Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Structural effects of monovalent anions on polymorphic lysozyme crystals

M. C. Vaney,[a]† I. Broutin,[a] P. Retailleau,[a] A. Douangamath,[a] S. Lafont,[a] C. Hamiaux,[b] T. Prangé,[b] A. Ducruix[a] and M. Riès-Kautt[a]*

[a]Laboratoire de Cristallographie et RMN Biologiques, CNRS–UMR 8015, Faculté de Pharmacie, 4 Avenue de l'Observatoire, 75006 Paris, France, and [b]LURE, Centre Universitaire Paris-Sud, Bâtiment 209D, BP 34, 91898 Orsay CEDEX, France

† Present address: Laboratoire de Génétique des Virus, CNRS UPR 9053, 1 Avenue de la Terrasse, 91190 Gif/Yvette, France.

Correspondence e-mail: ries@pharmacie.univ-paris5.fr

Received 15 December 2000
Accepted 13 March 2001

PDB References: HEWL/NapTS, 1b0d; HEWL/NaI, 1b2k; HEWL/$NaNO_3$, 1hf4; HEWL/KSCN, 1lcn.

Understanding direct salt effects on protein crystal polymorphism is addressed by comparing different crystal forms (triclinic, monoclinic, tetragonal and orthorhombic) for hen, turkey, bob white quail and human lysozymes. Four new structures of hen egg-white lysozyme are reported: crystals grown in the presence of NapTS diffracted to 1.85 Å, of NaI to 1.6 Å, of $NaNO_3$ to 1.45 Å and of KSCN to 1.63 Å. These new structures are compared with previously published structures in order to draw a mapping of the surface of different lysozymes interacting with monovalent anions, such as nitrate, chloride, iodide, bromide and thiocyanate. An analysis of the structural sites of these anions in the various lysozyme structures is presented. This study shows common anion sites whatever the crystal form and the chemical nature of anions, while others seem specific to a given geometry and a particular charge environment induced by the crystal packing.

1. Introduction

The ability of proteins to bind ions or small molecules has been proposed to explain the behaviour of protein solutions in the presence of various anions (Riès-Kautt & Ducruix, 1997). The nature of the anions drastically affects the solubility of proteins as well as their interaction potential in undersaturated solutions. The solubility curves of basic hen egg-white lysozyme (HEWL) (pI 11.2) crystallized at pH 4.5 in the presence of different salts (Riès-Kautt & Ducruix, 1989, 1991) are partly represented in Fig. 1 and have shown that the Hofmeister anion series is reversed, becoming

$$\mathrm{SCN^- \simeq pTS^- > I^- > NO3^- > Br^- > Cl^- > CH_3CO_2^- \simeq H_2PO_4^-},$$

where pTS^- stands for *para*-toluenesulfonate. Furthermore, anions were suspected to induce crystal polymorphism. A systematic study to detect anions on the protein surface has been undertaken in order to understand whether (i) anions simply interact with residues of opposite charge, changing the net charge of the protein and its solubility, or (ii) anions modify intermolecular contacts in the protein crystal, leading to polymorphism, or (iii) anions simply compete with water molecules.

Sulfate and phosphate anion sites have already been reviewed for protein structures in the PDB (Chakrabarti, 1993), as well as some chloride ions. This analysis has shown that the functional groups involved in the binding of these anions are mostly guanidinium from arginines and main-chain peptide groups, although polar or other basic residues may participate in these interactions. Recently, Dauter *et al.* (2000) demonstrated that many proteins can bind bromide or iodide

© 2001 International Union of Crystallography
Printed in Denmark – all rights reserved

Table 1
Lysozyme structures containing monovalent anions.

Where a PDB code is given in parentheses this means that the nitrate, chloride or sodium ions are localized at the same sites as in the structure mentioned by the PDB code. a.u., asymmetric unit; BPR, bromophenol red; BPB, bromophenol blue.

Type	PDB code	Space group	No. of molecules per a.u.	Resolution (Å)	Title	Number and type of ions in a.u.
HEWL	193l	$P4_32_12$	1	1.33	Earth-grown crystal	1 Cl^- + 1 Na^+
HEWL	194l	$P4_32_12$	1	1.40	Space-grown crystal	1 Cl^- + 1 Na^+
HEWL	1azf	$P4_32_12$	1	1.80	Wild type	5 Br^-
HEWL	1at5	$P4_32_12$	1	1.80	Succinimide residue tri-*N*-acetylchitotriose	1 Cl^- + 1 Na^+ (193l)
HEWL	1b0d	$P4_32_12$	1	1.85	Wild type	1 pTS^-
HEWL	—	$P4_32_12$	1	2.05	Soaked in NapTS solution	1 pTS^-
HEWL	—	$P4_32_12$	1	2.00	Wild type	1 BPR
HEWL	—	$P4_32_12$	1	2.00	Wild type	1 BPB
HEWL	1lz8	$P4_32_12$	1	1.80	Wild type	8 Cl^- + 1 Na^+
HEWL	1lz9	$P4_32_12$	1	1.80	Wild type	6 Br^- + 1 Na^+
HEWL	1hf4	$P2_1$	2	1.45	Space-grown crystal	8 NO_3^-
HEWL	1b2k	$P2_1$	2	1.60	Wild type	19 I^-
HEWL	1lkr	$P2_1$	2	1.60	Anisotropic refinement	17 I^-
HEWL	1lcn	$P2_1$	2	1.63	Wild type	3 SCN^-
HEWL	1lma	$P2_1$	1	1.75	Low-humidity form	2 NO_3^-
HEWL	5lym	$P2_1$	2	1.80	Wild type	6 NO_3^-
HEWL	3lzt	$P1$	1	0.92	Data collection at 120 K	6 NO_3^- + 3 ACT^-
HEWL	4lzt	$P1$	1	0.95	Data collection at 295 K	6 NO_3^-
HEWL	1lks	$P1$	1	1.10	Anisotropic refinement	6 NO_3^-
HEWL	—	$P1$	1	1.80	Neutron diffraction	4 NO_3^-
HEWL	2lzt	$P1$	1	1.97	Wild type	5 NO_3^-
HEWL	1lzn	$P1$	1	2.00	Neutron diffraction	4 NO_3^-
TEWL	1tew	$P6_122$	1	1.65	Wild type	1 SCN^-
BWQL	1dkk	$P2_1$	2	1.90	Wild type	4 NO_3^-
HL	1jsf	$P2_12_12_1$	1	1.15	Anisotropic refinement	7 NO_3^-
HL	1lz1	$P2_12_12_1$	1	1.50	Wild type	2 NO_3^-
HL	1lzr	$P2_12_12_1$	1	1.50	Tetraacetyl-chitotetraose	2 Cl^-
HL	1lmt	$P2_12_12_1$	1	1.60	Triacetyl-chitotriose peptide insertion	2 Cl^- (1lzr)
HL	1jka	$P2_12_12_1$	1	1.66	Mutant (E35D)	2 NO_3^-
HL	1jkb	$P2_12_12_1$	1	1.66	Mutant (E35A)	2 NO_3^- (1jka)
HL	1jkc	$P2_12_12_1$	1	1.60	Mutant (W109F)	2 NO_3^- (1jka)
HL	1jkd	$P2_12_12_1$	1	1.80	Mutant (W109A)	2 NO_3^- (1jka)
HL	1lz5	$P2_12_12_1$	1	1.80	Peptide insertion	2 Cl^- (1lzr)
HL	1lz6	$P2_12_12_1$	1	1.80	Peptide insertion	2 Cl^- (1lzr)

anions at their surface in a very short time period. This was advocated as a useful process in structure determination. The present study is focused on lysozyme, in particular HEWL, a monomeric protein of 129 amino acids which is extensively studied as a model protein by crystallographers and protein chemists as well as those working with protein folding. Since its first crystallization (Alderton & Fevold, 1946) and first three-dimensional structure (Blake *et al.*, 1965), new crystallization and structural features have been constantly under investigation. Recently, several HEWL structures have been refined at high resolution in different crystal forms: tetragonal HEWL/NaCl at 1.33 Å (Vaney *et al.*, 1996), triclinic HEWL/$NaNO_3$ at 0.92 Å (Walsh *et al.*, 1998) and orthorhombic HEWL/NaCl at 1.50 Å (Berthou *et al.*, 1983; further refined as reference 1aki in the PDB) in addition to a number of mutants. The 'standard' HEWL tetragonal crystal structure, usually obtained in sodium chloride, has been widely investigated in many laboratories. The hydration network and chloride ions present at the surface of the molecule were also determined, thus giving a large amount of information. All the lysozyme structures of interest for hydration-network analyses so far determined by X-ray diffraction are reported in Table 1.

