PROTEIN STRUCTURE REPORT

Crystal structure of *Streptococcus pneumoniae* Sp1610, a putative tRNA methyltransferase, in complex with S-adenosyl-L-methionine

Hai Minh Ta and Kyeong Kyu Kim*

Department of Molecular Cell Biology, Samsung Biomedical Research Institute, Sungkyunkwan University School of Medicine, Suwon 440-746, Korea

Received 12 November 2009; Revised 9 December 2009; Accepted 11 December 2009

DOI: 10.1002/pro.319

Published online 5 January 2010 proteinscience.org

Abstract: *Streptococcus pneumoniae* Sp1610, a Class-I fold S-adenosylmethionine (AdoMet)-dependent methyltransferase, is a member of the COG2384 family in the Clusters of Orthologous Groups database, which catalyzes the methylation of N¹-adenosine at position 22 of bacterial tRNA. We determined the crystal structure of Sp1610 in the ligand-free and the AdoMet-bound forms at resolutions of 2.0 and 3.0 Å, respectively. The protein is organized into two structural domains: the N-terminal catalytic domain with a Class I AdoMet-dependent methyltransferase fold, and the C-terminal substrate recognition domain with a novel fold of four α -helices. Observations of the electrostatic potential surface revealed that the concave surface located near the AdoMet binding pocket was predominantly positively charged, and thus this was predicted to be an RNA binding area. Based on the results of sequence alignment and structural analysis, the putative catalytic residues responsible for substrate recognition are also proposed.

Keywords: methyltransferase; Sp1610; S-adenosyl-L-methionine; tRNA; crystal

Introduction

A variety of modifications on the nucleosides of transfer RNAs (tRNAs) is necessary to stabilize their tertiary structures and to improve their performance

in the decoding process.^{1,2} Whereas the modification within the anticodon loops is necessary for precise codon pairing and accurate recognition by the cognate aminoacyl-tRNA synthetases, those in the D- and T-loops facilitate the folding and stabilization of the L-shaped tertiary structure.^{3–5} Among these nucleoside modifications, methylations on the bases or riboses of nucleosides which are mediated by RNA methyltransferases (MTase) are noticeable. Methyltransferases utilize the ubiquitous cofactor AdoMet as a methyl donor to transfer the methyl group to the target nucleotides. The AdoMet-dependent MTases are classified into five classes (Class I–V) based on the overall topology and the mode of

Grant sponsor: 21C Frontier Functional Proteomics Program; Grant number: FPR08B2-270; Grant sponsor: Ubiquitome Research Program; Grant number: M105 33010001-05N3301-00100; Grant sponsor: National Research Laboratory Program; Grant number: NRL-2006-02287.

*Correspondence to: Kyeong Kyu Kim, Department of Molecular Cell Biology, Samsung Biomedical Research Institute, Sungkyunkwan University School of Medicine, Suwon 440-746, Korea. E-mail: kkim@med.skku.ac.kr

binding to AdoMet in the catalytic domain.^{6,7} In each class, however, the sequence and structure of the substrate recognition domains are diverse, since they recognize a variety of substrates, such as proteins, DNA, RNA, phospholipids, and small molecules.⁸

Streptococcus pneumoniae Sp1610 is a member of the COG2384 family classified in the Clusters of Orthologous Groups database. The members of this family are found in Gram-negative and Gram-positive bacteria, but no homologues are known in archaea and eukaryota. Sp1610 was identified as a Class I AdoMet-dependent MTase⁹ and, recently, its homologue, *Bacillus subtilis* YqfN, has been established as TrmK that methylates N¹-adenosine at the position 22 of the D-loop of tRNA.¹⁰ Because *B. subtilis* YqfN and *S. pneumoniae* Sp1610 share 42% sequence identity and are categorized into the COG2384 family, they are also considered to share a substrate. Sp1610 has been proposed to be a good target for the discovery of novel antibiotics, since it is well conserved in many bacterial pathogens¹⁰ and is essential for cell viability,^{9,11} but its homologues are not detected in humans.

