Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Structure of tetragonal hen egg-white lysozyme at 0.94 Å from crystals grown by the counter-diffusion method

Claude Sauter,[a]† Fermin Otálora,[b] José-Antonio Gavira,[b]‡ Olga Vidal,[b]§ Richard Giegé[a] and Juan Ma. García-Ruiz[b]*

[a]UPR 9002, Département Mécanismes et Molécules de la Synthèse Protéique et Cristallogenèse, Institut de Biologie Moléculaire et Cellulaire, 15 Rue René Descartes, F-67084 Strasbourg CEDEX, France, and [b]Laboratorio de Estudios Cristalográficos, IACT, CSIC-Universidad Granada, Campus Fuentenueva (Fac. Ciencias), 18002 Granada, Spain

† Present address: European Molecular Biology Laboratory, Meyerhofstrasse 1, Postfach 10.22.09, D-69012 Heidelberg, Germany.
‡ Present address: Materials Science Building, Laboratory for Structural Biology, UAH, Huntsville, AL-35899, USA.
§ Present address: Laboratoire Léon Brillouin, CEA-Saclay, F-91191 Gif-sur-Yvette, France.

Correspondence e-mail: jmgruiz@ugr.es

Received 21 February 2001
Accepted 30 May 2001

PDB Reference: tetragonal lysozyme, 1iee.

Very high quality crystals of tetragonal hen egg-white lysozyme were grown in the Advanced Protein Crystallization Facility (APCF) on board the Space Shuttle using a modified free-interface diffusion (FID) reactor designed *ad hoc* to have a longer diffusion path. This design allows the performance of true counter-diffusion experiments. Crystals were obtained under the classical chemical conditions defined 50 y ago with NaCl as a crystallizing agent and acetate pH 4.5 as a buffer. Counter-diffusion crystallization allows a 'physical' instead of chemical optimization of growth conditions: indeed, this method screens for the best supersaturation conditions in a single trial and yields crystals of very high quality. A complete diffraction data set was collected at atomic resolution from one of these crystals using synchrotron radiation at the DESY–EMBL beamlines. The overall $R_{\rm merge}$ on intensities in the resolution range 31–0.94 Å was 5.2% and the data were 98.9% complete. Refinement was carried out with the programs *CNS* and *SHELX*97 to a final crystallographic R factor of 12.26% for 72 390 reflections. A mean standard uncertainty in the atomic positions of 0.024 Å was estimated from inversion of blocked least-squares matrices. 22 side chains show alternate conformations and the loop 59–75 adopts in the same crystal packing two conformations that were observed for either triclinic or tetragonal lysozyme in previous high-resolution studies. In addition to 255 water molecules, the crystallizing agent (one hexacoordinated sodium ion and five chloride anions) participates in the ordered lysozyme hydration shell.

1. Introduction

Once a chemical milieu (pH, additives, type of precipitating agent, purity of the sample *etc.*) has been selected for the crystallization of a biological macromolecule, the quality of the growing crystals will depend on the kinetics of the nucleation and growth processes. Such kinetics are controlled by the value of the supersaturation and more precisely by its variation with time (*i.e.* by the starting salt and macromolecule concentration and the dynamics of mass transport). The method of protein crystallization by counter-diffusion of the precipitating agent gives the opportunity to test different supersaturated states in a single trial, providing a 'self-screening' for the best growth conditions (García-Ruiz, Moreno *et al.*, 1998). To implement this method, the requisite is to remove convective fluid motion in the growth reactor. In order to assure a diffusion mass transport of the reactants, counter-diffusion methods can be implemented either under low- gravity conditions or alternatively on the ground using

either gelled solutions or very thin capillary volumes (García-Ruiz et al., 1999).

We report here the quality of hen egg-white lysozyme crystals grown by the counter-diffusion method in the ESA Advanced Protein Crystallization Facility (APCF) during the STS-95 mission of the NASA Space Shuttle, which eventually led to diffraction data at the highest resolution ever collected to date from tetragonal crystals from this protein. This model protein for crystal growth was selected because many of its physical and chemical properties in aqueous solutions (solubility, diffusivity, refractive index, density, viscosity) are well known. This was an absolute requisite for our space experiment, whose main goal was to record the supersaturation in the crystallization reactor by time-lapse Mach–Zehnder interferometry. In addition, we were looking to compare the quality of the space-grown crystals with the best crystals ever grown on ground by any method. For this aim lysozyme is clearly a good target model, as it has been extensively crystallized on-ground and a large number of X-ray data sets are available.

The present paper reveals some new structural features at atomic resolution and discusses the usefulness of the counter-diffusion method for growing high-quality crystals in microgravity experiments and its extrapolation to on-ground implementation of this technique.

2. Methods

2.1. Crystallization

The crystallization experiment was performed using a free-interface diffusion reactor designed *ad hoc* (Fig. 1). It was designed to allow experiments in which the time needed to transport the precipitating agent to the end of the protein reservoir is sensibly higher than the induction time for nucleation. The extended-length reactor (FID-XL) consists of two 67 mm long chambers containing the protein solution and the precipitating agent solution, respectively. They are connected by a plug used to activate and deactivate the experiment. This reactor allows a counter-diffusion regime for up to 15 d when used with small-molecule salts as the precipitating agent.

