We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

4,800
Open access books available

122,000
International authors and editors

135M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

CLARIVATE ANALYTICS BOOK CITATION INDEX INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

Chapter 3

Bifunctional Prokaryotic DNA-Methyltransferases

Dmitry V. Nikitin, Attila Kertesz-Farkas,
Alexander S. Solonin and Marina L. Mokrishcheva

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/51025

1. Introduction

Restriction-modification systems (RMS) are prokaryotic tools against invasion of foreign DNAs into cells [1]. They reduce horizontal gene transfer, thus stimulating microbial biodiversity. Usually, they consist of a restriction endonuclease (REase) and a modification DNA methyltransferase (MTase) enzyme recognising the same short 4-8 nucleotide sequence. MTase is responsible for methyl group transfer to adenine or cytosine nucleotides within the target sequence, thus preventing its hydrolysis by cognate REase. Up to now, more than 20 000 different RMS have been collected in the REBASE, the database holding all known, and many putative, RMS [2]. Many of these RMS have head-to-tail gene orientation, thus providing, by our hypothesis, the possibility of gene fusion through point mutations or genome rearrangements such as deletions, insertions, inversions or translocations. These events could be responsible for the origin of bifunctional restriction enzymes of type IIC [3] such as AloI, BcgI, BseMII, BseRI, BsgI, BspLU11III, CjeI, Eco57I, HaeIV, MmeI, PpiI, TstI and TspWGI; bifunctional MTases such as FokI and LlaI, and regulatory SsoII-related MTases [4].

In our previous work we proved the possibility of fully functional hybrid polypeptide origin through gene fusion, taking as an example Eco29kI RMS. In the given RMS, the REase gene precedes MTase gene and their Stop and Start codons overlap. By site-directed mutagenesis we joined these two ORFs into one and characterised the resulting protein, carrying both REase and MTase activities [5]. Its REase activity was decreased three times and the optima of the catalytic reaction changed, whereas MTase activity turned out to be intact [5-7]. The bifunctional enzyme could be changed as a result of evolution, leading to further divergence of its properties and functions in the cell. By our hypothesis, this example could serve as a molecular mechanism of new bifunctional RMS origin. In the current work, based on genomic data and their bioinformatics analysis, we aimed to prove that gene fusion could play

an important role in evolution of metyltransferases and in origin of multidomain eukaryotic methyltransferases.

For the current work we searched protein databases as described above and found 76 new potential bifunctional MTases. Their structural organisation is presented and discussed in the report. Beside this, we analysed structural organisation of 627 non-putative prokaryotic DNA-methyltransferases available out of 980 methyltransferases collected up to now in REBASE. The most frequently observed structural type, other than "canonical" MTases, represents SsoII-related methyltransferases, capable to serve as transcriptional autoregulatory proteins. These data provide additional evidences that gene fusion might play an important role in evolution of methyltransferases, restriction-modification systems and other DNA-modifying proteins. We discuss the general consequences of a hypothetical protein fusion event with methyltransferases and RMS enzymes.

1.1. Method used

For this study 10619 methyltransferases and 3250 restriction enzyme sequences were downloaded from the REBASE database [2] and searched against the non-redundant protein database which was downloaded from the NCBI's ftp site. The similarity search was carried out by BLAST version 2.2.23+ [8] on a Linux server using the BLAST default parameters. The local pairwise alignments hits were then filtered using the following criteria: a match to a methyltransferase enzyme was kept if its E-value was less than 1e-140, the sequence identity of the aligned region was greater than 80% and the subject sequence was at least twice longer than the query. By this algorithm we found 272 candidate hits. A match to a restriction enzyme was kept if its E-value was less than 1e-140, the sequence identity within the aligned region was greater than 80% and the length of the subject sequence was at least 1.5 times longer than the query. By this algorithm we found 28 candidate matches. The candidate matches then were manually analysed.

2. Research course

Methyltransferase fusions with a restriction endonuclease

Newly found potential bifunctional restriction and modification enzymes are presented in Table 1. With our BLAST search we succeeded to filter from NCBI database 22 fusions of a DNA methyltransferase with a restriction endonuclease, carrying both endonuclease and methyltransferase domains in one polypeptide. As can be seen from Table 1, the enzymes were grouped according to their domain organisation as it was presented in Conserved Domain Database [9]. As could be judged from their domain organisation, 20 new bifunctional REases are thought to represent the fusion of a REase with MTase and target recognition subunits of the type I restriction-modification systems (R-M-S structure), having similar organisation with the known type IIC bifunctional enzymes such as AloI [10], CjeI [11], MmeI [12], PpiI [13], TstI [13] and TspGWI [14]. Type I RMS enzymes are multisubunit proteins that function as a single protein complex, consisting of R, M and S subunits [3]. The S subu-

nit is the specificity subunit that determines which DNA sequence is recognised. The R subunit is essential for cleavage (restriction) and the M subunit catalyses the methylation reaction. Their protein products are marked as HsdS, HsdR and HsdM, respectively. Covalent linking of these subunits in one polypeptide is not thought to interfere with their catalytic activities, giving an opportunity for successful fusion. Currently, the REBASE contains more than 8000 entries corresponding to Type I RMS. Hypothetically, any of these RMS could be joint, naturally or artificially, giving a new bifunctional RMS.

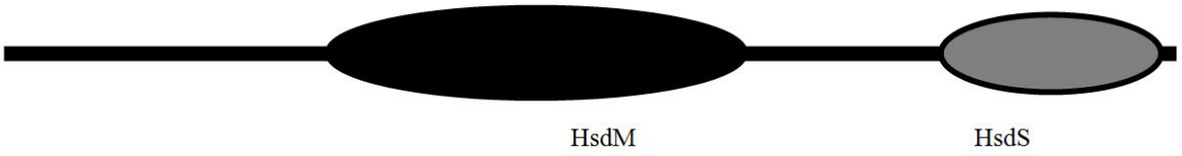
One RMS from *Bacteroides sp*. D22, probably originated from the fusion of type III enzymes. It has conserved motifs similar to Eco57I protein, which consists of Mod and Res subunits of type III enzymes [15]. Type III systems are composed of two genes (mod and res) encoding protein subunits that function in one protein complex either in DNA recognition and modification, Mod, or restriction, Res [3]. As in the case of the type I enzymes, in-frame fusion might not affect their normal functioning, thus, being also favourable for new bifunctional protein origin.

