Manuscript III

Manuscript

Initial characterization of ParI, an orphan C5-DNA methyltransferase from *Psychrobacter arcticus* 273-4

Miriam Grgic[1], Gro Elin Kjæreng Bjerga[2], Adele Kim Williamson[1], Bjørn Altermark[1] and Ingar Leiros[1]

[1]Department of chemistry, University of Tromsø, The Arctic University of Norway, Tromsø, Norway

[2] University of Bergen, Bergen, Norway

Abstract

Background

DNA methyltransferase are important enzymes in most living organisms, from single cell bacteria to eukaryotes. In prokaryotes, most DNA methyltransferases are members of the bacterial type II restriction modification system: The DNA methyltransferases methylates host DNA, thereby protecting it from cleavage by the restriction endonucleases that recognize the same sequence. In addition to being part of restriction modification systems, DNA methyltransferases can act as orphan enzymes having important role in controlling gene expression, DNA replication, cell cycle and DNA post replicative mismatch repair. In higher eukaryotes, DNA methylation is involved in the regulation of gene expression, maintenance of genome integrity, parental imprinting, chromatin condensation as well as being involved in development of diseases and disorders.

DNA-MTases have several potential applications in biotechnology for example in targeted gene silencing and in DNA-protein interaction studies

Results

According to bioinformatical analyses, the *par*I gene from *Psychrobacter arcticus* 273-4, encoding a cytosine specific DNA methyltransferase, is likely of phage origin. In this work, ParI was expressed and purified in a McrBC negative *E. coli* strain in fusion with a hexahistidine affinity tag and maltose binding protein solubility protein. The obtained recombinant ParI was monomeric with a molecular mass estimated to 54 kDa. The melting temperature of the protein was determined to be 53 ºC in ThermoFlour assay and 54 ºC in differential scanning calorimetry, with no secondary structures detectable at 65 ºC. Recombinant ParI showed activity in two *in vivo* assays.

Conclusions

In this work, the first recombinant expression of an orphan MTase from a psychrophilic bacterium, *P. arcticus* 273-4 was done. Despite originating from psychrophilic organism strong cold adapted features, in terms of low temperature activity and heat lability, could not be detected in ParI and due to that some additional investigations should be done.

Keywords
DNA methylation, DNA methyltransferase; ParI; *Psychrobacter arcticus* 273-4*;* cold-adaptation.

Background

The methylated DNA bases, N4-methylcytosine, N6-methyladenine and C5-methylcytosine, are considered additional bases of the genetic code that carry epigenetic information not encoded in the DNA sequence itself (reviewed in (1)).

Generally, DNA methylation has an important role in protein-DNA interaction by either enhancing or disrupting binding of proteins to DNA. In eukaryotes, DNA methylation is involved in several vital cellular processes involving gene expression, such as parental imprinting and cellular differentiation and development (1-4). Additionally, changes in methylation patterns are closely related to the development of various diseases (5-7). The enzymes responsible for these modifications are DNA methyltransferases (DNA-MTases). These enzymes are catalyzing the transfer of a methyl group from S-adenosyl-L-methionine (SAM) to DNA bases. In contrast to the eukaryotic DNA-MTases, most of the prokaryotic DNA-MTases are part of a host protection system as members of the restriction modification (RM) systems (8, 9). The main function of DNA-MTases in the RM system is methylation of host DNA and thereby conferring protection from cleavage by restriction endonucleases (REase) that recognize the same specific DNA sequence (10). Some DNA-MTases that are not accompanied by REase, so-called orphan DNA-MTases that have roles in controlling gene expression, DNA replication, cell cycle and directing post-replicative mismatch repair on newly synthesized DNA strands by *de novo* methylation. These latter activities are mostly coordinated by the N6-DNA-MTases, most notably Dam and CcrM (11-14). Additionally, methylation of DNA by Dam have an important role in DNA replication (11).

The C5-DNA-MTases modifies carbon 5 on cytosine (15).Whereas in eukaryotes C5-DNA-MTases are part of multi-domain proteins and complexes, in prokaryotes methylation of DNA is carried out by the single-domain proteins (16). When three-

dimensional protein structures of C5-DNA-MTases are compared, both prokaryotic and eukaryotic enzymes are similar in structural organization. All are two domain-proteins, organized into a large and a small domain with a marked cleft in between the domains (17). Despite their structural similarity, the sequence similarity among C5-DNA-MTases is low with the exception of ten conserved motifs. In addition to these motifs, all C5-DNA-MTases possess a variable region, the target recognition domain (TRD), which is involved in sequence recognition. Most of the motifs are located in the larger domain, while the TRD comprises most of the small domain. These motifs are involved either in SAM binding (motifs I and X, with additional help from motifs II, III, V), DNA binding (motifs VI and VIII and TRD) or catalysis (motif IV) (17-20).

Psychrobacter arcticus 273-4 is a gram-negative bacteria discovered in 20000-30000 years old permafrost soil at Kolyma (Siberia). It can grow at temperatures from -10 to 28 ºC, with a generation time of 3.5 days at temperatures below zero (-2.5 ºC) (21, 22). On the protein level, *P. arcticus* possesses many typical features for cold-adapted bacteria, such as less hydrophobic, aliphatic and acidic amino acids, fewer Pro residues, and low Arg and increased Lys contents (21), as well as several cold shock proteins (21-23).

In this study a characterization of an orphan C5-DNA-MTase ParI, from psychrophilic bacterium *Psychrobacter arcticus* was done. Enzyme was characterized in terms of thermal stability and activity.