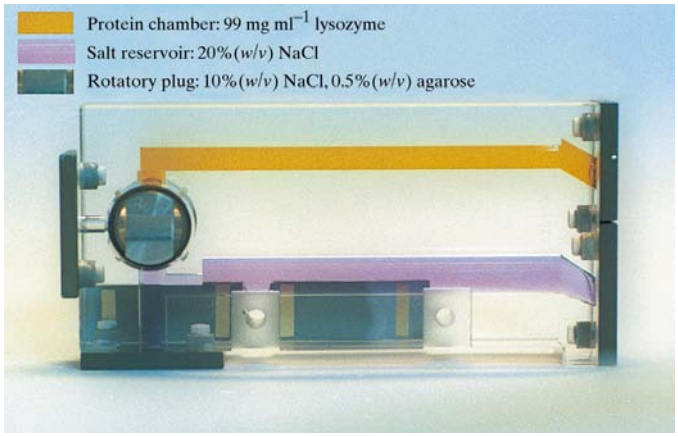

Figure 1
Free-interface diffusion extended-length (FID-XL) APCF reactor, designed to be used in long-chamber counter-diffusion experiments during STS-95 mission and currently available for experimentation in the APCF. The two coloured volumes are the protein (yellow) and salt (violet) chambers. Initial concentration of chemicals during the growth experiments are indicated.

The starting chemical conditions are shown in Fig. 1. A Mach–Zehnder interferometer monitored the changes in refractive index arising from concentration gradients inside the protein chamber and a time-lapse video microscope recorded the crystal nucleation and growth. The experiment was performed at 293 ± 0.1 K inside the thermostatted APCF box. Hen egg-white lysozyme from Seikagaku (ref: 100940; lot No. E96302) was used as purchased without further purification. More details on the crystallization experiment are provided in García-Ruiz *et al.* (2001).

2.2. Data collection

Diffraction data were collected using synchrotron radiation at EMBL beamlines X11 and BW7B at the DORIS storage ring, DESY, Hamburg on a 345 mm MAR Research imaging-plate scanner. A lysozyme crystal about 800 µm in length was soaked for a few seconds in a crystallization solution containing 40% glycerol before flash-cooling it in the nitrogen stream. One should note that the cooling procedure was not optimized to minimize the crystal mosaicity, but was used to reduce radiation damage during data collection. Thus, a mosaicity value of 0.47° is not surprising and cannot be considered as a quality indicator. For technical reasons (beamline availability) the data collection was started on beamline X11. The crystal was then recovered in liquid nitrogen and transferred to beamline BW7B. Three data sets (300 images in total; Table 1) were collected at 110 K on this single frozen crystal with different exposure times and to different resolutions in order to cover the whole range of intensities.[1] Data were integrated, scaled and merged using the *DENZO* and *SCALEPACK* programs (Otwinowski & Minor, 1997). An anomalous dispersion signal was detected using the program *LOCHVAT* (Dumas, 1994*a*,*b*); it was used to localize chloride ions but was not taken into account during the refinement. The overall $R_{\rm merge}$ on intensities for all data in the resolution range 31–0.9 Å was 5.2% and the data were 98.9% complete (Table 1). An overall B factor of 8.7 Å^2 was estimated from the Wilson plot using the program *TRUNCATE* from the *CCP*4 package (Collaborative Computational Project, Number 4, 1994).

2.3. Structure refinement

The starting model for refinement was the tetragonal structure of HEWL solved at 100 K to 1.7 Å resolution and having unit-cell parameters close to those described here (PDB code 1lsa; Kurinov & Harrison, 1995). The solvent molecules were removed from the starting model and a *B*

[1] Supplementary data have been deposited in the IUCr electronic archive (Reference: li0396). Services for accessing these data are described at the back of the journal.

Table 1
Data-collection and processing statistics.

Values in parentheses refer to the outermost resolution shell.

Synchrotron beamline	X11	BW7B	BW7B
Wavelength (Å)	0.9057	0.8337	0.8337
Crystal-to-detector distance (mm)	110	130	330
Maximum resolution (Å)	0.95	0.94	1.75
Oscillation step (°)	0.5	0.5	2
No. of images	90	90	120
Space group	$P4_32_12$		
Unit-cell parameters (Å)	a = 77.06, c = 37.22		
Crystal mosaicity (°)	0.47		
No. of observations	402070		
No. of unique reflections	72390		
Resolution range (Å)	0.94–31 (0.94–0.96)		
Completeness (%)	98.9 (88.6)		
Completeness, $I > 2\sigma(I)$ (%)	88.5 (53.3)		
$R_{\rm merge}$† (%)	5.2 (32.9)		
$\langle I/\sigma(I)\rangle$	15.5 (2.6)		

† $R_{\rm merge} = \sum_{hkl}\sum_i |\langle I\rangle - I_i|/\sum_{hkl}\sum_i (I_i)$.

factor of 10 Å^2 was assigned to all atoms. The refinement process utilized the programs *CNS* and *SHELX*97 (Brunger *et al.*, 1998; Sheldrick & Schneider, 1997) with stereochemical restraints based on those of Engh & Huber (1991). All reflections were included with no σ cutoff; 5% of the data were randomly selected and omitted during refinement for cross-validation analysis by means of the free R factor (Brünger, 1992). The main steps of the refinement progress are summarized in Table 2.

50 steps of rigid-body refinement were first performed using *CNS* in the resolution interval 31.0–2.0 Å. Several rounds of conjugate-gradient least-squares (CGLS) minimization including a bulk-solvent correction were carried out and the resolution was increased stepwise from 2.0 to 0.94 Å. At this stage, individual isotropic B factors were refined and water molecules developing sensible hydrogen bonds were added in the difference Fourier map (peaks > 4σ), leading to a model with R and $R_{\rm free}$ of 26.2 and 27.2%, respectively.

