seconds required for expansion. We thus conclude that the solvent that fills the expanding solvent cavities within the unit cells flows from elastically and plastically deformed crystal regions with excess solvent.

Evidence for the role of solvent transport in unit-cell expansion may be provided by the temperature- and glycerol-dependent timescale of the unit-cell expansion following

Figure 9
(a, b) Total solvent-cavity volume in the cubic unit cell of apoferritin and (c, d) the ratio of solvent-cavity volume located outside the apoferritin shell to the volume inside versus temperature for glycerol-free crystals with Gln82 in the 'out' configuration (a, c) and the 'in' configuration (b, d). Closed and open symbols indicate values determined before and after unit-cell expansion, respectively. Solid lines indicate values obtained from slowly cooled (0.1 K s^{-1}) rather than abruptly cooled ($\sim 300 \text{ K s}^{-1}$) crystals. The solvent-cavity volume and the fraction of the total cavity volume located outside the apoferritin shell drops with decreasing temperature and both increase during cold cell expansion. This suggests that the primary structural changes during cold expansion occur at the interface between apoferritin shells rather than within them.

cooling, shown in Fig. 7. Between temperatures of 220 and 260 K, the viscosity of bulk glycerol–water mixtures increases by roughly an order of magnitude, as the glycerol concentration increases from 0 to 40% (v/v), and by a similar factor for concentrations in this range on cooling from 260 to 220 K (Trejo González *et al.*, 2011). These trends are qualitatively consistent with the observed expansion timescales, although the viscosities should be modified by nanoconfinement within the protein structure, and other factors (for example, thermally activated structural relaxations) could affect expansion timescales.

Additional evidence for the role of solvent transport comes from how cold expansion affects mosaicities. For crystals soaked in glycerol solutions with concentrations of 20% and below cold cell expansion causes the mosaicities to drop, while

for crystals soaked in 40% solutions the mosaicities frequently increased or showed no change during expansion. For glycerol concentrations below $\sim 20\text{--}30\%$ (v/v) the bulk-solvent contraction is smaller than that of the solvent cavities occurring during abrupt cooling, so solvent should initially be expelled from the unit cells, increasing the mosaicity, and then re-enter during cold expansion, decreasing the mosaicity. For crystals soaked in 40% (v/v) glycerol the bulk-solvent contraction is larger than the initial solvent-cavity contraction, so during cooling solvent flows into the unit cell, depleting it from other crystal regions and disordering them. During cold expansion, even more solvent must flow into the unit cell, further depleting other crystal regions and disordering them, so the mosaicity should increase. The actual flows into and out of the unit cell are somewhat uncertain because both the

Figure 10

Percentage change in (a, b) the linear dimension of the cubic unit cell, (c, d) the radial distance from the centre of volume of the apoferritin shell to the centre of volume of a ferritin monomer within the shell, (e, f) the distance between centres of volume of ferritin monomers within a dimer and (g, h) the distance between centres of volume of adjacent shells (roughly, the distance between protein shells) versus temperature for glycerol-free cubic apoferritin crystals. Results for crystals with Gln82 in the 'out' (a–d) and 'in' (e–h) conformations are shown separately. Closed and open symbols indicate values before and after cold expansion, respectively, and solid lines represent values obtained from crystals that were slowly rather than abruptly cooled. Expansion has by far the largest effect on the shell–shell distance, with little effect on dimensions within the shell.

Figure 11

Atomic-level changes to apoferritin monomers in a dimer pair (a) during cooling from room temperature to 220 K and (b) during post-cooling expansion at $T = 220$ K. Only structures with Gln82 in the 'out' conformation are used. The blue arrowed line connects the centres of volumes of the monomers and the long red arrow points in the direction of the centre of the spherical apoferritin shell (comprising 24 monomers). Colouring indicates the normalized displacements $\widehat{\epsilon}_i^{a,b}$ between atomic positions in the compared structures, calculated using (7). Short red arrows indicate the direction of the displacements and the arrow lengths are 50 times larger than the actual displacements. The overall shift of atomic positions owing to the decrease in separation of monomers on cooling and the increase during expansion has been subtracted out and is not reflected in the colour/arrow scheme. The mean and standard deviation of the normalized displacements $\widehat{\epsilon}_i^{a,b}$ are (a) 0.36 ± 0.15 and (b) 0.21 ± 0.11 .

internal glycerol concentration and the solvent contraction are modified by preferential hydration of and nanoconfinement by protein surfaces.