From a crystallographic point of view, monoatomic ions may be easily confused with water molecules when inspecting electron-density maps. Furthermore, the selected force field used in the refinement programs may bias the results: a very low refined temperature factor and a significant remaining $F_o - F_c$ peak in the density map may be characteristic of an ion modelled as a water molecule. An elegant way to tackle this problem is to exploit the anomalous signal when relevant (Dauter *et al.*, 1999; Dauter & Dauter, 1999).

Another way to recognize such ions is to inspect the chemical environment and the contacts with polar atoms of the protein, depending on the nature and the size of the ions. An Na^+ ion requires more and shorter contacts than a water molecule and at least one Na^+ is now identified in the HEWL structure (Vaney *et al.*, 1996). Disordered structures cannot allow such subtle observations. High resolution (generally better than 2 Å) is required to discriminate water molecules from spherical ions (halides, Na^+ or K^+ cations or rare gases Kr, Xe), except for large or non-spherical ions such as SCN^-, NO_3^- and *a fortiori* pTS^-.

The present study focuses on interactions of *monovalent* anions at the surface of lysozyme in an attempt to understand the crystalline polymorphism induced by the nature of these salts. The sites of monovalent anions of four refined structures of HEWL are described and compared with those observed previously in other lysozyme structures in order to identify and formulate some common rules. The four new HEWL lysozyme structures discussed here are tetragonal HEWL/NapTS diffracting to 1.85 Å and three monoclinic forms, HEWL/NaI (1.6 Å), HEWL/$NaNO_3$ (1.45 Å) and HEWL/KSCN (1.63 Å) (Table 2). In addition, some turkey egg-white (TEWL), bob white quail (BWQL) and human (HL) lysozymes will also be mentioned in the discussion.

2. Materials and methods

2.1. Purification of lysozyme

Thrice-crystallized HEWL chloride (Cat. No. L-2879, batches 51H7150 and 73H7045, Sigma) was characterized as containing less than 0.3%(*w*/*w*) contaminant proteins (Riès-

Kautt *et al.*, 1997). This lysozyme was first dialyzed then deionized according to the above-mentioned work and stored freeze-dried. The isoionic protein powder was then dissolved in 50 m*M* sodium or potassium acetate buffer at pH 4.5 or brought to the pH of crystallization by addition of the acid corresponding to the anion to be studied. This standard procedure was applied to all batches of lysozyme used throughout.

2.2. Crystallizations

The crystallization conditions for the four HEWL complexes described in this paper differed in concentration, pH and crystallization techniques.

2.2.1. Tetragonal HEWL/NapTS crystals. Crystallization drops were prepared by mixing 5 µl of 40.4 mg ml^{-1} deionized lysozyme in 50 m*M* NaOAc buffer at pH 4.5 and 5 µl of reservoir. The drops were subjected to vapour diffusion at 291 K against a reservoir of 90 m*M* NapTS also in 50 m*M* NaOAc buffer at pH 4.5.

2.2.2. HEWL-soaked-NapTS crystals. Crystals of HEWL were obtained from standard HEWL crystallization conditions in the presence of 0.7 m*M* NaCl in 50 m*M* NaOAc buffer at pH 4.5 and at 291 K. For this purpose lysozyme was only dialyzed, because as the ions present are identical to those of crystallization only their concentration needs to be fixed. The HEWL/NaCl crystals were then soaked in a 1 m*M* solution of NapTS for 24 h.

2.2.3. Monoclinic HEWL/NaI crystals. Isoionic lysozyme was acidified with hydroiodic acid to a pH of 4.5; no buffer was used. Hanging drops were set up containing equal volumes of 50 mg ml^{-1} protein solution and reservoir. The drops were placed over wells containing 0.14 *M* of NaI (also adjusted to pH 4.5 with HI) at 291 K.

2.2.4. Monoclinic HEWL/$NaNO_3$ crystals. Previous work (Walsh *et al.*, 1998) showed that in the presence of sodium nitrate and over a period of one week to one month, monoclinic HEWL crystals obtained at 296 K slowly dissolve, while the triclinic form appears in the solution. However, at 277 K this monoclinic–triclinic transition does not proceed. In our experiments, crystals of HEWL/$NaNO_3$ were grown using the dialysis method under a microgravity environment during the USML-2 mission launched on 20 October 1995 aboard the Space Shuttle Columbia. The crystallization setup was the Advanced Protein Crystallization Facility (APCF) developed by the European Space Agency (Bosch *et al.*, 1992). A 10 mg ml^{-1} protein solution was prepared in 50 m*M* sodium acetate buffer at pH 5. A sodium nitrate concentration gradient was applied from 0 to 290 m*M* at 293 K. Of the several monoclinic crystals obtained in this setup, one crystal with dimensions 1 × 1 × 0.4 mm was used for data collection.

2.2.5. Monoclinic HEWL/KSCN crystals. Monoclinic thiocyanate crystals were grown at 291 K in hanging drops by the vapour-diffusion method. HEWL/KSCN crystals were obtained from solutions having a final concentration of 7.7 mg ml^{-1} isoionic protein dissolved in 50 m*M* potassium acetate buffer at pH 4.5 and 190 m*M* KSCN in the same buffer.

2.3. X-ray data collection and structural refinements

The data sets for the HEWL/NaI, HEWL/KSCN and HEWL/$NaNO_3$ structures were collected at room temperature using synchrotron radiation at LURE (Orsay, France) on the wiggler beamline DW32 (Fourme *et al.*, 1992) at a wavelength of 0.901 Å on a 180 mm diameter MAR Research image-plate detector. Data sets from the HEWL/NapTS and HEWL-soaked-NapTS crystals were recorded on an R-AXIS IIC detector with a Rigaku generator at a wavelength of 1.54 Å (Cu $K\alpha$; 40 kV, 50 mA). Table 2 summarizes data-reduction and refinement statistics for the four HEWL structures presented here. The significantly higher $R_{\rm sym}$ factor in the case of the 'space' HEWL/$NaNO_3$ crystal is probably a consequence of the delay in processing the crystals after the mission.

The starting models for the refinements of HEWL/NapTS and HEWL-soaked-NapTS structures were taken from the refined atomic coordinates at 1.33 Å resolution (Vaney *et al.*, 1996; PDB entry 193l). For HEWL/NaI and HEWL/KSCN structures, the position of the lysozyme molecules were determined by molecular replacement using the *AMoRe* program (Navaza, 1994) with 193l as the starting model. The translation between the two molecules of the asymmetric unit was found to be almost ($\frac{1}{2}$, 0, $\frac{1}{2}$) of the cell. Rigid-body refinements were then applied to the two solutions for each molecule and gave correlation coefficients of 54.1 and 72.1 with *R* factors of 40.3 and 31.4% for the HEWL/NaI and HEWL/KSCN structures, respectively. In the case of the HEWL/$NaNO_3$ structure, PDB entry 5lym was used as the

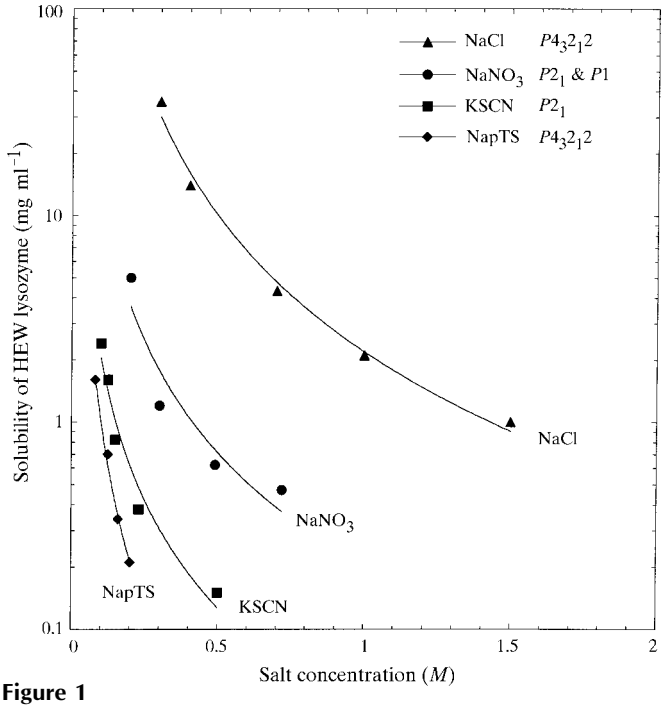

Figure 1
Solubility curves of lysozyme in different salt conditions.

Table 2
Summary of the data-collection and refinement statistics of lysozyme structures grown in the presence of different ions.