Results and discussion

Bioinformatical analysis of the *parI* gene

To acquire additional information about ParI, bioinformatics investigations were performed. By searching through REBASE (http://rebase.neb.com/rebase/rebase.html) (9), several putative enzymes and partners in the RM systems were identified in *P. arcticus* 273-4 (Table 2). *P. arcticus* possess one complete type I RM system (uncharacterized) as well as one solitary type II DpnI-like REase and one solitary C5-DNA MTase (ParI) (Table 1). The counterparts for functional type II RM systems were not identified for the orphan enzymes. In addition, *P. arcticus* also possess an incomplete type IV McrBC-like REase, where the C subunit of the canonical two subunits (B+C) complex is missing which makes

the McrB non-functional (24, 25). An evolutionary explanation for this may lie in the fact that the substrate for a functional McrBC enzyme is a methylated cytosine, and as *P. arcticus* contains a putatively functional ParI enzyme, an active McrBC would lead to self-DNA restriction and eventually cell death. Evolutionary, the ParI could have been part of a functional type II RM system that over time has lost its REase member as it became redundant (26). We speculate that ParI may have survived the evolutionary pressure because it methylates at the same recognition site as members of other type II RM system found in *P. arcticus,* thereby complementing methylation and protection of host DNA to make it more effective. A similar occurrence can be found in *E. coli* RM systems, where Dcm, an orphan C5-DNA MTase have the same recognition sequence as the EcoRII RM system (CCWGG) and one of the hypothesis that explains this is that Dcm serves as back up for methylation of *E. coli* DNA (26, 27).

Table 1. *P. arcticus* RM systems based on data from the REBASE database.

Type	Protein type[1]	Name
	M	M.ParAORF415AP
	M	M.ParAORF415BP
I	S	S.ParAORF415P
	R	ParAORF415P
	R	ParAORF723P
	M	M.ParAORF877P
	S	S.ParAORF877P
	R	ParAORF877P
	R	ParAORF142P[2]
II	M	M.ParAORF401P
	R	ParAORF401P
	M	M.ParAORF980P[3]
IV	R	ParAMcrBP

[1] R (restriction), M (modification), and S (specificity)

[2] DpnI-like (GATC specificity)

[3] ParI

By looking into genomes available in the REBASE database, we noticed that the *par*I gene, is part of an operon containing four genes, from gene 0981 to gene 0978, which are all designated as hypothetical proteins (Figure 1). This gene arrangement is further supported by the prediction of a promoter in front of gene 0981 and two Rho-independent terminators after gene 0978 (Figure 1).

In order to detect if the *par*I gene or surrounding genes are of phage origin we performed an analysis by the PHAST program (28). If the analyzed region contains only genes/proteins of a known phage, the region automatically has a completeness score of 150 (the maximum) and is considered intact. The strong sequence similarity of gene 0979 to a phage protein indicates that these genes may have been horizontally transferred and are of phage origin, although they are no longer part of an intact temperate phage. The genome of *P. arcticus* contains three prophage regions, including one complete temperate phage and two regions where phage genes are clustered (Table 3).

Table 2. Identified prophage regions in the *P. arcticus* 273-4 genome.

Region number	Length (Kb)	Completeness	Score[2]	#CDS[1]	Region position	Possible phage	%GC
1	42.4	intact	150	40	542631-585093	Escherichia phage HK75	44.79%
2	9	questionable	70	10	756847-765902	Prophage Xanthomonas axonopodis pv. citri str. 306	37.62%
3	12.3	incomplete	50	16	1191891-1204227	Burkholderia phage KS10	43.88%

[1] The number of coding sequences

[2] Criteria for scoring prophage regions (as intact, questionable, or incomplete) is determined by using three methods where number of certain phage organism is compared to total number of CDS of the region and based on that it is designated as intact, incomplete or questionable**.**

In addition, gene 0979 shows strong similarity to a phage protein, indicating that these genes have been horizontally transferred and are of phage origin, although they are no longer part of an intact temperate phage (table 3).

Table 3. Function of the genes surrounding *par*I.

Gene/designation	COG	Pfam	Function
0978			Hypothetical protein (unique)
0979			Hypothetical protein (phage related, Psymv2, orf4)
0980/*par*I	0270L		C-5 cytosine specific DNA methyl transferase
0981/*rec*T			Similar to RecT
0982	5377L		Phage related nuclease (recombination endonuclease)
0983			Hypothetical protein
0984			Hypothetical protein
0985	3620K		DNA binding protein (helix-turn-helix)

Structural and sequence homology of ParI with other methyltransferases

By sequence alignment of the ParI protein sequence with other C5-DNA-MTases sequences, all ten conserved motifs specific for these enzymes were identified (Figure 2). In motif IV (PCQ, catalytic motif), which is generally strongly conserved among C5-DNA-MTases, ParI contains a Asp in place of the conserved Pro (17). Generally, the Pro in motif IV has a role in orientation of the activated cytosine and SAM so that the methyl group transfer can be achieved (29). BLAST searches revealed that the ParI has highest sequence similarity (64%) to an undescribed C5-DNA-MTase from *Acinetobacter baumannii* including also Pro to Asn substitutionin motif IV. These both bacteria belong to same family, *Moraxellaceae,* which can explain this higher similarity (30).

Cloning, expression and purification of ParI

Initially, the *par*I gene was cloned and expressed as a His tagged protein from pDest17 vector (pHParI). Early experiments showed that solubility of this recombinant ParI protein was low (data not shown). To increase yield and solubility the *par*I gene was sub cloned in fusion to maltose binding protein (MBP) in the pHMGWH vector, generating the pHMParI construct. The MBP solubility tag is known to improve solubility on difficult-to-express proteins (31). The pHMParI construct was transformed to the McrBC-negative *E. coli* T7 Express strain for recombinant expression and purification. The *mcr* loci have been shown to reduce recovery of methylated sequences from an organism with methylcytosine in cloning experiments (32). Due to ParI being a putative C5-DNA methyltransferase, the DNA methylated by ParI would be a substrate for the McrBC gene product and would

therefore get restricted. For protein production, a 7 hrs expression time at 20 ºC was more successful than a 16-20 hrs expression time at the same temperature (data not shown). In the latter condition, no protein could be detected during purification. Based on SDS-PAGE analysis of protein integrity after lysis, we concluded that soluble protein of the recombinant His-MBP-tagged ParI was obtained after a 7 hrs expression (Figure 3A). The His-MBP fusion partner was successfully removed by tobacco etch virus (TEV) protease treatment (Figure 3B). There was a substantial loss of protein in TEV protease digestion- and subsequent purification steps, indicating instability. Typically, we obtained 2.5 mg of pure, recombinant ParI per liter of LB expression culture. The gel filtration confirmed the theoretical protein mass and its monomeric state (Figure 4). No DNA contamination in the sample of purified protein was observed.

