Effects of protein-crystal hydration and temperature on side-chain conformational heterogeneity in monoclinic lysozyme crystals

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

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

Received 10 October 2017

Accepted 3 January 2018

Edited by Z. S. Derewenda, University of Virginia, USA

Keywords: protein crystallography; protein structure; crystal dehydration; conformational heterogeneity; lysozyme; variable-temperature crystallography; protein–solvent interactions.

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

The modulation of main-chain and side-chain conformational heterogeneity and solvent structure in monoclinic lysozyme crystals by dehydration (related to water activity) and temperature is examined. Decreasing the relative humidity (from 99 to 11%) and decreasing the temperature both lead to contraction of the unit cell, to an increased area of crystal contacts and to remodeling of primarily contact and solvent-exposed residues. Both lead to the depopulation of some minor side-chain conformers and to the generation of new conformations. Side-chain modifications and main-chain r.m.s.d.s associated with cooling from 298 to 100 K depend on relative humidity and are minimized at 85% relative humidity (r.h.). Dehydration from 99 to 93% r.h. and cooling from 298 to 100 K result in a comparable number of remodeled residues, with dehydration-induced remodeling somewhat more likely to arise from contact interactions. When scaled to equivalent temperatures based on unit-cell contraction, the evolution of side-chain order parameters with dehydration shows generally similar features to those observed on cooling to $T = 100$ K. These results illuminate the qualitative and quantitative similarities between structural perturbations induced by modest dehydration, which routinely occurs in samples prepared for 298 and 100 K data collection, and cryocooling. Differences between these perturbations in terms of energy landscapes and occupancies, and implications for variable-temperature crystallography between 180 and 298 K, are discussed. It is also noted that remodeling of a key lysozyme active-site residue by dehydration, which is associated with a radical decrease in the enzymatic activity of lysozyme powder, arises due to a steric clash with the residue of a symmetry mate.

1. Introduction

Water, with its unique physicochemical properties, plays a central role in determining the structure and function of proteins and other biomolecules. Water is essential to protein folding/unfolding, to the stabilization of higher-order structures and in determining catalytic efficiencies (Rupley & Careri, 1991). The internal and large-scale conformational dynamics of proteins are coupled to the dynamics of surface-ordered (hydration) and bulk solvent (Fenimore *et al.*, 2004; Frauenfelder *et al.*, 2009). Protein function ceases when the surrounding solvent is depleted to only one or two hydration layers (Frauenfelder *et al.*, 2009).

X-ray crystallography has long been used to study the effects of crystal and protein hydration (Crowfoot & Riley, 1938; Perutz, 1946, 1954; Einstein & Low, 1962; Huxley & Kendrew, 1953). Protein crystals contain large amounts of solvent, typically 30–70% by volume (Kantardjieff & Rupp,

© 2018 International Union of Crystallography

2003; Matthews, 1968), that fills intermolecular spaces within the unit cell as well as internal protein cavities. Solvent molecules preferentially occupy specific sites near the protein surface, especially within the first hydration shell and at low temperatures, allowing them to be crystallographically detected, but most solvent tends to be positionally disordered and so contributes only to the average solvent electron density.

Crystal solvent content depends on relative humidity (r.h.) or, equivalently, the activity of water, a thermodynamic variable. Relative humidity determines how proteins pack within the crystal, the structure of individual proteins, and also the solvent structure. Dehydration, both intentional and accidental, decreases the unit-cell volume, and crystal order and diffraction resolution often degrade but sometimes improve. Commercial and homemade systems for adjusting hydration levels to optimize diffraction properties are now in regular use (Bowler *et al.*, 2015; Sanchez-Weatherby *et al.*, 2009; Kieferbauer *et al.*, 2000) and diffraction improvements have been obtained at relative humidities as low as 74% (Krauss *et al.*, 2012). Dehydration below ~85% r.h. often leads to first-order structural phase transitions involving the cooperative loss of a finite amount of water and the appearance of a new lattice (Salunke *et al.*, 1985; Einstein & Low, 1962; Perutz, 1946; Huxley & Kendrew, 1953). Diffraction outcomes are strongly affected by the rate of dehydration, which affects the ability of the protein and crystal to adjust to the evolving equilibrium state. Rapid dehydration often leads to spatial inhomogeneities and plastic lattice failure, while slow dehydration can 'supercool' a given lattice to relative humidities well below that at which it becomes thermodynamically unstable (Dobrianov *et al.*, 2001). Disorder created by dehydration can often be largely or fully reversed and the original crystal packing restored by rehydration (Salunke *et al.*, 1985; Dobrianov *et al.*, 2001; Kieferbauer *et al.*, 2000).

Previous crystallographic studies of the effects of hydration on protein structure have mainly focused on changes to the main-chain conformation. For example, orthorhombic and monoclinic crystals of bovine pancreatic ribonuclease A dehydrated to 79% r.h. showed rearrangement of surface loops and a small change in the hinge angle between subdomains relative to the native hydrated structure (Bell, 1999; Radha Kishan *et al.*, 1995). Dehydration of cytomegalovirus immediate-early 1 protein triggered a structural transition involving remodeling of the tertiary structure in addition to a change in space group (Klingl *et al.*, 2015). The most detailed studies have been performed using various crystal forms of hen egg-white lysozyme, which can retain crystalline and molecular order at humidities as low as 0.01% (Salunke *et al.*, 1985; Kodandapani *et al.*, 1990; Madhusudan *et al.*, 1993; Nagendra *et al.*, 1998; Harata & Akiba, 2006). The monoclinic form diffracts to the highest resolution over the widest range of relative humidities. The most pronounced main-chain changes occur in regions involved in substrate binding (Madhusudan *et al.*, 1993). However, the effects of dehydration on side-chain conformations and solvent structure have not been systematically studied.

The importance of the conformational ensemble of a protein in folding (Vendruscolo, 2007), molecular recognition (Lange *et al.*, 2008), catalysis (Boehr *et al.*, 2006) and allosteric regulation (Gunasekaran *et al.*, 2004) has long been recognized (DePristo *et al.*, 2004; Furnham *et al.*, 2006). As of 2010, only ~5% of PDB-deposited crystallographic structures were refined against multiple-conformer models (Lang *et al.*, 2010). In the last decade, this situation has improved as methods for identifying and modeling main-chain and side-chain conformational heterogeneity have been developed and have become easier to use. Low electron-density levels previously assumed to be noise, in fact, can contain abundant information about minority occupancy conformers. Analysis software including *Ringer* (Lang *et al.*, 2010, 2014), which samples and maps electron density around dihedral angles to identify peaks associated with low-occupancy alternative conformers of side-chain rotamers, and *qFit* (van den Bedem *et al.*, 2009; van den Bedem & Fraser, 2015), which autobuilds up to four conformations per residue, have facilitated the generation of multiconformer models. Bond-order parameters S^2 , which were originally developed to describe NMR data (Brueschweiler & Wright, 1994), have been generalized for crystallography to characterize both harmonic and nonharmonic (conformational) side-chain fluctuations (Fenwick *et al.*, 2014).

A primary application of these tools to date has been to characterize the effects of temperature. A comparison of PDB-deposited structures of 30 proteins at room temperature and $T = 100$ K indicated that cryocooling causes substantial remodeling of side-chain conformations (Fraser *et al.*, 2011). For at least 24 of these proteins the cryoprotectants used in cryogenic data collection were absent from the crystals used for room-temperature measurements, complicating the interpretation of structural differences. The conformational heterogeneity of a Michaelis complex of DHFR was found to be significantly distorted by cryocooling (Keedy *et al.*, 2014). Analysis of crystallographic data from the human proline isomerase cyclophilin A (cypA) acquired at seven temperatures between 298 and 100 K revealed how side-chain conformations and bond-order parameters evolve with temperature across the entire molecule. The evolution was found to be highly heterogeneous: bond-disorder parameters $1 - S^2$ have a 'hockey-stick' shape and the 'transition' temperatures above which order decreases more rapidly have a broad distribution between ~180 and ~270 K. Analysis of diffraction data collected to high doses from three different proteins (including cypA) at room temperature indicates that radiation damage has only small effects on conformational heterogeneity compared with those of temperature (Russi *et al.*, 2017). However, the full physical and biological significance of temperature-dependent conformational changes remains unclear.

Water activity (relative humidity), like temperature, is a thermodynamic variable. Some of the gross effects of dehydration, such as unit-cell contraction, are similar to those of cooling. Moreover, unintentional dehydration routinely occurs during crystal handling, mounting and storage in both

room-temperature and $T = 100$ K crystallographic practice. Unintentional dehydration is a primary cause of non-isomorphism of identically prepared crystals at room temperature and of unit-cell variations at cryogenic temperatures (Farley *et al.*, 2014). Could dehydration account for some conformational changes that have been attributed to cooling? To what extent does dehydration mimic the effects of cooling? How do the effects of these two thermodynamic perturbations on protein conformational ensembles and solvent structure as they exist in crystals differ?

