Acta Crystallographica Section D
**Biological Crystallography**
ISSN 0907-4449

# Crystal packing of plant-type L-asparaginase from *Escherichia coli*

**Karolina Michalska,[a] Dominika Borek,[a,b] Alejadra Hernández-Santoyo[a] and Mariusz Jaskolski[a,c]***

[a]Department of Crystallography, Faculty of Chemistry, A. Mickiewicz University, Poznan, Poland, [b]Department of Biochemistry, University of Texas Southwestern Medical Center at Dallas, Dallas, TX, USA, and [c]Center for Biocrystallographic Research, Institute of Bioorganic Chemistry, Polish Academy of Sciences, Poznan, Poland

Correspondence e-mail: mariuszj@amu.edu.pl

Received 8 November 2007
Accepted 21 December 2007

**PDB Reference:** T179A mutant of EcAIII, 2zak, r2zaksf.

Plant-type L-asparaginases hydrolyze the side-chain amide bond of L-asparagine or its $\beta$-peptides. They belong to the N-terminal nucleophile (Ntn) hydrolases and are synthesized as inactive precursor molecules. Activation occurs *via* the autoproteolytic release of two subunits, $\alpha$ and $\beta$, the latter of which carries the nucleophile at its N-terminus. Crystallographic studies of plant-type asparaginases have focused on an *Escherichia coli* homologue (EcAIII), which has been crystallized in several crystal forms. Although they all belong to the same $P2_12_12_1$ space group with similar unit-cell parameters, they display different crystal-packing arrangements and thus should be classified as separate polymorphs. This variability stems mainly from different positions of the EcAIII molecules within the unit cell, although they also exhibit slight differences in orientation. The intermolecular interactions that trigger different crystal lattice formation are mediated by ions, which represent the most variable component of the crystallization conditions. This behaviour confirms recent observations that small molecules might promote protein crystal lattice formation.

## 1. Introduction

L-Asparaginase hydrolyzes L-asparagine to L-aspartate and ammonia. There are two groups of evolutionarily distinct enzymes that catalyze this reaction: bacterial-type and plant-type L-asparaginases (Borek & Jaskolski, 2001; Michalska & Jaskolski, 2006). Plant-type L-asparaginases can be further divided into two subclasses defined by their potassium-(in)dependence. The best studied plant asparaginase is LlA, the potassium-independent enzyme from yellow lupin (*Lupinus luteus*; Borek *et al.*, 1999, 2004; Michalska *et al.*, 2006). Representatives of the potassium-independent enzymes are found not only in the plant kingdom, but also in bacteria (Borek & Jaskolski, 2001). The best studied enzyme in this group is EcAIII, which is encoded by the *iaa*A (or *ybi*K) gene in *Escherichia coli* (Borek & Jaskólski, 2000; Borek, 2001; Michalska *et al.*, 2005).

Plant-type L-asparaginases exhibit a relatively low affinity for L-asparagine ($K_m$ = 3.9 m*M* for EcAIII), especially when compared with their order-of-magnitude higher affinity for isoaspartyl (or $\beta$-aspartyl) peptides ($K_m$ = 0.14 m*M* for EcAIII; Borek *et al.*, 2004; Larsen *et al.*, 2001). The latter activity, which is only found in potassium-independent L-asparaginases (Bruneau *et al.*, 2006), is responsible for the degradation of aberrant isoaspartyl peptides arising spontaneously in proteins on prolonged storage.

**Table 1**
Crystal structures of EcAIII deposited in the PDB.

All structures belong to space group $P2_12_12_1$, with either one $(\alpha\beta)_2$ heterotetramer or a dimer of precursor molecules in the asymmetric unit. In all structures, the asymmetric unit contents have been standardized to the all-positive octant of the unit cell with minimum distance from the origin (see text).

| | 2zal | 1k2x | 1jn9 | 1t3m | T179A |
|---|---|---|---|---|---|
| Crystallization conditions | 0.08 *M* $CaCl_2$, 0.1 *M* Tris–HCl pH 8.5, 13% PEG 400, 17% PEG 4K, 0.1 *M* Na L-Asp | 0.2 *M* $MgCl_2$, 0.1 *M* Tris–HCl pH 8.5, 15% PEG 400, 20% PEG 4K | 0.2 *M* $CaCl_2$, 0.1 *M* Tris–HCl pH 8.5, 15% PEG 400, 20% PEG 4K | 0.3 *M* $Mg(NO_3)_2$, 0.1 *M* Tris–HCl pH 6.5†, 15% glycerol, 14% PEG 4K | 0.2 *M* $MgCl_2$, 0.1 *M* Tris–HCl pH 8.5, 15% PEG 4K |
| Unit-cell parameters (Å) | $a$ = 49.9, $b$ = 77.3, $c$ = 147.5 | $a$ = 50.3, $b$ = 77.6, $c$ = 148.2 | $a$ = 61.9, $b$ = 70.4, $c$ = 148.9 | $a$ = 66.3, $b$ = 71.6, $c$ = 149.5 | $a$ = 51.4, $b$ = 77.8, $c$ = 147.9 |
| Unit-cell volume (Å$^3$) | 568686 | 578411 | 649365 | 710642 | 591440 |
| Matthews coefficient (Å$^3$ Da$^{-1}$) | 2.30 | 2.17 | 2.43 | 2.66 | 2.17 |
| Direction cosines of the molecular NCS dyad | $\cos\varepsilon_1$ = 0.945<br>$\cos\varepsilon_2$ = −0.309<br>$\cos\varepsilon_3$ = 0.109 | $\cos\varepsilon_1$ = 0.995<br>$\cos\varepsilon_2$ = −0.071<br>$\cos\varepsilon_3$ = −0.065 | $\cos\varepsilon_1$ = 0.902<br>$\cos\varepsilon_2$ = −0.414<br>$\cos\varepsilon_3$ = −0.124 | $\cos\varepsilon_1$ = 0.954<br>$\cos\varepsilon_2$ = −0.299<br>$\cos\varepsilon_3$ = −0.021 | $\cos\varepsilon_1$ = 0.947<br>$\cos\varepsilon_2$ = −0.313<br>$\cos\varepsilon_3$ = 0.073 |
| Direction cosines of the Na(B)–Na(A) vector‡ | $\cos\varphi_1$ = 0.165<br>$\cos\varphi_2$ = 0.161<br>$\cos\varphi_3$ = −0.973 | $\cos\varphi_1$ = 0.072<br>$\cos\varphi_2$ = 0.188<br>$\cos\varphi_3$ = −0.980 | $\cos\varphi_1$ = −0.016<br>$\cos\varphi_2$ = 0.245<br>$\cos\varphi_3$ = −0.969 | $\cos\varphi_1$ = 0.065<br>$\cos\varphi_2$ = 0.283<br>$\cos\varphi_3$ = −0.957 | $\cos\varphi_1$ = 0.131<br>$\cos\varphi_2$ = 0.178<br>$\cos\varphi_3$ = −0.975 |
| Fractional coordinates of Na(B)—Na(A) midpoint | 0.037, 0.055, 0.414 | 0.008, 0.017, 0.299 | 0.086, 0.058, 0.195 | 0.092, 0.047, 0.200 | 0.042, 0.056, 0.415 |
| Distance of the Na(B)—Na(A) midpoint from the origin (Å) | 61.3 | 44.4 | 29.5 | 30.7 | 61.6 |

† As given in Prahl *et al.* (2004). The value cited in the PDB deposition with accession code 1t3m is 7.5. ‡ The two sodium ions are coordinated by residues 60–70 from chains corresponding to subunit $\alpha$ of the mature EcAIII molecule.

Structurally and mechanistically, plant-type asparaginases belong to the family of N-terminal nucleophile (Ntn) hydrolases, the distinctive feature of which is the utilization of an N-terminal nucleophilic Thr, Ser or Cys residue in catalysis (Brannigan *et al.*, 1995). Ntn hydrolases are represented for example, by aspartylglucosaminidases (Oinonen *et al.*, 1995; Guo *et al.*, 1998), the proteasome subunit $\beta$ (Löwe *et al.*, 1995; Groll *et al.*, 1997) or taspase1 (Tas1; Khan *et al.*, 2005). Their activity is oriented towards amide bonds in a wide range of substrates of different sizes. The enzymes are expressed as inactive precursors requiring autoproteolytic cleavage of a single peptide bond for maturation. As a result of autoproteolysis, either a propeptide is removed (Seemüller *et al.*, 1996) or two subunits, $\alpha$ and $\beta$, are created, generating the N-terminal nucleophile (Michalska *et al.*, 2005; Duggleby *et al.*, 1995; Guan *et al.*, 1998). In the latter case, the subunits form one $(\alpha\beta)$ functional domain with the catalytic residue at the N-terminus of subunit $\beta$.

Despite high sequence variability, the $(\alpha\beta)$ heterodimers share a common sandwich-like $\alpha\beta\beta\alpha$ topology (Oinonen & Rouvinen, 2000). The inner part of this fold is formed by two parallel $\beta$-sheets that face each other and are surrounded by two layers of $\alpha$-helices. The two $\beta$-sheets are of unequal extent, consisting of eight and four $\beta$-chains. The smaller sheet is formed entirely within the $\beta$ subunit, while the larger one comprises $\beta$-chains from both subunits. The nucleophilic residue is located in a deep cleft in the middle of the larger $\beta$-sheet.

