

Mutational Analysis Defines the Roles of Conserved Amino Acid Residues in the Predicted Catalytic Pocket of the rRNA:m⁶A Methyltransferase ErmC'

Gordana Maravić^{1,2*}, Marcin Feder³, Sándor Pongor¹, Mirna Flögel² and Janusz M. Bujnicki^{3*}

¹Protein Structure and Bioinformatics Group
International Centre for Genetic Engineering and Biotechnology (ICGEB)
Padriciano 99, 34012 Trieste
Italy

²Department of Biochemistry and Molecular Biology, Faculty of Pharmacy and Biochemistry
University of Zagreb, Ante Kovačića 1, 10000 Zagreb
Croatia

³Bioinformatics Laboratory
International Institute of Molecular and Cell Biology
Trojdena 4, 02-109 Warsaw
Poland

Methyltransferases (MTases) from the Erm family catalyze S-adenosyl-L-methionine-dependent modification of a specific adenine residue in bacterial 23 S rRNA, thereby conferring resistance to clinically important macrolide, lincosamide and streptogramin B antibiotics. Despite the available structural data and functional analyses on the level of the RNA substrate, still very little is known about the mechanism of rRNA:adenine-N⁶ methylation. Only predictions regarding various aspects of this reaction have been made based on the analysis of the crystal structures of methyltransferase ErmC' (without the RNA) and their comparison with the crystallographic and biochemical data for better studied DNA:m⁶A MTases. To validate the structure-based predictions of presumably essential residues in the catalytic pocket of ErmC', we carried out the site-directed mutagenesis and studied the function of the mutants *in vitro* and *in vivo*. Our results indicate that the active site of rRNA:m⁶A MTases is much more tolerant to amino acid substitutions than the active site of DNA:m⁶A MTases. Only the Y104 residue implicated in stabilization of the target base was found to be indispensable. Remarkably, the N101 residue from the "catalytic" motif IV and two conserved residues that form the floor (F163) and one of the walls (N11) of the base-binding site are not essential for catalysis in ErmC'. This somewhat surprising result is discussed in the light of the available structural data and in the phylogenetic context of the Erm family.

© 2003 Elsevier Ltd. All rights reserved.

Keywords: site-directed mutagenesis; antibiotic resistance; methylation; reaction mechanism; molecular evolution

*Corresponding authors

Introduction

The methylation of the N⁶ amino group of an adenine residue in 23 S ribosomal RNA (A2058 in *Escherichia coli*, A2085 in *Bacillus subtilis*) confers the resistance to clinically important macrolide, lincosamide and streptogramin B (MLS) antibiotics. Post-transcriptional generation of N⁶-methyladenine (m⁶A) or N⁶,N⁶-dimethyladenine (m²A) is carried out by S-adenosylmethionine (AdoMet)-dependent enzymes from the Erm (erythromycin (Em) ribosome methyltransferase) family.¹ This

modification has been to date reported mainly in Gram-positive bacteria, and only in a few pathogenic and commensal Gram-negative species,^{2–4} even though its target is conserved in all bacterial rRNA sequences.⁵ The methylatable adenine is located in a highly conserved region that participates in the formation of the peptidyl transferase center, which is the target of the MLS antibiotics.⁶ The presence of m⁶A or m²A markedly reduces the affinity between MLS antibiotics and the ribosome, which is the main factor of the MLS resistance phenotype.¹ The introduction of Em into clinical practice has increased the rate of evolution of bacterial resistance to the MLS antibiotics by increasing the number of *erm* genes in certain species and by horizontal transfer to unrelated genera.^{7,8}

The Erm methyltransferases (MTases) are related to the KsgA/Dim1p family, which groups together

Abbreviations used: Erm, erythromycin ribosome methyltransferase; MLS, macrolide, lincosamide and streptogramin B; MTase, methyltransferase; EM, erythromycin.

E-mail addresses of the corresponding authors: gordana@pharma.hr; iamb@genesilico.pl

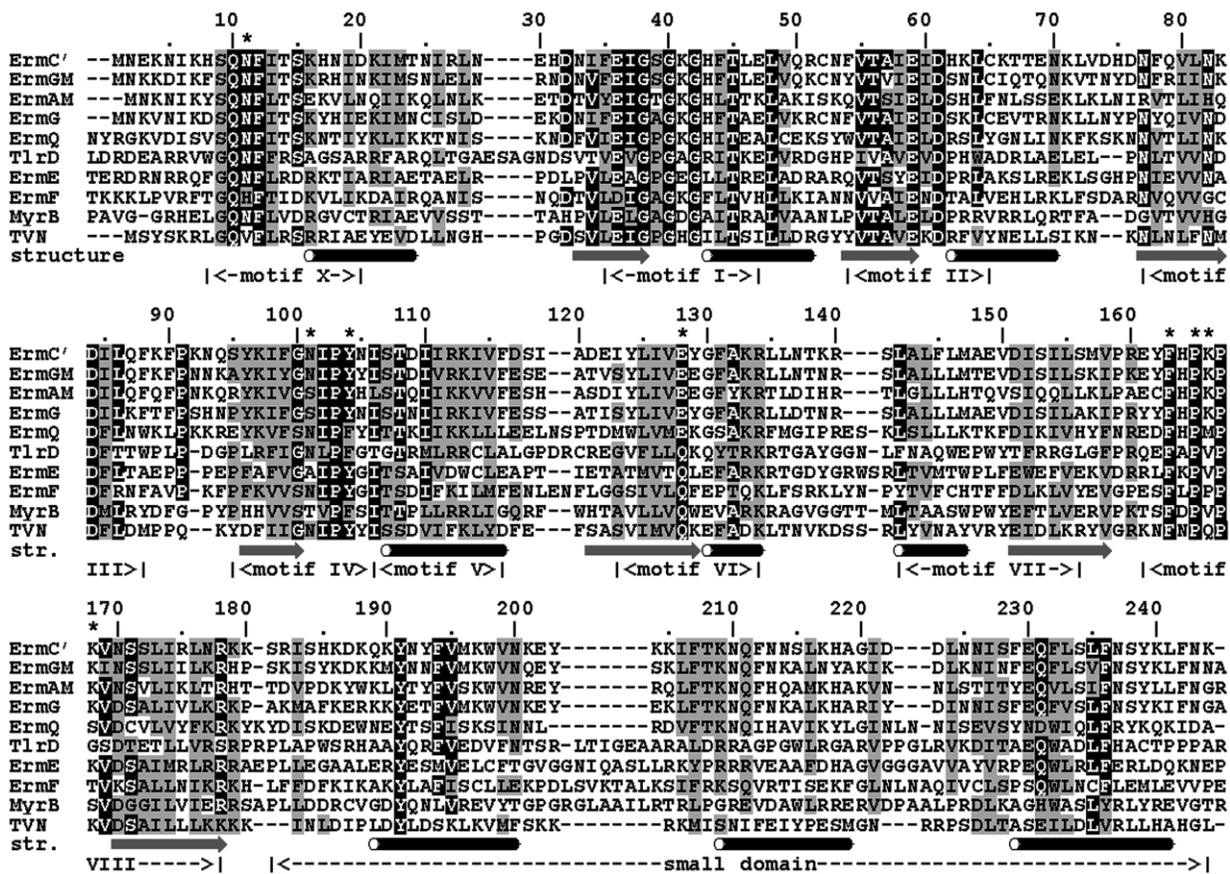


Figure 1. Multiple sequence alignment of the Erm family. Conserved amino acid residues are shown in black, physico-chemically similar residues are shown in gray. Secondary structure elements observed in *ErmC'* are indicated as tubes (helices) and arrows (strands). Conserved motifs are indicated below the alignment. Residues studied by mutagenesis in this work are indicated by asterisks (*).