Here, we report high-resolution (1.1–1.9 Å) measurements and multiconformer analysis of hydration-dependent side-chain conformations in monoclinic lysozyme crystals from relative humidities between 99 and 11% r.h., corresponding to crystal solvent contents ranging from 38 to 18%(v/v), respectively, at both 298 and 100 K. Detailed structural comparisons were facilitated by using carefully prepared and characterized crystals of otherwise identical composition. Substantial side-chain remodeling occurs on dehydration. Solvent-exposed residues are significantly more affected than buried residues, in part correlated with the large increase in the number of crystal contacts as the packing of the protein becomes denser. The largest side-chain conformation changes occur at humidities near ~75%, where significant depletion of water within the first hydration shell begins. Dehydration to 93% r.h. has similar effects to cooling to 100 K: dehydration causes a 5.4% contraction of the unit-cell volume and remodels 12.6% of side chains, compared with a 6.8% unit-cell contraction and 13.6% side-chain remodeling on cooling, with only solvent-exposed residues affected in both cases. These results complicate the interpretation of differences between room-temperature and cryogenic temperature conformational ensembles when the precise crystal history is not known and/or is not carefully controlled; they suggest that modest dehydration can serve as a proxy for studying many of the effects of temperature; and they emphasize the importance of packing environment in selecting targets for variable-temperature crystallographic studies.

2. Materials and methods

2.1. Crystallization

Monoclinic lysozyme crystals were grown at room temperature by the hanging-drop vapor-diffusion method. Lyophilized lysozyme powder (Sigma–Aldrich, St Louis, Missouri, USA) was dissolved in 50 mM sodium acetate buffer solution at pH 4.5 at a concentration of 10–15 mg ml⁻¹, and the reservoir solutions were composed of 2.8–3.0%(w/v) (~0.35 M) sodium nitrate in the same buffer solution. Crystallization drops of 6–10 µl in volume, prepared by mixing equal volumes of protein and reservoir solution, were suspended over ~500 µl reservoir solution, and rectangular plate-like crystals, mostly in clusters, were obtained within a week. The largest crystals, which were 300–600 µm in the long dimension, were used to facilitate the collection of complete crystallographic data sets at room temperature with minimal radiation damage.

2.2. Crystal dehydration

Crystals were dehydrated to seven relative humidities between 99 and 11% r.h. by equilibrating them at $T = 293$ – 298 K with mother-liquor reservoir solution ($\sim 99.2 \pm 0.5\%$ r.h.) and saturated solutions of potassium nitrate ($93.4 \pm 0.8\%$ r.h.), potassium chloride ($85.3 \pm 0.7\%$ r.h.), sodium chloride ($75.4 \pm 0.3\%$ r.h.), sodium bromide ($57.7 \pm 0.9\%$ r.h.), magnesium chloride ($33.1 \pm 0.2\%$ r.h.) and lithium chloride ($11.3 \pm 0.5\%$ r.h.) (Rockland, 1960; Young, 1967; Greenspan, 1977). Each crystal was harvested from its mother liquor using a microfabricated polymer loop, and excess mother liquor was carefully removed from the crystal and loop by blotting. The crystal and loop were then inserted into a polymer capillary (MicroRT system, MiTeGen, Ithaca, New York, USA) containing ~20 µl saturated salt solution and equilibrated for 24–48 h. The distance between the crystal and the solution was ~1 cm, and a few extra salt grains were added to the solution to ensure that it remained saturated.

Crystals that are rapidly dehydrated to 75% r.h. and below often undergo catastrophic plastic failure, leading to loss of diffraction (Dobrianov *et al.*, 2001). This is most likely to occur due to inhomogeneous stresses that develop as the crystal lattice undergoes its structural phase transition near 85% r.h. To reduce the rate of equilibration during dehydration to 75% r.h. and below, crystals were transferred to and then harvested from oil [NVH oil (Cargille, Cedar Grove, New Jersey, USA) or LV cryo-oil (MiTeGen, Ithaca, New York, USA)] before dehydration, and this was found to reproducibly yield high-quality diffraction. On the other hand, crystal dehydration to 85% and 93% r.h. can be excessively slowed by the surrounding mother liquor, so excess mother liquor was carefully removed. Crystals were equilibrated in the capillary tubes for ~24–48 h in all cases, which was experimentally determined to be sufficient for equilibration and to yield unit-cell volumes that were consistent with those from previous studies (Nagendra *et al.*, 1995; Harata & Akiba, 2006).

2.3. X-ray data collection

X-ray diffraction data were collected on MacCHESS station A1 at the Cornell High Energy Synchrotron Source (CHESS) using a 19.6 keV X-ray beam with a diameter of 100 µm. Diffraction images were recorded using either an ADSC Quantum 210 or a Dectris PILATUS 6M detector. Complete room-temperature data sets were collected from 2–5 different crystals at each relative humidity, and the models were refined using the highest resolution data sets at each relative humidity.

To compare the effects of dehydration with the effects of cryocooling, additional X-ray data sets were collected at $T = 100$ K using a nitrogen-gas cryostream from native monoclinic lysozyme crystals (99% r.h.) and from crystals dehydrated at 298 K to 85, 75, 58 and 11% r.h. No penetrating cryoprotectants were used, to ensure that the crystal composition was identical for 298 and 100 K data collection. Each crystal was looped from its drop and excess mother liquor was removed by blotting. The crystal was then transferred to LV CryoOil, harvested using a microfabricated polyimide loop,

Table 1

Data-collection and refinement statistics for room-temperature data collected using monoclinic lysozyme crystals.

Values in parentheses are for the highest resolution shell.

	99% r.h.	93% r.h.	85% r.h.	75% r.h.	58% r.h.	33% r.h.	11% r.h.
Wavelength (Å)	0.6299	0.6309	0.6309	0.6309	0.6309	0.6309	0.6299
Space group	$P2_1$	$P2_1$	$P2_1$	$P2_1$	$P2_1$	$P2_1$	$P2_1$
Unit-cell parameters							
a (Å)	28.13	27.62	27.40	26.73	26.36	25.32	25.33
b (Å)	63.33	62.84	62.61	58.44	57.21	56.84	56.39
c (Å)	60.72	58.96	58.79	31.10	31.00	30.75	30.46
β (°)	90.61	91.27	90.86	111.81	111.54	109.55	109.76
Solvent content (%)	34.9	31.2	30.2	22.0	19.1	15.6	14.1
Total reflections	706703	174070	253172	112196	53402	28031	43782
Multiplicity	12.4 (5.3)	3.3 (2.4)	3.4 (2.6)	3.3 (2.7)	3.2 (2.2)	2.8 (2.1)	5.1 (2.8)
Completeness (%)	98.9 (98.9)	99.3 (93.2)	98.0 (98.0)	99.6 (96.9)	97.9 (82.1)	95.2 (84.1)	96.6 (78.8)
Mosaicity (°)	0.09–0.56	0.13–0.66	0.14–0.28	0.33–0.51	0.48–0.64	0.57–2.45	0.53–1.14
CC _{1/2}	1.000 (0.589)	0.996 (0.593)	1.000 (0.725)	0.996 (0.804)	0.980 (0.242)	0.957 (0.163)	0.996 (0.788)
Mean $I/\sigma(I)$	45.3 (1.0)	19.5 (0.9)	20.7 (1.1)	24.8 (2.0)	19.5 (0.5)	9.3 (0.5)	35.2 (20.6)
Resolution [at $I/\sigma(I) = 2$] (Å)	1.35	1.31	1.19	1.11	1.54	1.94	1.70
$R_{\text{merge}}/R_{\text{meas}}$ (%)	5.6/5.8	10.4/12.3	4.3/5.1	3.5/4.2	5.7/6.6	16.4/19.4	7.0/7.8
Wilson B factor (Å ²)	20.4	14.5	11.9	11	22.6	26.3	28.9
Refinement resolution range (Å)	28.13–1.26	27.61–1.24	24.69–1.14	16.09–1.11	16.00–1.38	23.86–1.61	23.84–1.69
$R_{\text{work}}/R_{\text{free}}$ (%)	12.9/15.8	14.2/17.2	12.2/15.2	10.5/13.2	17.3/22.6	23.8/29.1	22.0/26.6
Test set for R_{free} (%)	5	5	5	5	5	5	5
R.m.s.d. bonds (Å)	0.015	0.013	0.026	0.011	0.005	0.003	0.003
R.m.s.d. angles (°)	1.46	1.36	1.46	1.37	0.73	0.003/0.76	0.003/0.73
Ramachandran favored/outliers (%)	97.3/0.3	98.3/0.0	96.9/0.0	96.1/0.0	95.4/0.5	96.3/0.7	94.6/0.7
Rotamer outliers (%)	2.6	0.2	0.9	0.4	0.4	2.61	6.3
Clashscore	3.8	0.9	4.2	3.4	1.1	3.8	6.2
B factors (Å ²)							
Overall	28.4	19.7	14.8	13.6	28.6	32.7	38.7
Macromolecule	27.9	18.9	13.9	13.0	28.2	32.5	38.6
Solvent	35.1	32.8	30.4	25.9	39.0	38.0	40.8
Waters (first hydration shell/total)	128/137	177/180	212/215	90/97	53/54	40/42	15/18

