Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Crystallization and preliminary crystallographic studies of a new L-asparaginase encoded by the *Escherichia coli* genome

Dominika Borek[a]* and Mariusz Jaskólski[a,b]

[a]Department of Crystallography, Faculty of Chemistry, A. Mickiewicz University, Grunwaldzka 6, 60-780 Poznan, Poland, and [b]Center for Biocrystallographic Research, Institute of Bioorganic Chemistry, Polish Academy of Sciences, Noskowskiego 12/14, 61-704 Poznan, Poland

Correspondence e-mail: borek@leucine.ibch.poznan.pl

Received 19 June 2000
Accepted 19 July 2000

A new *Escherichia coli* L-asparaginase belonging to the class of Ntn amidohydrolases has been crystallized using the vapour-diffusion method and PEG 4000 as the precipitant. The crystals belong to the orthorhombic space group $P2_12_12_1$ (unit-cell parameters a = 50.3, b = 77.6, c = 148.2 Å) and diffract to 1.65 Å resolution. The structure has been solved by molecular replacement using aspartylglucosaminidase from *Flavobacterium meningosepticum* as the search model. The asymmetric unit contains four protein chains composed into a dimer of $\alpha\beta$ heterodimers, where the subunits α and β are the product of autoproteolytic cleavage of the immature protein.

1. Introduction

L-Asparaginases (E.C. 3.5.1.1) are hydrolases which catalyze the breakdown of L-asparagine to L-aspartate. In *E. coli* and in many other microbes (Bonthron & Jaskólski, 1997), two L-asparaginases have been reported, probably sharing the same catalytic mechanism involving a catalytic triad, TKD (Rao *et al.*, 1996), that could be considered a variant of the canonical triad characterizing serine proteases (SHD). There is also another group of enzymes with L-asparaginase activity, glycosylasparaginases (E.C. 3.5.1.26), belonging to a novel class of hydrolases which use a processed N-terminal threonine, serine or cysteine as both a polarizing base and a nucleophile in the catalysis (Brannigan *et al.*, 1995). Glycosylasparaginases, or aspartylglucosaminidases, cleave the β-aspartylglucosamine bond of asparagine-linked oligosaccharides, converting the asparagine residue to aspartic acid, and require for their activity both a free α-amino and α-carboxy group on the Asn substrate (Tanaka *et al.*, 1973; Kaartinen *et al.*, 1992). Apart from their primary activity, they also show a weaker L-asparaginase activity (Noronkoski *et al.*, 1997; Liu *et al.*, 1998). cDNA encoding glycosylasparaginases has been found in a number of species (Fisher *et al.*, 1990; Liu *et al.*, 1996; Tarentino *et al.*, 1995; Tenhunen *et al.*, 1995). Another intriguing aspect of these enzymes is that a single-chain precursor is processed by intramolecular proteolysis to generate the newly exposed N-terminal threonine and an active Ntn hydrolase comprised of α and β subunits (Guan *et al.*, 1996). Cleavage is essential for glycosylasparaginase activity and occurs at the amide bond in the consensus site (Ikonen *et al.*, 1993; Fisher *et al.*, 1993; Tarentino *et al.*, 1995) before a threonine residue (Fig. 1). This conserved threonine is thought to play a central role in both hydrolase activities [autoproteolysis (Guan *et al.*, 1996) and final hydrolytic activity (Kaartinen *et al.*, 1991; Fisher *et al.*, 1993)]. In addition, it is postulated that plant asparaginases also have an evolutionary relationship to glycosylasparaginases (Lough *et al.*, 1992; Aronson, 1996; Fig. 1).

In 1997, an ORF corresponding to a gene *ybik* was identified in the *E. coli* genome and deposited as putative asparaginase (Blattner *et al.*, 1997). The corresponding amino-acid sequence has 31% identity and 50% homology to *F. meningosepticum* glycosylasparaginase (FmG) and 43% identity and 60% homology to *Lupinus luteus* L-asparaginase (Borek *et al.*, 1999). This putative protein has been overexpressed and it has been demonstrated that it undergoes a proteolytic cleavage exposing T179 (Fig. 1) and that the final product has L-asparaginase activity. This paper reports the crystallization and preliminary crystallographic characterization of the mature protein EcAIII.