enzymes that dimethylate two adjacent adenosine bases near the 3' end of 16S rRNA.⁹ The inactivation of the bacterial enzyme KsgA leads to resistance to the aminoglycoside antibiotic kasugamycin, but marginally affects the cell growth. Methylation exerted by Dim1p, the eukaryotic ortholog of KsgA, is dispensable, but the protein itself is essential, presumably due to its additional function as a quality controller in the ribosome synthesis.¹⁰

Our knowledge of the structure–function relationships in the rRNA:m⁶A MTases is limited mainly to members of the biomedically important Erm family. The structure of ErmAM has been determined by NMR,¹¹ followed by the solution of the crystal structure of ErmC' and its complexes with the cofactor and cofactor analogs.^{12,13} The structures of both Erm MTases are nearly identical.¹⁴ They comprise two structural domains. The larger N-terminal catalytic domain exhibits a typical $\alpha/\beta/\alpha$ sandwich architecture common to various AdoMet-dependent MTases that modify DNA, RNA, proteins, and small molecules.^{15,16} The smaller C-terminal domain consists of three α -helices and has been suggested to function as an

rRNA-binding domain.¹¹ This structure is unique among the MTases and is unrelated to known RNA-binding domains in other proteins.^{17,18} Recently it has been found that the mtTFB family of mitochondrial transcription factors exhibits the same structure as Erm and KsgA MTases.¹⁹ It has been also shown that the human mtTFB1 protein methylates the tandem adenine residues in the conserved stem–loop of human mitochondrial 12S rRNA and it can functionally complement the *E. coli* KsgA MTase,²⁰ suggesting an interesting link with aminoglycoside-induced deafness syndrome in humans.²¹

Despite the structural data available for ErmAM, ErmC' and mtTFB as well as functional analyses on the level of protein–RNA interactions in the Erm family, still very little is known about the enzymatic mechanism of m⁶A formation in the RNA. On the basis of comparisons of sequences and structures it was suggested that the mechanism of RNA:m⁶A methylation may be nearly identical with that of DNA:m⁶A MTases.¹³ In particular, the similarity of the RNA and DNA-adenine-N⁶ methylation mechanism was inferred from identity of the key residues in motif IV (NIPY in ErmC'), shown to be important for DNA MTases M.TaqI (NPPY)^{22,23} and M.EcoRV (DPPY).^{24,25} However, this prediction was based on the analysis of the

† <http://www.bioinfo.de/isb/1999-01/0016/>

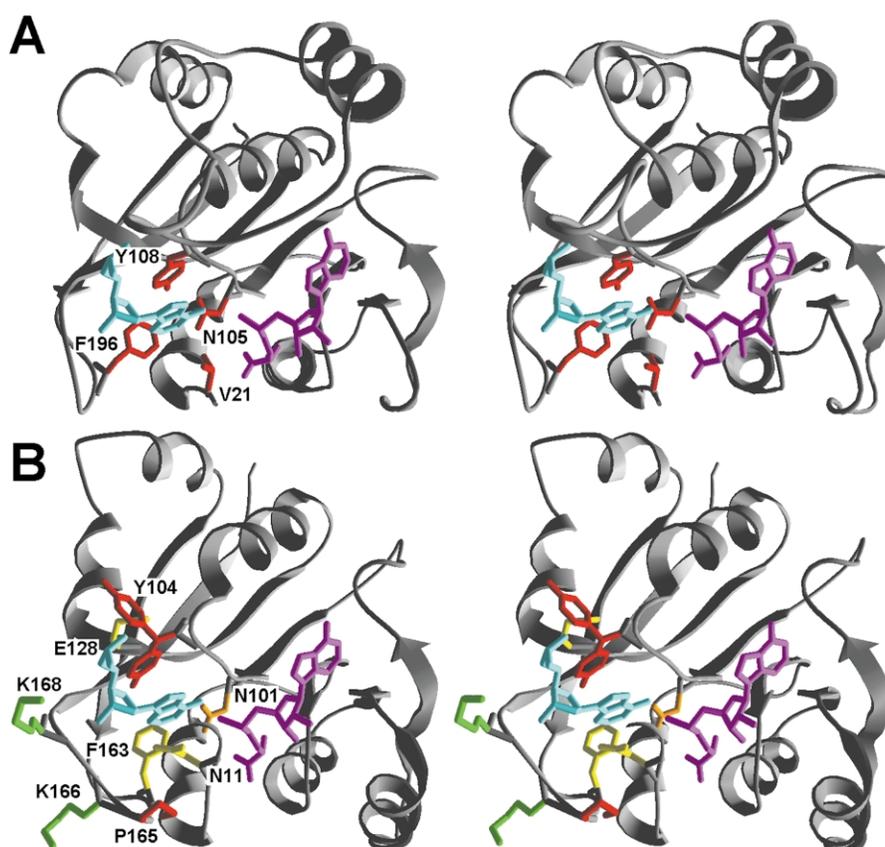


Figure 2. Comparison of the crystal structures of (A) M.TaqI and (B) ErmC'. For clarity, only the common catalytic domains are shown. The protein backbone is shown in the “ribbons” representation. AdoMet is shown in pink. Target adenine is shown in cyan. A, The coordinates of M.TaqI and the target adenine are taken from the crystal structure of M.TaqI complexed with DNA and a cofactor analog (PDB entry code 1g38),²³ while AdoMet is obtained from the M.TaqI-AdoMet cocrystal structure (PDB entry code 2adm).⁴⁵ Amino acid residues suggested to be essential for base-binding and catalysis are shown in red. B, The coordinates of ErmC' and AdoMet are taken from the structure of ErmC' in complex with cofactor AdoMet (PDB entry code 1qao),¹³ while coordinates of the base are derived from the superimposed M.TaqI structure (A). Residues analyzed in this work are shown in the wireframe representation. Color-coding indicates the effect of the mutation (compare with Figure 3): red, *in vitro* activity (A) < 5%; orange, 5% < A < 20%; yellow, 20% < A < 50%; green, A > 50%. Y104 is shown in two orientations, everted from the active site is the conformation observed in the ErmC'-AdoMet crystal structure,¹³ the stacking conformation with respect to the target adenine is hypothesized, based on the M.TaqI-DNA cocrystal structure (A).

ErmC' crystal structures with the cofactor and the cofactor analogs but without the target nucleic acid or the methylatable base.¹³ To our knowledge, no mutagenesis studies have been carried out to validate the presumably essential character of residues in the base-binding pocket of ErmC'. In order to understand the mechanism of RNA:m⁶A methylation, we performed site-directed mutagenesis of the predicted catalytic residues of ErmC' and studied the function of the mutants *in vitro* and *in vivo*. The sequence analysis allowed us to interpret these results not only in the light of the available structural data, but also in the phylogenetic context of the Erm family.

Results

The recently solved crystal structure of ErmC' complexed with AdoMet revealed amino acid residues involved in cofactor binding and suggested

the potential active site of the enzyme. On the basis of multiple sequence alignment of the Erm family members (Figure 1) and comparison of crystallographic structures of M.TaqI²³ and ErmC'¹³ (Figure 2), we have chosen several conserved residues (Asn¹¹, Asn¹⁰¹, Tyr¹⁰⁴, Glu¹²⁸, Phe¹⁶³, Pro¹⁶⁵, Lys¹⁶⁶ and Lys¹⁶⁸) to elucidate their role in AdoMet and target adenine binding and/or in catalysis. All amino acid residues were replaced with Ala by site-directed mutagenesis. We have tested the ErmC' mutants for their ability to generate Em resistance in Em-sensitive *E. coli* *in vivo*. We have also determined their kinetic parameters and RNA-binding affinities *in vitro*, and compared these parameters with those of the wild-type (wt) enzyme.