The refinement was continued with *SHELX*97 based on intensities (*i.e.* squared structure-factor amplitudes); each step consisted of 10–30 cycles of CGLS minimization. Default values of the program for distance, planarity and chiral restraints and a diffuse solvent model were used throughout. New ordered solvent molecules were found in the difference Fourier map (peaks > 4σ, $B_{\rm eq}$ < 60 Å^2). $3F_o - 2F_c$ and $F_o - F_c$ electron-density maps were calculated after each step and the model (protein and solvent) was checked and rebuilt using the program *O* (Kleywegt & Jones, 1997). The R and $R_{\rm free}$ factors calculated after the first run, treating B factors isotropically, were 23.7 and 26.5%, respectively. Refinement of anisotropic displacement parameters (ADPs) of protein atoms dramatically decreased R and $R_{\rm free}$ to 16.5 and 20.6%, respectively. ADPs of solvent molecules were restrained to be approximately isotropic. At this stage, clear electron density was observed near several residue side chains, which were subsequently modelled in two conformations. In relation to the side-chain disorder, the main chain was also found to adopt two conformations in sections 17–19, 77–78, 86–88 and 108–110. The largest backbone variations are observed for residues 58–62 and 68–75 in the 59–75 loop. The occupancy factors of solvent atoms were not refined, except those of water molecules associated with alternate protein conformations or sharing close positions (d < 2 Å). A sodium ion was added at the binding site described in previous studies using the Fourier difference map as a guide. Five chloride anions were detected as described by Dauter *et al.* (1999) by computing an anomalous ($\Delta F_{\rm ano}$,$\varphi_{\rm calc} - 90°$) Fourier map in the 10–1 Å resolution range with the data collected at λ = 0.907 Å using the programs *SFALL*, *CAD* and *FFT* from the *CCP*4 package (Collaborative Computational Project, Number 4, 1994).

In the final steps H atoms were added to the model according to geometrical criteria. This yielded a decrease of 0.8% in the R factor and of 1.6% in the $R_{\rm free}$ factor without increasing the number of parameters. The occupancy parameters were refined for all solvent atoms during the final steps, yielding a 0.4% fall in the R and $R_{\rm free}$ factors. When convergence of the CGLS minimization was achieved, all data were included in the refinement, leading to a final R factor of 11.64% for 64 761 reflections with $F_o > 4\sigma(F_o)$ and 12.26% for all data (72 390 reflections) in the resolution range 31–0.94 Å.

2.4. The final model

The final model for 1001 lysozyme atoms consists of a total of 1229 protein atom sites, one sodium ion, five chloride ions and 255 water molecules. 22 side chains were modelled in two conformations as well as six main-chain sections. The quality of the density map was generally excellent: the complete protein backbone could be seen and only a few side chains

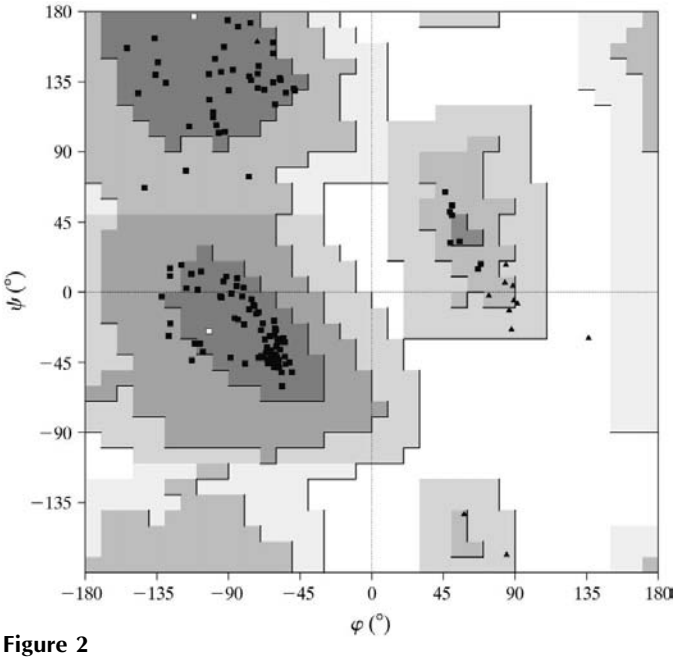

Figure 2
Ramachandran diagram of the final model. A total of 136 φ/ψ pairs are plotted including residues having double main-chain conformation (17–19, 59–62, 69–75, 77–78, 86–88 and 109–110). Glycine residues are represented as triangles and Arg73 residues from conformations *A* and *B* of loop 59–75 as open squares.

Table 2
Refinement progress.

The refinement was performed with the software *CNS* (C) and *SHELX*97 (S). All reflections in the indicated resolution range were included in the calculation with a 5% test set for cross-validation. In the final step, all data were used (72 390 reflections). The number of protein (N_{prot}), solvent (N_{H_2O}) and ion (N_{ion}) atoms in the model is indicated.

Refinement step	Prog.	Res. (Å)	N_{prot}	N_{H_2O}	N_{ion}	R/R_{free} (%)
Initial model (1lsa)	C	31–2.0	1001			54.56/54.57
Rigid-body adjustment	C	31–2.0	1001			36.56/35.32
Coordinate refinement	C	31–2.0	1001			30.49/34.68
Coordinate refinement	C	31–1.5	1001			31.66/33.80
Coordinate and *B* refinement	C	31–0.94	1001			30.14/31.01
Solvent molecules	C	31–0.94	1001	132		26.19/27.25
Isotropic B_{eq} refinement	S	31–0.94	1001	132		23.66/26.53
Solvent molecules	S	31–0.94	1001	177		22.36/25.53
ADP refinement	S	31–0.94	1001	176		16.48/20.61
Ions, side-chain conformers	S	31–0.94	1014	262		15.04/19.41
New conformers	S	31–0.94	1045	269	6	14.79/19.11
Two conformations for 59–75 loop	S	31–0.94	1190	242	6	13.49/17.12
H atoms	S	31–0.94	1201	249	6	12.74/15.53
H_2O occupancies	S	31–0.94	1201	255	6	12.33/15.09
Final model	S	31–0.94	1201	255	6	12.26