The melting temperature of ParI was found to be 53 ºC

As initial experiments showed that there was a substantial loss of protein during purification, we attempted to optimize the protein stability. ThermoFluor assays were conducted to assess the general stability of the protein. The melting temperature of the protein was determined to be 53 ºC in HEPES buffer at pH 8 (Figure 5). In this experimental condition the melting curve for ParI was gradual and stretches over a long temperature range which are typically observed in less stable proteins (33).

In DSC experiments ParI unfolded in a single unfolding event at 54 ºC

To further investigate the thermal stability of the protein, Differential Scanning Calorimetry (DSC) measurements using a N-DSC III (Calorimetry sciences corp.) calorimeter were performed. The DSC measurement indicated that ParI unfolds in a single unfolding event at a temperature of 54 ºC with ΔH being 82.81 kcal/mol and ΔS 0.2411 kcal/ (molK) (Figure 6). The melting temperature found by DSC is in agreement with the ThermoFluor measurements (Figure 5). In addition, we observed that ParI unfolds at 49-59 ºC (Figure 6), which is in agreement with our ThermoFluor results. The thermal denaturation of ParI was irreversible and after the run, aggregation of the precipitated protein was observed in the calorimeter cells.

Secondary structure analysis of ParI by circular dichroism analysis

To investigate the composition of the secondary structure of ParI, and to determine if it is in agreement with predicted secondary structure (Figure 2) we preformed circular dichroism (CD) analysis. Based on the obtained results, the content of secondary structures in ParI was calculated. All secondary structure elements (α-helix, β-sheet and loops) have characteristic CD spectra and based on those the secondary structure of the protein can be calculated (34). The spectrum for an α-helix protein have two negative bands of similar magnitude at 222 and 208 nm, and a positive band at ~ 190 nm. The spectrum for a β-sheet protein has a negative band between 210-220 nm and a positive band between 195 - 200 nm, while the spectrum for a loop has a negative band of great magnitude at around 200 nm. According to CD-calculation ParI contains 20% α-helices 30% β-strands and 12% turns in measurements at 15 ºC and 35 ºC (Figure 7). Furthermore, there are 35-40% disorder regions or loops present in the structure (Figure 6). This supports the secondary structure as predicted by PsiPred (Figure 1). When the protein was analyzed at 65 ºC, no secondary structure could be detected. These results correlate with our ThermoFluor and DSC measurements of the recombinant ParI protein in which the melting temperatures were found to be, 53 ºC and 54 ºC, respectively. When SAM was added to the protein no change in secondary structures and stability could be observed in CD measurements at 15 ºC indicating that binding of SAM does not cause major structural changes to the recombinant protein.

DNA from ParI-expressing *E. coli* cells was restricted by McrBC *in vitro*

To reveal whether ParI was active when recombinantly produced in *E. coli*, we used an *in vivo* assay where we screened for McrBC sensitivity on genomic DNA in the presence or absence of ParI. The two expression cultures of the MTase-deficient *E. coli* T7 Express were induced, one transformed with the pHParI expression construct (recovers the MTase activity through ParI) and the other transformed with the empty vector (ParI-deficient). Subsequently, genomic DNA was extracted from cultivated cells. The genomic DNA was then used as substrates in a McrBC restriction assay where DNA containing methyl cytosines in one or both DNA strands would be restricted. In this assay, we found that

genomic DNA from the ParI-expressing *E. coli* cells was restricted by McrBC, whereas genomic DNA from the ParI-deficient *E. coli* cells was not restricted (Figure 8).

ParI is able to fully or partially methylate the *P. arcticus* genomic DNA

In another assay, we investigated the McrBC sensitivity on genomic DNA from the native ParI host, *P. arcticus,* and compared it to the C5-DNA-MTase deficient homologue *P.* sp. PRwf-1. After treatment with McrBC enzyme, genomic DNA isolated from *P.* sp. PRwf-1 was intact while a smear could be observed in the lane with genomic DNA from *P. arcticus* (Figure 8, right panel). Our results indicate that genomic DNA from *P. arcticus* was methylated *in vivo*. According to REBASE, *P. arcticus* contains another C5-DNA MTase, in addition to ParI, encoded by gene *par*401 (Table 1). It is possible that this enzyme is responsible for the methylation we observe. In light of the results from recombinant ParI (Figure 9, left panel) we can assume that ParI is able to fully or partially methylate the *P. arcticus* genomic DNA, thus, making it sensitive to McrBC.

In yet another *in vivo* experiment, two types of *E. coli* expression strains were transformed with the pHParI construct. One of the strains used was the BL21 Star (DE3), which eliminates methylated DNA due to the presence of a functional McrBC system, and the other strain was the McrBC-deficient *E. coli* T7 Express. Both transformants were plated on LB plates containing 100 μg/ml ampicillin and 0.2 mM IPTG for induction of recombinant ParI expression and incubated overnight at 37 ºC. On the plate containing BL21 Star (DE3) cells, no colonies could be observed, while on the plate containing T7 expression *E. coli* colonies were observed (Table 1). The explanations for the lack of growth of BL21 Star (DE3) cells is likely that the genomic DNA has been methylated by ParI, which leads to restriction by the host's McrBC system and ultimately cell death.

Table 3. *In vivo* methylation and restriction assay.

