Published online 9 November 2010 *Nucleic Acids Research, 2011, Vol. 39, No. 5* ***1903–1918***
doi:10.1093/nar/gkq1033

Structural basis for the methylation of A1408 in 16S rRNA by a panaminoglycoside resistance methyltransferase NpmA from a clinical isolate and analysis of the NpmA interactions with the 30S ribosomal subunit

Nilofer Husain[1], Sonja Obranić[2], Lukasz Koscinski[3], J. Seetharaman[4], Fedora Babić[2], Janusz M. Bujnicki[3,5,*], Gordana Maravić-Vlahoviček[2,*] and J. Sivaraman[1,*]

[1]Department of Biological Sciences, 14 Science drive 4, National University of Singapore, Singapore 117543, [2]Faculty of Pharmacy and Biochemistry, Department of Biochemistry and Molecular Biology, University of Zagreb, Ante Kovačića 1, 10000 Zagreb, Croatia, [3]Laboratory of Bioinformatics, Institute of Biotechnology and Molecular Biology, Faculty of Biology, Adam Mickiewicz University, Umultowska 89, 61-614 Poznan, Poland, [4]X4 Beamline, Brookhaven National Laboratory, Upton, NY, USA and [5]Laboratory of Bioinformatics and Protein Engineering, International Institute of Molecular and Cell Biology, Trojdena 4, 02-109 Warsaw, Poland

Received September 3, 2010; Revised October 5, 2010; Accepted October 10, 2010

ABSTRACT

NpmA, a methyltransferase that confers resistance to aminoglycosides was identified in an *Escherichia coli* clinical isolate. It belongs to the kanamycin–apramycin methyltransferase (Kam) family and specifically methylates the 16S rRNA at the N1 position of A1408. We determined the structures of apo-NpmA and its complexes with *S*-adenosylmethionine (AdoMet) and *S*-adenosylhomocysteine (AdoHcy) at 2.4, 2.7 and 1.68 Å, respectively. We generated a number of NpmA variants with alanine substitutions and studied their ability to bind the cofactor, to methylate A1408 in the 30S subunit, and to confer resistance to kanamycin *in vivo*. Residues D30, W107 and W197 were found to be essential. We have also analyzed the interactions between NpmA and the 30S subunit by footprinting experiments and computational docking. Helices 24, 42 and 44 were found to be the main NpmA-binding site. Both experimental and theoretical analyses suggest that NpmA flips out the target nucleotide A1408 to carry out the methylation. NpmA is plasmid-encoded and can be transferred between pathogenic bacteria; therefore it poses a threat to the successful use of aminoglycosides in clinical practice. The results presented here will assist in the development of specific NpmA inhibitors that could restore the potential of aminoglycoside antibiotics.

INTRODUCTION

Aminoglycosides are bactericidal antibiotics that are widely used in treatment against severe infectious diseases caused by Gram-negative and Gram-positive bacteria. They bind to the 30S subunit of the ribosome at the A-site and induce a conformational change that is responsible for the loss of fidelity of the protein translation process (1). Due to the indiscriminate and excessive use of these antibiotics, pathogenic bacteria have acquired from antibiotic producers and further evolved several mechanisms of resistance to these antibiotics such as: alteration of the antibiotic to inactivate it, mutation on the small ribosomal subunit, which modifies the target site of the antibiotic, or methylation at specific sites on the 16S rRNA that prevent the binding of the antibiotic to the ribosome (2). Mutation of A to G at position 1408 in the 16S rRNA has been found in clinical isolates of mycobacteria (3). However, methylation provides high-level resistance to most aminoglycosides used in the clinical practice. It is carried out by specific methyltransferases that are found in two classes of bacteria: the antibiotic producers and the

*To whom correspondence should be addressed. Tel: +65 6516 1163; Fax: +65 6779 2489; Email: dbsjayar@nus.edu.sg
Correspondence may also be addressed to Janusz M. Bujnicki. Tel: +48 22 597 0750; Fax: +48 22 597 0715; Email: iamb@genesilico.pl
Correspondence may also be addressed to Gordana Maravić-Vlahoviček. Tel: +385 1 639 4448; Fax: +385 1 639 4400; Email: gordana@pharma.hr

pathogens. In antibiotic producers, the gene encoding the methyltransferase is usually found on the chromosome, whereas in pathogens it is typically present on an extrachromosomal plasmid (4).

There are two families of methyltransferases that are responsible for the resistance against aminoglycosides. These are the Arm (aminoglycoside resistance methyltransferase) family (5) and the Kam (kanamycin–apramycin methyltransferase) family (6), which specifically methylate the 16S rRNA at the N7 position of G1405, and the N1 position of A1408, respectively. Kam enzymes confer high-level resistance to broad spectrum aminoglycosides. KamA and KamB have been found in antibiotic producers, while NpmA from *Escherichia coli* clinical strain ARS3 is the first methyltransferase found in a clinical isolate that can methylate 16S rRNA at m^1A1408 (7). In addition, the *npmA* gene was found on a plasmid, so its rapid horizontal transfer among bacterial pathogens is expected, which poses a threat to the successful use of aminoglycoside antibiotics in treatment of infectious diseases.

Here, we report the crystal structure of NpmA and its complex with AdoMet and AdoHcy, along with structure-guided functional and modeling studies. Residues conserved in the Kam family were substituted with alanine and the resulting NpmA variants were analyzed for their methylation activity, for their ability to bind AdoMet and AdoHcy (by isothermal titration calorimetry), and for their ability to confer resistance to kanamycin *in vivo*. In addition, we analyzed the interaction of NpmA with the small ribosomal subunit by footprinting analyses and computational docking. These studies provide insight into the mechanism of substrate recognition and methylation of A1408 by NpmA. The NpmA structure and the knowledge of its interaction with the substrate will serve as a starting point for the development of specific inhibitors of the Kam family of methyltransferases that could reinstate aminoglycoside sensitivity in aminoglycoside-resistant pathogens.

MATERIALS AND METHODS

Cloning and site-directed mutagenesis

The *npmA* gene synthesis was ordered from the Epoch Biolabs Inc. following the sequence published in ref. (7). The *npmA* gene was subsequently recloned into NdeI and EcoRI sites of the pET-25b (+) vector with the addition of a C-terminal non-cleavable $(His)_6$-tag. The amino acid sequence of NpmA is devoid of methionine. In order to facilitate the structure determination through SelMet-labeled protein, we mutated four codons for leucine residues (L31, L90, L128 and L196) to methionine codons by inverse PCR (8) using Phusion® high-fidelity DNA polymerase (Finnzymes). Mutations leading to alanine substitutions of selected conserved amino acids (D30, D55, E88, P106, W107, W109, F177, S195, W197, K199, R200 or R205) were introduced either by using Platinum® Pfx DNA polymerase (Invitrogen) in the PCR-overlapping method (9) or by inverse PCR using Phusion® high-fidelity DNA polymerase (Finnzymes). All introduced mutations were verified by DNA sequencing. Deoxyoligonucleotides used in this work are listed in Supplementary Table S1.

Expression and purification

E. coli BL21 (DE3) was transformed with recombinant vector pET-25b(+) carrying the *npmA* gene for protein expression. *E. coli* cells were cultured in 1 l of LB medium supplemented with ampicillin (100 mg/l) at 37°C until the OD_{600} reached 0.5–0.6 AU. The expression was induced with 150 μM IPTG and cells continued to grow at 20°C overnight. Cells were harvested by centrifugation (9000*g*, 30 min, 4°C), and the pellet was resuspended in 40 ml of lysis buffer [50 mM HEPES sodium pH 7.5, 250 mM NaCl, 10% glycerol, 0.5% Triton-X 100, 5 mM β-mercaptoethanol (BME) and 1 protease inhibitor tablet (Roche)].

After sonication the cell lysate was centrifuged at 39191*g* for 30 min at 4°C. The supernatant was allowed to bind to the Ni-NTA agarose beads (Qiagen) for 1 h at 4°C and was subsequently washed with wash buffer I (50 mM HEPES sodium, pH 7.5, 250 mM NaCl, 10% glycerol, 5 mM BME and 5 mM imidazole, pH 7.5), wash buffer II (50 mM HEPES sodium, pH 7.5, 1.0 M NaCl, 10% glycerol, 5 mM BME and 10 mM imidazole, pH 7.5) and wash buffer III (50 mM HEPES sodium, pH 7.5, 250 mM NaCl, 10% glycerol, 5 mM BME and 15 mM imidazole, pH 7.5). Finally, the protein was eluted in three steps using buffer A (50 mM HEPES sodium, pH 7.5, 250 mM NaCl, 10% glycerol, 5 mM BME, 200 mM imidazole, pH 7.5), buffer B (50 mM HEPES sodium, pH 7.5, 250 mM NaCl, 10% glycerol, 5 mM BME and 250 mM imidazole, pH 7.5) and buffer C (50 mM HEPES sodium, pH 7.5, 250 mM NaCl, 10% glycerol, 5 mM BME and 300 mM imidazole, pH 7.5). The eluted protein was loaded onto a size exclusion column (Superdex 200, GE Healthcare) equilibrated with the buffer containing 20 mM HEPES sodium, pH 7.5, 250 mM NaCl, 5% glycerol, 5 mM BME and 10 mM $MgCl_2$. The eluted NpmA was concentrated up to 11 mg/ml. A similar protocol was adopted to purify the selenomethionine (SelMet) labeled NpmA using the M9 medium (10).

NpmA variants with alanine substitutions used for analysis in functional assays were expressed and purified following a simplified procedure. *E. coli* BL21(DE3)pLysS cells harboring the expression vector pET-25b(+) with different *npmA* variants were grown in LB medium supplemented with ampicillin (100 mg/l) and chloramphenicol (34 mg/l) at 37°C until the OD_{600} of 0.6–0.8. The expression was induced with 1 mM IPTG and carried out at 37°C for 3 h with the addition of 3% ethanol. Cells were pelleted by centrifugation (3500 *g*, 10 min, 4°C), frozen at −80°C and the pellet was resuspended in 40 ml of lysis buffer [50 mM NaH_2PO_4, pH 8.0, 300 mM NaCl, 10 mM imidazole and a cocktail of protease inhibitors (Roche)]. Following sonication, the debris was removed by centrifugation at 20 000 *g* for 20 min at 4°C. Clarified supernatant was incubated for an hour at 4°C with the Ni-NTA Agarose (Qiagen) equilibrated with the lysis

buffer. The resin with the bound NpmA variant was subsequently washed with the wash buffer (50 mM NaH_2PO_4, pH 8.0, 300 mM NaCl, 20 mM imidazole). The protein was eluted with the elution buffer (50 mM NaH_2PO_4, pH 8.0, 250 mM NaCl, 5% glycerol, 5 mM BME, 200 mM imidazole) and dialyzed against 20 mM Tris–HCl pH 7.5, 250 mM NH_4Cl, 10 mM $MgCl_2$, 6 mM BME, 10% glycerol. All NpmA variants were purified to 95% homogeneity judged by the SDS-polyacrylamide gel electrophoresis.