One of the found RMS from *Arthrospira maxima* CS-328 has a different domain organisation, belonging to R-M type of structure (type II REase and MTase fusion). It could be suggested that it originated from fusion of head-to-tail oriented Type II REase and MTase genes. The principal possibility of a new bifunctional RMS origin by this mechanism was proven by in-frame joining of type IIEco29kI REase and MTase genes [5]. The resulting RMS was capable of defending host cells from phage invasion, although 100 times less effectively than the wild type. In a similar way, a new bifunctional RMS could appear from other head-to-tail oriented RMS of type II such as AccI, BanI, Bsp6I, BsuBI, Cfr9I, DdeI, EagI, EcoPI, EcoP15, EcoRI, FnuDI, HaeIII, HgiBI, HgiCI, HgiCII, HgiDI, HgiEI, HgiGI, HhaII, HincII, HindIII, HinfI, HpaI, MboII, MwoI, NcoI, NdeI, NgoMI, NgoPII, NlaIII, PaeR7I, RsrI, SalI, Sau3A, Sau96I, TaqI, TthHB8I, XbaI, and XmaI [16]. Type II restriction enzymes and modification enzymes work separately and their fusion could create steric difficulties for their functionality. In fact, in the case of the RM.Eco29kI enzyme, its REase activity decreased three times in comparison with initial R.Eco29kI nuclease. This is perhaps why it is the only found example of natural bifunctional RMS originating from type II RMS.

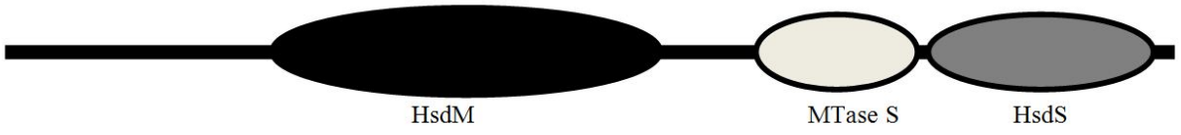
RMS
ZP_01728785.1, 1307 aa, type II restriction-modification enzyme, *Cyanothece sp*. CCY0110]
YP_892534, 1285 aa, restriction and modification enzyme, *Campylobacter fetus subsp. fetus* 82-40.
YP_001482568, 1190 aa, hypothetical protein C8J_0992, *Campylobacter jejuni subsp. Jejuni* 81116.
ZP_00368052, 1343 aa, type I restriction modification enzyme, *Campylobacter coli* RM2228.
ADC28633, 1364 aa, restriction modification enzyme, *Campylobacter jejuni subsp. Jejuni* IA3902.
ZP_01967949,1255 aa, hypothetical protein RUMTOR_01515, *Ruminococcus torques* ATCC
27756.

ZP_06253288, 1297 aa, putative type I restriction modification DNA specificity domain protein, *Prevotella copri* DSM 18205.

ZP_06373928, 1080 aa, hypothetical protein, *Campylobacter jejuni subsp. jejuni* 1336.

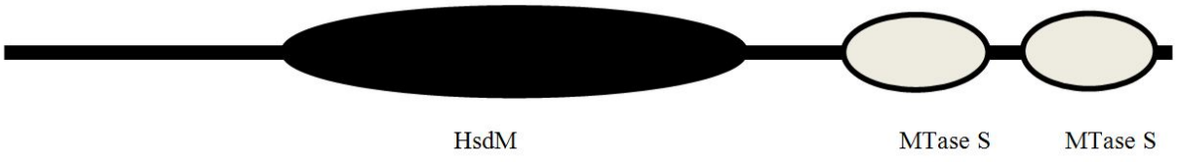


YP_002344446, 1339 aa, restriction modification enzyme, *Campylobacter jejuni subsp. Jejuni* NCTC 11168.



YP_003633850, 1134 aa, restriction modification system DNA specificity domain protein, *Brachyspira murdochii* DSM 12563.

ZP_06409683, 889 aa, putative type II restriction-modification enzyme, *Clostridium hathewayi* DSM 13479.

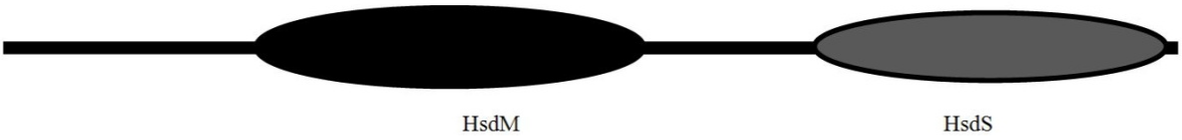


ZP_06998902.1, 1053 aa, type IIS restriction endonuclease, *Bacteroides sp*. D22.

ZP_03658113, 1171 aa, type II restriction-modification enzyme, *Helicobacter cinaedi* CCUG 18818.

ZP_03656081, 1322 aa, restriction modification enzyme, *Helicobacter canadensis* MIT 98-5491.

http://dx.doi.org/10.5772/51025



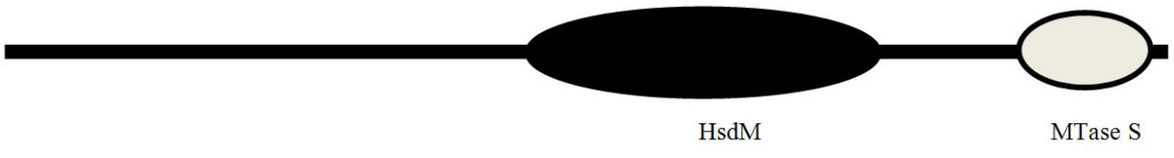
NP_860965, 1164 aa, type I restriction/modification enzyme, *Helicobacter hepaticus* ATCC 51449.

YP_003127864, 1068 aa, N-6 DNA Methylase, *Methanocaldococcus fervens* AG86.

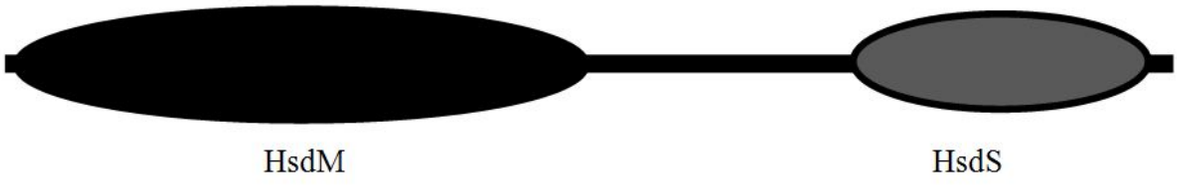
YP_001783948, 1110 aa, N-6 DNA Methylase, *Haemophilus somnus* 2336.

ZP_04582604, 894 aa, type II restriction-modification enzyme, *Helicobacter winghamensis* ATCC BAA-430.

ZP_01253792, 1020 aa, type II restriction-modification enzyme, *Psychroflexus torquis* ATCC 700755.



ZP_07401090, 737 aa, type II restriction-modification enzyme, *Campylobacter coli* JV20.