The structures of Sp1610 and YqfN were modeled on the basis of NMR-studies and threading,⁹ and homology modeling,¹⁰ respectively. However, the modeled structures do not provide the atomic details of the structures of these MTases. In particular, the C-terminal recognition domain cannot be modeled correctly, owing to the lack of sequence homology with any known structures. Therefore, in order to achieve a better understanding of the mechanisms relevant to substrate recognition and further application in the development of novel antibiotics, it is first necessary to assess the atomic details of Sp1610. In this study, we have solved the crystal structures of *S. pneumoniae* Sp1610 in the apo- and the AdoMet-bound forms at 2.0 and 3.0 Å, respectively. This study provided the structural characterization of the first COG2384 family enzyme at high resolution. The results of structural analyses and structure-based sequence alignments allowed us to propose the active site residues, which will facilitate further characterization of this group of enzymes, and hopefully also the development of novel antibacterial agents.

Results and Discussion

Overall structure

Sp1610 exists as a monomer both in solution and in the $P2_12_12_1$ crystal, that is composed of the N-terminal catalytic core domain (residues 1–157) and the C-terminal domain (residues 158–223) [Fig. 1(A)]. The catalytic core domain comprises the canonical class I Rossmann-like methyltransferase fold consisting of a central twisted seven-stranded β -sheet ($\beta 3$ – $\beta 2$ – $\beta 1$ – $\beta 4$ – $\beta 5$ – $\beta 7$ – $\beta 6$) [Fig. 1(A), marine cartoon] flanked by two bundles of helices on both sides [Fig. 1(A), orange

cartoon]. The first five strands of the β -sheet are parallel, whereas the remaining two strands are antiparallel. The catalytic core domain harbors the binding sites for AdoMet, and thus exhibits a high level of sequence conservation among its homologues in the COG2384 family (Fig. 2). Conversely, the C-terminal domain, which contains the four helices $\alpha 6$, $\alpha 7$, $\alpha 8$, and $\alpha 9$, evidences low sequence homology [Fig. 1(A), yellow cartoon]. Eight molecules found in the asymmetric unit of the $P3$ crystal of Sp1610 exhibited minimum conformational changes within the rmsd range of 0.23 Å for the C α atoms. Recently, the New York SGX Structural Genomic Organization deposited the crystal structure of an uncharacterized protein from *Listeria monocytogenes* (DUF633; PDB code 3GNL) which shares high sequence (41 %) and structural homology (0.9 Å rmsd for 223 Ca atoms) with Sp1610.

The DALI¹⁴ search revealed the similarity between the catalytic core domains of Sp1610 and PrmA, a ribosomal protein L11 methyltransferase of *Escherichia coli* (PDB code 2NXE; rmsd of 1.35 Å for 150 equivalent C α positions; Z-score of 16.4), and Trm5, a tRNA m¹G37 methyltransferase of *Methanocaldococcus jannaschii* (PDB code 2ZZM; rmsd of 1.42 Å for 157 equivalent C α positions; Z-score of 15.9). The structural homology detected in the catalytic domains verified the identification of Sp1610 as a Class I AdoMet-dependent MTase. However, no known MTase was identified as the structural homologue of Sp1610 when the C-terminal domain of Sp1610 was subjected to the DALI search. The chaperone protein Hscb and the signal recognition particle SRP have Z-scores of 6.0 and 5.6, respectively, but these scores were not significantly high and they were not functionally related to MTase. These results suggest that Sp1610 harbors a novel fold in its C-terminal domain. Therefore, these results also indicate that the C-terminal domain might not have been correctly predicted in previous modeling studies of Sp1610 and YqfN.^{9,10} For example, the Sp1610 model was predicted to have the highest structural homology with ErmC'MTase.⁹ However, the C-terminal domains of ErmC and Sp1610 cannot be overlapped, owing to the differences between their overall folds⁹ (Fig. 1).

The AdoMet binding site

AdoMet binds to the long and wide pocket located in the center of the N-terminal domain so that the donor-methyl group protrudes outward from the pocket (Fig. 1). The binding of AdoMet is not likely to induce large conformational changes in Sp1610, because the rmsd between the apo- and the ligand-bound structures is 0.54 Å. However, Glu46 and Arg5 moved toward AdMet to form ionic and hydrogen bonds with cofactor [Fig. 1(B)]. AdoMet is stabilized further by several residues in the cofactor binding pocket [Fig.