Crystallization and data collection

Purified NpmA was complexed with the cofactor *S*-adenosylmethionine (AdoMet) and reaction by-product *S*-adenosylhomocysteine (AdoHcy) at 1:12 (NpmA:AdoMet/AdoHcy) molar ratio. Crystallization trials were carried out at room temperature by hanging-drop vapor-diffusion method using crystallization screens from Hampton Research (Aliso Viejo, CA, USA) and Jena Bioscience screens (Jena, Germany). Initially crystals of NpmA and NpmA–AdoMet/AdoHcy complexes were obtained from the Jena Bioscience screens. Subsequently these conditions were optimized to obtain the diffraction quality crystals. The best crystals of NpmA alone or with AdoMet/AdoHcy were obtained by mixing 1 µl of concentrated protein with 1 µl crystallization solution [0.1 M citrate pH 5.5, 20% (w/v) PEG 4000 and 10% (v/v) 2-propanol for NpmA and NpmA-AdoMet complex; 30% (w/v) PEG 4000, 0.1 M sodium acetate, pH 4.6 and 0.2 M ammonium sulfate for NpmA-AdoHcy complex] and were grown up to 8 days.

Prior to data collection, crystals were briefly soaked in a cryoprotectant consisting of the paraton and mineral oil in a 1:1 ratio, and flash cooled in a N_2 cold stream (100 K). A complete single wavelength anomalous dispersion (SAD) data set was collected for the NpmA–AdoHcy complex at the beamline X8C, NSLS, Brookhaven National Laboratory using a Quantum4-CCD detector (Area Detector Systems Corp Poway, CA, USA) to 1.9 Å resolution. The NpmA data set was collected using the rotating anode generator (Bruker X8 Proteum) equipped with PLATINUM135 CCD detector, whereas the high-resolution NpmA–AdoHcy and NpmA–AdoMet complex datasets were collected at beamline BL13B1, NSRRC (Taiwan) using a ADSC Quantum-315 CCD Area Detector. All the data sets were processed and scaled using the program HKL2000 (11). The crystallographic and refinement statics are provided in Table 1.

Structure solution and refinement

The expected eight selenium sites in one asymmetric unit of the NpmA–AdoHcy complex crystal were located using the program Solve (12). The overall figure of merit was 0.77. The resultant electron density map was of good quality to automatically build over 80% of the model using wARP (13). The remaining part of the model was manually built using the program COOT (14). Alternating cycles of model building and refinement using the programs CNS and REFMAC was carried out (15, 16) with appropriate entries made in the dictionary for AdoHcy. The final *R*-factor is 0.21 (R_{free} is 0.24) up to 1.9 Å resolution. The structure of apo-NpmA, NpmA–AdoMet complex and the high-resolution NpmA–AdoHcy complex (1.68 Å) was solved by molecular replacement method using MOLREP (17) with the NpmA coordinates of the NpmA–AdoHcy complex as a search model. The overall geometry of the final models was analyzed by PROCHECK (18).

Isothermal titration calorimetry (ITC)

NpmA in a buffer consisting of 20 mM HEPES sodium, pH 7.5, 250 mM NaCl, 5% glycerol, 5 mM BME and 10 mM $MgCl_2$ was titrated against *S*-adenosyl methionine (AdoMet/AdoHcy, MP Biomedicals) solution prepared by dissolving the AdoMet in the same buffer as the protein to a final concentration of 100 mM. The ITC experiments were carried out using VP-ITC calorimeter (Microcal, LLC) at 20°C using 0.02–0.07 mM protein in the sample cell and 2–5 mM AdoMet/AdoHcy in the injector. All samples were thoroughly degassed and centrifuged. Injection volumes of 8–10 µl per injection were used for different experiments and for each experiment the heat of dilution for the ligand was measured. To restore the baseline, successive injections were separated by at least 4 min. The ITC data was analyzed by a single site-fitting model using Origin 7.0 (OriginLab Corp.) software.

Determination of the minimum inhibitory concentration of kanamycin

Kanamycin minimum inhibitory concentration (MIC) was determined in *E. coli* BL21(DE3)pLysS cells carrying recombinant pET-25b(+) vector with different *npmA* variants using the MIC dilution assay with small modifications of a procedure described previously. In brief, overnight cultures of BL21 (DE3) pLysS cells carrying mutant genes in pET-25b (+) vector were diluted 100-fold in fresh LB medium supplemented with ampicillin (100 mg/l) and chloramphenicol (34 mg/l) and grown until the OD_{600} of 1. Two hundred microliter of LB medium containing ampicillin, chloramphenicol and various concentration of kanamycin (0.5, 1, 2, 4, 8, 16, 32, 64, 128, 256, 512 and 1024 mg/l) was inoculated with 10^5 cells and grown in microtiter plates in total volume for 18 h at 37°C. MIC was determined as a minimal concentration of kanamycin that prevents the visible growth of bacteria.

Isolation of 30S ribosomal subunits

Tetra-$(His)_6$-tagged 70S ribosomes were isolated from *E. coli* JE28 cells (obtained from Dr Suparna Sanyal from Uppsala University, Sweden) using nickel-affinity chromatography, as a modification of a method described in ref. (19). Cells were harvested at OD_{600} of 1, cooled slowly to 4°C, pelleted and resuspended in the lysis buffer [20 mM Tris–HCl, pH 7.6, 10 mM $MgCl_2$, 150 mM KCl, 30 mM NH_4Cl, 0.5 g/l lysozyme, 10 mg/l DNaseI, protein inhibitor cocktail (Roche)]. The suspension was lysed by sonication and cell debris was removed by two centrifugation cycles at 20 000*g* for 10 min. Lysate was applied to a HiTrap™ Chelating HP column (1 ml, GE Healthcare) equilibrated with the lysis buffer. Column

Table 1. Crystallographic data and refinement statistics

Data set	AdoHcy		AdoMet	Apo-NpmA[b]
	SAD	High resolution[b]		
Cell parameters (Å, °) and Space group	a = 64.42, b = 64.42, c = 108.50, $P4_3$	a = 64.98, b = 64.98, c = 108.42 $P4_3$	a = 75.78, b = 64.42, c = 104.05, β = 100.08 $P2_1$	a = 64.86, b = 64.86, c = 108.61, $P4_3$
Data collection				
Resolution range (Å)	50–1.9 (1.93-1.90)	50–1.68 (1.71–1.68)	50–2.7 (2.75–2.70)	50–2.3 (2.34–2.3)
Wavelength (Å)	0.979	1.000	1.000	1.541
Observed reflections	422 988	381 987	75 750	153 501
Unique reflections	35 024	51 311	26 821	19 348
Redundancy	12.1	7.4	2.9	8.0
Completeness (%)	100 (100%)	100 (100%)	97.6 (95.2%)	96.5 (76.3%)
Overall ($I/\sigma I$)	17.0	12.7	11.6	10.8
R_{sym}[a]	0.08 (0.35)	0.07 (0.43)	0.13 (0.31)	0.102 (0.26)
Refinement and quality				
Resolution range (Å)	20–1.90	20–1.68	50–2.7	15–2.4
R_{work}[c] (no. of reflections)	0.21 (31 257)	0.196 (41 928)	0.252 (21 257)	0.20 (14 201)
R_{free}[d] (no. of reflections)	0.24 (1745)	0.232 (4701)	0.284 (2677)	0.26 (1576)
RMSD bond lengths (Å)	0.008	0.005	0.010	0.006
RMSD bond angles (°)	1.09	1.226	1.398	1.15
Average B-factors[e] (Å^2)				
Main chain (# atoms)	20.14 (852)	19.39 (860)	44.41 (844)[f]	30.04 (852)
Side chain (# atoms)	22.19 (745)	26.46 (748)	46.40 (728)[f]	36.64 (741)
Ligand (# atoms)	20.05 (27)	17.25 (27)	24.35 (28)	–
Waters (# atoms)	32.52 (321)	35.25 (489)	34.225 (74)	40.37 (259)
Ramachandran plot				
Most favored regions (%)	96.3	95.3	87.9	92.7
Additional allowed regions (%)	3.7	4.2	11.9	7.3
Generously allowed regions (%)	0.0	0.6	0.1	0.0
Disallowed regions (%)	0.0	0.0	0.0	0.0

[a]$R_{sym} = \Sigma|Ii - <I>|/\Sigma|Ii|$, where I_i is the intensity of the i-th measurement, and $<I>$ is the mean intensity for that reflection.
[b]For all models, reflections with $I > \sigma I$ was used in the refinement.
[c]$R_{work} = \Sigma||F_{obs}|-|F_{calc}||/\Sigma|F_{obs}|$, where F_{calc} and F_{obs} are the calculated and observed structure factor amplitudes, respectively.
[d]R_{free} = as for R_{work}, but for 10% (5% for NpmA–AdoHcy SAD data) of the total reflections chosen at random and omitted from refinement.
[e]Individual B-factor refinements were calculated.
[f]Four molecules (chains A–D) in the asymmetric unit. D chains shows high B-factor values and it is not included in the averaging. However, it has continuous well-defined electron density map.

was washed with the lysis buffer until A_{260} reached the baseline. Tetra-$(His)_6$-tagged 70S ribosomes were eluted with the elution buffer (20 mM Tris–HCl, pH 7.6, 10 mM $MgCl_2$, 150 mM KCl, 30 mM NH_4Cl, 150 mM imidazole) and dialyzed four times for 10 min against the low-Mg^{2+} buffer (20 mM Tris–HCl, pH 7.6, 1 mM $MgCl_2$, 150 mM KCl, 30 mM NH_4Cl) to enable separation of subunits and then applied to a HiTrap™ Chelating HP column (1 ml, GE Healthcare) equilibrated with the same buffer. As 30S subunits lack His-tag, they were collected in the flow through and concentrated by ultrafiltration (Amicon Ultra-15 Centrifugal Filter Unit with Ultracel-30 membrane, cutoff 30 kDa).

Methylation assay

The activity of NpmA and its variants was determined using a modification of a procedure described in Axis® Homocysteine Enzyme Immuno Assay kit (Axis-Shield Diagnostics Ltd.). The concentration of AdoHcy produced in the NpmA-catalyzed methyl transfer is proportional to the concentration of methylated RNA and production of AdoHcy (or indirectly m^1A1408) and is monitored in a competitive ELISA assay using anti-AdoHcy antibodies. Reactions for *in vitro* methylation contained 4 pmol of 30S subunits, 100 pmol of AdoMet (Sigma) and 8 pmol of NpmA. All reactions were carried out for 45 min at 37°C in 50 µl of 20 mM Tris–Cl, pH 6.5, 100 mM NH_4Cl, 10 mM $MgCl_2$, 6 mM BME with 10 % glycerol. Reactions were stopped by the addition of 1 µl of 0.5 M EDTA. Twenty-five microliter of reaction mixtures were then transferred to the microtiter plate wells coated with immobilized AdoHcy-BSA conjugate and incubated 30 min with 200 µl of mouse monoclonal anti-AdoHcy antibodies. Subsequently, wells were washed four times with 350 µl of the wash buffer provided in the kit and incubated 20 min with 100 µl of rabbit anti-mouse antibodies conjugated with the horseradish peroxidase. Following the wash step, the detection was initiated with the addition of 100 µl of the peroxidase substrate. Reaction was terminated after 20 min by the addition of 100 µl of sulphuric acid and absorbance was measured on 450 nm using a microplate reader. Concentration of AdoHcy was determined using AdoHcy calibration curve that was prepared using control AdoHcy samples provided in the kit. Duplicate samples were measured in three independent experiments in all cases.