Supplementary Fig. S3 shows estimates of the volume fraction of internal solvent present at room temperature that cannot be accommodated within the solvent-cavity volume deduced from refined structures of the initial cold crystal and the crystal after unit-cell expansion. The solvent volume at each temperature was estimated as described in Moreau *et al.* (2019) by assuming that the first hydration layer does not contract and that the remaining solvent has the same contraction as the bulk liquid. The volumes of solvent transported during both cooling and cold expansion are at least a few percent of the crystal volume. Since the crystals were typically a few hundred micrometres in size, were that much solvent to be present at the crystal surface it would easily be visible and would [for glycerol concentrations of 20% (v/v) and below] crystallize immediately. We are thus confident that most of the solvent transported during cooling and cold expansion remains within the crystal in defective regions that contribute to the observed rise in mosaicity.

Indirect evidence for the role of solvent transport in creating protein crystal disorder has been provided by X-ray topography, which has directly imaged the mosaic domain structure that forms on cooling (Kriminski *et al.*, 2002). The occasional success of 'cryoannealing' (Harp *et al.*, 1999), in which the brief warming of a cold crystal to room temperature and then re-cooling sometimes leads to reductions in mosaicity and improvement in overall crystal order, may be due in part to adjustment of the internal water content of the crystal via transport of solvent to or from the substantial bulk solvent that surrounds crystals in typical mounting practice (Juers & Matthews, 2004). The brief warming may also allow the relaxation of elastically deformed lattice regions and the healing of smaller plastically deformed regions, leaving a smaller number of plastically deformed regions to take up excess solvent on recooling and thereby reducing average disorder.

Although solvent transport into and out of unit cells correlates with changes in crystal mosaicity during cold expansion, we cannot conclude based on the present evidence that they are solely responsible for these changes or for mosaicity increases during cooling. Kinetic trapping of incompletely relaxed protein and lattice conformations during abrupt cooling should, even in the absence of differential contraction between solvent cavities and internal solvent, create disorder that increases the mosaicity relative to both its room-temperature and fully relaxed cold crystal states, and that increases the *B* factors relative to the relaxed cold state. However, the magnitudes of the estimated solvent flows, at least a few percent of the total solvent volume, are so large that their contribution to disorder must be substantial.

4.2. Correlation between solvent contraction, unit-cell contraction and crystal mosaicity

Direct efforts to demonstrate a relation between cold crystal mosaicity and the expected contraction of the internal

solvent on cooling between room temperature and $T = 77$ or 100 K have so far yielded unconvincing results. Juers *et al.* (2018) found no clear variation of mosaicity with bulk-solvent contraction for ice-free crystals of at least seven out of nine different protein crystal systems; only α -lactalbumin showed weakly suggestive evidence of a mosaicity minimum versus unit-cell contraction. However, in these careful experiments any trends may have been obscured by the relatively large lower bound on measurable mosaicity owing to the large incident beam divergence of the laboratory X-ray source used. Our recent data for crystals plunge-cooled in liquid nitrogen and measured at $T = 100$ K suggests a decrease in mosaicity with increasing glycerol concentration for both thaumatin and apoferritin, with crystals soaked in 40% (v/v) glycerol giving the lowest values (Moreau *et al.*, 2019). However, for both apoferritin (Fig. 4) and thaumatin crystals cooled in gas streams to temperatures between 180 and 260 K, glycerol-free crystals typically show the smallest cold mosaicities. Disorder owing to kinetic trapping of incompletely relaxed protein conformations may impose a lower bound on low-temperature mosaicities, and this disorder might grow as glycerol concentrations increase.

There is a reliable correlation between unit-cell contraction on cooling and the bulk contraction of the solution in which the crystals are grown or soaked. Juers and coworkers observed a near-linear relation between unit-cell contraction between room temperature and $T = 100$ K and the bulk-solvent contraction in eight different protein crystal systems (Juers *et al.*, 2018). The present data for apoferritin in Fig. 3 and the data for thaumatin (Moreau *et al.*, 2019) show that both the unit-cell and solvent-cavity volume contractions increase with increasing glycerol concentration at fixed temperature between 200 and 260 K.