were poorly defined. A final run of blocked-matrix refinement was performed to estimate the coordinate errors. To this end, the protein was split into 11 overlapping blocks of 15 residues, with an overlap of two residues between successive blocks. Positional parameters and ADPs were refined in each of these blocks. The overall r.m.s. error for all protein atoms is 0.024 Å. The quality of the structure was assessed using *PROCHECK* (Laskowski *et al.*, 1993), *SHELXPRO* (Sheldrick & Schneider, 1997) and *PARVATI* (Merritt, 1999). The lysozyme backbone shows good stereochemistry: all the φ/ψ angles are in the core region of the Ramachandran plot (Fig. 2). The mean ω value is 179.8°, with a standard deviation of 6.5°, and the χ_1 angles of side chains with a single conformation are clustered in three populations: −65.4, 66.4 and 182.5°, with standard deviations of 12.4, 3.7 and 8.5°, respectively. The final model was compared with previous structures after least-squares superposition of the backbone atoms (N, C^{α}, C) using the program *LSQMAN* (Kleywegt & Jones, 1997). Fig. 2 was prepared using *PROCHECK*, Figs. 3, 4 and 5 using *SETOR* (Evans, 1993) and Fig. 6 using *rastep* (Merritt & Bacon, 1997).

3. Results and discussion

3.1. A crystallization experiment that seeks out the best crystals

The quality of the crystals obtained in a crystallization experiment is a result of the optimization of the supersaturation and supersaturation rate. In the counter-diffusion method, the protein and the precipitating agent solution are placed one in front of the other either in direct contact or are separated by an intermediate chamber (in our experiment, a gel plug) working as a physical buffer. The method works properly when the starting concentration of precipitating agent is selected to be very high and the protein chamber is very long. Because the salt diffuses faster than the protein, the salt molecules will invade the protein chamber. When these solutions come into contact, the supersaturation is very high and a massive precipitation of small low-quality crystals is obtained at the beginning of the protein chamber. Its formation depletes the concentration of protein in the neighbouring zones. As the salt continues to diffuse forward in the protein chamber, new precipitation takes place, this time at lower supersaturation, which yields a lower nucleation density and therefore crystals of larger size. Iteration of this process provokes successive precipitation phenomena at lower supersaturation as the precipitation front moves ahead in the protein chamber. This yields precipitation zones of fewer crystals of larger size and higher quality. This is the reason why, unlike the classical drop and batch methods, the counter-diffusion technique explores a large number of crystallization conditions in one single experiment. The crystallization pathway always evolves towards equilibrium and therefore the experiment self-searches for the optimal crystallization scenario (García-Ruiz, 1991). As predicted by computer simulation (Otálora & García-Ruiz, 1996), this automatic search of the best nucleation and growth conditions occurs because of the development of a travelling supersaturation wave across the protein chamber. Interferometric data obtained during the crystallization experiments in space are discussed elsewhere (García-Ruiz *et al.*, 2001) and confirmed that crystallization does proceed in the correct regime; the existence of such a wave of precipitation was observed for the first time.

Counter-diffusion methods that automatically optimize nucleation and growth conditions are not exclusive to space experiments. They can also be implemented on the ground using long protein chambers in which mass transport proceeds by diffusion that are long enough to avoid homogenization during the nucleation induction time. Gels and capillaries have been demonstrated to have similar effects on the physics of the counter-diffusion setup (Robert *et al.*, 1999). Even though X-ray capillaries of less than 100 µm in diameter are needed to remove effectively convective flow, larger capillaries filled with a gel can be conveniently used to produce the same effect (García-Ruiz, Gavira *et al.*, 1998). In such a case, the gel will unavoidably incorporate into the growing protein crystals, but X-ray diffraction studies of these crystals yield unexpected high-quality diffraction patterns (Gavira, 2000) which makes the use of gels very promising in the future.

3.2. The overall HEWL structure

HEWL forms globular and compact particles rigidified by four disulfide bonds and the present atomic resolution structure shows the same fold as the starting model (1lsa; Kurinov & Harrison, 1995) and the previous 1.33 and 1.4 Å resolution structures (193l and 194l; Vaney *et al.*, 1996). Nevertheless, the combination of high-quality crystals and cryocrystallography

leads to a much more accurate description of the HEWL tetragonal crystal form. For well defined parts of the structure (Fig. 3), individual protein atoms were identified in the $(3F_o - 2F_c)$ Fourier map by resolved peaks at a density level of 5σ (3.35 e Å^{-3}). The overall r.m.s. coordinate error estimated from blocked-matrix least-squares refinement is 0.024 Å for all protein atoms and 0.013 Å for main-chain atoms with unity occupancy. The quality of the final model is summarized in Table 3. A plot of the φ/ψ angles in the polypeptide backbone is shown in Fig. 2: no non-glycine residue was found outside the most favoured region of the Ramachandran diagram.