Analysis of rRNA by primer extension

Total rRNA was extracted from *E. coli* BL21(DE3)pLysS cells carrying the expression vector pET-25b(+) with and without the variants of the *npmA* gene using GenElute Mammalian Total RNA Miniprep Kit (Sigma). Briefly, cells were harvested in early log phase (OD_{450} of 0.4–0.5), lysed by sonication and cell debris was removed by two centrifugation cycles at 20 000*g* for 10 min. Total RNA was then prepared from the lysates using the RNA cleanup protocol. A 5′-Cy5-modified deoxynucleotide primer, complementary to 16S rRNA from nucleotides 1459 to 1479, was hybridized to the rRNA and extended with AMV reverse transcriptase (Finnzymes) (20). Sequencing reactions were performed on rRNA isolated from *E. coli* cells lacking NpmA. Extension products were separated on 6% denaturing polyacrylamide gels and were visualized on Typhoon 9210 imager (GE Healthcare).

Chemical probing of 30S subunit–NpmA complex

Chemical probing was performed with DEPC and CMCT as described in (21). Briefly, 10 pmol of 30S subunits were incubated with a 2-fold molar excess of NpmA and 1 mM AdoHcy in either modification buffer A (50 mM HEPES-KOH, pH 8.0, 10 mM $MgCl_2$, 100 mM KCl, 5 mM dithiothreitol) or B (200 mM HEPES-KOH, pH 7.8, 10 mM $MgCl_2$, 100 mM KCl) for 30 min at 37°C in 50 µl reaction mixtures. Control reactions contained no AdoHcy. The complexes were modified with either 50 µl of CMCT (42 mg/ml in modification buffer A) for 10 min at 37°C, or 1 µl of DEPC for 30 min at room temperature. RNA was purified immediately with a GenElute Mammalian Total RNA Miniprep Kit (Sigma), RNA cleanup protocol. Chemical modifications were detected by primer extension analysis using AMV reverse transcriptase (Finnzymes) and 5′-Cy5 modified deoxyoligonucleotides complementary to the 16S rRNA sequences 1459–1479, 939–955 and 817–833. The cDNA products of primer extension reactions were separated on 6% polyacrylamide sequencing gel and visualized on Typhoon 9210 imager (GE Healthcare).

Bioinformatics analyses

To identify orthologs of NpmA, we used its sequence as a query to search the non-redundant protein database with PSI-BLAST (22). Sequences with significant similarity to NpmA were retrieved and aligned using MUMMALS (23). The alignment was refined based on the results of structure predictions obtained from the GeneSilico MetaServer (24). Based on structural superpositions onto NpmA, the sequences of TrmB, TrmK, TrmI and Sgm proteins were added to the alignment for comparison.

A model of the wt NpmA protein (with SelMet reverted to Leu) was constructed using SwissModel (25). GRAMM 1.03 (26) was used to generate 10 000 alternative docking models (poses) between the NpmA–AdoMet complex and the crystal structure of the small ribosomal subunit (Protein Data Bank code 2I2P). The search was restricted to the area within 50 Å radius from the A1408 nucleotide residue. The following parameters were used: energy scores for repulsion and attraction were 30 and 0, respectively, grid size 64 Å, grid step 2.8 Å. The NpmA structure was rotated with 10° angle intervals. Subsequently, all 10 000 poses were filtered using the Filtrest3D server (27), to rank them according to the proximity of the N1 atom of A1408 and the methyl group of AdoMet bound to NpmA. Top 100 variants have been clustered with MAXCLUSTER (http://www.sbg.bio.ic.ac.uk/~maxcluster/index.html), and a member of the largest cluster that exhibited best shape complementarity and minimal number of clashes between NpmA and the 30S subunit was chosen as the representative model.

RESULTS

Overall structure

The structure of NpmA–AdoHcy complex was solved by the SAD experiment, while the structures of the apo-NpmA and NpmA–AdoMet complex were solved by molecular replacement using NpmA–AdoHcy coordinates as a search model. The electron density for the first and the last residue, loop region (aa 144–157) and the C-terminal His-tag are not well defined and these residues are not included in the model (Figure 1). The models were refined to low *R*-values with good stereochemical parameters (Table 1). In apo-NpmA and NpmA–AdoHcy structures the asymmetric unit consists of two molecules, whereas in the NpmA–AdoMet complex, the asymmetric unit has four complex molecules. However, crystal contacts are not very extensive;

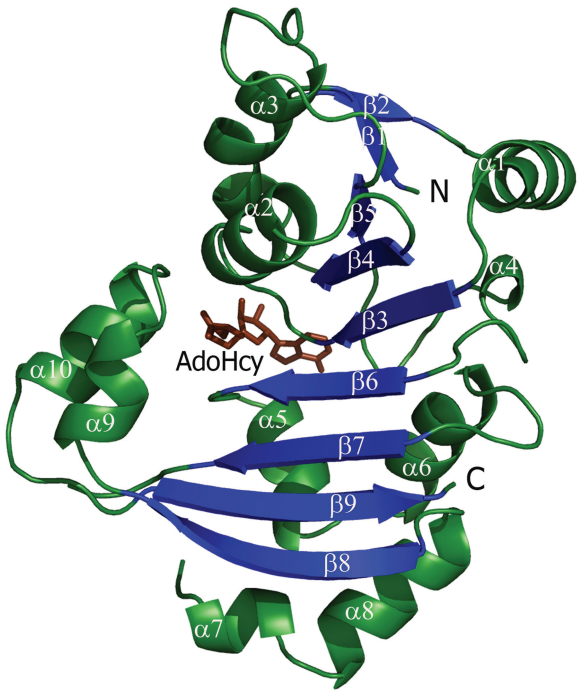

Figure 1. Structure of NpmA. Ribbon representation of NpmA-cofactor complex. The N- and C-termini are labeled. The cofactor AdoHcy is depicted in brown. This figure and the following figures in this article have been prepared by PyMol (14).

suggesting that they are not biologically relevant and the gel filtration experiment showed that NpmA exists as a monomer in solution (Supplementary Figure S1).

NpmA adopts a Class I methyltransferase fold consisting of a seven-stranded β sheet (β5↑–β4↑–β3↑–β6↑–β7↑–β9↓–β8↑). A search for structurally similar proteins within the PDB database was performed with DALI (28). NpmA shows the highest structural similarity to TrmB from *Bacillus subtilis*, PDB code 2FCA (29), with rmsd of 3.1 Å for 179 Cα atoms (Z-score = 16.5; 13% sequence identity) and to a number of its bacterial and eukaryotic orthologs from the TrmB/Trm8 family. TrmB is a guanosine-N^7 methyltransferase that specifically methylates the G46 nucleotide in tRNAs. A common feature of NpmA and TrmB that distinguishes them from all other Class I MTases identified so far is that the N-terminal helix (corresponding to motif X) folds onto the opposite part of the central β-sheet (Figure 1) compared to other enzymes of this fold (30). The space vacated by the N-terminal helix is instead occupied by a TrmB/NpmA-specific insertion between strands β8 and β9, which in NpmA comprises two short helices, and in TrmB one helix and a loop. Thus, NpmA and TrmB should be considered closely homologous enzymes sharing the same idiosyncrasies within the RFM superfamily. According to DALI searches, NpmA shares structural similarities with other RFM-fold proteins with predicted or determined methyltransferase activity, but they are all more distant than members of the TrmB family (Supplementary Table S2).

NpmA–AdoMet/AdoHcy interactions

The ligands in NpmA–AdoMet and NpmA–AdoHcy complexes display a similar mode of binding (Figure 2; Supplementary Figure S3). The highly conserved consensus motif G × G × G (G32 × G × G36) in the loop connecting the β3 to its successive α helix forms the bottom of the cofactor-binding site, while residues D30, N38, P56,

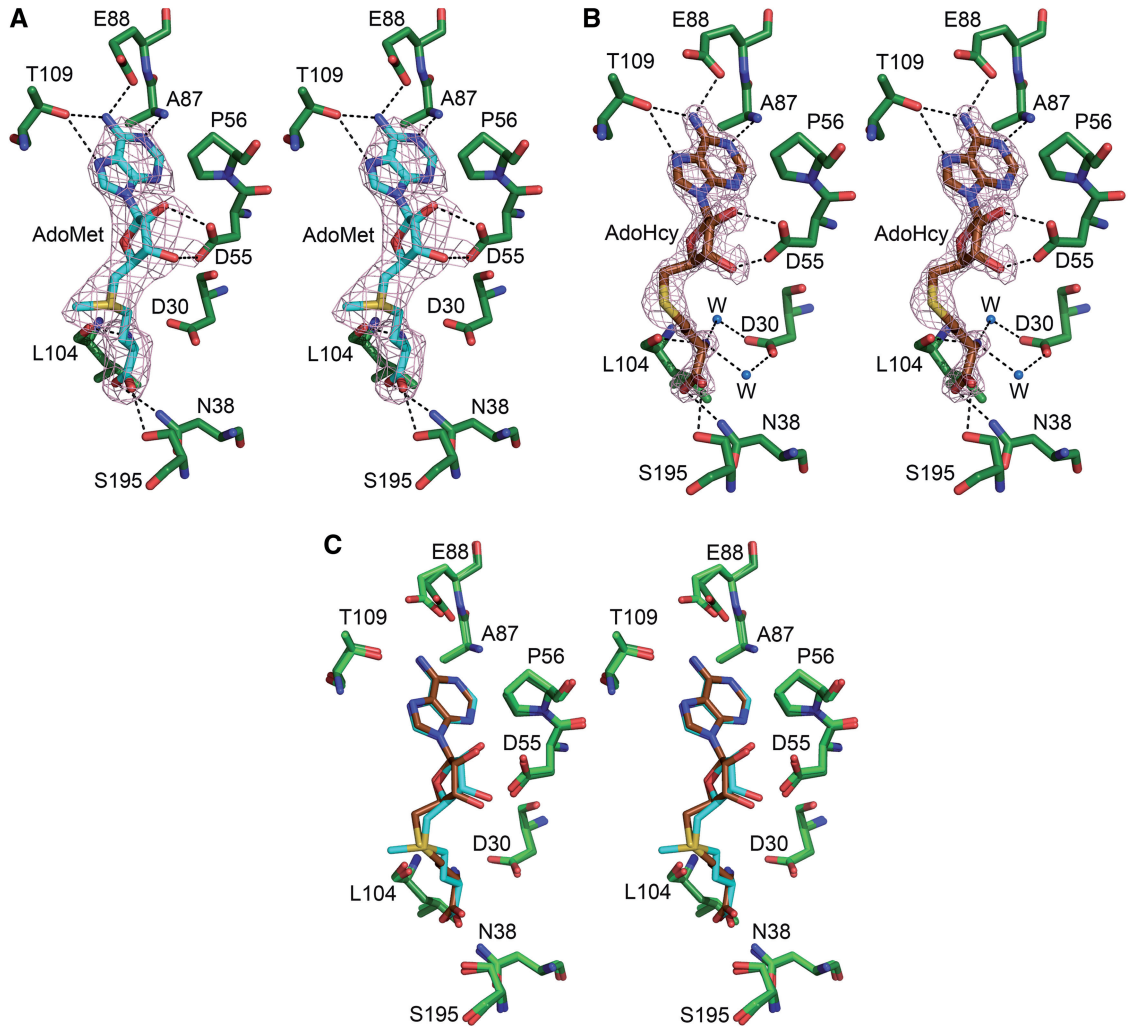

Figure 2. Stereo view of the cofactor binding site in the active site region (**A**) NpmA–AdoMet complex: the F_o-F_c simulated annealing (SA)-omit map of the cofactor AdoMet contoured at 2σ is shown. AdoMet is shown in cyan. (**B**) NpmA–AdoHcy complex: the F_o-F_c SA-omit map of the AdoHcy contoured at 3σ is shown. AdoHcy is shown in brown. Ligands and all atoms within 3Å were omitted prior to refinement. (**C**) Stereo view of the superposition of the ligand binding site of NpmA–AdoMet and NpmA–AdoHcy complexes. AdoHcy is shown in brown and AdoMet in cyan, respectively. The residue G32 that makes hydrogen bonding contact with carboxypropyl moiety of AdoMet/AdoHcy is not shown for the clarity.