	Number of colonies on plate	
Strain	Experiment 1	Experiment 2
BL21Star(DE3)	0	0
McrBC-negative *E. coli* T7 Express strain	460	580

Bisulfite conversion

In an attempt to elucidate the specificity of the ParI enzyme, we performed an *in vitro* bisulfite conversion assays that would provide information about the recognition site. Briefly, in these assays, methylated cytosines can be separated from non-methylated cysteins as they are converted to thymines due to the chemical reaction of bisulfite conversion. Before proceeding with bisulfite conversion, the DNA that was utilized as substrate was incubated with ParI, for 16-20hrs. After incubation, bisulfite conversion was performed and the DNA was subsequently amplified by PCR and cloned into a cloning vector. The vector was transformed into *E. coli* DH5α, and plasmid was isolated from clones and sequenced. We did this in three separate experiments using different batches of freshly purified ParI. In total, we analyzed twelve sequences. Out of these, four showed a full methylation on all cytosines, indicating that the ParI enzyme cannot discriminate between different recognition sites, which might be caused by non-optimal assay conditions. These samples were considered these false positives. In the remaining eight samples, the methylation pattern varied between a consistent methylation in certain positions and only partial methylation at other positions (Figure 9B). Based on these data, we could not find any consistent pattern that could point towards the exact recognition site specific to ParI. A more frequent methylation could be observed when methylated cytosine was surrounded by adenine or thymine. These results are consistent with our findings in the *in vivo* assays where ParI-methylated genomic DNA was sensitive to McrBC which digests at (G/A)^{m}C sites (Figure 9). In addition, conversion data showed that most methylated cytosines were part of the (T/G)C dinucleotide. To conclude, the exact methylation pattern could not be elucidated from our data, but the motif is likely to include (T/G)C, in an A-rich surrounding.

Conclusions

In this paper, the C5 DNA-MTase, ParI, from *P. arcticus* 273-4 was partly characterized. To our knowledge, this is the first characterization of an orphan 5-DNA-MTase from an psychrophilic bacterium. The C5-DNA-MTase could not be expressed in standard *E. coli* expression strains due to the existence of McrBC enzymes that were specific against the

methylated cytosines. A successful expression was obtained by using a McrBC-negative *E. coli* expression strain. Initially, the gene encoding the ParI protein was cloned in fusion to a His tag. Expression trials showed that solubility of this recombinant ParI protein was low (data not shown). To improve solubility, we cloned ParI in fusion with an MBP solubility tag. The change of construct contributed to an improvement in protein yield, solubility and stability of the protein. Using two different *in vivo* assays, we showed that the protein was methylating bacterial genomic DNA. Based on biochemical characterization, recombinant ParI has an unfolding temperature at 53 and 54 ºC, which was confirmed with ThermoFluor and DSC experiments respectively. In addition, this result was confirmed by CD data that showed no secondary structures at 65 ºC. Since *P. arcticus* is a psychrophilic bacterium it was assumed that ParI would also exhibit cold adaptation features, such as higher activity at lower temperature and heat lability, but since we did not compare our results to a native enzyme or enzymes from a mesophilic homologue this question remains unresolved.

Based on several methylation assays, in which the ParI-methylated genomic DNA was treated with the commercially available McrBC, activity of ParI was detected *in vivo*. In order to elucidate the specific sequence that ParI recognize and methylate, a bisulfite assay was performed. These data were inconclusive, but we found indicationgs that a (T/G)C dinucleotide in an adenine rich surrounding is making up part of the DNA recognition sequence that is methylated by ParI. To determine the exact site that ParI is methylating, more experiments are required.

Competing interests

There is no conflict of interest between the authors.

Authors' contributions

IL, BA and GEKB conceived the study and participated in its design and coordination, helped in data analysis and reviewed the manuscript, MG performed all experiments, interpreted the data and wrote the manuscript; AKW helped in designing and performed the CD experiment. All authors read and approved the final manuscript.

Acknowledgements

This study was supported by the University of Tromsø, the arctic university of Norway and Norwegian research council as part of the MARZymes initaitve.

Methods

Bioinformatical analysis of the *parI* gene

Sequence similarity search was performed using the BLAST program (35) provided by the NCBI (http://blast.ncbi.nlm.nih.gov/Blast.cgi) A multiple sequence alignment where ParI was compared to bacterial C5 DNA-MTases was made using the T-coffee web server (36). The alignment was rendered using the ESPRIPT server (37). Secondary structure predications were analysed by PsiPred (38).

In order to determine whether ParI was of phage origin we analysed the genomic DNA from *P. arcticus* 273-4 by the PHAST server (28).

Cloning, expression and purification

The gene encoding ParI, *parI* [GenBank ID: 71038525] was amplified from *Psychrobacter arcticus* 273-4 (DSMZ) with primers synthesized by Sigma-Aldrich (Table 3) and Phusion polymerase (NEB) in a PCR reaction following protocol for Phusion polymerase. Amplified gene was cloned into the pDest17 vector containing TEV protease-removable N-terminal hexahistidine (His) tag (resulting in pHParI) and pHMGWA vector with the downstream TEV protease-removable N-terminal hexahistidine (His) tag and maltose binding protein (MBP). Construct is here termed pHMParI. Both constructs were cloned using the Gateway technology (Life Technologies) following protocol provided by manufacturer (39) and the primers listed in Table 4. The pHParI construct was used for assay experiments while pHMParI was used for protein production. Newly cloned constructs were transformed by heat-shock method, into McrBC-negative *E. coli* T7 Express strain (NEB) for recombinant expression and purification. These cells lack the McrBC/Mrr restriction endonucleases and thus do not restrict methylated DNA (genotype *fhuA2 lacZ::T7 gene1 [lon] ompT gal sulA11 R(mcr-73::miniTn10*--TetS*)2 [dcm] R(zgb-210::Tn10*--TetS*) endA1 Δ(mcrC- mrr)114::IS10*).The protein was expressed in 1L cultures of Lysogeny broth (LB) media and induced with isopropyl-β-D-1-thiogalactopyranoside (IPTG, VWR) at a final concentration of 0.5 mM. After 7 h expression at 20 ºC, the cells were harvested by centrifugation at 6000 rpm for 30 min at RT and resuspended in lysis buffer (50 mM Tris pH 8.0, 750 mM NaCl, 5 % (v/v) glycerol,

