Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Structure of the isoaspartyl peptidase with L-asparaginase activity from *Escherichia coli*

Adam Prahl,[a] Marzena Pazgier,[a] Mahdi Hejazi,[b] Wolfgang Lockau[b] and Jacek Lubkowski[a]*

[a]Macromolecular Crystallography Laboratory, National Cancer Institute at Frederick, Frederick, MD 21702, USA, and [b]Institut für Biologie, Humboldt-Universität zu Berlin, D-10115 Berlin, Germany

Correspondence e-mail: jacek@ncifcrf.gov

Received 2 December 2003
Accepted 12 February 2004

PDB Reference: isoaspartyl peptidase, 1t3m, r1t3msf.

The crystal structure of the *Escherichia coli* enzyme (EcAIII) with isoaspartyl dipeptidase and L-asparaginase activity has been solved and refined to a resolution of 1.65 Å, with crystallographic *R*-factor and $R_{\rm free}$ values of 0.178 and 0.209, respectively. EcAIII belongs to the family of N-terminal hydrolases. The amino-acid sequence of EcAIII is homologous to those of putative asparaginases from plants. The structure of EcAIII is similar to the structures of glycosylasparaginases. The mature and catalytically active form of EcAIII is a heterotetramer consisting of two α-subunits and two β-subunits. Both of the equivalent active sites present in the EcAIII tetramer is assisted by a metal-binding site. The metal cations, modelled here as Na^+, have not previously been observed in glycosylasparaginases. This reported structure helps to explain the inability of EcAIII and other plant-type asparaginases to hydrolyze N^4-(β-*N*-acetylglucosaminyl)-L-asparagine, the substrate of glycosylasparaginases.

1. Introduction

During the sequencing the genome of the *Escherichia coli* strain K-12, a new gene, *ybiK*, was annotated as a putative asparaginase (Blattner *et al.*, 1997). The protein encoded by *ybiK*, referred to here as EcAIII, has 33% amino-acid sequence identity to *Flavobacterium meningosepticum* glycosylasparaginase (FmG) and 43% identity to the plant enzyme from *Lupus luteus*. EcAIII was subsequently overexpressed and crystallized (Borek & Jaskolski, 2000) and shown to have L-asparaginase activity. The same authors also reported that EcAIII undergoes autoproteolysis in a similar manner to glycosylasparaginases, becoming an N-terminal (Nth) hydrolase with a Thr residue as the nucleophile. Subsequently, Hejazi *et al.* (2002) found that EcAIII possesses isoaspartyl activity (hydrolyzing β-aspartyl dipeptides to L-aspartic acid and a second amino acid), whereas it is unable to hydrolyze N^4-(β-*N*-acetylglucosaminyl)-L-asparagine (GlcNAc-Asn), the primary substrate of the related glycosylasparaginases. The amino-acid sequence similarity and substrate specificity led to the classification of EcAIII as a plant-type asparaginase. Compared with glycosylasparaginases, plant-type enzymes have been less extensively studied (Lea *et al.*, 1978, 1984; Lea & Miflin, 1980; Dickson *et al.*, 1992; Lough, Chang *et al.*, 1992; Lough, Reddington *et al.*, 1992; Borek *et al.*, 1999). It is known that in leguminous plants they play a role in the processes associated with the assimilation of atmospheric nitrogen (Dickson *et al.*, 1992; Lough, Chang *et al.*, 1992; Borek *et al.*, 1999). Recent reports include the results of activity studies for the *Salmonella enterica* enzyme (Larsen *et al.*, 2001) and homologous proteins from *Synechocystis* sp. PCC 6803, *Anabaena* sp. PCC 7120, and *Arabidopsis thaliana* (Hejazi *et al.*, 2002). Whereas structures of human glycosylasparaginase (PDB code 1apz) and FmG (PDB code 1ayy) have been reported (Oinonen *et al.*, 1995; Guo *et al.*, 1998), structural information on plant-type enzymes is quite limited. In addition to the preliminary crystallographic report on EcAIII (Borek & Jaskolski, 2000), two structures of this enzyme have recently been deposited in the Protein Data Bank (PDB codes 1jn9 and 1k2x); however, they were not accompanied with published analyses. In this report, we present the crystal structure of EcAIII at 1.65 Å resolution.

