Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Cloning, purification, crystallization and X-ray analysis of the *Escherichia coli* pyrimidine nucleoside hydrolase YeiK

Barbara Giabbai and Massimo Degano*

DIBIT Scientific Institute S. Raffaele, I-20132 Milan, Italy

Correspondence e-mail: degano.massimo@hsr.it

Received 30 October 2003
Accepted 8 December 2003

The *E. coli yeiK* gene product is homologous to members of the nucleoside hydrolase family of enzymes, the physiological role of which in bacteria is still unknown. Here, the cloning, expression in milligram quantities and enzymatic characterization of YeiK as a pyrimidine-specific nucleoside hydrolase is reported. Crystals of YeiK diffract X-rays to a resolution of 1.7 Å and belong to the triclinic crystal system in space group $P1$, with unit-cell parameters $a = 44.81$, $b = 85.71$, $c = 90.68$ Å, $\alpha = 112.95$, $\beta = 101.95$, $\gamma = 85.92^\circ$.

1. Introduction

Nucleoside hydrolases (NHs) are central enzymes of the purine-salvage pathway in purine-auxotrophic organisms such as trypanosomes (Hammond & Gutteridge, 1984). These parasites lack the purine nucleoside phosphorylase activity found in prokaryotes and higher eukaryotes and rely instead on NH activity. NHs catalyze the hydrolysis of the N-glycosidic bond between the anomeric C atom of ribose and the purine or pyrimidine base, yielding free ribose and base (Miller *et al.*, 1984). The nucleobase is then recycled for nucleotide biosynthesis. Because of this metabolic difference, NH enzymes are considered to be attractive targets for the development of specific anti-trypanosomal drugs. NH proteins from kinetoplastids have been extensively characterized, revealing the existence of three classes with distinct substrate specificities. The non-specific NHs (inosine/uridine-preferring; IU-NH), whose epitome is the IU-NH from *Crithidia fasciculata*, hydrolyze both purine and pyrimidine nucleosides (Parkin *et al.*, 1991). Purine-specific NHs (inosine/adenosine/guanosine-preferring; IAG-NH), which have a strict specificity for inosine, adenosine and guanosine (Parkin, 1996; Pelle *et al.*, 1998), are almost exclusively encountered in trypanosomes. Finally, a guanosine/inosine-preferring NH (GI-NH) from *C. fasciculata* has been purified and characterized (Estupinan & Schramm, 1994). The structures of both IU- and IAG-NH enzymes show a common structural framework: a central β-sheet core surrounded by α-helices and topped by a four-helix cap. Substantially distinct active-site features, however, determine the substrate specificity (Degano *et al.*, 1996; Versees *et al.*, 2001). While the aromatic ring of the IU-NH inhibitor phenyliminoribitol interacts loosely with the surrounding protein residues (Degano *et al.*, 1998), the adenine ring of adenosine is stacked between tryptophan residues in the IAG-NH (Versees *et al.*, 2002). The NHs with different specificity also differ in the mechanism for leaving-group activation in the hydrolytic reaction: a histidine residue is involved in IU-NHs (Gopaul *et al.*, 1996), while an as yet unidentified proton donor balances the charge developing at the purine ring in IAG-NHs (Versees *et al.*, 2002).

The central common feature of all NHs so far characterized is a cluster of aspartate residues in the protein N-terminal region that chelate a calcium ion to assist in substrate binding and catalysis. Homology-based database searches identify more than 130 genes bearing this aspartate cluster and thus encoding putative NH-like proteins in both prokaryotes and eukaryotes. NH proteins are thus not confined to parasitic protozoa, but rather have a widespread distribution that hints at a conserved biological role. Functional characterizations of these gene products have confirmed their activity as NH enzymes. Expression of the three *Escherichia coli* NH-like genes augments nucleoside hydrolase activity in total cell extracts (Petersen & Moller, 2001). Recently, an NH homologue from *Campylobacter jejuni* was shown to be inactive towards commonly found nucleosides, while an NH gene from *Caenorhabditis elegans* encoded an enzyme with IAG-NH-like specificity (Versées *et al.*, 2003). The biological role of these enzymes in bacteria and higher eukaryotes is yet to be determined, since the recycling of purine and pyrimidine bases is performed by the nucleoside phosphorylase-catalyzed reaction. Structural information on non-parasitic NHs could provide valuable insights into their substrate specificity, catalytic mechanism and biological function.

