Acta Crystallographica Section F
Structural Biology Communications
ISSN 2053-230X

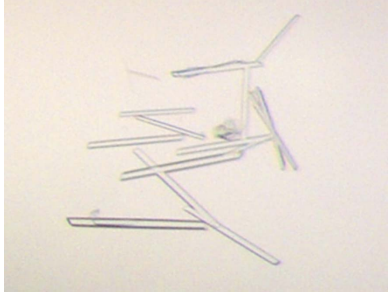
Overproduction, crystallization and preliminary X-ray crystallographic analysis of *Escherichia coli* tRNA N^6-threonylcarbamoyladenosine dehydratase

Sunmin Kim,[a] Keon Young Kim,[a] Jeong Kuk Park,[a] Byung Il Lee,[b] Yun-Gon Kim[c] and SangYoun Park[a]*

[a]School of Systems Biomedical Science, Soongsil University, Seoul 156-743, Republic of Korea, [b]Biomolecular Function Research Branch, Division of Convergence Technology, Research Institute, National Cancer Center, Goyang, Gyeonggi 410-769, Republic of Korea, and [c]Department of Chemical Engineering, Soongsil University, Seoul 156-743, Republic of Korea

Correspondence e-mail: psy@ssu.ac.kr

Received 10 July 2014
Accepted 17 September 2014



Escherichia coli tRNA N^6-threonylcarbamoyladenosine dehydratase (TcdA), previously called CsdL or YgdL, was overproduced and purified from *E. coli* and crystallized using polyethylene glycol 3350 as a crystallizing agent. X-ray diffraction data were collected to 2.70 Å resolution under cryoconditions using synchrotron X-rays. The crystals belonged to space group $P2_1$, with unit-cell parameters $a = 65.4$, $b = 96.8$, $c = 83.3$ Å, $\beta = 111.7^\circ$. According to the Matthews coefficient, the asymmetric unit may contain up to four subunits of the monomeric protein, with a crystal volume per protein mass (V_M) of 2.12 Å^3 Da^{-1} and 42.1% solvent content.

1. Introduction

Transfer RNAs (tRNAs) undergo various post-transcriptional modifications necessary for proper function during translation (reviewed by El Yacoubi *et al.*, 2012). A well characterized alteration is the wobble modification, which includes 5-methoxycarbonylmethylation and 2-thiolation of U34. N^6-Threonylcarbamoyladenosine (t^6A), N^6-methylation and 2-methylthiolation at the A37 position of ANN-decoding tRNAs have also been well documented. All of these modifications were shown to stabilize the codon–anticodon interactions and to be necessary for precise decoding of translation (El Yacoubi *et al.*, 2012; Schweizer *et al.*, 1969; Murphy *et al.*, 2004). Interestingly, *Escherichia coli* t^6A has recently been shown to undergo further modification into a cyclic N^6-threonylcarbamoyladenosine (ct^6A) through an ATP-dependent dehydration reaction catalyzed by tRNA threonylcarbamoyladenosine dehydratase (TcdA; Miyauchi *et al.*, 2013). TcdA is widely conserved in bacteria in the phyla Proteobacteria (classes Betaproteobacteria, Gammaproteobacteria and Deltaproteobacteria), Verrucomicrobia and Firmicutes (classes Clostridia and Bacilli) (Fig. 1). The *tcdA* gene is located downstream of and transcribed in the opposite direction to the *csdA* and *csdE* genes encoding CsdA (cysteine desulfurase) and CsdE (sulfur acceptor), respectively, from the sulfur-generating and sulfur-utilizing cysteine sulfinate desulfinase (CSD) system. Hence, TcdA was previously named CsdL (Trotter *et al.*, 2009). Experiments suggest that CsdA, CsdE and TcdA are functionally related (Trotter *et al.*, 2009; Bolstad *et al.*, 2010; Miyauchi *et al.*, 2013).