2. Experimental procedures

2.1. Expression, purification and oligomerization analysis

Cloning and expression experiments were performed as described previously (Hejazi *et al.*, 2002). After lysis of the *E. coli* cell pellet, the EcAIII present in the supernatant was captured using a HiTrap chelating HP column (Pharmacia) and eluted with an imidazole gradient. Fractions containing EcAIII were dialyzed against 20 m*M* Tris, 5 m*M* EDTA buffer pH 8.5, applied onto a DEAE Sepharose FF column (Pharmacia) and eluted with an NaCl gradient. The final purification

Table 1
Data-collection and refinement statistics for EcAIII.

Values in parentheses correspond to the highest resolution shell.

Data collection	
Resolution range (Å)	30.0–1.65 (1.71–1.65)
R_{merge}	0.086 (0.528)
No. independent reflections	84176
Redundancy	4.7
Completeness (%)	98.9 (97.2)
Average $I/\sigma(I)$	12.1 (1.99)
Refinement	
No. reflections	
Working set	70692
Test set	3982
Resolution range (Å)	10.0–1.65
R_{work}/R_{free}	
All reflections	0.173/0.216
For $F > 4\sigma(F)$	0.163/0.204
No. non-H atoms	5017
No. water molecules	681
No. heteroatoms	14
Average B factor (Å^2)	
All non-H atoms	26.9
Protein atoms	
No. discretely disordered residues	21
R.m.s.d.s from ideality	
Bonds (Å)	0.009
Angle distances (Å)	0.025
Zero chiral volumes (Å^3)	0.045
Non-zero chiral volumes (Å^3)	0.058

was performed by size-exclusion chromatography using a Sephacryl S-200 FF column (Pharmacia). The identity of pure EcAIII was confirmed by SDS–PAGE, mass spectrometry and N-terminal sequencing of either of the two chains. Using gel electrophoresis under non-denaturing conditions, dynamic light scattering and size-exclusion chromatography with reference proteins, we consistently found that EcAIII forms heterotetramers in solution (results not shown).

2.2. Crystallization, X-ray data collection and processing

Crystallization experiments were carried out at room temperature using the hanging-drop vapour-diffusion method. Droplets were prepared by mixing equal volumes of protein (5 mg ml^{-1} in 20 m*M* HEPES buffer pH 7.5) and reservoir solution. After optimization, prism-shaped crystals (dimensions 0.5 × 0.2 × 0.1 mm) were obtained from solution containing 14%(*w*/*v*) PEG 4000, 15%(*v*/*v*) glycerol, 0.3 *M* $Mg(NO_3)_2$, 0.1 *M* bis-Tris–HCl pH 6.5. X-ray diffraction data were collected using a synchrotron-radiation source (beamline 22BM on the SER-CAT station at the Advanced Photon Source, Argonne National Laboratory, IL, USA), with the wavelength tuned to 0.98 Å, and intensities were recorded on a MAR CCD detector (X-ray Research, Germany). Data reduction and scaling were performed using the *HKL*2000 program (Otwinowski & Minor, 1997) and the identity of the space group was unambiguously deduced from the systematic extinctions observed for the scaled intensities. Crystals of EcAIII belong to space group $P2_12_12_1$, with unit-cell parameters $a = 66.32$, $b = 71.64$, $c = 149.58$ Å. The X-ray data-collection and processing statistics are shown in Table 1.