and inserted into a capillary tube containing salt solution. After equilibration for 24–48 h, a few frames of X-ray data were first collected at 298 K with the capillary tube in place and with the nitrogen cryostream blocked using an ‘air-blade’ shutter, to ensure that a unit cell consistent with the target dehydrated state had been obtained and that the diffraction was to high resolution. The capillary tube was then removed and the cryostream unshuttered to cool the sample, and a complete and redundant data set was collected. The maximum time that the sample was exposed to ambient humidity (20–45% r.h.) air or dry N₂ while its temperature was warm enough (>230 K) for dehydration to occur was less than 3 s. Given the relatively large crystal sizes, this was too short for appreciable changes in crystal hydration. At the lowest relative humidity of 11%, the 100 K unit cell for each crystal was smaller than the room-temperature unit cell recorded before the capillary tube was removed, indicating that no significant rehydration occurred. High-quality, ice-free $T = 100$ K data with consistent unit cells were obtained for all relative humidity values without the use of penetrating cryoprotectants. Data-collection and refinement statistics are given in Tables 1 and 2.

2.4. Data processing

Data frames were indexed and scaled using *SCALEPACK* and *DENZO* (Otwinowski & Minor, 1997), and molecular replacement and model refinement were performed using

PHENIX (Adams *et al.*, 2010). PDB entries 2d4k and 2d4j were used as initial models for crystals with relative humidities above and below the ~85% r.h. structural phase transition, respectively. Water molecules were modeled in all structures, and for data with resolution better than 1.4 Å anisotropic B factors were refined and H atoms were modeled.

Standard refinement programs, by default, fit a single conformer model to the experimental electron-density map. Initial multiconformer models were generated using *qFit* (van den Bedem *et al.*, 2009), which builds up to four alternative conformations per residue where they are locally supported by the real-space electron-density map. *Ringer* was used to detect additional rotameric conformations of each side chain, and the *qFit* models were then manually inspected, modified and validated using *Coot* (Emsley *et al.*, 2010). Final model validation was performed using *MolProbity* in the *PHENIX* software package (Chen *et al.*, 2010). The resulting multiconformer models led to substantial reductions in R_{free} values (up to ~2%), especially for high-resolution structures.

Protein volumes within the unit cell were calculated as the volume enclosed by the solvent-excluded surface (SES) using the program *3vee* with a probe size of 1.4 Å and a grid size of 0.07 Å. The SES is calculated by ‘rolling’ a fixed-diameter probe over the surface of a voxelized three-dimensional model of the protein (Voss & Gerstein, 2010). H atoms were added (if absent) and alternate conformations of side chains were removed (as the software assumes a fixed occupancy of 1 for all side chains) from the model before using *3vee*. The

Table 2
Data-collection and refinement statistics for data collected from monoclinic lysozyme crystals at $T = 100$ K.
Values in parentheses are for the highest resolution shell.

	99% r.h. cryo	85% r.h. cryo	75% r.h. cryo	58% r.h. cryo	11% r.h. cryo
Wavelength (Å)	0.6307	0.6307	0.9768	0.9768	0.6307
Space group	$P2_1$	$P2_1$	$P2_1$	$P2_1$	$P2_1$
Unit-cell parameters					
a (Å)	27.39	26.74	26.22	26.23	25.03
b (Å)	62.09	62.09	58.01	57.16	56.41
c (Å)	59.28	58.21	30.67	30.69	30.30
β (°)	90.36	90.32	111.64	111.35	109.14
Solvent content (%)	30.2	27.2	20.9	16.7	13
Total reflections	52483	74056	28918	23476	7418
Multiplicity	6.7 (4.0)	7.8 (4.5)	8.0 (4.2)	6.0 (3.4)	2.5 (2.1)
Completeness (%)	97.0 (97.0)	97.0 (97.0)	96.9 (95.1)	98.4 (96.8)	96.0 (96.0)
Mosaicity (°)	0.43–0.61	0.41–1.34	0.46–0.65	0.32–0.85	0.74–1.56
$CC_{1/2}$	0.996 (0.762)	0.996 (0.858)	1.000 (0.957)	0.996 (0.909)	0.992 (0.885)
Mean $I/\sigma(I)$	34.0 (1.7)	36.5 (1.8)	40.3 (4.4)	34.7 (3.6)	10.2 (0.6)
Resolution [at $I/\sigma(I) = 2$] (Å)	1.26	1.10	1.17	1.25	1.80
$R_{\text{merge}}/R_{\text{meas}}$ (%)	6.8/7.4	6.5/6.9	5.5/6.1	7.3/8.1	9.2/9.8
Wilson B factor (Å ²)	11.9	7.5	11.7	14.5	20.9
Refinement resolution range (Å)	27.50–1.26	29.10–1.10	23.24–1.16	28.58–1.24	28.62–1.77
$R_{\text{work}}/R_{\text{free}}$ (%)	13.4/16.7	12.9/15.2	11.2/14.3	14.3/17.9	18.2/23.1
Test set for R_{free} (%)	5	5	5	5	5
R.m.s.d. bonds (Å)	0.032	0.017	0.010	0.008	0.0087
R.m.s.d. angles (°)	1.77	1.68	1.14	1.00	1.16
Ramachandran favored/outliers (%)	98.2/0.0	98.6/0.0	96.0/0.4	97.4/0.0	95.3/2.3
Rotamer outliers (%)	0.5	0.9	0.83	0.59	3.7
Clashscore	3.6	5.5	5.2	4.7	5.6
B factors (Å ²)					
Overall	16.8	11.0	13.9	19.5	22.7
Macromolecule	15.6	9.2	13.1	18.6	22.4
Solvent	27.8	20.6	24.0	29.1	29.7
Waters (first hydration shell/total)	320/326	363/384	133/143	102/106	35/41

fractional solvent-accessible surface area for each residue was calculated using the web server at <http://cib.cf.ocha.ac.jp/bitool/ASA/>. The solvent volume within the unit cell was calculated by subtracting the SES volume of protein from the unit-cell volume. These results were compared with those obtained using a Matthews coefficient analysis.

3. Experimental results and analysis

3.1. Unit-cell, protein and solvent volumes

Monoclinic crystals of hen egg-white lysozyme undergo a structural transition near 85% r.h. in which the two crystallographically independent molecules in the native asymmetric unit become equivalent, resulting in a halving of the unit cell (Salunke *et al.*, 1985). The crystals can maintain excellent order down to 5% r.h., corresponding to a Matthews coefficient-derived solvent content of 9% by volume (Nagendra *et al.*, 1998). Fig. 1 shows how the unit-cell volume, protein volume and solvent volume per protein molecule vary with relative humidity at 298 and 100 K. At 298 K the unit-cell volume decreases by 24% between 99 and 11% r.h. On cooling to 100 K, the unit-cell volume contracts by 6.8% at 99% r.h., decreasing to only 1.3% at 11% r.h. The SES protein volume (Supplementary Table S1) decreases by only 1–2% over this relative humidity range and by 0.5–2.4% at fixed relative humidity (with no systematic dependence on relative humidity) on cooling to 100 K. At 99% r.h., nearly all of the

Figure 1
Crystallographic unit-cell volume (blue squares), protein volume [red triangles; calculated as the volume enclosed by the solvent-excluded surface (SES) area of the protein molecule] and solvent volume (orange circles; given by the difference between unit-cell and protein volume) per protein molecule for monoclinic lysozyme crystals *versus* relative humidity. Dark and light symbols are values at room temperature and $T = 100$ K, respectively. The discontinuity in unit-cell and solvent volume between 85 and 75% r.h. is due to a structural transition in which the unit-cell volume is roughly halved. The dashed lines are guides to the eye through the room-temperature data points.

6.8% unit-cell contraction on cooling to $T = 100$ K is due to the solvent, whereas at 11% r.h. the cell, protein and solvent have comparably small contractions. Note that cooling the native (99%) structure to 100 K has a similar effect on unit-cell and solvent volume as does dehydration to $\sim 85\%$ r.h. Note that the Matthews coefficient analysis overestimates the solvent content at all relative humidities and temperatures, with the largest fractional errors occurring at low relative humidity and $T = 100$ K. These errors arise because our calculated $T = 298$ K lysozyme density (1.41 g ml^{-1} for the native structure) is larger than the assumed density (1.35 g ml^{-1}) in, for example, *PHENIX* (Quillin & Matthews, 2000; Fischer *et al.*, 2004), because the protein density is larger (1.44 g ml^{-1}) at 100 K, and because errors in protein density have a larger effect on solvent fraction at low solvent fractions.