Structural and biochemical studies of EcAIII (Borek, 2001; Prahl *et al.*, 2004; Michalska *et al.*, 2005) and LlA (Michalska *et al.*, 2006) revealed that the proteins form $(\alpha\beta)_2$ dimers of heterodimers with two active sites sculpted around the nucleophilic threonine residue (Thr179 in EcAIII). The structure of EcAIII in complex with the reaction product L-aspartate (Michalska *et al.*, 2005) is of particular importance as it reveals the binding mode of a molecule that closely mimics the natural substrate L-asparagine and explains the observed substrate specificity. The most discriminating role is played by the conserved Arg207, which forms a strong salt bridge with the $\alpha$-carboxylate group of the substrate. In the structure of LlA crystallized without an L-Asp ligand the position of the $\alpha$-carboxy group is occupied by a chloride anion, which forms analogous hydrogen bonds to the guanidinium group of the Arg anchor. Superposition of plant-type asparaginases with taspase1 reveals that the arginine residue may act as a switch, presenting for hydrolysis either a $\beta$-Asn-type substrate or $\alpha$-Asp-peptide bonds. The required conformation of the arginine side chain is defined by interactions at the $(\alpha\beta)$–$(\alpha\beta)$ interface, explaining the oligomeric structure of this group of enzymes (Michalska *et al.*, 2006).

Currently, mature EcAIII is represented in the Protein Data Bank (Berman *et al.*, 2000) by four crystallographic models with the following PDB accession codes: 1k2x (1.65 Å resolution), 1jn9 (2.27 Å; Borek, 2001), 1t3m (1.65 Å; Prahl *et al.*, 2004) and 2zal (1.90 Å; Michalska *et al.*, 2005). The first three structures correspond to EcAIII alone, while 2zal represents EcAIII in complex with L-aspartate. All these structures were solved in the orthorhombic space group $P2_12_12_1$ with similar unit-cell parameters (Table 1). However, this similarity is quite misleading as these forms exhibit different crystal-packing modes, which have not been analyzed in detail so far. It has recently been proposed that combining relatively narrow ranges of basic conditions with a variety of small molecules as additives is an efficient strategy for macromolecular crystallization (McPherson & Cudney, 2006). Here, we demonstrate that with the same precipitating agent (polyethylene glycol)

**Table 2**
Data-collection and processing statistics for T179A EcAIII.

Values in parentheses are for the last resolution shell.

| | |
|---|---|
| Radiation source | X12, EMBL Hamburg |
| Wavelength (Å) | 1.000 |
| Temperature of measurements (K) | 100 |
| Space group | $P2_12_12_1$ |
| Unit-cell parameters (Å) | $a$ = 51.4, $b$ = 77.8, $c$ = 147.9 |
| Mosaicity (°) | 0.6 |
| Resolution range (Å) | 25.0–2.0 (2.07–2.00) |
| Oscillation step (°) | 0.3 |
| No. of images | 600 |
| Reflections collected | 115601 |
| Unique reflections | 38010 (2 743) |
| Completeness (%) | 94.0 (69.1) |
| Redundancy | 3.0 (2.1) |
| $\langle I\rangle/\langle\sigma(I)\rangle$ | 9.2 (2.1) |
| $R_{\rm merge}$† | 0.103 (0.400) |

† $R_{\rm merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle|/\sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the intensity of observation $i$ of reflection $hkl$.

**Table 3**
Refinement statistics for T179A EcAIII.

| | |
|---|---|
| Resolution (Å) | 25.0– 2.0 |
| No. of reflections in working set | 36439 |
| No. of reflections in test set | 1524 |
| $R$†/$R_{\rm free}$ | 0.197/0.253 |
| No. of atoms | |
| Protein | 4299 |
| Ligand | 8 |
| Solvent | 140 |
| $Na^+$ | 2 |
| $Cl^-$ | 2 |
| R.m.s. deviations from ideality | |
| Bond lengths (Å) | 0.016 |
| Bond angles (°) | 1.55 |
| Average $B$ factor (Å$^2$) | 23.6 |
| Residues in Ramachandran plot (%) | |
| Most favoured regions | 90.5 |
| Allowed regions | 9.5 |

† $R = \sum_h ||F_o| - |F_c||/\sum_h |F_o|$ for all reflections, where $F_o$ and $F_c$ are the observed and calculated structure factors, respectively. $R_{\rm free}$ was calculated analogously for the test reflections, which were randomly selected and excluded from the refinement.

and different salts, several polymorphs can be obtained. We analyze the packing arrangements of the EcAIII molecules in the different crystal forms of the mature protein and of a mutated T179A precursor protein and try to correlate mother-liquor additives with the interactions between the crystal structure components.

## 2. Methods

### 2.1. Structural models

This work is based on the analysis of five crystal structures of EcAIII. Four of them have been deposited in the PDB under accession codes 1k2x, 1jn9, 1t3m (EcAIII alone) and 1seo (enzyme–product complex). However, the present analysis of crystal packing has revealed an important component of the solvent area that was not included in the original 1seo deposit. The structure has therefore been re-refined with proper inclusion of the newly identified L-Asp/$Ca^{2+}$ cluster in the solvent region. The new model has been submitted to the PDB with accession code 2zal, replacing the original 1seo structure. The present analysis is based on the new coordinate set, ignoring the original 1seo deposit.

The fifth coordinate set corresponds to an unpublished structure of a mutated variant of EcAIII in which the catalytic Thr179 residue has been replaced by alanine. The T179A mutant is catalytically inactive both in processing external substrates and in autocatalytic maturation. The T179A EcAIII protein therefore corresponds to an immature precursor with all 321 residues in one uninterrupted polypeptide chain. This coordinate set has been deposited in the PDB with code 2zak.

To avoid confusion connected with the different systems of chain identifiers used in the PDB files, we will use the following consistent scheme. The letter A (or B) represents a polypeptide chain corresponding to one precursor molecule. The inclusion of labels, as in A$\alpha$ and A$\beta$ (or B$\alpha$ and B$\beta$), defines the subunits $\alpha$ and $\beta$ created after autoproteolysis of the precursor molecule. In a mature EcAIII protein, the A$\alpha$, A$\beta$, B$\alpha$ and B$\beta$ chains would correspond to chain identifiers A, B, C and D, respectively.

### 2.2. Software used for structure comparison and analysis

The crystal contacts of the EcAIII molecules in all crystal forms were analyzed using the *WHAT IF* program (Vriend, 1990) and extended with calculations in *AREAIMOL* (Lee & Richards, 1971) from the *CCP*4 package (Collaborative Computational Project, Number 4, 1994) of the molecular surfaces buried on crystal packing. Superpositions of the molecular $C^\alpha$ traces were calculated in *LSQKAB* (Kabsch, 1976).

For TLS analysis, the *TLSMD* server (Painter & Merritt, 2006*a*,*b*; http://skuld.bmsc.washington.edu/~tlsmd/) was used to assess the best TLS model for each structure. The number of TLS groups in each case was determined from a manual analysis of the refinement results *versus* the number of TLS segments. Before submitting a given PDB file for *TLSMD* analysis, ten cycles of refinement without TLS parameters were calculated in *REFMAC*5 (Murshudov *et al.*, 1997) with the stereochemical weights (Engh & Huber, 2001) adjusted to produce an $R$ factor close to that in the original PDB deposit. This procedure was applied because some of the deposited structures were refined using incompatible TLS models. The result of the analysis indicated that it is generally beneficial to use more than one TLS segment per polypeptide chain. We decided to use eight segments per chain as a reasonable compromise between the number of parameters and the improvement of the refinement results.

### 2.3. Crystallization and structure determination of T179A EcAIII

The mutagenesis, overexpression and purification of the T179A variant of EcAIII will be described elsewhere. The crystallization of the mutant protein was carried out by the hanging-drop vapour-diffusion method at 292 K. The drops contained a 1:1 volume ratio of protein at 7 mg ml$^{-1}$ in a

solution composed of 50 m*M* Tris–HCl pH 8.5 and 100 m*M* NaCl and a range of precipitating buffers from Hampton Crystal Screen I. Crystals appeared within 1 d in a solution containing 200 m*M* $MgCl_2$, 100 m*M* Tris–HCl pH 8.5, 30%(*w*/*v*) PEG 4K. After optimization, the best single crystals were obtained with the PEG 4K concentration adjusted to 15%. X-ray diffraction data extending to 2 Å were collected at beamline X12 (EMBL) of the DESY synchrotron. The crystal was cryoprotected in a 1:1 mixture of the mother liquor and 50% PEG 400 prior to flash-cooling in a nitrogen-gas stream at 100 K. The data set was processed using *HKL*-2000 (Otwinowski & Minor, 1997). Data-collection and processing statistics are summarized in Table 2. The structure was solved by molecular replacement using *MOLREP* (Vagin & Teplyakov, 1997) with the *1k2x*[1] model as a probe. The asymmetric unit contains two precursor molecules (chains A and B) forming a dimer. The refinement was performed in *REFMAC*5 (Murshudov *et al.*, 1997) with manual corrections according to the electron-density maps performed in *Coot* (Emsley & Cowtan, 2004). The structure was refined to a final *R* value of 0.197 and an $R_{free}$ of 0.253. The model contains residues 2–163 and 178–313 for chain A, and 2–158 and 178–313 for chain B. The missing residues are not visible in the electron-density maps owing to disorder. This applies in particular to the residues preceding Gly178 (forming the C-terminal end of the '$\alpha$-subunit portion'), which are definitely present in this single-chain molecule but still display high flexibility as in the mature protein, in which this 'C-terminus' is free. The refinement statistics are given in Table 3.