HEWL	NapTS	NaCl, NapTS soaked	NaI	$NaNO_3$ 'space'	KSCN
PDB code	1b0d	None	1b2k	1hf4	1lcn
Space group	$P4_32_12$	$P4_32_12$	$P2_1$	$P2_1$	$P2_1$
Unit-cell parameters					
a (Å)	78.93	79.06	27.73	27.94	28.01
b (Å)	78.93	79.06	62.79	62.73	63.18
c (Å)	38.28	37.97	59.84	60.25	60.41
β (°)	—	—	90.10	90.76	90.02
Z	8	8	4	4	4
Data collection					
No. of unique reflections	10484	7323	26152	35206	25539
Resolution range (Å)	13.5–1.85	15–2.07	13.3–1.60	10.0–1.45	15.8–1.63
R_{sym} (%)	4.9	5.5	9.5	12.3	2.8
Completeness (%)	97.5	94.7	96.7	95.5	96.8
Refinement					
Resolution limits (Å)	8–1.84	8–2.07	8–1.60	6–1.45	15.83–1.63
R factor (%)	20.2	20.8	19.84	21.5	19.75
R_{free} factor (%)	24.1	26.8	23.7	25.6	23.7
No. of reflections used	10461	7291	26000	31225	25515
No. of protein atoms	1001	1001	2002	2002	2002
No. of solvent molecules	86	62	302	246	136
No. of heteroatoms	11	11	19	28	9
No. of monovalent ions	1	1	19	7	3
R.m.s.d.s from ideality					
Bond lengths (Å)	0.008	0.007	0.004	0.006	0.006
Bond angles (°)	1.32	1.37	1.17	1.35	1.35
Improper angles (°)	1.13	1.06	0.60	1.14	1.18
Dihedral angles (°)	23.3	23.4	25.2	23.8	23.8
Average $\langle B \rangle$ factors (Å^2)					
$\langle B \rangle$ Wilson	21.4	18.2	15.2	13.7	18.5
Main-chain atoms	15.0	13.8	12.2	11.9	19.5
Side-chain atoms	19.4	17.0	14.7	15.5	23.6
Water molecules	33.6	28.0	29.0	30.4	32.9
Ions	17.3 (pTS^-)	45.9 (pTS^-)	33.4 (I^-)	25.8 (NO_3^-)	47.6 (SCN^-)

starting model, corresponding to the monoclinic form (Rao & Sundaralingam, 1996).

All the protein structures were refined with the conjugate-gradient facility described in the *X-PLOR* program (Brünger, 1992), using the accurate parameter files described by Engh & Huber (1991).

The ions and water molecules were located after successive cycles of Fourier maps ($2F_o - F_c$ and $F_o - F_c$ maps at their nominal resolution) together with graphic inspections and reconstructions.

2.4. Structure comparisons

Before studying lysozyme structures from other species, preliminary sequence alignments were undertaken to decide whether or not these lysozymes can be included in a structural comparison with HEWL. Alignments were made with *FASTA* (Pearson, 1996) at the Infobiogen site (GIS Infobiogen, France). For example, 1dkk (BWQ lysozyme) and HEWL sequences share 96.9% sequence identity over the whole amino-acid sequence, while goose egg-white lysozymes were excluded from the analysis (153l and 154l; Weaver *et al.*, 1995) because of the non-overall homologous three-dimensional structure with HEWL.

With the expection of the four new structures described in this paper, all the lysozyme structures included in this study were extracted from the Protein Data Bank release of March 2000 (Abola *et al.*, 1987; Bernstein *et al.*, 1977). References for the lysozyme structures we have used are as follows: hen egg-white lysozyme, 5lym (Rao & Sundaralingam, 1996), 2lzt (Ramanadham *et al.*, 1990), 193l and 194l (Vaney *et al.*, 1996), 1lkr, 1lks (Steinrauf, 1998), 3lzt and 4lzt (Walsh *et al.*, 1998), 1azf (Lim *et al.*, 1998), 1lz8 (Dauter *et al.*, 1999) and 1lz9 (Dauter & Dauter, 1999); turkey egg-white lysozyme, 1tew (Howell, 1995); bob quail egg-white lysozyme, 1dkk (Chacko *et al.*, 1996); human lysozyme, 1jka, 1jkb, 1jkc and 1jkd (Muraki *et al.*, 1997), 1lmt (Yamada *et al.*, 1995), 1lz5 and 1lz6 (Yamada *et al.*, 1993) and 1lzr (Song *et al.*, 1994). Other coordinates, kindly provided by their authors, were included in the list: water molecules and nitrate ions for the human lysozyme structure (1lz1) at 1.5 Å resolution (Blake *et al.*, 1983), two hen egg-white lysozyme structures complexed with neutral bromophenol red (BPR) and with neutral bromophenol blue (BPB) (Madhusudan & Vijayan, 1992) and three sets of HEWL coordinates studied using neutron diffraction: 1lzn (Bon *et al.*, 1999), crystallized in dimethyl sulfoxide (Lehmann *et al.*, 1985) and in ethanol (Lehmann & Stansfield, 1989). Although lysozyme structures crystallized in other media have been published (using different concentrations of urea; Pike & Acharya, 1994) their coordinates were not deposited with the PDB. These structures were not included in this paper. Since the preparation of this work, several other lysozyme structures have become available, particularly a number of mutants, some of them containing characterized anions and some others refined with all density peaks as waters. They were not included in this paper: Cl^- in 1bg1, 1c10, 1dpw, 1dpx, 1qtk (HEWL) or nitrate in 1rem (HL).

For the sake of comparison, superimposition of each molecule of lysozyme was performed by alignments based on their secondary structure using the *SARF* program (Alexandrov, 1996). Symmetry-related ions were generated for each asymmetric unit containing lysozyme using the *CONTACT* and *PBSET* programs from the *CCP*4 package (Collaborative Computational Project, Number 4, 1994). Within the list of symmetry-related counterions, only those interacting directly with the molecule(s) in the asymmetric unit were retained for further analysis. The hydrogen-bond distances between anions and lysozyme atoms are taken within 2.6 and 3.4 Å, except iodide atoms which were analyzed in the range 3–4.5 Å. The lysozyme molecules and counterions from the various struc-

tures are displayed in Fig. 2 using the *GRASP* program (Nicholls *et al.*, 1991).

The atomic coordinates and structure factors of the structures HEWL/NapTS, HEWL/NaI, HEWL/$NaNO_3$ and HEWL/KSCN were deposited with the PDB (reference codes 1b0d, 1b2k, 1hf4 and 1lcn, respectively). The HEWL-soaked-NapTS coordinates were not submitted to the PDB since the HEWL/NapTS structure is at a higher resolution, showing the pTS^- ion at the same position but with lower atomic temperature factors.

3. Results and structure descriptions

3.1. Tetragonal HEWL/NapTS

Since our first studies on lysozyme solubilities, SCN^- and pTS^- were found to be the most efficient anions for decreasing the solubility of HEWL at pH 4.5 and are thought to interact with positively charged residues (Riès-Kautt & Ducruix, 1991). Although the decrease in solubility arising from SCN^- could be transposed to the crystallization of other basic proteins such as BPTI (Hamiaux *et al.*, 1999) and erabutoxin b (Saludjian *et al.*, 1992), pTS^- seems to present a dual behaviour. The solubility of the protein may increase or decrease depending on the interactions through the hydrophilic sulfonic group or the hydrophobic toluene group.

To understand how pTS^- interacts with the protein, large HEWL crystals were grown using NapTS as a crystallization agent in the range 50–200 m*M* and the structure was refined. A single fully occupied pTS^- ion is observed between three symmetry-related lysozyme molecules, displaying five contacts with the three molecules plus additional contacts with the solvent. Hydrogen bonding and hydrophobic contacts of the anion are shown in Fig. 3.

In order to check whether the pTS^- molecule lies within a preferential site, a tetragonal HEWL/NaCl crystal was soaked in a 1 m*M* NapTS solution at 291 K. In the refined structure of the HEWL-soaked-NapTS, a pTS^- ion is located in the same site as the previously observed site in the HEWL/NapTS structure but with a lower occupancy factor. However, the pTS^- ion observed in the soaked crystal has looser interactions with the protein and in particular with the atoms of symmetry-related lysozyme molecules. As shown in Table 2, the average $\langle B \rangle$ value for the pTS^- atoms is higher in the soaked crystal structure than in the co-crystallized structure. The pTS^- itself displays equivalent conformations. The same observation holds for the protein: the r.m.s. deviations between HEWL/NapTS and HEWL-soaked-NapTS structures are 0.19 Å (C^α atoms) and 0.76 Å (all protein atoms), respectively.