However, as shown in Fig. 6, for apoferritin crystals that are slowly cooled to temperatures above 200 K or that have undergone cold expansion, the net unit-cell contraction on cooling is independent of glycerol concentration and thus solvent contraction.

The different behaviours versus glycerol concentration/bulk-solvent contraction observed during abrupt cooling, slow cooling and cold expansion suggest one possible reason why a simple correlation between mosaicity and solvent contraction has not been observed. Protein crystals can be considered as poroelastic materials (Biot, 1941). As discussed in detail by Juers *et al.* (2018), excess solvent expansion or contraction relative to the solvent cavities generates excess pressure that, on short timescales, will drive cavity expansion and, on long time scales, will drive solvent flow. Consequently, the strongest correlation between solvent-cavity contraction and solvent contraction (and between unit-cell and solvent contractions) on cooling should be observed when crystals are cooled so fast to temperatures well below the glass-transition temperature of the internal solvent (for example, well below ~ 200 K) that little flow can occur before the solvent vitrifies. When crystals are cooled very slowly or held at fixed temperatures above ~ 200 K, cavity and solvent contraction should be uncorrelated. This neglects any chemical effects of different solvent

compositions on, for example, protein conformation and contraction behaviour.

4.3. Why does the unit cell contract and then expand?

For cubic apoferritin crystals, the initial unit-cell volumes obtained following fast cooling are substantially smaller than those obtained via slow cooling to the same temperature or by allowing crystal relaxation at fixed temperature above 200 K. This implies that the initial fast-cooled state reflects a kinetically favoured but non-equilibrium configuration. During cooling, the amplitudes of atomic motions within local energy minima decrease and minor conformers separated by small barriers from major conformers are depopulated. The time-scales for these relaxations are short compared with the time interval during fast cooling and the sample remains above the protein–solvent glass transition. These local relaxations should in general lead to contraction of the protein and of the crystal lattice.

However, larger scale, cooperative changes in protein conformation and lattice packing towards new equilibrium configurations will in general occur much more slowly. These changes will be driven in part by temperature-dependent changes in interaction strengths. For example, as temperature decreases, the pH, the pK_a values of side chains and the water activity all change. The hydrophobic interaction largely responsible for protein folding and for the formation of some crystal contacts weakens owing to increased tetrahedral ordering of water and a reduction in the entropic cost of solvent ‘caging’ around hydrophobic residues, and this can lead to cold unfolding/denaturation of proteins in solution (Privalov, 1990). Cooperative structural changes have much slower kinetics than the quenching of local motions, especially within the constrained environment of the crystal, and so limited evolution toward new minima can occur during fast cooling. The present results show that as long as the internal solvent remains liquid, substantial cooperative evolution of the protein and its lattice towards a new temperature-dependent equilibrium can occur at temperatures as low as 220 K.

Several other mechanisms that could give rise to cold unit-cell expansion can be ruled out. Firstly, unit cells can expand when internal ice forms. For glycerol concentrations below 20% (v/v), diffraction from internal ice is sharp and easily detected and its intensity rapidly saturates after nucleation. Internal ice was not observed (Fig. 2) and cannot explain the observed expansions in nominally ice-diffraction-free crystals. For larger glycerol concentrations, ice grows more slowly and the ice-grain size becomes much smaller, so that initially weak and diffuse ice diffraction may be difficult to detect in the diffuse diffraction background.

Secondly, migration of excess solvent initially present at the surface of the crystal or from ambient humid gas to the crystal interior is known to cause crystal expansion at and near room temperature. This is believed to be the most important mechanism involved in ‘cryo-annealing’ protocols (Harp *et al.*, 1999; Juers & Matthews, 2004). However, for all of the crystals

reported here external solvent was carefully removed and replaced with oil. Any substantial volumes of surface solvent (estimated using optical measurements to be much less than 0.1% of the crystal volume) rapidly formed ice I_h , especially for glycerol-free crystals, for which we observed the greatest reductions in mosaicity and B factor during cell expansion. The surrounding oil provides an effective barrier to water diffusion to or from the ambient gas. The ambient gas was the cold, dry, flowing N_2 of the cryostream that, when inadequate oil was present, rapidly dehydrated the crystals and caused the unit cells to shrink, not expand.