3.2. Solvent content and solvent ordering

Fig. 2 shows the crystal solvent content, as determined from the unit-cell volume and *3vee* analysis, as a function of relative humidity at both 298 and 100 K, as well as the resolution of the refined crystallographic data sets. The $T = 298$ K solvent content decreases from 0.37 to $\sim 0.34(v/v)$ between 99 and 85% r.h., abruptly drops to $0.26(v/v)$ as a result of the structural transition and then decreases to $\sim 0.18(v/v)$ at 11% r.h. Dehydration to 75% r.h. leads to a small improvement in resolution. Further dehydration degrades the resolution by a variable amount that tends to decrease as the rate of

Figure 2

Crystal solvent content (derived from analysis of the solvent-excluded protein volume as in Fig. 1) and data-set resolution as a function of relative humidity. Red and blue symbols represent values at room temperature and 100 K, respectively. The dashed lines are guides to the eye through the room-temperature data points. The enzymatic activity of lysozyme ceases below 0.2 g of water per gram of protein molecule, corresponding to 21% (v/v) solvent content. Inset: ratio of the crystallographically detected hydration-shell volume to the total solvent volume as a function of relative humidity for the five highest resolution data sets, assuming unit occupancy for all modeled waters. Near 75% r.h., the ratio of the hydration solvent volume to the total solvent volume at cryogenic temperature has a maximum. The decrease in hydration-shell volume between 75 and 58% r.h. is due in part to an increase in overall and hydration-shell B factors so that fewer waters were modeled. Truncation of all data sets to $1.5 \text{ \AA}$ resolution had little effect on hydration-shell volumes.

Table 3

Fraction of water sites modeled at 298 K that are conserved on cooling to 100 K versus relative humidity.

Relative humidity	Fraction of waters conserved (298→100 K)
99	0.85
85	0.81
75	0.84
58	0.74
11	0.50

dehydration decreases and that also depends on how the shape and initial perfection of each crystal affects its response to stresses during the structural transition near 85% r.h. (Dobrianov *et al.*, 2001).

The internal crystal solvent can be crudely separated into hydration water (water contained within the first hydration shell that interacts strongly with the protein and stabilizes its secondary and tertiary structure) and bulk-like water. Dehydration first depletes bulk-like water before affecting hydration water. The enzymatic activity of lysozyme ceases at below 0.2 g of water per gram of lysozyme molecules (Rupley *et al.*, 1983; Rupley & Careri, 1991). This corresponds to a $\sim 22\%$ (v/v) solvent content which, for monoclinic lysozyme, occurs below the structural transition at $\sim 75\%$ r.h.

In crystallography, hydration waters are more likely to be found at particular sites near the protein surface, especially at $T = 100$ K. Many hydration-shell water sites in lysozyme are invariant with respect to the crystal space group (Madhusudan & Vijayan, 1991). The inset in Fig. 2 shows the ratio of the crystallographically detected hydration-shell volume to the total solvent volume as a function of relative humidity for the five highest resolution data sets (99–58% r.h.) at both 298 and 100 K. Waters were modeled only if the resulting B factor was less than $60 \text{ \AA}^2$. The hydration-shell volume was determined using the *3vee* server by subtracting the solvent-excluded volume of the protein and modeled waters (at 100% occupancy) and of the protein only. The detected hydration-shell volume roughly doubles on cooling to 100 K at a fixed relative humidity and the average B factor of the modeled waters decreases from ~ 35 to $\sim 28 \text{ \AA}^2$ for the native crystal. This reflects increased solvent ordering and localization when thermal fluctuations are reduced (Nakasako, 2004; Nakasako *et al.*, 1999). At $\sim 75\%$ r.h., corresponding to the threshold for enzymatic activity, almost all disordered ‘bulk’ solvent has been depleted, and further dehydration depletes hydration water.

Table 3 indicates the extent to which modeled water sites are conserved on cooling at fixed relative humidity. A water site is deemed to be ‘conserved’ between two structures if it is displaced by no more than a cutoff distance of $1.2 \text{ \AA}$ after least-squares superposition of the structures. The cutoff distance to identify conserved water sites in the literature varies from 1 to $2.4 \text{ \AA}$, with values between 1 and $1.5 \text{ \AA}$ being most common (Poornima & Dean, 1995; Shaltiel *et al.*, 1998; Zhang & Matthews, 1994). For the present data, varying the threshold between 1 and $1.5 \text{ \AA}$ produced only small fractional increases/decreases in the number of conserved sites. At fixed

relative humidity, the fraction of water sites detected at 298 K that are conserved at 100 K decreases from $\sim 85\%$ at 99% r.h. to $\sim 74\%$ at 58% r.h., and then drops to only $\sim 50\%$ at 11% r.h. Waters at the active site are mostly conserved down to 75% r.h. and then begin to disappear on further dehydration.

3.3. Crystal contacts

Increased protein packing due to the removal of crystal solvent increases the number and area of intermolecular contacts. Various criteria can be used to determine whether two atoms in different molecules within a crystal are in contact, for example, if their separation is below a threshold of 4–6 Å (Dasgupta *et al.*, 1997; Lorimer *et al.*, 2009; Carugo & Argos, 1997). Fig. 3 shows the evolution of the number and fraction of residues involved in crystal contacts *versus* relative humidity at $T = 298$ and 100 K, assuming that atoms are in contact if their center-to-center distance is less than the sum of their van der Waals radii plus 0.25 Å (Juers & Matthews, 2001). As shown in Supplementary Fig. S2, qualitatively similar behavior is observed if a fixed separation threshold of 4 Å is used.

At 298 K the fraction of residues involved in crystal contacts increases from $\sim 32\%$ at 99% r.h. to $\sim 50\%$ at 11% r.h. Cooling to 100 K increases the 99% r.h. contact fraction to 40%, but has no significant effect at 11% r.h. Comparing the effects of cooling and dehydration on the native (99% r.h.) structure, cooling to 100 K produces a comparable increase in crystal contacts as dehydration to

Figure 3

Number of residues involved in crystal contacts *versus* relative humidity at room temperature (red) and $T = 100$ K (blue). The dashed red line is a guide to the eye for the room-temperature data. As the relative humidity decreases, the protein molecules become more densely packed, increasing the number of contact residues. The number of contacts detected depends on the assumed cutoff separation between residues. Here, two atoms are considered to be in contact if their center-to-center distance is less than the sum of their van der Waals radii plus 0.25 Å, for example 0.2 Å. Supplementary Fig. S2 shows the number of contact residues determined using a fixed 4 Å cutoff. The dashed line is a guide to the eye for the room-temperature data points.

Table 4

Modifications of the 103 lysozyme residues having χ_1 rotamers due to dehydration and cryocooling relative to the native (99% r.h.) room-temperature structure.

Modifications were determined by visual inspection of *Ringer* plots as in Fig. 4. Minor conformers above the 0.3σ electron-density threshold were deemed to be suppressed if their density fell below the threshold or decreased substantially relative to the major conformer. A minor conformer was deemed to be enhanced if it became detectable above the threshold or if its above-threshold density increased while that of the major conformer decreased. A new conformation was identified if there was a major change in position of the major conformer or of a high-density minor conformer. This qualitative assessment more readily identifies changes in minor conformations than does the Pearson correlation coefficient of Fig. 6, as the latter weights large electron-density features more heavily.

	Suppressed	Enhanced	New conformer
99→93% r.h.	11	2	4
99→85% r.h.	14	6	3
99→75% r.h.	14	10	5
99% r.h.→ $T = 100$ K	19	3	3
85% r.h.→ $T = 100$ K	12	6	2
75% r.h.→ $T = 100$ K	8	7	3
58% r.h.→ $T = 100$ K	11	10	3
11% r.h.→ $T = 100$ K	11	9	12

85% r.h., consistent with the similar changes in unit-cell volume that these perturbations produce.

3.4. Side-chain remodeling: *Ringer* analysis

Intermolecular crystal contacts affect side-chain heterogeneity, as contacts formed by electrostatic interactions impose some local restrictions on the side-chain conformations of the residues involved. Previous analysis using *Ringer* has shown that cryocooling induces changes in electron density sampled in rings around side-chain dihedral angles χ , and these have been interpreted as evidence for conformational remodeling of side-chain rotamers. As shown in Fig. 4 for between 99 and 85% r.h., dehydration induces similar changes, including the suppression of minor (low-occupancy) conformers (Fig. 4b), the enhancement of minor conformers (Fig. 4c) and the appearance of new conformations (Fig. 4d). Lysozyme has 129 residues, of which 103 have side-chain rotamers with a dihedral χ_1 angle and rotamer. As shown in Table 4, on dehydration from 99 to 85% r.h. 12 residues showed suppression of a minor conformer, seven showed systematic enhancement of a minor conformer, three adopted new conformers and the remaining 81 residues showed no change. Cryocooling of the native (99% r.h.) crystal causes the suppression of more minor conformers, the enhancement of fewer minor conformers and the generation of a comparable number of new conformations. As relative humidity is decreased, cryocooling still causes substantial side-chain remodeling, even though the change in unit-cell volume on cooling becomes smaller and smaller.