Figure 1. The crystal structure of Sp1610. A: Ribbon diagram of Sp1610. The secondary structure elements are colored in orange (α -helices in the N-terminal domain), yellow (α -Helices in the C-terminal domain), marine (β -strands in the N-terminal domain), and forest green (loops). The N- and C-terminal regions and each secondary structure element were labeled. B: The AdoMet-binding site. The AdoMet (yellow) and enzyme residues (forest green) are shown as stick models. The $F_o - F_c$ omit map contoured at 3σ represents the bound AdoMet. Hydrogen or ionic bonds are indicated by black dotted lines. Glu46 and Arg5 in the apo-form, which undergo conformational changes upon AdoMet binding, are indicated by a marine color. C: Electrostatic surface representation of Sp1610. The red and blue areas represent negatively and positively charged regions, respectively. AdoMet in the pocket is shown in the yellow stick model. The putative substrate binding pocket is highlighted in the dotted oval box. D: Close-up view of the AdoMet- and putative substrate binding pockets. AdoMet (yellow), AdoMet-binding residues (forest green) and the residues forming the substrate binding pocket (marine) are shown as stick models.

1(B)]: the adenine moiety of AdoMet makes hydrophobic contacts with Val22, Val47, Asn74, Leu97, Met93, and Ile101; the N⁶ atom of adenine and Asn74 form a hydrogen bond; the hydroxyl groups in the ribose ring form hydrogen bonds with Glu46; the carboxylate group is recognized by Arg5 via the charge-enhanced hydrogen bond; the amino group of AdoMet forms hydrogen bonds with the backbone carbonyl groups of Gly23 and Ala91. The residues involved in AdoMet binding are highly conserved in the COG2384 family (Fig. 2).

Members belonging to the Class I AdoMet-dependent MTases family including Sp1610 are expected to have similar AdoMet binding modes. In general, MTases harbor three conserved binding motifs: motif 1 containing the GXGXGG sequences is involved in binding to the amino acid portion; motif 2 containing an acidic and a hydrophobic residue is responsible for binding to the ribose moiety and forming the hydrophobic pocket that stabilizes the adenine ring, respectively; motif 3 binds to the adenine ring via ionic or hydrogen bonds with D/E/N/Q

Figure 2. Multiple sequence alignment of Sp1610 and its homologues in the COG2384 family. The amino acid sequence of *S. pneumoniae* Sp1610 is aligned with those of the representative members of the COG2384 family (the sequence alignment of all COG2384 family members can be seen in Ref. 10). The secondary structure elements and the residue numbers of Sp1610 were indicated above the sequence alignment. α , β , and η indicate α -helices, β -strands and 3_{10} -helices, respectively. The highly conserved and moderately conserved residues are represented by red boxes and red characters, respectively. The AdoMet- and putative substrate binding residues are highlighted above sequences by green and blue dots, respectively. The residues marked with black stars are found in the positively charged area covering the concave surface between the two domains. Sequences were aligned using CLUSTALW¹² and colored with ESPrpt.¹³ Abbreviations: Sp, *Streptococcus pneumoniae*; Bs, *Bacillus subtilis*; Ps, *Pseudomonas syringae*; Ef, *Enterococcus faecium*; Dr, *Deinococcus radiodurans*; Mg, *Mycoplasma genitalium*; Lm, *Listeria monocytogenes*.

residues.¹⁵ These three motifs in Sp1610 were also identified and shown to recognize each part of AdoMet (Figs. 2 and 3) although their sequences with canonical class I AdoMet-dependent MTases does not match very well.

The substrate binding site

The C-terminal domain, together with the N-terminal catalytic domain, forms a concave surface that is surrounded by positively charged residues and located close to the AdoMet binding pocket [Fig. 1(C)]. Such concave surfaces containing positively charged residues are commonly utilized for DNA/RNA recognition. Therefore, this region is hypothesized to be a putative substrate binding site,

which may bind to the negatively-charged phosphate backbone of tRNA via charge-charge interactions.

The AdoMet-dependant methylation involves the direct transfer of the methyl group from AdoMet to the substrate, and thus the target base needs to be positioned close to the methyl group. In the current crystal structure, the methyl moiety of AdoMet points toward a small pocket consisting of the Arg5, Gly92-Met93-Gly94-Gly95, Gln118-Pro119-Asn120, and Tyr150 residues [Fig. 3(B)], which are very well conserved in the members of the COG2384 family (Fig. 2). These results strongly indicate that the residues located in this pocket might perform important functions, such as base recognition.

