Acta Crystallographica Section F
Structural Biology and Crystallization Communications
ISSN 1744-3091

Purification, crystallization and preliminary X-ray crystallographic analysis of 23S RNA m^{2}G2445 methyltransferase RlmL from *Escherichia coli*

Kai-Tuo Wang, Linglong Ma, Jie Nan, Xiao-Dong Su and Lanfen Li*

The National Laboratory of Protein Engineering and Plant Genetic Engineering, School of Life Sciences, Peking University, Beijing 100871, People's Republic of China

Correspondence e-mail: lilf@pku.edu.cn

Received 18 May 2010
Accepted 31 August 2010

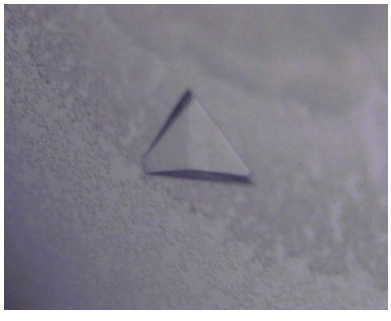
The RlmL (YcbY) protein in *Escherichia coli* is an rRNA methyltransferase that is specific for m^{2}G2445 modification of 23S RNA. The *rlmL* gene was cloned into the expression vector pET28a and expressed in the host *E. coli* strain BL21 (DE3). Recombinant protein with a six-histidine tag was purified by Ni^{2+}-affinity chromatography followed by gel filtration. Crystals were grown using the hanging-drop vapour-diffusion method and a detergent was used as an additive to improve diffraction quality. The final crystals diffracted to 2.2 Å resolution. The crystals belonged to space group $P2_1$, with unit-cell parameters a = 73.6, b = 140.8, c = 102.9 Å, β = 102.3°. The crystal has a most probable solvent content of 62.8% with two molecules in the asymmetric unit.

1. Introduction

The ribosome is the most complicated macromolecular machinery in the cell. In all kingdoms of life, post-transcriptional modifications of ribosomal RNA (rRNA) are required for both ribosome maturation and fine-tuning of the protein-synthesis process (Grosjean, 2005). Prokaryotic rRNA modifications are mostly conducted by specific methyltransferases and pseudouridine synthases. Modification of *Escherichia coli* rRNA has been comprehensively mapped compared with other organisms (Rozenski *et al.*, 1999). A total of 36 nucleotides are modified in *E. coli* rRNA, 11 of which are from 16S rRNA and 25 of which are from 23S rRNA. Most of the enzymes responsible for rRNA modification have been identified in *E. coli*, with only a few exceptions (Purta *et al.*, 2009).

It was found that the *rlmL* gene (previously known as *ycbY*) in *E. coli* encodes a methyltransferase specific for m^{2}G2445 modification of 23S rRNA (Lesnyak *et al.*, 2006). Most prokaryotic rRNA modifications occur in functionally important regions, for example the peptidyltransferase centre of the large subunit or the decoding centre of the small subunit (Decatur & Fournier, 2002). The nucleotide G2445 lies on helix 74 of 23S RNA, which is close to the peptidyltransferase centre in the large ribosomal subunit. Helix 74 is ultra-conserved among all kingdoms of life, particularly the C2065–G2445 pair (Cannone *et al.*, 2002). It is presumed that methylation can prevent base-triplet formation and maintain the C2065–m^{2}G2445 base pairing so that the peptidyltransferase centre maintains an optimal conformation. A knockout experiment showed retarded growth in minimal medium. It has been observed that the RlmL protein interacts with the pre-50S intermediate during the early stage of large ribosomal subunit assembly from 23S rRNA (Sergiev *et al.*, 2007; Jiang *et al.*, 2007).

RlmL is an *S*-adenosyl-L-methionine (SAM) dependent methyltransferase (MTase) that uses SAM as a methyl-group donor. To date, 19 of the 24 rRNA methyltransferases in *E. coli* have been identified (Purta *et al.*, 2009) and 13 of the proteins (or their homologues) have been structurally characterized. However, the structural characteristics of RlmL or homologous proteins are still unavailable. RlmL is a large (~80 kDa) multi-domain protein. A clusters of orthologous groups (COG; Tatusov *et al.*, 1997) search suggested that RlmL contains two distinct methyltransferase domains (Purta *et al.*, 2008). The N-terminal half of RlmL belongs to the COG0116 family and the



C-terminal half belongs to the COG1092 family. Although the functions of the COG0116 and COG1092 families are not fully understood, it is interesting that both families are annotated as methyltransferases.