No correlation is observed between nitrate-ion position and occupancy and side-chain remodeling. The number of modeled nitrates per lysozyme molecule varied with relative humidity: at $T = 300$ K it increased from 2.5 (five in the asymmetric unit) at 99% r.h. to four at 85% r.h., decreasing to three at 75% r.h. and zero at 11% r.h. This variation largely

follows the variation in diffraction resolution, and so does not necessarily reflect variations in the actual numbers of nitrates or their ordering.

3.4.1. Contact versus noncontact residues. To understand how the local environment of a side chain affects its remodeling during dehydration or cooling, residues were first categorized as contact or noncontact based upon their final (dehydrated and/or cooled) side-chain state. To assess whether the χ_1 side-chain rotamer was remodeled, the Pearson correlation coefficient between *Ringer* plots for the initial and final states was calculated for each residue (Keedy *et al.*, 2015). Pearson correlation coefficients can range from -1 to 1 , and are independent of the absolute value of the electron density. They are only weakly dependent on *B* factors and resolution, and blurring or sharpening of the electron density has only a small effect. Correlation coefficients below 0.85 but well above 0 generally indicate the remodeling of minority conformations, whereas negative values usually correspond to changes in the majority conformation. Side chains with correlation coefficients below 0.85 , a conservative cutoff, were thus considered to be remodeled.

Fig. 5(a) gives the number of altered contact and noncontact residues, relative to the native (99% r.h., 298 K) structure, versus the relative humidity at 298 and 100 K. At

room temperature, the majority of altered residues are involved in crystal contacts, although the fraction of altered noncontact residues grows substantially below 58% r.h. Cooling the native crystal to 100 K alters roughly the same number of residues as does dehydrating to 93% r.h., although cooling affects more noncontact residues.

3.4.2. Solvent-exposed and buried residues. Solvent accessibility might also be expected to influence which residues are altered. Solvent accessibility was assessed by calculating the solvent-accessible surface area (SASA) using the *PyMOL* function *get_area* with a default probe size of $1.4 \text{ \AA}$ and then normalizing by the maximum surface area of each residue type to obtain the fractional SASA. The maximum surface area of each residue type was estimated using an isolated tripeptide Gly-*X*-Gly in which *X* is the residue of interest. In previous work, fractional SASA values below 5% and below 20% were most commonly used to define the cutoff for buried residues (Zhang *et al.*, 2009; Miller *et al.*, 1987; Chen & Zhou, 2005). A survey of 37 different protein structures found that 15% of all residues in small proteins and 32% in large proteins were buried assuming a 5% cutoff (Miller *et al.*, 1987). Lysozyme is a small protein and Supplementary Fig. S3 plots the number of buried residues versus the fractional SASA cutoff. The number of buried residues versus the cutoff

exhibits a change in slope near 3%. With a 3% cutoff buried residues comprise 22–24% of all residues for all humidities, consistent with the survey. Residues with fractional SASA values smaller than 3% were thus defined as buried.

Fig. 5(b) gives the number of altered residues that are solvent-accessible and buried in the final state. Dehydration to relative humidity values of between 93 and 58%, as well as cryocooling of these dehydrated samples, has almost no effect on buried residues.

3.4.3. Overall side-chain correlations. Fig. 6(a) shows overall pairwise Pearson correlation coefficients for χ_1 side-chain rotamers, which provide an overall measure of the similarity of side-chain conformations. These were obtained by averaging pairwise side-chain correlations calculated from *Ringer* plots over the 103 residues having χ_1 side-chain rotamers, between structures at 298 and 100 K at different relative humidities. As one would expect, the correlations decrease as the difference in relative humidity increases. As seen in the leftmost column of Fig. 6(a), cooling the native structure ('h99') to 100 K ('h99c') gives the same correlation coefficient as dehydrating the native structure at

Figure 4

Example plots of $2mF_o - DF_c$ electron density (σ) for χ_1 rotamers as a function of dihedral angle determined using *Ringer* for three humidities above the unit-cell transition at room temperature. The assumed 0.3σ noise level for each map is indicated by the dashed line. (a) Val29 maintains its conformation during dehydration. (b) An alternative conformation of Asp18 in the native structure (99% r.h.) is systematically suppressed with decreasing humidity. (c) An alternate conformation of Asn77 is enhanced with decreasing humidity. (d) Arg125 develops a new conformation at lower humidity. In this and subsequent figures, 99% r.h. is abbreviated as h99, *etc.*

298 K to 85% r.h. ('h85'). As seen by inspecting the diagonal boxed entries in Fig. 6(a), for crystals dehydrated to the same relative humidity, the largest correlation between room-temperature and 100 K structures occurs for 85% r.h., which is also the humidity that yields the highest resolution 100 K data set. Similar patterns are evident in Supplementary Fig. S4 for χ_2 , which is calculated by averaging over 72 residues having χ_2 rotamers. Similarly, as indicated by the entries within the horizontal rectangles in Fig. 6(a) and Supplementary Fig. S4,

Figure 5

Number of altered residues, of the 103 lysozyme residues having χ_1 rotamers, due to dehydration and cryocooling, relative to the native (99% r.h.) structure, categorized according to (a) whether or not the residues were involved in crystal contacts in the final, non-native structure and (b) whether the residues were solvent-exposed or buried. Here, side chains were deemed to be altered if the Pearson correlation coefficient between *Ringer* curves for the native and non-native structures was less than 0.85. A large majority of the altered side chains are involved in crystal contacts and/or are solvent-exposed in non-native room-temperature and 100 K structures, although noncontact and buried residues are substantially perturbed at 33 and 11% r.h. Dehydrating the native crystal to 93% r.h. alters a comparable number of residues as does cryocooling to 100 K.

the χ_1 rotamers of the native (99% r.h.) structure at 100 K are most strongly correlated with those in the 85% r.h. room-temperature structure, and for χ_2 rotamers with the 93% r.h. room-temperature structure. This demonstrates the structural similarity between the effect of cryocooling and dehydration to 85–93% r.h. of the native crystal in terms of the side-chain conformations and heterogeneity. Fig. 6(b) shows the room-temperature molecular structure in the native state color-coded according to the χ_1 correlation of each residue with the 75% r.h. structure.

3.4.4. Backbone r.m.s.d. Lysozyme does not exhibit major structural rearrangement of its backbone on dehydration or cooling, but smaller changes are evident, especially in turn regions. Fig. 6(c) shows the pairwise backbone root-mean-squared deviation (r.m.s.d.) between all dehydrated and cooled structures. R.m.s.d. values increase with increasing humidity difference. As with side-chain correlations, cryocooling the native structure produces overall changes in the backbone that are comparable to those obtained by dehydration to 85% r.h. at room temperature. Surprisingly, 298 K backbones at 58, 33 and 11% r.h. have smaller r.m.s.d. values with the native structure than with 93–75% r.h. structures. A similar although less strong trend is seen in Fig. 6(a) for side-chain correlations. As is the case with side-chain correlations discussed above, the smallest backbone perturbation caused by cryocooling at a fixed relative humidity is observed for 85% r.h.

3.5. Multiconformer modeling and side-chain order parameters

Overall side-chain conformational heterogeneity can be characterized using bond-order parameters S^2 , which include contributions from both harmonic and nonharmonic motions as well as from static disorder. The order parameter S^2 corresponding to a dihedral angle χ_1 can be decomposed into two components: S^2_{ortho} , which is associated with displacements around a conformer at a given χ_1 value, and S^2_{angular} , which is associated with displacements between conformers at different χ_1 values. At each relative humidity, *qFit*, *Ringer* and manual building were used to construct a multiconformer model. Order parameters S^2 were then calculated for χ_1 rotamers about bond vectors $C^\beta-X'$, where X is C, O or S, for most residues, about $C^\alpha-C^\beta$ for Ala and about $C^\alpha-H^{\alpha 2}$ for Gly. Supplementary Fig. S5 shows sample plots of order parameters S^2 and their component S^2_{ortho} versus relative humidity for several residues near the active site of lysozyme for relative humidity between 99 and 75%, for which the dataset resolution was high and nearly constant. Truncation of all data sets to 1.4 Å resolution had little effect on these results. Of the 103 residues with side-chain rotamers, 76 showed an increase in S^2 with decreasing relative humidity, and nearly all showed an increase in S^2_{ortho} .

3.6. Modulation of the active site

The active site of lysozyme, which in the native state is largely hydrated by ordered water molecules, undergoes

significant remodeling on dehydration. The most pronounced change occurs at Trp62, which undergoes motion in substrate binding (Kelly *et al.*, 1979; Johnson & Phillips, 1965). Trp62 is somewhat disordered in the native crystal, and on dehydration to 85% r.h. it loses one of its minor conformers while maintaining its major conformation. At relative humidity values below that of the crystal lattice transition, Trp62 undergoes a large conformation change that has been suggested to be similar to its behavior during substrate binding (Nagendra *et al.*, 1996; Kodandapani *et al.*, 1990).