Figure 3. Structural comparison of Sp1610 and Trm5. **A:** Ribbon diagrams of the catalytic domain of Sp1610 (light blue) and Trm5 (light cyan) are superimposed in stereo view. The N-terminus and the secondary structure elements are labeled. The AdoMet molecules of Sp1610 and Trm5 are also indicated in yellow and orange stick models, respectively. **B:** The AdoMet- and putative substrate binding pockets of Sp1610 (left) and those of Trm5 (right) are compared in the same view with Figure 3(A). AdoMet (yellow for Sp1610-bound and orange for Trm5-bound), guanosine 37 (red), the residues involved in AdoMet binding (forest for both Sp1610 and Trm5) and the residues for substrate binding (blue for Sp1610 and cyan for Trm5) are shown in stick models and labeled. **C:** The structure-guided sequence alignment of the catalytic domain of Sp1610 and Trm5. The residues in motifs 1, 2, and 3 involved in AdoMet recognition are shown with green boxes. The residues involved in the substrate binding are indicated by blue boxes in Trm5. The putative substrate binding residues in Sp1610 are marked by cyan boxes.

Comparison of the active sites of Sp1610 and Trm5

To further investigate the substrate binding mode of Sp1610, its activity was compared with that of Trm5, as their N-terminal catalytic domains share a high degree of structural homology (see the section

on overall structure). The position and orientation of AdoMet in both structures are almost identical, suggesting the conservation of their binding modes, in which two acidic residues are strongly involved in cofactor binding, and the methyl group points in the same direction [Figs. 3(A,B)]. However, the residues

within the putative nucleoside binding sites of Sp1610 and Trm5 do not overlap completely, although the relative orientation between the cofactor and base recognition sites is well-conserved. In Trm5, substrate recognition is attributable to the hydrogen bonds between the 37th guanosine of tRNA and Arg145, Asn265, and Tyr177, as well as to the hydrophobic packing between the guanine base and Tyr177 and Pro267.¹⁶ In this configuration, the N¹ atom of guanine is positioned proximally to the methyl group of AdoMet [Fig. 3(B)]. Considering the orientation of the methyl group of AdoMet to the pocket [Fig. 3(B)] and the high sequence conservation of the residues in the pocket (Fig. 2), it can be assumed that this pocket is a key element in the recognition of nucleoside and that Arg5 and Tyr150 of Sp1610 may functionally correspond to Arg145 and Tyr177 of Trm5. Although Arg5 interacts with AdoMet in the cofactor-bound structure, it is possible that the side chain of Arg5 move to the pocket when Sp1610 binds to the substrate [Fig. 1(B)].

The NPPY motif has been shown to be important for the recognition of the substrate nucleoside in many amino-MTases.¹⁵ In the NPPY motif consisting of Asn265-Leu266-Pro267-Lys269 of Trm5, Asn265 forms a hydrogen bond with the exocyclic 2-amino group of G37 and Pro267 stacks with the G37 guanine ring¹⁶ (Fig. 3). The structure-based sequence alignment and structural overlap between Trm5 and Sp1610 showed that the Gly92-Met93-Gly94-Gly95 in Sp1610 is aligned with the NPPY motif in spite of the sequence variation [Fig. 3(C)]. Therefore, this sequence motif is considered to take a similar role of NPPY motif in Sp1610. Interestingly, the structural superimposition of Sp1610 and Trm5 revealed an additional motif (Gln118-Pro119-Asn120) located in close proximity to the NPPY motif of Trm5 [Fig. 3(B)]. Considering the sequence conservation of this additional motif in COG2384 family proteins, it appears highly likely to participate in substrate recognition.

In conclusion, we have solved the crystal structure of Sp1610, the first crystal structure of a member of the COG2384 family, in its apo- and AdoMet-bound forms. Whereas its AdoMet binding mode is similar to other members of the Class I AdoMet MTases, the oval fold in the C-terminal domain and the putative target binding site differ from those of other known MTases. On the basis of the sequence alignment and structure superimposition of Sp1610 with Trm5, we identified the residues possibly participating in substrate recognition. The biochemical identification of the real substrate of Sp1610 and the structural study of its complex with Sp1610 will allow for complete elucidation of the reaction mechanism of this family of MTases, and will allow for its further application in the development of an antibacterial agent.