T109, E88, A87, D55, G32, L104 and S195 delineate the cofactor-binding site (Figure 2). The residues D30, G32 and D55 are highly conserved in Class I AdoMet dependent MTases and are involved in cofactor binding. D30 coordinates the carboxypropyl moiety via two water molecules and D55 coordinates both hydroxyls of the ribose moiety. The residues equivalent to E88, A87 and L104 of NpmA have also been found to interact with the cofactor in other MTases (31). In particular the side chain of E88 (like its equivalents in other related enzymes) coordinates the exocyclic amino group of the AdoMet adenine moiety, while L104 and A87 are engaged in hydrogen bonds with the cofactor via their backbone. The adenine moiety is stacked between the side-chains of P56 and L110, and its N7 atom is additionally coordinated by the side chain of T109. Besides, residues N38, G32, L104 and S195 interact with the carboxypropyl moiety of AdoMet/AdoHcy. In total, NpmA makes 10 direct hydrogen bonds with AdoMet/AdoHcy. The cofactor binding site in apo-NpmA is highly similar to that of the NpmA–AdoMet/AdoHcy complex, with the exception of the D55, whose side chain exhibits a slightly different orientation than in the cofactor-bound complexes.

In order to study the predicted cofactor-binding role of residues D30, D55, E88, T109 and S195, we used site-directed mutagenesis to generate alanine substitutions in the respective positions. We analyzed the cofactor binding ability and methylation activity of the corresponding protein variants *in vitro* by a methylation assay (Figure 3) and *in vivo* by comparing them with the wt enzyme in their ability to render bacteria resistant to kanamycin (Table 2). We also identified the presence of methylated A1408 in cells expressing different NpmA variants by primer extension on 16S rRNA (Figure 4). Interestingly, the only residue we found critical for the NpmA activity is D30. The D30A variant was almost completely unable to render bacteria resistant to kanamycin, its methylation activity *in vitro* was reduced to nearly 20% activity of the native enzyme and primer extension analysis revealed that m^1A1408 was present in minimal amounts in 16S rRNA isolated from cells expressing D30A. Surprisingly, D55 that coordinates the ribose moiety of the cofactor via two hydrogen bonds was found to be important, but not crucial for the NpmA methyltransferase activity. The kanamycin MIC and enzyme activity of D55A were reduced to 50% compared to the wt NpmA, but the reverse transcriptase stop due to the m^1A modification on nucleotide A1408 in the primer extension analysis was clearly visible. Furthermore, E88A, T109A and S195A variants showed no substantial difference from the wt enzyme in their methyltransferase activity, as shown by all *in vivo* and *in vitro* activity tests.

To gain more insight into AdoMet/AdoHcy binding ability of unexpectedly active D55A, E88A, T109A, S195A and D30A protein variants versus the wt enzyme, we analyzed these proteins by ITC (Figure 5, Supplementary Figure S5 and Table 3). The dissociation constant (K_d) for AdoMet and AdoHcy of the wt NpmA was determined to be 20 μM with $N = 1.1$ and 0.6 μM with $N = 1.0$, respectively. The K_d for AdoMet (18 μM) is comparable to that of the Sgm methyltransferase, whereas the K_d for AdoHcy is 500 times higher than that of Sgm (21).

Table 2. Effects of NpmA single mutations on kanamycin resistance

NpmA variant	Target	Kanamycin MIC (mg/l)
Negative control (empty vector)	–	4–8
NpmA-wt	–	1024
D30A	AdoMet binding	8
D55A	AdoMet binding	512
E88A	AdoMet binding	1024
P106A	A1408 binding/catalysis	1024
W107A	A1408 binding/catalysis	8
T109A	AdoMet binding	1024
F177A	RNA binding	1024
S195A	AdoMet binding	1024
W197A	A1408 binding/catalysis	8
K199A	RNA binding	1024
R200A	RNA binding	1024
R205A	RNA binding	1024

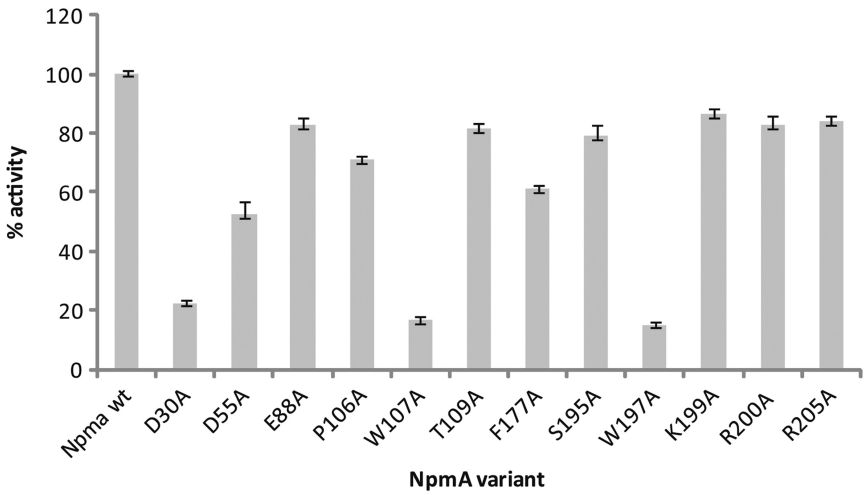

Figure 3. Relative activities of NpmA variants expressed with respect to the wild-type enzyme.

The T109A variant retained its binding ability, although weaker, with K_d values of 0.4 mM and 2 μM for AdoMet and AdoHcy, respectively. Interestingly, the K_d of S195A for AdoHcy was found to be 33 μM, but we observed a complete loss of AdoMet binding, while the NpmA variants D30A, D55A and E88A completely lost their ability to interact with both AdoMet and AdoHcy. It is surprising that S195A, E88A and D55A variants, for which we could not detect cofactor binding *in vitro*, showed remarkable activity *in vivo*. However, the ITC experiments were performed in the absence of the substrate, the 30S subunit.

NpmA-target adenine interactions

Spatially adjacent and conserved residues P106 and W107 from motif IV, F177 from motif VII and W197 from motif VIII were predicted to form the adenosine binding pocket. Notably W107 side chain adopts different conformation in NpmA–AdoMet and NpmA–AdoHcy complexes. In AdoMet complex, the indole ring of W107 moves closer to the methyl group of the cofactor to enhance the interactions, whereas the lack of methyl group in AdoHcy keeps the side chain of W107 away from AdoHcy (Supplementary Figure S2). The results of our functional tests confirmed the importance of the two Trp residues, as the *in vitro* methylation activity of W107A and W197A variants was significantly reduced (to ~15% of the wt enzyme activity). Interestingly, m^1A1408 could not be detected in 16S rRNA from the cells expressing the altered proteins and they failed to cause any kanamycin resistance in bacterial cells under the conditions tested. P106A and F177A on the other hand, were as active *in vivo* as the wt enzyme according to the amount of the methylated 16S rRNA (Figure 4) and kanamycin resistance (Table 2). We observed a decrease of activity in the *in vitro* methylation assay, where P106A and F177A showed ~70 and 60% activity, respectively, compared to the wt NpmA. This effect might be a consequence of the protein instability outside the living cell, especially in the case of F177A, where we observed a lower yield of the purified protein (data not shown). W107 and W197 are positioned similarly to aromatic residues in the active sites of many other nucleic acid MTases that are known to stabilize the target base via stacking interactions e.g. DNA:m^6A MTase M.TaqI (32) or 23S rRNA:$m^{6,6}$A2058 MTase ErmC′ (33). Our structural and functional analyses suggest that these two Trp residues function in a similar manner in NpmA, by stabilizing the target adenosine in a favorable position for the catalytic step. On the other hand, the side-chains of P106 and F177 seem to have a secondary and non-essential role in this process.

Guiding of the target base into the catalytic center of MTases that modify nucleic acids is often accompanied by interactions of positively charged residues exposed on the enzyme surface with either DNA/RNA backbone or with the nucleotide residues located next to the target (34–36). We therefore analyzed three conserved residues K199, K200 and R205 that are in the vicinity of the adenosine binding pocket. Surprisingly, all three residues turned out to be completely dispensable and no difference could be found between the variants with individual alanine substitutions and the wt NpmA. Our results differ from those obtained for KamB enzyme, where the corresponding R195 was found to be very important and R196 and R201 were indispensable for the MTase activity *in vivo* (37).

RNA footprinting of NpmA interaction with the 30S ribosomal subunit

To obtain information about interactions between NpmA and the 30S subunit, which was found to be the only

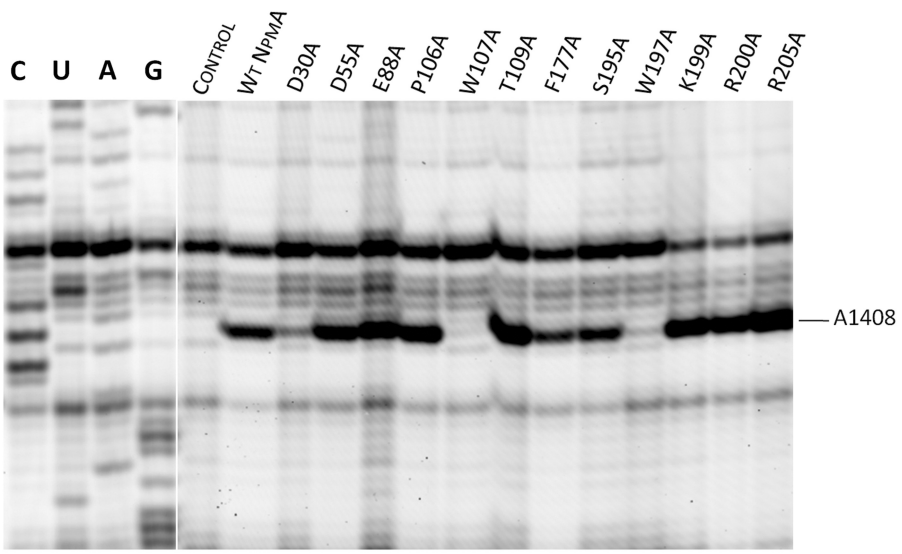

Figure 4. Gel image of a primer extension on rRNA isolated from cells expressing either wild-type or mutant forms of the NpmA enzyme. rRNA was purified from the cells actively expressing an NpmA variant and extended with a primer complementary to the 1459–1479 region in 16S rRNA. Dideoxy sequencing lanes (CUAG) were used to locate the site of modification. Control rRNA was purified from cells expressing no NpmA. The NpmA variant is indicated above the corresponding lane on the gel. The position of the reverse transcriptase stop due to the m^1A modification on nucleotide A1408 is noted on the right side of the gel image.