### 2.4. Re-refinement of 1seo (2zal)

Careful analysis of difference electron-density maps revealed that the solvent region of the crystalline EcAIII–L-Asp complex contains a large volume of unassigned electron density. It could be confidently interpreted as an $(\text{L-Asp}^{2-})_3{\cdot}(Ca^{2+})_5$ cluster. The 1seo structure was re-refined (in *REFMAC*5) using the original diffraction data and including the cluster in the set of atomic coordinates. The statistics of the new refinement (Table 4), the temperature factors and the intramolecular and intermolecular geometry all confirm the correctness of this model (2zal).

**Table 4**
Re-refinement statistics for 2zal.

For comparison, the previous parameters are given in parentheses.

| | |
|---|---|
| Resolution (Å) | 20.0–1.9 (20.0–1.9) |
| No. of reflections in working set | 42200 (42200) |
| No. of reflections in test set | 1372 (1372) |
| $R$†/$R_{free}$ | 0.157/0.188 (0.164/0.189) |
| No. of atoms | |
| Protein | 4281 (4267) |
| Ligand | 61 (18) |
| Solvent | 248 (270) |
| $Na^+$ | 2 (6) |
| $Ca^{2+}$ | 5 (1) |
| $Cl^-$ | 1 (1) |
| R.m.s. deviations from ideality | |
| Bond lengths (Å) | 0.014 (0.014) |
| Bond angles (°) | 1.34 (1.39) |
| Average *B* factor (Å$^2$) | 13.3 (14.7) |
| Residues in Ramachandran plot (%) | |
| Most favoured regions | 92.9 (92.4) |
| Allowed regions | 7.1 (7.6) |

† $R = \sum_h \big||F_o| - |F_c|\big| / \sum_h |F_o|$ for all reflections, where $F_o$ and $F_c$ are the observed and calculated structure factors, respectively. $R_{free}$ was calculated analogously for the test reflections, which were randomly selected and excluded from the refinement.

## 3. Results

### 3.1. EcAIII structures deposited in the PDB

The X-ray structures of EcAIII deposited in the PDB (1k2x, 1jn9, 1t3m, 2zal and T179A) correspond to crystals obtained under crystallization conditions that differed in ionic composition (Table 1). The crystals of 2zal and 1jn9 were grown in the presence of calcium chloride, while for the crystallization of 1k2x, T179A and 1t3m magnesium ions were used with chloride and nitrate counterions, respectively. In the mother liquor of the 1t3m crystal a small amount of chloride ions were also present that were introduced during Tris buffer preparation. Among the crystal forms, the crystallization conditions for 2zal were unique in that the solution contained significant amounts of sodium L-aspartate. Since the L-Asp molecule, a product of the catalyzed reactions, is bound in the active site, the 2zal structure represents an EcAIII–L-Asp complex, while the other depositions correspond to EcAIII alone.

### 3.2. Inorganic ions in the crystal structures of EcAIII

Some of the ions identified in the crystal structures of EcAIII are buried within the protein core or are exposed on the surface but do not participate in crystal contacts (intramolecular ions), while others mediate contacts between symmetry-related protein molecules (intermolecular ions). Among the intramolecular ions, the most prominent position is held by a sodium cation that is coordinated in all EcAIII forms by a loop comprising residues 60–70 of each $\alpha$ subunit. Contrary to the prevalent pattern, in which the signature of the protein is expressed by the utilization of side chains in metal coordination, the octahedral coordination sphere of EcAIII is established exclusively by main-chain carbonyl groups. Application of the bond-valence method (Brese & O'Keeffe, 1991; Müller *et al.*, 2003) and *B*-factor analysis have allowed unambiguous assignment of $Na^+$ ions at all those sites. Moreover, the cavity formed by the 60–70 loop is specifically designed for sodium ions as even a high concentration of calcium (in the crystallization conditions of 1jn9) did not lead to $Na^+ \rightarrow Ca^{2+}$ replacement, even though the coordination geometry appears to be theoretically acceptable for calcium. Moreover, in the immature *T179A* molecule the $Na^+$ ion is bound by the same structural element, preserving the

[1] Throughout this article, a PDB code in roman font designates the complete crystal structure or the PDB deposition. Italic font is used to designate the protein molecule in the given crystal structure. The same convention applies to the T179A/*T179A* symbols used to designate the $P2_12_12_1$ crystal structure/molecule of the immature T179A mutant of EcAIII.

conformation adopted in the wild-type structures. The presence of the $Na^+$ coordination complex even in the crystals obtained in the absence of sodium ions in the crystallization solutions (1k2x, 1jn9 and 1t3m) as well as the structural context in which the ion binding occurs both suggest that the $Na^+$ ions may be coordinated at the protein-folding stage. The sodium-binding loop is located in close proximity to the active site, where it supports the proper positioning of the nucleophilic Thr179 residue, while its own conformation is stabilized by the central metal cation.

The other cases of intramolecular ions are less specific and clearly depend on the crystallization conditions. For example, the *1k2x* molecule binds seven chloride anions on its surface using residues Ser147A$\alpha$[2] and Ser147B$\alpha$, Arg104B$\alpha$, Arg262B$\beta$, Glu293A$\beta$ and Glu293B$\beta$, with the additional participation of water molecules. Two of those chloride anions are also bound by Arg262B$\beta$ and Glu293B$\beta$ in the *1jn9* structure. Intramolecular nitrate ions, which were utilized during the crystallization of the 1t3m form, were identified in two places. One of them is located near His119B$\alpha$ and Arg76A$\alpha$, while the other $NO_3^-$ ion occupies the active site near Arg207A$\beta$, mimicking the $\alpha$-carboxylate group of a substrate/product molecule, as in the EcAIII–L-Asp complex (2zal). However, the presence of these nitrate ions at full occupancy is doubtful as they have poor electron density and *B* factors that are much higher (70–85 Å$^2$) than the surrounding residues (15–20 Å$^2$). The *T179A* mutant molecule does not bind any intramolecular ions apart from the two $Na^+$ ions mentioned above.

Intermolecular chloride anions were found in all EcAIII structures except 1t3m. In both crystal forms grown in the presence of calcium ions (2zal and 1jn9), the $Ca^{2+}$ ions were found to mediate interactions between symmetry-related protein molecules. In particular, in 2zal they participate in the formation of the large cluster that also includes aspartate anions. In 1t3m, an intermolecular nitrate ion was found in a bridge between crystallographic copies of the protein molecule, but this ion, like the other $NO_3^-$ ions, has poor electron density and extremely high *B* factors. Details concerning ion-mediated lattice interactions are described in §3.4.

### 3.3. Orientation and position of the EcAIII molecule in the asymmetric unit

Although all the $P2_12_12_1$ X-ray structures of EcAIII deposited in the PDB have the same pattern of unit-cell parameters, they show a scatter (of up to 32%) that is dependent on the crystallization conditions (Table 1). In all cases, the asymmetric unit contains chains A and B. In the wild-type EcAIII structures they are divided into subunits A$\alpha$, A$\beta$, B$\alpha$ and B$\beta$ and organized into one functional $(\alpha\beta)_2$ dimer of heterodimers with an internal non-crystallographic twofold axis. In the *T179A* molecule, the unprocessed chains A and B form a homodimer corresponding to one dimer of heterodimers of the mature protein. Since in all the structures the polypeptide chain corresponding to the $\alpha$ subunit of the mature protein contains a constant and rigid sodium coordination complex, the $Na^+$ cations can be used to represent the entire AB molecule and its translation within the unit cell (Fig. 1). Moreover, the Na(B)—Na(A) vector, together with the pseudo-dyad of the heterotetramer, can be used to describe the orientation of the AB molecule in the unit cell. The orientation of the Na(B)—Na(A) vectors relative to the crystallographic **a**, **b** and **c** directions is defined in Table 1 by the angles $\varphi_1$, $\varphi_2$ and $\varphi_3$ (or their direction cosines), while the orientation of the NCS dyad is defined by the angles $\varepsilon_1$, $\varepsilon_2$ and $\varepsilon_3$, respectively.