The effect of mutations on ErmC' function *in vivo*

To analyze quantitatively the effect of single

Table 1. Effects of ErmC' single mutations on erythromycin resistance

ErmC' variant	Em MIC (mg/l)
Negative control (empty pUC18)	80
Wild-type	>2560
N11A	2560
N101A	80
Y104A	80
E128A	2560
F163A	1280
P165A	640–1280
K166A	2560
K168A	>2560

mutations on ErmC' function *in vivo*, we have determined the minimal inhibitory concentrations (MICs) of Em for *E. coli* DH5 α cells harboring pUC18 plasmids with mutant *ermC'* genes. The Em MIC for the wt enzyme was >2560 mg/l, whereas for the negative control it was 80 mg/l. As shown in Table 1, mutants N101A and Y104A were totally unable to render DH5 α cells resistant to Em, thus giving us first indication that N101 and Y104 might play a crucial role in catalytic mechanism of ErmC'. Mutants F163A and P165A still mediated Em resistance, although at reduced level, suggesting that these amino acid residues are important, but not indispensable for the ErmC' activity. Results obtained for P165 are supported with the data presented for the Erm(B), where corresponding P164 was exchanged with serine giving MIC of 640 mg/l.²⁶ Finally, mutants N11A, E128A, K166A and K168A exhibited little or no difference in generating Em resistance when compared to the wt enzyme, indicating that these mutations do not affect severely the proper function of ErmC' *in vivo*.

RNA-binding affinity of purified ErmC' mutants

To investigate whether selected amino acid residues are involved in substrate binding, we have compared the RNA-binding affinities of purified mutant proteins with that of the wt enzyme. We have expressed and purified all proteins as described in Materials and Methods, with the

exception of F163A, which turned out to be fully insoluble when expressed at 30 °C and its expression was therefore carried out at 25 °C. In the filter binding assay we have determined the apparent dissociation constants of the ErmC' variants to a 32-mer RNA oligonucleotide that mimics the adenine loop in domain V of 23 S rRNA of *B. subtilis* and was used previously in the kinetic studies of the wt ErmC'.¹³ It is noteworthy that only 40–60% of the obtained complex was retained on the nitrocellulose membrane. This is in agreement with the observation made by Su & Dubnau, who have studied the binding of 23 S rRNA to native ErmC'.²⁷

Table 2 summarizes the apparent dissociation constant values obtained in this work. Most of the ErmC' mutants bind RNA with affinities similar to the wt protein (N11A, N101A, Y104A, P165A and K168A), hence we can exclude the possibility that the inactivity of N101A, Y104A and the lower activity of the P165A mutant shown *in vivo* are due to defects in RNA binding. Mutants E128A and F163A bind RNA substrate slightly better than the wt ErmC'. In the case of E128A this might be mere consequence of the removal of the negative charge. On the other hand, K166A mutant shows very mild decrease of the affinity towards the RNA substrate, which might be indication that this residue contributes to substrate-binding by electrostatic interactions with the negatively charged phosphate groups.

Kinetic characterization of catalytically active ErmC' variants

ErmC' has been shown to dimethylate its RNA substrate in two consecutive, random bi-bi reactions.²⁸ To gain detailed insight into biochemical function of individual amino acid residues selected for mutagenesis, we have determined kinetic parameters of ErmC' mutants for both the AdoMet cofactor and the RNA substrate. Interestingly, all the mutants proved to retain the catalytic activity *in vitro*, with the sole exception of Y104A, for which no activity could be measured (Table 2; Figure 3), in agreement with the predicted essential

Table 2. Summary of the *in vitro* characterization of ErmC' variants

Mutants	K_d^{app} ($M \times 10^{-7}$)	AdoMet			RNA		
		K_M ($M \times 10^{-6}$)	k_{cat} ($s^{-1} \times 10^{-2}$)	Relative k_{cat}/K_M	K_M ($M \times 10^{-6}$)	k_{cat} ($s^{-1} \times 10^{-3}$)	Relative k_{cat}/K_M
WT	1.2 \pm 0.3	2.8 \pm 0.2	6.6 \pm 0.1	1.00	0.7 \pm 0.2	6.4 \pm 0.4	1.00
N11A	1.3 \pm 0.2	4.0 \pm 0.7	2.9 \pm 0.1	0.32	0.8 \pm 0.1	2.3 \pm 0.1	0.33
N101A	1.1 \pm 0.4	13.9 \pm 2.9	3.3 \pm 0.3	0.10	0.8 \pm 0.2	0.7 \pm 0.1	0.11
Y104A	1.1 \pm 0.2	–	–	–	–	–	–
E128A	0.8 \pm 0.3	4.5 \pm 0.4	4.8 \pm 0.1	0.46	0.3 \pm 0.1	1.9 \pm 0.1	0.59
F163A	0.9 \pm 0.2	2.0 \pm 0.3	3.2 \pm 0.1	0.67	1.4 \pm 0.3	3.7 \pm 0.6	0.28
P165A	1.5 \pm 0.5	18.5 \pm 9.2	1.6 \pm 0.3	0.04	1.3 \pm 0.4	0.4 \pm 0.1	0.04
K166A	2.0 \pm 0.6	3.1 \pm 0.4	5.4 \pm 0.2	0.74	4.1 \pm 1.3	22.0 \pm 3.2	0.58
K168A	1.7 \pm 0.5	4.0 \pm 0.6	5.4 \pm 0.2	0.59	1.4 \pm 0.2	8.7 \pm 0.4	0.70

Experiments were done at least three times in duplicate to assess the error intervals.