The present data show that this conformation change results from a steric clash with a symmetry mate. Fig. 7 shows the two

side chains Trp62 and Arg73 at the active site of native (orange) and 75% r.h. (blue) structures, as well as the nearby side chains of a symmetry mate at 75% r.h. (cyan). In the native crystal, Arg73 has one alternate conformation and both conformations freely extend away from the center of the reference protein molecule. On dehydration from 85 to 75% r.h., a significant loss of solvent associated with the lattice transition leads to a tighter packing of protein molecules that forces Arg73 to change its conformation. This conformation change causes an $\sim 90^\circ$ rotation in the side chain of Trp62 to prevent a steric clash. The low-humidity conformation, which stacks Arg73 onto the Trp62 ring, is also stabilized by a favorable cation– π interaction.

3.7. Dehydration- and cryocooling-induced changes in tetragonal thaumatin crystals

A more limited investigation of the effects of dehydration and cryocooling was conducted using crystals of tetragonal thaumatin. These crystals were grown in hanging drops comprised of equal volumes of 40 mg ml⁻¹ protein in 0.1 M sodium acetate buffer pH 6.5 and reservoir solution consisting of 0.5 M potassium sodium tartrate in the same buffer. The buffer gave a relative humidity for the native crystal of 99%, and dehydration using saturated solutions of potassium nitrate, zinc sulfate and sodium chloride gave relative humidity values of 93, 89 and 75%, respectively. Complete and redundant X-ray data sets were collected from these crystals and from cryoprotectant-free native crystals cooled to 100 K,

Figure 6

(a) Pairwise Pearson correlation coefficients for *Ringer* plots at χ_1 as in Fig. 4, averaged over all 103 residues, for all possible pairs of dehydrated and cryocooled structures. Data sets collected at 100 K are labelled with ‘-c’. The leftmost column indicates correlations with the native structure at room temperature; the diagonal boxed entries indicate correlations between room-temperature and 100 K structures for crystals dehydrated to the same relative humidity, and the horizontal rectangle highlights correlation between the 100 K native structure and dehydrated structures at room temperature. (b) Lysozyme structure color-coded according to the Pearson correlation coefficient between the native and 75% r.h. room-temperature structures, calculated using the χ_1 *Ringer* plot for each residue. All residues with lower correlation coefficients are located on the surface, while buried residues generally have high correlation coefficients. The active site and its vicinity exhibit intermediate correlation. (c) Backbone r.m.s.d. between all possible pairs of dehydrated and cryocooled structures. Only backbone atoms were used for the alignment of the two structures and calculation of the r.m.s.d.

and these data were modeled and refined as for lysozyme. Supplementary Table S2 gives the resulting refinement statistics.

Thaumatococcus has 206 residues compared with 129 for lysozyme, and its tetragonal crystals have a Matthews coefficient-derived solvent content of 58.5%, compared with 37.5% for monoclinic lysozyme. As shown in Supplementary Fig. S6, the unit cell contracts by about 18% between 99 and 75% r.h. and by about 4% on cooling to 100 K, compared with contractions of ~ 20 and 7% for monoclinic lysozyme. For both proteins, nearly all of the contraction is due to a reduction in solvent volume, and this contraction is associated with an increase in the fraction of residues involved in crystal contacts (Supplementary Fig. S7).

As shown in Supplementary Fig. S8, the absolute number and fraction of residues whose χ_1 rotamers are altered by dehydration and by cooling are much smaller for tetragonal thaumatococcus than for monoclinic lysozyme: at 75% r.h. the altered fractions are ~ 0.07 and ~ 0.19 , respectively. The fraction of altered residues that are involved in crystal contacts (0.6 versus 0.88 at 75% r.h.) is much smaller for tetragonal thaumatococcus. As shown in Supplementary Fig. S9, the pairwise Pearson correlation coefficients for *Ringer* plots at χ_1 , averaged over all residues with side-chain rotamers, are larger for all relative humidity pairs and all room–cryo temperature pairs for thaumatococcus than for lysozyme. Moreover, an inspection of *Ringer* plots and correlation coefficients for residues flagged as altered shows that side-chain conformational changes are on average smaller for thaumatococcus than for lysozyme. These results are consistent with the much larger

solvent content [0.49 versus 0.22(v/v) at 75% r.h.] and less intimate contact between molecules in thaumatococcus.

Supplementary Figs. S8 and S9 show that cryocooling the native state of thaumatococcus to 100 K produces a comparable increase in the number of contact residues and a comparable fraction of altered residues as does dehydration to 89% r.h. Furthermore, Supplementary Fig. S8(a) shows that the ratio of noncontact to contact residues altered by cooling to 100 K and by dehydration to 89% r.h. are similar. This suggests a connection between conformational perturbations caused by cooling and dehydration.

4. Discussion

4.1. Unit-cell, protein and solvent volumes

The use of carefully prepared, cryoprotectant-free crystals with identical solvent compositions for room-temperature and 100 K measurements allows quantitative comparison of unit-cell, protein and solvent volumes. Both dehydration from 99% r.h. to as low as 11% r.h. and cooling from 298 to 100 K lead to only small (1 and 2%, respectively) changes in protein volume, as determined from the calculated volume of the SES. On the other hand, the calculated solvent volume at 99% r.h. decreases on cooling to 100 K by $\sim 14\%$ (from the SES volume) or $\sim 19\%$ (from the Matthews coefficient analysis). At 11% r.h., the calculated solvent-volume contractions are much smaller ($\sim 0\%$ from the SES or $\sim 9\%$ from the Matthews coefficient) and have much larger uncertainties ($\sim 5\text{--}7\%$ of the solvent volume) because the solvent volumes are much smaller. SES protein volumes are calculated assuming a single conformer with no thermal motion, and so do not include the effects of alternate conformers or finite atomic *B* factors. To a first approximation (assuming fixed, temperature-dependent bond lengths and no overlap of nonbonded atomic van der Waals spheres), protein motions should not affect the instantaneous and thus time-averaged protein volume. As an upper bound on the possible effects of conformational heterogeneity, SES volumes were calculated assuming that all side-chain conformers were populated with unit occupancy. This increased the 298 K SES volumes by $\sim 5\%$ at 99% r.h. and $\sim 1\%$ at 11% r.h.

4.2. Comparison of dehydration and cooling

Previous work has explored how cryocooling modifies unit-cell and protein volumes (Juers & Matthews, 2001), overall protein structure (Tilton *et al.*, 1992; Teeter *et al.*, 2001) and a protein's ensemble of side-chain conformations (Lang *et al.*, 2010; Fraser *et al.*, 2011; Keedy *et al.*, 2014, 2015). The present work shows that the effects of modest dehydration on these properties are similar, both superficially and in many details. Cooling native (99% r.h.) monoclinic lysozyme crystals to 100 K has similar effects on unit-cell and protein volume, on the number of crystal contacts, on backbone structure (as measured by the r.m.s.d.) and on side-chain structure (as measured by the χ_1 and χ_2 correlation coefficients) as does dehydration to 85% r.h. Cooling native crystals to 100 K modifies a comparable fraction of side-chain conformations at

Figure 7

The important active-site residue Trp62 undergoes a significant conformational change between the native (orange) and 75% r.h. (blue) structures. In the native crystal, Arg73 freely extends away from the active site. Dehydration brings the protein molecules closer together. In the 75% r.h. structure, a steric clash with the side chains of an adjacent molecule (cyan) changes the conformation of Arg73, which causes a 90° rotation of Trp62.

χ_1 as dehydration to 93%; both overwhelmingly affect solvent-exposed residues, while cooling may have slightly more effect on noncontact residues. Cooling and dehydration both cause a similar variety of side-chain modifications, including the suppression of minor conformers, the enhancement of minor conformers and the appearance of new conformations. The same general trends are observed in tetragonal thaumatin crystals, although side-chain modifications at a given relative humidity are less pronounced due to the much larger solvent content and reduced protein packing.

This similarity complicates the interpretation of differences between structures and especially side-chain conformational ensembles determined at room and cryogenic temperatures (Fraser *et al.*, 2011), especially when the detailed crystallization, harvesting, mounting and cryocooling histories of each crystal are either not known or not identical. Protein crystals regularly undergo modest dehydration during harvesting, mounting, storage and data collection at room temperature. Anecdotal evidence, such as substantial variations in deposited unit-cell volumes between similarly prepared crystals, suggests that dehydration was particularly frequent when room-temperature data collection involved drawing crystals into glass capillaries, wicking away excess liquid and carefully sealing both ends of the capillary with wax. Dehydration is likely to have been a primary cause of room-temperature crystal non-isomorphism; crystals that are reproducibly hydrated (by insertion into humidity-controlled gas streams or by equilibration with saturated salt solutions) show highly reproducible (to within 0.1%) unit-cell parameters (Farley & Juers, 2014). Moreover, dehydration to

93% r.h. and somewhat lower typically has no visible effects on crystals (for example, they do not crack as long as dehydration occurs gradually) and usually does not cause major degradation of diffraction, and so can be hard to detect. Consequently, careful control of crystal hydration for both room and cryogenic temperature data collection is essential if the effects of these two perturbations are to be deconvolved.