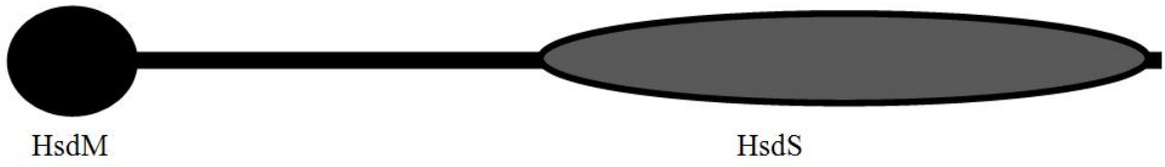
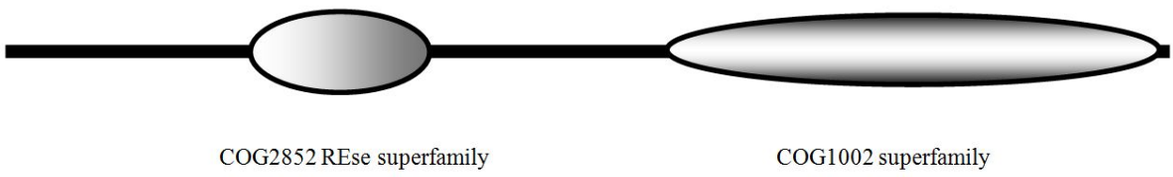
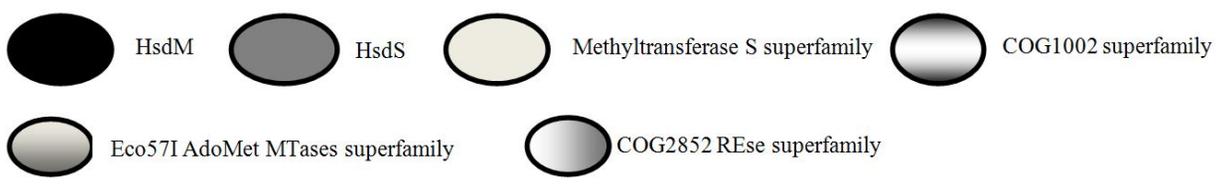
ZP_06009534, 727 aa, restriction and modification enzyme, *Campylobacter fetus subsp. venerealis* str. Azul-94.

RM
ZP_03272741.1, 1473 aa, protein of unknown function DUF559, *Arthrospira maxima* CS-328.



Table 1. Methyltransferase fusions with a restriction endonuclease. Domains are predicted and presented as in Conserved Domain Database [9]. Different domains are shown by different fillings and their classification is shown under the table. Short description for each found protein includes Gene Bank accession number, length in amino acids, current name in the database and host strain information.

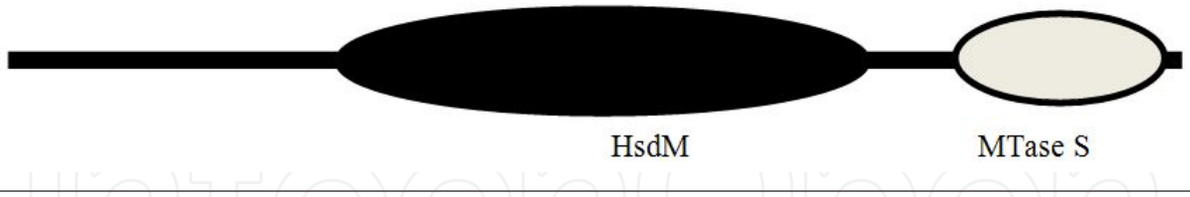
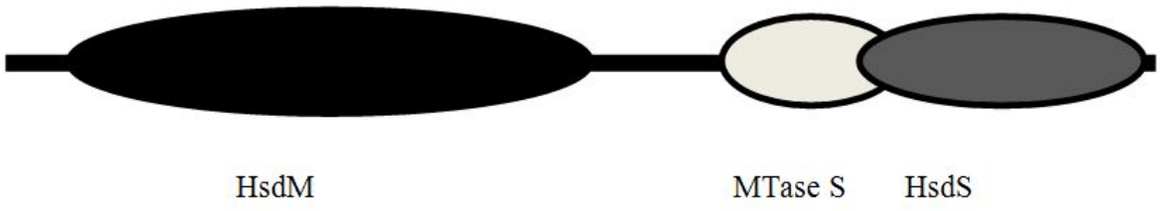
The newly found potential type IIC proteins are good candidates to expand on the current list of 12 bifunctional enzymes: AloI[10], BcgI [17], BseMII [18], BseRI [19], BspLU11III [20], CjeI[11], Eco57I [15], HaeIV [21], MmeI [12], PpiI [13], TstI [13] and TspGWI [14]. Taking into consideration intensiveness with what new microbial genomes have been sequencing during the last decade, new bifunctional RMS could be discovered very soon.

3. Fusions between two DNA methyltransferases

As shown in Table 2, in contrast to the situation with bifunctional MTase - REase fusions, among 54 newly found fusions between two methyltransferases 49, apparently, are joining of two different methyltransferase ORFs (M-M structure) and 5 of methyltransferase (HsdM) and target recognition subunit (HsdS, M-S structure). 11 M-M type proteins represent interesting examples of dcm and dam methyltransferase fusion. In this case dam corresponds not only to one particular MTase, but to a conserved domain common for DNA adenine methyltransferases, as adopted from the Conserved Domain Database web site [9]. In a similar way, dcm corresponds to a conserved domain common for DNA cytosine methyltransferases. These MTases catalyse methyl group transfers to different nucleotide bases, adenine in

the case of dam, and cytosine in the case of dcm. It could be suggested that originally they belonged to two different genes, and were joint in-frame occasionally. The post-segregation killing effect of restriction-modification enzymes prevents RMS from being lost [22], thus promoting maintenance of a fused ORF and its spreading in bacterial populations. The next chapter will be devoted to a more detailed analysis of this effect. The other 23 bifunctional MTases of M-M type probably originated from a similar joining of two dam methyltransferase genes (Table 2). Further evolution of bifunctional MTases depends on their involvement in RMS functioning. If both activities are critical for the RMS work, for example, modifying two different bases of an asymmetric recognition sequence [23], they will be maintained as elements of this RMS. If the activity of at least one MTase domain of the bifunctional enzyme is redundant, it could accumulate mutations and, after many generations, reduce or gain new substrate specificity and function in the cell history.

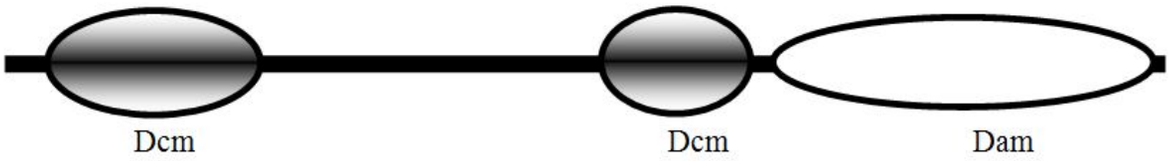
MS
CBL28228.1, 825 aa, Type I restriction-modification system methyltransferase subunit, *Synergistetes bacterium* SGP1.
ZP_04875770, 760 aa, N-6 DNA Methylase family, *Aciduliprofundum boonei* T469.
YP_002961053, 777 aa, hypothetical protein MCJ_005510, *Mycoplasma conjunctivae* HRC/581.