For the purpose of this comparison, the atomic coordinates of the EcAIII molecules deposited in the PDB have been

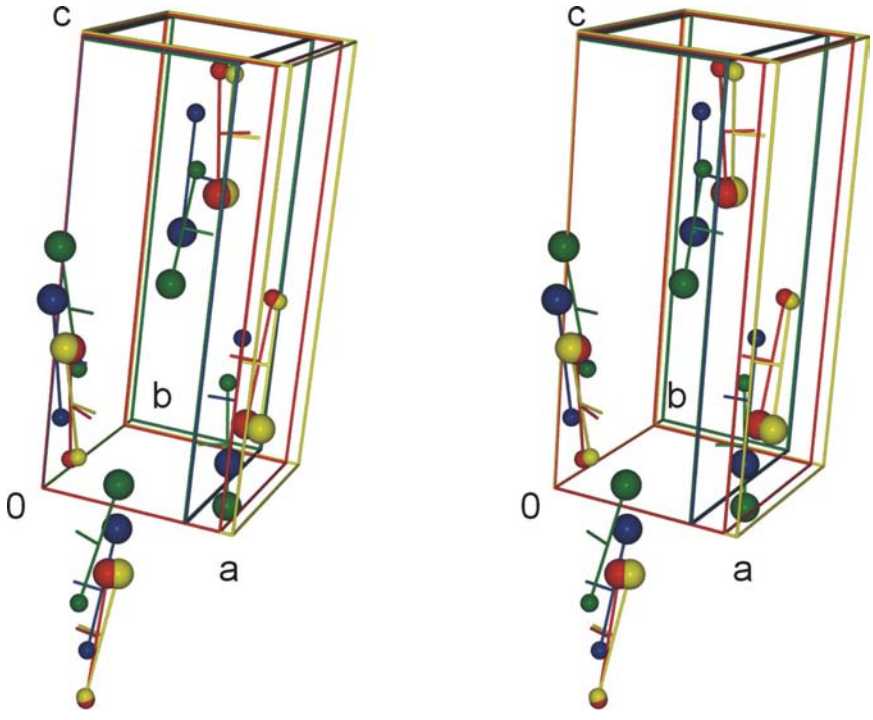


**Figure 1**
Crystal packing of EcAIII structures (stereo). The asymmetric unit contains molecules A and B (unprocessed in the case of the T179A mutant or processed into subunits $\alpha$ and $\beta$ in the remaining structures), creating one functional AB module. Each molecule is represented by a sphere corresponding to a sodium ion coordinated in a loop comprising residues Leu60–Ile70 in the $\alpha$-subunit sequence. The $Na^+$ ions coordinated by chain A are represented by a larger radius, while the ions coordinated by chain B have a smaller radius. The two spheres from one EcAIII molecule are connected to indicate the asymmetric unit contents. The short dash near the centre of each Na(B)—Na(A) vector represents the direction of the NCS dyad of the EcAIII molecule. The colour code is as follows: green, 2zal; blue, 1k2x; red, 1jn9; yellow, 1t3m. The green and blue unit cell boxes overlap exactly. The T179A molecule is not shown as it is exactly overlapped by 2zal.

[2] The numbering corresponds to structures 1k2x, 1jn9, 2zal and T179A. The symbol following the residue number represents the chain designation used in this paper.

examined in order to identify such a crystallographic copy of the protein molecule (designated as the principal position $x$, $y$, $z$) for which the midpoint of the Na(B)—Na(A) vector was closest to the origin and located in the all-positive octant of the unit cell. The 1t3m structure was chosen as a reference, because among all the cases the deposited *1t3m* molecule is already at position $x$, $y$, $z$. *2zal* could also be chosen, but it is located farther away from the origin.

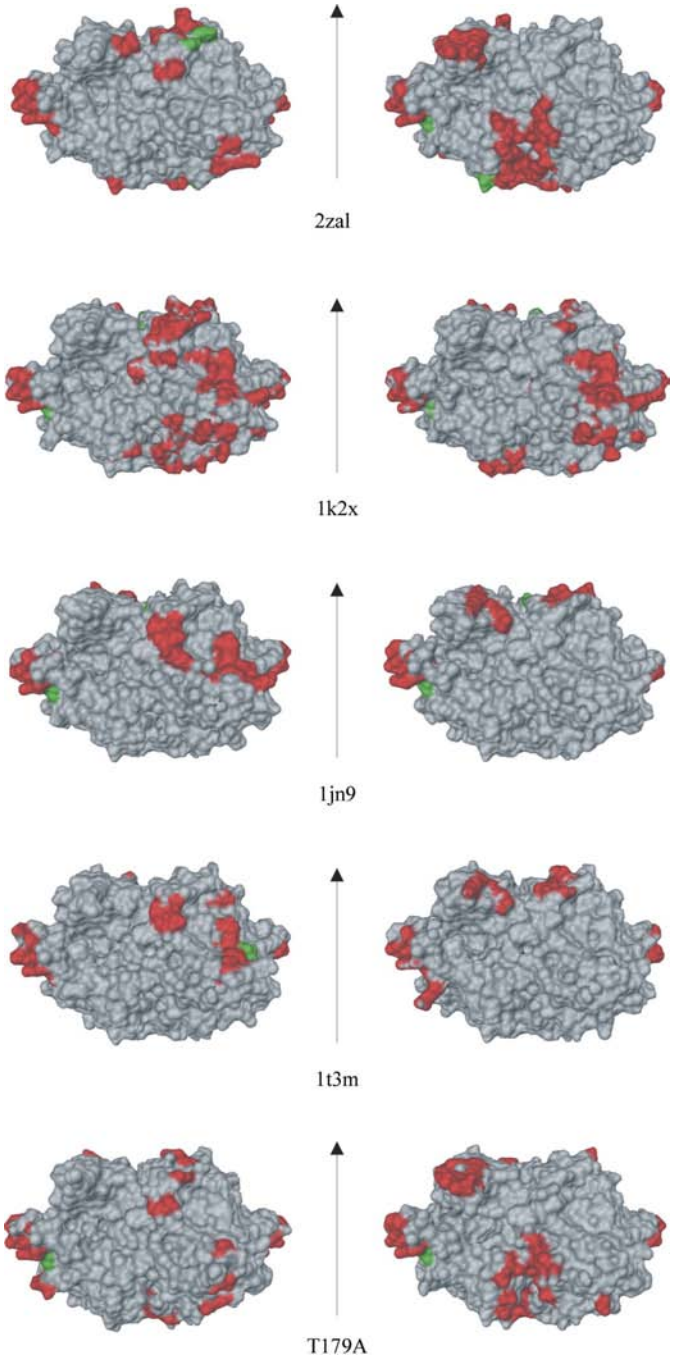


**Figure 2**
Crystal contacts mapped onto the solvent-accessible surface of EcAIII. For each protein, two orientations are shown differing by 180° rotation. The red colour shows residues participating in protein–protein interactions; the green colour represents residues interacting with symmetry-related ions.

The analysis reveals a clear similarity of the crystal-packing modes of 1jn9 and 1t3m, except for a **b**/2 translation, which could be attributed to a different choice of unit-cell origin. Accordingly, the deposited coordinates of *1jn9* have been translated by $-\mathbf{b}/2$. The case of structure 1k2x is more complicated because after the application of the $\frac{1}{2}-x, -y, \frac{1}{2}+z$ symmetry operation the molecule had to be shifted by $-\mathbf{a}/2$ to achieve a consistent origin. From the molecular-replacement solution of the T179A EcAIII precursor structure, it was clear that the packing of the molecules follows the pattern of 2zal but the molecular-replacement-generated model required shifting by $(\mathbf{a}+\mathbf{c})/2$.

To compare the translation and orientation of the EcAIII molecules in their respective unit cells, for each $x$, $y$, $z$ molecule represented by a pair of intramolecular $Na^+$ cations the orientation of the Na(B)—Na(A) vector (approximately normal to the NCS twofold axis) was determined and the distance of its midpoint from the origin was calculated (Table 1). The results show that in all cases the molecules have a very similar (within 6° of each other) orientation with respect to the **c** axis. The highest variation in the orientation of the molecules is observed with respect to the **a** axis, for which the $\varphi_1$ angle varies from 81° for *2zal* up to 91° for *1jn9*. The overall orientation of the *1jn9* molecule is very similar to that of the *1t3m* molecule, with the largest difference (5°) found for $\varphi_1$. The molecules *1t3m* and *1k2x* have a nearly identical $\varphi_1$ angle, but their $\varphi_3$ angles differ by 6°. Moreover, they adopt very different translations along two lattice directions (**a** and **c**). The unit-cell parameters of the 2zal and 1k2x crystals are almost the same, but the orientations of the molecules differ by up to 5° ($\varphi_1$) and they have very different translations along **c**. The orientation of the Na(B)—Na(A) vector of the *T179A* molecule is nearly the same as for *2zal*.

The molecular NCS dyad of the AB dimer is aligned with the crystallographic **a** axis within 26°. The $\varepsilon_2$ angle shows similar deviations and varies from 94° for *1k2x* to 114° for *1jn9*. The most conserved orientation of the dyad is observed with respect to the **c** axis, for which the $\varepsilon_3$ angle ranges from 84° (*2zal*) to 97° (*1jn9*). The NCS dyad of the *1k2x* molecule is distinct in that it is nearly perfectly aligned with the **a** axis.