3.2. Monoclinic HEWL/NaI

While this work was in progress, another monoclinic HEWL/NaI structure at 1.6 Å was published (Steinrauf, 1998; PDB code 1lkr). The 1lkr structure shows 17 I^-, whereas we observe 19 I^- ions. However, both the crystallization and refinement conditions were different: 10 mg ml^{-1} lysozyme at pH 8 with ~330 m*M* NaI was used for 1lkr, while the crystal used in this work was grown from 50 mg ml^{-1} lysozyme at pH 4.2, although the NaI concentration was lower (140 m*M*). The 1lkr structure was refined with anisotropic thermal parameters and loose constraints in such a way that it cannot be structurally compared with the present HEWL/NaI complex.

The monoclinic HEWL/NaI contains two molecules of lysozyme in the asymmetric unit. The 19 I^- ions identified interact mostly with polar atoms of the protein and are generally found at the interface of lysozyme molecules (Fig. 2*a*).

Almost all of the 19 iodide atoms found in our structure present characteristic residual peaks in the difference Fourier maps corresponding to an anisotropic behaviour. However, owing to the limit in resolution, all atoms were treated isotropically.

Five common I^- sites are shared between molecules *A* and *B* of the asymmetric unit (this corresponds to ten of the 19 ions observed).

As already observed for 1lkr, the loop 65–75 of molecule *B* has a distinct conformation compared with that usually observed in other monoclinic structures and also in the *P*1, $P2_12_12_1$ or $P4_32_12$ crystal forms. This loop is perfectly defined in the electron density, with an average temperature factor of 19.0 $Å^2$ over all the atoms of molecule *A*, while $\langle B \rangle$ is 13.6 $Å^2$ in molecule *B*. The r.m.s.d.s between the C^α atoms of molecules *A* and *B* is 1.56 Å, with a maximum displacement of 9.15 Å for Gly71 (Fig. 4). The presence of I^- 403 interacting with Gly71 O near loop 65–75 may explain the different loop conformations.

This new loop conformation induces different packing interactions with symmetry-related molecules and leads to a stacking of Arg73 with Trp62, both perfectly defined in the $2F_o - F_c$ density map. In the triclinic HEWL/$NaNO_3$ structures (2lzt, 3lzt, 4lzt), guanidinium and aromatic contacts from Arg73 and Trp62 residues were also observed, but the loop 65–75 remains in the 'standard' conformation. These planar stacking interactions of arginine and aromatic side chains had previously been noticed in various structures other than lysozymes (Burley & Petsko, 1986; Flocco & Mowbray, 1994; Levitt & Perutz, 1988). In our case, the presence of I^- 403 constrains the loop in such a way that Arg73 lies in a disallowed ($\varphi = 42°$, $\psi = -116°$) region of the Ramachandran diagram.

3.3. Monoclinic HEWL/$NaNO_3$

Seven NO_3^- ions were first clearly identified in the structure of monoclinic HEWL/$NaNO_3$ after refinement. Of the seven NO_3^- ions, four could be considered as belonging to specific lysozyme sites as they lie at the same position between molecules *A* and *B* (in sites 2 and 4). For example, NO_3^- 801*B* is located between molecules *A* and *B* of the asymmetric unit and NO_3^- 801*A* is located between molecule *B* and a symmetry-related molecule *A*. The two positions are 'equivalent' and are described as site 4 in Table 3 and Fig. 5

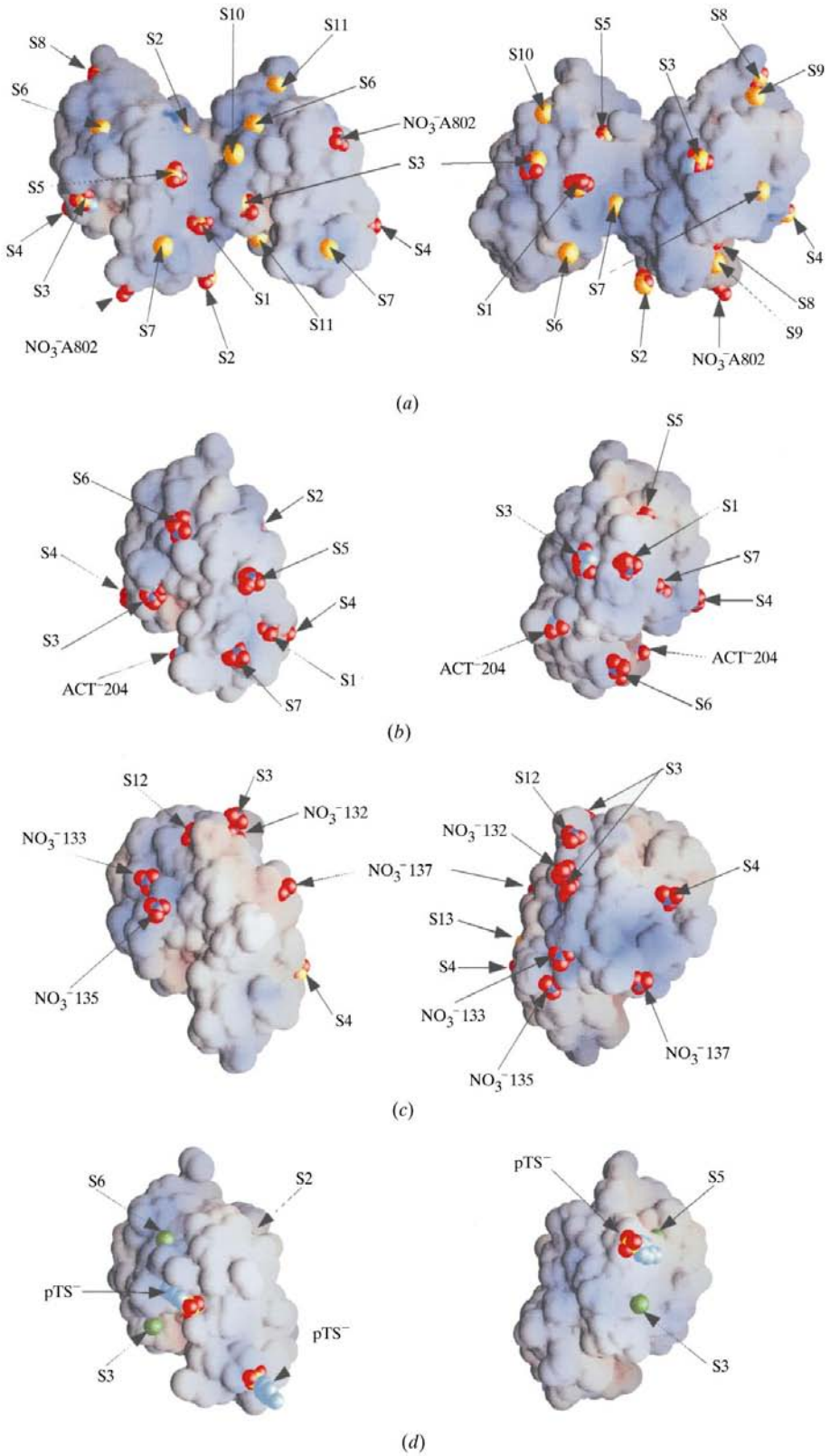

Figure 2
The anionic sites at the surface of the lysozyme molecule. Each of the common sites are labelled with the letter S, while the unique sites are labelled with the name of the specific anion observed. Left and right parts represent the head and tail views of the asymmetric unit in (*a*) the dimeric monoclinic lysozyme with iodine (yellow spheres) and nitrate ions, (*b*) triclinic lysozyme, (*c*) monoclinic monomeric HEWL/$NaNO_3$ and (*d*) the tetragonal form of HEWL/$NapTs^-$.

(*RIBBONS* program; Carson, 1987). An eighth NO_3^- ion was localized during the analysis of the common anionic sites.

The nitrate ion NO_3^- 802*B* located near Trp62 of molecule *B* seems to force the flipping of Trp62 side chain compared with the position of the same residue in molecule *A*.

Several NO_3^- ions were found in other lysozyme structures, essentially in the tetragonal and monoclinic forms. Their binding to the protein and their common sites are listed in Table 3.

3.4. Monoclinic HEWL/KSCN

The structure of monoclinic HEWL/KSCN (two molecules per asymmetric unit) refined at 1.63 Å resolution exhibits some of the characteristics mentioned previously in the HEWL/NaI structure, with loop 65–75 of molecule *B* adopting the alternate conformation. However, in HEWL/KSCN the loop is not well defined and some electron density is missing along the main chain, as evidenced by an average $\langle B \rangle$ factor over the ten residues of 36.1 $Å^2$ (compared with a value of 21.5 $Å^2$ for the whole protein).