YP_002004068, 785 aa, N-6 DNA methylase, *Candidatus phytoplasma mali*.

ZP_04643292, 907 aa, putative restriction-modification enzyme, *Lactobacillus gasseri* 202-4.

MM
YP_006537.1, 754 aa, DNA methyltransferase Dmt, *Enterobacteria* phage P1.
ZP_07502038, 862 aa, putative DNA methyltransferase, *Escherichia coli* M605.

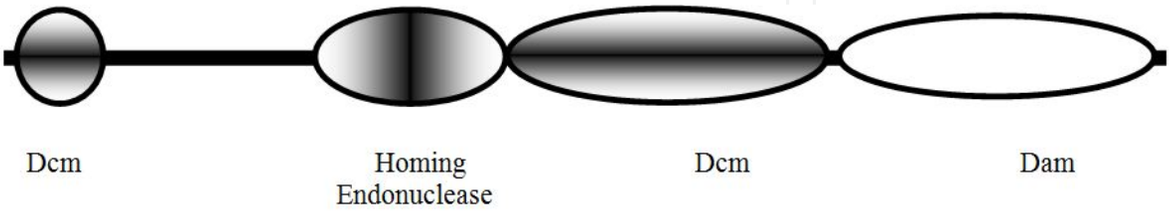
ZP_07381018, 808 aa, DNA adenine methylase, *Pantoea sp*. AB.

ZP_06715601, 722 aa, DNA C5 cytosine methyltransferase, *Edwardsiella tarda* ATCC 23685.

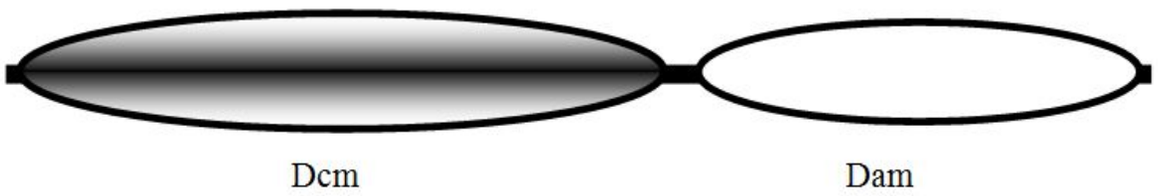
ZP_02701644, 819 aa, DNA methyltransferase Dmt, *Salmonella enterica subsp. enterica* serovar Newport str. SL317.



YP_003237972, 1013 aa putative DNA methyltransferase, *Escherichia coli* O111:H- str. 11128.



AF503408_62, 672 aa, adenine DNA methyltransferase Dam, *Enterobacteria* phage P7.

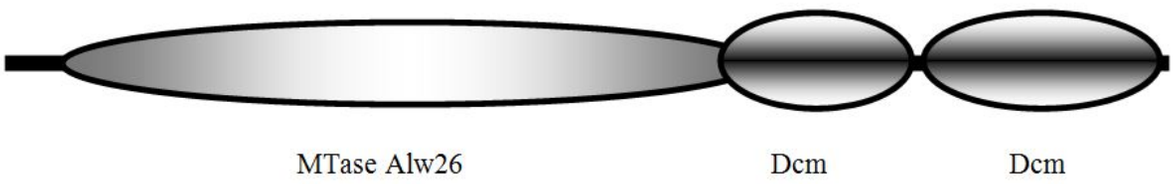


AF458984_2, 952 aa, N6 adenine and C5 cytosine DNA methyltransferase, *Acinetobacter lwoffii*.

AAQ72364, 1007 aa, methylase fusion protein, *Geobacillus stearothermophilus*.

AAS09913, 1068 aa, BsmBI M1-M2 methyltransferase fusion protein, *Geobacillus stearothermophilus.*

AF458983_2, 1061 aa, N6 adenine and C5 cytosine DNA methyltransferase, *Hafnia alvei*.



ACZ62643, 687 aa, BtsCIM Methylase, *Ureibacillus thermosphaericus*.

ZP_05132801, 653 aa, DNA methyltransferase, *Clostridium sp*. 7_2_43FAA.

ZP_02080986, 664 aa, hypothetical protein CLOLEP_02452, *Clostridium leptum* DSM 753.

P29347, 653 aa, Adenine-specific methyltransferase StsI, *Streptococcus sanguinis*.

YP_002744130, 627 aa, DNA adenine Methylase, *Streptococcus equi subsp. zooepidemicus*.

YP_002746827, 620 aa, DNA adenine Methylase, *Streptococcus equi subsp. equi* 4047.

ZP_04008314, 618 aa, adenine DNA methyltransferase, *Lactobacillus salivarius* ATCC 11741.

ZP_00231287, 608 aa, DNA methyltransferase, *Listeria monocytogenes* str. 4b H7858.

ZP_03958097, 625 aa, possible site-specific DNA methyltransferase, *Lactobacillus ruminis* ATCC 25644.

YP_003532979, 617 aa, adenine methyltransferase, *Erwinia amylovora* CFBP1430.

YP_002650601, 617 aa, adenine specific DNA methyltransferase, *Erwinia pyrifoliae* Ep1/96.

ZP_06871587, 670 aa, adenine methyltransferase, *Fusobacterium nucleatum subsp. nucleatum* ATCC 23726.

ZP_04573763, 670 aa, adenine methyltransferase, *Fusobacterium* sp. 7_1.

ZP_06560574, 722 aa, DNA adenine methylase, *Megasphaera genomosp*. type_1 str. 28L.

ZP_06290558, 729 aa, modification methylase, *Peptoniphilus lacrimalis* 315-B.

ZP_04452030, 711 aa, hypothetical protein GCWU000182_01325, *Abiotrophia defectiva* ATCC 49176.

ZP_05861343, 722 aa, putative adenine methylase, *Jonquetella anthropi* E3_33 E1.

ZP_04977178, 722 aa, adenine methyltransferase, *Mannheimia haemolytica* PHL213.

ACZ68466, 710 aa, modification Methylase, *Staphylococcus aureus*.

ADD81212, 700 aa, DNA methyltransferase, *Enterococcus faecalis*.

ZP_07642163, 716 aa, modification Methylase, *Streptococcus mitis* SK597.

ZP_01834649, 716 aa, adenine methyltransferase, *Streptococcus pneumonia* SP23-BS72.

NP_602723, 496 aa, adenine-specific methyltransferase, *Fusobacterium nucleatum subsp. nucleatum* ATCC 25586.

MTase D12 Dam

ZP_05849684.1, 752 aa, twin-argninine leader-binding protein DmsD, *Haemophilus influenzae* NT127.

AAT40808, 518 aa, type III restriction-modification system methyltransferase-like protein, *Haemophilus influenzae*.

ZP_04736658.1, 724 aa, putative type III restriction/modification system modification Methylase, *Neisseria gonorrhoeae* PID332.