### 3.4. Crystal contacts

A contact was detected if the distance between the atomic van der Waals spheres (Vriend, 1990) was less than 0.5 Å. The residues participating in interactions with symmetry-related molecules were then mapped onto the solvent-accessible protein surfaces (Fig. 2). A comparison of the patterns obtained clearly shows that there are a few residues that form the same crystal contacts in all structures. This group includes residues that participate in the interactions between molecules related by the screw axis along **c**: $x, y, z$ (1) and $\frac{1}{2}-x, -y, \frac{1}{2}+z$ (2) [and also $-x, \frac{1}{2}+y, \frac{1}{2}-z$ (3) and $\frac{1}{2}+x, \frac{1}{2}-y, -z$ (4)]. Specifically, Arg17A$\alpha$ from molecule (1) interacts with Met20B$\alpha$, Leu22B$\alpha$ and Glu25B$\alpha$ from molecule (2) and there are reciprocal interactions involving the same chains. Another common contact between (1) and (2) occurs with the partici-

**Table 5**
$Ca^{2+}$ coordination spheres in the 2zal structure.

Distances are in Å.

| Ca1 | | Ca2 | | Ca3 | | Ca4 | | Ca5 | |
|---|---|---|---|---|---|---|---|---|---|
| Asp1f $O^{\delta2}$ | 2.34 | Asp1f O | 2.53 | Glu258(A$\beta$) $O^{\varepsilon1}$ | 2.57 | Asp1f $O^{\delta1}$ | 2.40 | Asp1f $O^{\delta2}$ | 2.30 |
| Asp1f N | 2.52 | Asp3f $O^{\delta1}$ | 2.69 | Glu258(A$\beta$) $O^{\varepsilon2}$ | 2.48 | Asp1f $O^{\delta2}$ | 2.57 | Asp188(A$\beta$)† $O^{\delta1}$ | 2.33 |
| Asp1f O | 2.49 | Asp3f $O^{\delta2}$ | 2.52 | Asp2f‡ O | 2.37 | Asp3f O | 2.60 | Asp2f‡ $O^{\delta2}$ | 2.33 |
| Asp3f $O^{\delta2}$ | 2.43 | Asp188(A$\beta$)† $O^{\delta2}$ | 2.29 | Glu137(A$\alpha$)‡ $O^{\varepsilon1}$ | 2.51 | Gly253(A$\beta$) O | 2.26 | Asp2f‡ N | 2.53 |
| Asp3f O | 2.52 | Wat† | 2.37 | Glu137(A$\alpha$)‡ $O^{\varepsilon2}$ | 2.40 | Asp2f‡ OXT | 2.38 | Asp2f‡ OXT | 2.53 |
| Asp3f N | 2.52 | Wat† | 2.38 | Wat | 2.32 | Glu137(A$\alpha$)‡ $O^{\varepsilon1}$ | 2.37 | Gly135(A$\alpha$)‡ O | 2.36 |
| Wat | 2.41 | Wat† | 2.47 | Wat | 2.44 | Wat‡ | 2.50 | Wat‡ | 2.41 |
| Wat | 2.30 | | | Wat‡ | 2.39 | Wat | 2.39 | | |
| Bond valence§ | | | | | | | | | |
| 2.28 | | 1.92 | | 2.31 | | 2.36 | | 2.24 | |

† Residues from molecule at $\frac{1}{2}+x, \frac{1}{2}-y, -z$. ‡ Residues from molecule at $x-1, y, z$. § Calculated according to Müller *et al.* (2003)

pation of the C-termini of the $\beta$ subunits, but this interaction shows some variability and involves a $Cl^-$ ion in all structures except 1t3m.

All other intermolecular interactions differ in the EcAIII structures, except for the 1jn9/1t3m and 2zal/T179A pairs, which show a high degree of conservation. The close relationship between 1jn9 and 1t3m is not perturbed by the structure-specific ion-mediated interactions involving $Ca^{2+}$ in 1jn9 and $NO_3^-$ in 1t3m.

The intermolecular contacts in structures 2zal or T179A and 1k2x cannot be compared with either 1jn9 or 1t3m or even with each other, despite the close similarity of their unit-cell parameters. However, the 2zal and T179A structures have similar interactions involving the C-terminal $\alpha$-helix of chain A$\alpha$ from molecule (1), which forms a complex network of hydrogen bonds with chains A$\beta$, B$\alpha$ and B$\beta$ from molecule $x-\frac{1}{2}, \frac{1}{2}-y, 1-z$. In 1jn9 and 1t3m much less elaborate contacts of this helix are also found, but they connect molecule (1) with chain A$\alpha$ of molecule (3). In 1k2x there are similar crystal contacts, but there are also unique interactions.

In the mature EcAIII structures the linker peptide cannot be modelled because the C-terminus of subunit $\alpha$ is either degraded (Borek *et al.*, 2004) or disordered. Among these structures, the C-terminal helix of subunit $\alpha$ in 2zal shows the least degree of disorder. Its stabilization may be attributed to the interactions described above. One would expect that in the intact T179A precursor molecule it might be possible to trace the entire linker region connecting subunits $\alpha$ and $\beta$. Unfortunately, in this case the linker also has no well defined electron density and could not be traced much farther than in the 2zal structure. Only one of the two additional residues visible in the T179A structure in the C-terminal helix of the $\alpha$-subunit sequence participates in intermolecular interactions. This suggests that the linker does not participate in crystal contacts that would stabilize its intrinsically flexible conformation in any of the EcAIII structures.

In 2zal and T179A the N-terminus of chain A$\alpha$ (or A) also participates in the formation of crystal contacts. Additionally, in 2zal a Tris molecule bridges a chain of protein molecules generated by the lattice translation along **a**. Another unique feature of the 2zal structure is a set of lattice-stabilizing bonds that involve the $Ca^{2+}$ cations and the interstitial L-Asp molecules (see below).

### 3.5. Intermolecular L-aspartate/$Ca^{2+}$ ion cluster in the 2zal structure

Three intermolecular aspartate ions and five $Ca^{2+}$ ions form a complicated network of interactions connecting three symmetry-related protein molecules (Fig. 3) in the 2zal structure. Interestingly, there are only very few direct contacts between the L-Asp molecules and the protein. The coordination spheres of the calcium cations include seven or eight ligands. Chemically, the ligands are either O atoms contributed by the L-Asp anions, protein side-chain carboxylate groups, main-chain carbonyl groups and water molecules or N atoms from the aspartate molecules. The latter interactions demonstrate that the amino groups involved in calcium coordination cannot be protonated, which is in agreement with the pH of the crystallization buffer (8.5). Each of the calcium cations has a different coordination sphere, reflecting the plasticity of these cations as coordination centres (Table 5). Four modes of $Ca^{2+}$–$COO^-$ chelation can be distinguished. In the monodentate mode, the cation interacts with just one O atom from a carboxylate group, while in the bidentate mode both O atoms from the same $COO^-$ group act as ligands. In the $\alpha$ mode, the cation is chelated by one O atom from the $\alpha$-carboxylate and by the amino group in the $\alpha$ position. The fourth mode, termed here as $\alpha/\beta$, occurs when an additional electron-rich substituent is present at the $\beta$ position. This is the case for aspartate, where a carboxylate group is attached in the $\beta$ position. One $Ca^{2+}$ cation can therefore interact with as many as three functional groups of one L-Asp molecule.

Ca1 is coordinated by two water molecules and two L-Asp molecules (Asp1f and Asp3f), each of which acts as an $\alpha/\beta$ chelator. The coordination sphere of Ca2 is formed by three water molecules and three carboxylate groups, two of which are monodentate (the side chain of Asp188A$\beta$ and the $\alpha$-$COO^-$ group of Asp1f) and one bidentate ($\beta$-$COO^-$ from Asp3f). The Ca3 cation has similar ligands, but in this case only one carboxylate group is monodentate ($\alpha$-$COO^-$ from Asp2f), while two others are bidentate (the side chains of

Glu258A$\beta$ and Glu137A$\alpha$). Ca4 is surrounded by two water molecules, one main-chain carbonyl group (Gly253A$\beta$) and four carboxylate groups, one bidentate ($\beta$-COO$^-$ from Asp1f) and three monodentate ($\alpha$-COO$^-$ from Asp2f, $\alpha$-COO$^-$ from Asp3f and the side chain of Glu137A$\alpha$). The last calcium cation, Ca5, is chelated by Asp2f in $\alpha/\beta$ mode, two monodentate carboxylate groups (the side chain of Asp188A$\beta$ and $\beta$-COO$^-$ from Asp1f), one main-chain carbonyl group (Gly135A$\alpha$) and one water molecule. Ca3, Ca4 and Ca5 mediate interactions between protein molecules related by lattice translation along **a**.

Each of the intermolecular aspartate anions is involved in binding calcium cations. Asp1f chelates four cations, while Asp2f and Asp3f chelate three cations each. Because of the intricate system of connections, the L-Asp/Ca$^{2+}$ cluster resembles a salt crystal formed within a protein crystal. In the crystal structure of calcium aspartate, (L-Asp$^{2-}$)·(Ca$^{2+}$)·2H$_2$O (Schmidbaur *et al.*, 1989), the Ca$^{2+}$ cation is coordinated in $\alpha/\beta$ mode by one aspartate molecule, in bidentate mode by $\beta$-COO$^-$ from another crystallographic L-Asp ligand and in monodentate mode by $\beta$-COO$^-$ from a third aspartate. Two water molecules complete the octadentate coordination sphere. Each of the aspartate molecules interacts with three Ca$^{2+}$ cations. None of the cations present in the 2zal structure exactly reproduces these interactions. The closest analogue is Ca5, but its coordination sphere comprises only seven atoms and none of the ligands is bidentate.