A strong density peak was also found at the same location as I^- 403 in the HEWL/NaI structure, but is difficult to ascertain as a SCN^- ion. It was decided to keep Wat66 in the deposited PDB file at this position, but the presence of a disordered thiocyanate may not be excluded.

Finally, three unambiguous SCN^- ions were located in two sites: SCN^- 201 is bound to molecule *B* (site 1) but not found at the equivalent position in molecule *A*; SCN^- 202 and SCN^- 203 stand at the interface between two lysozyme molecules (site 4) and occupy the same site in molecules *A* and *B*.

4. Discussion: structural comparisons and description of the anion sites

The methods used to compare the anion sites in the lysozyme structures

are detailed in §2. The interactions of the ions with the lysozyme molecules were tentatively classified by pooling sites for which distances and types of residues involved are comparable, as listed in Table 3. Fig. 2 gives an overview of the distribution of the ions at the surface of HEWL for monoclinic, triclinic, tetragonal and orthorhombic forms.

To rationalyze, these anionic sites may be divided into three main categories: firstly, some common (lysozyme-specific) binding sites independent of the space groups but related to the nature of side chains at the surface; secondly, space-group dependent sites for which anions sit in places generated by crystal packings; finally, sites that are specific to the anion itself.

Nevertheless, there is no simple rule to decide which is which. Indeed, a given site is generally not the result of a single effect but more often a combination of the three. Hence, there is no such 'universal' anionic site, *i.e.* a site that is observed in all the structures. Because these effects vary from a crystal structure to another, some of them are more frequently observed, whereas others are observed only once.

4.1. Lysozyme-specific sites

In $P2_1$, iodide ions fill almost all anionic sites (from sites 1 to 11, excluding sites 12, 13 and 14). As mentioned above, in this space group some monovalent ions, in particular those at sites 2–7, are 'duplicated', *i.e.* a given ion interacts with the two lysozyme molecules of the asymmetric unit at the same place, whatever their environment: they are 'pure' specific sites of the lysozyme molecule. This is seen particularly for sites 1–4, which accommodate various kind of anions whatever their size (Fig. 5). These sites are also observed in the triclinic forms (sites 1, 4, 6, 7).

Site 1 (Fig. 5) exists in all $P1$ and $P2_1$ structures listed in Table 1, except for 1lma ($P2_1$ with one molecule per asymmetric unit). In this site, NO_3^- and I^- anions interact with Cys80 N and Ser81 N (Table 3). In the $P1$ form, NO_3^- additionally interacts with Arg21 N^{ε} of the other molecule in the asymmetric unit. To compare the nitrate site 1 in a monoclinic form with another nitrate of the triclinic form, the symmetry-related lysozyme molecules must be generated in this latter case. They show that equivalent locations exist for NO_3^- 202 and NO_3^- 512.

Site 2 exists in four crystal forms (triclinic, monoclinic, tetragonal and hexagonal). In addition to I^- or NO_3^- anions, it also accommodates Br^- (1azf) or SCN^- (1tew). This site appears to be specific to lysozyme and is probably important for modulating crystal polymorphism.

In site 3 the presence of a Cl^- (*i.e.* 193l and 194l structures) or a Br^- ion (1azf) interacting with Asn113 $N^{\delta 2}$ and with the symmetry-related phenolic O atom of Tyr23 seems specific for the $P4_32_12$ packing. However, the crystal packing of HEWL in $P2_1$ (5lym, HEWL/NaI and 1lkr) shares common anion sites involving the side chain Asn113. Despite the fact that monoclinic forms have different crystal packings compared with the tetragonal forms, the presence of the residue Asn113 at the surface of the protein combined with other basic residues from symmetry-related proteins (or from the second molecule of the asymmetric unit) leads to a strong polar environment (for example Arg*A*114 for 5lym and Lys*A*96* for HEWL/NaI) favourable for the binding of an anion. Only the $P1$ structure 1lks (1.10 Å) contains a nitrate ion at this site, while the

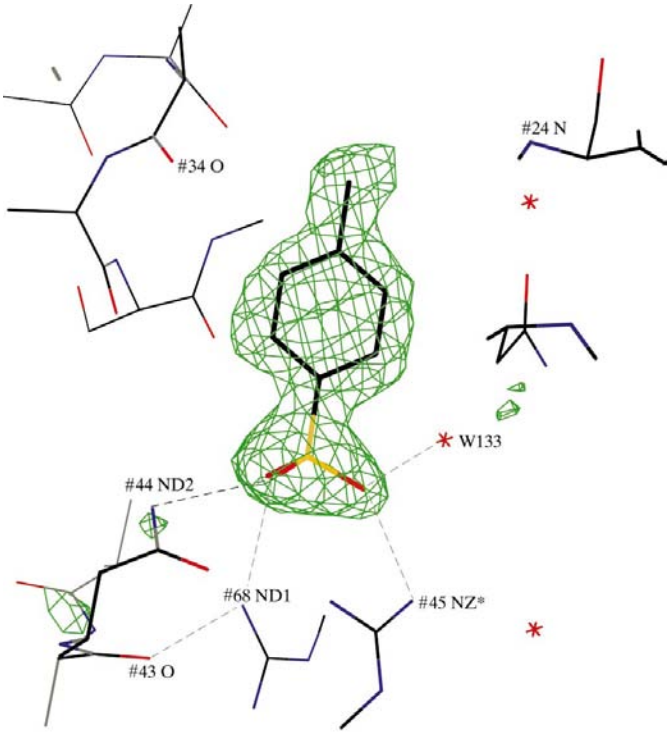

Figure 3
Details of the interactions around the pTs^- anion (site 14 in 1b0d). The omit map is calculated from the refined theoretical phases; contouring is at 2σ above the mean background level.

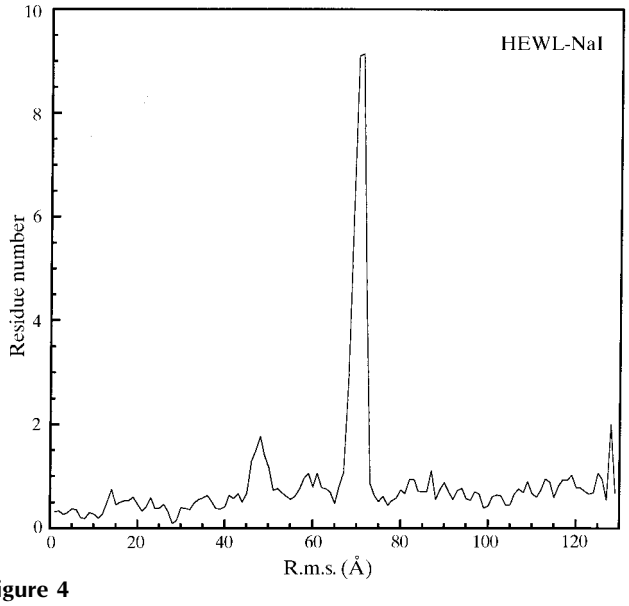

Figure 4
R.m.s. deviations *versus* the sequence number between the two independent molecules of the HEWL/NaI monoclinic structure showing the non-equivalence of loops 65–75.

Table 3
Common anionic sites in lysozyme structures.

Residues in bold indicate, within a given crystalline form, those always interacting with the monoanion. Comparing different crystalline forms, the common residues that interact with the anion are vertically aligned for clarity. The fifth column denotes the number of positions found for a given anion.