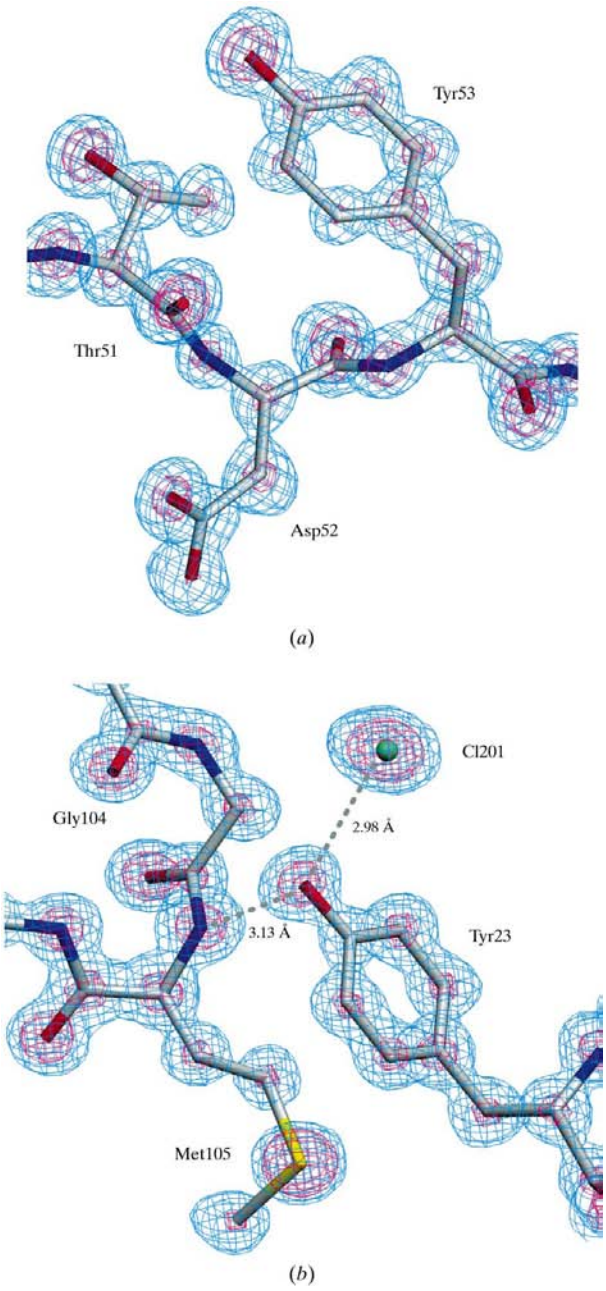

Figure 3
Typical examples of density maps in well defined lysozyme parts. (*a*) A region in β-sheet including Asp52, a residue that participates in the catalytic activity of the enzyme. (*b*) Chloride 201 bound to Tyr23 and Asn113 of a symmetry-related molecule (not shown), first identified by Vaney *et al.* (1996). The final $3F_o - 2F_c$ electron-density map is contoured at a level of 1.2 e Å^{-3} (1.8σ) in blue and of 3 e Å^{-3} (4.5σ) in purple.

Table 3
Final model.

Resolution range (Å)	31–0.94
R factor for 72 390 reflections (%)	12.26 (11.63)†
Free *R* factor for 3655 reflections (%)	15.09 (14.35)†
Observation-to-parameter ratio	5.3
Number of protein atoms (non-H)	1229
Number of solvent, sodium and chloride sites	255, 1, 5
Mean B_{eq} (Å^2) and *A*‡	
Overall	14.4, 0.46
Protein	13.0, 0.47
Solvent	21.1, 0.40
Ions	14.1, 0.36
R.m.s. deviation from ideal geometry§	
Bond lengths (Å)	0.016
Angle distances (Å)	0.037
Deviation from planes (Å)	0.031
Chiral volumes (Å^3)	0.11
Ramachandran plot statistics¶	
Residues in most favoured regions (%)	85.4
Residues in additional allowed regions (%)	14.6

† *R* factors in brackets are for reflections with $F_o > 4\sigma(F_o)$. ‡ *A* values were determined using *PARVATI* (Merritt, 1999). The anisotropy $A = E_{min}/E_{max}$ is defined as the ratio of the smallest and the largest eigenvalues of the ADP matrix. § Values determined using *SHELXPRO* (Sheldrick & Schneider, 1997). ¶ Values determined using *PROCHECK* (Laskowski *et al.*, 1993).

The high observation-to-parameter ratio at atomic resolution (<1.2 Å) allows anisotropic refinement (Sheldrick, 1990) and such a treatment significantly improved the model (Murshudov *et al.*, 1999) as indicated by a 5% fall in both the *R* and R_{free} factors. The average value of anisotropy (A = 0.46, σ = 0.16) in the HEWL structure (Table 3) is similar to that observed so far for atomic resolution structures (Merritt, 1999). After taking ADPs into account, it became clear from the electron-density inspection that about 20% of the residues had more than one conformation, either for the side-chain or for the main-chain atoms (see below). Density peaks corresponding to H atoms were also observed in the difference map near the main-chain atoms (N, C$^\alpha$) of many residues and in the vicinity of a few side-chain atoms (Fig. 4). Thus, considering H atoms as a riding model led to an additional 1.6% decrease in the R_{free} factor.

As a consequence of higher resolution, an increased number of solvent sites (255 H_2O) was observed in comparison to the 1.33 Å study (142 H_2O). In the early steps of the refinement the occupancy factors of solvent atoms were fixed to 1 for well defined molecules or to 0.5 for those showing high U_{eq} (>0.5 Å^2). Partial occupancies were refined for water molecules sharing close locations (36 H_2O), thus forming alternate hydration networks, or linked to side-chain conformers (38 H_2O). When the refinement reached the convergence, water molecules essentially in the second hydration shell, *i.e.* without any direct contact to the protein, were characterized by low density peaks (1–1.5σ), suggesting partial occupation of many solvent sites. Hence, in a final step the occupancy factors of all solvent atoms were refined and, as in the case of triclinic lysozyme (Walsh *et al.*, 1998), this led to a 0.4% decrease in both the *R* and the R_{free} factors. The solvent

content of the final model is equivalent to 170 fully occupied water molecules.

3.3. Amino acids in multiple conformations

A second conformer was identified for a total of 22 side chains both in the $3F_o - 2F_c$ and difference Fourier maps. Side-chain static disorder concerns different types of residues: Arg (14, 21, 45, 61, 112, 125, 128), Asn (19, 59, 77), Asp (119, 87), Gln (121), Ile (55, 58, 78), Lys (1, 97), Ser (81, 85, 86) and Val (109). Only four of these residues (1, 59, 86, 109) were identified at 1.33 Å resolution; 13 are also observed in the triclinic structure that contains 25 residues with alternate conformations (Walsh *et al.*, 1998). Flexible residues such as Arg61, Arg73, Arg128 and Gln121 show poorly defined side chains, as does Trp63, a residue known for its flexibility and its sensitivity to radiation damage (see Walsh *et al.*, 1998). As in previous HEWL studies, the main chain of the last two residues (Arg128 and Leu129) was modelled in one conformation even though the density suggested alternate possibilities. Indeed, these residues are close to disordered regions (18–19, 69–70, 128–129) of symmetry-related molecules and this may induce some structural heterogeneity in the C-terminus.