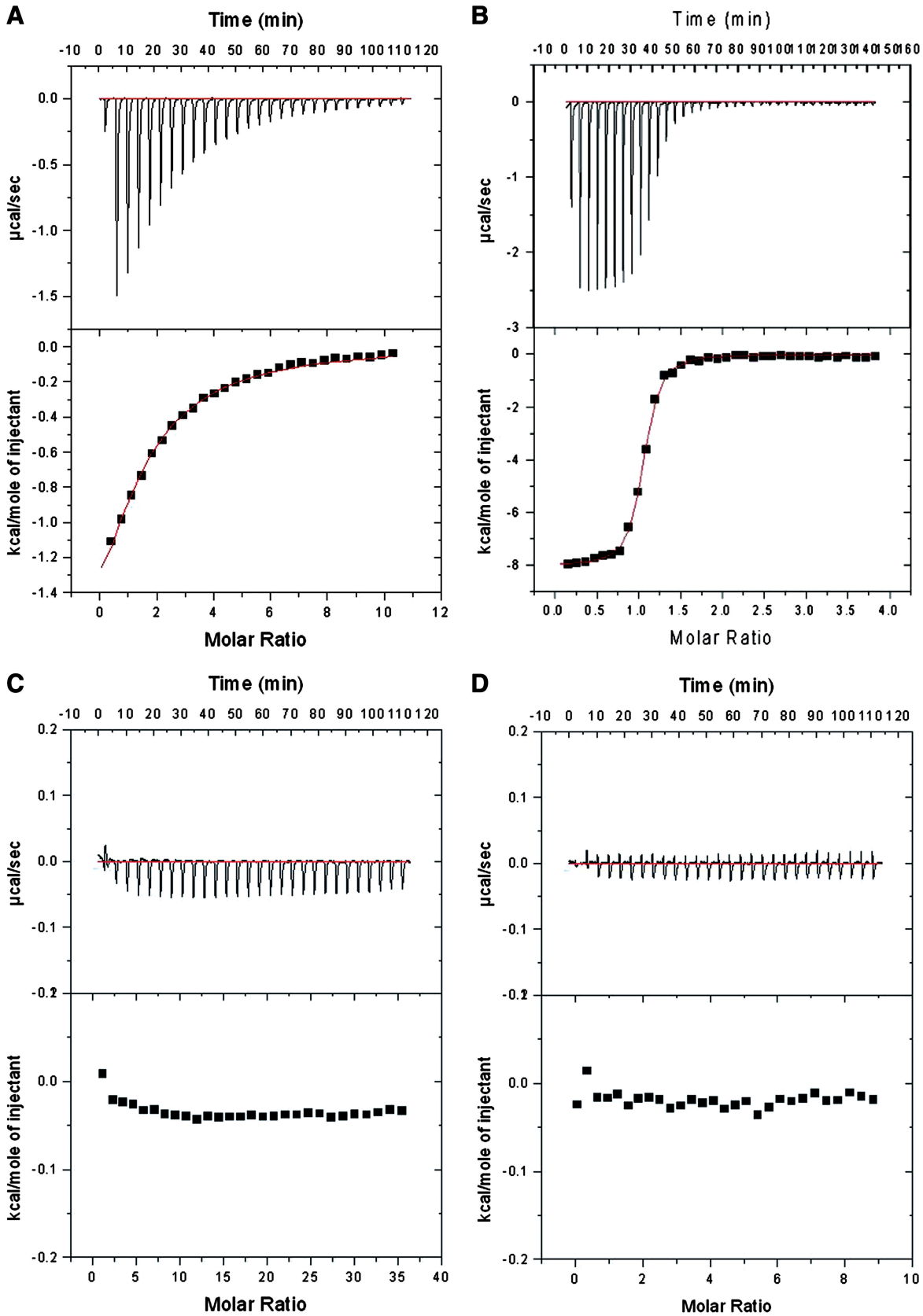

Figure 5. ITC experiments. (**A**) NpmA–AdoMet titration. (**B**) NpmA–AdoHcy titration. (**C**) D30A–AdoMet titration. (**D**) D30A–AdoHcy titration. The upper panels show the injection profile after baseline correction and the bottom panels show the integration (heat release) for each injection (except the first one). The ITC profile for other variants such as D55A, E88A that are similar to D30A (no binding) are provided in the Supplementary section. Similarly the remaining variants mentioned in the main text (e.g. T109A and S195A) are given in the Supplementary section (Supplementary Figure S5).

substrate methylatable *in vitro* (7), we performed DEPC (specific for A) and CMCT (specific for G and U) probing of 16S rRNA. The summary of the results is shown in Figures 6 and 7. Chemical modifications in 16S rRNA isolated from the MTase-treated 30S subunits were detected by primer extension. We used primers complementary to the 16S RNA nucleotides 1459–1479, 939–955 and 817–833. This enabled us to examine the regions that are proximal to the target A1408 in the fully assembled 30S subunit and thus could form the putative binding site for NpmA. Chemical probing allowed us to monitor the changes in the reactivity of nucleotides upon NpmA binding, where nucleotides became either less reactive (potentially protected from the water-soluble reagent) or more reactive (potentially more exposed to the water-soluble reagent). In addition, we also observed a different pattern of reverse transcriptase stops upon NpmA binding at nucleotides other than those expected to react with either DEPC or CMCT. This effect is often observed within regions with A and C residues and may be caused e.g. by structural changes in the RNA after its isolation from the 30S particles that were incubated with NpmA, resulting from chemical modifications. A distinct section containing the majority of nucleotides that become protected upon NpmA binding was found in helix 24 (between nucleotides 810 and 760). In addition, a few protected nucleotides were also found in helices 22 and 23 (nucleotides 754, 706, 682, 665) and in helix 25 (nucleotide 880). In helices 26 and 26a, we observed a few nucleotides that become exposed upon NpmA binding. Binding of NpmA results in protection of nucleotides in helices 42 and 43 (residues 1299–1362) and exposure of residues in helix 41. It is remarkable that the target nucleotide residue A1408 was not found to be protected by NpmA. Instead, two distinct regions upstream of A1408 (including nucleotides 1388–1390, 1392, 1393, 1395, 1397 in helix 28 and 1404, 1406 and 1407 in the top part of helix 44) showed enhanced reactivity as a result of NpmA-induced structural changes.

Model of the NpmA-30S subunit complex

To predict the three-dimensional mode of protein–substrate interactions, we docked the NpmA–AdoMet complex to the 30S subunit structure, with a restraint imposed as to minimize the distance between the methyl group of AdoMet and the target nucleoside A1408

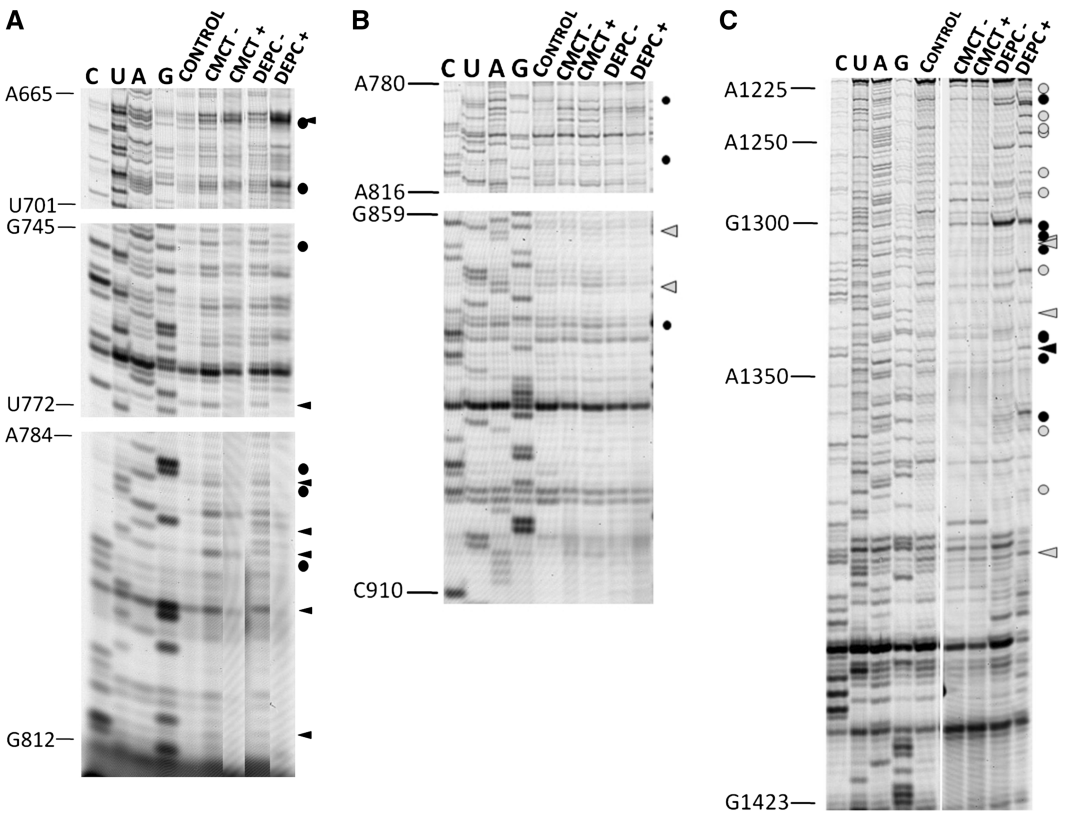

Figure 6. Fluorescent image of a primer extension gel showing chemical footprinting of NpmA methyltransferase on the 30S ribosomal subunit. The 30S subunit was incubated at 37°C with 2-fold molar excess of NpmA with or without the presence of SAH. Samples were modified with CMCT and DEPC. 16S rRNA was purified from the subunits and analysed by primer extension from primer 817–833 (**A**), 939–955 (**B**) and 1459–1479 (**C**). Each lane on the gel images is marked as follows: dideoxy sequencing lanes (CUAG) were used to locate the sites of modifications; control rRNA was purified from the chemically untreated 30S subunits; CMCT- and DEPC- show primer extension on rRNA isolated from 30S subunits treated with the chemical, but not incubated with SAH and NpmA; CMCT+ and DEPC+ show primer extension on rRNA isolated from 30S subunits treated with the chemical and incubated with SAH and NpmA. Nucleotides that were protected upon NpmA binding are marked with a black triangle (CMCT) or a black dot (DEPC). Nucleotides that were exposed upon NpmA binding are marked with a grey triangle (CMCT) or a grey dot (DEPC). Helix numbering is indicated in boxes. Intervening lanes were removed for the clarity of the presentation.