Sites	Type of ions	Space group	Lyso-zyme	Intra/ Inter	Hydrogen bonds (side- or main-chain atom)						Charged residues			Other interactions			
1	NO_3^-, I^-	$P2_1$	HEW, BWQ	2/3	Arg21	Gly22		Asp66	**Cys80**	**Ser81**	Arg21			Asn65 O			
	SCN^-	$P2_1$	HEW	0/1			Tyr53			Ser81							
	NO_3^-	$P1$	HEW	0/6	**Arg21**				Cys80	**Ser81**				Arg21 O			
2	NO_3^-, I^-	$P2_1$	HEW	6/3		Gly71	Ser72	**Ile88**	**Ser86**	Thr89	Arg14	His15		Thr69 O	Gly71 O	Ser86 O	Asp87 $O^{\delta1}$
	NO_3^-	$P1$	HEW	5/0				**Ile88**				His15					**Asp87 $O^{\delta1}$**
	Br^-, Cl^-	$P4_32_12$	HEW	1/0				Ile88			Arg14	His15					
	SCN^-	$P6_122$	TEW	0/1	Arg14						Arg14						
3	NO_3^-, I^-, ACT^-	$P2_1$	HEW	2/4	Arg21			Asn113			Arg21	Lys96	Arg114	Gly16 O	Asn19 O	Asn113 $O^{\delta1}$	
	NO_3^-, ACT^-	$P1$	HEW	0/2				Asn113	Arg114				Arg114				
	Br^-, Cl^-	$P4_32_12$	HEW	0/3			**Tyr23**	**Asn113**									
	NO_3^-	$P2_12_12_1$	Human	0/1							Lys97 (Lys96 in HEWL)						
4	NO_3^-	$P2_1$	HEW	1/7		Asn65	**Asn74**	Asn77	Arg112		Arg112	**Lys116**		Asn77 O	Arg112 O		
	SCN^-	$P2_1$	HEW	0/2		Asn65	Asn74				Arg112	Lys116					
	NO_3^-, Br^-	$P1$	HEW	0/6		Asn65	**Asn74**					**Lys116**					
	Br^-, Cl^-	$P4_32_12$	HEW	2/0		Asn65	Asn74										
	NO_3^-, Cl^-	$P2_12_12_1$	Human	0/6	**Asn27**	**Asn66**	Asn75										
5	NO_3^-, I^-	$P2_1$	HEW	3/2	**Ser24**	Leu25	**Gly26**				Gln41	Gln121		Val120 O			
	NO_3^-	$P1$	HEW	1/2	**Ser24**	**Leu25**	**Gly26**				**Gln41**	Gln121					
	Br^-, Cl^-	$P4_32_12$	HEW	1/0	Ser24		Gly26										
6	I^-	$P2_1$	HEW	2/2							**Arg5**	**Lys33**		Asp101 O			
	NO_3^-	$P1$	HEW	2/4								**Lys33**	Arg73				
	Br^-, BPR, Cl^-	$P4_32_12$	HEW	1/1								**Lys33**					
7	I^-	$P2_1$	HEW	0/4	**Tyr23**	Arg45	Asn77		**Asn106**		Arg45		Lys116				
	NO_3^-	$P1$	HEW	0/3	**Tyr23**			**Met105**	**Asn106**		**Arg45**	Arg68					
8	I^-, NO_3^-	$P2_1$	HEW	2/2	Trp63	Gly126	Arg128				Arg61	Arg128					
9	I^-	$P2_1$	HEW	0/2	**Trp62**	**Arg61**								**Gln121 O**			
10	I^-	$P2_1$	HEW	0/2							**Lys13**	**Arg114**					
11	I^-	$P2_1$	HEW	0/2	**Asn103**						Arg112	Arg128		Cys6 S^{γ}	**Gly126 O**		
12	NO_3^-	$P2_1$	HEW	0/1	Arg5	Cys6	Glu7		Asn106								
	NO_3^-	$P2_12_12_1$	Human	1/1		**Cys6**	**Glu7**	Arg14			Arg14						
13	Cl^-	$P2_12_12_1$	Human	0/3	**Ser80** (Pro79 in HEWL)						**Arg107** (Asn106 in HEWL)						
14	pTS^-, Cl^-	$P4_32_12$	HEW	0/2	Asn44	Arg45	Arg68							Gly22 O			

presence of an acetate ion is described for 3lzt. Despite the higher resolution of the 4lzt (0.95 Å), no NO_3^- ion is observed at this site. Only one structure in $P2_12_12_1$ contains a nitrate ion: the high-resolution structure of 1jsf (1.15 Å).

In the $P2_1$ structures, the anionic site 4 (Fig. 5) appears twice and near the two independent molecules of the asymmetric unit. It concerns almost all the $P2_1$ crystal structures except 5lym and 1lma (one molecule per asymmetric unit). On the other hand, all six $P1$ crystal structures include a nitrate ion here. For the $P4_32_12$ space group only the 1azf structure is listed; the human lysozyme ($P2_12_12_1$) crystal form contains an anion at this site. For the monoclinic and the triclinic structures the packing shows that the same symmetry-related residues are involved in the contacts with the anions. These include the side chains of Asn65, Asn74 and Asn77 and the main chain (Ile78) of one molecule and Lys116, Arg112 of the other molecule of the asymmetric unit ($P2_1$) or of the symmetry-related molecule ($P1$). In the orthorhombic crystal form (HL only), the packing is different and the residues concerned are now Asn66, Asn75 (corresponding to Asn65 and Asn74 in HEWL) and Asn27 of a symmetry-related molecule.

In site 5, two of the triclinic forms (4lzt and 1lks) show the presence of a nitrate ion, but 3lzt (0.92 Å) does not. In the monoclinic form (iodide), this site as well as site 2 are occupied and thus can be considered as 'intrinsic' to the lysozyme molecule(s). However, only the 5lym structure contains a unique nitrate ion at an intramolecular position within the molecule *A*. In molecule *B*, the residue Gln 121*B* interacts with Ser24*B* O^{γ} so that the atoms $O^{\varepsilon1}$ and $N^{\varepsilon2}$ are localized at the nitrate position. For 1lma the residue Gln121 prevents the fixation of a nitrate ion, but in the 1dkk and HEWL/$NaNO_3$ structures a water molecule stands at the position of NO_3^- 514. In the tetragonal crystal form, only the 1azf structure contains an anion at that site.

In all the $P1$ structures site 6 is fully occupied by a nitrate ion. Both HEWL/NaI structures (space group $P2_1$) show the existence of a double iodide site. Each iodide anion is bound to one lysozyme molecule and to another molecule of the asymmetric unit involving hydrogen bonding with two basic

residues, Arg5 and Lys33. None of the other HEWL/NO_3^- structures contain a NO_3^- ion, in spite of the fact that the $P2_1$ packing could easily accommodate this anion.

4.2. Space-group dependent sites

Sites 7 to 11 are more frequently observed in one particular crystal form of lysozyme when several forms are compared with the others. The packing effect arises from the different environment probed by the lysozyme molecules in the packing. This is especially true for the monoclinic form, where two molecules are found in the asymmetric unit. Their orientations with respect to their symmetry-related mates also generate additional non-equivalent anionic sites 8, 9, 10 and 11 observed only at the surface of one monomer and not the other and which are not observed in P_1, $P2_12_12_1$ and $P4_32_12$ space groups.

It should be taken into account that sites 9, 10 and 11 are specific iodide sites as they are only found in the two HEWL iodide structures, while site 8 also accommodates a nitrate. Although the anion NO_3^- 150 of the 3lzt structure is near site 9 (O1 is about 2.1 Å from I^- 415) and the aromatic plane of Trp62 in 3lzt moves from 90° from its position in HEWL/NaI or 1lkr structures, NO_3^- 150 was not included in site 9 because of the distance being too great. This may be an example of an anionic site displaced by conformational changes of some residues.

Sites 1 and 7, which were previously termed as 'specific' to lysozyme, should be specially mentioned. When comparing the monoclinic $P2_1$ (two molecules per asymmetric unit) and triclinic $P1$ (one molecule per asymmetric unit) space groups, the relative orientation of the lysozyme molecules along the crystallographic c axis are similar, as pointed out by Salemme *et al.* (1988), so equivalent sites may be built in both cases. However, 1lma ($P2_1$, one molecule per asymmetric unit) is the only $P2_1$ structure containing no anion at these sites. Generating symmetry-related lysozyme molecules shows that sites 1 and 7 are narrowed owing to packing considerations in such a way that no nitrate ion could fill these positions. Further, the tetragonal forms and $P2_12_12_1$ HL could not accommodate a nitrate ion at that site either, also because of a tight packing arrangement.

As a consequence, their absence does not imply that they are non-specific to lysozyme, but means they are 'packing forbidden'.

4.3. Particular (non-specific) sites

Some anions in the various lysozyme structures were found to have one or two ions in sites 12 to 14. This concerns one nitrate ion in the HEWL/$NaNO_3$ (this work), one nitrate anion and three acetate ions in the 3lzt structure, the BPB molecule in the HEWL/BPB structure, the *para*-toluenesulfonate in HEWL/NapTS plus one chloride (1lz8) and five nitrate ions in the HL structure (1lz1 and 1jsf).

Site 13 concerns only HL structures ($P2_12_12_1$) and a chloride ion. The intermolecular interactions with the chloride are with Ser80 N, Arg107 $N^{\eta 2}$ and water molecules. Arg107 is not conserved in HEWL (replaced by Asn106). The same holds for site 14.

The other common chloride sites described in HL structures 1lmt (Cl^- 145), 1lzr (Cl^- 131), 1lz5 and 1lz6 (Cl^- 131) in the $P2_12_12_1$ crystal form were not included in the present study because of the non-homologous amino-acid sequence of HL with HEWL for residues involved in the binding. They have no equivalent in HEWL but can alternatively be considered as HL-specific.