2.3. Structure solution and refinement

The structure of EcAIII was solved by the molecular-replacement method using the *AMoRe* program (Navaza, 1993). The search model was based on the heterotetramer of FmG (PDB code 1ayy) after replacing all residues other than Gly and Ala by Ser. Molecular replacement in the resolution range 14.0–3.5 Å gave a clear solution characterized by values of the correlation coefficient and the crystallographic R factor of 0.494 and 0.449, respectively. The bulk of the structure refinement was completed with the program *CNS* (Brünger *et al.*, 1998) using all experimental intensities. R_{free} was calculated based on 5.3% of randomly selected data excluded from refinement. This protocol resulted in the building of a complete model that, in addition to the protein, included nearly all solvent molecules and two structurally equivalent Na^+ ions. During the final calculations using the program *SHELXL* (Sheldrick & Schneider, 1997), the B factors of all S atoms and the Na^+ ions were refined using the anisotropic model. The final model refined at 1.65 Å resolution corresponds to R-factor and R_{free} values of 0.173 and 0.216, respectively. The refinement statistics are shown in Table 1.

3. Results

3.1. General description of the structure

The quaternary structure of EcAIII is very similar to structures of glycosylasparaginases (Oinonen *et al.*, 1995; Guo *et al.*, 1998). The asymmetric unit of EcAIII consists of a heterotetramer formed of two α-subunits and two β-subunits. In the final model, each heterotetramer contains 158 residues (1–158) in each of the α-subunits and 134 residues (179–312) in each of the β-subunits. No electron density is observed for the last 20 residues of the α-subunits and for the last 12 residues of the β-subunits. The core of each heterodimer, which bears the active site of the enzyme, has an $\alpha\beta\beta\alpha$-sandwich topology (Fig. 1), which is typical for Nth hydrolases (Artymiuk, 1995; Brannigan *et al.*, 1995; Oinonen & Rouvinen, 2000). Within this core, the two β-sheets are packed against each other and flanked by layers of α-helices. This mixed β-sheet consists of eight β-strands ($S_\alpha4$, $S_\alpha3$, $S_\alpha2$, $S_\beta2$, $S_\beta1$, $S_\alpha1$, $S_\beta7$ and $S_\beta8$; subscripts annotate the subunit location) with topology 3 1