10 mM $MgCl_2$) supplemented with 1x complete protease inhibitor cocktail (Roche). The cells were disrupted by a cell disruptor (Constant Systems, Ltd.) using a pressure of 1.35 kbar in four cycles. The lysate was cleared by centrifugation at 20000x*g* for 30 min at 4 ºC after which it was incubated with HL/SAN DNase (ArticZymes) for 1h at 4ºC. Affinity purification of recombinant ParI was carried out on a 5 ml HisTrap HP column (GE Healthcare) equilibrated with buffer A (50 mM Tris pH 8.0, 750 mM NaCl, 5 % (v/v) glycerol and 10 mM imidazole) using an ÄKTA purifier (GE healthcare). The bound protein was eluted across a gradient of 0–100 % buffer B (50 mM Tris pH 8.0, 750 mM NaCl, 5 % v/v glycerol and 500 mM imidazole). The purity of the protein was evaluated on SDS-PAGE. Separation of the His-MBP fusion partners from the recombinant ParI protein was done in the buffer C (50 mM Tris pH 8.0, 200 mM NaCl, 5 % (v/v) glycerol, 1 mM DTT and 0.5 mM EDTA) supplemented with TEV protease in a 1:10 ratio to ParI and incubated overnight at 4˚C. Buffer exchange was done using a HiPrep 26/60 desalting column (GE Healthcare). The recombinant ParI protein was released during a second step of HisTrap affinity purification. To determine the oligomeric state of the protein, the protein was separated on a HiLoad 16/60 Superdex 200 prep grade gel filtration chromatography column (GE Healthcare) in buffer C. The protein concentration was determined by measuring the absorbance at 280nm using a Nanodrop spectrophotometer (NanoDrop Technologies). The theoretical extinction coefficient for the protein is 63995 $M^{-1}cm^{-1}$ as calculated by the ProtParam tool (Expasy web page, http://web.expasy.org/protparam/). Pure protein was concentrated using 10MWCO Amicon Ultra Centrifugal filters (Merck Millipore). Identity of recombinant protein was determined by MS/MS by the Proteomics facility, University of Tromsø.

Table 4. List of PCR primers used for cloning of *par*I.

Primer name	Sequence (5'-3')[1]
psyc_0980_NHisTEV_F1	TTCGAAAACCTGTATTTTCAGGGCATGAAATCCTTATTACCG
Adapt_N-term_F2	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCGAA AACCTG
psyc_0980_Rev1	TACAGCGAGAAAATAGCATAACACCCAGCTTTC
Adapt-NHis-Rev2	TAACACCCAGCTTTCTTGTACAAAGTGGTCCCC

[1] The underlined sequences represent gene-specific regions.

Thermofluor assay

To assess the stability of the protein, a fluorescence-based thermal shift (ThermoFlour) assay was used (33). The buffer used in this experiment was 50mM HEPES pH 8. The final concentration of ParI in the reaction was 0.5 mg/ml. SyproOrange SYPRO® Orange Protein Gel Stain (Sigma-Aldrich) was diluted 1:75 from the starting concentration. The assay was performed in a volume of 25 µl and was run in a MiniOpticon real-time PCR system (BioRad) in a temperature range from 5 ºC to 95 ºC with 1 ºC increment every third second.

Differential Scanning Calorimetry

Denaturation curves were recorded at temperatures between 5 ºC and 95 ºC using a scan rate of 1 ºC/min in a N-DSC III (Calorimetry sciences corp.) calorimeter. The protein was dialyzed against the DSC-buffer (50 mM HEPES pH 8.0, 200 mM NaCl) and used at a final concentration of 1.4 mg/ml. The DSC-buffer was used as a reference.

Circular Dichroism

The recombinant ParI protein was dialyzed overnight at 4 ºC against CD-buffer (10 mM Tris pH 8.0 and 100 mM NaF). The samples were filtered through a 0.45 µm pore size filter (Spin X Costar) to remove precipitate and diluted to a final concentration of 0.15 mg/ml. 80 mM SAM diluted in milliQ water was added per 800 µl of either protein or CD-buffer. Data was collected on a J-810 CD spectrophotometer Jasco Analytical instruments)using a 1mm path length cuvette using the following settings: sensitivity was 100 mdeg, datapitch 0.5 nm, scan speed 50 nm/min, response 2.0 s, bandwidth 1nm, accumulation 3 scans, units

CD mdeg. Three scans were recorded and averaged. Three scans of buffers were also recorded in absence or presence of SAM. The measurements were done at three temperatures 15 ºC, 35 ºC, 65 ºC. Measurement at 15 ºC was done in the presence and absence of SAM, while measurements at 35 ºC and 65 ºC were done only without SAM. In the data analysis, each set of spectra corresponding to each condition was analyzed with two different programs, ContinLL and Selcon3 using the SP170 (optimized for the wavelength range between 190 and 240 nm) reference set, which gave the proportions of secondary structure by fitting the recorded spectra to the spectra of the set of known proteins (40-43).

In vivo and *in vitro* MTase genome assay

Two *in vivo* activity assays were performed. In the first assay, genomic DNA purified from bacterial samples using GenElute Gel extraction Kit (Sigma Aldrich) was used as substrates. *P. arcticus* was grown at 15 ºC for 3 days and the *P.* sp. PRwf-1, a C5-DNA MTase negative mesophilic member of the *Psychrobacter* genus, was grown at 37 ºC overnight before genomic DNA extraction. 50 ml cultures of McrBC-negative *E. coli* T7 Express strain (NEB) transformed with the pHParI construct were grown until it reached late log phase (OD 0.7-0.9) and induced with IPTG. Protein was expressed at 20 ºC for 7h. Genomic DNA from all samples was treated with McrBC endonuclease (NEB) in 1x NEB2 buffer (NEB) supplemented with 200 µg/ml BSA and 1 mM GTP, and incubated at 37 ºC for 1h. The reaction mix was separated by 1 % agarose gel electrophoresis and the DNA was post-stained with RedSafe (iNtRON Biotechnology). As a negative control, untransformed cells were used. As a positive control pUC19 vector methylated with M.SssI (NEB) methyltransferase was used.

In a second *in vivo* assay, adapted from (44) the pHParI construct was transformed into two different *E. coli* expression strains, BL21Star (DE3) (Life Technologies) and the McrBC-negative T7 Express (NEB) that lacks the same genes. Cells were transformed as previously described and plated on LB plates containing 100 µg/ml ampicillin and 0.2 mM IPTG. The plates were incubated at 37 ºC overnight.