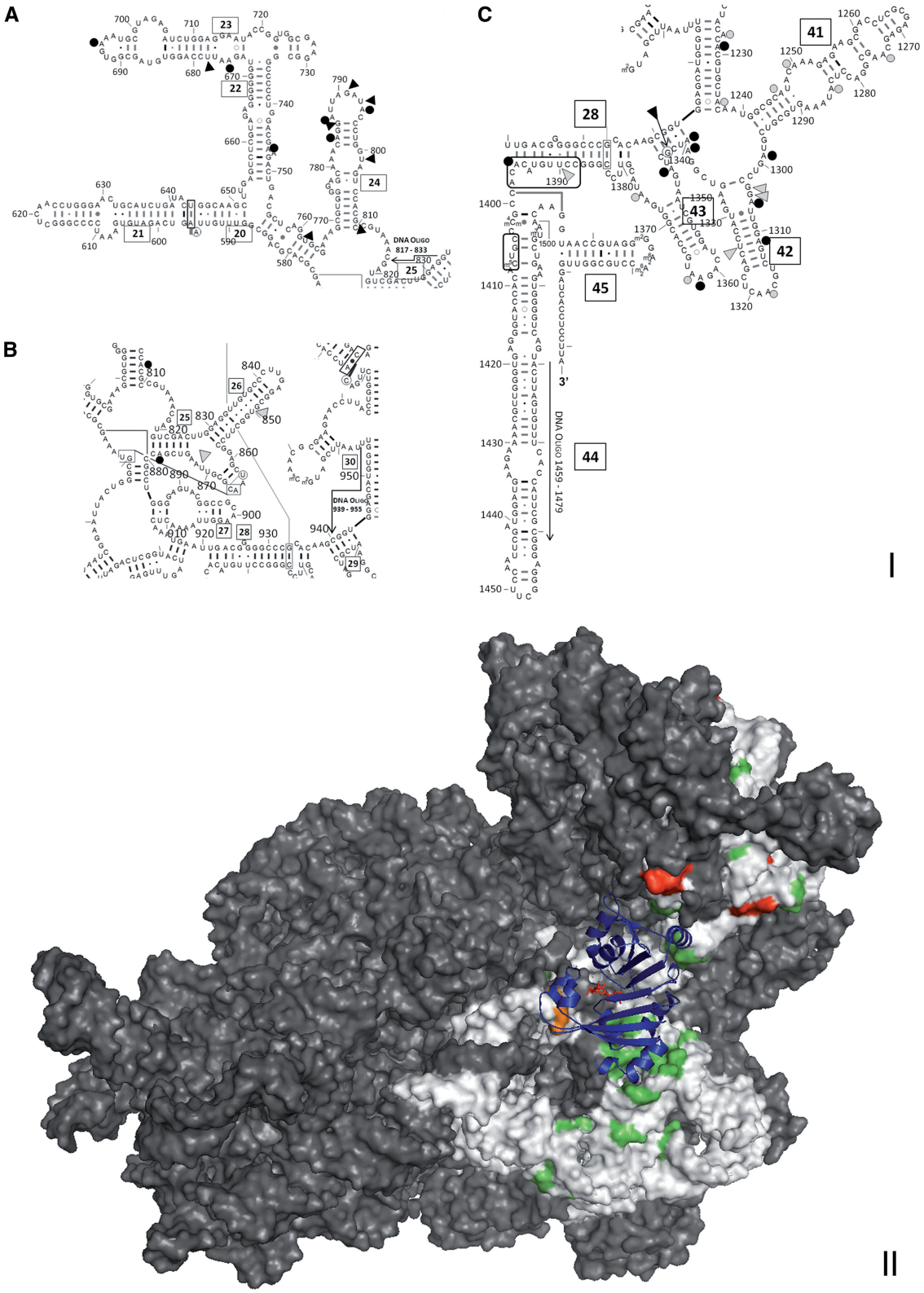

Figure 7. Panel I. Schematic representation of chemical footprinting of NpmA methyltransferase on the 30S ribosomal subunit. Samples were modified with CMCT and DEPC. RNA was purified from 30S subunits and analysed by primer extension from primer 817–833 (**A**), 939–955 (**B**) and 1459–1479 (**C**). Nucleotides found protected upon NpmA binding are marked with a black triangle (CMCT) or a black dot (DEPC). Nucleotides found exposed upon NpmA binding are marked with a grey triangle (CMCT) or a grey dot (DEPC). Regions showing enhanced non-specific reactivity with chemicals (indicating local structural changes caused by NpmA binding) are circled. Helix numbering is indicated in boxes. Panel II. Mapping of footprinting data onto the 30S structure and the model of NpmA-30S complex. Components of the 30S subunit are shown in the surface representation. NpmA is shown as a solid ribbon in blue. Ribosomal proteins and regions of 16S rRNA not analyzed by footprinting are indicated in light grey. Results of the footprinting experiments are color coded as follows: residues with unchanged reactivity are shown in white, residues relatively more reactive in the footprinting experiment are shown in red, residues less reactive in the footprinting experiment are shown in green. The target A1408 is shown in orange.

(see 'Materials and Methods' section for details). In the representative model obtained after clustering of best-fitting poses (Figure 7), NpmA and the 30S subunit fit together relatively tightly, suggesting that no large-scale conformational changes are required for the binding to occur. NpmA is predicted to interact directly only with 16S rRNA, and not with ribosomal proteins. In our docking model most of the contacts with the RNA are mediated by the positively charged N-terminal region (with the 3′ major 16S rRNA domain) and by the catalytic face of NpmA (with the central and 3′ minor 16S rRNA domains). No direct contacts are formed with the 5′ domain. The region comprising residues 144–157 of NpmA, for which coordinates could not be built in the crystal structure, is located close to the target base, but its anchor points remain exposed on the protein surface in the complex. We predict that this region is unstructured in the unbound form of NpmA, but may become structured upon formation of the NpmA-30S subunit complex. While the methylatable N1 atom of A1408 does not make a direct contact with the enzyme, the target base can be positioned within the active site of NpmA (between W107 and W197 residues) after being simply rotated out of the helix. It is worthwhile to mention that in the 30S subunit structure, A1408 is stacked within helix 44, but not base-paired, hence its flipping may be relatively easy. In summary, the docking model supports the interpretation of the experimental data that only limited conformational changes in the 30S subunit, including A1408 base flipping, are required for NpmA to methylate its target in 16S rRNA.

DISCUSSION

The nucleotide residue A1408 in 16S rRNA of the 30S ribosomal subunit is a crucial 'hot spot' of aminoglycoside resistance. The universally conserved A1408 has a critical role in binding of aminoglycoside antibiotics to the A site of 16S rRNA. Structural studies have shown that the hydroxyl group on sugar ring I of 4,6- and 4,5-disubstituted 2-deoxystreptamines forms hydrogen bonds with the N1 and N6 atoms of A1408 (38–40). The N1 methylation of A1408 by NpmA apparently disrupts the interaction of the 16S rRNA with the drug, as indicated in footprinting experiments (7) thereby conferring resistance. The recent detection of the Kam family member NpmA in a clinical isolate (7) indicates an imminent threat to the utility of aminoglycosides as antibacterial drugs and warrants attention. The structure of NpmA and its complex with the cofactors and substrate holds key to understanding its mode of action and potentially to develop specific inhibitors of the A1408 methylation activity that could reinstate the activity of aminoglycoside antibiotics.

Based on structural analyses we identified residues in NpmA potentially important for cofactor-binding and catalysis and carried out site-directed mutagenesis to obtain protein variants with alanine substitutions in the corresponding positions. As predicted, the ITC experiments revealed that NpmA variants with cofactor-binding residues substituted by Ala have completely lost or exhibited significantly reduced ability to bind AdoMet or AdoHcy (Figure 5, Supplementary Figure S5 and Table 3). We were, however, greatly surprised by the fact that the variant D55A retained the ability to carry out m^1A1408 methylation *in vitro*, as well as to render the host-cell resistant to aminoglycosides, because substitutions of this conserved residue in other RNA methyltransferases, e.g. D55 in KamB (37) or D156 in Sgm (41) has lead to drastic reduction of these enzymes' activity both *in vivo* and *in vitro*. We speculate that while the D55A variant of NpmA cannot form a binary complex with the cofactor, it may bind AdoMet and methylate the target when it is already engaged in interactions with the substrate. For some related methyltransferases, e.g. Dam enzyme from phage T4 phage, it has been demonstrated that the binding of the methyl group donor or the nucleic acid substrate reciprocally improve the affinity for each other in the ternary complex (42). This suggests that NpmA may also exhibit significantly increased affinity for AdoMet when pre-bound to the 30S subunit. Other enzymes, however, e.g. Dam from *E. coli* (which is very closely related to T4 Dam), exhibit similar affinity for AdoMet in the presence or in the absence of the substrate (43). Apparently, large variations of parameters concerning the binding of the ligands and their mutual interactions are not uncommon among related methyltransferases. We speculate that in NpmA, binding of the 30S substrate increases the affinity for AdoMet to such extent that it can rescue the ligand binding ability of the D55A variant, while the similar effect has not been observed for KamB. Our results show that the relative importance of 'the same' residues in closely related proteins may vary. This study highlights the importance of comparative biochemical and functional analyses of proteins that belong to the same family and suggests that structural and functional characterization of just single members of each protein family (as in the structural genomics approach) may be insufficient to understand sequence-structure–function relationships in proteins. The finding that the side chain of D55 residue is dispensable for the *in vivo* activity of NpmA, and hence that mutations altering that residue may appear in the nature, has to be taken into consideration in analyses aiming at structure-based design of inhibitors.

NpmA binding to the 30S subunit

NpmA was found to successfully methylate only fully assembled 30S subunits, while the entire 70S ribosome

Table 3. ITC results

NpmA variant	K_d for AdoMet	K_d for AdoHcy
Wild-type	20 μM	0.6 μM
D30A	No binding	No binding
D55A	No binding	No binding
T109A	0.4 mM	2 μM
S195A	No binding	33 μM
E88A	No binding	No binding

or naked 16S rRNA could not be methylated (7). This reactivity towards 30S subunits is analogous to aminoglycoside resistance MTases from the Arm family, ArmA and Sgm, which methylate G1405 (44,45). Here we present the analysis of interactions between a member of the Kam family and small ribosomal subunit as its substrate. Using DEPC and CMCT chemical probing and subsequent primer extension analysis we were able to identify helices 24 and 42 with associated loop as primary regions that interact with NpmA. In addition, two short sections of 16S rRNA 5′ from the target nucleotide residue A1408 showed increased non-specific reactivity to chemicals when isolated from 30S subunits incubated with the NpmA. This might be the result of a local structural change in the rRNA caused by NpmA binding, suggesting that NpmA-30S interaction could be associated with the top of helices 44 and 28 as well.

We have recently reported a similar footprinting and docking analysis for Sgm MTase from the Arm family that modifies G1405 (21). While G1405 and A1408 are positioned very closely in rRNA, suggesting a similar mode of binding for Sgm and NpmA, we could not clearly define the region(s) of 16S rRNA that form distinctive interactions with Sgm. Instead, protected nucleotides were scattered over a wide area, together with a large number of nucleotides with the activity enhanced upon Sgm binding, indicating complex structural changes of a rather global character that enable the otherwise buried G1405 base to become available for methylation by Sgm. Unlike Sgm, NpmA binding to the 30S subunit does not lead to global structural rearrangements; instead we observe only local changes. Remarkably, the target base A1408 was not found to be protected upon NpmA binding, which suggests that it may be flipped-out and more solvent accessible in the complex with the enzyme than in the 30S subunit before modification. Its binding may be stabilized in the NpmA active site by residues W107 and W197, found here to be essential for the methylation to occur *in vivo*. The conformation of W107 side chain is different in complexes with AdoMet and AdoHcy, which suggests that limited conformational changes upon binding may occur also on the protein side. In the crystal structures of NpmA, the side-chains of these Trp residues are too close to each other to enable binding of the target base, but a simple change to alternative rotamers allows the protein to form a binding pocket to accommodate the adenine moiety (data not shown). The necessity of minor conformational adjustments has to be taken into consideration in the future analyses aiming at structure-based design of inhibitors.