Here, we report the preliminary crystallographic analysis of the *E. coli* RlmL protein. Hopefully, our structural study will help us to understand the true function and catalytic selectivity of this protein family.

2. Materials and methods

2.1. Cloning, expression and purification

The *E. coli rlmL* gene (gi:16128915; GeneID 945564) was cloned into the pET28a (Novagen) expression vector using the conventional cloning method to form an N-terminal six-His-tagged clone. The primers used were the forward primer 5′-CGCGGATCCATGAATTCTCTGTTTGCCAGTACG-3′ and the reverse primer 5′-CCGCTCGAGTCAGGCTGCGGTAATCAGCC-3′, which contained sequences for *Bam*HI and *Xho*I restriction sites, respectively. *E. coli* strain BL21 (DE3) cells (Invitrogen) were transformed with the expression vector bearing *rlmL* and were grown at 310 K in 1 l Luria–Bertani (LB) medium containing 50 µg ml^{-1} kanamycin. After the culture had been grown to an OD_{600} of 0.6, induction was carried out by the addition of isopropyl β-D-1-thiogalactopyranoside to 0.5 m*M* at 291 K followed by incubation for an additional 20 h.

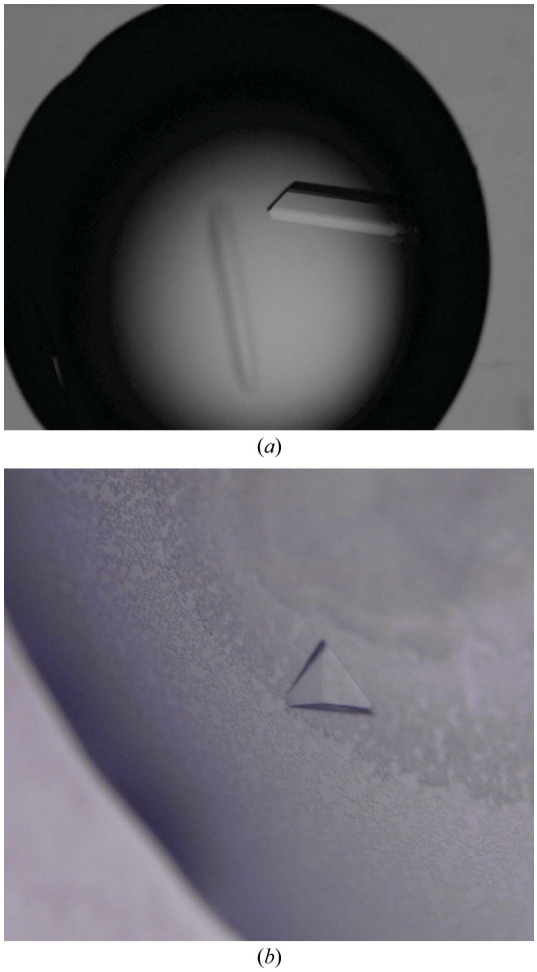


Figure 1
Crystals of the *E. coli* RlmL protein. 2 µl SAH-saturated RlmL protein was mixed with 2 µl reservoir solution consisting of 0.1 *M* HEPES pH 7.0, 12%(*w*/*v*) PEG 8000 using the hanging-drop vapour-diffusion method. (*a*) Rod-shaped crystals obtained without addition of detergent (approximate dimensions 0.05 × 0.1 × 0.6 mm). (*b*) Plate-shaped crystal obtained using *n*-octanoylsucrose (at a final concentration of ∼6%) as an additive after microseeding (approximate dimensions 0.05 × 0.05 × 0.2 mm).