The N-terminal nucleotide-binding Rossmann fold of TcdA is predicted to belong to the eukaryotic ubiquitin-activating (E1-like) superfamily of proteins with a C-terminal zinc-binding domain containing four conserved cysteines (Trotter *et al.*, 2009; Fig. 1). Prokaryotes contain two other E1-like proteins, ThiF and MoeB, which function as adenylation/thiocarboxylation enzymes during molybdopterin and thiamin biosynthesis, respectively (Duda *et al.*, 2005; Lake *et al.*, 2001). TcdA and ThiF/MoeB proteins share significant identity within the N-terminal Rossmann fold (TcdA and ThiF have 29.0% identity with 131 residues overlapping; TcdA and MoeB have 32.8% identity with 122 residues overlapping; Fig. 1), but TcdA lacks conserved cysteines within the zinc-binding site of the ThiF/MoeB family (Fig. 1). In fact, secondary-structure prediction (Combet *et al.*, 2000) of the TcdA C-terminus indicates a very different fold compared with that of ThiF/MoeB (Fig. 1). Since the zinc-binding site in the C-terminal region of ThiF and MoeB is pivotal

Table 1
Macromolecule-production information.

Source organism	*E. coli*
DNA source	Genomic DNA
Forward primer†	GGCAGCCATATGTCTGTGGTAATTAGTGA
Reverse primer‡	TTCGGATCCTTAACCCTGACGCGCCG
Cloning vector	pET-28a
Expression vector	pET-28a
Expression host	*E. coli* BL21 (DE3)
Complete amino-acid sequence of the construct produced§	MGSSHHHHHHSSGLVPRGSHMSVVISDAWRQRFGGTARLYGE-KALQLFADAHICVVGIGGVGSWAAEALARTGIGAITLIDM-DDVCVTNTNRQIHALRDNVGLAKAEVMAERIRQINPECRV-TVVDDFVTPDNVAQYMSVGYSYVIDAIDSVRPKAALIAYC-RRNKIPLVTTGGAGGQIDPTQIQVTDLAKTIQDPLAAKLR-ERLKSDFGVVKNSKGKLGVDCVFSTEALVYPQSDGTVCAM-KATAEGPKRMDCASGFGAATMVTATFGFVAVSHALKKMMA-KAARQG

† The *Nde*I site is underlined. ‡ The *Bam*HI site is underlined. § The non-native His_6 tag and thrombin site are underlined.

for substrate recognition of ThiS and MoaD, respectively, both of which share a ubiquitin-like fold, the novel C-terminal fold of TcdA may create a surface required for tRNA interaction. Because ThiF and MoeB utilize ATP for substrate adenylation and subsequent substrate thiocarboxylation, TcdA may catalyze tRNA dehydratase activity in a similar mechanistic manner.

We have initiated overproduction, crystallization and diffraction experiments on *E. coli* TcdA. Diffracting crystals were obtained by hanging-drop vapour diffusion and a complete 2.70 Å resolution X-ray diffraction data set was collected using synchrotron radiation.

2. Materials and methods

2.1. Macromolecule production

Full-length *E. coli* TcdA (residues 1–268) was amplified from genomic *E. coli* DNA by polymerase chain reaction (PCR) and cloned into pET-28a vector (Novagen) using *Nde*I and *Bam*HI restriction sites (Table 1). The N-terminal His_6-tagged protein was overproduced in *E. coli* BL21 (DE3) cells (Novagen). Transformed cells were grown at 37°C in Luria–Bertani (LB) medium to an OD_{600} of ∼0.8 in the presence of 25 µg ml^{-1} kanamycin. Overproduction of the recombinant protein was induced by the addition of 0.5 m*M* isopropyl *β*-D-1-thiogalactopyranoside (IPTG) at 22°C, and cells were further cultivated for 16 h.