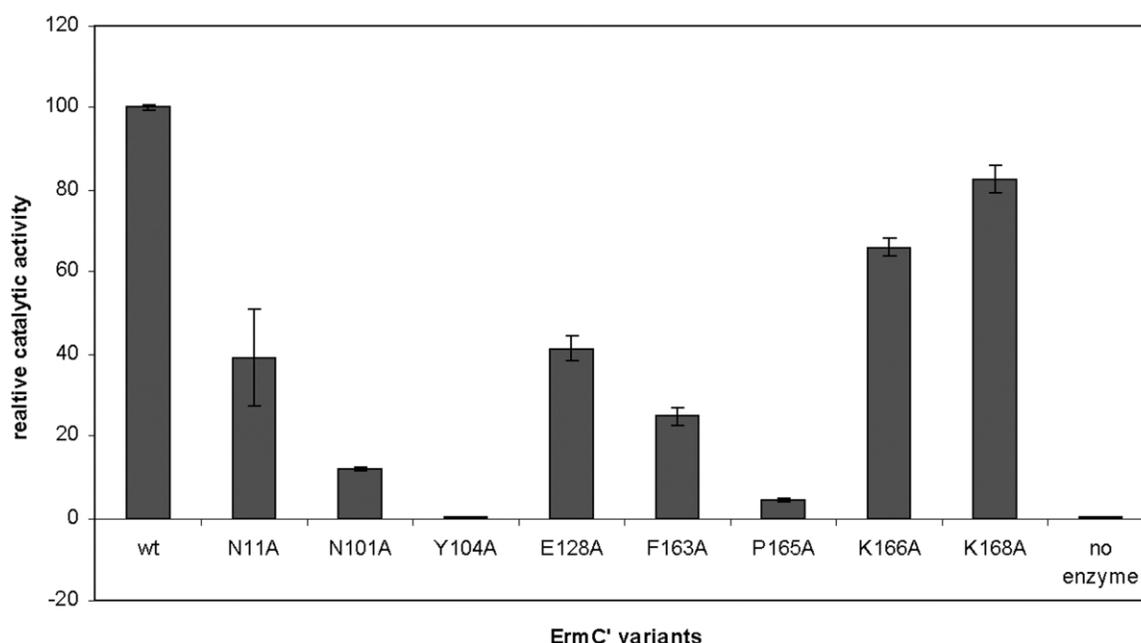


Figure 3. Relative catalytic activities of *ErmC'* variants. RNA oligonucleotide was methylated in the 50 μ l reaction mixture containing 1.1 μ M RNA, 0.2 μ M methyltransferase *ErmC'*, 0.13 μ M [3 H]AdoMet and one unit of RNasin in reaction buffer (50 mM Tris-HCl (pH 7.5), 40 mM KCl, 4 mM MgCl₂, 10 mM DTT) for 40 minutes at 25°C. Methylated RNA was TCA-precipitated, dried and counted for radioactivity. All experiments were carried out at least in triplicate. Relative activities of different *ErmC'* mutants are expressed with respect to the wild-type enzyme. Error bars indicate the maximal deviation between all results obtained for each mutant.

character of this residue and its involvement in stabilization of the target base (Figure 2). It is intriguing that the MTase activity of the N101A mutant *in vitro* is reduced only to 10% of the wt enzyme, while we found it is inactive *in vivo*. Both the apparent dissociation constant and K_M for RNA of N101A mutant are unchanged with respect to the wt *ErmC'*, suggesting that this substitution does not interfere with RNA binding, despite the interaction of N101 with the target adenine. On the other hand, the substantially decreased turnover number is accompanied with a great elevation of K_M for AdoMet, indicating that decreased catalytic efficiency is mediated by inability of N101A mutant to properly bind the cofactor.

The residue E128 was selected for mutagenesis because of its relative conservation in the Erm family (Figure 1) and spatial proximity to the essential residue Y104 in the crystal structure of *ErmC'* (Figure 2). The E128A mutant shows increased K_M for RNA substrate and a decrease in the K_d^{APP} , suggesting that E128 does not have an important role in target adenine binding. However, substitution of its carboxylate side-chain with alanine resulted in decreased k_{cat} , indicating its role in catalysis, which might be mediated through the interaction with vicinal Y104. It is interesting though, that this mutation does not greatly affect the *ErmC'* function *in vivo*, but this could be explained with different stability of mutant protein *in vivo* and *in vitro*.

N11 is well conserved in the Erm family (Figure 1). Its side-chain is directed away from AdoMet

and forms a wall of the predicted catalytic pocket (Figure 2). The N11A mutant has a similar affinity for the cofactor and substrate as the wt enzyme type, but its turnover number is threefold lower, which suggests that this amino acid is important (albeit not essential) for catalysis.

The opposite side of the predicted adenine-binding site comprises amino acid residues from the conserved motif VIII (Figures 1 and 2). We have analyzed four residues from this motif, F163, P165, K166 and K168. P165 is conserved in all Erm MTases (Figure 1). The catalytic activity of the P165A mutant *in vitro* is severely impaired followed by almost sevenfold increase in K_M for AdoMet and twofold increase in K_M for RNA. F163 is conserved among the Erm family members (Figure 1) and in nearly all DNA:m⁶A γ -MTs.²⁹ In both M.TaqI and *ErmC'* its side-chain forms the floor of the adenine-binding pocket (Figure 2). The exchange of this residue with alanine gave interesting results: the mutation does not affect the affinity towards AdoMet, thus excluding its involvement in cofactor binding. However, slightly decreased K_M for RNA accompanied with decreased k_{cat} result in decreased catalytic activity. Interestingly, the apparent dissociation constant showed elevated affinity towards the RNA substrate, but K_M value decreased, implying that stabilization of target adenine in ternary complex might be reduced despite apparent higher RNA-binding activity, which then leads to a lower turnover number.

Basic residues K166 and K168 are positioned close to the adenine-binding site and we presumed

they could support RNA binding even though they are not fully conserved in the Erm family (Figure 1). Kinetic parameters show that alanine substitutions in these positions do not interfere with cofactor binding, but strongly influence substrate binding, especially in the case of K166A where K_M for RNA is sevenfold higher when compared to the wt enzyme. Still, catalytic efficiency is not drastically diminished because the decrease in affinity towards RNA is compensated with a higher turnover number. K166 and K168 are clearly not very important for the catalytic step, nevertheless, with their positive charge they may create a favorable environment for RNA binding.

Discussion

The residues N101 and Y104 are located in the motif IV (101-NIPY), positioned between the cofactor-binding site and the predicted adenine-binding site in the ErmC' crystal structure¹³ (Figures 1 and 2). A relaxed form of this motif (N/D/S)-(I/P)-P-(Y/F/W/H) is generally conserved in the majority of MTases that methylate exocyclic amino groups of adenine, cytosine, or guanine in DNA or RNA.^{16,29–31} Structural studies suggested that the residue in the first position of the motif (N101 in ErmC') forms a hydrogen bond to the target amino group and may participate in its deprotonation during or after the methyl group transfer.^{23,32} Mutagenesis studies on DNA MTases have shown that residues in this position are indispensable for catalysis and that different amino acid residues are not interchangeable between the individual enzymes (i.e. the mutual interchanges of Asn, Asp, and Ser usually resulted in complete elimination of the catalytic activity.^{24,33,34} Interestingly, recent quantum mechanical studies of the DNA:m⁶A MTase M.TaqI complex with DNA and docked AdoMet²³ suggested that the optimal orientation of AdoMet and the methylatable base in the catalytic pocket is much more important than hydrogen bonding between Asn and the target amino group.²³ As described by structural and biophysical analyses on DNA:m⁶A MTases, the target base is typically stabilized by π -stacking between a pair of conserved aromatic residues (corresponding to Y104 and F163 in ErmC'). In the crystal structure of M.TaqI–DNA complex, the target adenine forms face-to-face and edge-to-face interactions with Tyr in motif IV and Phe in motif VIII, respectively.²³ Alanine substitutions of these residues in M.TaqI showed activity <1% of the wt enzyme,²² and similar results were also obtained for M.EcoRV.²⁵

The results of our study show, on the contrary to the expectations, that the active site of RNA:m⁶A MTase ErmC' is evidently more tolerant to alanine substitutions than observed for DNA:m⁶A MTases. Only the Y104 residue in ErmC' is absolutely essential for the activity *in vivo* and *in vitro*, while mutants N101A and F163A retain considerable

activity (N101A *in vitro* and F163 both *in vivo* and *in vitro*). Remarkably, the analysis of the multiple sequence alignment of the Erm family reveals that the predicted catalytic Asn residue is not universally conserved between all members (Figure 1). Some of the well-characterized and undoubtedly functional Erm MTases have a “naturally” occurring Ala (ErmE from *Saccharopolyspora erythraea*³⁵ or ErmA from *Aeromicrobium erythreum*³⁶) or Thr (MyrB from *Micromonospora griseorubida*³⁷) in the first position of motif IV. These observations reinforce the notion that the Asn residue is not essential for the rRNA:m⁶A methylation, even though it may be essential for DNA:m⁶A methylation in the γ -class of MTases.²³ That we were not able to observe the *in vivo* activity of the N101A mutant, may be explained by the additional negative effect on the ErmC' stability this mutation could have *in vivo*.