2. Materials and methods

2.1. Cloning and purification

The DNA coding for EcAIII was amplified from *E. coli* JM108-K12 strain by PCR and cloned into the pET11d vector (Novagen) (Borek, unpublished results). The construct was transferred into the expression host *E. coli* BL21(DE3) and protein expression was induced by the addition of isopropyl-β-D-thiogalactoside. The bacteria were harvested by centrifugation and stored at 253 K. The protein was purified from the soluble fraction

Table 1
X-ray data-collection statistics.

Values for the last resolution shell, 1.68–1.65 Å, are given in parentheses.

Space group	$P2_12_12_1$
Unit-cell parameters (Å)	a = 50.3, b = 77.6, c = 148.2
Crystal dimensions (mm)	0.2 × 0.2 × 0.1
Mosaicity (°)	0.5
Temperature (K)	100
Resolution range (Å)	20.0–1.65
No. of observations	736323
No. of unique reflections	70249
Completeness (%)	99.4 (89.9)
$\langle I/\sigma(I)\rangle$	17.2 (2.5)
R_{sym} (%)	7.5 (47.6)

of the disrupted cells by fractionating with polyethylene glycol (PEG) 8000, column chromatography on DEAE cellulose DE-52 (Whatman) and gel filtration on Sephacryl 300 (Pharmacia).

2.2. Crystallization

All crystallization trials were carried out at room temperature using the hanging-drop vapour-diffusion method with droplets prepared by mixing equal volumes of protein and reservoir solutions. The initial trials were performed using the commercially available sparse-matrix Crystal Screen kits (Jancarik & Kim, 1991). The protein concentration (determined by the Bradford method; Bradford, 1976) for initial screening was 15 mg ml^{-1} in 20 m*M* Tris–HCl pH 8.5. Droplets with an initial volume of 4 µl and a reservoir volume of 1 ml were used in the crystallization experiments. In the initial crystallizations, small needle-shaped crystals appeared after 2 h in a solution containing 0.2 *M* $MgCl_2$, 30% PEG 4000, 0.1 *M* Tris–HCl buffer pH 8.5. After optimization, larger prismatic crystals of dimensions 0.3 × 0.2 × 0.1 mm could be grown within 3 d using 0.1 *M* Tris–HCl, 15–19%(*w*/*v*) PEG 4000, 10–15%(*v*/*v*) PEG 400, 0.15–0.25 *M* $MgCl_2$ (Fig. 2).

2.3. Data collection and processing

A complete diffraction data set extending to 1.65 Å resolution was measured from a single crystal at the EMBL X31 beamline, DESY, Hamburg. The crystal was mounted in a nylon-fibre loop and flash-frozen in a nitrogen-gas stream (Teng, 1990). As the concentration of polyethyleneglycols used in the crystallization experiment was fairly high, no additional cryoprotectant was necessary. The intensity data were collected at a wavelength of 1.0442 Å using a 300 mm MAR image-plate system (MAR Research). A total of 144 images were collected with 1° oscillations. A sample image is shown in Fig. 3. Indexing and integration of the images was performed in *DENZO* and scaling of the intensity data in *SCALEPACK* from the *HKL* package (Otwinowski & Minor, 1997). Structure factors were derived from the reflection intensities using the program *TRUNCATE* from the *CCP*4 (Collaborative Computational Project, Number 4, 1994) suite. Table 1 summarizes the data-collection statistics. The synchrotron data consist of 736 323 measurements of 70 249 unique reflections, with an R_{merge} of 7.5%. Diffraction symmetry and systematic absences indicated that the crystal belongs to the orthorhombic space group $P2_12_12_1$, with unit-cell parameters a = 50.3, b = 77.6, c = 148.2 Å.

