Acta Crystallographica Section D
Biological Crystallography
ISSN 1399-0047

Weak protein–cationic co-ion interactions addressed by X-ray crystallography and mass spectrometry

Philippe Bénas,[a] Nicolas Auzeil,[b] Laurent Legrand,[c] Franck Brachet,[a] Anne Regazzetti[b] and Madeleine Riès-Kautt[a]*‡

[a]Laboratoire de Cristallographie et RMN Biologiques, Faculté des Sciences Pharmaceutiques et Biologiques, Université Paris Descartes, UMR 8015 CNRS, 4 Avenue de l'Observatoire, 75270 Paris CEDEX 06, France, [b]Laboratoire de Chimie-Toxicologie Analytique et Cellulaire EA 4463, Faculté des Sciences Pharmaceutiques et Biologiques, Université Paris Descartes, 4 Avenue de l'Observatoire, 75270 Paris CEDEX 06, France, and [c]Institut des NanoSciences de Paris (INSP), UMR 7588 CNRS/UPMC (Université Paris 6), 4 Place Jussieu, 75252 Paris CEDEX 05, France

‡ Present address: 3P5, Institut Cochin, Université Paris Descartes, CNRS (UMR 8104), Inserm U1016, 22 Rue Méchain, 75014 Paris, France.

Correspondence e-mail: madeleine.ries@inserm.fr

Received 20 February 2014
Accepted 15 May 2014

The adsorption of Rb^+, Cs^+, Mn^{2+}, Co^{2+} and Yb^{3+} onto the positively charged hen egg-white lysozyme (HEWL) has been investigated by solving 13 X-ray structures of HEWL crystallized with their chlorides and by applying electrospray ionization mass spectrometry (ESI-MS) first to dissolved protein crystals and then to the protein in buffered salt solutions. The number of bound cations follows the order $Cs^+ < Mn^{2+} \simeq Co^{2+} < Yb^{3+}$ at 293 K. HEWL binds less Rb^+ ($q_{tot} = 0.7$) than Cs^+ ($q_{tot} = 3.9$) at 100 K. Crystal flash-cooling drastically increases the binding of Cs^+, but poorly affects that of Yb^{3+}, suggesting different interactions. The addition of glycerol increases the number of bound Yb^{3+} cations, but only slightly increases that of Rb^+. HEWL titrations with the same chlorides, followed by ESI-MS analysis, show that only about 10% of HEWL binds Cs^+ and about 40% binds 1–2 Yb^{3+} cations, while the highest binding reaches 60–70% for protein binding 1–3 Mn^{2+} or Co^{2+} cations. The binding sites identified by X-ray crystallography show that the monovalent Rb^+ and Cs^+ preferentially bind to carbonyl groups, whereas the multivalent Mn^{2+}, Co^{2+} and Yb^{3+} interact with carboxylic groups. This work elucidates the basis of the effect of the Hofmeister cation series on protein solubility.

PDB references: HEWL in 0.5 *M* $MnCl_2$, data collected at room temperature, 4neb; HEWL in 1.1 *M* $MnCl_2$, data collected at room temperature, 4nfv; HEWL in 1.9 *M* CsCl, data collected at room temperature, 4ng1; dialyzed HEWL batch-crystallized in 1.9 *M* CsCl, data collected at 100 K, 4ng8; previously deionized HEWL crystallized in 1.0 *M* RbCl, data collected at 125 K, 4ngi; dialyzed HEWL batch-crystallized in 1.0 *M* RbCl, data collected at 100 K, 4ngj; previously deionized HEWL batch-crystallized in 0.2 *M* $CoCl_2$, 4ngk; previously deionized HEWL batch-crystallized in 1.0 *M* $CoCl_2$, 4ngo; previously deionized HEWL batch-crystallized in 0.6 *M* $CoCl_2$, 4ngl; previously deionized HEWL batch-crystallized in 0.5 *M* $YbCl_3$, 4ngv; dialyzed HEWL batch-crystallized in 0.5 *M* $YbCl_3$, data collected at 100 K, 4ngw; dialyzed HEWL batch-crystallized in 0.75 *M* $YbCl_3$, data collected at 100 K, 4ngy; previously deionized HEWL crystallized in 0.48 *M* $YbCl_3$/30%(*v*/*v*) glycerol, data collected at 125 K, 4ngz

1. Introduction

Studies in physical chemistry have shown that ions affect a large number of the properties of salt solutions (Collins & Washabaugh, 1985) according to the Hofmeister series (Hofmeister, 1888; Kunz *et al.*, 2004), including protein solubility, macromolecular conformations (Von Hippel & Schleich, 1969) and more generally protein–protein interactions in solution. Ion-specific effects have been attributed to their ability to structure water molecules, giving rise to the classification of ions as chaotropes and kosmotropes. The former are those that disrupt the water network, or 'structure breakers', such as K^+, Cs^+, NH_4^+, SCN^- and I^-, whereas the latter, such as SO_4^{2-}, HPO_4^{2-}, Mg^{2+} and Li^+, are called 'structure makers' and are able to reorder water molecules around themselves.

As far as protein solubility and crystallization are concerned, the question was whether the salts act indirectly by structuring or perturbing the protein hydration shell or directly as ion pairs with ionic sites on the protein surface. Indeed, for a long time the salt effect was attributed to perturbation of the protein hydration shell (Tanford, 1961; Arakawa & Timasheff, 1984) by salts competing for their own hydration, before direct interactions of bulk solution ions with charged groups at the surface of the protein were considered

to form ion pairs (Riès-Kautt & Ducruix, 1997; Collins, 2006; Zhang & Cremer, 2006). Ion pairing further relies on a concept according to which ions are more likely to bind to each other according to the law of matching water affinity (Collins & Washabaugh, 1985; Washabaugh & Collins, 1986) or the hard/soft concept (Pearson, 1987). The identification of the nature of these pairs gives access to new tools that are able to be generalized for protein crystallization. Reviews of the effects of ionic strength on ion–protein interactions have underlined the importance of their chemistry (Collins, 2012).

We have already revisited the effect of anions in previous solubility studies (Riès-Kautt & Ducruix, 1989) and crystallographic analyses (Vaney *et al.*, 2001) of the positively charged hen egg-white lysozyme (HEWL). This led us to demonstrate that anions directly interact as counter-ions with the protein surface of HEWL and are responsible for the observed solubility decrease when increasing the salt concentration and mostly when changing the nature of the anion (Riès-Kautt & Ducruix, 1997) from acetate to thiocyanate. This observation could be confirmed for the crystallization of other positively charged proteins, such as BPTI (Hamiaux *et al.*, 1999) and toxins (Ménez & Ducruix, 1990, 1993; Saludjian *et al.*, 1992).

We have shown that the adsorption of counter-ions decreases the effective protein net charge and consequently the protein solubility in aqueous salt solutions (Riès-Kautt & Ducruix, 1991). The effectiveness of the investigated anions appears to follow the reverse of Hofmeister's series, which is linked to the better affinity of soft ions to interact with positively charged sites of the protein. Conversely, we suggested that the adsorption of co-ions, if it occurs, should increase the apparent net charge of the protein and hence its solubility. We effectively measured such a behaviour for the positively charged HEWL crystallized in the presence of several monovalent, divalent and trivalent cation chlorides and observed a solubility increase when lysozyme was crystallized with multivalent cations (Bénas *et al.*, 2002). Results published for HEWL crystallized with $NiCl_2$ (Li *et al.*, 2005) also showed increased solubility of this protein above 0.5 *M*. In the literature, however, these effects have been better documented for anions than for cations. Among the latter, monovalent cations have been more investigated than multivalent cations, as in Hofmeister's original work.

It must be emphasized that this work does not address the binding of cations with high association constants that have a structural or catalytic function in proteins, since lysozyme does not require any cations for its catalytic activity. We deal here with cations interacting as co-ions with solvent-exposed sites of protein surfaces in general. Such weak interactions are more difficult to characterize than are strong biochemical associations. We therefore undertook a further investigation of cation adsorption using two very different but complementary biophysical approaches: X-ray crystallography and mass spectrometry. We aimed to define a possible stoichiometry for cations bound to HEWL by these two approaches, possibly estimating the relative abundance of the free protein and the HEWL–cation complexes, and eventually to describe their binding sites.

Crystallographic structure determination allows ions or molecules adsorbed onto protein molecules to be observed depending on their occupancy factor q. The latter is the fraction of the asymmetric units in the irradiated crystal volume in which they are present at a given position and averaged over the data-collection time. In this approach the thermal agitation factor B is another important and correlated parameter in assigning solvent molecules. In the present study we also have performed careful data processing and analysis as well as precise calculations (Weiss *et al.*, 2002) in order to overcome the difficulty in discriminating between ions and water molecules. Furthermore, the anomalous signal contribution of the studied cations has been taken into account as described by Dauter and Dauter for anions bound to HEWL (Dauter & Dauter, 1999). The best possible electron-density assignments for the solvent molecules were finally obtained in conjunction with a careful analysis of the chemical environment.

Electrospray ionization mass spectrometry (ESI-MS) has also proven to be an approach for revealing noncovalent binding, including metal ions binding to proteins as reviewed by Potier *et al.* (2005). As early as 1994, we observed the adsorption of sulfate and phosphate onto lysozyme when studying the effect of anions on protein solubility (Riès-Kautt *et al.*, 1994). This has also been applied to demonstrate that HEWL can bind up to six Zn^{2+} or eight Cu^{2+} cations in water at pH 6.9 (Moreau *et al.*, 1995).

ESI-MS can identify ions that have too low an occupancy or too high a temperature factor and hence are hidden in the bulk solvent in X-ray structures. However, it cannot provide information either about their location on the protein or about ions trapped at protein interfaces owing to crystal packing. Nevertheless, cation interactions occurring at high ionic strength during HEWL crystallization cannot be directly investigated by ESI-MS analysis, since the high salt content hampers the ionization process. We have therefore tested two different ESI-MS analysis approaches. In the first series, crystals from the same crystallization batches as used for the X-ray and solubility studies were dissolved in water–MeOH [85:15(*v*:*v*)]. In the second series, a diluted HEWL solution was titrated with increasing concentrations of Cs^{+}, Mn^{2+}, Co^{2+} and Yb^{3+} chlorides.

The two combined ESI-MS approaches are compared with the crystallographic results to provide insights into the stoichiometry of HEWL complexes with cations as well as the relative efficiency of cation binding to HEWL.

2. Materials and methods

2.1. Reagents and solution preparations

HEWL (Sigma catalogue No. L6876, batch 73H7045) was desalted to its isoionic state, *i.e.* with the only possible counter-ions being H^{+} and OH^{-}, as previously detailed in Retailleau, Ducruix *et al.* (1997). The isoionic protein solution was then acidified to pH 4.5 by the addition of about ten molar equivalents of HCl, allowing all cations to be tested as chloride

salts. The purity of the HEWL was checked by SDS–PAGE followed by enhanced silver staining (Thomas *et al.*, 1996), showing a single band, and by ESI-MS (M_r = 14 305 ± 2). The crystals used for data collection at 100 K were grown using HEWL (Fluka Analytical 62971-50G-F, batch 0001356468) without further purification.