Materials and Methods

Protein expression and purification

The gene encoding Sp1610 from *S. pneumoniae* was cloned into the pVFT3S vector (Korean patent 10-0690230), which harbors a Tobacco Etch Virus (TEV) protease cleavage site between the N-terminal 6His-thioredoxin (Trx) and Sp1610. The plasmid was then transformed into *E. coli* BL21(DE3) (Novagen, WI) for expression. Cells were grown in Luria-Bertani medium to an OD⁶⁰⁰ of 0.6 before induction with 1 mM isopropyl- β -D thiogalactopyranoside, and were grown for an additional 24 h at 18°C. Harvested cells were resuspended in resuspension buffer (50 mM Tris-HCl pH 8.0, 500 mM NaCl and 20 mM imidazole), sonicated, and centrifuged for 40 min at 20,000 rpm. The supernatant was then applied to a Ni-NTA column (GE Healthcare, NJ) and proteins were eluted with a linear gradient of imidazole from 50 mM–1.0M. Fractions containing Sp1610 were dialyzed against TEV cutting buffer (25 mM Tris-HCl pH 8.0, 100 mM NaCl and 1 mM DTT), and treated with TEV protease to remove the N-terminal 6His-Trx tag. The resultant solution was applied to a Ni-NTA column to remove cleaved tags. The final stage of purification was conducted via size exclusion chromatography using a Superdex200 column (GE Healthcare, NJ) equilibrated in 25 mM Tris-HCl pH 8.0 and 100 mM NaCl. Pooled fractions containing Sp1610 were concentrated to 20–30 mg/mL with YM-10 Centricon (Millipore, MA) for crystallization. The SeMet-labeling Sp1610 was expressed and purified in the same fashion as described for native protein.

Crystallization and data collection

The initial crystallization screening was conducted at 22°C via the microbatch method using Crystal Screen I, II, and SalRX (Hampton Research, CA) and Wizard I and II (Emerald Biostructures, WA) kits. Protein drops were prepared by mixing 1 μ L of reservoir solution with 1 μ L of Sp1610 solution. Sp1610 was crystallized under two initial conditions: CSII-14 (2.0M ammonium sulfate, 0.1M tri-sodium citrate pH5.6, 0.2M sodium/potassium tartrate) and CSII-28 (1.6M tri-sodium citrate pH 5.6). Further optimization of the initial conditions was required to obtain diffraction-quality crystals by altering the concentration of both precipitants and protein via the hanging-drop method. Finally, native SP1610 was crystallized as a large trigonal prism within a week, with a reservoir consisting of 1.6M tri-sodium citrate at pH 5.6. The crystal belonged to the *P*3 trigonal space group, with the following unit cell parameters: $a = b = 142.6$ Å, $c = 148.1$ Å. Eight molecules are present in an asymmetric unit, with 70% of the crystal volume occupied by solvent. The diffraction-quality crystals from the SeMet-labeled SP1610 were obtained from 2.0M ammonium sulfate, 0.1M tri-sodium citrate pH 5.6, 0.2M

Table I. Summary of Data Collection and Refinement Statistics of Sp1610 (Values in Parentheses Refer to the Last Resolution Shell)