Bisulfite reaction

The bisulfite conversion was carried out using the EZ DNA Methylation Kit (Zymo Research) and following the manufacturer's protocol. The methylated and non-methylated pUC19 DNA Set (Zymo Research) was used as substrate. Non-methylated pUC19 was incubated 16-20 hrs or less than 8 h at room temperature for ParI methylation. To enable cloning and sequencing, the uracil bases were converted to thymine by *Taq* DNA polymerase (Ampliqon) amplification using the primers provided with the kit. A TOPO-TA cloning kit (Life Technologies) was used for cloning of the PCR products into the pCR-TOPO vector. DNA was transformed into DH5α competent cells and plated on LB plates containing ampicillin (100 µg/ml) and grown overnight at 37 ºC. Colonies were used to inoculate 5 ml liquid LB media containing 100 µg/ml ampicillin and incubated overnight at 37 ºC with shaking at 220 rpm. Plasmid was isolated using QIAprep Spin Miniprep Kit (Qiagen) and sequenced in PCR reaction with M13 primers using Sanger sequencing and analyzed at the DNA sequencing core facility at the University of Tromsø. The sequences were analyzed in BioEdit (45).

References

1. Jeltsch, A. (2002) Beyond Watson and Crick: DNA methylation and molecular enzymology of DNA methyltransferases. *Chembiochem* **3**, 274-293
2. Li, E., Bestor, T. H., and Jaenisch, R. (1992) Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* **69**, 915-926
3. Bird, A. (2002) DNA methylation patterns and epigenetic memory. *Genes & development* **16**, 6-21
4. Takebayashi, S.-i., Tamura, T., Matsuoka, C., and Okano, M. (2007) Major and essential role for the DNA methylation mark in mouse embryogenesis and stable association of DNMT1 with newly replicated regions. *Molecular and cellular biology* **27**, 8243-8258
5. Rogers, J. M., Ellis-Hutchings, R. G., and Lau, C. (2007) Epigenetics and the Developmental Origins of Health and Disease. In *Comprehensive Toxicology (Second Edition)* pp. 69-88, Elsevier, Oxford
6. Labrie, V., Pai, S., and Petronis, A. (2012) Epigenetics of major psychosis: progress, problems and perspectives. *Trends in Genetics* **28**, 427-435
7. Lu, Q. (2013) The critical importance of epigenetics in autoimmunity. *Journal of autoimmunity*
8. Wilson, G. G. (1991) Organization of restriction-modification systems. *Nucleic acids research* **19**, 2539-2566
9. Roberts, R. J., Vincze, T., Posfai, J., and Macelis, D. (2009) REBASE --a database for DNA restriction and modification: enzymes, genes and genomes. *Nucleic acids research* **38**, D234-D236
10. Wilson, G. G. (1988) Type II restriction-modification systems. *Trends in Genetics* **4**, 314-318
11. Reisenauer, A., Kahng, L. S., McCollum, S., and Shapiro, L. (1999) Bacterial DNA methylation: a cell cycle regulator? *Journal of bacteriology* **181**, 5135-5139
12. Gromova, E. S., and Khoroshaev, A. V. (2003) Prokaryotic DNA methyltransferases: the structure and the mechanism of interaction with DNA. *Molecular Biology* **37**, 260-272
13. Løbner-Olesen, A., Skovgaard, O., and Marinus, M. G. (2005) Dam methylation: coordinating cellular processes. *Current opinion in microbiology* **8**, 154-160
14. Heithoff, D. M., Sinsheimer, R. L., Low, D. A., and Mahan, M. J. (1999) An essential role for DNA adenine methylation in bacterial virulence. *Science* **284**, 967-970
15. Jeltsch, A., and Gumport, R. I. (2004) DNA Methyltransferases, Bacterial. In *Encyclopedia of Biological Chemistry* pp. 644-651, Elsevier, New York
16. Song, J., Rechkoblit, O., Bestor, T. H., and Patel, D. J. (2011) Structure of DNMT1-DNA complex reveals a role for autoinhibition in maintenance DNA methylation. *Science* **331**, 1036-1040
17. Kumar, S., Cheng, X., Klimasauskas, S., Mi, S., Posfai, J., Roberts, R. J., and Wilson, G. G. (1994) The DNA (cytosine-5) methyltransferases. *Nucleic acids research* **22**, 1