It is noteworthy that the amide NH₂ group of N101 forms a hydrogen bond with the carboxyl group of the AdoMet methionine moiety¹³ (Figure 4). In accordance with this finding, the N101A mutant of ErmC' has fivefold increased K_M for AdoMet (Table 2). These observations suggest that N101 of ErmC' does not participate in the key catalytic step such as hybridization change of the attacking adenine N⁶ nitrogen toward sp³, as it was suggested for the DNA:m⁶A methylation mechanism. We suggest that the main role of N101 is to minimize the transition state energy by the stabilization of the AdoMet and the target base in the optimal orientation. Another important element essential for the stabilization of the target base, which is universally conserved in all MTases, is the backbone carbonyl oxygen of the residue following N101 (I102 in ErmC'; Figure 4). The relative importance of the Asn residue probably depends on the subtle modifications of the catalytic pocket and on the relative stabilization of the target base and the cofactor required by different enzymes to perform the catalysis. This explains why in some adenine MTases the Asn residue (or Asp in some DNA:m⁶A MTases and mRNA:m⁶A MTases) is essential, while in the others it is dispensable. Our results are in perfect agreement with conclusions of a recent quantum mechanical examination of the factors controlling the catalytic efficiency of M.TaqI, that significant rate enhancement is provided by the optimal positioning of AdoMet and adenine in the active site and not by the hydrogen bonding of functional groups to adenine *per se*.³⁸

The different behavior of the F163A mutant of ErmC' (activity reduced only ca 4-fold) compared to the F196 mutant of M.TaqI (activity reduced 400-fold) may be explained by the different orientation of the phenylalanyl side-chain with respect to the methylated base. The comparison with the M.TaqI–DNA complex reveals that the side-chains of F163 in ErmC' and F196 in M.TaqI are attached to the polypeptide backbone exhibiting different conformation (Figure 2). Unlike the Y104 side-chain of ErmC', which can be rotated to mimic the

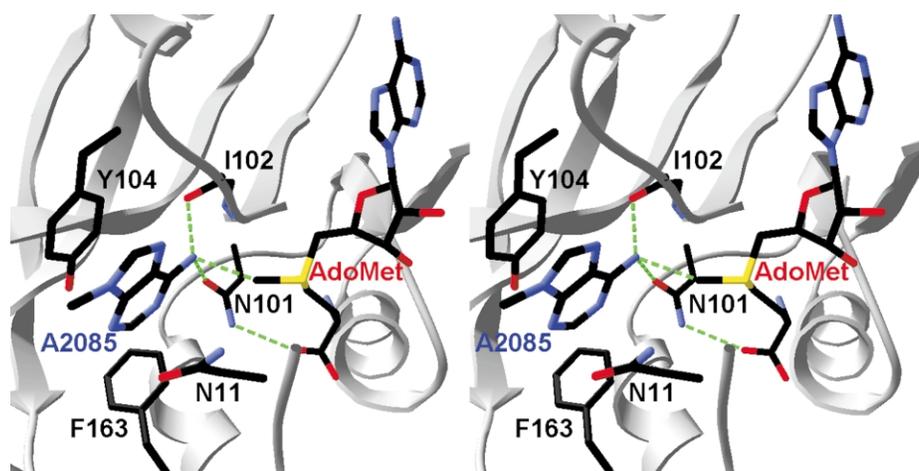


Figure 4. Predicted interactions between the residues in the active site of *ErmC'*, AdoMet and the target adenine (A2085 in *B. subtilis* 23 S rRNA). Green dotted lines indicate hydrogen bonds and the interaction between the methyl group of AdoMet and the attacking amino group of adenine. In addition to side-chains interacting with the target base, the backbone of I102, providing the key carbonyl oxygen group is shown. C, N, O, and S atoms are shown in black, blue, red, and yellow, respectively.

orientation of Y108 and face-to-face stacking with the adenine in *M.TaqI* (Figure 2), the F163 side-chain of *ErmC'* cannot be rotated to superimpose onto F196 of *M.TaqI* unless a conformational rearrangement of the catalytic pocket occurs. Although a significant conformational change upon RNA binding by *ErmC'* cannot be excluded, it remains plausible that F163 interacts with the

target adenine differently than F196 in *M.TaqI*.²³ The results of our kinetic analyses of the F163A mutant suggest that F163 enhances both the binding of the RNA substrate and the catalysis itself, probably by contributing to the stabilization of the target base in the flipped-out conformation.

The mutational analysis of quite well conserved residues N11 and E128, located relatively closely

Table 3. Deoxyoligonucleotides employed

Deoxyoligo-nucleotide	Description	Sequence
Oligo-1	Sense PCR primer for <i>ermC'</i> cloning	AAAACCTGCAGTATAAAATTTAACGATCAC (PstI)
Oligo-2	Antisense PCR primer for <i>ermC'</i> cloning	CGCGGATCCCCCTACGAGGTTGTGTCG (BamHI)
Oligo-3	Sense PCR primer for the introduction of NdeI site	AAGAGGGTTCATATGAACGAGAAAA (NdeI)
Oligo-4	Antisense PCR primer for the introduction of NdeI site	CTCGTTCATATGAACCCCTCTTTATTT (NdeI)
Oligo-5	Sense PCR primer for the introduction of XhoI site and His ₆ -Tag	CACCATCACCATCACCATTAACCTCGAGGTTAAGG GATGCATAAACTGC (His ₆ -Tag; XhoI)
Oligo-6	Antisense PCR primer for the introduction of XhoI site and His ₆ -Tag	CTCGAGTTAATGGTGATGGIGATGGIGCTTATTA ATAATTTATAGCTATTG (XhoI; His ₆ -Tag)
Oligo-7	Sense PCR primer for N11A	ATAAAACACAGTCAAGCCITTTATTACTTCAAAC
Oligo-8	Antisense PCR primer for N11A	GTTTTGAAGTAATAAAGGCTTGACTGTGTTTTAT
Oligo-9	Sense PCR primer for N101A	CCTATAAAATATTTGGTGCTATACCTTATAACAT AAGT
Oligo-10	Antisense PCR primer for N101A	ACTTATGTTATAAGGTATAGCACCAAATATTTTAC AGG
Oligo-11	Sense PCR primer for Y104A	TTTGTAATATACCTGCTAACATAAGTACGGATATA
Oligo-12	Antisense PCR primer for Y104A	TATATCCGTACTTATGTTAGCAGGTATATTACAAA
Oligo-13	Sense PCR primer for E128A	TTAATCGTGGCATACGGGTTTTG
Oligo-14	Antisense PCR primer for E128A	CCCGTATGCCACGATTAATAAAT
Oligo-15	Sense PCR primer for F163A	GGTTCCAAGAGAATATGCTCATCCTAAACCTAA AGTG
Oligo-16	Antisense PCR primer for F163A	CACTTTAGGTTTAGGATGAGCATATTCTCTTGG AACC
Oligo-17	Sense PCR primer for P165A	AGAGAATATTTTCATGCTAAACCTAAAGTGAATAGC
Oligo-18	Antisense PCR primer for P165A	GCTATTCACITTTAGGTTTAGCATGAAAATATTCTCT
Oligo-19	Sense PCR primer for K166A	TCATCCTGCACCTAAAGTGAATAG
Oligo-20	Antisense PCR primer for K166A	CTATTCACITTTAGGTTGACAGGATGA
Oligo-21	Sense PCR primer for K168A	ATCCTAAACCTGCAGTGAATAGC
Oligo-22	Antisense PCR primer for K168A	GCTATTCACITGACAGGTTTAGGAT

to the active site (Figure 2), revealed that they are not essential for RNA binding or catalysis. The N11 residue forms a wall of the predicted catalytic pocket, in some *Erm* sequences it is substituted by His or Val (Figure 1). We suspect that the modest deficiency in activity of the N11A mutant (35% of the wt) results from slight destabilization of the flipped-out adenine. E128 is also strongly conserved in the *Erm* family; among the known sequences, it is interchangeable only with Gln (Figure 1). Remarkably, its carboxylate side-chain is partially buried in the hydrophobic environment and does not make any polar interactions with the neighboring residues or the polypeptide backbone, but packs closely against the aromatic ring of Y104. Kinetic parameters of the E128A mutant suggest that the identity of the Glu side-chain (or Gln, which is isosteric to Glu with respect to non-hydrogen atoms) may be important for assuming the optimal conformation by the essential Y104 residue and not for making direct contacts with the substrate. In particular, we speculate that E128 may facilitate the rotation of the Y104 side-chain, which is directed away from the binding site in the crystal structure of *ErmC*,¹³ but probably assumes the M.TaqI-like orientation upon substrate-binding (Figure 2).