ZP_01788156.1, 687 aa, twin-argninine leader-binding protein DmsD, *Haemophilus influenzae* 3655.

YP_001292975.1, 724 aa, twin-argninine leader-binding protein DmsD, *Haemophilus influenzae* PittGG.

YP_001291080.1, 713 aa, putative type III restriction/modification system modification methylase, *Haemophilus influenzae* PittEE.

ZP_01789924.1, 681 aa, twin-argninine leader-binding protein DmsD, *Haemophilus influenzae* PittAA.

ZP_04736658.1, 724 aa, putative type III restriction/modification system modification methylase, *Neisseria gonorrhoeae* PID332.

ZP_04719045.1, 736 aa, putative type III restriction/modification system modification methylase, *Neisseria gonorrhoeae* 35/02.
YP_975326.1, 744 aa, putative type III restriction/modification system modification methylase, *Neisseria meningitidis* FAM18.
ZP_05849684.1, 752 aa, twin-argninine leader-binding protein DmsD, *Haemophilus influenzae* NT127.
ZP_05986443.1, 701 aa, type III restriction/modification enzyme, methylase subunit, *Neisseria lactamica* ATCC 23970.
ABI85522.1, 717 aa, DNA methylase M.Hin1056ModP-2, *Haemophilus influenzae*.
YP_207780.1, 756 aa, putative type III restriction/modification system modification methylase, *Neisseria gonorrhoeae* FA 1090.
ZP_05986443.1, 701 aa, type III restriction/modification enzyme, *Neisseria lactamica* ATCC 23970.

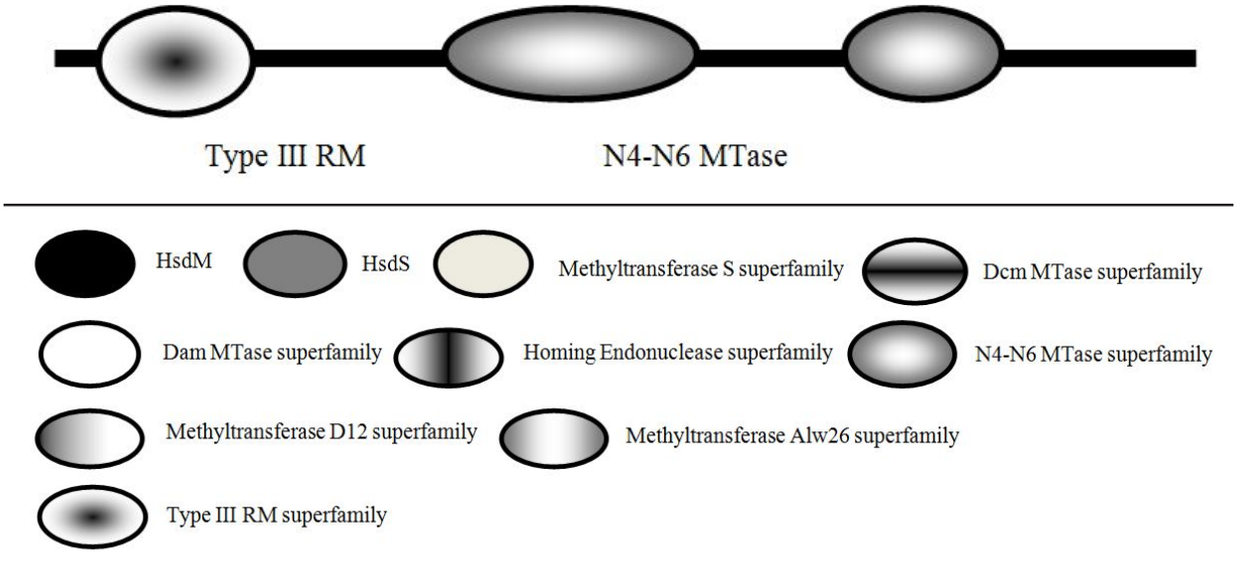


Table 2. Fusions between two DNA methyltransferases. Domains are predicted and presented as in Conserved Domain Database [9]. Different domains are shown by different fillings and their classification is shown under the table. Short description for each found protein includes Gene Bank accession number, length in amino acids, current name in the database and host strain information.

For example, DNMT1, a DNA methyltransferase 1 from *H. sapiens* [24], contains a conserved domain of m5C MTases at its C-terminal domain. The N-terminal and central parts of the enzyme include several different domains such as DMAP binding domain, replication foci domain, zinc finger domain and two bromo adjacent homology domains (Figure 1 a). It could be suggested that several gene fusion events were involved in its evolutionary. This hypothesis is supported by existence of simpler homologs of DNMT1 such as, for example, M.AimAII (Figure 1 b) and M.AimAI (Figure 1 d) from *Ascobolus immersus*, and M.NcrNI from *Neurospora crassa* (Figure 1 c), looking like not completely assembled DNMT1 with one or several domains missing. 15 other M-M type bifunctional MTases from Table 2 represent the fusion of a conserved pfam12564 type III RMS 60 aa domain with pfam01555 N4-N6 MTase domain, characteristic both for N4 cytosine-specific and N6 adenine-specific DNA

http://dx.doi.org/10.5772/51025

methylases. The pfam01555 conserved domain could be found both in type II MTases, such as M.KpnI, and type III Mod proteins, such as EcoP1I and EcoP15I. In contrast, a 60 aa conserved pfam12564 domain could be found only in several type III enzymes. Its addition could influence the biochemical properties of the corresponding proteins. To establish the character of its influence experimentally in the future, it would be necessary to compare the biochemical properties of MTases containing the pfam12564 60 aa sequence, with MTases not containing it.