18. Reinisch, K. M., Chen, L., Verdine, G. L., and Lipscomb, W. N. (1995) The crystal structure of Haelll methyltransferase covalently complexed to DNA: An extrahelical cytosine and rearranged base pairing. *Cell* **82**, 143-153
19. Cheng, X., Kumar, S., Posfai, J., Pflugrath, J. W., and Roberts, R. J. Crystal structure of the Hhal DNA methyltransferase complexed with S-adenosyl-l-methionine. *Cell* **74**, 299-307
20. Sankpal, U. T., and Rao, D. N. (2002) Structure, function, and mechanism of HhaI DNA methyltransferases. *Crit Rev Biochem Mol Biol* **37**, 167-197
21. Ayala-del-Río, H. L., Chain, P. S., Grzymski, J. J., Ponder, M. A., Ivanova, N., Bergholz, P. W., Di Bartolo, G., Hauser, L., Land, M., and Bakermans, C. (2010) The genome sequence of Psychrobacter arcticus 273-4, a psychroactive Siberian permafrost bacterium, reveals mechanisms for adaptation to low-temperature growth. *Applied and environmental microbiology* **76**, 2304-2312
22. Bakermans, C., Ayala-del-Rio, H. L., Ponder, M. A., Vishnivetskaya, T., Gilichinsky, D., Thomashow, M. F., and Tiedje, J. M. (2006) Psychrobacter cryohalolentis sp. nov. and Psychrobacter arcticus sp. nov., isolated from Siberian permafrost. *Int J Syst Evol Microbiol* **56**, 1285-1291
23. Bergholz, P. W., Bakermans, C., and Tiedje, J. M. (2009) Psychrobacter arcticus 273-4 uses resource efficiency and molecular motion adaptations for subzero temperature growth. *Journal of bacteriology* **191**, 2340-2352
24. Raleigh, E. A. (1992) Organization and function of the mcrBC genes of Escherichia coli K12. *Molecular microbiology* **6**, 1079-1086
25. Sutherland, E., Coe, L., and Raleigh, E. A. (1992) McrBC: a multisubunit GTP-dependent restriction endonuclease. *Journal of Molecular Biology* **225**, 327-348
26. Seshasayee, A. S. N., Singh, P., and Krishna, S. (2012) Context-dependent conservation of DNA methyltransferases in bacteria. *Nucleic acids research* **40**, 7066-7073
27. Militello, K. T., Simon, R. D., Qureshi, M., Maines, R., Van Horne, M. L., Hennick, S. M., Jayakar, S. K., and Pounder, S. (2012) Conservation of Dcm-mediated cytosine DNA methylation in Escherichia coli. *FEMS microbiology letters* **328**, 78-85
28. Zhou, Y., Liang, Y., Lynch, K. H., Dennis, J. J., and Wishart, D. S. (2011) PHAST: A Fast Phage Search Tool. *Nucleic Acids Research*
29. Kossykh, V. G., Schlagman, S. L., and Hattman, S. (1995) Function of Pro-185 in the ProCys of conserved motif IV in the EcoRII [cytosine-C5]-DNA methyltransferase. *FEBS letters* **370**, 75-77
30. ROSSAU, R., VAN LANDSCHOOT, A., GILLIS, M., and DE LEY, J. (1991) Taxonomy of Moraxellaceae fam. nov., a New Bacterial Family To Accommodate the Genera Moraxella, Acinetobacter, and Psychrobacter and Related Organisms. *International Journal of Systematic Bacteriology* **41**, 310-319
31. Sun, P., Tropea, J. E., and Waugh, D. S. (2011) Enhancing the solubility of recombinant proteins in Escherichia coli by using hexahistidine-tagged maltose-binding protein as a fusion partner. In *Heterologous Gene Expression in E. coli* pp. 259-274, Springer

32. Raleigh, E. A., Murray, N. E., Revel, H., Blumenthal, R. M., Westaway, D., Reith, A. D., Rigby, P. W. J., Elhai, J., and Hanahan, D. (1988) McrA and McrB restriction phenotypes of some E. coli strains and implications for gene cloning. *Nucleic acids research* **16**, 1563-1575
33. Boivin, S., Kozak, S., and Meijers, R. (2013) Optimization of protein purification and characterization using Thermofluor screens. *Protein expression and purification* **91**, 192-206
34. Greenfield, N. J. (2006) Using circular dichroism spectra to estimate protein secondary structure. *Nature protocols* **1**, 2876-2890
35. Boratyn, G. M., Camacho, C., Cooper, P. S., Coulouris, G., Fong, A., Ma, N., Madden, T. L., Matten, W. T., McGinnis, S. D., Merezhuk, Y., Raytselis, Y., Sayers, E. W., Tao, T., Ye, J., and Zaretskaya, I. (2013) BLAST: a more efficient report with usability improvements. *Nucleic Acids Research* **41**, W29-W33
36. Notredame, C., Higgins, D. G., and Heringa, J. (2000) T-Coffee: A novel method for fast and accurate multiple sequence alignment. *J Mol Biol* **302**, 205-217
37. Robert, X., and Gouet, P. (2014) Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Research* **42**, W320-W324
38. Jones, D. T. (1999) Protein secondary structure prediction based on position-specific scoring matrices. *Journal of molecular biology* **292**, 195-202
39. Katzen, F. (2007) GatewayÂ® recombinational cloning: a biological operating system. *Expert Opinion on Drug Discovery* **2**, 571-589
40. Provencher, S. W., and Gloeckner, J. (1981) Estimation of globular protein secondary structure from circular dichroism. *Biochemistry* **20**, 33-37
41. Sreerama, N., and Woody, R. W. (1993) A self-consistent method for the analysis of protein secondary structure from circular dichroism. *Analytical biochemistry* **209**, 32-44
42. Van Stokkum, I. H. M., Spoelder, H. J. W., Bloemendal, M., Van Grondelle, R., and Groen, F. C. A. (1990) Estimation of protein secondary structure and error analysis from circular dichroism spectra. *Analytical biochemistry* **191**, 110-118
43. Sreerama, N., Venyaminov, S. Y. U., and Woody, R. W. (1999) Estimation of the number of helical and strand segments in proteins using circular dichroism spectroscopy. *Protein Science* **8**, 370-380
44. Sankpal, U. T., and Rao, D. N. (2002) Mutational analysis of conserved residues in HhaI DNA methyltransferase. *Nucleic acids research* **30**, 2628-2638
45. Hall, T. A. (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. In *Nucleic acids symposium series* Vol. 41 pp. 95-98

Figure legends.

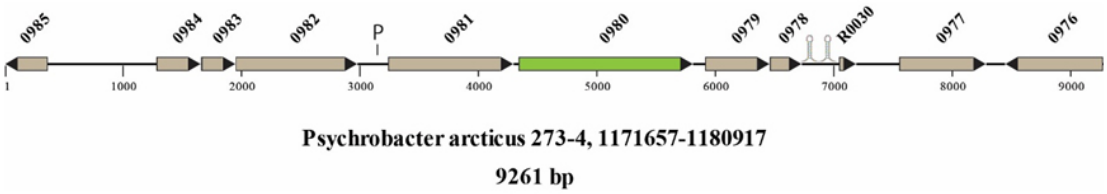


Figure 1. Gene arrangement surrounding the *par*I gene 0980. Cartoon is depicting the genome region 1171657-1180917 of *P. arcticus* 273-4. A promoter (P) was predicted upstream of gene 0981 (green) and two Rho-independent terminators (hairpins) can be predicted downstream gene 0978. Cartoon is drawn to scale.