4.3. Why do cooling and dehydration generate similar structural perturbations?

Temperature affects chemical properties such as pH, the pK_a values of ionizable groups and their charges, and 'mechanical' properties such as the amplitude of thermal vibrations about each conformation, the occupancy of alternative conformations and the rate of interconversion between them. As temperature decreases, average bond lengths shrink, higher energy conformations are depopulated and the protein becomes more compact. More flexible regions at the protein surface that tend to be involved in crystal contacts contract more than the protein as a whole, decreasing the volume of the unit cell more than the protein volume, and increasing the number and area of contacts between molecules. Both protein compaction and increased protein–protein contacts stiffen the protein as whole, which further reduces the amplitude of thermal motions. These also introduce steric hindrances that can block minor or major conformations or lead to new conformations.

Dehydration affects chemical properties, including the concentrations of solutes within the solvent channels and the pH. By removing solvent, dehydration increases the number and area of contacts between molecules. These stiffen the protein as a whole, reducing the amplitude of thermal motions at fixed temperature. They introduce steric hindrances that can block minor or major conformations or lead to new conformations. As solvent is depleted from the first hydration layer, protein function ceases, in part due to a loss of conformational flexibility facilitated by mobile surface waters.

The strong similarity in the conformational responses of the protein to cooling and to modest dehydration suggests that, at least for monoclinic lysozyme crystals, the global character of these responses is dominated by the effects of increased molecular packing. If this is the case, then one might expect that plots of harmonic and overall conformational heterogeneity *versus* unit-cell size would be similar, regardless of whether the unit-cell changes were caused by temperature or dehydration. Fig. 8 shows example plots of

Figure 8
Disorder parameters $1 - S^2$ and $1 - S^2_{\text{ortho}}$ versus unit-cell volume for a selection of residues at χ_1 , as deduced using multiconformer refinement from room-temperature data sets collected at 99, 93, 85 and 75% r.h. For comparison, the disorder parameter for each residue and the cell volume of the native crystal at $T = 100$ K are also shown. The same data, plotted as S^2 versus relative humidity, are shown in Supplementary Fig. S5.

$1 - S^2$ and $1 - S_{\text{ortho}}^2$ for χ_1 rotamers at $T = 298$ K versus unit-cell volume, for relative humidity values between 99 and 75%, using the same data as in Supplementary Fig. S5. The right-most data point corresponds to the native 99% r.h. crystal at 298 K. For comparison, $1 - S^2$ and the unit-cell volume of the native crystal at 100 K are indicated. Plots for the majority of side chains have a 'hockey stick' shape, with the rate of increase of disorder with unit-cell volume being larger at larger unit-cell volumes. This is similar to the behavior that is observed when the side-chain order parameters of cyclophilin A were plotted versus temperature between 300 and 100 K (Keedy *et al.*, 2015). In both proteins, the majority of χ_1 rotamers have a single dominant conformation that does not change, so S^2 and S_{ortho}^2 are similar. Some do show conformational rearrangements that affect S_{angular}^2 and S^2 , but S_{ortho}^2 still generally increases with decreasing unit-cell volume. These results are consistent with the notion that unit-cell volume and protein packing dominate global aspects of the effects of both cooling and dehydration on protein conformation.

As previously established (see, for example, Radha Kishan *et al.*, 1995), the fact that unit-cell contraction may drive many of the conformational changes that are observed on cooling or dehydration in crystallography does not imply that these changes are biologically irrelevant. Crystal contacts can mimic some effects of ligand binding at active sites or of regulator binding at remote sites, and chains of steric hindrance triggered by increased contact may illuminate allosteric networks. Since crystal dehydration can generate much larger changes in unit-cell volume and in area of intermolecular contact than can cooling, it can be a highly useful probe of conformational interactions within proteins when adequate diffraction resolution is preserved.

Fig. 9 shows how the effects of dehydration and cooling can be understood and distinguished in terms of energy landscapes and occupancies. Both cause increased packing density, molecular stiffening and steric clashes that translate into changes in the local free-energy landscape and into changes in occupancies within that landscape. However, there are no obvious hydration-related analogs to the Boltzmann factor, or to thermal depopulation of minor conformers and reduced amplitudes of thermal motions within fixed energy landscapes. In general, distinguishing purely thermal effects from those associated with energy landscape changes, and distinguishing landscape changes intrinsic to the protein from those associated with the crystalline environment, require careful analysis of the environment of each residue.

5. Conclusions

Improved computational and modeling tools are facilitating the study of protein conformational ensembles using crystallography, and the analysis of these ensembles to identify functionally important motions and interactions. The electron density associated with minor conformations is generally weak and so can be difficult to positively identify, especially at the resolutions near and above 2 Å that are typical of many

targets of biological interest. Dehydration and cooling (as well as hydrostatic pressure; Collins *et al.*, 2011) provide easily controlled continuous crystal and molecular perturbations, and the resulting systematic changes in electron density should in many cases allow 'real' weak density features to be distinguished from artefacts.

The present results, which were obtained using crystals in carefully prepared reference states and subject to well defined perturbations, allow detailed and quantitative comparison of the local and global effects of dehydration and cryocooling on protein structure and conformational heterogeneity. In low-solvent-content, relatively tightly packed monoclinic lysozyme crystals, the effects of dehydration and cooling are strongly similar, suggesting that the changes in intermolecular and intramolecular packing interactions associated with unit-cell contraction dominate the observed behavior. Only modest dehydration, as might be expected to occur in past crystallographic practice (for example when using glass capillaries) or current practice (for example when using small crystals), is required to mimic many of the effects of cooling from room temperature to 100 K. Dehydration combined with differences/uncertainties in cryoprotectant concentrations thus complicates the interpretation of differences between room-temperature and 100 K structures.

Figure 9

The effects of dehydration and cooling on side-chain energy landscapes, shown as plots of free energy versus side-chain rotamer angle χ (black lines). Red lines indicate occupancy (electron density) at room temperature. The dashed blue line in (a) indicates occupancy at 100 K. (a) A potential with two wells. (b) A steric clash eliminates one well. (c) Stiffening of the potential of (b) due to, for example, increased packing density and steric encroachment. (d) A potential with a well at a new position relative to (a)–(c). Dehydration from state (a) at fixed temperature may lead to (b) elimination of a minor conformer, increasing S_{angular}^2 , (c) reduced amplitude of thermal motion about the major conformation, increasing S_{ortho}^2 , and (d) the appearance of a new major conformation. Cooling at fixed hydration can lead to similar changes in energy landscape associated with increased packing, and in addition to depopulation of the minor conformer in (a) and reduced amplitudes of thermal motion in (a)–(d), without affecting the underlying energy landscape.

Weak density features associated with minor conformers are more likely to be perturbed by functionally interesting interactions, but they are also more likely to be disrupted by changes in crystal packing. Here, temperature-dependent rather than hydration-dependent crystallography has a clear advantage, in that dehydration effects are dominated by crystal-packing interactions while temperature also directly affects occupancies and motion amplitudes. High-solvent-content, weakly packed crystals are ideal targets for temperature-dependent structural studies, assuming that crystals with adequate order can be obtained and cooled without ice formation. When these are not available, comparison of cryocooled and dehydrated structures of cryoprotectant-free crystals having comparable unit-cell contractions relative to the native, room-temperature state may highlight aspects of the conformational landscape of the protein that are not specific to the crystalline environment.

Acknowledgements

RET acknowledges a financial conflict of interest, as some of the tools used in this work were manufactured by MiTeGen LLC, in which he holds 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). All X-ray data collection was performed at the Cornell High-Energy Synchrotron Source (CHESS), which is supported by the National Science Foundation and the National Institutes of Health/National Institute of General Medical Sciences under NSF award DMR-0936384, using the Macromolecular Diffraction at CHESS (MacCHESS) facility, which is supported by award GM-103485 from the National Institutes of Health through its National Institute of General Medical Sciences.