Cells were harvested using centrifugation at 4500*g* for 10 min (4°C). A 1 l equivalent of bacterial cell pellets were resuspended in 50 ml ice-cold lysis buffer (20 m*M* Tris pH 7.5, 500 m*M* NaCl, 5 m*M* imidazole) and lysed on ice by sonication (200 W, 60% power, 10 min, 0.5 s on/0.5 s off). Lysates were centrifuged at 70 000*g* for 30 min (4°C) and supernatants were poured over a 5 ml nickel–nitrilotriacetic acid agarose (Ni–NTA; Qiagen) gravity column. The protein-bound column was washed with five column volumes of wash buffer (20 m*M* Tris pH 7.5, 20 m*M* imidazole, 500 m*M* NaCl) and the protein was eluted with elution buffer (20 m*M* Tris pH 7.5, 200 m*M* imidazole, 500 m*M* NaCl). The protein-containing fractions (approximate total volume of 10 ml) were pooled and 50 µl of 0.25 U µl^{-1} bovine thrombin (Invitrogen) was added. After incubation for 16 h at 4°C, the sample was further purified using a HiLoad 26/60 Superdex 200 size-exclusion chromatography (SEC) column (column volume 318 ml) which was pre-equilibrated with SEC buffer (50 m*M* Tris pH 7.5, 150 m*M* NaCl, 2 m*M* DTT). The entire ∼10 ml from a 1 l equivalent of bacterial culture was loaded using a 10 ml super-loop connected to an ÄKTA FPLC system (GE Healthcare). The elution profile of TcdA showed two major peaks, which were separately concentrated by ultracentrifugation (Amicon). The approximate molecular masses of the two TcdA fractions were estimated using standard molecular-mass markers run on analytical SEC (Superdex 200 10/300 GL, GE Healthcare). Final protein concentrations [first elution peak (peak 1), 33 mg ml^{-1}; second elution peak (peak 2), 8.5 mg ml^{-1}] were estimated by A_{280} with a molar extinction coefficient (19 000 M^{-1} cm^{-1}) calculated based upon the number of tryptophan and tyrosine residues (Gill & von Hippel, 1989). Purity and homogeneity were assessed using SDS–PAGE analysis (Fig. 2).

Figure 1
Sequence alignment of *E. coli* ThiF, MoeB, TcdA (CsdL) and TcdA (CsdL)-like proteins from various prokaryotic organisms. Conserved residues in all ThiF, MoeB and TcdA are shaded. In ThiF and MoeB, conserved cysteines are boxed and the catalytically essential as well as the zinc-binding cysteines are shown with asterisks. In TcdA, the cysteines of *E. coli* TcdA that recur in the different TcdA orthologues are boxed. The secondary-structure elements and labels of ThiF and MoeB are shown above the alignment. The predicted secondary-structure elements of the TcdA C-terminus are also shown. (Species codes: *Ec*, *Escherichia coli*; *Am*, *Akkermansia muciniphila*; *Nm*, *Neisseria meningitidis*; *Gs*, *Geobacter sulfurreducens*; *Ca*, *Caloramator australicus*; *Sa*, *Staphylococcus aureus*.)

Table 2
Crystallization.

Method	Vapour diffusion (hanging drop)
Plate type	Linbro 24-well
Temperature (°C)	22
Protein concentration (mg ml^{-1})	8.5
Buffer composition of protein solution	50 m*M* Tris pH 7.5, 150 m*M* NaCl, 2 m*M* DTT
Composition of reservoir solution	20–30%(*m*/*v*) PEG 3350, 0.15 *M* DL-malic acid pH 7.0
Volume and ratio of drop	2 µl (1:1 ratio of protein and reservoir)
Volume of reservoir (µl)	500

The concentrated proteins in SEC buffer were flash-cooled and stored in liquid nitrogen.