Commercial deionized and thrice-distilled water (Meram, France) was used to prepare the solutions. All chloride salts were of ACS grade with purity greater than 99%. CsCl was purchased from Bethesda Research Laboratory (Maryland, USA), RbCl and $YbCl_3.6H_2O$ from Aldrich, $MnCl_2.4H_2O$ and $CoCl_2.6H_2O$ from Merck and ammonium acetate from Fluka (Saint-Quentin Fallavier, France). AP Normapur glycerol (VWR) was used at a purity of ≥99.5%.

All salt stock solutions were filtered through 0.22 µm filter systems (Polylabo 22676, Nalgene 90621) and their concentrations were checked by measurement of the refractive index. Glycerol stock solutions were filtered through 0.45 µm Polylabo–Nalgene filter units.

Lysozyme was filtered through a 0.22 µm Millipore filter (Millex-GV4). The protein concentration was determined by UV absorption at 280 nm, using $\varepsilon_{0.1\%}$ = 2.66 l g^{-1} cm^{-1}, and was averaged from two or three measurements falling within 10% accuracy.

HEWL (2 µ*M*) solutions for the ESI-MS titration experiments were prepared in 20 m*M* ammonium acetate buffer adjusted to pH 4.5 with diluted HCl. Cation solutions (20–600 µ*M* final salt concentration) were mixed with HEWL and stored at 20°C for 30 min before analysis.

2.2. HEWL crystallization in cation chloride solutions

Crystallization conditions are summarized in Table 1. The pH of crystallizing drops was checked to be 4.5 ± 0.1 using a micro-electrode. No buffer was added in order to limit investigations to the effect of chlorides and the given cations under study. All crystallizations were performed at 18 ± 0.1°C in a thermally regulated incubator.

All batch crystallizations were set up using the microbatch method in order to ensure the actual salt and protein concentrations. For convenience of crystal mounting in capillaries, 10 µl microbatches were set up as sitting drops (Emerald Bio crystallization plates) over a well containing the same salt-solution concentration as that within the crystallization drops.

Crystal growth in the presence of glycerol was performed in 9 + 9 µl sitting drops by vapour diffusion. The salt and protein concentrations indicated in Table 1 are those expected at equilibrium, *i.e.* twice the known initial concentration.

2.3. Crystallography

2.3.1. X-ray data collection and processing. One single crystal was used for each data collection and details can be found in Table 1. All but the X-ray data collected at 100 K were processed using programs from the *HKL* package v.1.96.2. X-ray diffraction patterns were indexed and integrated using *DENZO* and the data were further reduced using *SCALEPACK* (Otwinowski & Minor, 1997). The data collected at 100 K were indexed, integrated and reduced using the *XDS* package (Kabsch, 2010*a*,*b*). Friedel's law was assumed to be false in the data-reduction step.

The reduced reflection files were converted to MTZ format using *SCALEPACK2MTZ* from the *CCP*4 suite (Winn *et al.*, 2011) or using *XDSCONV* from the *XDS* package and were then brought to an absolute scale using the *CCP*4 program *TRUNCATE* (French & Wilson, 1978).

As shown in Table 1, the overall statistics are satisfactory. The value of 86% for the completeness of the 4ngl data set is explained by a data collection that was limited to 90° while the crystal orientation would have required a total rotation angle of 160°.

2.3.2. Crystal structure refinement and manual rebuilding. The crystal structures were solved by straightforward molecular replacement using the polypeptide chain of PDB entry 193l (Vaney *et al.*, 1996) as an initial model. Molecular replacement for crystals at room temperature or at 125 K was conducted with *AMoRe* (Navaza, 1994) and the structures of crystals at 100 K were solved by *Phaser* (McCoy *et al.*, 2007). For all of the anomalous data sets, a total of seven iterative cycles of structure refinement using *CNS* (*Crystallography and NMR System*) v.1.1 (Brünger *et al.*, 1998) and manual rebuilding in either *O* (v.7.0.1 or 8.2; Jones *et al.*, 1991) or *Coot* (Emsley *et al.*, 2010) were performed. *CNS* refinements were based on structure-factor amplitudes (as a cross-validated data set by free *R* flags) and maximum-likelihood target functions. The bulk-solvent *B* factor was assumed to be anisotropic in the refinement cycles. Briefly, each *CNS* cycle consisted of the following: file generation (atomic coordinates and structure file) and optimization of the X-ray *versus* geometry weights followed by simulated-annealing molecular dynamics on torsion angles and then individual *B*-factor refinement. Alternate conformation occupancies of the protein were then refined along with the occupancies of the solvent molecules, namely protein-bound ions and water molecules. Electron-density maps (both standard and σ_A-weighted; Read, 1986) were generated using the calculated phases from the output model.

A final refinement cycle was performed with up-to-date refinement software (*REFMAC* v.5.8) from the *CCP*4 suite using the efficient refinement and automated rebuilding programs available on the *PDB_REDO* web server (Joosten *et al.*, 2011).

The usual statistics of crystallographic refinement and geometrical parameters were monitored (see Table 2).

2.3.3. Electron-density inspection and atom discrimination. Water molecules, cations and chlorides were placed in the electron-density maps based on several criteria. The first was the size of the electron-density blobs to be filled in, for instance if the identification of well defined metal sites is easy according to their large number of electrons. The situation becomes less evident at low occupancies and with overlapping binding sites, so that confusion with chlorides or water molecules may possibly occur. Peak heights have therefore been

Table 1
Crystallization conditions and data-collection statistics.

Values in parentheses are for the outermost shell.

Salt	RbCl		CsCl		$MnCl_2$	
PDB code	4ngj	4ngi	4ng1	4ng8	4neb	4nfv
Crystallization method and conditions						
Method	Batch	VD†	Batch	Batch	Batch	Batch
[Salt] (*M*)	1.00	1.00‡	1.90	1.90	0.50	1.10
[HEWL] (mg ml^{-1})	25	124‡	17.5	25	64.25	64.25
[Glycerol] [%(*v*/*v*)]	0	30‡	0	0	0	0
Data collection						
Temperature (K)	100	125	293	100	293	293
Wavelength (Å)	0.9755	1.5418	1.3880	0.9755	0.9464	1.3880
Unit-cell parameters§						
a (Å)	79.02	79.04	79.44	78.94	79.19	79.61
c (Å)	37.22	36.79	38.12	37.07	37.72	37.89
Mosaicity (°)	0.07	0.64	0.13	0.05	0.14	0.13
Resolution limits (Å)	39.50–1.10 (1.17–1.10)	19.76–1.70 (1.76–1.70)	34.36–1.82 (1.89–1.82)	35.30–1.09 (1.16–1.09)	27.31–1.48 (1.53–1.48)	35.00–1.63 (1.69–1.63)
Total No. of reflections¶	395459	75465	73833	373306	166934	215410
No. of unique reflections¶	90874	23492	20871	92973	38160	28993
Multiplicity¶	4.4 (4.1)	3.2 (1.6)	3.5 (3.4)	4.0 (3.7)	4.4 (4.2)	7.4 (6.9)
$\langle I/\sigma(I)\rangle$	22.9 (4.4)	15.1 (3.1)	12.7 (3.9)	20.9 (4.6)	26.3 (3.3)	23.8 (4.6)
Completeness (%)	99.6 (98.9)	96.8 (98.3)	99.9 (100.0)	99.8 (99.7)	99.9 (100.0)	100.0 (100.0)
$R_{\rm merge}$†† (%)	4.1 (32.3)	8.8 (22.7)	10.9 (32.0)	4.4 (33.3)	4.5 (40.4)	7.5 (37.0)

Salt	$CoCl_2$			$YbCl_3$			
PDB code	4ngk	4ngl	4ngo	4ngv	4ngw	4ngy	4ngz
Crystallization method and conditions							
Method	Batch	Batch	Batch	Batch	Batch	Batch	VD†
[Salt] (*M*)	0.20	0.60	1.00	0.50	0.50	0.75	0.48‡
[HEWL] (mg ml^{-1})	64.25	64.25	32.75	64.25	75	100	124‡
[Glycerol] [%(*v*/*v*)]	0	0	0	0	0	0	30‡
Data collection							
Temperature (K)	293	293	293	293	100	100	125
Wavelength (Å)	0.9464	0.9464	0.9464	0.9464	1.3851	1.3851	1.5418
Unit-cell parameters§							
a (Å)	79.29	79.28	79.30	78.91	78.67	78.67	78.59
c (Å)	37.73	37.53	37.54	37.70	37.13	37.15	37.02
Mosaicity (°)	0.15	0.16	0.16	0.15	0.10	0.08	0.45
Resolution limits (Å)	27.33–1.50 (1.55–1.50)	28.03–1.52 (1.57–1.52)	27.26–1.58 (1.60–1.58)	27.26–1.64 (1.70–1.64)	35.18–1.37 (1.45–1.37)	35.18–1.35 (1.43–1.35)	19.65–1.70 (1.76–1.70)
Total No. of reflections¶	168114	147202	132745	76367	156981	202503	124625
No. of unique reflections¶	36300	30500	31351	27665	46458	48641	24026
Multiplicity¶	4.6 (4.2)	4.8 (4.2)	4.2 (4.2)	2.8 (2.8)	3.4 (3.1)	4.2 (3.8)	5.2 (2.5)
$\langle I/\sigma(I)\rangle$	36.6 (4.1)	28.7 (3.7)	26.2 (4.7)	13.0 (3.4)	13.0 (3.8)	13.6 (3.2)	23.3 (3.8)
Completeness (%)	98.5 (97.4)	86.5 (90.7)	99.9 (100.0)	99.8 (100.0)	99.6 (98.3)	99.7 (99.1)	99.7 (97.7)
$R_{\rm merge}$†† (%)	3.5 (38.6)	4.2 (35.8)	4.9 (30.9)	6.7 (29.8)	7.4 (34.1)	8.2 (36.9)	6.7 (20.2)

† Vapour diffusion with a 30% final concentration of glycerol. ‡ Concentration expected at equilibrium. § All crystals belonged to the tetragonal space group $P4_32_12$ with one molecule per asymmetric unit. ¶ Considering all the anomalous pairs F^+ and F^-. †† $R_{\rm merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$.

compared using the sulfur peaks of cysteine as internal standards following the method proposed by Weiss *et al.* (2002). Theoretical electron-density maps have similarly been computed including either the investigated ion or a water molecule. Further occupancy and temperature-factor refinements using one or the other of the models have nevertheless shown that the method remains indicative owing to the strong correlation between occupancies and temperature factors. The proper identification was then based on either the presence or absence of an anomalous signal in the corresponding Bijvoet difference Fourier maps on one hand and the chemical environment of the electron-density blob to be attributed on the other. The divalent and trivalent cations studied in this work have rather short binding distances and hence are quite easily assigned in electron-density maps, although use of the anomalous signal was required for the proper assignment of alternate positions at site A1 for Mn^{2+} and Yb^{3+}. In contrast, Cs^+ ions ($Z = 55$) have binding distances that cover the range of hydrogen bonds up to 3.85 Å, but a low occupancy correlated with a high thermal agitation can blur proper identification. Assignments were therefore greatly helped by the electron-density peaks found in the Bijvoet difference Fourier maps, despite the large number of Cs^+ electrons. Rb^+ ions have binding distance lengths comparable to hydrogen bonds and interact with atoms bearing lone-pair electrons as hydrogen would. They are thus very hard to detect assuming Friedel's law to be true, *i.e.* neglecting their anomalous contribution to the observed structure factors.