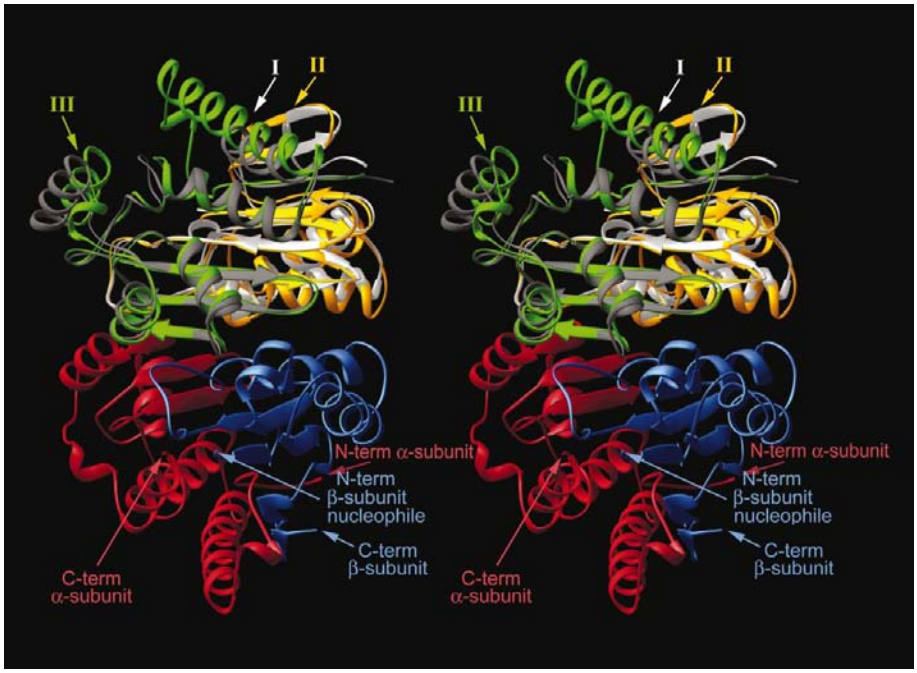

Figure 1
Stereo representation of the EcAIII heterotetramer. The approximate twofold axis relating two heterodimers (shown in red and blue and in green and yellow) of EcAIII lies near the plane of this figure. The upper dimer is superimposed on the corresponding motif of FmG (shown in white and grey). Three regions, indicated I, II and III, are structurally most divergent for EcAIII and FmG. Particularly striking is the difference between the first α-helices in the α-domains of both structures.

1X −4 1 −3 −1 (as determined by the program *PROMOTIF*; Hutchinson & Thornton, 1996). Except for $S_\alpha4$ and $S_\alpha3$, all the β-strands are antiparallel. The second β-sheet is formed of four antiparallel β-strands ($S_\beta3$, $S_\beta4$, $S_\beta5$ and $S_\beta6$) with topology 1 1 1. While the larger sheet is flanked by five α-helices that are located within the α-subunit, the three α-helices from the β-subunit form the layer adjacent to the smaller four-stranded β-sheet. The heterotetramer of EcAIII is stabilized by a number of hydrogen bonds and hydrophobic contacts between two α/β-dimers.

3.2. The metal-ion sites

Inspection of the positive $F_o - F_c$ electron-density map revealed two equivalent well defined spherical peaks (∼30 standard deviations in height) located in the two α-subunits. As shown in Fig. 2, the direct environment of each peak is formed of six main-chain O atoms (Leu60, Glu61, Cys63, Phe66, Ala68 and Ile70) at distances ranging between 2.3 and 2.8 Å from its centre, excluding water, a hydrated proton or a negatively charged ion as the occupant of this site. Despite the significance of magnesium ions in the crystallization of EcAIII, it is unlikely that these sites are occupied by Mg^{2+}. In proteins, Mg^{2+} ions are usually coordinated by electronegative octahedrally arranged atoms at distances of 2.0–2.1 Å from the central metal (Castagnetto *et al.*, 2002). Although our X-ray data do not allow us to unambiguously resolve the identity of the metal ion, preliminary refinements with various metal ions modelled for this peak indicated that the atomic number of the actual metal ion should be comparable to that of Mg^{2+}. Taking into account the distances from the metal ion to coordinating O atoms, the geometry of the coordination sphere (Nayal & Di Cera, 1996; Harding, 2002) and its likely atomic number, we modelled this ion as Na^+. Although the exact role of Na^+ is not quite clear, it appears that the conformation of the fragment Leu60–Ala68 must be stabilized by a small cation buried within it. Since the overall folds of EcAIII, FmG and human glycosylasparaginase are quite similar, we compared the Na^+ site in EcAIII with equivalent fragments of the other two enzymes. Interestingly, we found that the conformations of the equivalent fragments in EcAIII and FmG are very similar (including five of the O atoms coordinating the Na^+ ion in EcAIII; see Fig. 2), yet the presence of a metal ion was not reported in FmG (Guo *et al.*, 1998). In turn, the conformation of this fragment in human glycosylasparaginase is somewhat different and is stabilized by a disulfide bond (Oinonen *et al.*, 1995). These observations further support the role of Na^+ ions in structure stabilization of EcAIII.