3.4. Two conformations for the loop 59–75

In spite of the overall rigidity of lysozyme, large variations have been observed in the flexible 59–75 loop, which adopts two main conformations depending on the space group and the presence of the substrates (for reviews, see Vaney *et al.*, 1996; Steinrauf, 1998). After the first step of anisotropic refinement, a second conformation based on that observed for this loop in triclinic crystals was introduced to satisfy the additional density still present in the $3F_o - 2F_c$ map. This contributed to the elimination of remaining strong peaks in the difference map and yielded a decrease in both R factors (Table 2). Alternate positions concern two backbone sections, 58–62 and 68–75, which undergo linked movements. The segment connecting these sections (63–67) is stabilized by a disulfide bridge (Cys64–Cys80) and presents only one conformation. Overall occupancy factors were assigned to the atoms of each loop conformer (including bound ions and solvent molecules) and their sum was restrained to unity during subsequent refinement steps as implemented in *SHELX*.

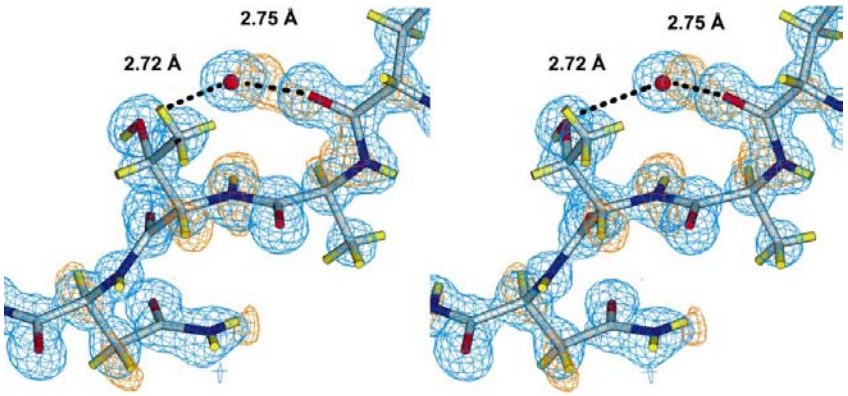

Figure 4
Hydrogen location in the Fourier difference map. This stereoview shows the residues Asn42, Thr43 and Ala44. Density peaks in the difference map (computed before the addition of H atoms to the model) clearly indicate their location. This is true for the well ordered water molecule. $3F_o - 2F_c$ (in blue) and $F_o - F_c$ (in orange) maps are contoured at levels of 1.2 e Å^{-3} (1.8σ) and 0.21 e Å^{-3} (2.5σ), respectively.

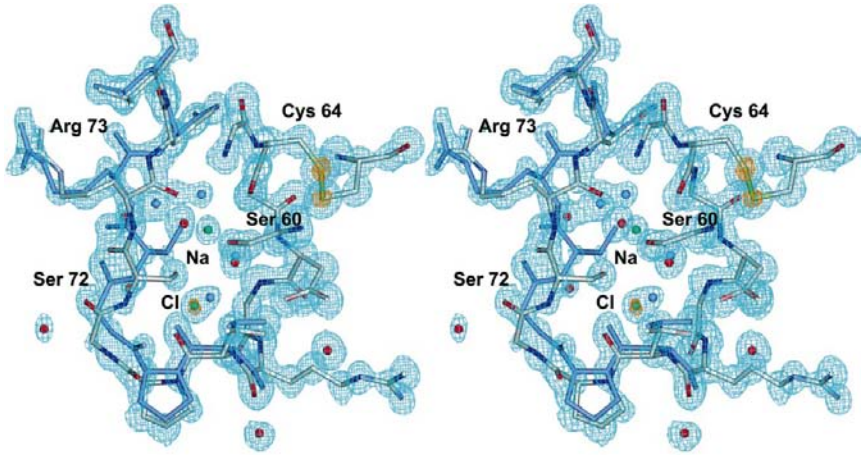

Figure 5
The two conformations for the loop 59–75. Conformation *A* (indicated in white) binds a sodium ion (Na200) and a chloride ion (Cl203) drawn in green; the four residues that bind the Na ion are highlighted (see also Fig. 6). The loop in conformation *B* (indicated in purple) does not bind any ions but it is stabilized by alternate water molecules (in purple). R.m.s.d.s between conformation *A*, conformation *B* and other known structures are given in Table 4. The contour levels for the $3F_o - 2F_c$ map (in blue) and for the anomalous difference map (in orange) which localizes Cl and S atoms, are 1.5 and 4σ, respectively.

As shown in Fig. 5, two conformations are simultaneously present in the same crystal packing and have equivalent occupancy factors (0.48 and 0.52 for conformations *A* and *B*, respectively). They superimpose well on the two types observed in the latest high-resolution structures (Table 4). As observed in previous studies, the main difference is a flip of the peptide bond Arg73–Asn74 (Fig. 2). Conformation *A*, present in the starting model 1lsa, is related to the binding of a sodium and its chloride counterion in the tetragonal space group, as shown by Vaney *et al.* (1996) and Dauter *et al.* (1999). The CO group of Arg73 participates in the coordination of the sodium ion. In conformation *B*, the CO group is rotated toward the solvent, no ion is observed and the loop is exclusively

Table 4
Comparison of loop 59–75 conformations.