2.2. Crystallization

Conditions for obtaining TcdA crystals were screened using commercial solutions (Hampton Research) in Linbro 24-well hanging-drop crystallization plates set up at 22°C (Table 2). Drops were set up by mixing equal 1 µl volumes of the reservoir solution and the protein solution in SEC buffer and equilibrating against 500 µl reservoir solution. Crystals were only obtained from the second major peak of the SEC elution. Crystals which grew in 20–30%(*m*/*v*) polyethylene glycol (PEG) 3350, 0.15 *M* DL-malic acid pH 7.0 were transferred to a cryoprotectant solution consisting of reservoir solution supplemented with 15% glycerol prior to flash-cooling in liquid nitrogen for storage and transportation to the synchrotron facility.

2.3. Data collection and processing

X-ray diffraction data were collected at −173°C using a CCD detector (ADSC Quantum 315r) on beamline 5C of the Pohang Light Source (PLS; Pohang, Republic of Korea). The crystal was rotated through a total of 140° with 1.0° oscillation range per frame. Data were processed in space group $P2_1$ using *HKL*-2000 (Otwinowski & Minor, 1997; Table 3). *Phaser* was used for molecular-replacement phasing methods (McCoy *et al.*, 2005).

3. Results and discussion

Recombinant *E. coli* TcdA was overproduced in *E. coli* as a soluble protein and purified with an overall yield of ∼15 mg per litre of LB culture. Transmembrane helix prediction (*TMHMM* server v.2.0) suggested the presence of a single transmembrane segment with repeating hydrophobic residues (TcdA residues 235–257). However, the full-length protein behaved as a soluble protein without any addition of detergents. The SEC elution profile of TcdA on HiLoad 26/60 Superdex 200 (330 ml column volume) showed two major peaks with elution volumes at 242 ml (peak 1) and 258 ml (peak 2). SDS–PAGE analysis of the separately concentrated proteins under the two peaks indicated sizes corresponding to monomeric TcdA, with no clear difference in the two molecular masses (Fig. 2). The estimated molecular masses of the two states based on standard protein markers were ∼45 kDa (peak 1) and ∼35 kDa (peak 2).

Crystals were obtained only from TcdA from peak 2. Dilution of TcdA of peak 1 (33 mg ml^{-1}) to the concentration of TcdA of peak 2 (8.5 mg ml^{-1}) showed no effect on crystal formation. Clustered crystals of TcdA were grown in several crystallization conditions although single crystals grew using a reservoir solution of 20–30%(*w*/

Table 3
Data collection and processing.

Values in parentheses are for the outer shell.

Diffraction source	Beamline 5C, PLS
Wavelength (Å)	0.987
Temperature (°C)	−173
Detector	ADSC Quantum 315r
Crystal-to-detector distance (mm)	300
Rotation range per image (°)	1
Total rotation range (°)	140
Exposure time per image (s)	1
Space group	$P2_1$
a, *b*, *c* (Å)	65.4, 96.8, 83.3
α, β, γ (°)	90.0, 111.7, 90.0
Mosaicity (°)	0.88
Resolution range (Å)	50.0–2.70 (2.80–2.70)
Total No. of reflections	76187
No. of unique reflections	26246
Completeness (%)	97.0 (99.1)
Multiplicity	2.9 (3.0)
$\langle I/\sigma(I)\rangle$	14.2 (3.8)
$R_{\rm r.i.m.}$†	0.121 (0.473)
Overall *B* factor from Wilson plot (Å^2)	39.2

† The redundancy-independent merging *R* factor $R_{\rm r.i.m.}$ was estimated by multiplying the conventional $R_{\rm merge}$ value by the factor $[N/(N-1)]^{1/2}$, where *N* is the data multiplicity.