Thirdly, initial solvent contraction might occur if the pressure within the crystal during cooling became sufficient to drive the solvent into a high-density amorphous (HDA) ice (at low temperature) or a high-density liquid (HDL) phase. The observed position of the solvent (and NVH oil) diffraction peaks (Fig. 2) remained constant during the cell expansion, and in glycerol-free crystals is consistent with the expected diffraction from normal (low-density) water or LDA. The solvent-peak position does not shift and is unambiguously inconsistent at all times with the peak position expected for HDA (Kim *et al.*, 2008).

Fourthly, a temperature-dependent chemical potential for glycerol or salt within the crystal could lead to transient changes in the unit cell following cooling. For example, if preferential hydration and glycerol exclusion increase with decreasing temperature, then diffusion of glycerol out of and water into the crystal following cooling could lead to changes in the unit cell. The sign of such changes is not obvious.

Finally, we note that not all apoferritin crystals examined at temperatures between 220 and 260 K showed cold expansion on the 200 s timescale of our experiments. For those that did not, both the unit-cell contraction and the mosaicity increase on cooling were much larger (for example a factor of two at 240 K) than for nominally identically prepared crystals that showed cold expansion. The difference in behaviour could be owing to slight dehydration by the dry N_2 gas stream during cooling if, for example, the oil thinned near a crystal corner or edge or owing to some difference in crystallization conditions.

5. Conclusions

The present results provide the clearest evidence to date for internal solvent transport in protein crystals during cooling, and insight into the roles of this transport and of kinetically quenched structural relaxations in generating crystal disorder during cooling. During abrupt cooling, local vibrations and small conformational motions freeze out, leading to contraction of the unit cell and of the solvent cavities within it. At ambient pressure, the equilibrium volume of the internal solvent generally contracts by a different percentage to the equilibrium volume of the solvent cavities themselves. For extremely large ($>10^4$ K s $^{-1}$) cooling rates, solvent flow during the cooling time before vitrification should be limited, internal pressure within the solvent cavity should grow, and solvent-cavity and solvent contractions should track. For cooling at the rates typically achieved using cold gas streams or hand plun-

ging in liquid nitrogen ($\sim 10^2$ – 10^3 K s $^{-1}$), solvent flows into or out of unit cells until it vitrifies, reducing or eliminating the correlation between solvent cavity and solvent contractions, but creating elastic and plastic deformations in regions where excess solvent accumulates (or from which it is depleted) that increase crystal disorder. At much slower cooling rates (<1 K s $^{-1}$), solvent may have time to flow to or from the crystal surface without substantial accumulation within/depletion from internal regions and without associated increases in crystal disorder.

Abrupt cooling quenches slower timescale, larger length-scale cooperative relaxations of the protein–solvent system as it attempts to evolve towards its temperature-dependent equilibrium, and this also creates crystal disorder. However, at final temperatures near and above the protein–solvent glass transition temperature (~ 200 K) and as long as the internal solvent remains liquid, substantial evolution of both conformation and lattice packing can occur on timescales of seconds to tens of seconds that increase with decreasing temperature. In the case of apoferritin, these lead to a reversal of solvent flows and a dramatic reordering of the crystal at fixed low temperature, including in crystals that have no cryoprotectants other than salt present in the mother liquor.

Perhaps the most interesting, if obvious, conclusion from the present work is the importance of time in variable/multi-temperature studies of protein structure. Although short equilibration times (for example, a few seconds) at a given temperature should be sufficient to examine the freeze-out of side-chain conformers, exploring cooperative conformational changes will generally require much longer equilibration times.