		R.m.s.d.† (Å)		
	Arg73 φ/ψ (°)	Conformation *A*	Conformation *B*	Heteroatoms bound to the loop
Conformation *A*	−110.1/173.9	—	0.61	Na^+, Cl^-, 2 H_2O
1lsa‡	−120.2/158.5	0.20	0.53	3 H_2O
193l§	−100.3/169.0	0.19	0.54	Na^+, 3 H_2O
1lz8¶	−117.7/173.1	0.16	0.56	Na^+, Cl^-, 2 H_2O
Conformation *B*	−101.5/−22.8	0.61	—	3 H_2O
3lzt††	−110.1/−24.3	0.67	0.29	3 H_2O

† R.m.s.d. for N, CA and C positions in residues 57–76 after least-squares superposition of their main-chain atoms. ‡ Tetragonal HEWL (120 K) solved at 1.7 Å resolution (Kurinov & Harrison, 1995); starting model. § Tetragonal HEWL (279 K) solved at 1.33 Å resolution (Vaney *et al.*, 1996). ¶ Tetragonal HEWL (100 K) solved at 1.54 Å resolution (Dauter *et al.*, 1999). †† Triclinic HEWL (120 K) solved at 0.92 Å resolution (Walsh *et al.*, 1998).

stabilized by an alternate network of water molecules as in triclinic crystals (Walsh *et al.*, 1998).

3.5. Cation and anion binding sites

As mentioned above, one sodium cation is associated with conformation *A* of loop 59–75 (Fig. 5). Noticeably, when the loop backbones from conformation *A* in lysozyme structures 193l and 1lz8 are superimposed (Table 4), the sodium position is perfectly conserved. The hexagonal coordination geometry of the cation predicted from valence calculations (Nayal & Di Cera, 1996) involves four protein atoms, Ser72 O^{γ}, Ser60 O, Cys64 O and Arg73 O, and two solvent molecules. As illustrated in Fig. 6, the sodium ligands form a distorted octahedron (Fig. 6) originally described by Vaney *et al.* (1996). A chloride counter-ion (Cl203, see below) is bridged to the cation by a water molecule. This halide site is akin to that seen in 1lz8.

HEWL is a basic protein and binds various anions (reviewed by Vaney *et al.*, 2001): nitrates, acetates, iodides or chlorides were recently described at high resolution in the triclinic, monoclinic and orthorhombic crystal forms (Walsh *et al.*, 1998; Steinrauf, 1998; Oki *et al.*, 1999). In the tetragonal environment, chloride and bromide anions were observed by Vaney *et al.* (1996), Lim *et al.* (1998) and Dauter *et al.* (1999). The latter study identified eight chlorides on the basis of their anomalous signal and the authors utilized the single-wavelength anomalous dispersion from S and Cl atoms at 1.54 Å to solve the structure of HEWL. Although only a weak anomalous signal was expected in the present case ($\delta f'$ values for S and Cl at λ = 0.907 Å are 0.20 and 0.25 electrons, respectively, *i.e.* 2.8 times lower than at 1.54 Å), anomalous differences were used to compute a (ΔF_{ano}, φ_{calc} − 90°) Fourier map as proposed by Dauter *et al.* (1999). In addition to density peaks for the ten lysozyme S atoms (eight Cys and two Met), five peaks were observed in the solvent shell with a density higher than four times the r.m.s.d. of the anomalous map (see Fig. 5). They correspond to five of the eight chloride sites characterized by Dauter *et al.* (1999). The environment of these halides was essentially as described in the latter study and the same numbering was used: chloride 201 (Fig. 3*b*) is bound to residues Tyr23 (OH group, distance of 2.98 Å) and Asn113 (N^{δ}, 3.41 Å) of two symmetry-related molecules: chloride 202 is linked to Ser24 (O^{γ}, 3.06 Å) and Gly26 (N, 3.13 Å), chloride 203 is associated with conformation *A* of loop 59–75 as illustrated in Figs. 5 and 6, chloride 204 interacts with Ile88 (N, 3.30 Å) and His15 (N^{ε}, 3.42 Å) and chloride 205 interacts with Lys33 (N^{ζ}, 3.52 Å). The five anions also have water molecules as ligands and the electron density indicates that the chloride sites are not fully occupied and that, as for chloride 203 in loop 59–75 (Fig. 5), alternative solvent networks are present. Final B_{eq} values and occupancy factors of chlorides 201–205 are 10.0, 14.5, 10.7, 20.2, 19.0 $Å^2$ and 0.70, 0.75, 0.48, 0.44, 0.75, respectively. Chlorides 206, 207 and 208 were associated with residues 65, 73 and 68, respectively (Dauter *et al.*, 1999). These ions are not seen in the present study, probably owing to their location in a region close to the flexible 59–75 loop where the electron density for the solvent shell is not well defined. Thus, weakly bound chlorides, if present, are unlikely to be detected from their anomalous signal.

4. Concluding remarks

The production of crystals that yield diffraction at atomic resolution provides many advantages for crystallographic studies. In such a case, the model refinement requires only very few manual adjustment, mainly during the latest steps, and is carried on in a series of well defined steps (see Sevcik *et al.*, 1996; Sheldrick & Schneider, 1997; Walsh *et al.*, 1998). As

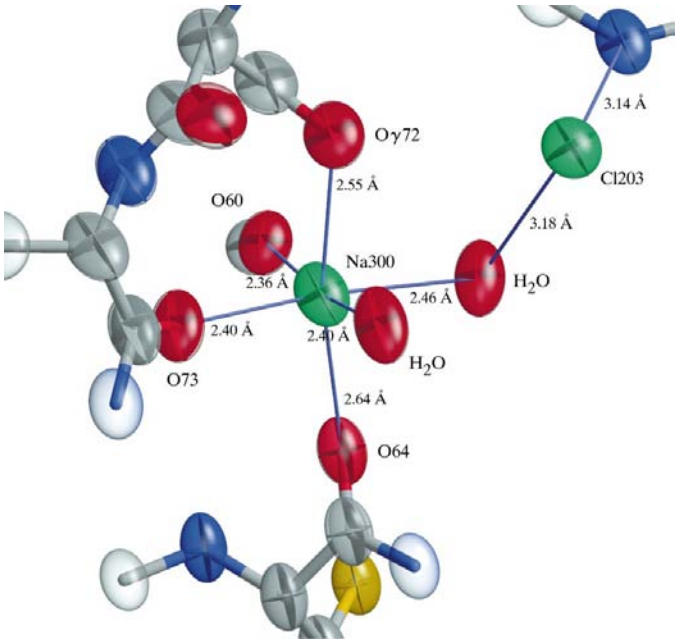

Figure 6
Coordination ligands of the sodium and chloride ions in loop 59–75. Ellipsoids were drawn with a 30% probability using the software *rastep* (Merritt & Bacon, 1997).