In general, the interactions of anions with the protein atoms involve basic residues such as Arg and Lys, hydrophilic residues (especially Asn) and a few other residues such as Ser, Gly, Ile and main-chain atoms through the hydrogen-bonding scheme (see Table 3). Usually, water molecules are strongly involved in interactions with anions, although some anions are not observed to interact with solvent because the second layer of solvent is more difficult to characterize. This concerns several nitrate ions, bromide, thiocyanate and iodide ions distributed among all the anionic sites.

More than 50% of the contacts between anions and lysozyme are engaged through three residues, Asn, Arg and Lys, showing clearly that hydrophilic residues, especially those which are positively charged, are the preferred environment for anions (Fig. 6).

One of the most important questions that can be addressed concerns the influence of anions on the crystal form packing and how they favour a given form over another. Lysozyme-specific sites are not implicated in the polymorphism as they always favour the same charge distribution at the surface (except for bulky anions). Conversely, the single unique sites may be relevant as they are the result of non-specific interactions between the numerous polar sites on the protein surface.

In this context, small ions as Cl^- and Br^- (halide ions) show a tendency to induce tetragonal packing, while large monovalent ions of low charge density such as I^-, NO_3^- and SCN^- appear to favour monoclinic packings.

5. Conclusions

The present work aims to embody known data for the environment of the protein, which is not only surrounded by the water layers but also by counterions, a fact which is often put aside. Although some of the ions can be considered as 'sticky' (attractive forces), none of them presents a binding constant which can be measured by chemical methods. This work shows that some of the locations at the surface of the protein are favored by ions, contrary to a random distribution. This is influenced by the charges of the side chains, packing considerations and the chemical specificity of each anion. The lysozyme molecule has the great advantage of crystallizing in a large variety of crystal systems ranging from triclinic to hexagonal and presents a number of common anionic sites that can be rationalized. Although no site is common to all structures, most of them are observed in a majority of crystal structures whatever their nature. All these sites may exist if three main features are present: a sequence-dependent

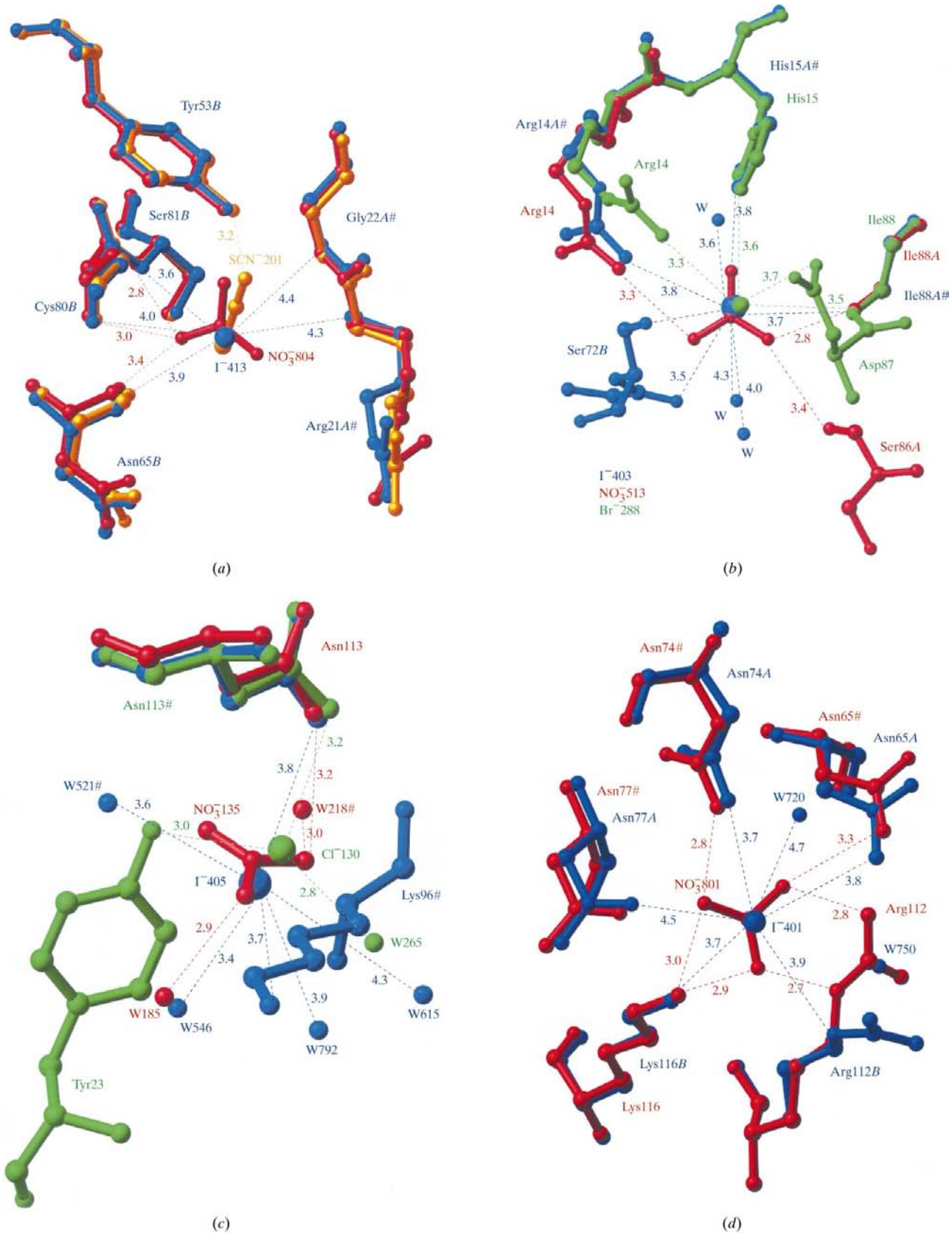

Figure 5
Four of the main ionic sites in lysozyme. (*a*) Site 1, a common site for nitrate (in red), iodide (in blue) and thiocyanate (in yellow) ions. (*b*) Site 2 shares a nitrate, a bromide and a iodide anion; HEWL/NaI in blue, 5lym in red and 1azf in green. (*c*) Site 3 includes an iodide (HEWL/NaI in blue), a nitrate (1lks in red) and a chloride (193l in green). (*d*) Site 4 (HEWL/$NaNO_3$, red; HEWL/NaI, blue) showing modifications induced in the amino-acid side chains that line the cavity.

(protein-specific) term, a packing-dependent and a anion-specific (unique) terms. All these parameters combined together favour binding tight enough to give a visible signature in electron-density maps and to drive the building of a thermodynamically stable crystal form.

It is evident that more high-resolution structures will be needed to fully characterize the ionic environment in this model protein. A recent similar approach has also been applied to the bovine pancreatic trypsin inhibitor (BPTI) crystallized in acidic media (Hamiaux *et al.*, 1999, 2000). BPTI is also a good example of counterion influence on polymorphism and temperature dependencies. In fact, BPTI crystallized in presence of SCN^- shows an increasing solubility with temperature, whereas this is reversed in the presence of Cl^- and sulfate. In the case of lysozyme, the use of very high resolution experimental phases should definitively help to characterize the distribution of anions at the surface and to apply electrostatic mapping concepts.

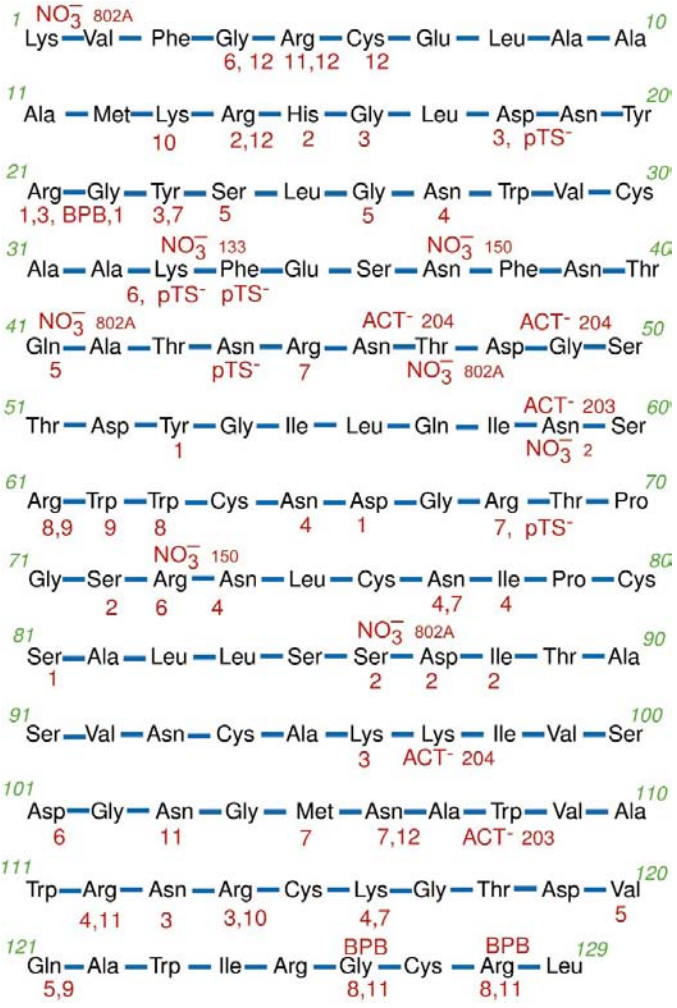

Figure 6
All anionic sites (red numbers) aligned on the lysozyme sequence. The unique sites are labelled according to the chemical formula of the anion taken from the corresponding PDB file.