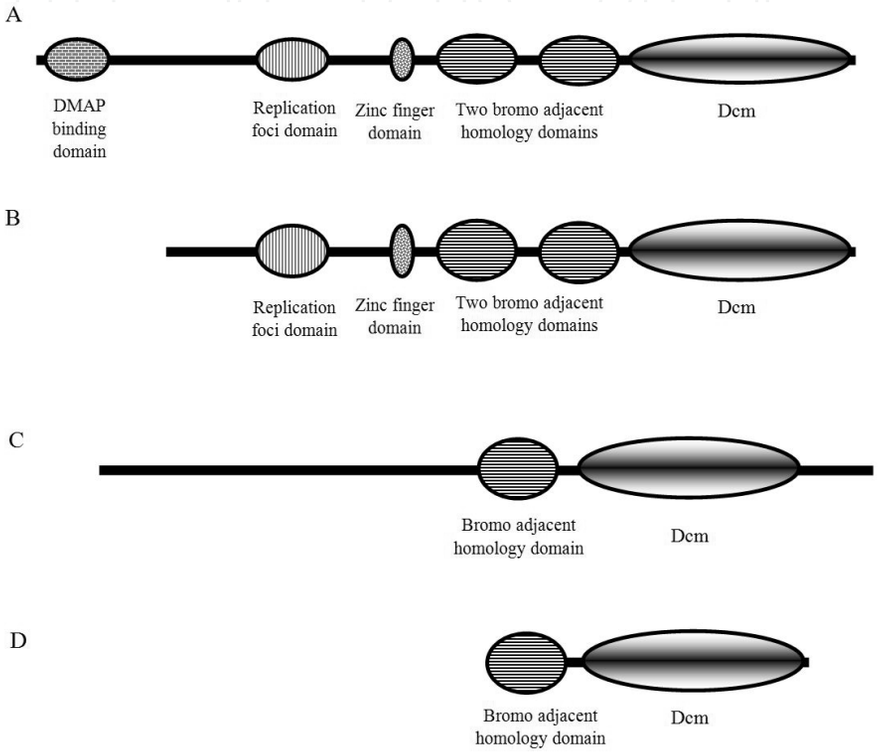


Figure 1. Domain organisation of DNMT1, a DNA methyltransferase 1 from *H. sapiens* (a); M.AimAII, a methyltransferase from *Ascobolus immersus* (b), M.NcrNI from *Neurospora crassa* (c) and M.AimAI from *Ascobolus immersus* (d). Domains are predicted and presented as in Conserved Domain Database [9]. Different domains are shown by different fillings.

4. Fusion of RMS enzymes with a hypothetical protein

Figure 2 shows the consequences of a RMS protein fusion with a hypothetical protein, X. The upper part of the figure represents normal RMS functioning, when host DNA is protected by an MTase and foreign DNA is degraded by a cognate REase. In the case of RMS loss, MTase and REase enzymes will be diluted following cell divisions, host DNA will become unprotected and, finally, degraded by the residual REase activity.

This effect is known as post-segregation killing [22] and is possibly due to REase activity lasting longer than MTase activity after the RMS loss. The lower part of the figure illustrates

the situation of the fusion of some protein X with one of the RMS enzymes. If this joining will not affect seriously the enzyme activities, the RMS will continue to protect cells from foreign DNA invasion. In the case of RMS loss, the same mechanism will lead to host DNA degradation. In this way, a fused protein with RMS enzymes will be supported and spread among bacterial populations. In fact, during a BLAST search we could see close homologues (>98%) of bifunctional enzymes such as, for example, CjeI, spread among numerous strains of the *Campylobacter* group of microorganisms. These observations could illustrate our hypothesis. Another example of the propagation of a RMS-fused protein is provided by M.SsoII-related enzymes. These MTases represent fusion of a regulatory protein with a C5-cytosine methyltransferase [4] and can be found in 63 different microbial taxa.

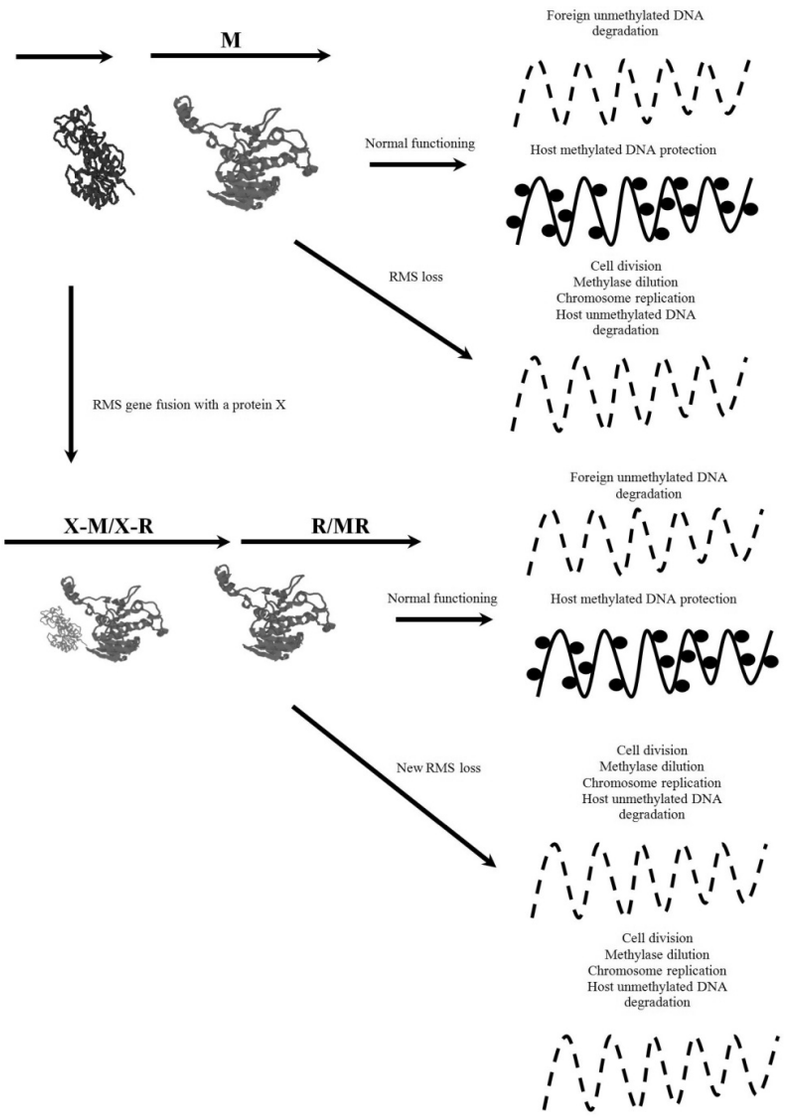


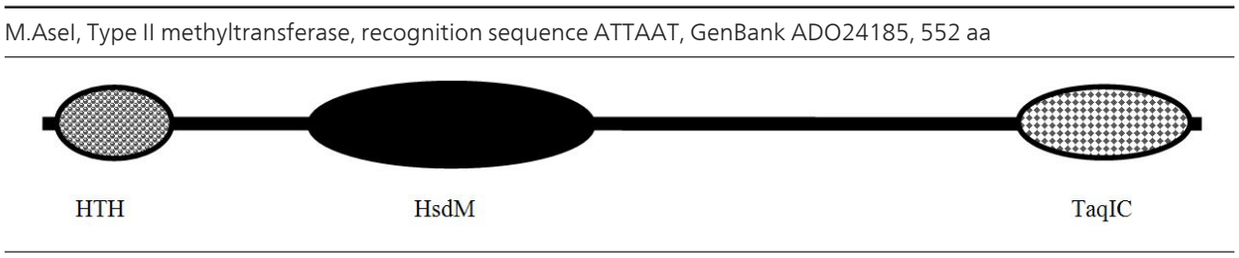
Figure 2. Schematic representation of post segregation killing effect responsible for maintaining RMS and proteins joint with one of its enzymes. The figure summarizes different outcomes of a hypothetical protein X joining to one of a RMS enzymes. If this joining is neutral or positive for the RMS functioning, it will be maintained and spread; if detrimental, it will be eliminated. Black filled circles show methylated nucleotides, interrupted lines - degraded DNA.

http://dx.doi.org/10.5772/51025

In another case, if fusion with a hypothetical protein is detrimental for the activity of one of the RMS enzymes, an outcome will depend on which of the enzymes is affected. If an MTase activity will be reduced, the corresponding RMS will be eliminated due to host DNA degradation by cognate REase. If a REase activity is disturbed, the corresponding RMS will become non-functional, will not be supported by post-segregation killing mechanism and, after many generations, could disappear or take on different functions in the cell. Another, less probable, scenario is possible if a joining with a hypothetical protein would improve the properties of RMS enzymes. In this situation, the corresponding RMS would protect host cells more effectively and that would increase their selective advantage over competitive microbial populations, which, in turn, could lead to a wider distribution of the RMS carrying the fused protein.