seen in the present study of HEWL, an atomic resolution structure is much more accurate. It brings a wealth of details such as a better description of flexibility in the macromolecule through the refinement of ADPs and the modelling of conformational disorders, an enhanced description of the solvent network and ion-binding sites. Recent technological advances in data collection (third-generation synchrotron sources, two-dimensional area detectors and cryocooling methods) have contributed to the explosion of atomic resolution studies (Merritt, 1999).

Current advances in data-collection methodology and instrumentation are shifting the limiting factor for high-quality diffraction data towards the intrinsic quality of the crystal sample. In that view, it is vital to develop crystallization strategies to increase the potential of the crystalline material already available. The automatic optimization of nucleation and growth conditions provided by counter-diffusion methods, either on ground or in space, can be a significant advance in this direction.

We gratefully acknowledge discussions with E. Ennifar, P. Dumas and Lopez-Jaramillo. We thank the European Molecular Biology Laboratory at the DESY storage ring, Hamburg for the beam time allocated to this project. The flight opportunity was provided by the European Space Agency. We thank the Astrium's APCF team (Friedrichhafen, Germany) for helping during the design of the FID extended-length reactors. Financial support was provided by CICYT (Spain), CNRS (France) and the EU project European Bio-Crystallogenesis Initiative (BIO4-CT98-0086).

References

Brünger, A. T. (1992). *Nature (London)*, **355**, 472–475.
Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* D**54**, 905–921.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Dauter, Z., Dauter, M., de La Fortelle, E., Bricogne, G. & Sheldrick, G. M. (1999). *J. Mol. Biol.* **289**, 83–92.
Dumas, P. (1994*a*). *Acta Cryst.* A**50**, 526–537.
Dumas, P. (1994*b*). *Acta Cryst.* A**50**, 537–546.
Engh, R. A. & Huber, R. (1991). *Acta Cryst.* A**47**, 392–400.
Evans, E. (1993). *J. Mol. Graph.* **11**, 134–138.
García-Ruiz, J. M. (1991). *Key Eng. Mater.* **58**, 87–106.
García-Ruiz, J. M., Gavira, J. A., Otálora, F., Guasch, A. & Coll, M. (1998). *Mater. Res. Bull.* **33**, 1593–1598.
García-Ruiz, J. M., Moreno, A., Otálora, F., Viedma, C., Rondón, D. & Zautscher, F. (1998). *J. Chem. Ed.* **75**, 442–446.
García-Ruiz, J. M., Novella, M. L. & Otálora, F. (1999). *J. Cryst. Growth*, **196**, 703–710.
García-Ruiz, J. M., Otálora, F., Novella, M. L., Gavira, J. A., Sauter, C. & Vidal, O. (2001). In the press.
Gavira, J. A. (2000). PhD thesis, University of Granada, Spain.
Kleywegt, G. J. & Jones, T. A. (1997). *Methods Enzymol.* **277**, 525–545.
Kurinov, I. V. & Harrison, R. W. (1995). *Acta Cryst.* D**51**, 98–109.
Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–290.
Lim, K., Nadarajah, A., Forsythe, E. L. & Pusey, M. L. (1998). *Acta Cryst.* D**54**, 899–904.
Merritt, E. A. (1999). *Acta Cryst.* D**55**, 1109–1117.
Merritt, E. A. & Bacon, D. J. (1997). *Methods Enzymol.* **277**, 505–524.
Murshudov, G. N., Vagin, A. A., Lebedev, A., Wilson, K. S. & Dodson, E. J. (1999). *Acta Cryst.* D**55**, 247–255.
Nayal, M. & Di Cera, E. (1996). *J. Mol. Biol.* **256**, 228–234.
Oki, H., Matsuura, Y., Komatsu, H. & Chernov, A. A. (1999). *Acta Cryst.* D**55**, 114–121.
Otálora, F. & García-Ruiz, J. M. (1996). *J. Cryst. Growth*, **169**, 361–367.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Robert, M. C., Vidal, O., García-Ruiz, J. M. & Otálora, F. (1999). *Crystallization of Nucleic Acids and Proteins: a Practical Approach*, edited by A. Ducruix & R. Giegé. p. 331. Oxford: IRL Press.
Sevcik, J., Dauter, Z., Lamzin, V. S. & Wilson, K. S. (1996). *Acta Cryst.* D**52**, 327–344.
Sheldrick, G. M. (1990). *Acta Cryst.* A**46**, 467–473.
Sheldrick, G. M. & Schneider, T. R. (1997). *Methods Enzymol.* **277**, 319–343.
Steinrauf, L. K. (1998). *Acta Cryst.* D**54**, 767–780.
Vaney, M. C., Broutin, I., Retailleau, P., Douangamath, A., Lafont, S., Hamiaux, C., Prangé, T., Ducruix, A. & Riès-Kautt, M. (2001). *Acta Cryst.* D**57**, 929–940.
Vaney, M. C., Maignan, S., Riès-Kautt, M. & Ducruix, A. (1996). *Acta Cryst.* D**52**, 505–517.
Walsh, M. A., Schneider, T. R., Sieker, L. C., Dauter, Z., Lamzin, V. S. & Wilson, K. S. (1998). *Acta Cryst.* D**54**, 522–546.