The results of footprinting experiments with Sgm and NpmA are in perfect agreement with the docking studies. In our previous work (21) we found that the Sgm structure cannot be docked to its target G1405 without very severe steric clashes with the 30S subunit, which implies that significant conformational changes upon its binding must occur. On the other hand, NpmA can be easily docked to its target A1408, to form a relatively close-fitting complex with the 30S subunit. Therefore we postulate that Sgm and NpmA use very different mechanisms to access their target bases.

Regions of 16S rRNA that we found to be involved in interactions with NpmA come in close proximity and form functional sites in fully assembled 30S subunits, including the decoding site and site of association with 50S subunit, which clearly explains why NpmA cannot act on 70S ribosomes. In addition, regions identified as NpmA binding sites partly overlap with the KsgA-30S interaction sites identified by directed and solution hydroxyl radical footprinting that include helices 2, 11, 24, 27, 28, 44 and 45 and loop 790 (46). KsgA is a housekeeping dimethyltransferase that modifies A1518 and A1519 (both to $m^{6,6}A$) in helix 45 of 16S rRNA. It acts as a regulator of ribosome biogenesis and binds to immature 30S subunits in a translationally inactive conformation (47). NpmA is known to act on mature 30S subunits, but it has not been determined whether it can methylate any of the reconstitution intermediates (RIs) in the 30S subunit assembly. The first intermediate, RI, lacks tertiary ribosomal proteins S2, S3, S10, S14 and S21 in central and 3′-domain and is not competent to form a small ribosomal subunit. rRNA in RI undergoes an extensive conformational rearrangement to form RI*, which can then bind the remaining five ribosomal proteins and form 30S subunit (48). Holmes and Culver have performed an extensive analysis of conformational changes in 16S rRNA during the assembly of small ribosomal subunit (49) and mapped structural differences between the assembly intermediates RI and RI* (50). The majority of conformational changes related to the central and 3′-domain of the 30S subunit occur during transition from RI to RI*. These changes include functional sites on 30S, where NpmA binds and recognizes its target, hence we may propose that the earliest stage that supports the NpmA action could be a nearly assembled RI* of the 30S subunit.

Relationship of NpmA to other MTases; insight from structural analyses

The availability of a high-resolution NpmA structure allowed us to establish its relationship to other MTase families. NpmA is a member of the RFM superfamily and exhibits a class I MTase fold with the exceptional topological rearrangement involving the N-terminal helix. The arrangement of the N-terminal helix in NpmA is similar to those found in the structure of tRNA:m^7G56 MTase TrmB, a representative of a protein family that appears to be the closest homolog of Kam-family MTases (Supplementary Figure S6). It is very intriguing that an RNA:m^1A MTase (NpmA) appears to be more closely related to an RNA:m^7G MTase (TrmB) than to other m^1A MTases, for which structures have been determined, including tRNA:m^1A58 MTase TrmI (51) and tRNA:m^1A22 MTase TrmK (52,53). NpmA is also structurally distinct from Arm/RmtB/Sgm enzymes that form m^7G1405 in 16S rRNA (Supplementary Figure S6). G1405 is very close to A1408, the site of methylation of NpmA. While NpmA/Kam and Arm/RmtB/Sgm methyltransferases belong to the same RFM superfamily and methylate bases that are very close to each other in the 30S subunit leading to a similar

phenotypic effect, i.e. resistance to aminoglycoside antibiotics, their catalytic domains are not very closely related and presumed substrate-binding domains are completely different.

Based on structural superposition we generated a multiple sequence alignment involving all members of the Kam family (including NpmA studied in this work), as well as individual representatives of m^1A MTase families TrmI and TrmK, and m^7G MTases TrmB and Sgm (Supplementary Figure S4). Motifs I and II involved in the binding of the common cofactor AdoMet exhibit residues that are universally conserved between these enzymes (in particular carboxylates involved in coordination of the methionine and ribose moieties). On the other hand, motifs IV, VI, and VIII that commonly harbor residues involved in binding of the target base and catalysis are completely different. In the active sites, traces of similarity are observed only between NpmA and TrmB enzymes: Y193 is homologous and most likely functionally equivalent to the base-binding residue W197 of NpmA. Both proteins also share a common Phe residue (F116 in TrmB and F105 in NpmA), which stacks face to edge with the above-mentioned base-binding aromatic residue and thereby stabilizes it. Other than that, however, the active sites of TrmB and NpmA are different, consistent with the activity on a different substrate: the N7 atom of guanosine versus the N1 atom of adenosine.

The absence of common residues in motifs responsible for catalysis suggest that m^1A MTases acting on different RNA targets have non-homologous active sites and most likely developed independently from different ancestors within the RFM superfamily. In particular, our analysis suggests that the Kam family evolved from the TrmB/Trm8 family of m^7G MTases. This is in agreement with the results of our phylogenetic analysis of the entire RFM superfamily, in which the Kam, TrmI and TrmK families of m^1A MTases do not form a single branch on the tree and most likely have evolved independently (J.M.B. unpublished data). This situation resembles the polyphyletic evolution of m^7G MTases, whose five families (Arm/Sgm, TrmB/Trm8, RsmG, Abd1/Ecm1 and Bud23) are predicted to be independent evolutionary inventions with the same fold but different active sites (21). Likewise, 2′-*O*-ribose MTases include several families of enzymes with different specificities (e.g. RrmJ-like, HEN1 and Trm13) that exhibit different active sites anchored to the common RFM fold (54) or even to a different fold (55). Recently, a family of RNA MTases of the SPOUT fold have been characterized, in which some members exhibit m^1G specificity (56), while others form m^1A but not m^1G or exhibit dual specificity and generate m^1A or m^1G (57). This variety of protein families and protein folds with different active sites able to carry out RNA:m^1A methylation suggests that this type of reaction is probably relatively easy for an enzyme to evolve (e.g. from an MTase of a different specificity). On the other hand, the uniqueness of the adenosine-binding site of MTases from the Kam family suggests that they may be good targets for the development of Kam-specific inhibitors. The structural analyses presented in this work provide a stepping stone for structure-based virtual screening to find candidates for such compounds.

ACCESSION NUMBER

Coordinates of apo-NpmA, NpmA–AdoMet and NpmA–AdoHcy were deposited in the Protein Data Bank (http://www.pdb.org) under accession codes 3P2I, 3P2K and 3P2E, respectively. The PDB code for the NpmA–AdoHcy complex SAD data set is 3PB3.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

Data for this study were measured at beam lines X12C and X29 of the National Synchrotron Light Sources, BNL and beamline BL13B1, NSRRC (Taiwan). We thank Dr Suparna Sanyal from Uppsala University, Sweden for the gift of *E. coli* JE28 cells. We are grateful to the Ruder Bošković Institute for the use of the Typhoon imager. Ana Jurić and Eva Kukolj (University of Zagreb) are thanked for excellent technical assistance.

FUNDING

Biomedical Research Council (BMRC) of Singapore (to J.S. and his team); A*STAR (grant R154000362305); Croatian Ministry of Science (grant 006-0982913-1219 to G.M.V. and her team); ICGEB (grant CRP/CRO08-02); EU FP6 (grant #043682 'EuroPharm'); Polish Ministry of Science and Higher Education (grants N301 2396 33 and 188/N-DFG/2008/0 to J.M.B. and his team); EU FP6 (grant EURASNET, LSHG-CT-2005-518238); N.H. is a PhD student in receipt of a research scholarship from the National University of Singapore (NUS). Funding for open access charge: Singapore, BMRC; A*STAR (grant R154000362305).

Conflict of interest statement. None declared.

NOTE ADDED IN PROOF

During submission of this article the crystal structures of KamB-AdoHcy and NpmA-AdoMet were independently published by another group (Macmaster *et al.*, *Nucleic Acids Res.* published electronically 2010 Jul 17, 10.1093/nar/gkq627). The other group's structures were on hold in PDB, which prevented us from making detailed comparisons.

REFERENCES

1. Pfister,P., Hobbie,S., Vicens,Q., Bottger,E.C. and Westhof,E. (2003) The molecular basis for A-site mutations conferring aminoglycoside resistance: relationship between ribosomal susceptibility and X-ray crystal structures. *Chembiochem*, **4**, 1078–1088.