With three −2 charged aspartate anions and five calcium cations, the total charge of the (L-Asp)$_3$/(Ca$^{2+}$)$_5$ cluster is +4. The interface formed by three protein molecules where the cluster is located contains three side-chain carboxylate groups, reducing the total charge in this area to +1.

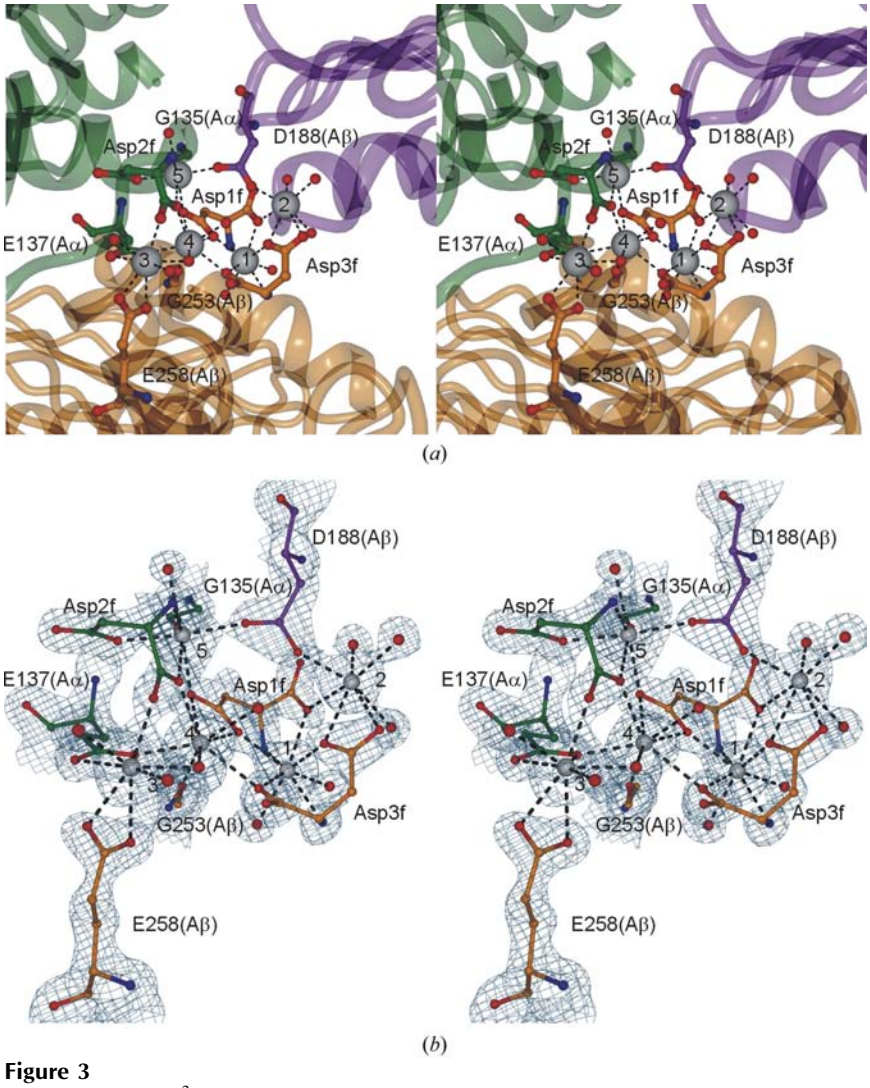


**Figure 3**
The L-aspartate/Ca$^{2+}$ cluster found between symmetry-related protein molecules in the 2zal structure (stereo). (*a*) The protein molecule at position $x, y, z$ is shown in orange, that at $\frac{1}{2}+x$, $\frac{1}{2}-y$, $-z$ in purple and that at $x-1, y, z$ in green. The symbols in parentheses indicate the polypeptide chain. The larger grey spheres represent calcium cations, while the smaller red spheres correspond to water molecules in the Ca$^{2+}$ coordination spheres. (*b*) A magnification of the L-Asp/Ca$^{2+}$ coordination complex shown in the same view as in (*a*), with the corresponding $2F_o - F_c$ map contoured at the 1.6$\sigma$ level.

### 3.6. Analysis of surface contacts

Analysis of the surface area buried on crystal packing (Table 6) shows that in all crystal forms except 1k2x chain A$\alpha$ is the main component of the packing interfaces, contributing 55% of all surface-contact area in the case of 1jn9 and 1t3m. Not surprisingly, chains A$\beta$ and B$\beta$, which form the core of the $(\alpha\beta)_2$ molecule, do not create extensive interactions (5–15%). Despite the internal symmetry of the heterotetramer, chain B$\alpha$ makes less extensive lattice contacts than chain A$\alpha$. The reason for this asymmetry is the imperfect (relative to crystallographic symmetry) orientation of the NCS twofold axis. 1k2x is unusual because the partition of the packing surfaces between chains A$\alpha$, A$\beta$, B$\alpha$ and B$\beta$ is almost perfectly symmetrical, with A$\alpha$ and B$\alpha$ each contributing one third of all interactions and A$\beta$ and B$\beta$ each contributing about 15% of the interaction surface. This symmetry follows from the fact that the NCS dyad in 1k2x is nearly parallel to **a**. At the same time, analysis of the individual contacts shows that this symmetry is achieved despite slightly different specific interactions (Supplementary Table S1[3]). In fact, combined analysis of the packing surfaces and buried surfaces contributed by individual residues shows that the packing of 1jn9 and 1t3m is very similar, which is in agreement with the NCS dyad directions, the similarity of their unit-cell parameters and the low r.m.s. deviations of their C$^\alpha$ super-

[3] Supplementary material has been deposited in the IUCr electronic archive (Reference: DZ5124). Services for accessing this material are described at the back of the journal.

**Table 6**
Surface area buried on packing interactions.

| | 1jn9 | 1k2x | 2zal | 1t3m | T179A |
|---|---|---|---|---|---|
| Total solvent-accessible area with generated symmetry-related atoms (Å$^2$) | 18443.5 | 17071.2 | 16486.0 | 18013.6 | 17693.4 |
| Total solvent-accessible area without considering symmetry-related atoms (Å$^2$) | 20659.8 | 20207.0 | 20286.0 | 20148.1 | 20778.6 |
| Total area differences owing to presence of symmetry-related atoms (Å$^2$) | −2216.3 | −3135.8 | −3800.3 | −2134.5 | −3085.2 |
| Chain A$\alpha$ area difference (Å$^2$)/% of total contact surface | −1211.3/54.6 | −1039.0/33.1 | −1602.5/42.2 | −1162.1/54.4 | −1961.3/63.6 |
| Chain A$\beta$ area difference (Å$^2$)/% of total contact surface | −115.7/5.2 | −471.4/15.0 | −714.4/18.8 | −103.5/4.8 | |
| Chain B$\alpha$ area difference (Å$^2$)/% of total contact surface | −628.9/28.4 | −1041.6/33.2 | −614.3/16.2 | −698.6/32.7 | −943.8/30.6 |
| Chain B$\beta$ area difference (Å$^2$)/% of total contact surface | −105.1/4.7 | −469.1/15.0 | −429.8/11.3 | −101.6/4.8 | |
| Heteroatom area difference (Å$^2$) | $Cl^-$ 403, −54.9<br>$Cl^-$ 405, −71.6<br>$Ca^{2+}$ 407, −1.0<br>$Ca^{2+}$ 408, −10.3<br>$Ca^{2+}$ 409, −17.1 | $Cl^-$ 803, −59.1<br>$Cl^-$ 804, −55.6 | Asp502, −159.9<br>Asp503, −44.9<br>Asp504, −39.4<br>Trs505, −71.9<br>Trs506, −28.6<br>$Ca^{2+}$ 803, −8.2<br>$Ca^{2+}$ 804, −7.5<br>$Ca^{2+}$ 805, −4.6<br>$Ca^{2+}$ 806, −19.7<br>$Cl^-$ 807, −54.6 | $NO_3^-$ 902, −51.1<br>$NO_3^-$ 903, −17.6 | Trs1, −105.9<br>$Cl^-$ 4, −39.7<br>$Cl^-$ 5, −34.5 |
| Heteroatoms (% of total contact-area difference) | 6.7 | 3.6 | 11.6 | 3.2 | 5.8 |

position (see below). It is interesting to note that almost 12% of the interaction surfaces of 2zal are created by heterogroups introduced by crystallization conditions.