5. Domain organisation of non-putative DNA methyltransferases from REBASE

We analyse structural organisation of 627 non-putative prokaryotic DNA-methyltransferases collected up to now in REBASE, a major database of restriction-modification enzymes [2]. We succeeded to download sequences of 627 prokaryotic methyltransferases out of 980 non-putative MTases enlisted in REBASE on 01.12.2011 (for their detailed description see Supplementary materials). Out of 627, 190 sequences belong to dcm type of DNA-methyltransferases; 172 - to N6-N4 type; 99 - to HsdM type and 78 - to dam-related enzymes according to Conserved Domain Database [9]. We found that the most frequently observed structure, other than "canonical" methyltransferases with conserved motives responsible for binding with AdoMet and a methyl group transfer, represents C5-methyltransferase core domain fusion with a regulatory DNA-binding protein and up to now includes 18 potential enzymes (Table 3). These SsoII-related methyltransferases carry additional DNA-binding HTH-domains and they are capable to serve as transcriptional autorepressors [4]. Among these 18 SsoII-like MTases HTH-motif can be located in majority of cases on N-terminal part, and in three proteins - in the middle of their polypeptide chains (M.Esp1396I, M.PflMI and M.SfiI; Table 3). Ability for autoregulation was not confirmed for majority of these SsoII-like MTases and will require some experimental proofs, which could be considered as perspective future directions of research. The fact of SsoII-like enzymes propagation among different bacterial taxa could illustrate well our analysis of a RMS protein fusion with a hypothetical protein, described in the previous chapter.



M.AvaII, Type II methyltransferase, recognition sequence GGWCC, GenBank BA000019, 477 aa
M.BadAII, Type II methyltransferase, recognition sequence CCNGG, GenBank AP009256, 383 aa
M.Csp68KI, Type II methyltransferase, recognition sequence GGWCC, GenBank CP000806, 461 aa
M.Ecl18kI, Type II methyltransferase, recognition sequence CCNGG, GenBank Y16897, 379 aa
M.Eco47II, Type II methyltransferase, recognition sequence GGNCC, GenBank X82105, 417 aa
M.Kpn2kI, Type II methyltransferase, recognition sequence CCNGG, GenBank AF300473, 379 aa
M.MspI, Type II methyltransferase, recognition sequence CCGG, GenBank X14191, 418 aa
M.Sau96I, Type II methyltransferase, recognition sequence GGNCC, GenBank X53096, 430 aa
M1.ScrFI, Type II methyltransferase, recognition sequence CCNGG, GenBank U89998, 389 aa
M.SenPI, Type II methyltransferase, recognition sequence CCNGG, GenBank U65460, 379 aa
M.SinI, Type II methyltransferase, recognition sequence GGWCC, GenBank J03391, 461 aa
M.SsoII, Type II methyltransferase, recognition sequence CCNGG, GenBank M86545, 379 aa
M.StyD4I, Type II methyltransferase, recognition CCNGG, GenBank D73442, 379 aa

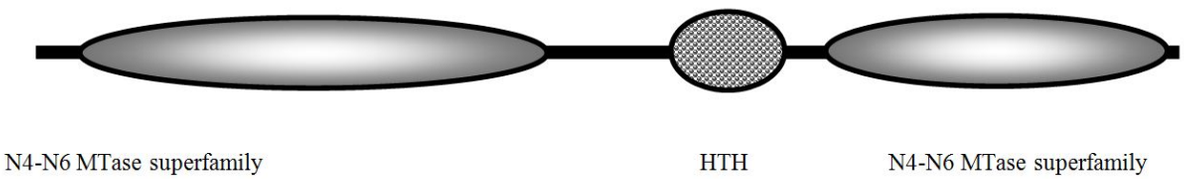
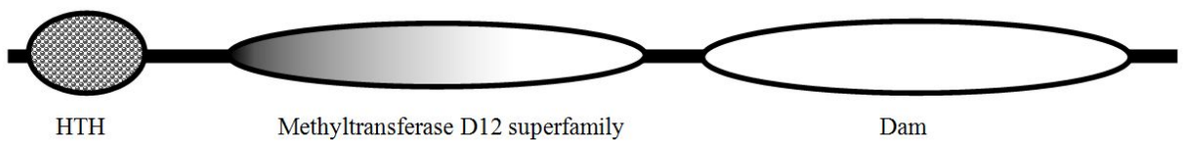
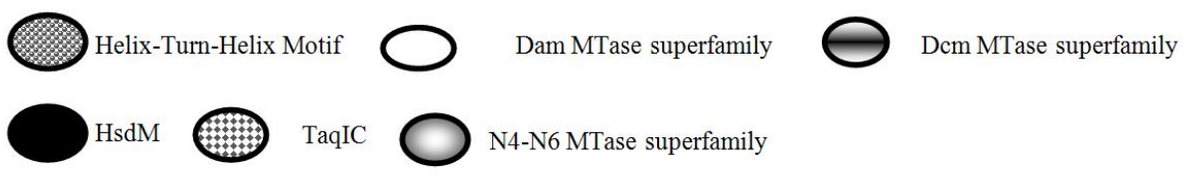
M.Esp1396I, Type II methyltransferase, recognition sequence CCANNNNNTGG, GenBank AF527822, 332 aa
M.PflMI, orphan methyltransferase, recognition sequence CCANNNNNTGG, GenBank ADZ31427, 363 aa
M.SfiI, Type II methyltransferase, recognition sequence GGCCNNNNNGGCC, GenBank AF039750, 421 aa

M1.SfaNI, Type II methyltransferase, recognition sequence GCATC, GenBank GU565605, 700 aa



Table 3. Domain organisation of SsoII-like prokaryotic DNA methyltransferases, collected in REBASE. Domains are predicted and presented as in Conserved Domain Database [9]. Different domains are shown by different fillings.

http://dx.doi.org/10.5772/51025

6. Conclusion

Here we report finding 76 new bifunctional methyltransferases. The majority of the found joint proteins with a nuclease are thought to be fusions of a restriction nuclease with methylase and target recognition subunits of type I restriction-modification systems (R-M-S structure). The majority of the found joint proteins between two methylases appears to be dam-dcm and dam-dam enzyme fusions (M-M structure). Similar proteins could serve as structural intermediates for multidomain eukaryotic methyltransferase evolution. We suggest that a hypothetical protein fusion with a restriction-modification enzyme can promote its propagation in bacterial populations. Altogether, our data illustrate a role of gene fusion in restriction-modification enzyme evolution.