We have also studied the function of residues outside the catalytic pocket. P165 is one of two conserved Pro residues in the motif VIII that presumably stabilize the loop, which serves as a "floor" of the adenine-binding site.¹³ P165 also makes contacts with the N-terminal tail of *ErmC*, in particular with N11 (Figure 2). The catalytic activity of the P165A mutant is severely impaired and kinetic parameters suggest that this substitution affects both AdoMet and RNA binding. Pro165 is found in *cis*-conformation in *Erm* MTases and is fully conserved throughout the family. The substantial effect of the P165A mutation on the *ErmC* activity may result from a significant deformation of both the cofactor-binding site and the catalytic pocket. Upon replacement of the conserved *cis*-Pro residue with Ala, the structural conformation of the adenine-binding loop is lost, which in turn may also reflect to the orientation of the neighboring F163 residue found to be important for the stabilization of the target adenine. Finally, residues K166 and K168 are not universally conserved in the *Erm* family (Figure 1); however, their localization in the motif VIII and exposition to the surface of the protein suggested they may be involved in RNA binding. Nonetheless, K166A and K168A mutants manifested only minor decrease of the activity *in vitro* and no visible defect *in vivo*, suggesting that they are not important for substrate binding and catalysis.

In conclusion, our results from a mutational analysis of the predicted catalytic pocket of *ErmC* indicate that Y104 is essential for the rRNA:m⁶A MTase activity and suggest that it is involved in the stabilization of the target adenine. This function is much less pronounced for F163 and N11,

which form the floor and one of the walls of the base-binding pocket. Our data also indicate that the Asn residue from the "catalytic" motif IV is not essential for catalysis in RNA:m⁶A MTases. We suggest that N101 participates in stabilization of the optimal orientation of the cofactor and the methylatable amino group rather than in any essential "chemical" step of the reaction mechanism. These somewhat unexpected results suggest that the active site of rRNA:m⁶A MTases is much more tolerant to amino acid substitutions than the active site of DNA:m⁶A MTases. In the absence of *ErmC*-RNA co-crystal structure, our data will serve as a useful platform to guide further analyses aiming at elucidation of interactions between the MTase, the cofactor, and the substrate, which may be essential for successful development of inhibitors against this biomedically important enzyme.

Materials and Methods

Bacterial strains and plasmids

B. subtilis BD1167 carrying naturally occurring plasmid pIM13 with the *ermC* gene³⁹ was kindly provided by Dr David Dubnau, New York University School of Medicine. *E. coli* DH5 α and BL21(DE3)pLysS and the expression vector pET-25b(+) were obtained from Novagen. Cloning vector pUC18 was from Amersham Biosciences.

Gene cloning and site-directed mutagenesis

The *ermC* gene together with its native promoter and terminator was amplified from the *B. subtilis* plasmid pIM13. Simultaneously, four new restriction sites were introduced using PCR-overlapping method. PstI and BamHI sites were introduced at 5' and 3' end of the construct respectively, while NdeI and XhoI sites were introduced to flank the *ermC* coding sequence. Additionally, codons for His₆-Tag were introduced immediately before the STOP codon to facilitate protein purification. Obtained construct was cloned into PstI and BamHI sites of pUC18 vector. Site-directed mutagenesis was carried out described by either PCR-overlapping method or QuikChange protocol (Stratagene). All introduced changes as well as the absence of unwanted mutations were confirmed by DNA sequencing. For the needs of protein expression and purification, coding sequences of mutant genes were recloned into NdeI and XhoI sites of the expression vector pET-25b(+). Oligonucleotides used in this work are listed in Table 3.

Protein expression and purification

E. coli BL21(DE3)pLysS cells carrying mutant genes in the pET-25b(+) vector were grown at 30 °C in Luria-Bertani medium supplemented with 100 μ g/ml of ampicillin and 50 μ g/ml of chloramphenicol. At A_{600} of 1.0, the expression was induced with 1 mM IPTG and carried out for five hours at 30 °C. Proteins were purified in a two-step chromatographic procedure. Affinity chromatography was carried out on a HiTrap Chelating column (Amersham Biosciences) described by manufacturer's instructions. Mutant proteins were eluted with a

linear gradient of imidazole (0.1 M–0.3 M imidazole in 50 mM phosphate buffer, pH 8.0, containing 0.3 M NaCl). Fractions containing partially purified proteins were pooled and diluted with TDGM buffer (50 mM Tris–HCl (pH 8.0), 5 mM DTT, 10% (v/v) glycerol, 10 mM MgCl₂) to reduce the ionic strength and then applied to a cation-exchange HiTrap SP column (Amersham Biosciences) equilibrated with TDGM buffer. *ErmC'* variants were eluted with a linear NaCl gradient (0.2 M–0.6 M NaCl in TDGM buffer). Obtained pure proteins (purity >95%) were concentrated on the YM-10 (Amicon) membrane and stored at –80 °C in TDGM buffer containing 150 mM NaCl.

Determination of Em MICs

Erythromycin MICs were determined in *E. coli* DH5 α essentially as described previously.²⁶ Briefly, overnight cultures of DH5 α cells carrying mutant genes in pUC18 vector were diluted 1:25 in fresh 2YT medium supplemented with 100 μ g/ml of ampicillin and grown until the A₅₀₀ of 0.7–0.8. Following 100-fold dilution, a 5 μ l aliquot was applied to 2YT plates which contained 100 μ g/ml of ampicillin and various concentrations of Em (80, 160, 320, 640, 1280 and 2560 mg/l). Plates were grown for 18 hours at 37 °C and MIC was determined as a minimal concentration of Em that inhibits confluent growth.

Filter-binding assay

RNA-binding properties of mutant proteins were determined using a synthetic RNA oligonucleotide (32-mer CGCGACGGACGGA²⁰⁸⁵AAGACCCCUAUCC GUCGCG, hairpin structure) designed to mimic the adenine loop in domain V of *B. subtilis* 23 S rRNA (residues 2073–2090 and 2638–2651) and used previously in studies of the wt enzyme.¹³ The oligonucleotide substrate was labeled radioactively using adenine-5'-[γ -³²P]triphosphate (Amersham Biosciences) and bacteriophage T4 polynucleotide kinase (New England Biolabs). RNA (10 nM) was incubated with 50–1500 nM *ErmC'* variants in binding buffer (40 mM Tris–HCl (pH 7.6), 40 mM KCl, 4 mM Mg(OAc)₂, 10 mM DTT, 1 mM EDTA, 0.2 mg/l of BSA) with the addition of one unit of RNasin (Promega) per reaction mixture. Binding reactions were carried out in reaction volume of 20 μ l for 25 minutes at 37 °C. Nitrocellulose filter sheets Optitran BA-S 83 (pore size 0.22 μ m) from Schleicher & Schuell were preincubated for two hours in binding buffer. Presoaked filters were placed into a dot-blot apparatus (96 wells, Schleicher & Schuell). Wells were washed with 100 μ l of binding buffer immediately before the samples were applied. To minimize non-specific background retention of RNA, only six samples were applied to the wells at a time. Fifteen microliters of reaction mixture was vacuum filtered and wells were immediately flushed with 100 μ l of binding buffer. After drying, filters were exposed overnight to the intensifying screen and the amounts of bound complexes were determined using Cyclon Phosphoimager System (Canberra Packard). Binding curves and determination of the apparent dissociation constants were done using the program SigmaPlot, version 8.0 (SPSS, Inc.). Experiments were repeated at least three times in duplicate.