Table 2
Refinement statistics.

Values in parentheses are for the outermost shell.

Salt	RbCl		CsCl		$MnCl_2$	
PDB code	4ngj	4ngi†	4ng1	4ng8	4neb	4nfv
Resolution range (Å)	55.88–1.10 (1.13–1.10)	55.89–1.70 (1.74–1.70)	56.17–1.82 (1.87–1.82)	55.82–1.09 (1.12–1.09)	55.99–1.48 (1.52–1.48)	56.29–1.63 (1.67–1.63)
No. of reflections in working set	43312 (3103)	11728 (778)	10471 (757)	44207 (3168)	18522 (1334)	14176 (1007)
R_{free} test-set size (%)	9.7 (9.5)	8.2 (7.6)	8.1 (7.0)	9.8 (9.6)	9.8 (10.4)	9.7 (10.4)
R_{cryst}‡ (%)	14.3 (20.6)	17.6 (25.1)	15.8 (20.9)	15.2 (21.3)	12.2 (14.4)	14.8 (19.8)
R_{free}§ (%)	17.3 (21.5)	21.5 (30.2)	19.0 (22.2)	17.6 (23.9)	15.1 (21.7)	17.9 (19.7)
Mean B (Å^2)	15.1	21.8	20.0	16.2	22.3	22.7
k_{sol} (e$^-$ Å^{-3})	0.31	0.39	0.34	0.34	0.34	0.37
B_{sol} (Å^2)	45.5	41.1	35.7	48.6	35.7	38.6
No. of non-H atoms						
Protein	1096	1080	1036	1064	1076	1093
Cations	1	1	1	9	2	4
Chloride	2	1	1	2	1	1
Water	165	134	83	155	62	79
R.m.s. deviations						
Bonds (Å)	0.014	0.011	0.011	0.017	0.011	0.020
Angles (°)	1.68	1.36	1.35	1.54	1.35	1.78
Ramachandran plot analysis, residues in (%)						
Most favoured regions	96.97	98.06	97.46	96.26	98.11	98.08
Additionally allowed regions	3.03	1.94	2.54	3.74	1.89	1.92
Disallowed regions	0.00	0.00	0.00	0.00	0.00	0.00

Salt	$CoCl_2$			$YbCl_3$			
PDB code	4ngk	4ngl	4ngo	4ngv	4ngw	4ngy	4ngz†
Resolution range (Å)	56.06–1.50 (1.54–1.50)	56.06–1.52 (1.56–1.52)	56.07–1.58 (1.62–1.58)	55.79–1.64 (1.68–1.64)	55.62–1.37 (1.41–1.37)	55.63–1.35 (1.39–1.35)	55.57–1.70 (1.74–1.70)
No. of reflections in working set	17945 (1282)	15129 (1160)	15639 (1152)	13870 (999)	22474 (1600)	23404 (1681)	11957 (849)
R_{free} test-set size (%)	7.7 (8.1)	7.6 (6.2)	7.7 (7.3)	7.9 (8.0)	9.7 (9.2)	9.7 (10.6)	9.4 (9.8)
R_{cryst}‡ (%)	12.2 (15.1)	15.0 (21.8)	14.9 (16.6)	17.1 (20.5)	16.3 (21.0)	19.1 (27.0)	17.8 (26.0)
R_{free}§ (%)	15.9 (22.7)	16.3 (25.7)	7.6 (18.3)	18.6 (22.7)	18.8 (27.8)	21.6 (26.1)	20.5 (32.1)
Mean B (Å^2)	22.9	23.2	22.9	21.7	19.3	21.3	21.6
k_{sol} (e$^-$ Å^{-3})	0.34	0.33	0.35	0.34	0.33	0.35	0.40
B_{sol} (Å^2)	36.4	32.2	38.7	35.0	37.5	43.5	39.7
No. of non-H atoms							
Protein	1068	1074	1054	1065	1116	1146	1029
Cations	1	1	1	4	3	3	3
Chloride	1	1	1	1	2	2	3
Water	58	61	70	77	130	145	129
R.m.s. deviations							
Bonds (Å)	0.012	0.014	0.011	0.006	0.012	0.019	0.011
Angles (°)	1.34	1.52	1.39	1.05	1.31	1.70	1.30
Ramachandran plot analysis, residues in (%)							
Most favoured regions	98.13	98.02	97.27	97.20	94.57	97.83	97.54
Additionally allowed regions	1.87	1.98	2.73	2.80	5.43	2.17	2.46
Disallowed regions	0.00	0.00	0.00	0.00	0.00	0.00	0.00

† In the presence of 30% glycerol. ‡ $R_{cryst} = \sum_{hkl} ||F_{obs}| - |F_{calc}|| / \sum_{hkl} |F_{obs}|$, where $F_{calc}(hkl)$ is the calculated structure-factor amplitude and $F_{obs}(hkl)$ is the observed structure-factor amplitude. § R_{free} is calculated as R_{cryst}, where the F_{obs} are taken from a test set that were excluded from the refinement (the size of the test set given in the table).

Depending on the incident wavelength, the anomalous dispersive contribution and the corresponding f'' can be as low as 0.46% and 0.22 e$^-$ for S atoms and 0.19% and 0.28 e$^-$ for chloride ions, respectively. Nevertheless, in all crystal structures reported here the anomalous signal was clearly visible in the electron-density maps for each S and Cl atom, and in particular for the disulfide bridges and the well characterized Cl^- bound to the OH group of Tyr23. As for cations, the lowest dispersive contribution is always higher than 1% and the lowest expected f'' (1.23 e$^-$) found for the $MnCl_2$/HEWL crystal collected at a radiation wavelength of 0.94 Å is still more than five times larger than the lowest f'' expected for the S atoms. After data processing to include the anomalous contributions we were hence able to unambiguously assign the ions bound to the protein by comparing the observed anomalous signal for the S atoms and the Tyr23-bound Cl^- with those at other locations and by evaluating their chemical environment.

Fig. 2 was prepared using *PyMOL* (Schrödinger) and Figs. 8–11 using *Discovery* 4 (Discovery Studio Modeling Environment, Release 4.0, Accelrys Software Inc., San Diego, USA).

2.4. Mass spectrometry

ESI-MS analyses were performed with an ion-trap mass spectrometer working in positive-ion mode (Thermo Fisher Scientific, France). The *Xcalibur* (Thermo Fisher Scientific) software was used for data acquisition and analysis. The voltage of the spray needle was set to 4.0 kV, that of the electrospray capillary was 46 V and that of the tube lens offset was 46 V, and the capillary temperature was 200°C. The spray was stabilized with nitrogen as the sheath gas (60 arbitrary units) and auxiliary gas (20 arbitrary units). These optimized parameters gave the best sensitivity without disrupting noncovalent interactions in the gas phase. The syringe-pump flow rate was set to 5 µl min^{-1}. The acquisition time was fixed at 1 min. Three micro-scans were collected to produce a unique scan and the duration of the accumulation of ions in the analyser was 50 ms. Acquisition was performed in full scan mode between m/z 200 and 2000 and mass spectra were acquired and displayed in a profile data type. After each acquisition, a Surveyor HPLC pump (Thermo Fisher Scientific) was used to wash the acquisition system with water at a flow rate of 200 µl min^{-1} for 20 min. The *BioWorks Browser* software (v.3.0; Thermo Fisher Scientific) was used to calculate an average mass by deconvolution for free HEWL and for HEWL–cation complexes.

Crystals grown in the same series of experiments as those used for X-ray diffraction were dissolved in 200 µl H_2O prior to analysis.

The second series of experiments were run with HEWL solutions in ammonium acetate, *i.e.* not from dissolved crystals. After each run containing a cation chloride salt, a spectrum was acquired for a 2 µ*M* HEWL solution made with 20 m*M* ammonium acetate buffer as the sole component in order to verify the complete disappearance of HEWL–ion adducts. Before the next acquisition the system was finally washed with water at a flow rate of 200 µl min^{-1} for a further 20 min. For each solution, mass spectra were acquired three times.

Several assumptions were used and have to be reported to specify the validity range for quantitative analysis in titration experiments. It is assumed that the signal response for each individual species is proportional to its concentration in the gas phase and, by inference, in solution (Whittal *et al.*, 2000). It is also anticipated that free and cation-bound proteins give the same signal response. It seems reasonable to consider that cation binding does not drastically affect the protein conformation since the shape and the charge-state distribution for free and cation-bound HEWL are identical. Thus, it was assumed that both free HEWL and HEWL in complex with cations had equivalent desorption and ionization behaviours. Consequently, the various HEWL species recorded differ only by mass increments corresponding to the binding of ions. Under this assumption, the relative ion abundances recorded for the free and complexed HEWL are considered to be a direct measurement of their relative concentrations in solution.

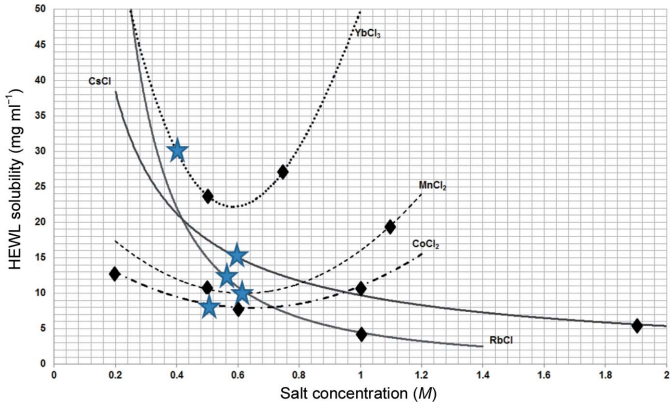

Figure 1
Indication of the crystals used for X-ray diffraction (diamonds) and for ESI-MS (stars) experiments on the HEWL solubility curves (pH 4.5 and 25°C).

3. Results and discussion

The HEWL crystals used for the X-ray diffraction and ESI-MS experiments (diamonds and stars, respectively, in Fig. 1) were selected from crystallization experiments aimed at determining the solubility curves (Bénas *et al.*, 2002).

For the X-ray crystal analysis, a single crystal with good diffraction properties was chosen for each type of cation from batches corresponding to the decreasing part of the HEWL solubility curves. Further HEWL structures have been determined at higher $MnCl_2$, $CoCl_2$ and $YbCl_3$ concentrations where the solubility increases in order to find a possible explanation for the particular solubility behaviour observed with the multivalent cations.

All crystals belonged to the tetragonal space group $P4_32_12$ (see Table 1 for data-collection statistics) and the protein structures are almost identical: the average root-mean-square deviation for the coordinates (⟨r.s.m.d.⟩) is 0.19 ± 0.09 Å considering the protein main-chain atoms and 0.63 ± 0.14 Å considering all protein atoms.

The crystals for the ESI-MS measurements were chosen at the highest salt concentration compatible with the technique and as close as possible to the conditions of the crystals for X-ray diffraction.

3.1. Types and numbers of ions identified by X-ray crystallography

Table 3 summarizes the nature and the numbers of ions identified by X-ray crystallography along with the corresponding crystallization and data-collection conditions. The total number of ions, q_{tot}, is the sum of the individual q values over all sites where the ion is observed in a given structure. The details of the binding sites and the chemical environment are further discussed in §3.3.