The present study reports the cloning, expression, purification, enzymatic characterization and crystallization of YeiK protein from *E. coli*. The *yeiK* gene product indeed

encodes an NH enzyme, with specificity towards pyrimidine nucleosides (EC 3.2.2.8). A possible therapeutic application of the enzyme in cancer gene therapy also emerges from the hydrolytic activity of YeiK on the widely used chemotherapeutic agent 5-fluorouridine (Malet-Martino *et al.*, 2002). In this approach, a foreign gene or protein is directed to cancer cells and its activity sensitizes the tumour to the drugs that are co-administered. The biochemical and structural characterization of enzymes with prodrug-activating activities could greatly enhance the range of anticancer compounds that can be used *in vivo*. Triclinic crystals of YeiK diffracting to 1.7 Å resolution were obtained, placing the cornerstone for a high-resolution structural determination of this enzyme.

2. Materials and methods

2.1. Cloning

Genomic DNA from the *E. coli* K12 derivative DH5α strain was isolated using standard protocols. Primers 5′-CACC**CATATG**GAAAAGAGAAAAATTATTC-3′ and 5′-ATT**GAATTC**TTAATGGGTTTTGATGTAGC-3′ were designed from the published gene sequence to amplify the full-length gene including *Nde*I and *Eco*RI restriction sites, respectively. The PCR product of approximately 1.0 kbp was ligated to the pET101/D-TOPO (Invitrogen) vector and transformed into TOP10 chemically competent *E. coli* cells. The *yeiK* gene was then subcloned into the expression vector pET28b (Novagen) after digestion with the *Nde*I and *Eco*RI restriction enzymes. The plasmid was amplified in DH5α-competent *E. coli* cells. Both DNA strands of the cloned gene were completely sequenced using the automated dideoxy method.

2.2. Expression and purification

The pET28-YEIK plasmid, encoding the protein fused to a hexahistidine tag at the N-terminus, was transformed into BL21(DE3) (Novagen) competent cells for protein production. Cells were grown at 310 K in LB medium containing 50 µg ml^{-1} kanamycin. An overnight 5 ml culture was diluted in 200 ml of medium. Protein expression was induced at a culture OD_{600} of 0.6 by addition of 0.5 m*M* isopropyl-β-D-1-thiogalactopyranoside (Sigma). Cells were harvested after 3 h by centrifugation and resuspended in 1/10th volume of a buffer containing 20 m*M* Tris–HCl pH 8.0, 50 m*M* NaCl, 1 mg ml^{-1} lysozyme, 20 µg ml^{-1} ribonuclease A, 2 µg ml^{-1} deoxyribonuclease I and Complete EDTA-free protease-inhibitor cocktail (Roche). Cells were incubated at 310 K for 10 min and disrupted by sonication at 277 K using a Bandelin sonifier. Cellular debris was removed by centrifugation at 25 000*g* for 30 min at 277 K. Ni–NTA resin (Qiagen) was added to the clear supernatant and incubated for 1 h with gentle shaking at 277 K and then loaded onto a polypropylene column. Unbound proteins were removed by washing the resin with 100 column volumes of washing buffer (20 m*M* Tris–HCl pH 7.5, 0.5 *M* NaCl, 10 m*M* imidazole). Bound proteins were eluted with the same buffer containing 0.5 *M* imidazole. The affinity-purified YeiK was injected onto a Superdex 200 10/30 size-exclusion chromatography column (Amersham Pharmacia) equilibrated with 20 m*M* HEPES pH 7.4, 150 m*M* NaCl. Purified YeiK was digested overnight at room temperature with thrombin (Sigma) at a ratio 1:1500(*w*:*w*) after dialysis against a buffer containing 20 m*M* Tris–HCl pH 8.3, 150 m*M* NaCl, 2.5 m*M* EDTA. The digested protein was loaded onto a Superdex 200 chromatography column as described above to remove the protease and the cleaved His tag. The quantitative removal of the affinity tag was verified by Western blot analysis using a horseradish peroxidase-conjugated anti-His_6 antibody (Santa Cruz) and ECL detection.

Table 1
Data-collection and processing statistics.

Values in parentheses refer to the highest resolution shell.