The cells were harvested by centrifugation at 6000*g* for 10 min. The pellet was resuspended in lysis buffer containing 20 m*M* Tris–HCl pH 8.0 and 500 m*M* NaCl and lysed by sonication on ice. The cell lysate was centrifuged at 40 000*g* for 40 min to remove cell debris and the supernatant was applied onto a 5 ml Ni^{2+}-chelating column (HiTrap Chelating HP, GE Healthcare). Loose-binding impurities were removed using washing buffer containing 20 m*M* Tris–HCl pH 8.0, 500 m*M* NaCl and 50 m*M* imidazole. The target-protein peak was eluted using elution buffer (20 m*M* Tris–HCl pH 8.0, 500 m*M* NaCl and 250 m*M* imidazole), collected and concentrated using a centrifugal ultrafiltration device (Amicon Ultra, Millipore) with a molecular-mass cutoff of 30 000 Da. The protein was further purified by gel filtration using a Superdex 75 column (HiLoad, 120 ml column volume, GE Healthcare) pre-equilibrated with buffer consisting of 20 m*M* Tris–HCl pH 8.0 and 200 m*M* NaCl. The protein purity was examined by SDS–PAGE at each step. The purified protein was directly concentrated for crystallization-screening setup using the Amicon Ultra device.

2.2. Crystallization and data collection

Crystallization-screening trials for the RlmL protein were performed at 293 K using the hanging-drop vapour-diffusion method. For initial screening, the protein was diluted with the same buffer as used in gel filtration (20 m*M* Tris–HCl pH 8.0 and 200 m*M* NaCl) to ∼20 mg ml^{-1} as measured by the Bradford assay (Bradford, 1976) using bovine serum albumin as a standard protein. Initial crystallization conditions were screened using the commercial crystallization screening kits Index, Crystal Screen and Crystal Screen 2 (Hampton Research, USA). In each trial, 1 µl protein solution was mixed with 1 µl reservoir solution and equilibrated against 500 µl reservoir solution.

X-ray diffraction data were collected on the BL17U-MX beamline at Shanghai Synchrotron Radiation Facility (SSRF), Shanghai, People's Republic of China using a MAR 225 charge-coupled device (CCD) detector. The crystal was flash-frozen with liquid nitrogen and kept at 100 K in a nitrogen stream to eliminate radiation damage. A total of 360 frames were collected with 1° oscillation per frame. 17%(*v*/*v*) glycerol in the reservoir condition was used as a cryoprotectant. Diffraction data were processed using the *HKL*-2000 program (Otwinowski & Minor, 1997).

3. Results and discussion

RlmL was successfully cloned into the pET28a expression vector and the final vector was verified by DNA sequencing. The molecular mass of the target protein (702 amino-acid residues) was calculated as 78.9 kDa from the sequence. The recombinant protein has an MGSSHHHHHHSSGLVPRGSHMASMTGGQQMGRGS tag at the N-terminus, leading to a final molecular weight of ∼82 kDa as detected on SDS–PAGE. The protein was expressed in a soluble form when induced at a lower temperature (291 K). Recombinant protein with a six-His tag was purified to homogeneity using a two-step chromatographic procedure.

No protein crystals were observed on initial screening using the purified RlmL protein. A few protein crystals appeared when 2 m*M* *S*-adenosylmethionine (SAM) was added to the protein during

Table 1
Data-collection and processing statistics.

Values in parentheses are for the highest resolution shell.

Wavelength (Å)	0.97916
Space group	$P2_1$
Unit-cell parameters (Å, °)	a = 73.6, b = 140.8, c = 102.9, $\alpha = \gamma$ = 90.0, β = 102.3
Resolution (Å)	50–2.20 (2.24–2.20)
$R_{\rm merge}$† (%)	12.1 (68.0)
$I/\sigma(I)$	16.4 (1.6)
Completeness (%)	96.8 (88.2)
No. of observed reflections	689131
No. of unique reflections	99787
Molecules per asymmetric unit‡	2
Solvent content (%)	62.8

† $R_{\rm merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where the summation is over all reflections. ‡ Most probable value.

screening, which is common for crystal growth of SAM-dependent MTases. After optimization, large rod-shaped single crystals were obtained using a reservoir consisting of 0.1 *M* 4-(2-hydroxyethyl) piperazine-1-ethanesulfonic acid (HEPES) pH 7.0 and 12%(*w*/*v*) polyethylene glycol (PEG) 8000. However, these crystals only diffracted to around 10 Å resolution after dehydration and cryo-protection.