3.3. The active sites of EcAIII: comparison with glycosylasparaginses

As shown in Fig. 2, the topology of the active site in EcAIII is very similar to that previously published for FmG and for human glycosylasparaginase (not shown). The deep funnel-like active-site cavity of EcAIII is located between the two subunits of the heterodimer. The nucleophilic centre, Thr179, is surrounded by several residues, including Thr197, Arg207, Asp210, Thr230 and Gly231, which are conserved in both glycosyl and plant-type asparaginases. These residues were shown to be important for catalysis in glycosylasparaginases (Oinonen *et al.*, 1995; Guo *et al.*, 1998) and they are likely to play a similar role in EcAIII. Two residues, Trp11 and Phe13 (in FmG), which are close to the catalytic Thr in glycosylasparaginases and participate in binding the carbohydrate fragment of the substrate, are not found in the N-terminal regions of the plant-type asparaginases. The extensive structural similarity between the active sites in EcAIII and glycosylasparaginases, however, does not explain the differences in the substrate specificities of both families of enzymes, particularly the inability of EcAIII and other plant-type asparaginases to proteolytically degrade asparagine-linked glycoproteins (Hejazi *et al.*, 2002). A possible answer may be the size of the active-site cavity, which in EcAIII is substantially smaller than in glycosylasparaginases.

Comparison of the active-site cavities in EcAIII and FmG (Fig. 1) indicates three clear structural differences between these two enzymes. The locations of the last two β-strands in the β-subunit and the last α-helix in the α-subunit are closer to the centre of the active site in EcAIII (marked II and III in Fig. 1). Particularly striking, however, is the size of the first α-helix in the α-subunit, which is significantly longer in EcAIII when compared with FmG (marked I in Fig. 1). Therefore, the smaller active site of EcAIII is unable to accommodate glycoprotein substrates that carry a bulky sugar component. In order to understand the basis of affinity of the plant-type asparaginases for the isoaspartyl dipeptidase substrates (Hejazi *et al.*, 2002), additional structural and mutation experiments are needed.

3.4. Comparison with other EcAIII structures

During the writing of this report, two structures of EcAIII refined at resolutions of 2.3 Å (PDB code 1jn9) and 1.65 Å (PDB code 1k2x) were deposited in the Protein Data Bank by Jaskólski and coworkers,

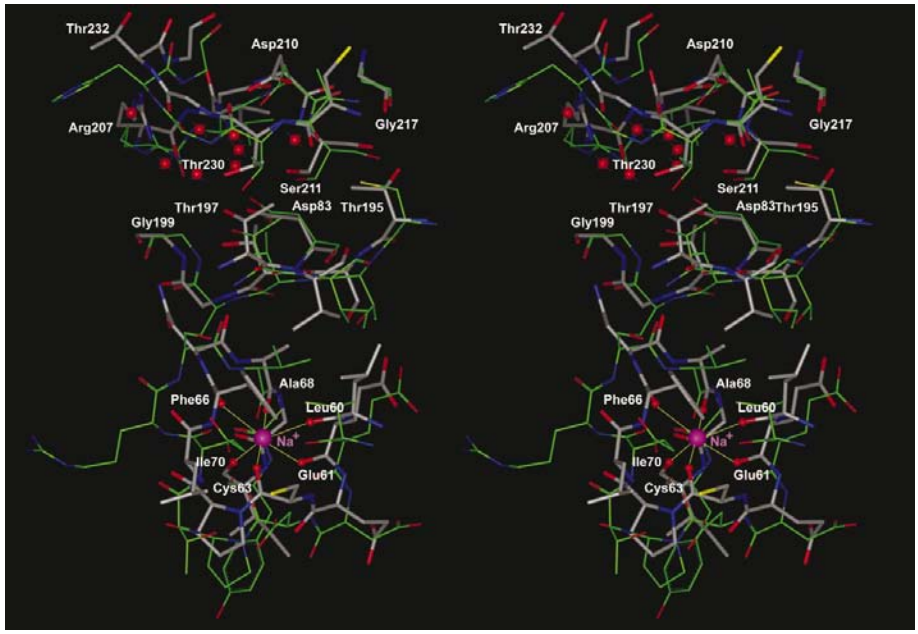

Figure 2
The superposition of the active sites of EcAIII (grey) and FmG (green). The adjacent metal cation-binding site in EcAIII is also shown. The active-site residues in EcAIII implicated to be important for the activity of the glycosylasparaginases (including Thr197, the N-terminal residue in the β-subunit) are labelled, as are the residues that form the Na^+-binding site. Water molecules are shown as small red spheres.