The potential of protein crystals for the study of cold denaturation is particularly intriguing. Most proteins are expected to unfold at sufficiently low temperature, and this may be a major issue in the cold storage and recovery of proteins and biologic medications and in the cryopreservation of cells and tissues. Studies of unfolding in solution have been constrained by the formation of ice near 273 K and so have often required the addition of denaturants such as urea to promote unfolding at higher temperatures. Nanoconfined water within protein crystals can be maintained in a (super-cooled) liquid state at temperatures down to ~ 200 K for tens of seconds or more using at most small cryoprotectant concentrations. While full unfolding cannot occur within the constraints of the protein lattice, cooperative conformation changes associated with weakening of the hydrophobic interaction that gives rise to unfolding should be observable and may allow differences in low-temperature stability between proteins to be evaluated.

Acknowledgements

All X-ray data collection was performed at the Cornell High Energy Synchrotron Source (CHESS), which was supported by the National Science Foundation under award DMR-1332208, using the Macromolecular Diffraction at CHESS (MacCHESS) facility, which was supported by award GM-

103485 from the National Institute of General Medical Sciences, National Institutes of Health. RET acknowledges a financial conflict of interest, as some of the tools used in this work were provided by MiTeGen LLC, in which he has a significant financial interest.

Funding information

This work was supported by the NSF under award MCB-1330685. DWM acknowledges additional support from Cornell University's Molecular Biophysics Training Grant (NIH T32GM0082567) and from the NIH (R01-GM127528).

References

- Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L.-W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C. & Zwart, P. H. (2010). *Acta Cryst.* **D66**, 213–221.
- Atakisi, H., Moreau, D. W. & Thorne, R. E. (2018). *Acta Cryst.* **D74**, 264–278.
- Biot, M. (1941). *J. Appl. Phys.* **12**, 155–164.
- Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. & Richardson, D. C. (2010). *Acta Cryst.* **D66**, 12–21.
- Clegg, S. L., Ho, S., Chan, C. & Brimblecombe, P. (1995). *J. Chem. Eng. Data*, **40**, 1079–1090.
- Doster, W. (2010). *Biochim. Biophys. Acta*, **1804**, 3–14.
- Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* **D66**, 486–501.
- Fahy, G. M. & Wovk, B. (2015). *Methods Mol. Biol.* **1257**, 21–82.
- Fenwick, R. B., van den Bedem, H., Fraser, J. S. & Wright, P. E. (2014). *Proc. Natl Acad. Sci. USA*, **111**, E445–E454.
- Fraser, J. S., van den Bedem, H., Samelson, A. J., Lang, P. T., Holton, J. M., Echols, N. & Alber, T. (2011). *Proc. Natl Acad. Sci. USA*, **108**, 16247–16252.
- Glycerine Producers Association (1963). *Physical Properties of Glycerine and Its Solutions*. New York: Glycerine Producers Association.
- Hare, D. E. & Sorensen, C. M. (1987). *J. Chem. Phys.* **87**, 4840–4845.
- Harp, J. M., Hanson, B. L., Timm, D. E. & Bunick, G. J. (1999). *Acta Cryst.* **D55**, 1329–1334.
- Hempstead, P. D., Yewdall, S. J., Fernie, A. R., Lawson, D. M., Artymiuk, P. J., Rice, D. W., Ford, G. C. & Harrison, P. M. (1997). *J. Mol. Biol.* **268**, 424–448.
- Juergens, D. H., Farley, C. A., Saxby, C. P., Cotter, R. A., Cahn, J. K. B., Holton-Burke, R. C., Harrison, K. & Wu, Z. (2018). *Acta Cryst.* **D74**, 922–938.
- Juergens, D. H. & Matthews, B. W. (2001). *J. Mol. Biol.* **311**, 851–862.
- Juergens, D. H. & Matthews, B. W. (2004). *Acta Cryst.* **D60**, 412–421.
- Kabsch, W. (2010). *Acta Cryst.* **D66**, 125–132.
- Keedy, D. A., Hill, Z. B., Biel, J. T., Kang, E., Rettenmaier, T. J., Brandão-Neto, J., Pearce, N. M., von Delft, F., Wells, J. A. & Fraser, J. S. (2018). *Elife*, **7**, e36307.
- Keedy, D. A., Kenner, L. R., Warkentin, M., Woldeyes, R. A., Hopkins, J. B., Thompson, M. C., Brewster, A. S., Van Benschoten, A. H., Baxter, E. L., Uervirojnangkoorn, M., McPhillips, S. E., Song, J., Alonso-Mori, R., Holton, J. M., Weis, W. I., Brunger, A. T., Soltis, S. M., Lemke, H., Gonzalez, A., Sauter, N. K., Cohen, A. E., van den Bedem, H., Thorne, R. E. & Fraser, J. S. (2015). *Elife*, **4**, e07574.
- Kim, C. U., Chen, Y.-F., Tate, M. W. & Gruner, S. M. (2008). *J. Appl. Cryst.* **41**, 1–7.
- Kriminski, S., Caylor, C. L., Nonato, M. C., Finkelstein, K. D. & Thorne, R. E. (2002). *Acta Cryst.* **D58**, 459–471.