2.4. Structure solution

A Patterson self-rotation function was calculated from the native data in order to detect non-crystallographic symmetry (NCS) using the *POLARRFN* program from the *CCP*4 suite (Collaborative Computational Project, Number 4, 1994). However, despite trying several resolution limits and integration radii, no clear NCS elements could be identified. For molecular-replacement calculations, the program *EPMR* (Kissinger *et al.*, 1999) was used. It samples the rotation and translation space using genetic algorithms. As a probe, a polyalanine model of FmG (Guo *et al.*, 1998; PDB code 2gaw) was used. The possible solutions were characterized by very poor correlation coefficient (CC) and *R* values. For instance, using data in the resolution range 15.0–3.0 Å, the CC and the *R* values for the best solution after rigid-body refinement were 18.3 and 59.6%, respectively. However, since the packing of the molecules in the unit cell (inspected using the program *O*; Jones *et al.*, 1991) was plausible and the electron-density maps looked promising, further refinement was undertaken using the program *REFMAC* (Murshudov *et al.*, 1997). After mutating the recognizable fragments of the structure to the EcAIII sequence and correcting the major differences between the model and the crystal structure, several cycles of positional and *B*-factor refinement against all data in the resolution range 10–1.65 Å brought the *R* and R_{free} factors to 29.2 and 33.1%, respectively. Further steps of model building and refinement are in progress.

```
                                        ↓
E. coli[1]                  EK---------QKMG--TVGAVALD
L. luteus[2]                EKELPLTNGDSQI-G--TVGCVAVD
F. meningosepticum[3]       ENH------D-------TIGMIALD
H. sapiens[4]               ETE-----DDR---GHDTIGMVVIH
```

Figure 1
Sequences of several L-asparaginases/glycosylasparaginases in the region of the autoprocessing (arrow) and active site. 1, Blattner *et al.* (1997); 2, Borek *et al.* (1999); 3, Tarentino *et al.* (1995); 4, Fisher *et al.* (1990).

3. Results and discussion

The protein used for crystallization is not the original product of the *E. coli ybik* gene (33 kDa) but the autocatalytically processed product consisting of two subunits α and β with respective molecular masses of 19 and 14 kDa deduced from SDS–PAGE electrophoresis. As demonstrated by this structure analysis, the N-terminal residue of the second subunit is a threonine (Thr179; Fig. 1), providing structural grounds for classifying EcAIII as an Ntn hydrolase.

In the crystallization experiments it was found that addition of magnesium chloride was necessary for the formation of larger diffraction-quality crystals. Typical single crystals of EcAIII grown in the conditions described above have a prismatic habit but are nearly always malformed, *i.e.* they have voids or holes at one end of the prism (Fig. 2). These defects do not seem to have any serious adverse effect on the diffraction pattern and the crystals usually show good diffraction to relatively high angles (Fig. 3). However, they pose a problem on flash-freezing, leading to crystal cracking and very high mosaicity. This problem can be circumvented by selecting smaller specimens with minimal defects or larger ones in which only the good portion is exposed to the X-rays. The present data were collected using the former approach (mosaicity 0.5°).

Analysis of the Matthews volume (Matthews, 1968) indicated that the asymmetric unit is most likely to contain two copies of the $\alpha\beta$ heterodimer, the most obvious protein aggregate to consider. Such an assumption yields V_M = 2.17 Å^3 Da^{-1},

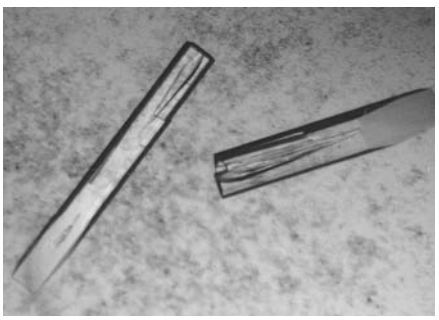

Figure 2
Single crystals of EcAIII.

Figure 3
Diffraction image recorded for an EcAIII crystal. Crystal oscillation $\Delta\varphi = 1.0^\circ$. The maximum resolution at the edge of the detector is 1.65 Å.

corresponding to 42% solvent content. From the similarity of the amino-acid sequences and the maturation pattern, it appeared that the structure of FmG (Guo *et al.*, 1998) should be a good model for solving the structure of EcAIII by molecular replacement. Since FmG also forms $(\alpha\beta)_2$ heterotetramers in its crystal structure, the complete tetramer was used as the molecular-replacement probe. Unfortunately, the Patterson self-rotation function did not reveal any obvious non-crystallographic twofold axis as would be expected for an FmG-like model. Similarly, MR calculations with the FmG model did not produce any obvious (high-correlation) results. The 'best' solution was rather unpromising considering the low correlation coefficient and high R values. However, it oriented the molecular dyad of the FmG model parallel to the a axis, had a plausible packing arrangement and, when subjected to atomic refinement, resulted in a steep fall in the R factor. Furthermore, in over half of the structure the MR-phased maps had excellent appearance, even revealing some of the side chains not present in the polyalanine FmG model. The model was therefore refined further, with a rapid extension of the resolution to a limit of 1.65 Å and with modelling of the side chains clearly visible in about 50% of the structure. It also became evident that about 40% of the structure shows significant differences with respect to the FmG model. Those areas have been rebuilt, leading to the current model. The best electron-density maps correspond to the N-terminal threonines 179 in both subunits β. Several residues are missing at the C-termini of all subunits. As already indicated by the orientation of the MR model, the molecular dyad relating the two $\alpha\beta$ heterodimers of the EcAIII molecule deviates from the a direction by less than 6°. This explains why no NCS could be detected in the Patterson map, where it was masked by the orthorhombic symmetry of the space group.