although no published reports are available. These structures correspond to different crystal forms of EcAIII. When superimposed, the deposited structures are nearly identical to that reported by us (root-mean-square deviations for equivalent atoms vary between 0.15 and 0.28 Å depending on the selection of atoms). Only a few minor differences can be found between these structures and that described here. The deposited structures differ in the presence of Ca^{2+} and Cl^{-} ions (not identified by us), one additional residue at the C-terminus of each of the α-subunits, slight conformational differences of the termini (but not the catalytic Thr179) and a few solvent-exposed turns and modification of the side chain of Cys63 (not observed in our structure). Interestingly, the authors of the deposited structures also modelled Na^{+} ions (not Mg^{2+}) in the metal-binding sites adjacent to the active sites of EcAIII.

References

Artymiuk, P. J. (1995). *Nature Struct. Biol.* **2**, 1035–1037.

Blattner, F. R., Plunkett, G. III, Bloch, C. A., Perna, N. T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J. D., Rode, C. K., Mayhew, G. F., Gregor, J., Davis, N. W., Kirkpatrick, H. A, Goeden, M. A., Rose, D. J., Mau, B. & Shao, Y. (1997). *Science*, **277**, 1453–1474.

Borek, D. & Jaskolski, M. (2000). *Acta Cryst.* D**56**, 1005–1007.

Borek, D., Podkowinski, J., Kisiel, A. & Jaskolski, M. (1999). *Plant. Physiol.* **119**, 1568.

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.

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.

Castagnetto, J. M., Hennessy, S. W., Roberts, V. A., Getzoff, E. D., Tainer, J. A. & Pique, M. E. (2002). *Nucleic Acids Res.* **30**, 379–382.

Dickson, J. M. J. J., Vincze, E., Grant, M. R., Smith, L. A., Rodber, K. A., Farnden, K. J. F. & Reynolds, P. H. S. (1992). *Plant Mol. Biol.* **20**, 333–336.

Guo, H. C., Xu, Q., Buckley, D. & Guan, C. (1998). *J. Biol. Chem.* **273**, 20205–20212.

Harding, M. M. (2002). *Acta Cryst.* D**58**, 872–874.

Hejazi, M., Piotukh, K., Deutzmann, R., Volkmer-Engert, R. & Lockau, W. (2002). *Biochem. J.* **364**, 129–136.

Hutchinson, E. G. & Thornton, J. M. (1996). *Protein Sci.* **5**, 212–220.

Larsen, R. A., Knox, T. M. & Miller, C. G. (2001). *J. Bacteriol.* **138**, 3089–3097.

Lea, P. F., Farnden, L. & Miflin, B. J. (1978). *Phytochemistry*, **17**, 217–222.

Lea, P. F., Festenstein, G., Hughes, S. J. & Miflin, B. J. (1984). *Phytochemistry*, **3**, 511–514.

Lea, P. F. & Miflin, B. J. (1980). *The Biochemistry of Plants*, Vol. 5, pp. 290–293. New York: Academic Press.

Lough, T. J., Chang, K. S., Carne, A., Monk, B. C, Reynolds, P. H. & Farnden, K. J. (1992). *Phytochemistry*, **31**, 1519–1527.

Lough, T. J., Reddington, B. D., Grant, M. R., Hill, D. F., Reynolds, P. H. & Farnden, K. J. (1992). *Plant Mol. Biol.* **19**, 391–399.

Navaza, J. (1993). *Acta Cryst.* D**49**, 588–591.

Nayal, M. & Di Cera, E. (1996). *J. Mol. Biol.* **256**, 228–234.

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.

Sheldrick, G. M. & Schneider, T. R. (1997). *Methods in Macromolecular Crystallography*, edited by D. Turk & L. Johnson, pp. 72–81. Amsterdam: IOS Press.