3.1.1. Data collected at room temperature. The number of bound cations in the structures corresponding to the decrease in solubility with increasing salt concentration follows the order $Cs^+ < Co^{2+} \simeq Mn^{2+} < Yb^{3+}$. Only one binding site (site E) is observed for Cs^+. It binds with an occupancy of 0.23, although the crystal was grown in 1.9 *M* CsCl. The multivalent

Table 3
Numbers of ions identified in the HEWL structures.

Salt	RbCl		CsCl		$MnCl_2$		$CoCl_2$			$YbCl_3$			
PDB code	4ngj	4ngi†	4ng1	4ng8	4neb	4nfv	4ngk	4ngl	4ngo	4ngv	4ngw	4ngy	4ngz†
[Salt] (*M*)	1.00	1.00	1.90	1.90	0.50	1.10	0.20	0.60	1.00	0.50	0.50	0.75	0.48
Data-collection temperature (K)	100	125	293	100	293	293	293	293	293	293	100	100	125
All cations q_{total}	**0.72**	**0.76**	**0.23**	**3.90**	**0.80**	**1.33**	**0.54**	**0.83**	**0.80**	**1.38**	**1.51**	**1.89**	**1.92**
Sites A													
A1 (Asp52)				0.28	0.31	0.30	0.54	0.83	0.80	0.67	0.55	0.75	0.79
						0.34							
						0.24				0.22			
A2 (Glu35)					0.49	0.45				0.10			
A3 (Asn46)				0.44									
				0.35									
Sites B													
B1(Asp101 $O^{\delta2}$)				0.33									
				0.32									0.67
B2 (Asp101 $O^{\delta1}$/O)				0.41									
B3 (Trp62)				0.44									
Site C (C-terminus)										0.39	0.62	0.72	0.46
Site D (Asp87)											0.34	0.42	
Site E (Asn77)	0.72	0.76	0.23	0.73									
Site F (Asn44)				0.60									
All Cl^- q_{total}	**1.98**	**1.00**	**0.92**	**1.84**	**0.98**	**1.00**	**1.00**	**1.00**	**1.00**	**0.99**	**1.95**	**2.00**	**2.87**
Site G (Tyr23)	1.00	1.00	0.92	0.98	0.98	1.00	1.00	1.00	1.00	0.99	1.00	1.00	0.99
Site H (Ser24)	0.98			0.86							0.95	1.00	0.93
Site I (Lys33)													0.95

† With a final concentration of 30% glycerol.

cations Co^{2+} (q_{tot} = 0.54), Mn^{2+} (q_{tot} = 0.83) and Yb^{3+} (q = 0.99) bind to site A, while an additional trivalent cation Yb^{3+} is found at site C with an occupancy of 0.39.

Checking the effect of higher salt concentrations, where the solubility rises, shows that more Mn^{2+} and Co^{2+} cations are bound: q_{tot} = 0.8 for Co^{2+} and 1.33 for Mn^{2+}. These cations are bound at the same sites as those seen at lower concentration, but have a higher q, and even populate alternate positions for Mn^{2+}.

Only one Cl^- is observed in all of these structures in binding site G, with an occupancy ranging from 0.92 to 1.00.

3.1.2. Data collected at 100 K. Data have been recorded at 100 K in order to compare the relative binding of the two monovalent cations with each other, as well as to evaluate the effect of temperature on the binding of Cs^+ and Yb^{3+}.

Site E is equally occupied by Rb^+ or the larger Cs^+ cation (Figs. 2*a* and 2*b*), with q = 0.72 and 0.73, respectively. The significant difference comes from the q_{tot} of the two cations, since Rb^+ is only present in site E, whereas six further Cs^+ cations are observed on the protein surface at site A1 (q = 0.28), site A3 (q = 0.79), site B1 (q = 0.65), site B2 (q = 0.41), site B3 (q = 0.44) and a new site, F, with q = 0.60. The total number of Cs^+ cations at 100 K is about 17 times that at room temperature. The crystals were both grown at room temperature, but the flash-cooling for data collection at 100 K seems to 'freeze' labile Cs^+ cations, as illustrated by the increase in q at site E from 0.23 to 0.73.

The number of bound cations follows the order $Rb^+ << Cs^+$. This corroborates the lower solubility curve of HEWL crystallized in RbCl than in CsCl (Fig. 1) at 25°C.

In contrast to Cs^+, the total number of Yb^{3+} cations at 100 K is nearly identical to that at room temperature for a 0.5 *M* salt concentration. This strongly suggests a different type of binding for the two types of cations, as highlighted by the structural analysis at the atomic level in §3.3.

There is a unique Yb^{3+} position at site A1 at 100 K instead of the two alternate positions observed at room temperature, and a new site (site D) becomes occupied by an Yb^{3+} cation (q = 0.34).

Comparing now the binding sites in the crystal grown at 0.5 *M* with that grown at 0.75 *M*, q_{tot} for Yb^{3+} increases from 1.51 to 1.89, whereas the binding of Cl^- is nearly identical (q_{tot} varying from 1.95 to 2.00), without any new sites for either the cation or the anion.

In these four structures at 100 K a second Cl^- anion is observed in site H close to Ser24, with q ranging from 0.86 to 1.00.

3.1.3. Effect of glycerol. A last set of structures were solved from crystals grown in the presence of 30%(*v*/*v*) glycerol, from which data were recorded under cryo-conditions, addressing the effect on ion binding of the presence of this often-used additive. One would expect that electrostatic interactions would be reinforced in the presence of glycerol, since the dielectric constant of the solvent is lowered.

For Rb^+ there is no significant difference in q_{tot} or the type or the number of binding sites in the presence and absence of glycerol. In contrast, q_{tot} for Yb^{3+} increases from 1.51 to 1.92 with 30% glycerol and the binding sites change: the occupancy at A1 is higher than in the absence of glycerol and a new Yb^{3+} site appears at B1 (Fig. 2*c*), while A2 and A3 are no longer occupied. The binding at site C is nearly unchanged.

For Cl^- anions, a third Cl^- site (at site I) close to Lys33 (q = 0.95) is observed in the presence of glycerol and $YbCl_3$.

3.2. Types and number of cations identified by ESI-MS

Three series of ESI-MS investigations were performed. Firstly, crystals from the same set of X-ray diffraction experiments were dissolved in parallel to the X-ray structure determinations in order to search for protein m/z increased by bound cations. Secondly, HEWL solutions were prepared with increasing cation concentrations in order to obtain further information on the various HEWL–cation complexes. Finally, an attempt at quantification gives a relative efficiency of the cations for interaction with the protein, although limited by the technical feasibility.

The binding of cations modifies both the mass m and the charge z of the protein complex. For each charge state, additional peaks appear incremented by the mass of the adsorbed

cations. A given charge state z arises from the sum of the n cations and z' protons H^+. Similarly, the measured mass m/z includes the mass of HEWL, the n adsorbed cations and z' protons.

3.2.1. Dissolved lysozyme–cation crystals. Since a high salt concentration hampers ionization and a satisfactory signal-to-noise ratio in MS, the HEWL crystals used were those grown at the highest ionic strength compatible with an interpretable MS spectra (indicated by stars in Fig. 1) and dissolved in 200 µl pure water prior to injection. All spectra exhibit +8 to +13 charge states for HEWL.

Monovalent Rb^+ and Cs^+. In the MS spectra of HEWL crystals grown in the presence of RbCl or CsCl, no additional ion peaks were observed up to 0.6 *M* salt, which is the limit of salt concentration for which the signal-to-noise remains interpretable.

Multivalent cations: Mn^{2+}, Co^{2+} and Yb^{3+}. In the spectra of HEWL crystals grown in 0.6 *M* $MnCl_2$, 0.5 *M* $CoCl_2$ or 0.4 *M* $YbCl_3$ (top, middle and bottom, respectively, in Fig. 3), the most populated charge state z is +10. This indicates a similar ionization of the three species and allows their comparison. The highest peak intensities correspond to the addition of 1–2 Mn^{2+} cations (n = 1–2), but to free lysozyme (n = 0) in the presence of the two other salts. Further binding of up to five adsorbed Mn^{2+} cations (n = 5), two Co^{2+} cations and three Yb^{3+} cations is observed.

These results are in agreement with the X-ray structures revealing that two Mn^{2+} cations and one Co^{2+} cation bind to sites A1 and A2. The additional cations detected by MS might be owing to nonspecific adsorption occurring in the MS source or to interactions that cannot be highlighted by crystallography, as this latter technique shows averaged electron density over the unit cells and the time of data collection. They might also be owing to the presence of MeOH added to enhance the signal and which reinforced electrostatic interactions as observed for glycerol by crystallography. This motivated us to address the binding of cations in solution as described below.

3.2.2. Lysozyme–cation complexes in ammonium acetate buffer solutions. The results obtained by ESI-MS analyses for dissolved crystals have been supplemented by analyses of HEWL in solution to which increasing amounts of salts have been added at relatively low concentrations in order to address a possible mechanism of adsorption.

When HEWL (2 µ*M*) is in an aqueous solution brought to pH 4.7 with HCl, the MS spectrum (Fig. 4, top) shows the classical expected profile (Riès-Kautt *et al.*, 1994) with +8 to +13 charge states with m/z 1101.82, 1193.60, 1301.87, 1431.71, 1590.81 and 1789.25. The average molecular weight is 14 305 ± 2 Da, consistent with the mass calculated from its chemical formula minus eight H atoms for the four disulfide bridges.

Several papers have correlated the number and values of the protein charged states with the ability of the protein to partially unfold in acidic or organic solvents or to preserve