	Sp1610/SAM		SelMet	
Data collection				
Space group	<i>P</i> 3		<i>P</i> 2 ₁ 2 ₁ 2 ₁	
Cell dimensions				
<i>a</i> , <i>b</i> , <i>c</i> (Å)	142.6		35.9	
	142.6		58.8	
	148.1		143.3	
α , β , γ (°)	90, 90, 120		90, 90, 90	
Molecules/AU	8		1	
		Peak	Inflection	Remote
Wavelength (Å)	1.0000	0.97888	0.97917	0.96395
Resolution (Å)	3.0	2.0	2.0	2.1
<i>R</i> _{sym} or <i>R</i> _{merge} ^a	8.7 (63.9)	9.7 (43.96)	10.1 (44.2)	11.7 (41.1)
<i>I</i> / σ <i>I</i>	22.8 (3.8)	14.7 (2.7)	14.5 (2.5)	11.0 (2.0)
Completeness (%)	100 (100)	90.5 (66.9)	89.7 (65.9)	88.9 (64)
Redundancy	5.7 (5.7)	3.1 (2.7)	3.0 (2.6)	3.0 (2.4)
Refinement				
Resolution (Å)	50–3.0	50–2.0		
No. reflections	64,055	97,709		
<i>R</i> _{work} ^b / <i>R</i> _{free} ^c	24.3/29.6	18.9/23.1		
No. atoms				
Protein	1749	1749		
Ligand/ion	108	—		
Water	—	177		
<i>B</i> -factors (Å ²)				
Protein	32.2	30.1		
Ligand/ion	35.4	—		
Water	—	39.2		
R.M.S.D ^c				
Bond lengths (Å)	0.02	0.02		
Bond angles (°)	2.0	1.7		

^a $R_{\text{symm}} = \sum_h \sum_i |I(h_i) - \langle I(h) \rangle| / \sum_h \sum_i I(h_i)$, where $I(h_i)$ is the single intensity of reflection h as determined by the i th measurement and $\langle I(h) \rangle$ is the mean intensity of reflections h .

^b $R_{\text{cryst}} (\%) = \sum |F_o - F_c| / \sum F_o$, where F_o is the observed structure factor amplitude, and F_c is the structure factor calculated from the model.

^c $R_{\text{free}} (\%)$ is calculated in the same manner as R_{cryst} using 5% of all reflections excluded from refinement stages using high resolution data.

^dR.M.S.D, Root-mean-square deviation.

sodium/potassium tartrate within a few days, and were identified to belong to the *P*2₁2₁2₁ space group, with the following unit cell parameters: *a* = 35.9 Å, *b* = 58.8 Å, *c* = 143.3 Å. One molecule was detected in an asymmetric unit with a solvent content of 55.0%. AdoMet was soaked into the crystals in the *P*3 space group.

The diffraction data for the SeMet crystals and the native crystals were collected at resolutions of 2.0–2.1 and 3.0 Å, respectively, using an ADSC Quantum 210 CCD detector at beam line AR-NW12 of the Photon Factory, Japan. Prior to data collection, the crystals were cryoprotected by adding glycerol to a final concentration of 20–25% (v/v) and flash-frozen in a cold nitrogen stream. The exposure times varied between 5 and 10 s for 1° oscillations. The diffraction data were processed and scaled with the *DENZO* and *SCALEPACK* programs from the *HKL* program suite.¹⁷ The data collection and refinement statistics are summarized in Table I.

Structure determination and refinement

The crystal structure of Sp1610 was solved by MAD phasing at a resolution of 2.0 Å. Three selenium binding sites, detected using the SOLVE¹⁸ program, were used for phase calculation. The MAD-phased electron density map had high quality and more than 90% residues were correctly assigned by the RESOLVE¹⁹ and ARP/WARP²⁰ automatic model building programs. Subsequent manual model building was conducted using the COOT program.²¹ The model was refined with the REFMAC program,²² including the bulk solvent correction. The final model consists of residues 1–223 with an *R* factor of 18.9% and an *R*-free of 23.1%. The two C-terminal residues (Val224–Lys225) are disordered in the crystal, and could not be built. The refined model, evaluated by the PROCHECK program,²³ has excellent stereochemistry, with 94.2 and 5.8% residues belonging to the most favored and additionally allowed regions, respectively.

The AdoMet-bound structure in the *P3* space group was solved via molecular replacement (MR) using the MOLREP program²⁴ with the SeMet structure as the search model. The $F_o - F_c$ electron-density maps calculated with MR phases and contoured at 3σ revealed unambiguous density of four AdoMet molecules bound to eight protein molecules within the asymmetric unit. The AdoMet molecules were introduced into $F_o - F_c$ difference maps using COOT and refined with REFMAC. The refinement statistics of the ligand-free and AdoMet-bound structures are provided in Table I. All structural figures were generated with the PYMOL program.²⁵ Coordinates and structure factors for the apo- and AdoMet-bound Sp1610s have been deposited with accession codes **3KR9** and **3KU1**, respectively.