Beamline	ESRF ID14-EH2
Wavelength (Å)	0.933
Resolution range (Å)	81.6–1.7 (1.79–1.70)
No. measured reflections	239378 (23873)
No. unique reflections	116872 (12529)
Completeness (%)	91.5 (63.8)
R_{sym}†	0.045 (0.217)
R_{meas}‡	0.064 (0.306)
$\langle I/\sigma(I)\rangle$	9.5 (3.0)

† $R_{sym} = \sum_h \sum_i |\langle I_h\rangle - I_{i,h}|/\sum_h I_i$, where i is the ith measurement of reflection h. ‡ R_{meas} is the multiplicity-weighted R_{sym} (Diederichs & Karplus, 1997).

2.3. Substrate specificity

NH activity was assayed using a colorimetric test as previously described (Parkin *et al.*, 1991), with the following modifications. Reactions times were adjusted to between 5 and 25 min to allow sufficient conversion to products. Activity measurements were repeated in triplicate, with the substrate concentration fixed at 5 m*M*. The absorbance expected from 100% conversion of substrate was measured using 5 m*M* ribose. Protein concentration was measured using the BCA method (Pierce).

2.4. Dynamic light scattering

Protein solutions were centrifuged at 20 000*g* for 30 min at 277 K to remove precipitates or dust particles. Dynamic light-scattering (DLS) measurements were performed at 277, 293 and 310 K on a DynaPro MS/X instrument with temperature control (Protein Solutions). Laser intensity, time of acquisition and number of measurements were optimized at each protein concentration to minimize the residual to the fitted scattering curves. Enzyme concentrations used ranged from 0.01 to 10 mg ml^{-1}. Data collection and deconvolution was performed using the *DYNAMICS* v.6 software.

2.5. Crystallization and data collection

Preliminary crystallization conditions were established using the hanging-drop vapour-diffusion method. Equal volumes of a protein solution at 10 mg ml^{-1} in 20 m*M* HEPES buffer pH 7.4 and 150 m*M* NaCl and precipitant solution (2 µl each) were mixed and placed on a cover slip over 700 µl of reservoir solution and incubated at 298 K. Several different conditions from Crystal Screens I and II and Index Screen (Hampton Research) yielded single crystals of different shapes. Diffraction-quality crystals grew in conditions optimized from Index Screen condition No. 73. Flat plate-like crystals typically appeared after 48 h from 100 m*M* Tris–HCl pH 8.5, 200 m*M* NaCl and 25%(*w*/*v*) PEG 3350 and grew to maximum dimensions of 0.3 × 0.3 mm. Crystals were transferred to an artificial mother liquor containing 25% glycerol for cryoprotection and were immediately flash-cooled in a nitrogen stream at 100 K. Crystals diffracted to 1.5 Å at beamline ID14-EH2 of the European Synchrotron Radiation Facility (Grenoble, France) using a Quantum 4 CCD detector from ADSC at a wavelength of 0.933 Å. Diffraction data to 1.7 Å were collected from a single crystal with the rotation method using 1° oscillations and 3 s exposure (Table 1). Data were indexed, integrated with *MOSFLM* (Leslie, 1992), scaled and reduced with *SCALA* (Evans, 1997; Collaborative Computational Project 4, 1994).

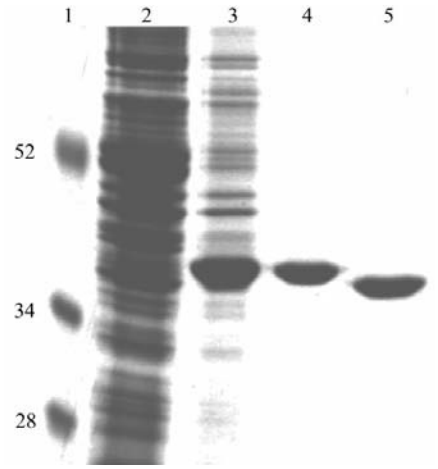

Figure 1
Expression and purification of YeiK. Denaturing SDS–PAGE analysis of YeiK at various stages of purification. Lane 1, molecular-weight markers (kDa); lane 2, BL21(DE3) cells, soluble fraction; lane 3, overexpression of YeiK in BL21(DE3)/pET28-YEIK, soluble fraction; lane 4, elution of YeiK from the Ni–NTA resin; lane 5, removal of the N-terminal His tag by enzymatic digestion.