The research of MJ was supported in part by an International Research Scholar's award from the Howard Hughes Medical Institute. DB acknowledges support from DESY, Hamburg under an exchange program with the Polish Academy of Sciences. We thank Dr Wojciech Rypniewski for assistance at the Hamburg synchrotron and Dr Dominique Housset for helpful discussions.

References

Aronson, N. N. Jr (1996). *Glycobiology*, **6**, 669–675.
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.
Bonthron, D. T. & Jaskólski, M. (1997). *Acta Biochim. Pol.* **44**, 491–504.
Borek, D., Podkowinski, J., Kisiel, A. & Jaskolski, M. (1999). *Plant Physiol.* **119**, 1568.
Bradford, M. M. (1976). *Anal. Biochem.* **72**, 248–254.
Brannigan, J. A., Dodson, G., Duggleby, H. J., Moody, P. C. E., Smith, J. L., Tomchick, D. R. & Murzin, A. G. (1995). *Nature (London)*, **378**, 416–419.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Fisher, K., Klein, M., Park, H., Vettese, M. B. & Aronson, N. N. Jr (1993). *FEBS Lett.* **323**, 271–275.
Fisher, K., Tollersrud, O. K. & Aronson, N. N. Jr (1990). *FEBS Lett.* **269**, 440–444.
Guan, C., Cui, T., Rao, V., Liao, W., Benner, J., Lin, C.-L. & Comb, D. (1996). *J. Biol. Chem.* **271**, 1732–1737.
Guo, H.-C., Xu, Q., Buckley, D. & Guan, C. (1998). *J. Biol. Chem.* **273**, 20205–20212.
Ikonen, E., Julkunen, I., Tollersrud, O.-K., Kalkkinen, N. & Peltonen, L. (1993). *EMBO J.* **12**, 295–302.
Jancarik, J. & Kim, S.-H. (1991). *J. Appl. Cryst.* **24**, 409–411.
Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* A**47**, 110–119.
Kaartinen, V., Mononen, T., Laatikainen, R. & Mononen, I. (1992). *J. Biol. Chem.* **267**, 6855–6858.
Kaartinen, V., Williams, J. C., Tomich, J., Yates, J. R. III, Hood, L. E. & Mononen, I. (1991). *J. Biol. Chem.* **266**, 5860–5869.
Kissinger, C. R., Gehlhaar, D. K. & Fogel, D. B. (1999) *Acta Cryst.* D**55**, 484–491.
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.
Liu, Y., Dunn, G. S. & Aronson, N. N. Jr (1996). *Glycobiology*, **6**, 527–536.
Liu, Y., Guan, C. & Aronson, N. N. Jr (1998). *J. Biol. Chem.* **273**, 9688–9694.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* D**53**, 240–255.
Noronkoski, T., Stoineva, I. B., Petkov, D. D. & Mononen, I. (1997). *FEBS Lett.* **412**, 149–152.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Rao, J. K. M., Gribskov, M., Lubkowski, J., Miller, M., Swain, A. L. & Wlodawer, A. (1996). In *Techniques in Protein Chemistry VII*, edited by R. M. Marshak. San Diego: Academic Press.
Tanaka, M., Kohno, M. & Yamashina, I. (1973). *J.Biochem.* **73**, 1285–1289.
Tarentino, A. L., Quinones, G., Hauer, C. R., Changchien, L.-M. & Plummer, T. H. Jr (1995). *Arch. Biochem. Biophys.* **316**, 399–406.
Teng, T. Y. (1990). *J. Appl. Cryst.* **23**, 387–391.
Tenhunen, K., Laan, M., Manninen, T., Palotie, A., Peltonen, L. & Jalanko, A. (1995). *Genomics*, **30**, 244–250.