Abbreviations

ORF - open reading frame; REase - restriction endonuclease; MTase - methyltransferase; RMS - restriction-modification system; TRD - target recognition domain; aa - amino acids.

Author details

Dmitry V. Nikitin[1,2*], Attila Kertesz-Farkas[2], Alexander S. Solonin[1] and Marina L. Mokrishcheva[1]

*Address all correspondence to: dvnikitin@rambler.ru

1 Institute of Biochemistry and Physiology of Microorganisms, Russian Academy of Sciences, Russia

2 ICGEB, Area Science Park, Italy

References

[1] Williams, R. (2003). Restriction endonucleases: classification, properties, and applications. *Molecular Biotechnology*, 23, 225-243.

[2] Vincze, T., Posfai, J., & Macelis, D. (2010). REBASE-a database for DNA restriction and modification: enzymes, genes and genomes. *Nucleic Acids Research*, 38, 234-236.

[3] Roberts, R. J., Belfort, M., Bestor, T., et al. (2003). A nomenclature for restriction enzymes, DNA methyltransferases, homing endonucleases and their genes. *Nucleic Acids Research*, 31, 1805-1812.

[4] Karyagina, A., Shilov, I., Tashlitskii, V., Khodoun, M., Vasil'ev, S., Lau, P., & Nikolskaya, I. (1997). Specific binding of SsoII DNA methyltransferase to its promoter region provides the regulation of SsoII restriction-modification gene expression. *Nucleic Acids Research*, 25, 2114-2120.

[5] Mokrishcheva, M., Solonin, A. S., & Nikitin, D. (2011). Fused eco29kIR- and M genes coding for a fully functional hybrid polypeptide as a model of molecular evolution of restriction-modification systems. *BMC Evolutionary Biology*, 11, 35.

[6] Nikitin, D., Mokrishcheva, M., Denjmukhametov, M., Pertzev, A., Zakharova, M., & Solonin, A. (2003). The construction of an overproducing strain, purification and biochemical characterization of the 6His-Eco29kI restriction endonuclease. *Protein Expression and Purification*, 30, 26-31.

[7] Nikitin, D., Mokrishcheva, M., & Solonin, A. (2007). 6His-Eco29kI methyltransferase methylation site and kinetic mechanism characterization. *Biochimica Biophysica Acta*, 1774, 1014-1019.

[8] Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, 215, 403-410.

[9] Marchler-Bauer, A., Lu, S., Anderson, J., et al. (2011). CDD: a Conserved Domain Database for the functional annotation of proteins. *Nucleic Acids Research*, 39, 225-229.

[10] Cesnaviciene, E., Petrusiyte, M., Kazlauskiene, R., Maneliene, Z., Timinskas, A., Lubys, A., & Janulaitis, A. (2001). Characterization of AloI, a Restriction-modification System of a New Type. *Journal of Molecular Biology*, 314, 205-216.

[11] Vitor, J., & Morgan, R. (1995). Two novel restriction endonucleases from Campylobacter jejuni. *Gene*, 157, 109-110.

[12] Morgan, R., Bhatia, T., Lovasco, L., & Davis, T. (2008). MmeI: a minimal Type II restriction-modification system that only modifies one DNA strand for host protection. *Nucleic Acids Research*, 36, 6558-6570.

[13] Jurenaite-Urbanaviciene, S., Serksnaite, J., Kriukiene, E., Giedriene, J., Venclovas, C., & Lubys, A. (2007). Generation of DNA cleavage specificities of type II restriction endonucleases by reassortment of target recognition domains. *Proceedings of National Academy of Sciences USA*, 104, 10358-10463.

[14] Zylicz-Stachula, A., Bujnicki, J., & Skowron, P. M. (2009). Cloning and analysis of a bifunctional methyltransferase/restriction endonuclease TspGWI, the prototype of a Thermus sp. enzyme family. *BMC Molecular Biology*, 10, 52.

[15] Rimseliene, R., Maneliene, Z., Lubys, A., & Janulaitis, A. (2003). Engineering of restriction endonucleases: using methylation activity of the bifunctional endonuclease Eco57I to select the mutant with a novel sequence specificity. *Journal of Molecular Biology*, 327, 383-391.

[16] Wilson, G. G. (1991). Organization of restriction-modification systems. *Nucleic Acids Research*, 19, 2539-2566.

[17] Kong, H., Roemer, S., Waite-Rees, P., Benner, J., Wilson, G., & Nwankwo, D. (1994). Characterization of BcgI, a New Kind of Restriction-Modification System. *Journal of Biological Chemistry*, 269, 683-690.

[18] Jurenaite-Urbanaviciene, S., Kazlauskiene, R., Urbelyte, V., Maneliene, Z., Petrusyte, M., Lubys, A., & Janulaitis, A. (2001). Characterization of BseMII, a new type IV restriction-modification system, which recognizes the pentanucleotide sequence 5′-CTCAG(N)(10/8)/. *Nucleic Acids Research*, 29, 895-903.

[19] Mushtaq, R., Naeem, S., Sohail, A., & Riazuddin, S. (1993). BseRI a novel restriction endonuclease from a Bacillus species which recognizes the sequence 5′...GAGGAG... 3′. *Nucleic Acids Research*, 21, 3585.

[20] Lepikhov, K., Tchernov, A., Zheleznaja, L., Matvienko, N., Walter, J., & Trautner, T. A. (2001). Characterization of the type IV restriction modification system BspLU11III from Bacillus sp. LU11. *Nucleic Acids Research*, 29, 4691-4698.

[21] Piekarowicz, A., Golaszewska, M., Sunday, A., Siwińska, M., & Stein, D. (1999). The HaeIV restriction modification system of Haemophilus aegyptius is encoded by a single polypeptide. *Journal of Molecular Biology*, 293, 1055-1065.

[22] Kobayashi, I., Nobusato, A., Kobayashi-Takahashi, N., & Uchiyama, I. (1999). Shaping the genome- restriction-modification systems as mobile genetic elements. *Current Opinion in Genetics and Development*, 9, 649-656.

[23] Madhusoodanan, U. K., & Rao, D. N. (2010). Diversity of DNA methyltransferases that recognize asymmetric target sequences. *Critical Reviews in Biochemistry and Molecular Biology*, 45, 125-145.

[24] Yen, R., Vertino, P., Nelkin, B., et al. (1992). Isolation and characterization of the cDNA encoding human DNA methyltransferase. *Nucleic Acids Research*, 20, 2287-2291.

IntechOpen