3. Results and discussion

The *E. coli yeiK* gene product was cloned and overexpressed in soluble recombinant form for biochemical and structural characterization. The high levels of expression in BL21(DE3) cells typically yielded 10 mg of pure protein from 200 ml of culture. The use of limiting quantities of Ni–NTA resin allowed a single-step purification of the recombinant protein. Removal of the N-terminal affinity tag was readily obtained through enzymatic digestion of the purified protein. A Western blot analysis showed a loss of reactivity towards an anti-His_6 antibody after digestion with thrombin, thus indicating a quantitative removal of the hexahistidine-containing peptide. The thrombin-digested YeiK corresponds in composition to the wild-type protein, with a Gly-Ser-His tripeptide added before the N-terminal methionine. The recombinant protein runs as a single band on a denaturing SDS–PAGE gel with an apparent molecular weight of 36 kDa, close to the value of 33 748 Da calculated based on the amino-acid composition (Fig. 1). In size-exclusion chromatography experiments performed at protein concentrations between 0.2 and 2 mg ml^{-1}, the protein elutes from a Superdex 200 column at a retention volume corresponding to a molecular weight of 143 kDa and is hence consistent with a homotetrameric quaternary structure. DLS measurements performed on YeiK at concentrations ranging from 0.01 to 10 mg ml^{-1} indicated a hydrodynamic radius of 4.8 ± 0.2 nm and a deduced molecular weight of 137 kDa, with a typical polydispersity of 15%.

The substrate specificity of YeiK defines the enzyme as a pyrimidine-specific nucleoside hydrolase. Purine nucleosides are not hydrolyzed to a measurable extent, with specific activities less than 1000-fold compared with those observed for pyrimidine nucleosides. 2′-, 3′- and 5′-deoxynucleosides are also not substrates for YeiK, thus confirming the general NH property in discriminating the presence of all ribose hydroxyls (Fig. 2). The differences between the affinity-tagged and thrombin-cleaved recombinant YeiK suggest an involvement of the N-terminal tag in the enzymatic activity. Alternatively, the N-terminal extension could hinder catalysis through a steric effect.

Of the several crystal forms obtained in the crystallization experiments, the highest quality diffraction was observed from the triclinic YeiK crystals (Fig. 3). These crystals belong to space group *P*1, with unit-cell parameters $a = 44.81$, $b = 85.71$, $c = 90.68$ Å, $\alpha = 112.95$, $\beta = 101.95$, $\gamma = 85.92^\circ$. The unit-cell volume is consistent with one YeiK tetramer in the crystal asymmetric unit with a Matthews coefficient of 2.3 $Å^3$ Da^{-1}, corresponding to 46% solvent content (Table 1). The high-resolution data set collected from these crystals paves the way for the determination of the crystal structure of YeiK. The sequence similarity between YeiK and the IU-NHs for which structures have been determined to date, those from *C. fasciculata* (PDB code 1mas; 40% identity) and *Leishmania major* (PDB code 1ezr; 42% identity), allows the use of the molecular-replacement technique for phase determination. The availability of a high-resolution model of the first bacterial nucleosidase could provide further insights into the biological role of this class of enzymes in non-parasitic organisms. Moreover, the structure of YeiK will represent the first characterization of a pyrimidine-specific NH, revealing the structural basis of the substrate specificity. Finally, the YeiK protein is highly active towards the widely used chemotherapeutic agent 5-fluorouridine, which requires activation *via* either phosphorylation at the O5′ atom or conversion to 5-fluorouracil *via* glycosidic bond cleavage. This finding prompts further studies for a possible use of this gene in cancer suicide gene therapy to augment cell sensitivity to the prodrug.