### 3.7. Superposition of the EcAIII molecules

The available crystallographic models of EcAIII were superposed and the pairwise r.m.s. deviations between the equivalent $C^\alpha$ atoms (Table 7) were calculated in order to examine the conformational differences and to analyze whether there is any correlation with crystal contacts. As shown in Fig. 4, the most flexible region of the structure is the C-terminal $\alpha$-helix of chain A$\alpha$, which is involved in lattice stabilization in all crystals, albeit through different contacts. The equivalent $\alpha$-helix from chain B$\alpha$ does not participate in crystal contacts. It adopts essentially the same conformation in *1k2x*, *1jn9*, *1t3m* and *T179A*, while it is slightly translated in *2zal*. Another part of the protein with apparent flexibility comprises residues Arg17–Glu25 in both chains (A$\alpha$ and B$\alpha$), which participate in the same type of interactions with symmetry-related molecules in all structures. Interestingly, in terms of secondary structure, the residues Arg17–Gln19 represent either a $3_{10}$-helix (in chains A$\alpha$ and B$\alpha$ of *1t3m* and *t179a* and in chain A$\alpha$ of *1k2x* and *2zal*) or an unstructured coil (in chains A$\alpha$ and B$\alpha$ of *1jn9* and in chain B$\alpha$ of *1k2x* and *2zal*). There are also two loops (Ala72–His80 and Gly199–Leu204) that display visible plasticity. In chain B, these loops adopt the same conformation in molecules *1k2x*, *1jn9*, *1t3m* and *T179A*, while in chain A they only superpose well for *1jn9*, *1t3m* and *T179A*. These parts of the protein do not form any crystal contacts; however, it is clear that their conformation depends on the position of the neighbouring C-terminal helix from chain A$\alpha$ or B$\alpha$, which is determined by intermolecular interactions.

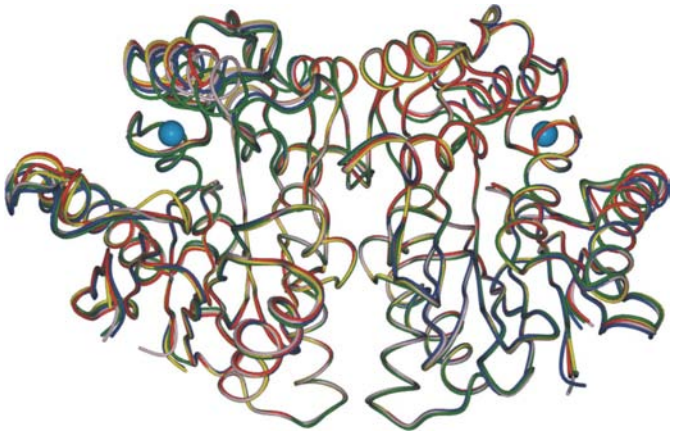

**Figure 4**
Superposition of the AB molecules of EcAIII found in the different crystal structures. Colour code: green, 2zal; blue, 1k2x; red, 1jn9; yellow, 1t3m; pink, T179A. The blue spheres represent the sodium cations coordinated by the loops comprising residues Leu60–Ile70.

The values of the r.m.s. deviations in Table 7 indicate that molecules *1t3m* and *1jn9* are identical within 0.39 Å, which is not surprising as these two variants of EcAIII adopt the same packing mode. *1k2x* shows almost the same level of similarity (about 0.5 Å) to all other polymorphs, but the agreement is not as good as for the *1t3m*/*1jn9* pair. The least superposable molecule is *2zal*, the closest analogue of which is *1k2x*. The overall conformation of the *T179A* precursor EcAIII is very similar to those of the mature proteins. However, despite the similar crystal packing of T179A and 2zal, the *T179A* molecule is more similar to the other EcAIII models than to *2zal*.

### 3.8. TLS analysis

The procedure of macromolecular refinement requires the selection of adequate model parameters, with an appropriate balance between the amount of experimental data and model complexity. At moderate resolution, an individual isotropic *B*-factor model provides a realistic data-to-parameter ratio.

**Table 7**
R.m.s. deviations (Å) between the corresponding $C^{\alpha}$ atoms of least-squares superposed EcAIII molecules.

The upper triangle is for AB/AB superpositions and the lower triangle for AB/BA superpositions. The diagonal shows superpositions within the AB molecules. The superpositions, which were calculated using *LSQKAB* (Kabsch, 1976), included residues 2–157 and 179–312.

| | 2zal | 1k2x | 1jn9 | 1t3m | T179A |
|---|---|---|---|---|---|
| 2zal | 0.61 | 0.55 | 0.72 | 0.66 | 0.65 |
| 1k2x | 0.72 | 0.54 | 0.54 | 0.51 | 0.48 |
| 1jn9 | 0.71 | 0.51 | 0.33 | 0.39 | 0.50 |
| 1t3m | 0.71 | 0.53 | 0.45 | 0.17 | 0.50 |
| T179A | 0.77 | 0.56 | 0.44 | 0.49 | 0.56 |

However, owing to the fractal nature of macromolecular structure, the use of only isotropic atomic displacement parameters is often not optimal. The conflict between the necessity for a more elaborate (*e.g.* anisotropic) model and the paucity of data can be solved, at least partially, by the introduction of anisotropic motion in the form of TLS (translation, libration and screw-rotation) parameters (Schomaker & Trueblood, 1968; Winn *et al.*, 2001), modelling concerted vibrations of larger groups of atoms. For the purpose of our analysis, we used an eight-segment TLS model (Painter & Merritt, 2006*a*,*b*).

There were small differences between the analyzed molecules regarding partitioning into TLS segments by the *TLSMD* server. However, despite the differences, the range of motion of the individual rigid-body fragments did not differ significantly (Supplementary Fig. S1). The TLS parameters reflect a balance between the area that is solvent-exposed and the area that is buried in packing interactions. A notable exception is the fragment Ser255–Arg262 of *1k2x*, which has a significantly different translational TLS component. We also note differences in the behaviour of loop Ala14–Ile29 and at the C-termini of chains A$\alpha$ and B$\alpha$ of all crystal forms. These fragments are solvent-exposed, which explains their relatively high TLS components, but the differences are probably connected with different interactions between symmetrical copies of the molecule, as it is this loop and the C-termini that are mostly responsible for packing interactions. Generally, a detectable correlation is observed between the number and type of packing interactions and the TLS behaviour.

## 4. Discussion and conclusions

In view of the above observations, the five EcAIII crystal structures deposited in the PDB should be treated as three polymorphic modifications (with allowance for the variations of the protein itself, solvent composition and L-Asp binding in the 2zal structure). This polymorphism (real or pseudo) can be attributed to the molecule's peculiar ability to pack with essentially the same lattice despite different translations and despite variations of the orientation of the molecular pseudo-dyad relative to the crystallographic screw axes, especially the screw axis along **a**, with which it is roughly aligned. The most consistent unit-cell parameter is *c*. The orientation of the molecules with respect to the **c** direction is also relatively stable. This common feature of all the structures reflects the similarity in the packing interactions along **c**, suggesting their importance for crystal lattice formation.

The variability or 'indeterminacy' of molecular translation and orientation observed in the EcAIII structures may give the term 'pseudopolymorphism' a different connotation, namely to mean that there may be a continuum of EcAIII crystal forms with a continuous variation of the position and orientation (**R**) of the asymmetric unit molecule. However, it appears that once a crystal seed with a given **R** has been formed, the crystal growth can preserve this selection and it is possible to form macroscopic crystals without translational/orientational disorder. It is of note on the other hand that crystallization experiments with EcAIII frequently lead to malformed or twinned crystals. It appears that the selection of a given 'polymorph' during crystallization may critically depend on slight variations of the crystallization conditions. However, it is difficult to speculate which component of the crystallization solutions plays the most crucial role in the selection of a given crystal-packing mode. In particular, the most similar 1jn9 and 1t3m structures were obtained using different mother liquors and therefore have different compositions of the solvent region within the crystals. It is also difficult to conclude why the crystal packing of T179A is different from that of 1k2x, as their crystals were obtained under very similar conditions. Instead, the T179A packing corresponds to that of 2zal, which has unique lattice interactions involving an additional (and clustered in a highly unusual way) complex of L-aspartate and calcium ions. The similarity between T179A and 2zal could be attributed to the presence of sodium cations in the mother liquor (either from storage buffer or from reservoir solution); however, such ions were not found beteween symmetry-related protein molecules in these crystals. On the other hand, the relatively high r.m.s. deviation of the $C^{\alpha}$ atoms of those two molecules suggests that a similar crystal-packing mode does not automatically guarantee the highest similarity of the molecular structures. The differences between the *T179A* and *2zal* molecules cannot be explained by the fact that *T179A* represents an immature precursor, as *T179A* shows the same degree of similarity (0.5 Å) to all the mature EcAIII molecules (except *2zal*) as the mature proteins show among themselves (Table 7). It is also unlikely that the binding of an aspartate molecule in the active site of 2zal is responsible for the slight modifications of the protein conformation because in such a case one would expect to observe, in the first case, some local alterations within the catalytic pocket that would be propagated to more distant parts of the molecule, something that is not observed here. However, we may speculate that the intermolecular L-Asp/$Ca^{2+}$ cluster in 2zal does influence the structure of the protein molecule. It is noteworthy that 2zal has the smallest unit-cell volume, despite the expanded composition of its contents. A comparison of the individual unit-cell parameters reveals that the *a* parameter of 2zal is smaller by 1.5 Å (3%) than that of T179A. It is possible that the presence of the positively charged L-Asp/$Ca^{2+}$ cluster attracts the protein molecules towards each other to saturate all possible inter-

actions. The stronger intermolecular interactions in turn result in a more 'strained' molecular conformation. The Tris molecule which occupies the site of the L-aspartate/$Ca^{2+}$ cluster in the T179A structure is not able to achieve the same effect and the molecules are less tightly packed. Although they generally have the same crystal contacts as in 2zal, the interactions are weaker and allow the protein molecule to adopt a more relaxed conformation. This explanation is supported by the data corresponding to buried area differences (Table 6), which show that *2zal* uses more of its available surface than *T179A* for crystal contacts. Moreover, the L-Asp/$Ca^{2+}$ cluster indeed links protein molecules that are related by the twofold screw axis along **a**.