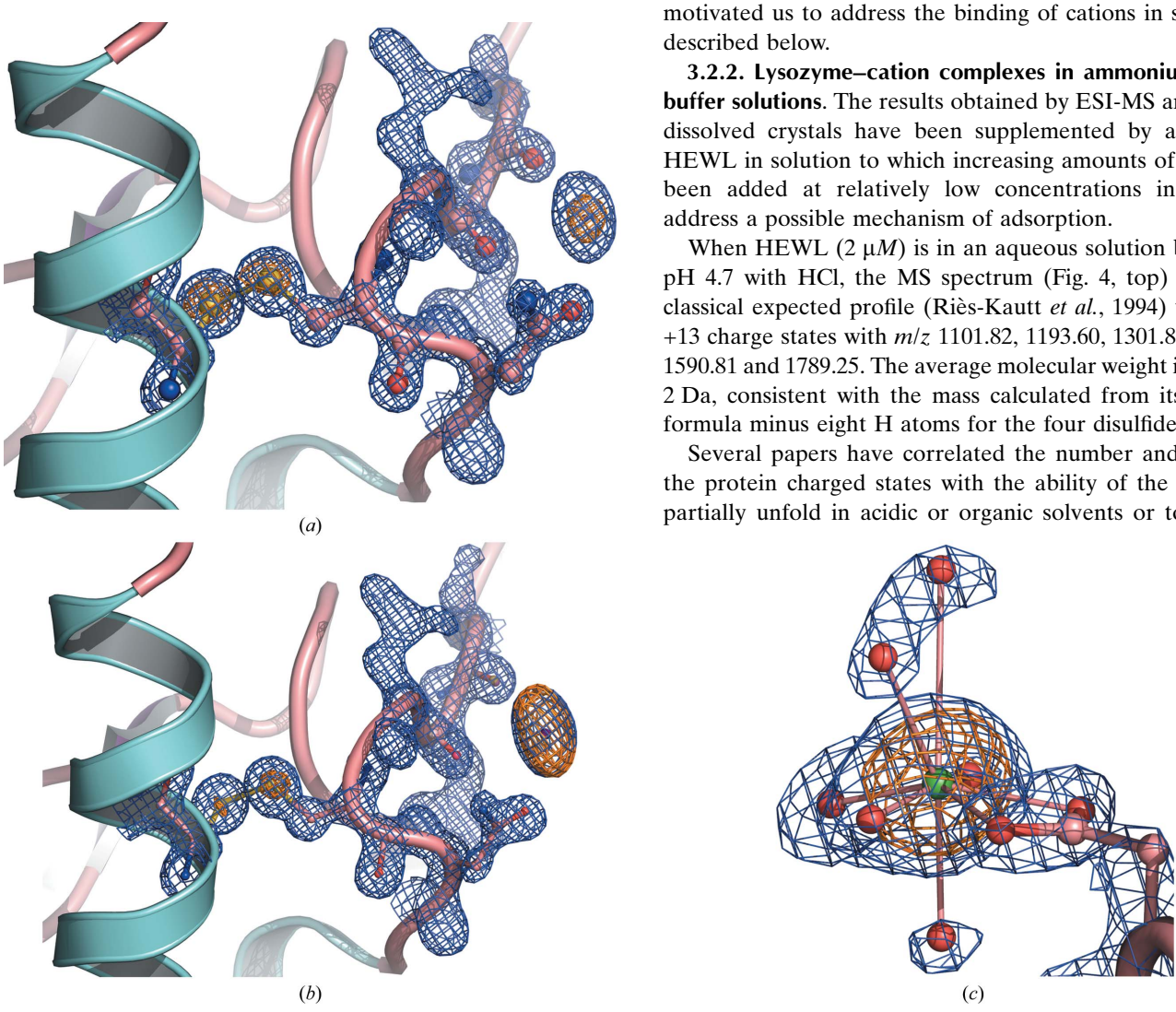

Figure 2
Anomalous difference and σ_A-weighted electron-density maps for Rb^+ (*a*) and Cs^+ (*b*) at binding site E and Yb^{3+} (*c*) at binding site B1. ($2mF_{obs} - DF_{calc}$) electron-density maps (blue) are contoured at 1.7σ. Anomalous difference maps (orange) are contoured at 3.6σ. (*a*, *b*) Electron-density maps are drawn around residues Arg73–Asn77, Cys94 and the cation. The disulfide bridge between Cys76 and Cys94 is also shown for comparison of anomalous signals. Cation-binding atoms of HEWL are shown: C═O of Arg73 and Asn74, as well as the side chain of Asn77 (for clarity the other atoms are hidden).

Table 4
Calculated and experimental molecular weights of [HEWL^{z+} + n cations].

		Molecular mass of HEWL + n cation(s) (Da)							
		Calculated			Experimental (±3)†				
					z = 8			z = 9	
Cation	Atomic mass	n = 0	n = 1	n = 2	n = 0	n = 1	n = 2	n = 0	n = 1
Cs	132.91	14305	14438	14571	14307	14441	—	14306	—
Mn	54.94		14360	14415	14304	14363	14414	14308	14361
Co	58.93		14364	14423	14306	14365	14425	14306	14365
Yb	173.04		14478	14651	14307	14479	14652	14308	14479

† Average value based on 3–12 ESI-MS measurements.

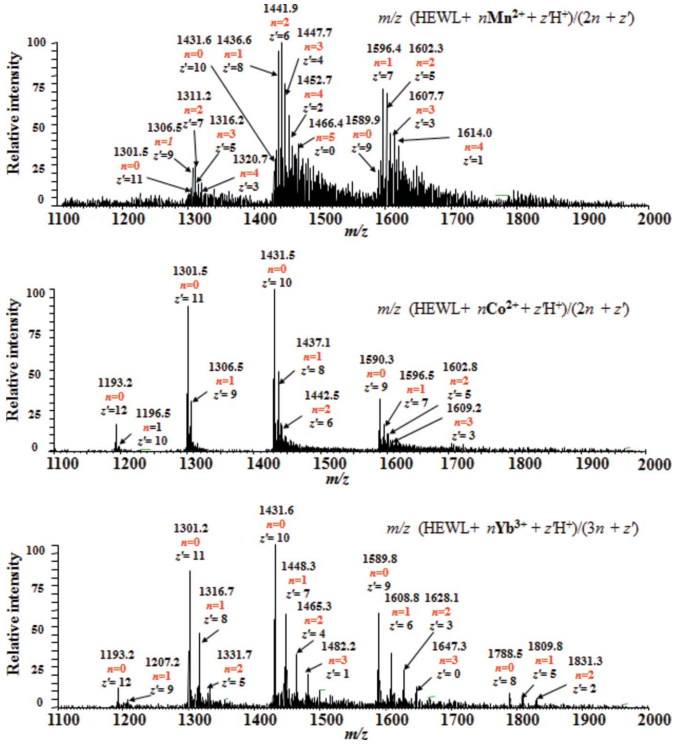

Figure 3
ESI-MS spectra of HEWL crystals dissolved in water.

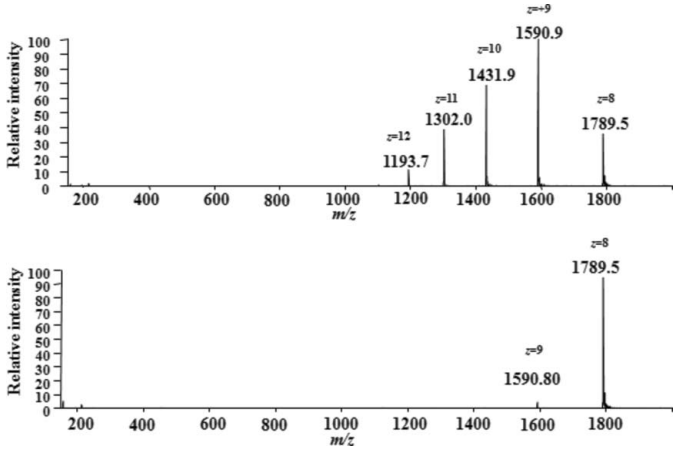

Figure 4
ESI-MS spectra of free HEWL in 85:15(*v*:*v*) water:MeOH brought to pH 4.5 with HCl (top) and in 20 m*M* ammonium acetate (bottom).

its native conformation in ammonium salts (Winston & Fitzgerald, 1997; Pramanik *et al.*, 1998; Heck & Van Den Heuvel, 2004) during transfer into the gas phase, including a recent study by ion-mobility MS (Angel, 2011). Although HEWL is known to be a very robust protein (Mao *et al.*, 2003), which undergoes no irreversible changes or loss of activity between pH 2.7 and 11 (Tanford & Wagner, 1954), we preferred to run the titration analysis in ammonium acetate. Indeed, the HEWL mass spectrum (Fig. 4, bottom) in 20 m*M* ammonium acetate buffer pH 4.5 exhibits only two charge states: z = 9 and z = 8. Deconvolution of 24 spectra leads to an experimental average mass of 14 306 ± 2, identical to that measured in HCl/water within the measurement accuracy, and is in agreement with no disruption of the disulfide bridges.

Comparing the two spectra, they mainly differ in the number of charged states (two instead of five) and in the most intense peak being +8 instead of +9. The reduced charge states observed in ammonium acetate might be attributed to ammonia participating in gas-phase proton-transfer reactions with multi-charged lysozyme (Lemaire *et al.*, 2001). Using ammonium acetate-buffered conditions yields a nearly single ionization state with an improved signal-to-noise ratio and is more favourable for investigating the adsorption by a semi-quantitative approach.

Ammonium-buffered HEWL solutions [20 m*M* ammonium acetate buffer pH 4.5 with 15%(*v*/*v*) methanol] were mixed with increasing concentrations of metal cation ranging from 20 to 600 µ*M*. Above 600 µ*M* salt the spectra become too perturbed by the presence of salt, especially when one recalls that the ionic strength equals three times the concentration of $MnCl_2$ or $CoCl_2$ and six times the salt concentration of $YbCl_3$.

At 100 µ*M* cation chlorides, corresponding to a HEWL:cation ratio of 1:50, the spectra (Fig. 5) show an ion peak corresponding to free lysozyme (n = 0) together with a peak for a HEWL complex with one cation (n = 1), except for Cs^+. Indeed, there is no peak owing to adsorption of Cs^+ at this salt concentration. In the presence of 350 µ*M* Mn^{2+}, Co^{2+} or Yb^{3+} (*i.e.* a HEWL:cation ratio of 1:175) a second peak appears corresponding to two adsorbed cations per protein molecule. Only a single cation is adsorbed in the case of Cs^+ at this concentration.

The experimental values of the different species are in agreement with the calculated values as shown in Table 4.

To conclude on the qualitative aspects, we were able to observe the adsorption of one and then two multivalent cations onto lysozyme molecules, but barely one monovalent Cs^+.

3.2.3. Titration curves. The signal intensities for free lysozyme (n = 0) and for HEWL–cation complexes (n = 1 or 2) were summed for the two protein charge states z = 9 and z = 8 over the whole range of salt concentrations. Fig. 6 shows the titration curves with the decrease of free HEWL and the

appearance of HEWL bearing one and two cations as a function of the salt concentration.

Increasing the concentration from 20 to 600 μ*M* Mn^{2+}, Co^{2+} or Yb^{3+} induces a decrease in free HEWL and the appearance of lysozyme complexes with one and then two bound cations. In contrast, increasing the Cs^{+} concentration produces only a slight decrease in free HEWL concomitant with the formation of 10% HEWL in complex with one Cs^{+}.

We must first emphasize that saturation was not reached in any of these MS experiments and therefore further quantitative conclusions, such as defining an association constant, for example, cannot be drawn. At the highest investigated salt concentration, 90% of the HEWL remains free in CsCl, 60% in $YbCl_3$ and approximately 30–40% in $MnCl_2$ and $CoCl_2$ (Fig. 7). The fraction of HEWL complex with one cation is $\leq$10% in CsCl but is $\simeq$30% for the multi-charged cations. Complexes with two cations account for 20% for Mn^{2+} or Co^{2+} and 10% for Yb^{3+}. Finally, HEWL complexes with three cations represent about 10% of the species with Mn^{2+} and Co^{2+} within experimental error.

From these results, we can estimate the total amount of each type of bound cation by multiplying the number *n* of cations by their fraction. For example, 30% of HEWL molecules bind one Mn^{2+} cation (0.3 Mn^{2+}), 20% bind two Mn^{2+} cations (0.4) and 10% bind three Mn^{2+} cations (0.3), giving a total of one bound Mn^{2+} per HEWL molecule on average. The same calculation gives about 0.1 Cs^{+} cation, 0.5 Yb^{3+} cation and about one Co^{2+} cation.