It is interesting that the diffraction quality of the RlmL protein crystal improved dramatically when using the detergent *n*-octanoyl-sucrose as an additive and the crystal shape changed from rod-shaped (Fig. 1*a*) to plate-shaped (Fig. 1*b*). Several other detergent molecules from Hampton Research Detergent Screen kits also acted in a similar way (data not shown). Detergents are essential for crystal screening of membrane proteins, but in many cases they have also been found to help in the crystallization of soluble proteins (Guan *et al.*, 2001). The best RlmL protein crystals were obtained using the same reservoir conditions as mentioned above. 2 µl *S*-adenosylhomocysteine-saturated protein solution, 2 µl reservoir solution and 1 µl detergent solution [30%(*w*/*v*) *n*-octanoylsucrose] were mixed for hanging-drop vapour-diffusion crystal growth. The microseeding procedure was used to improve crystal reproducibility. Briefly, the hanging drops were equilibrated against reservoir for 1 h and the crushed crystal microseeds were introduced into the drops with a cat whisker. The crystal diffracted to better than 2.2 Å resolution at the synchrotron and belonged to space group $P2_1$ as indicated by the systematic absences, with unit-cell parameters a = 73.6, b = 140.8, c = 102.9 Å, β = 102.3°. Cell-content analysis suggested that there are most probably two molecules in the asymmetric unit, with a Matthews coefficient of 3.30 Å^3 Da^{-1} (Matthews, 1968) and a corresponding solvent content of 62.8%. The crystallographic parameters and data-collection statistics are listed in Table 1.

There are no RlmL structures currently available in the PDB, but as mentioned above the N-terminal and C-terminal halves of RlmL belong to different COG families. It is suggested that this structure could be solved by the molecular-replacement method using initial models from the COG0116 and COG1092 families. We hope that the structural information will help in understanding the structure–function relationship of the RlmL protein.

This work was supported by funds from Peking University's 985 and 211 projects. We are grateful for the support from the SSRF (Shanghai, China) during data collection.

References

Bradford, M. M. (1976). *Anal. Biochem.* **72**, 248–254.
Cannone, J. J., Subramanian, S., Schnare, M. N., Collett, J. R., D'Souza, L. M., Du, Y., Feng, B., Lin, N., Madabusi, L. V., Muller, K. M., Pande, N., Shang, Z., Yu, N. & Gutell, R. R. (2002). *BMC Bioinformatics*, **3**, 2.
Decatur, W. A. & Fournier, M. J. (2002). *Trends Biochem. Sci.* **27**, 344–351.
Grosjean, H. (2005). *Fine-Tuning of RNA Functions by Modification and Editing.* Berlin/Heidelberg: Springer.
Guan, R., Wang, M., Liu, X. & Wang, D. (2001). *J. Cryst. Growth*, **231**, 273–279.
Jiang, M., Sullivan, S. M., Walker, A. K., Strahler, J. R., Andrews, P. C. & Maddock, J. R. (2007). *J. Bacteriol.* **189**, 3434–3444.
Lesnyak, D. V., Sergiev, P. V., Bogdanov, A. A. & Dontsova, O. A. (2006). *J. Mol. Biol.* **364**, 20–25.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Purta, E., O'Connor, M., Bujnicki, J. M. & Douthwaite, S. (2008). *J. Mol. Biol.* **383**, 641–651.
Purta, E., O'Connor, M., Bujnicki, J. M. & Douthwaite, S. (2009). *Mol. Microbiol.* **72**, 1147–1158.
Rozenski, J., Crain, P. F. & McCloskey, J. A. (1999). *Nucleic Acids Res.* **27**, 196–197.
Sergiev, P. V., Bogdanov, A. A. & Dontsova, O. A. (2007). *Nucleic Acids Res.* **35**, 2295–2301.
Tatusov, R. L., Koonin, E. V. & Lipman, D. J. (1997). *Science*, **278**, 631–637.