This work was supported by the European Biocrystallogenesis initiative [contract No. BIO4-CT98-0086(DG12-SSMI)]. We thank the European Space Agency (ESA), the Centre National d'Etudes Spatiales (CNES) and the Centre National de la Recherche Scientifique (CNRS) for financial support of this project. IB was supported by CNES. PR and SL were supported by ESA. We acknowledge R. Fourme and J. P. Benoit for the development of the W32 station at LURE (Orsay, France). We are grateful to Dr Alexandrov who made the *SARF* program available *via* an ftp site. We are indebted to C. C. F. Blake for providing the human lysozyme coordinates at 1.5 Å resolution (Blake *et al.*, 1983) and M. Vijayan for the hen egg-white lysozyme coordinates complexed with BPR and BPB (Madhusudan & Vijayan, 1992).

References

Abola, E. E., Bernstein, F. C., Bryant, S. H., Koetzle, T. F. & Weng, J. (1987). *Protein Data Bank* in *Crystallographic Databases: Information Content, Software Systems, Scientific Applications*, edited by F. H. Allen, G. Bergerhoff & R. Sievers, pp. 107–132. Bonn/Cambridge/Chester: IUCr.
Alderton, G. & Fevold, H. L (1946). *J. Biol. Chem.* **164**, 1–5.
Alexandrov, N. N. (1996). *Protein Eng.* **9**, 727–732.
Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. F., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T. & Tasumi, M. (1977). *J. Mol. Biol.* **112**, 535–542.
Berthou, J., Lifchitz, A., Artymiuk, P. & Jolles, P. (1983). *Proc. R. Soc London Ser. B*, **217**, 471–489.
Blake, C. C. F., Koenig, D. F., Mair, G. A., North, A. C. T., Philips, D. C. & Sarma, V. R. (1965). *Nature (London)*, **206**, 757–761.
Blake, C. C. F., Pulford, W. C. A. & Artymiuk, P. J. (1983). *J. Mol. Biol.* **167**, 693–723.
Bon, C., Lehmann, M. S. & Wilkinson, C. (1999). *Acta Cryst.* D**55**, 978–987.
Bosch, R., Lautenschlager, P., Potthast, L. & Stapelmann, J. (1992). *J. Cryst. Growth*, **122**, 310–316.
Brünger, A. T. (1992). *X-PLOR Manual. Version 3.1. A System for X-ray Crystallography and NMR.* New Haven/London: Yale University Press.
Burley, S. K. & Petsko, G. A. (1986). *FEBS Lett.* **203**, 139–143.
Carson, M. (1987). *J. Mol. Graph.* **5**, 103–106.
Chacko, S., Silverton, E. W., Smith-Gill, S. J., Davies, D. R., Shick, K. A., Xavier, K. A., Willson, R. C., Jeffrey, P. D., Chang, C. Y., Sieker, L. C. & Sheriff, S. (1996). *Proteins Struct. Funct. Genet.* **26**, 55–65.
Chakrabarti, P. (1993). *J. Mol. Biol.* **234**, 463–482.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Dauter, Z. & Dauter, M. (1999). *J. Mol. Biol.* **289**, 93–101.
Dauter, Z., Dauter, M., de La Fortelle, E., Bricogne, G. & Sheldrick, G. (1999). *J. Mol. Biol.* **289**, 83–92.
Dauter, Z., Dauter, M. & Rajashankar, K. R. (2000). *Acta Cryst.* D**56**, 232–237.
Engh, R. A. & Huber, R. (1991). *Acta Cryst.* A**47**, 392–400.
Flocco, M. M. & Mowbray, S. L. (1994). *J. Mol. Biol.* **235**, 709–717.
Fourme, R., Dhez, P., Benoit, J.-P., Kahn, R., Dubuisson, J.-M., Besson, P. & Frouin, J. (1992). *Rev. Sci. Instrum.* **63**, 982–987.
Hamiaux, C., Pérez, J., Prangé, T., Veesler, S., Riès-Kautt, M. & Vachette, P. (2000). *J. Mol. Biol.* **297**, 697–712.
Hamiaux, C., Prangé, T., Riès-Kautt, M., Ducruix, A., Lafont, S. & Veesler, S. (1999). *Acta Cryst.* D**55**, 103–113.
Howell, P. L. (1995). *Acta Cryst.* D**51**, 654–662.
Lehmann, M. S., Mason, S. A. & McIntyre, G. J. (1985). *Biochemistry*, **24**, 5862–5869.

Lehmann, M. S. & Stansfield, R. F. D. (1989). *Biochemistry*, **28**, 7028–7033.
Levitt, M. & Perutz, M. F. (1988). *J. Mol. Biol.* **201**, 751–754.
Lim, K., Nadarajah, A., Forsythe, E. L. & Pusey, M. L. (1998). *Acta Cryst.* D**54**, 899–904.
Madhusudan, M. & Vijayan, M. (1992). *Protein Eng.* **5**, 399–404.
Muraki, M., Goda, S., Nagahora, H. & Harata, K. (1997). *Protein Sci.* **6**, 473–476.
Navaza, J. (1994). *Acta Cryst.* A**50**, 157–163.
Nicholls, A., Sharp, K. A. & Honig, B. (1991). *Proteins Struct. Funct. Genet.* **11**, 281–296.
Pearson, W. R. (1996). *Methods Enzymol.* **266**, 227–258.
Pike, A. C. W. & Acharya, K. R. (1994). *Protein Sci.* **3**, 706–710.
Ramanadham, M., Sieker, L. C. & Jensen, L. H. (1990). *Acta Cryst.* B**46**, 63–69.
Rao, S. T. & Sundaralingam, M. (1996). *Acta Cryst.* D**52**, 170–175.
Riès-Kautt, M., Broutin, I., Ducruix, A., Shephard, W., Kahn, R., Chayen, N., Blow, D., Paal, K., Littke, W., Lorber, B., Théobald-Dietrich, A. & Giegé, R. J. (1997). *J. Cryst. Growth*, **181**, 79–96.
Riès-Kautt, M. M. & Ducruix, A. (1989). *J. Biol. Chem.* **264**, 745–748.
Riès-Kautt, M. M. & Ducruix, A. (1991). *J. Cryst. Growth*, **110**, 20–25.
Riès-Kautt, M. & Ducruix, A. (1997). *Methods Enzymol.* **276**, 23–59.
Salemme, F. R., Geneiser, L., Finzel, B. C., Hilmer, R. M. & Wendoloski, J. J. (1988). *J. Cryst. Growth*, **90**, 273–282.
Saludjian, P., Prangé, T., Navaza, J., Ménez, R., Guilloteau, J. P., Riès-Kautt, M. M. & Ducruix, A. (1992). *Acta Cryst.* B**48**, 520–531.
Song, H., Inaka, K., Maenaka, K. & Matsushima, M. (1994). *J. Mol. Biol.* **244**, 522–540.
Steinrauf, L. K. (1998). *Acta Cryst.* D**54**, 767–779.
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.
Weaver, L. H., Grütter, M. G. & Matthews, B. W. (1995). *J. Mol. Biol.* **245**, 54–68.
Yamada, T., Matsushima, M., Inaka, K., Ohkubo, T., Uyeda, A., Maeda, T., Titani, K., Sekiguchi, K. & Kikuchi, M. (1993). *J. Biol. Chem.* **268**, 10588–10592.
Yamada, T., Song, H., Inaka, K., Simada, Y., Kikuchi, M. & Matsushima, M. (1995). *J. Biol. Chem.* **270**, 5687–5690.