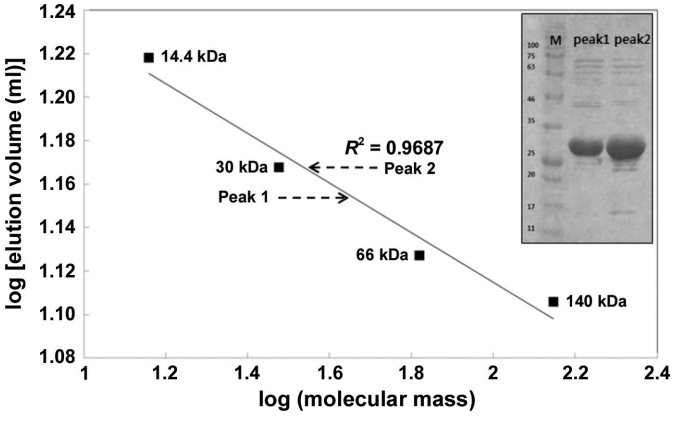


Figure 2
SEC analysis of polydisperse TcdA (protein standards: α-lactalbumin, 14.4 kDa; carbonic anhydrase, 30 kDa; albumin, 66 kDa; lactate dehydrogenase, 140 kDa) and SDS–PAGE analysis (inset) of final purified proteins (lane *M* contains molecular markers labelled in kDa).

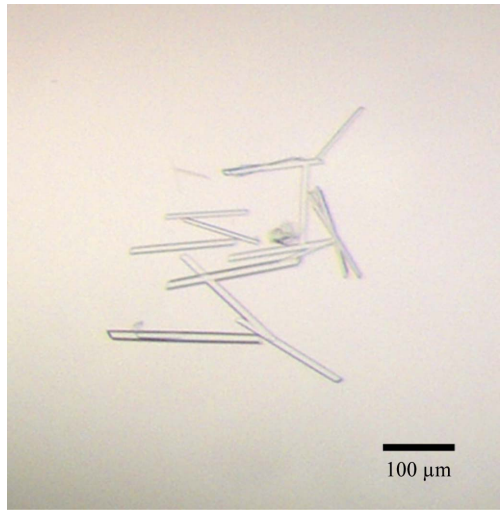


Figure 3
Crystals of *E. coli* TcdA.

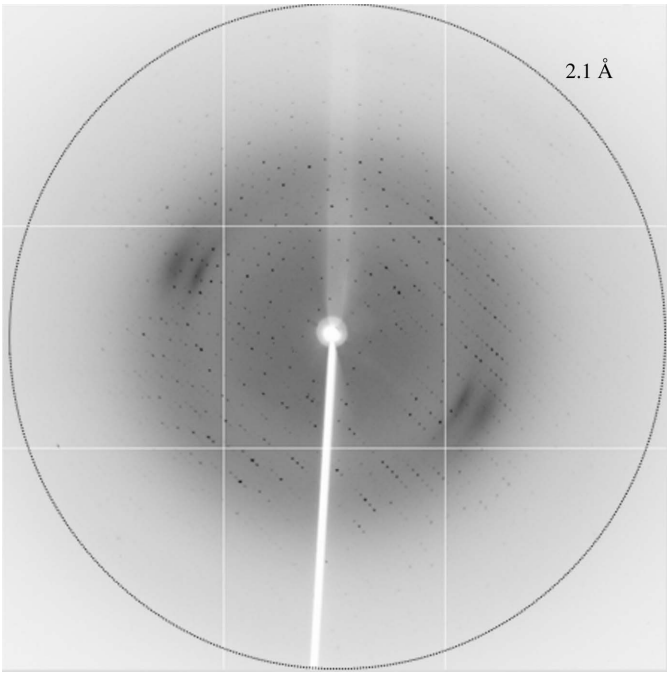


Figure 4
Representative diffraction image of the TcdA crystal.