These ESI-MS results at 600 μ*M* salt suggest that the number of bound cations follows the order $Cs^{+} << Yb^{3+} < Co^{2+} \simeq Mn^{2+}$. Although it might be tempting to compare these values with q_{tot} obtained from the X-ray structures, we must be aware that here the HEWL–cation complexes are studied in a gas phase far from the crystal–solution equilibrium and strongly depend on their ionization ability. Ions trapped owing to crystal packing, such as Yb^{3+} at the C-terminus (site C) are not expected to be observed using this approach. Moreover, we have already underlined that ESI-MS analysis requires some MeOH to enhance the signal and we have seen that the addition of organic species might act differently in the binding

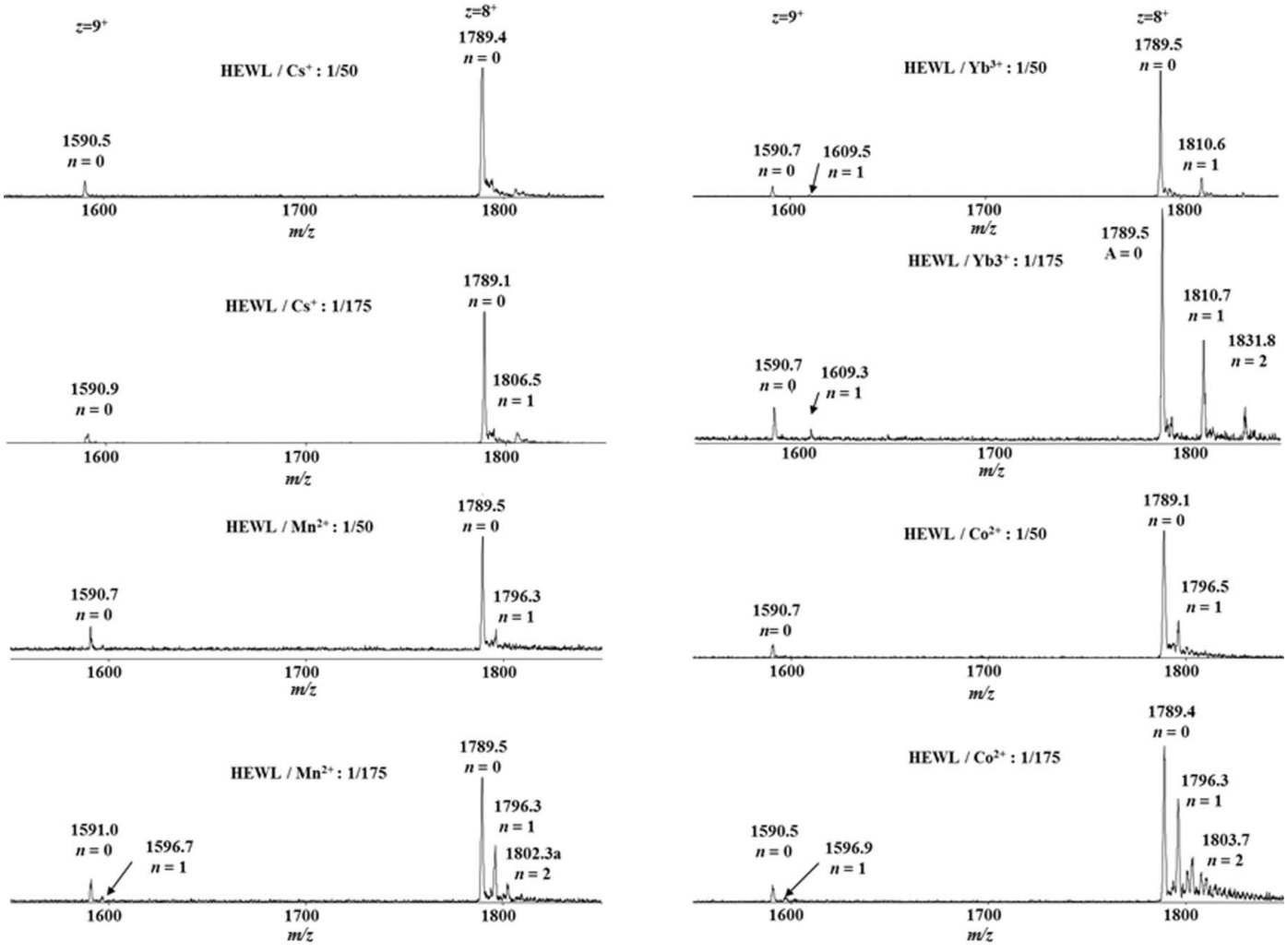

Figure 5
ESI-MS mass spectra of HEWL (2 μ*M*) in 20 m*M* ammonium acetate pH 4.5 buffer in 85:15(*v*:*v*) water:MeOH at two concentrations of added cations: protein:cation ratios of 1:50 (*i.e.* 100 μ*M* salt) and 1:175 (*i.e.* 350 μ*M* salt) as indicated. Cs^{+}, top left; Yb^{3+}, top right; Mn^{2+}, bottom left; Co^{2+}, bottom right.

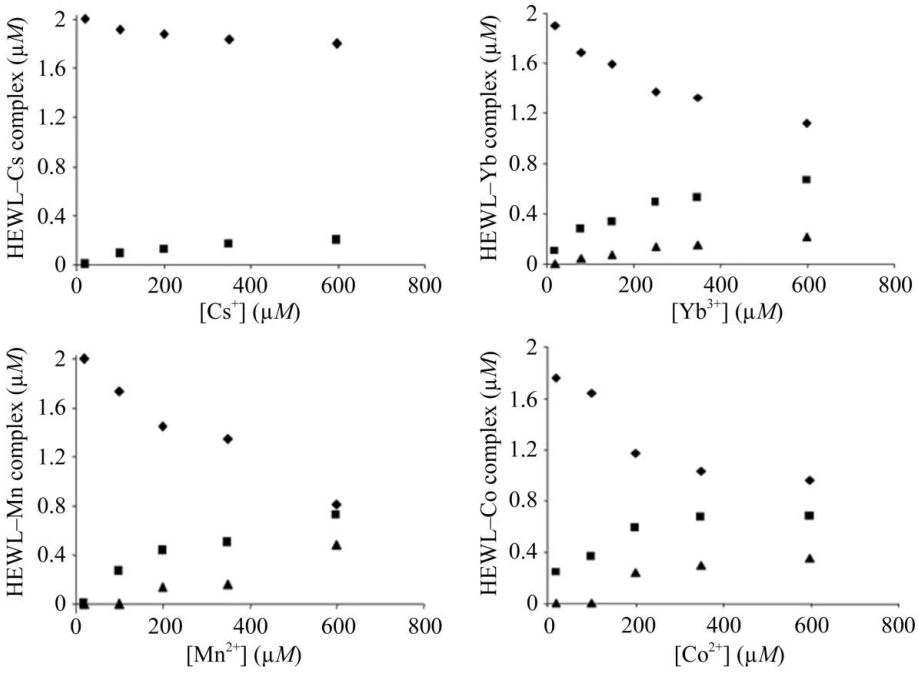

Figure 6
HEWL titration curves measured by ESI-MS. Concentration with 0 (diamonds), 1 (squares; $n = 1$) and 2 (triangles; $n = 2$) adsorbed cations *versus* salt concentration (μM) in solution: Cs^+, top left; Yb^{3+}, top right, Mn^{2+}, bottom left; Co^{2+}, bottom right.

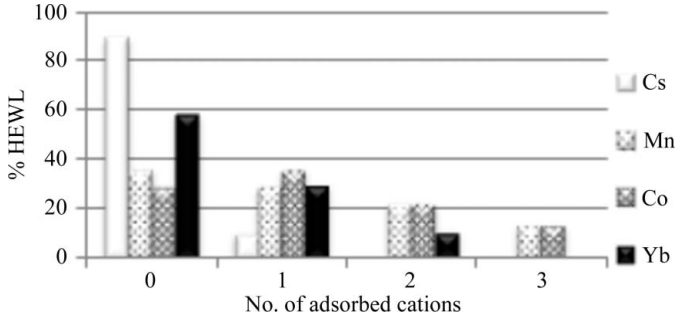

Figure 7
Percentage of free HEWL and as a complex with 1–3 cations at 600 μM salt.

of monovalent cations compared with multivalent cations for crystals grown in the presence of glycerol.

3.3. Ion-binding sites

The X-ray analysis allowed the identification of 13 different ion-binding areas (A–I) on the HEWL surface, ten binding cations and three binding chlorides, as illustrated in Fig. 8. The locations of the cation-binding sites (A–F) and Cl^- (H–I) are shown together with the electronegative (red) and positive (blue) surface areas of the protein. The areas A and B are subdivided into three binding sites each, as detailed below.

The shortest binding distances measured in the reported X-ray structures are listed in Table 5.

3.3.1. Binding zone A: Cs^+, Mn^{2+}, Co^{2+} and Yb^{3+}. The cations of binding zone A are found either close to the carboxyl of Asp52 (site A1), that of Glu35 (A2) or the carbonyl of Asn46 (A3). The glycoside hydrolase activity of lysozyme is known to involve Asp52 and Glu35 lying in the catalytic cleft (Vocadlo *et al.*, 2001), but no ions participate in the activity.

In the structures at room temperature, Co^{2+}, Mn^{2+} and Yb^{3+} are all bound to Asp52 $O^{\delta2}$ of site A1 (Fig. 9). Co^{2+}, the smallest of the three cations, binds only to this unique site (yellow ball), with a higher occupancy on increasing the ionic strength. It is close to five water molecules at distances of 2.0–2.5 Å. Asp52 $O^{\delta2}$ binds one Mn^{2+} cation at 0.5 *M* but three at 1.1 *M* (Fig. 9, right) in alternate positions. Mn^{2+} is also at a short distance from three water molecules. There are two alternate positions in site A1 for the largest Yb^{3+} cation and four water molecules are at short or medium distances depending on the data set.

In addition to site A1 a further Mn^{2+} cation is bound to the Glu35 carboxylic group (site A2) at 0.5 *M* as well as at 1.1 *M* $MnCl_2$. This site is also occupied by a Yb^{3+} cation (Fig. 9, left) but not by Co^{2+}. Site A2 is clearly separate from site A1, taking the distances into account and the sum of the refined occupancies, which would greatly exceed 100% in the crystal structure at 1.1 *M* $MnCl_2$.

For monovalent cations, no anomalous signal is observed in binding zone A among the structures at room temperature, unlike in the structure of HEWL in CsCl at 100 K. Indeed, three Cs^+ cations are identified: one at site A1 and two at A3. The latter are alternate positions interacting with the carbonyl of Asn46. In all of the structures presented in this work, only Cs occupies site A3.

3.3.2. Binding zone B: Cs^+ and Yb^{3+}. Binding zone B involves Asp101 and is about 16 Å from the catalytic cleft. Similarly to site A, it is subdivided into three sites: the cation is bound to Asp101 $O^{\delta2}$ at B1, to Asp101 $O^{\delta1}$ and to the carbonyl of the main chain at B2 and to Trp62 $C^{\delta1}$ at B3. In this zone, anomalous signals in difference Fourier maps were only observed in cryo-conditions for Cs^+ and Yb^{3+} (PDB entries 4ng8 and 4ngz, respectively).