References

- Adams, P. D. *et al.* (2010). *Acta Cryst.* **D66**, 213–221.
- Bedem, H. van den, Dhanik, A., Latombe, J.-C. & Deacon, A. M. (2009). *Acta Cryst.* **D65**, 1107–1117.
- Bedem, H. van den & Fraser, J. S. (2015). *Nature Methods*, **12**, 307–318.
- Bell, J. A. (1999). *Protein Sci.* **8**, 2033–2040.
- Boehr, D. D., McElheny, D., Dyson, H. J. & Wright, P. E. (2006). *Science*, **313**, 1638–1642.
- Bowler, M. W., Mueller, U., Weiss, M. S., Sanchez-Weatherby, J., Sorensen, T. L.-M., Thunnissen, M. M. G. M., Ursby, T., Gobbo, A., Russi, S., Bowler, M. G., Brockhauser, S., Svensson, O. & Cipriani, F. (2015). *Cryst. Growth Des.* **15**, 1043–1054.
- Brueschweiler, R. & Wright, P. E. (1994). *J. Am. Chem. Soc.* **116**, 8426–8427.
- Carugo, O. & Argos, P. (1997). *Protein Sci.* **6**, 2261–2263.
- 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.
- Chen, H. & Zhou, H.-X. (2005). *Nucleic Acids Res.* **33**, 3193–3199.
- Collins, M. D., Kim, C. U. & Gruner, S. M. (2011). *Annu. Rev. Biophys.* **40**, 81–98.
- Crowfoot, D. & Riley, D. (1938). *Nature (London)*, **141**, 521–522.
- Dasgupta, S., Iyer, G. H., Bryant, S. H., Lawrence, C. E. & Bell, J. A. (1997). *Proteins*, **28**, 494–514.
- DePristo, M. A., de Bakker, P. I. W. & Blundell, T. L. (2004). *Structure*, **12**, 831–838.
- Dobrianov, I., Kriminski, S., Caylor, C. L., Lemay, S. G., Kimmer, C., Kisselev, A., Finkelstein, K. D. & Thorne, R. E. (2001). *Acta Cryst.* **D57**, 61–68.
- Einstein, J. R. & Low, B. W. (1962). *Acta Cryst.* **15**, 32–34.
- Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* **D66**, 486–501.
- Farley, C., Burks, G., Siebert, T. & Juers, D. H. (2014). *Acta Cryst.* **D70**, 2111–2124.
- Farley, C. & Juers, D. H. (2014). *J. Struct. Biol.* **188**, 102–106.
- Fenimore, P. W., Frauenfelder, H., McMahon, B. H. & Young, R. D. (2004). *Proc. Natl Acad. Sci. USA*, **101**, 14408–14413.
- Fenwick, R. B., van den Bedem, H., Fraser, J. S. & Wright, P. E. (2014). *Proc. Natl Acad. Sci. USA*, **111**, E445–E454.
- Fischer, H., Polikarpov, I. & Craievich, A. F. (2004). *Protein Sci.* **13**, 2825–2828.
- 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.
- Frauenfelder, H., Chen, G., Berendzen, J., Fenimore, P. W., Jansson, H., McMahon, B. H., Stroe, I. R., Swenson, J. & Young, R. D. (2009). *Proc. Natl Acad. Sci. USA*, **106**, 5129–5134.
- Furnham, N., Blundell, T. L., DePristo, M. A. & Terwilliger, T. C. (2006). *Nature Struct. Mol. Biol.* **13**, 184–185.
- Greenspan, L. (1977). *J. Res. Natl Bur. Stand. Sect. A*, **81A**, 89.
- Gunasekaran, K., Ma, B. & Nussinov, R. (2004). *Proteins*, **57**, 433–443.
- Harata, K. & Akiba, T. (2006). *Acta Cryst.* **D62**, 375–382.
- Huxley, H. E. & Kendrew, J. C. (1953). *Acta Cryst.* **6**, 76–80.
- Johnson, L. N. & Phillips, D. C. (1965). *Nature (London)*, **206**, 761–763.
- Juery, D. H. & Matthews, B. W. (2001). *J. Mol. Biol.* **311**, 851–862.
- Kantardjiev, K. A. & Rupp, B. (2003). *Protein Sci.* **12**, 1865–1871.
- Keedy, D. A. *et al.* (2015). *Elife*, **4**, 7574.
- Keedy, D. A., van den Bedem, H., Sivak, D. A., Petsko, G. A., Ringe, D., Wilson, M. A. & Fraser, J. S. (2014). *Structure*, **22**, 1–12.
- Kelly, J. A., Sielecki, A. R., Sykes, B. D., James, M. N. G. & Phillips, D. C. (1979). *Nature (London)*, **282**, 875–878.
- Kiefersauer, R., Than, M. E., Dobbek, H., Gremer, L., Melero, M., Strobl, S., Dias, J. M., Soulimane, T. & Huber, R. (2000). *J. Appl. Cryst.* **33**, 1223–1230.
- Klingl, S., Scherer, M., Stamminger, T. & Muller, Y. A. (2015). *Acta Cryst.* **D71**, 1493–1504.
- Kodandapani, R., Suresh, C. & Vijayan, M. (1990). *J. Biol. Chem.* **265**, 16126–16131.
- Lang, P. T., Holton, J. M., Fraser, J. S. & Alber, T. (2014). *Proc. Natl Acad. Sci. USA*, **111**, 237–242.
- 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.
- Lange, O. F., Lakomek, N.-A., Fares, C., Schröder, G. F., Walter, K. F. A., Becker, S., Meiler, J., Grubmüller, H., Griesinger, C. & de Groot, B. L. (2008). *Science*, **320**, 1471–1475.
- Lorimer, D., Raymond, A., Walchli, J., Mixon, M., Barrow, A., Wallace, E., Grice, R., Burgin, A. & Stewart, L. (2009). *BMC Biotechnol.* **9**, 36.
- Madhusudan, Kodandapani, R. & Vijayan, M. (1993). *Acta Cryst.* **D49**, 234–245.
- Madhusudan & Vijayan, M. (1991). *Curr. Sci.* **60**, 165–170.
- Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
- Miller, S., Janin, J., Lesk, A. M. & Chothia, C. (1987). *J. Mol. Biol.* **196**, 641–656.
- Nagendra, H. G., Sudarsanakumar, C. & Vijayan, M. (1995). *Acta Cryst.* **D51**, 390–392.
- Nagendra, H. G., Sudarsanakumar, C. & Vijayan, M. (1996). *Acta Cryst.* **D52**, 1067–1074.

- Nagendra, H. G., Sukumar, N. & Vijayan, M. (1998). *Proteins*, **32**, 229–240.
- Nakasako, M. (2004). *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **359**, 1191–1206.
- Nakasako, M., Odaka, M., Yohda, M., Dohmae, N., Takio, K., Kamiya, N. & Endo, I. (1999). *Biochemistry*, **38**, 9887–9898.
- Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
- Perutz, M. (1946). *Trans. Faraday Soc.* **42**, B187–B195.
- Perutz, M. (1954). *Proc. R. Soc. London Ser. A*, **225**, 264–286.
- Poornima, C. S. & Dean, P. M. (1995). *J. Comput. Aided Mol. Des.* **9**, 521–531.
- Quillin, M. L. & Matthews, B. W. (2000). *Acta Cryst.* **D56**, 791–794.
- Radha Kishan, R. V., Chandra, N. R., Sudarsanakumar, C., Suguna, K. & Vijayan, M. (1995). *Acta Cryst.* **D51**, 703–710.
- Rockland, L. B. (1960). *Anal. Chem.* **32**, 1375–1376.
- Rupley, J. A. & Careri, G. (1991). *Adv. Protein Chem.* **41**, 37–172.
- Rupley, J. A., Gratton, E. & Careri, G. (1983). *Trends Biochem. Sci.* **8**, 18–22.
- Russi, S., González, A., Kenner, L. R., Keedy, D. A., Fraser, J. S. & van den Bedem, H. (2017). *J. Synchrotron Rad.* **24**, 73–82.
- Russo Krauss, I., Sica, F., Mattia, C. A. & Merlino, A. (2012). *Int. J. Mol. Sci.* **13**, 3782–3800.
- Salunke, D. M., Veerapandian, B., Kodandapani, R. & Vijayan, M. (1985). *Acta Cryst.* **B41**, 431–436.
- Sanchez-Weatherby, J., Bowler, M. W., Huet, J., Gobbo, A., Felisaz, F., Lavault, B., Moya, R., Kadlec, J., Ravelli, R. B. G. & Cipriani, F. (2009). *Acta Cryst.* **D65**, 1237–1246.
- Shaltiel, S., Cox, S. & Taylor, S. S. (1998). *Proc. Natl Acad. Sci. USA*, **95**, 484–491.
- Teeter, M. M., Yamano, A., Stec, B. & Mohanty, U. (2001). *Proc. Natl Acad. Sci. USA*, **98**, 11242–11247.
- Tilton, R. F. Jr, Dewan, J. C. & Petsko, G. A. (1992). *Biochemistry*, **31**, 2469–2481.
- Vendruscolo, M. (2007). *Curr. Opin. Struct. Biol.* **17**, 15–20.
- Voss, N. R. & Gerstein, M. (2010). *Nucleic Acids Res.* **38**, W555–W562.
- Young, J. F. (1967). *J. Appl. Chem.* **17**, 241–245.
- Zhang, X.-J. & Matthews, B. W. (1994). *Protein Sci.* **3**, 1031–1039.
- Zhang, H., Zhang, T., Chen, K., Shen, S., Ruan, J. & Kurgan, L. (2009). *Proteins*, **76**, 617–636.