McPherson & Cudney (2006) showed that multivalent molecules with negatively charged groups, especially carboxylates, are the most efficient compounds at triggering the crystallization of macromolecules by cross-linking charged side chains. The presence of a polar group in the molecule used as an additive may also be beneficial as it provides functionalities for anchoring hydrogen bonds. L-Aspartate, one of the components of the 2zal mother liquor, fulfils all the requirements for the perfect crystallization agent, so it is not surprising that it participates in lattice formation. However, one would expect that L-Asp would interact with positively charged protein side chains, which is not the case. Interestingly, it interacts with acidic residues *via* the bridging calcium ions. Thus, a counterion for L-Asp is needed for this type of binding. The extended network created by the L-Asp/$Ca^{2+}$ cluster is not a necessary condition for EcAIII crystallization but it provides an additional stabilization effect. The PEG component is a more likely factor that enhances nucleation and allows crystal growth.

With all the above analyses, it must be borne in mind that comparison of structural models based on different X-ray diffraction data sets may be sensitive to many external factors that could affect the final conclusion. Specific details of the crystallization experiment, cryoprotection, data-collection protocol and finally of all the stages of data analysis have their effect on the final model. At moderate resolution, it may be difficult to single out a particular parameter that caused a specific change and it may be even more problematic to answer the question about the significance of some structural differences; for instance, to decide whether different side-chain conformations between two models of the same protein are the consequence of a significant physical or chemical factor or simply stochastic fluctuations induced by cryogenic conditions.

The 2zal, T179A and 1k2x structures represent an interesting example of crystal forms for which the preservation of space-group symmetry and unit-cell isometry, which are the formal requirements of isomorphism, does not entail 'internal isomorphism' or isostructurality. A similar phenomenon has been reported for Lon protease, which crystallized in the $P6_5$ space group utilizing two distinct lattices which differed in the orientation of the molecule with respect to the sixfold screw axis (Dauter *et al.*, 2005). Another example of a similar non-isomorphism of protein crystals caused by a small rotation around a crystallographic axis was observed for the two forms of CysB, both with $P2_12_12$ symmetry and with the same unit-cell parameters (Verschueren *et al.*, 1999). For the crystals of human topoisomerase I, a continuum of structures can be formed in the $P2_1$ space group with unit-cell parameters that differ in the values of the $b$ and $\beta$ parameters (Redinbo *et al.*, 1999). Moreover, plant-type L-asparaginases are not the only example of enzymes with L-asparaginase activity for which the ability to form polymorphic crystals has been observed. Interestingly, bacterial-type L-asparaginases are also known to grow numerous polymorphs (Kozak *et al.*, 2002). In particular, the structures of periplasmic L-asparaginases from *E. coli* and *Erwinia chrysanthemi* crystallized in space groups $P6_522$ and $P6_122$, respectively, represent a rare case of quasi-enantiomorphic crystals with different packing modes (Jaskólski *et al.*, 2001).

This work was supported in part by a grant from the State Committee for Scientific Research (2 P04A 040 29) and by a subsidy from the Foundation for Polish Science to MJ. Some of the calculations were carried out in the Poznan Metropolitan Supercomputing and Networking Center.

## References

Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N. & Bourne, P. E. (2000). *Nucleic Acids Res.* **28**, 235–242.
Borek, D. (2001). PhD thesis. A. Mickiewicz University, Poznan, Poland.
Borek, D. & Jaskólski, M. (2000). *Acta Cryst.* D**56**, 1505–1507.
Borek, D. & Jaskolski, M. (2001). *Acta Biochim. Pol.* **48**, 893–902.
Borek, D., Michalska, K., Brzezinski, K., Kisiel, A., Podkowinski, J., Bonthron, D. T., Krowarsch, D., Otlewski, J. & Jaskolski, M. (2004). *Eur. J. Biochem.* **271**, 3215–3226.
Borek, D., Podkowinski, J., Kisiel, A. & Jaskolski, M. (1999). *Plant Physiol.* **119**, 1568–1569.
Brannigan, J. A., Dodson, G., Duggleby, H. J., Moody, P. C., Smith, J. L., Tomchick, D. R. & Murzin, A. G. (1995). *Nature (London)*, **378**, 416–419.
Brese, N. E. & O'Keeffe, M. (1991). *Acta Cryst.* B**47**, 192–197.
Bruneau, L., Chapman, R. & Marsolais, F. (2006). *Planta*, **224**, 668–679.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Dauter, Z., Botos, I., LaRonde-LeBlanc, N. & Wlodawer, A. (2005). *Acta Cryst.* D**61**, 967–975.
Duggleby, H. J., Tolley, S. P., Hill, C. P., Dodson, E. J., Dodson, G. & Moody, P. C. E. (1995). *Nature (London)*, **373**, 264–268.
Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.
Engh, R. A. & Huber, R. (2001). *International Tables for Crystallography*, Vol. F, edited by M. G. Rossmann & E. Arnold, pp. 382–392. Dordrecht: Kluwer Academic Publishers.
Groll, M., Ditzel, L., Löwe, J., Stock, D., Bochtler, M., Bartunik, H. D. & Huber, R. (1997). *Nature (London)*, **386**, 463–671.
Guan, C., Liu, Y., Shao, Y., Cui, T., Liao, W., Ewel, A., Whitaker, R. & Paulus, H. (1998). *J. Biol. Chem.* **273**, 9695–9702.
Guo, H. C., Xu, Q., Buckley, D. & Guan, C. (1998). *J. Biol. Chem.* **273**, 20205–20212.
Jaskólski, M., Kozak, M., Lubkowski, J., Palm, G. & Wlodawer, A. (2001). *Acta Cryst.* D**57**, 369–377.
Kabsch, W. (1976). *Acta Cryst.* A**32**, 922–923.
Khan, J. A., Dunn, B. M. & Tong, L. (2005). *Structure*, **13**, 1443–1452.

Kozak, M., Borek, D., Janowski, R. & Jaskólski, M. (2002). *Acta Cryst.* D**58**, 130–132.
Larsen, R. A., Knox, T. M. & Miller, C. G. (2001). *J. Bacteriol.* **183**, 3089–3097.
Lee, B. & Richards, F. M. (1971). *J. Mol. Biol.* **55**, 79–400.
Löwe, J., Stock, D., Jap, B., Zwickl, P., Baumeister, W. & Huber, R. (1995). *Science*, **268**, 533–539.
McPherson, A. & Cudney, B. (2006). *J. Struct. Biol.* **156**, 387–406.
Michalska, K., Brzezinski, K. & Jaskolski, M. (2005). *J. Biol. Chem.* **280**, 28484–28491.
Michalska, K., Bujacz, G. & Jaskolski, M. (2006). *J. Mol. Biol.* **360**, 105–116.
Michalska, K. & Jaskolski, M. (2006). *Acta Biochim. Pol.* **53**, 627–640.
Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* D**53**, 240–255.
Müller, P., Köpke, S. & Sheldrick, G. M. (2003). *Acta Cryst.* D**59**, 32–37.
Oinonen, C. & Rouvinen, J. (2000). *Protein Sci.* **9**, 2329–2337.
Oinonen, C., Tikkanen, R., Rouvinen, J. & Peltonen, L. (1995). *Nature Struct. Biol.* **2**, 1102–1108.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Painter, J. & Merritt, E. A. (2006*a*). *J. Appl. Cryst.* **39**, 109–111.
Painter, J. & Merritt, E. A. (2006*b*). *Acta Cryst.* D**62**, 439–450.
Prahl, A., Pazgier, M., Hejazi, M., Lockau, W. & Lubkowski, J. (2004). *Acta Cryst.* D**60**, 1173–1176.
Redinbo, M. R., Stewart, L., Champoux, J. J. & Hol, W. G. (1999). *J. Mol. Biol.* **292**, 685–696.
Schmidbaur, H., Bach, I., Wilkinson, D. L. & Müller, G. (1989). *Chem. Ber.* **122**, 1439–1444.
Schomaker, V. & Trueblood, K. N. (1968). *Acta Cryst.* B**24**, 63–76.
Seemüller, E., Lupas, A. & Baumeister, W. (1996). *Nature (London)*, **382**, 468–471.
Vagin, A. & Teplyakov, A. (1997). *J. Appl. Cryst.* **30**, 1022–1025.
Verschueren, K. H. G., Tyrrell, R., Murshudov, G. N., Dodson, E. J. & Wilkinson, A. J. (1999). *Acta Cryst.* D**55**, 369–378.
Vriend, G. (1990). *J. Mol. Graph.* **8**, 52–56.
Winn, M. D., Isupov, M. N. & Murshudov, G. N. (2001). *Acta Cryst.* D**57**, 122–133.