One Yb^{3+} cation interacts with both Asp101 $O^{\delta1}$ and Asp101 $O^{\delta2}$ in site B1 (Fig. 10). This cation (green ball in Fig. 2*c*) is bound to the O atoms (red balls in Fig. 2*c*) of water molecules and of the carboxylic group of Asp101 by a bidentate coordination, in contrast to the Yb^{3+} cation in A1. The two upper apical water molecules are in alternate positions.

Four Cs^+ cations occupy site B (Fig. 10): two bound to Asp101 $O^{\delta2}$ at B1, one to Asp101 $O^{\delta1}$ and one interacting with the Trp62 indole group at B3, with the latter being in an alternate position with one Cs^+ at B1.

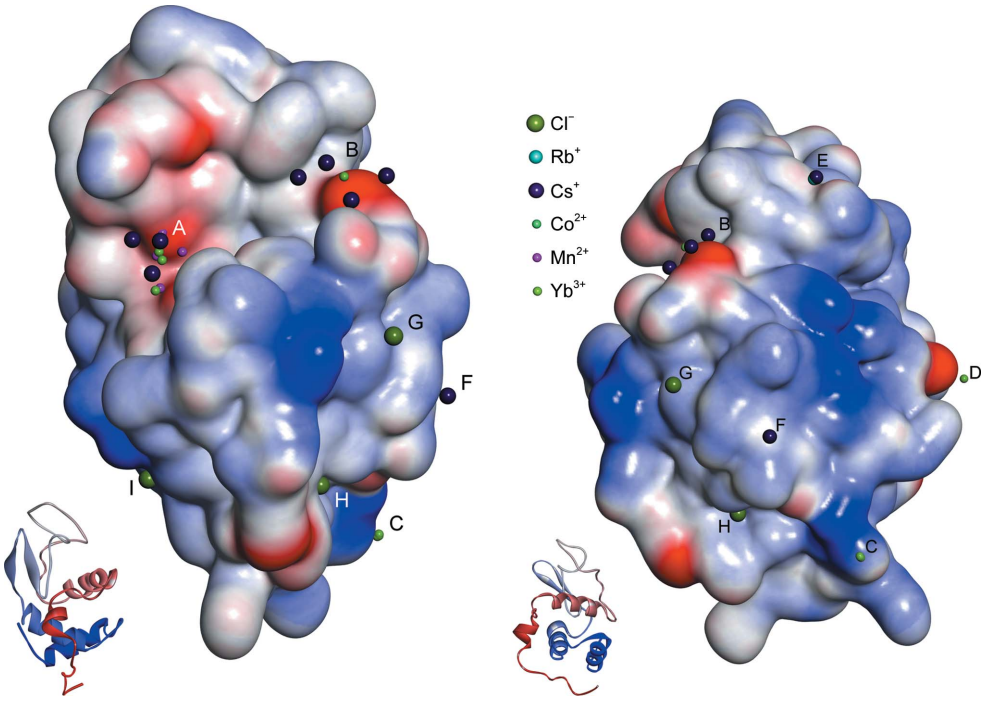

Figure 8
Ion-binding zones on HEWL. The mapping is a superimposition of the cation locations of the different HEWL structures onto PDB entry 4ngz represented by its electrostatic potential surface coloured from negative (red) to positive (blue) values. The two views are at 90° to each other. The ions are represented by a ball size proportional to their actual ionic radius and are coloured according to the chart. The insets at the bottom represent the HEWL secondary-structure elements coloured from the N-terminus (red) to the C-terminus (blue).

A major difference is observed between Cs^+ and Yb^{3+} since the binding of the former involves carbonyl groups whereas that of Yb^{3+} does not as it only binds to carboxylic acid groups.

3.3.3. Binding sites C and D: Yb^{3+}. In all HEWL structures with Yb^{3+} there is one Yb^{3+} cation bound to the C-termini of the two symmetry-related protein molecules (site C) as shown in Fig. 11. The occupancy is higher at 0.75 *M* than at 0.50 *M*, as indicated in Table 3.

Similarly to site C, only Yb^{3+} is observed bound to Asp87 $O^{\delta2}$ (site D), and only for the data recorded at 100 K.

3.3.4. Sites E and F: Rb^+ and Cs^+. The most interesting finding here is that only carbonyl groups bind the monovalent cations, as observed for binding zone B: Rb^+ and Cs^+ both interact with Asn74 O and Asn77 $O^{\delta1}$ as well as with the main-chain carbonyl of Arg73 at site E (Fig. 8). Binding site E appears to be specific for monovalent cations, since no anomalous signal can be detected in the structures involving divalent or trivalent cations.

Another Cs^+ is found in site F, interacting with the carbonyl groups of Asn19 and Asn44, with an occupancy of 0.60 at 100 K.

3.3.5. Chloride-binding sides G, H and I. In all of the structures solved in this work, a Cl^- anion is found bound to Tyr23 OH (site G). In structures obtained using data collected under cryoconditions, a second Cl^- anion is observed bound to Ser24 O^{γ} and Gly26 NH. A third chloride anion is finally identified interacting with Lys33 in the HEWL structure crystallized in 0.48 *M* $YbCl_3$, 30% glycerol, from which data were collected under cryoconditions (PDB entry 4ngz).

All of these sites G–I have previously been described as anion-binding sites of HEWL (Dauter *et al.*, 1999; Vaney *et al.*, 2001).

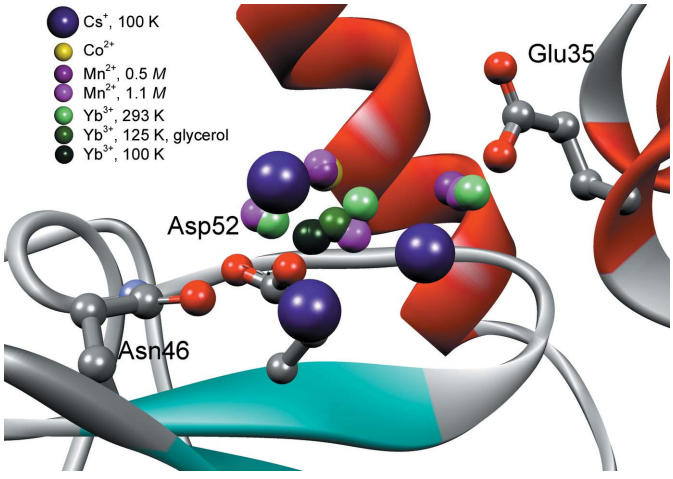

Figure 9
Cation localizations in binding zone A depending on the crystallization and/or data-collection conditions.

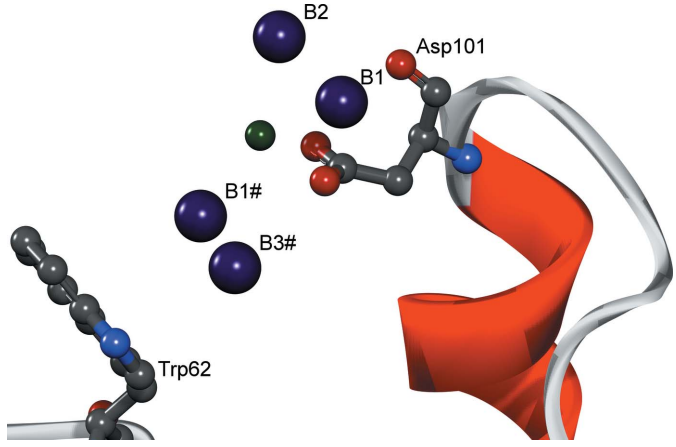

Figure 10
Binding zone B with four Cs^+ and one Yb^{3+} cations in the three sites involving Asp101 $O^{\delta1}$, Asp101 $O^{\delta2}$ and Asp101 O. (Cs^+, large indigo balls; Yb^{3+}, green). Alternate Cs^+ positions are marked #.

Table 5
Shortest ion-binding distances (Å) for each type of cation and for Cl^- in the different HEWL structures.

Salt	Rb^+		Cs^+		Mn^{2+}		Co^{2+}			Yb^{3+}			
PDB code	4ngj	4ngi	4ng1	4ng8	4neb	4nfv	4ngk	4ngl	4ngo	4ngv	4ngw	4ngy	4ngz
Sites A													
A1 (Asp52 $O^{\delta2}$)				3.23	2.18	2.29	2.17	2.02	1.98	2.15	1.97	1.95	2.13
						2.27				2.66			
						2.06							
A2 (Glu35 $O^{\delta2}$)					2.29	2.33				2.77			
A3 (Asn46 O^{δ})				2.72									
				3.04									
Sites B													
B1 (Asp101 $O^{\delta2}$)				3.15									2.35
				3.14									
B2 (Asp101 $O^{\delta1}$/O)				3.12									
B3 (Trp62 $C^{\delta1}$)				3.59									
Site C (C-terminus)										2.43	2.35	2.42	2.26
Site D (Asp87 $O^{\delta2}$)											2.57	2.52	
Site E (Asn77 O^{δ})	2.86	3.05	2.92	2.94									
Site F (Asn44 O^{δ})				3.02									
Cl^-													
Site G (Tyr23 OH)	3.02	3.07	2.92	3.01	2.95	2.99	2.89	2.89	3.01	2.94	3.00	3.03	2.98
Site H (Ser24 O^{γ})	3.07			3.11							3.03	3.10	3.08
Site I (Lys33 N^{ξ})													3.77

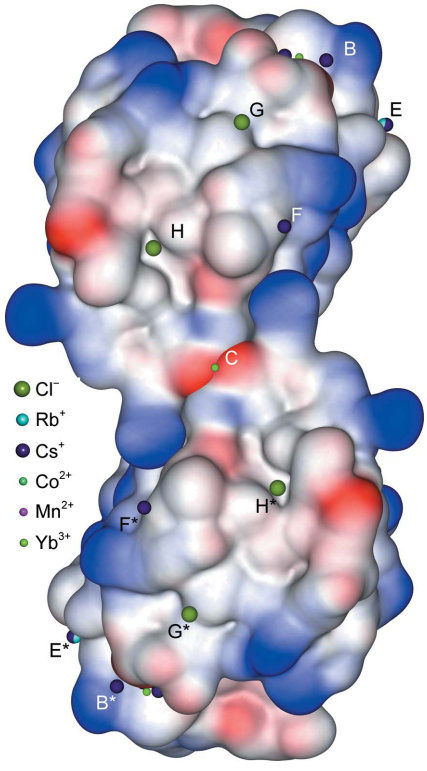

Figure 11
Yb^{3+} bound to site C involving the C-termini of two symmetry-related HEWL molecules. On the electrostatic potential surface of HEWL (same colour code as in Fig. 8) the cations are represented by a ball size proportional to their ionic radius and are coloured according to the chart.

4. Conclusions

Both the X-ray and the ESI-MS results presented in this study answer the first question addressed in this work by providing previously lacking experimental evidence to explain how these co-ions influence the solubility of the positively charged protein. This work therefore contributes to the elucidation of salt effects and extends Hofmeister's multivalent cation series. Cations interact more weakly as co-ions with proteins than do anionic counter-ions; the highest cation occupancy among the structures presented in this study remains lower than that of Cl^-. This is owing to their coordination requirements and of their different solvation in H_2O, since they interact with the O atoms of water molecules while anions point towards the H atoms.

Moreover, our results allow progress in the understanding of ion–protein interactions, especially in the relative contributions of hydration and electrostatics.