Methylation assay

Methylation of RNA *in vitro* was done according to previously described procedures^{13,40} with some modifications. RNA oligonucleotide was denatured at 90 °C for one minute and renatured by cooling slowly to room temperature. Reaction was carried out in methylation buffer (50 mM Tris–HCl (pH 7.5), 40 mM KCl, 4 mM MgCl₂, 10 mM DTT) containing 1.1 μ M RNA, 0.2 μ M MTase *ErmC'*, 0.13 μ M [*methyl*-³H]AdoMet (16.5 Ci/mmol) and one unit of RNasin in a total reaction volume of 50 μ l. [*methyl*-³H]AdoMet (82 Ci/mmol) was from Amersham Biosciences, non-radioactive AdoMet was from Sigma. All additions were performed at 0 °C and reaction mixtures were then transferred to a 25 °C water-bath for 40 minutes.

In kinetic experiments, reaction mixtures were preincubated at room temperatures before the addition of the enzyme. Kinetic parameters for AdoMet were determined using AdoMet concentrations from 1 μ M to 10 μ M. RNA kinetic parameters were determined using RNA concentration range from 0.1 μ M to 1.5 μ M. Reactions were stopped by the addition of 0.5 ml of 10% (w/v) trichloroacetic acid (TCA, Sigma) and the carrier RNA was added to facilitate the precipitation. The precipitated RNA was pelleted by centrifugation. RNA pellets were then washed with 1 ml of 10% TCA, dried and counted for radioactivity. RNA-free and enzyme-free blanks yielded 30–60 cpm under these conditions. All experiments were done at least three times in duplicates. Kinetic parameters were calculated from double reciprocal plots using the Kinetic module of the program SigmaPlot, version 8.0 (SPSS, Inc.).

Bioinformatics analyses

PSI-BLAST⁴¹ was used to search the non-redundant version of current sequence databases (nr) at NCBI, Bethesda, USA†. All full-length sequences were retrieved and realigned using the CLUSTALX program⁴² to the profiles obtained from the degapped multiple sequence alignments reported by PSI-BLAST. The CLUSTALX alignment was assessed on the atomic level, by homology modeling using MODELLER⁴³ and evaluation of the models by VERIFY3D.⁴⁴ If sequence–structure incompatibility was detected, the alignment was modified manually and the model-building and evaluation procedure was reiterated.

Acknowledgements

We are grateful to Dr David Dubnau for the generous gift of the *B. subtilis* BD1167. This work was supported by the ICGEB fellowship to G.M. and by the Ministry of Science of Republic of Croatia (#006320, #006611). J.M.B. is supported by the KBN grant 6P04B00519, the Young Investigator award from EMBO and Howard Hughes Medical Institute, and by a fellowship from the Foundation for Polish Science. The work at the ICGEB Protein Structure and Function Group was partly

† <http://www.ncbi.nlm.nih.gov>

supported by EU grant ENGEM QLK3-CT-2001-00448. We are indebted to Dr Stephen Douthwaite for the critical reading of the manuscript and useful comments. We also thank an anonymous referee for suggestions concerning the P165A mutant.

References

- Weisblum, B. (1995). Erythromycin resistance by ribosome modification. *Antimicrob. Agents Chemother.* **39**, 577–585.
- Tanaka, T. & Weisblum, B. (1975). Systematic difference in the methylation of ribosomal ribonucleic acid from gram-positive and Gram-negative bacteria. *J. Bacteriol.* **123**, 771–774.
- Roberts, M. C., Chung, W. O., Roe, D., Xia, M., Marquez, C., Borthagaray, G. *et al.* (1999). Erythromycin-resistant *Neisseria gonorrhoeae* and oral commensal *Neisseria* spp. carry known rRNA methylase genes. *Antimicrob. Agents Chemother.* **43**, 1367–1372.
- Chung, W. O., Werckenthin, C., Schwarz, S. & Roberts, M. C. (1999). Host range of the *ermF* rRNA methylase gene in bacteria of human and animal origin. *J. Antimicrob. Chemother.* **43**, 5–14.
- Wuyts, J., De Rijk, P., Van de, P. Y., Winkelmanns, T. & De Wachter, R. (2001). The European large subunit ribosomal RNA database. *Nucl. Acids Res.* **29**, 175–177.
- Harms, J., Schluenzen, F., Zarivach, R., Bashan, A., Gat, S., Agmon, I. *et al.* (2001). High resolution structure of the large ribosomal subunit from a mesophilic eubacterium. *Cell*, **107**, 679–688.
- Arthur, M., Brisson-Noel, A. & Courvalin, P. (1987). Origin and evolution of genes specifying resistance to macrolide, lincosamide and streptogramin antibiotics: data and hypotheses. *J. Antimicrob. Chemother.* **20**, 783–802.
- Shoemaker, N. B., Vlamakis, H., Hayes, K. & Salyers, A. A. (2001). Evidence for extensive resistance gene transfer among *Bacteroides* spp. and among *Bacteroides* and other genera in the human colon. *Appl. Environ. Microbiol.* **67**, 561–568.
- Van Buul, C. P. & Van Knippenberg, P. H. (1985). Nucleotide sequence of the *ksgA* gene of *Escherichia coli*: comparison of methyltransferases effecting dimethylation of adenosine in ribosomal RNA. *Gene*, **38**, 65–72.
- Lafontaine, D. L., Preiss, T. & Tollervey, D. (1998). Yeast 18 S rRNA dimethylase Dim1p: a quality control mechanism in ribosome synthesis? *Mol. Cell. Biol.* **18**, 2360–2370.
- Yu, L., Petros, A. M., Schnuchel, A., Zhong, P., Severin, J. M., Walter, K. *et al.* (1997). Solution structure of an rRNA methyltransferase (ErmAM) that confers macrolide-lincosamide-streptogramin antibiotic resistance. *Nature Struct. Biol.* **4**, 483–489.
- Bussiere, D. E., Muchmore, S. W., Dealwis, C. G., Schluckebier, G., Nienaber, V. L., Edalji, R. P. *et al.* (1998). Crystal structure of ErmC', an rRNA methyltransferase which mediates antibiotic resistance in bacteria. *Biochemistry*, **37**, 7103–7112.
- Schluckebier, G., Zhong, P., Stewart, K. D., Kavanaugh, T. J. & Abad-Zapatero, C. (1999). The 2.2 Å structure of the rRNA methyltransferase ErmC' and its complexes with cofactor and cofactor analogs: implications for the reaction mechanism. *J. Mol. Biol.* **289**, 277–291.
- Abad-Zapatero, C., Zhong, P., Bussiere, D. E., Stewart, K. & Muchmore, S. W. (1999). rRNA methyltransferases (ErmC' and ErmAM) and antibiotic resistance. In *S-Adenosylmethionine-dependent Methyltransferases: Structures and Functions* (Cheng, X. & Blumenthal, R. M., eds), pp. 199–226, World Scientific Inc, Singapore.
- 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**, 1–8.
- Fauman, E. B., Blumenthal, R. M. & Cheng, X. (1999). Structure and evolution of AdoMet-dependent MTases. In *S-Adenosylmethionine-dependent Methyltransferases: Structures and Functions* (Cheng, X. & Blumenthal, R. M., eds), pp. 1–38, World Scientific Inc, Singapore.
- Burd, C. G. & Dreyfuss, G. (1994). Conserved structures and diversity of functions of RNA-binding proteins. *Science*, **265**, 615–621.
- Draper, D. E. & Reynaldo, L. P. (1999). RNA binding strategies of ribosomal proteins. *Nucl. Acids Res.* **27**, 381–388.
- Schubot, F. D., Chen, C. J., Rose, J. P., Dailey, T. A., Dailey, H. A. & Wang, B. C. (2001). Crystal structure of the transcription factor sc-mtTFB offers insights into mitochondrial transcription. *Protein Sci.* **10**, 1980–1988.
- Seidel-Rogol, B. L., McCulloch, V. & Shadel, G. S. (2003). Human mitochondrial transcription factor B1 methylates ribosomal RNA at a conserved stem-loop. *Nature Genet.* **33**, 23–24.
- Prezant, T. R., Agapian, J. V., Bohlman, M. C., Bu, X., Oztas, S., Qiu, W. Q. *et al.* (1993). Mitochondrial ribosomal RNA mutation associated with both antibiotic-induced and non-syndromic deafness. *Nature Genet.* **4**, 289–294.
- Pues, H., Bleimling, N., Holz, B., Wolcke, J. & Weinhold, E. (1999). Functional roles of the conserved aromatic amino acid residues at position 108 (motif IV) and position 196 (motif VIII) in base flipping and catalysis by the N6-adenine DNA methyltransferase from *Thermus aquaticus*. *Biochemistry*, **38**, 1426–1434.
- Goedecke, K., Pignot, M., Goody, R. S., Scheidig, A. J. & Weinhold, E. (2001). Structure of the N6-adenine DNA methyltransferase M.TaqI in complex with DNA and a cofactor analog. *Nature Struct. Biol.* **8**, 121–125.
- Roth, M., Helm-Kruse, S., Friedrich, T. & Jeltsch, A. (1998). Functional roles of conserved amino acid residues in DNA methyltransferases investigated by site-directed mutagenesis of the EcoRV adenine-N6-methyltransferase. *J. Biol. Chem.* **273**, 17333–17342.
- Jeltsch, A., Roth, M. & Friedrich, T. (1999). Mutational analysis of target base flipping by the EcoRV adenine-N6 DNA methyltransferase. *J. Mol. Biol.* **285**, 1121–1130.
- Farrow, K. A., Lyras, D., Polekhina, G., Koutsis, K., Parker, M. W. & Rood, J. I. (2002). Identification of essential residues in the ErmB rRNA methyltransferase of *Clostridium perfringens*. *Antimicrob. Agents Chemother.* **46**, 1253–1261.
- Su, S. L. & Dubnau, D. (1990). Binding of *Bacillus*