References

- Dunin-Horkawicz S, Czerwoniec A, Gajda MJ, Feder M, Grosjean H, Bujnicki JM (2006) MODOMICS: a database of RNA modification pathways. *Nucleic Acids Res* 34:D145–D149.
- Sprinzel M, Horn C, Brown M, Ioudovitch A, Steinberg S (1998) Compilation of tRNA sequences and sequences of tRNA genes. *Nucleic Acids Res* 26:148–153.
- Martínez Giménez JA, Sáez GT, Seisdedos RT (1998) On the function of modified nucleosides in the RNA-world. *J Theor Biol* 194:485–490.
- Vermeulen A, McCallum SA, Pardi A (2005) Comparison of the global structure and dynamics of native and unmodified tRNAval. *Biochemistry* 44:6024–6033.
- Nakanishi K, Nureki O (2006) Recent progress of structural biology of tRNA processing and modification. *Mol Cells* 19:157–166.
- Schubert HL, Blumenthal RM, Cheng X (2003) Many paths to methyltransfer: a chronicle of convergence. *Trends Biochem Sci* 28:329–335.
- Kozbial PZ, Mushegian AR (2005) Natural history of S-adenosylmethionine-binding proteins. *BMC Struct Biol* 14:5–19.
- Martin JL, Fiona M, McMillan FM (2002) SAM (dependent) I AM: the S-adenosylmethionine-dependent methyltransferase fold. *Curr Opin Struct Biol* 12:783–793.
- Constantine KL, Krystek SR, Healy MD, Doyle ML, Siemers NO, Thanassi J, Yan N, Xie D, Goldfarb V, Yanchunas J, Tao L, Dougherty BA, Farmer BT, II (2005) Structural and functional characterization of CFE88: evidence that a conserved and essential bacterial protein is a methyltransferase. *Protein Sci* 14:1472–1484.
- Roovers M, Kaminska KH, Tkaczuk KL, Gigot D, Droogmans L, Bujnicki JM (2008) The YqfN protein of *Bacillus subtilis* is the tRNA: m¹A22 methyltransferase (TrmK). *Nucleic Acids Res* 36:3252–3262.
- Thanassi JA, Hartman-Neumann SL, Dougherty TJ, Dougherty BA, Pucci MJ (2002) Identification of 113 conserved essential genes using a high-throughput gene disruption system in *Streptococcus pneumoniae*. *Nucleic Acids Res* 30:3152–3162.
- Thompson JD, Higgins DG, and Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 22:4673–4680.
- Gouet P, Robert X, Courcelle E (2003) ESPript/ENDscript: extracting and redrawing sequence and 3D information from atomic structures of proteins. *Nucleic Acids Res* 31:3320–3323.
- Holm L, Sander C (1998) Touring protein fold space with Dali/FSSP. *Nucleic Acids Res* 26:316–319.
- Bheemanaik S, Reddy YV, Rao DN (2006) Structure, function and mechanism of exocyclic DNA methyltransferases. *Biochem J* 399:177–190.
- Goto-Ito S, Ito T, Kuratani M, Bessho Y, Yokoyama S (2009) Tertiary structure checkpoint at anticodon loop modification in tRNA functional maturation. *Nat Struct Mol Biol* 16:1109–1115.
- Otwiński Z, Minor W (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* 276:307–326.
- Terwilliger TC, Berendzen J (1999) Automated MAD and MIR structure solution. *Acta Crystallogr Sect D* 55:849–861.
- Terwilliger TC (2003) Automated main-chain model building by template matching and iterative fragment extension. *Acta Crystallogr Sect D* 59:38–44.
- Morris RJ, Perrakis A, Lamzin VS (2003) ARP/wARP and automatic interpretation of protein electron density maps. *Methods Enzymol* 374:229–244.
- Emsley P, Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr Sect D* 60:2126–2132.
- Murshudov GN, Vagin AA, Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr Sect D* 53:240–255.
- Laskowski RA, MacArthur MW, Moss DS, Thornton JM (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *J Appl Crystallogr* 26:283–291.
- Vagin A, Teplyakov A (2000) An approach to multi-copy search in molecular replacement. *Acta Cryst D* 56:1622–1624.
- DeLano WL (2002) The PyMOL molecular graphics system on world wide web. Available at: <http://www.pymol.org>.