4.1. The monovalent and multivalent cations bind to different types of protein binding sites

The large monovalent Rb^+ and Cs^+ cations (ionic radius 1.5–1.7 Å) preferentially interact with the carbonyl groups of the HEWL molecule, while the smaller multivalent Co^{2+}, Mn^{2+} and Yb^{3+} cations (0.7–0.9 Å) preferably bind to the carboxyl groups of the Asp and Glu side chains or of the main-chain C-terminus.

Rb^+ and Cs^+ are chaotropic ions with a low charge density and poor hydration. They interact more favourably with the carbonyl groups of the protein *via* an ion–dipole interaction than with the charged and more hydrated carboxylic acids, as predicted by the law of matching water affinity. The occupancy is reinforced in the structures from the cryo data sets. In the various X-ray structures involving Rb^+ or Cs^+ there is no water molecule at a short distance (<2.5 Å). The smaller Rb^+ is less bound than the larger Cs^+ according to the observations from the X-ray and ESI-MS analyses.

The multivalent Co^{2+}, Mn^{2+} and Yb^{3+} cations bind to carboxylic groups on the protein surface *via* Coulombic interactions, which are less sensitive to temperature than are

ion–dipole interactions. Carboxylic groups are known to be strongly hydrated with a high charge density, characterizing them as kosmotropes (Collins, 2004). With respect to inner sphere ion-pair formation, carboxylic groups are presumed to interact more favourably with small cations that also bear a high charge density (Collins, 2006), *i.e.* that are highly hydrated. Water molecules are indeed observed at a short distance (<2.5 Å) from the cation depending on the data set.

Overall, our experimental results provide experimental insights to explain the contribution of the chemistry of the salt effect in terms of favourable interactions of soft ions with soft protein sites or of hard ions with hard protein sites (Collins, 2012).

4.2. The efficiency of the cations in affecting protein solubility

HEWL solubility decreases until the salt concentration reaches about 0.5 *M*, whatever the cation. The solubility continues to decrease when increasing the monovalent cation concentration, but increases with multivalent cations. The different results presented in this study clarify the solubility behaviour, providing the following considerations.

The ESI-MS titration shows the binding of about 0.1 Cs^{+}, 0.5 Yb^{3+}, one Mn^{2+} and one Co^{2+} cations per HEWL molecule at 600 µ*M* salt, suggesting that the number of bound cations follows the order $Cs^{+} << Yb^{3+} < Mn^{2+} \simeq Co^{2+}$, in agreement with the law of matching water affinity, *i.e.* the order of their hydration.

The number of bound cations (q_{tot}) from the X-ray structure additionally includes cations from crystal packing. This is the case for Yb^{3+} at site C at the interface of two symmetrical HEWL molecules.

Furthermore, it must be underlined that binding a trivalent cation increases the apparent net charge of the resulting protein polyelectrolyte more, and hence its solubility, than a divalent cation does.

Owing to these considerations, the ranking of the cation effect on HEWL solubility becomes $Cs^{+} << Mn^{2+} \simeq Co^{2+} < Yb^{3+}$.

A noteworthy outcome of this work is the coherence between the two biophysical approaches to address ion–protein interactions, although cation binding to the protein is different in the gas phase inside the mass-spectrometer source in vacuum and in the solid phase of a crystal grown at high ionic strength.

4.3. The change in the protein polyelectrolyte with increasing ionic strength

When isoionic HEWL is brought to pH 4.5 it acquires a net charge of about +10 and has at least ten counter-ions. Its solubility with Cl^{-} as counter-ion is as high as 366 mg ml^{-1} (Retailleau, Riès-Kautt *et al.*, 1997). Increasing the salt concentration not only affects the hydration of the protein polyelectrolyte in solution but also changes its composition, since ions bind to specific sites of the protein. Along the phase diagram different HEWL complexes are formed because of the binding of an increasing number of counter-ions and co-ions, with their amount depending on their nature and on their own affinity towards the different types of protein binding sites. The overall net charge of the complex is therefore affected, as is its solubility.

The behaviour of a protein in solution can consequently no longer be considered only in terms of a given polyelectrolyte with its net charge and hydration shell, but has to be addressed at the atomic level of local competition of solvent molecules and ions, *i.e.* the contribution of the chemistry of the salt effect in protein crystallization in terms of favourable ion pairs of soft ions with soft protein sites or hard ions with hard protein sites according to the law of matching water affinity (Collins, 2012). Since soluble proteins expose the same amino acids to the solvent, the HEWL results on ion–protein interactions can be transposed to general rules for the crystallization of soluble proteins, as well as to their folding and stability.

Francine Libot is acknowledged for initial ESI-MS analysis of dissolved crystals in acetonitrile/water. PB was supported by the European Biocrystallogenesis initiative (contract No. BIO4-CT98-0086 (DG12-SSMI). We thank the Centre National de la Recherche Scientifique (CNRS) for financial support of this project. We acknowledge the Laboratoire pour l'Utilisation du Rayonnement Electromagnétique (LURE) and the European Synchrotron Radiation Facility (ESRF) for providing the synchrotron-radiation facilities, and particularly Dr C. Mueller-Dieckmann for assistance in using beamline ID29 at the ESRF. The authors are indebted to Professor T. Prangé, Professor C. W. Carter and Professor A. Ducruix for helpful scientific discussions.

References

Angel, L. A. (2011). *Eur. J. Mass Spectrom.* **17**, 207–215.
Arakawa, T. & Timasheff, S. N. (1984). *Biochemistry*, **23**, 5912–5923.
Bénas, P., Legrand, L. & Riès-Kautt, M. (2002). *Acta Cryst.* D**58**, 1582–1587.
Brünger, 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.
Collins, K. D. (2004). *Methods*, **34**, 300–311.
Collins, K. D. (2006). *Biophys. Chem.* **119**, 271–281.
Collins, K. D. (2012). *Biophys. Chem.* **167**, 43–59.
Collins, K. D. & Washabaugh, M. W. (1985). *Q. Rev. Biophys.* **18**, 323–422.
Dauter, Z. & Dauter, M. (1999). *J. Mol. Biol.* **289**, 93–101.
Dauter, Z., Dauter, M., de La Fortelle, E., Bricogne, G. & Sheldrick, G. M. (1999). *J. Mol. Biol.* **289**, 83–92.
Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* D**66**, 486–501.
French, S. & Wilson, K. (1978). *Acta Cryst.* A**34**, 517–525.
Hamiaux, C., Prangé, T., Riès-Kautt, M., Ducruix, A., Lafont, S., Astier, J. P. & Veesler, S. (1999). *Acta Cryst.* D**55**, 103–113.
Heck, A. J. & Van Den Heuvel, R. H. (2004). *Mass Spectrom. Rev.* **23**, 368–389.
Hippel, H. von & Schleich, T. (1969). *Structure Stability of Biological Macromolecules*, edited by F. E. Timasheff, pp. 418–574. New York: Marcel Dekker.
Hofmeister, F. (1888). *Arch. Exp. Pathol. Pharmakol.* **24**, 247–260.
Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* A**47**, 110–119.

Joosten, R. P., Joosten, K., Cohen, S. X., Vriend, G. & Perrakis, A. (2011). *Bioinformatics*, **27**, 3392–3398.
Kabsch, W. (2010*a*). *Acta Cryst.* D**66**, 125–132.
Kabsch, W. (2010*b*). *Acta Cryst.* D**66**, 133–144.
Kunz, W., Henle, J. & Ninham, B. W. (2004). *Curr. Opin. Colloid Interface Sci.* **9**, 19–37.
Lemaire, D., Marie, G., Serani, L. & Laprévote, O. (2001). *Anal. Chem.* **73**, 1699–1706.
Li, S. J., Matsuura, T., Tanaka, H., Nakagawa, A., Tsukihara, T. & Ataka, M. (2005). *J. Cryst. Growth*, **276**, 222–229.
Mao, D., Babu, K. R., Chen, Y.-L. & Douglas, D. J. (2003). *Anal. Chem.* **75**, 1325–1330.
McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
Ménez, R. & Ducruix, A. (1990). *J. Mol. Biol.* **216**, 233–234.
Ménez, R. & Ducruix, A. (1993). *J. Mol. Biol.* **232**, 997–998.
Moreau, S., Awadé, A. C., Mollé, D., Le Graet, Y. & Brulé, G. (1995). *J. Agric. Food Chem.* **43**, 883–889.
Navaza, J. (1994). *Acta Cryst.* A**50**, 157–163.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Pearson, R. G. (1987). *J. Chem. Educ.* **64**, 561.
Potier, N., Rogniaux, H., Chevreux, G. & Van Dorsselaer, A. (2005). *Methods Enzymol.* **402**, 361–389.
Pramanik, B. N., Bartner, P. L., Mirza, U. A., Liu, Y.-H. & Ganguly, A. K. (1998). *J. Mass Spectrom.* **33**, 911–920.
Read, R. J. (1986). *Acta Cryst.* A**42**, 140–149.
Retailleau, P., Ducruix, A. & Riès-Kautt, M. (2002). *Acta Cryst.* D**58**, 1576–1581.
Retailleau, P., Riès-Kautt, M. & Ducruix, A. (1997). *Biophys. J.* **73**, 2156–2163.
Riès-Kautt, M. M. & Ducruix, A. F. (1989). *J. Biol. Chem.* **264**, 745–748.
Riès-Kautt, M. M. & Ducruix, A. F. (1991). *J. Cryst. Growth*, **110**, 20–25.
Riès-Kautt, M. & Ducruix, A. (1997). *Method Enzymol.* **276**, 23–59.
Riès-Kautt, M., Ducruix, A. & Van Dorsselaer, A. (1994). *Acta Cryst.* D**50**, 366–369.
Saludjian, P., Prangé, T., Navaza, J., Ménez, R., Guilloteau, J. P., Riès-Kautt, M. & Ducruix, A. (1992). *Acta Cryst.* B**48**, 520–531.
Tanford, C. (1961). *Physical Chemistry of Macromolecules*, pp. 238–253. New York: John Wiley & Sons.
Tanford, C. & Wagner, M. L. (1954). *J. Am. Chem. Soc.* **76**, 3331–3336.
Thomas, B. R., Vekilov, P. G. & Rosenberger, F. (1996). *Acta Cryst.* D**52**, 776–784.
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, C., Riès-Kautt, M. & Ducruix, A. (1996). *Acta Cryst.* D**52**, 505–517.
Vocadlo, D. J., Davies, G. J., Laine, R. & Withers, S. G. (2001). *Nature (London)*, **412**, 835–838.
Washabaugh, M. W. & Collins, K. D. (1986). *J. Biol. Chem.* **261**, 12477–12485.
Weiss, M. S., Panjikar, S., Nowak, E. & Tucker, P. A. (2002). *Acta Cryst.* D**58**, 1407–1412.
Whittal, R. M., Ball, H. L., Cohen, F. E., Burlingame, A. L., Prusiner, S. B. & Baldwin, M. A. (2000). *Protein Sci.* **9**, 332–343.
Winn, M. D. *et al.* (2011). *Acta Cryst.* D**67**, 235–242.
Winston, R. L. & Fitzgerald, M. C. (1997). *Mass Spectrom. Rev.* **16**, 165–179.
Zhang, Y. & Cremer, P. S. (2006). *Curr. Opin. Chem. Biol.* **10**, 658–663.