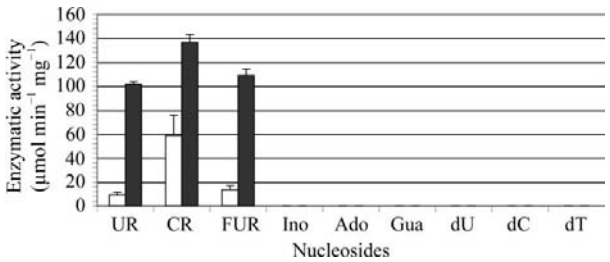

Figure 2
Enzymatic activity of affinity-tagged (white bars) and thrombin-treated YeiK (black) for various nucleosides. Values reported are corrected for non-enzymatic background hydrolysis. The specific activity of YeiK for pyrimidine and 2′-deoxynucleosides ranges between 10^{-2} and 10^{-4} μmol min^{-1} mg^{-1}, within the experimental error. Abbreviations: UR, uridine; CR, cytidine; FUR, 5-fluorouridine; Ino, inosine; Ado, adenosine, Gua, guanosine; dU, 2′-deoxyuridine; dC, 2′-deoxycytidine; dT, 2′-deoxythymidine.

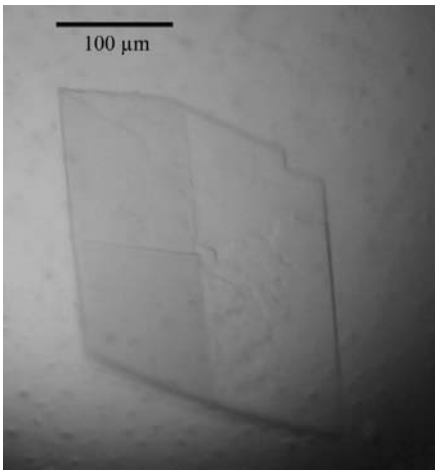

Figure 3
Plate-like triclinic crystal of recombinant YeiK from *E. coli*.

This study was supported by grants from Human Frontiers of Science Program and the Italian Foundation for Cancer Research. Support by ESRF beamline staff is gratefully acknowledged.

References

Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.

Degano, M., Almo, S. C., Sacchettini, J. C. & Schramm, V. L. (1998). *Biochemistry*, **37**, 6277–6285.
Degano, M., Gopaul, D. N., Scapin, G., Schramm, V. L. & Sacchettini, J. C. (1996). *Biochemistry*, **35**, 5971–5981.
Diederichs, K. & Karplus, P. A. (1997). *Nature Struct. Biol.* **4**, 269–275.
Estupinan, B. & Schramm, V. L. (1994). *J. Biol. Chem.* **269**, 23068–23073.
Evans, P. R. (1997). *Jnt CCP4/ESF–EACBM Newsl. Protein Crystallogr.* **33**, 22–24.
Gopaul, D. N., Meyer, S. L., Degano, M., Sacchettini, J. C. & Schramm, V. L. (1996). *Biochemistry*, **35**, 5963–5970.
Hammond, D. J. & Gutteridge, W. E. (1984). *Mol. Biochem. Parasitol.* **13**, 243–261.
Leslie, A. G. W. (1992). *Jnt CCP4/ESF–EACBM Newsl. Protein Crystallogr.* **26**, 27–33.
Malet-Martino, M., Jolimaitre, P. & Martino, R. (2002). *Curr. Med. Chem. Anti-Cancer Agents*, **2**, 267–310.
Miller, R. L., Sabourin, C. L., Krenitsky, T. A., Berens, R. L. & Marr, J. J. (1984). *J. Biol. Chem.* **259**, 5073–5077.
Parkin, D. W. (1996). *J. Biol. Chem.* **271**, 21713–21719.
Parkin, D. W., Horenstein, B. A., Abdulah, D. R., Estupinan, B. & Schramm, V. L. (1991). *J. Biol. Chem.* **266**, 20658–20665.
Pelle, R., Schramm, V. L. & Parkin, D. W. (1998). *J. Biol. Chem.* **273**, 2118–2126.
Petersen, C. & Moller, L. B. (2001). *J. Biol. Chem.* **276**, 884–894.
Versees, W., Decanniere, K., Pelle, R., Depoorter, J., Brosens, E., Parkin, D. W. & Steyaert, J. (2001). *J. Mol. Biol.* **307**, 1363–1379.
Versees, W., Decanniere, K., Van Holsbeke, E., Devroede, N. & Steyaert, J. (2002). *J. Biol. Chem.* **277**, 15938–15946.
Versées, W., Van Holsbeke, E., De Vos, S., Decanniere, K., Zegers, I. & Steyaert, J. (2003). *Acta Cryst.* D**59**, 1087–1089.