- subtilis* ermC' methyltransferase to 23 S rRNA. *Biochemistry*, **29**, 6033–6042.
28. Denoya, C. & Dubnau, D. (1989). Mono- and dimethylating activities and kinetic studies of the ErmC 23 S rRNA methyltransferase. *J. Biol. Chem.* **264**, 2615–2624.
 29. Malone, T., Blumenthal, R. M. & Cheng, X. (1995). Structure-guided analysis reveals nine sequence motifs conserved among DNA amino-methyltransferases, and suggests a catalytic mechanism for these enzymes. *J. Mol. Biol.* **253**, 618–632.
 30. Bujnicki, J. M. (2000). Phylogenomic analysis of 16 S rRNA:(guanine-N2) methyltransferases suggests new family members and reveals highly conserved motifs and a domain structure similar to other nucleic acid amino-methyltransferases. *FASEB J.* **14**, 2365–2368.
 31. Bujnicki, J. M., Feder, M., Radlinska, M. & Blumenthal, R. M. (2002). Structure prediction and phylogenetic analysis of a functionally diverse family of proteins homologous to the MT-A70 subunit of the human mRNA:m⁶A methyltransferase. *J. Mol. Evol.* **55**, 431–444.
 32. Gong, W., O'Gara, M., Blumenthal, R. M. & Cheng, X. (1997). Structure of PvuII DNA-(cytosine N4) methyltransferase, an example of domain permutation and protein fold assignment. *Nucl. Acids Res.* **25**, 2702–2715.
 33. Radlinska, M. & Bujnicki, J. M. (2001). Site-directed mutagenesis defines the catalytic aspartate in the active site of the atypical DNA:m⁴C methyltransferase M.NgoMXV. *Acta Microbiol. Pol.* **50**, 93–101.
 34. Roth, M. & Jeltsch, A. (2001). Changing the target base specificity of the EcoRV DNA methyltransferase by rational *de novo* protein-design. *Nucl. Acids Res.* **29**, 3137–3144.
 35. Uchiyama, H. & Weisblum, B. (1985). N-Methyl transferase of *Streptomyces erythraeus* that confers resistance to the macrolide-lincosamide-streptogramin B antibiotics: amino acid sequence and its homology to cognate R-factor enzymes from pathogenic bacilli and cocci. *Gene*, **38**, 103–110.
 36. Roberts, A. N., Hudson, G. S. & Brenner, S. (1985). An erythromycin-resistance gene from an erythromycin-producing strain of *Arthrobacter* sp. *Gene*, **35**, 259–270.
 37. Inouye, M., Morohoshi, T., Horinouchi, S. & Beppu, T. (1994). Cloning and sequences of two macrolide-resistance-encoding genes from mycinamicin-producing *Micromonospora griseorubida*. *Gene*, **141**, 39–46.
 38. Newby, Z. E., Lau, E. Y. & Bruce, T. C. (2002). A theoretical examination of the factors controlling the catalytic efficiency of the DNA-(adenine-N6)-methyltransferase from *Thermus aquaticus*. *Proc. Natl Acad. Sci. USA*, **99**, 7922–7927.
 39. Monod, M., Denoya, C. & Dubnau, D. (1986). Sequence and properties of pIM13, a macrolide-lincosamide-streptogramin B resistance plasmid from *Bacillus subtilis*. *J. Bacteriol.* **167**, 138–147.
 40. Zhong, P., Pratt, S. D., Edalji, R. P., Walter, K. A., Holzman, T. F., Shivakumar, A. G. & Katz, L. (1995). Substrate requirements for ErmC' methyltransferase activity. *J. Bacteriol.* **177**, 4327–4332.
 41. Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucl. Acids Res.* **25**, 3389–3402.
 42. Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F. & Higgins, D. G. (1997). The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucl. Acids Res.* **25**, 4876–4882.
 43. Sali, A. & Blundell, T. L. (1993). Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* **234**, 779–815.
 44. Luthy, R., Bowie, J. U. & Eisenberg, D. (1992). Assessment of protein models with three-dimensional profiles. *Nature*, **356**, 83–85.
 45. Labahn, J., Granzin, J., Schluckebier, G., Robinson, D. P., Jack, W. E., Schildkraut, I. & Saenger, W. (1994). Three-dimensional structure of the adenine-specific DNA methyltransferase M.TaqI in complex with the cofactor S-adenosylmethionine. *Proc. Natl Acad. Sci. USA*, **91**, 10957–10961.

Edited by K. Nagai

(Received 7 April 2003; received in revised form 30 June 2003; accepted 4 July 2003)