- Lang, P. T., Ng, H.-L., Fraser, J. S., Corn, J. E., Echols, N., Sales, M., Holton, J. M. & Alber, T. (2010). *Protein Sci.* **19**, 1420–1431.
- Li, A. J. & Nussinov, R. (1998). *Proteins*, **32**, 111–127.
- Loerting, T., Bauer, M., Kohl, I., Watschinger, K., Winkel, K. & Mayer, E. (2011). *J. Phys. Chem. B*, **115**, 14167–14175.
- Moreau, D. W., Atakisi, H. & Thorne, R. E. (2019). *IUCrJ*, **6**, 346–356.
- Privalov, P. L. (1990). *Crit. Rev. Biochem. Mol. Biol.* **25**, 281–305.
- Ringe, D. & Petsko, G. A. (2003). *Biophys. Chem.* **105**, 667–680.
- Shen, C., Julius, E. F., Tyree, T. J., Moreau, D. W., Atakisi, H. & Thorne, R. E. (2016). *Acta Cryst. D* **72**, 742–752.
- Teeter, M. M., Yamano, A., Stec, B. & Mohanty, U. (2001). *Proc. Natl Acad. Sci. USA*, **98**, 11242–11247.
- Tilton, R. F., Dewan, J. C. & Petsko, G. A. (1992). *Biochemistry*, **31**, 2469–2481.
- Trejo González, J. A., Longinotti, M. P. P. & Corti, H. R. (2011). *J. Chem. Eng. Data*, **56**, 1397–1406.
- Tyree, T. J., Dan, R. & Thorne, R. E. (2018). *Acta Cryst. D* **74**, 471–479.
- Val, N. de, Declercq, J.-P., Lim, C. K. & Crichton, R. R. (2012). *J. Inorg. Biochem.* **112**, 77–84.
- Vedula, L. S., Brannigan, G., Economou, N. J., Xi, J., Hall, M. A., Liu, R., Rossi, M. J., Dailey, W. P., Grasty, K. C., Klein, M. L., Eckenhoff, R. G. & Loll, P. J. (2009). *J. Biol. Chem.* **284**, 24176–24184.
- Vinothkumar, K. R. & Henderson, R. (2016). *Q. Rev. Biophys.* **49**, e13.
- Voss, N. R. & Gerstein, M. (2010). *Nucleic Acids Res.* **38**, W555–W562.
- Warkentin, M., Badeau, R., Hopkins, J. B. & Thorne, R. E. (2012). *Acta Cryst. D* **68**, 1108–1117.
- Warkentin, M., Hopkins, J. B., Badeau, R., Mulichak, A. M., Keefe, L. J. & Thorne, R. E. (2013). *J. Synchrotron Rad.* **20**, 7–13.
- Warkentin, M. & Thorne, R. E. (2009). *J. Appl. Cryst.* **42**, 944–952.
- Warkentin, M. & Thorne, R. E. (2010a). *Acta Cryst. D* **66**, 1092–1100.
- Warkentin, M. & Thorne, R. E. (2010b). *J. Struct. Funct. Genomics*, **11**, 85–89.
- Weik, M., Kryger, G., Schreurs, A. M. M., Bouma, B., Silman, I., Sussman, J. L., Gros, P. & Kroon, J. (2001). *Acta Cryst. D* **57**, 566–573.
- Weik, M., Ravelli, R. B. G., Silman, I., Sussman, J. L., Gros, P. & Kroon, J. (2001). *Protein Sci.* **10**, 1953–1961.