v) PEG 3350, 0.15 *M* DL-malic acid pH 7.0. Within 2 d, the crystals grew to approximately 20 × 20 × 250 µm (Fig. 3) and were sufficient in size for X-ray diffraction experiments. Most crystals which were screened for diffraction using synchrotron radiation showed anisotropic diffraction, with only one out of ten crystals giving isotropic diffraction in all directions of oscillation to an average limit of ∼2.7 Å resolution (Fig. 4). The varying PEG 3350 concentrations used for crystallization did not affect the overall diffraction quality. The crystals used for data collection were in a cryoprotectant solution consisting of reservoir solution supplemented with 15% glycerol. A total of 26 246 unique reflections were measured and merged in space group $P2_1$ as concluded from the systematic absences (unit-cell parameters $a = 65.4$, $b = 96.8$, $c = 83.3$ Å, $\beta = 111.7^\circ$). The merged data set was 99.1% complete overall with an $R_{\rm r.i.m}$ of 12.1% (50–2.70 Å). The statistics for the collected data are summarized in Table 3. According to the Matthews coefficient (Matthews, 1968), the asymmetric unit may contain up to four subunits of monomeric TcdA with a crystal volume per protein mass ($V_{\rm M}$) of 2.12 Å^3 Da^{-1} and a 42.1% solvent content. However, an asymmetric unit containing three subunits is also plausible, with a $V_{\rm M}$ of 2.83 Å^3 Da^{-1} and 56.6% solvent content.

A self-rotation function search did not reveal any peaks corresponding to noncrystallographic symmetry. Surprisingly, despite the significant sequence identity (29–32%) found between the N-terminal Rossmann folds of TcdA and MoeB/ThiF (Fig. 1), molecular replacement using search models of MoeB (PDB entry 1jw9; Lake *et al.*, 2001) and ThiF (PDB entry 1zkm; Duda *et al.*, 2005) either in the monomeric form or in the various noncrystallographic dimers (or tetramers) of ThiF and the crystallographic dimer of MoeB failed to give a probable solution (*Phaser* statistics of TFZ < 4). Other search models using homology or polyalanine models of ThiF showed no improvement in the molecular-replacement trials.

In the future, we plan to determine the TcdA structure using direct phasing. Because no standard *in vitro* assay currently exists for probing N^6-threonylcarbamoyladenosine dehydratase activity, additional information on the TcdA structure in complex with ATP, t_6A or even tRNA would be crucial in understanding the mechanism that governs TcdA activity.

This work was supported by the Soongsil University Research Fund of 2012.

References

Bolstad, H. M., Botelho, D. J. & Wood, M. J. (2010). *J. Proteome Res.* **9**, 5358–5369.
Combet, C., Blanchet, C., Geourjon, C. & Deléage, G. (2000). *Trends Biol. Sci.* **25**, 147–150.
Duda, D. M., Walden, H., Sfondouris, J. & Schulman, B. A. (2005). *J. Mol. Biol.* **349**, 774–786.
El Yacoubi, B., Bailly, M. & de Crécy-Lagard, V. (2012). *Annu. Rev. Genet.* **46**, 69–95.
Gill, S. C. & von Hippel, P. H. (1989). *Anal. Biochem.* **182**, 319–326.
Lake, M. W., Wuebbens, M. M., Rajagopalan, K. V. & Schindelin, H. (2001). *Nature (London)*, **414**, 325–329.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
McCoy, A. J., Grosse-Kunstleve, R. W., Storoni, L. C. & Read, R. J. (2005). *Acta Cryst.* D**61**, 458–464.
Miyauchi, K., Kimura, S. & Suzuki, T. (2013). *Nature Chem. Biol.* **9**, 105–111.
Murphy, F. V., Ramakrishnan, V., Malkiewicz, A. & Agris, P. F. (2004). *Nature Struct. Mol. Biol.* **11**, 1186–1191.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Schweizer, M. P., Chheda, G. B., Baczynskyj, L. & Hall, R. H. (1969). *Biochemistry*, **8**, 3283–3289.
Trotter, V., Vinella, D., Loiseau, L., Ollagnier de Choudens, S., Fontecave, M. & Barras, F. (2009). *Mol. Microbiol.* **74**, 1527–1542.