2. Kotra,L.P., Haddad,J. and Mobashery,S. (2000) Aminoglycosides: perspectives on mechanisms of action and resistance and strategies to counter resistance. *Antimicrob. Agents Chemother.*, **44**, 3249–3256.
3. Prammananan,T., Sander,P., Brown,B.A., Frischkorn,K., Onyi,G.O., Zhang,Y., Bottger,E.C. and Wallace,R.J. Jr (1998) A single 16S ribosomal RNA substitution is responsible for resistance to amikacin and other 2-deoxystreptamine aminoglycosides in *Mycobacterium abscessus* and *Mycobacterium chelonae*. *J. Infect. Dis.*, **177**, 1573–1581.
4. Douthwaite,S., Fourmy,D. and Yoshizawa,S. (2005) Nucleotide methylations in rRNA that confer resistance to ribosome-targeting antibiotics. *Topics Curr. Genet.*, **12**, 285–307.
5. Doi,Y. and Arakawa,Y. (2007) 16S ribosomal RNA methylation: emerging resistance mechanism against aminoglycosides. *Clin. Infect. Dis.*, **45**, 88–94.
6. Koscinski,L., Feder,M. and Bujnicki,J.M. (2007) Identification of a missing sequence and functionally important residues of 16S rRNA:m(1)A1408 methyltransferase KamB that causes bacterial resistance to aminoglycoside antibiotics. *Cell Cycle*, **6**, 1268–1271.
7. Wachino,J., Shibayama,K., Kurokawa,H., Kimura,K., Yamane,K., Suzuki,S., Shibata,N., Ike,Y. and Arakawa,Y. (2007) Novel plasmid-mediated 16S rRNA m^1A1408 methyltransferase, NpmA, found in a clinically isolated *Escherichia coli* strain resistant to structurally diverse aminoglycosides. *Antimicrob. Agents Chemother.*, **51**, 4401–4409.
8. Ochman,H., Gerber,A.S. and Hartl,D.L. (1988) Genetic applications of an inverse polymerase chain reaction. *Genetics*, **120**, 621–623.
9. Higuchi,R., Krummel,B. and Saiki,R.K. (1988) A general method of in vitro preparation and specific mutagenesis of DNA fragments: study of protein and DNA interactions. *Nucleic Acids Res.*, **16**, 7351–7367.
10. Doublie,S. (1997) Preparation of selenomethionyl proteins for phase determination. *Methods Enzymol.*, **276**, 523–530.
11. Otwinowski,Z.M.W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.*, **276**, 307–326.
12. Terwilliger,T.C. and Berendzen,J. (1999) Automated MAD and MIR structure solution. *Acta Crystallogr. D. Biol. Crystallogr.*, **55**, 849–861.
13. Hattne,J. and Lamzin,V.S. (2008) Pattern-recognition-based detection of planar objects in three-dimensional electron-density maps. *Acta Crystallogr. D. Biol. Crystallogr.*, **D64**, 834–842.
14. Emsley,P. and Cowtan,K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr. D. Biol. Crystallogr.*, **60**, 2126–2132.
15. Brunger,A.T., Adams,P.D., Clore,G.M., DeLano,W.L., Gros,P., Grosse-Kunstleve,R.W., Jiang,J.S., Kuszewski,J., Nilges,M., Pannu,N.S. *et al.* (1998) Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr. D. Biol. Crystallogr.*, **54**, 905–921.
16. Murshudov,G.N., Vagin,A.A. and Dodson,E.J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Cryst.*, **D53**, 240–255.
17. Vagin,A. and Teplyakov,A. (2010) Molecular replacement with MOLREP. *Acta Crystallogr. D. Biol. Crystallogr.*, **66**, 22–25.
18. Laskowski,R.A., Moss,D.S. and Thornton,J.M. (1993) Main-chain bond lengths and bond angles in protein structures. *J. Mol. Biol.*, **231**, 1049–1067.
19. Ederth,J., Mandava,C.S., Dasgupta,S. and Sanyal,S. (2009) A single-step method for purification of active His-tagged ribosomes from a genetically engineered *Escherichia coli*. *Nucleic Acids Res.*, **37**, e15.
20. Stern,S., Moazed,D. and Noller,H.F. (1988) Structural analysis of RNA using chemical and enzymatic probing monitored by primer extension. *Methods Enzymol.*, **164**, 481–489.
21. Husain,N., Tkaczuk,K.L., Tulsidas,S.R., Kaminska,K.H., Cubrilo,S., Maravic-Vlahovicek,G., Bujnicki,J.M. and Sivaraman,J. (2010) Structural basis for the methylation of G1405 in 16S rRNA by aminoglycoside resistance methyltransferase Sgm from an antibiotic producer: a diversity of active sites in m7G methyltransferases. *Nucleic Acids Res.*, **38**, 4120–4132.
22. Altschul,S.F., Madden,T.L., Schaffer,A.A., Zhang,J., Zhang,Z., Miller,W. and Lipman,D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.*, **25**, 3389–3402.
23. Pei,J. and Grishin,N.V. (2006) MUMMALS: multiple sequence alignment improved by using hidden Markov models with local structural information. *Nucleic Acids Res.*, **34**, 4364–4374.
24. Kurowski,M.A. and Bujnicki,J.M. (2003) GeneSilico protein structure prediction meta-server. *Nucleic Acids Res.*, **31**, 3305–3307.
25. Schwede,T., Kopp,J., Guex,N. and Peitsch,M.C. (2003) SWISS-MODEL: an automated protein homology-modeling server. *Nucleic Acids Res.*, **31**, 3381–3385.
26. Vakser,I.A., Matar,O.G. and Lam,C.F. (1999) A systematic study of low-resolution recognition in protein–protein complexes. *Proc. Natl Acad. Sci. USA*, **96**, 8477–8482.
27. Gajda,M.J., Tuszynska,I., Kaczor,M., Bakulina,A.Y. and Bujnicki,J.M. (2010) FILTREST3D: discrimination of structural models using restraints from experimental data. *Bioinformatics*, in press.
28. Holm,L., Kaariainen,S., Rosenstrom,P. and Schenkel,A. (2008) Searching protein structure databases with DaliLite v.3. *Bioinformatics*, **24**, 2780–2781.
29. Zegers,I., Gigot,D., van Vliet,F., Tricot,C., Aymerich,S., Bujnicki,J.M., Kosinski,J. and Droogmans,L. (2006) Crystal structure of *Bacillus subtilis* TrmB, the tRNA (m7G46) methyltransferase. *Nucleic Acids Res.*, **34**, 1925–1934.
30. Bujnicki,J.M. (1999) Comparison of protein structures reveals monophyletic origin of the AdoMet-dependent methyltransferase family and mechanistic convergence rather than recent differentiation of N4-cytosine and N6-adenine DNA methylation. *In Silico Biol.*, **1**, 175–182.
31. Kozbial,P.Z. and Mushegian,A.R. (2005) Natural history of S-adenosylmethionine-binding proteins. *BMC Struct. Biol.*, **5**, 19.
32. Goedecke,K., Pignot,M., Goody,R.S., Scheidig,A.J. and Weinhold,E. (2001) Structure of the N6-adenine DNA methyltransferase M.TaqI in complex with DNA and a cofactor analog. *Nat. Struct. Biol.*, **8**, 121–125.
33. Maravic,G., Feder,M., Pongor,S., Flogel,M. and Bujnicki,J.M. (2003) Mutational analysis defines the roles of conserved amino acid residues in the predicted catalytic pocket of the rRNA:m6A methyltransferase ErmC'. *J. Mol. Biol.*, **332**, 99–109.
34. Sankpal,U.T. and Rao,D.N. (2002) Structure, function, and mechanism of HhaI DNA methyltransferases. *Crit. Rev. Biochem. Mol. Biol.*, **37**, 167–197.
35. Alian,A., Lee,T.T., Griner,S.L., Stroud,R.M. and Finer-Moore,J. (2008) Structure of a TrmA-RNA complex: a consensus RNA fold contributes to substrate selectivity and catalysis in m5U methyltransferases. *Proc. Natl Acad. Sci. USA*, **105**, 6876–6881.
36. Lee,T.T., Agarwalla,S. and Stroud,R.M. (2005) A unique RNA Fold in the RumA-RNA-cofactor ternary complex contributes to substrate selectivity and enzymatic function. *Cell*, **120**, 599–611.
37. Macmaster,R., Zelinskaya,N., Savic,M., Rankin,C.R. and Conn,G.L. (2010) Structural insights into the function of aminoglycoside-resistance A1408 16S rRNA methyltransferases from antibiotic-producing and human pathogenic bacteria. *Nucleic Acids Res.*, doi:10.1093/nar/gkq627.
38. Carter,A.P., Clemons,W.M., Brodersen,D.E., Morgan-Warren,R.J., Wimberly,B.T. and Ramakrishnan,V. (2000) Functional insights from the structure of the 30S ribosomal subunit and its interactions with antibiotics. *Nature*, **407**, 340–348.
39. Vicens,Q. and Westhof,E. (2002) Crystal structure of a complex between the aminoglycoside tobramycin and an oligonucleotide containing the ribosomal decoding a site. *Chem. Biol.*, **9**, 747–755.
40. Francois,B., Russell,R.J., Murray,J.B., Aboul-ela,F., Masquida,B., Vicens,Q. and Westhof,E. (2005) Crystal structures of complexes between aminoglycosides and decoding A site oligonucleotides: role of the number of rings and positive charges in the specific binding leading to miscoding. *Nucleic Acids Res.*, **33**, 5677–5690.
41. Maravic Vlahovicek,G., Cubrilo,S., Tkaczuk,K.L. and Bujnicki,J.M. (2008) Modeling and experimental analyses reveal a two-domain structure and amino acids important for the activity of aminoglycoside resistance methyltransferase Sgm. *Biochim. Biophys. Acta*, **1784**, 582–590.

42. Evdokimov,A.A., Sclavi,B., Zinoviev,V.V., Malygin,E.G., Hattman,S. and Buckle,M. (2007) Study of bacteriophage T4-encoded Dam DNA (adenine-N6)-methyltransferase binding with substrates by rapid laser UV cross-linking. *J. Biol. Chem.*, **282**, 26067–26076.
43. Mashhoon,N., Carroll,M., Pruss,C., Eberhard,J., Ishikawa,S., Estabrook,R.A. and Reich,N. (2004) Functional characterization of *Escherichia coli* DNA adenine methyltransferase, a novel target for antibiotics. *J. Biol. Chem.*, **279**, 52075–52081.
44. Liou,G.F., Yoshizawa,S., Courvalin,P. and Galimand,M. (2006) Aminoglycoside resistance by ArmA-mediated ribosomal 16S methylation in human bacterial pathogens. *J. Mol. Biol.*, **359**, 358–364.
45. Cubrilo,S., Babic,F., Douthwaite,S. and Maravic Vlahovicek,G. (2009) The aminoglycoside resistance methyltransferase Sgm impedes RsmF methylation at an adjacent rRNA nucleotide in the ribosomal A site. *RNA*, **15**, 1492–1497.
46. Xu,Z., O'Farrell,H.C., Rife,J.P. and Culver,G.M. (2008) A conserved rRNA methyltransferase regulates ribosome biogenesis. *Nat. Struct. Mol. Biol.*, **15**, 534–536.
47. Desai,P.M. and Rife,J.P. (2006) The adenosine dimethyltransferase KsgA recognizes a specific conformational state of the 30S ribosomal subunit. *Arch. Biochem. Biophys.*, **449**, 57–63.
48. Sykes,M.T. and Williamson,J.R. (2009) A complex assembly landscape for the 30S ribosomal subunit. *Annu. Rev. Biophys.*, **38**, 197–215.
49. Holmes,K.L. and Culver,G.M. (2005) Analysis of conformational changes in 16 S rRNA during the course of 30 S subunit assembly. *J. Mol. Biol.*, **354**, 340–357.
50. Holmes,K.L. and Culver,G.M. (2004) Mapping structural differences between 30S ribosomal subunit assembly intermediates. *Nat. Struct. Mol. Biol.*, **11**, 179–186.
51. Roovers,M., Wouters,J., Bujnicki,J.M., Tricot,C., Stalon,V., Grosjean,H. and Droogmans,L. (2004) A primordial RNA modification enzyme: the case of tRNA (m^1A) methyltransferase. *Nucleic Acids Res.*, **32**, 465–476.
52. Roovers,M., Kaminska,K.H., Tkaczuk,K.L., Gigot,D., Droogmans,L. and Bujnicki,J.M. (2008) The YqfN protein of *Bacillus subtilis* is the tRNA: m^1A22 methyltransferase (TrmK). *Nucleic Acids Res.*, **36**, 3252–3262.
53. Ta,H.M. and Kim,K.K. (2010) Crystal structure of *Streptococcus pneumoniae* Sp1610, a putative tRNA methyltransferase, in complex with S-adenosyl-L-methionine. *Protein Sci.*, **19**, 617–624.
54. Feder,M., Pas,J., Wyrwicz,L.S. and Bujnicki,J.M. (2003) Molecular phylogenetics of the RrmJ/fibrillarin superfamily of ribose 2'-O-methyltransferases. *Gene*, **302**, 129–138.
55. Tkaczuk,K.L., Dunin-Horkawicz,S., Purta,E. and Bujnicki,J.M. (2007) Structural and evolutionary bioinformatics of the SPOUT superfamily of methyltransferases. *BMC Bioinformatics*, **8**, 73.
56. Jackman,J.E., Montange,R.K., Malik,H.S. and Phizicky,E.M. (2003) Identification of the yeast gene encoding the tRNA m^1G methyltransferase responsible for modification at position 9. *RNA*, **9**, 574–585.
57. Kempenaers,M., Roovers,M., Oudjama,Y., Tkaczuk,K.L., Bujnicki,J.M. and Droogmans,L. (2010) New archaeal methyltransferases forming 1-methyladenosine or 1-methyladenosine and 1-methylguanosine at position 9 of tRNA. *Nucleic Acids Res.*, doi:10.1093/nar/gkq451.