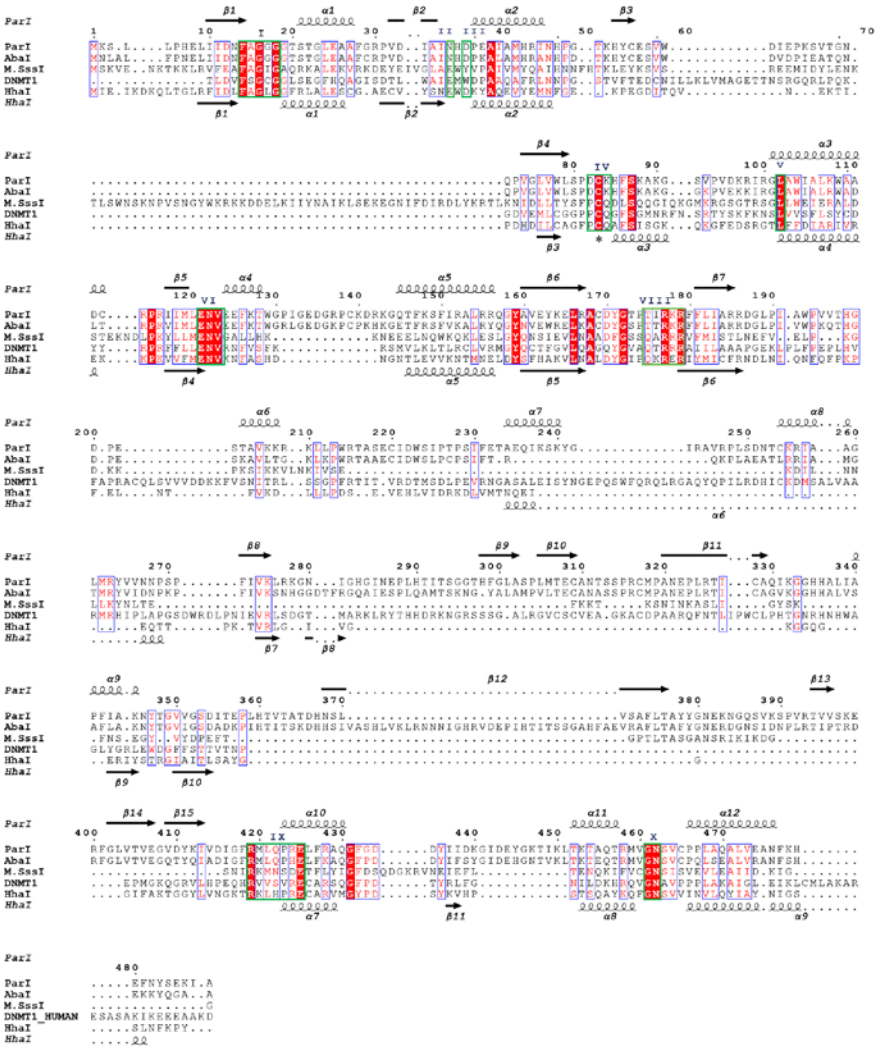


Figure 2. Multiple sequence alignment of bacterial C5 DNA-MTases and human DNMT1. Motifs VI and VIII are involved in DNA binding, motifs I, X, II, III and V are involved in SAM binding and the catalytic motif IV is marked with black asterisk. The top secondary structure elements for ParI are predicted using the PsiPred (38) and protein sequence for ParI. Identical residues, *i.e.* conserved motifs are marked by red blocks and in green squares. HhaI from *Haemophilus parahaemolyticus* [GenBank: 127455]; ParI from *Psychrobacter arcticus* 273-4 [GenBank:71065540]; M.SssI from *Spiroplasma* sp. [GenBank: 417325]; AbaUH5107 from *Acinetobacter baumannii* [GenBank: 446969424] and the catalytic C5 DNA-MTase domain of human DNMT1 [GenBank: 12231019]. The bottom secondary structure elements was rendered from HhaI [PDB: 1MHT].

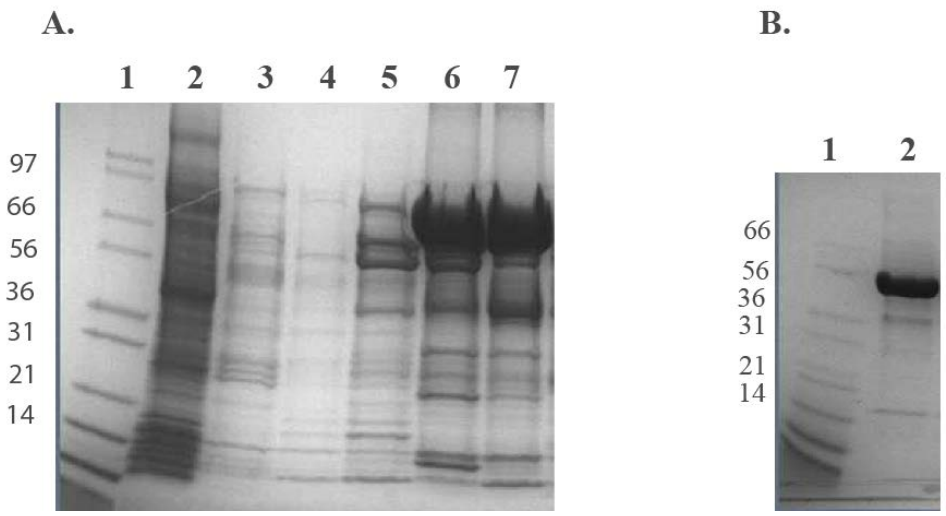


Figure 3. Purification of recombinant ParI by immobilized metal affinity chromatography. (A) Analysis of selected fractions from the initial HisTrap purification of the MBPHis ParI enzyme. Lane 1, Mark 12 molecular weight marker; lane 2, flow-through fraction from purification; lanes 3-7, fractions from purification with lanes 6 and 7 containing MBPHis ParI. Theoretically, the recombinant His-MBP-tagged ParI is estimated to be 94 kDa. (B) SDS gel picture from second histrap purification, after removing tag by cleavage with TEV protease. Lane 1, Mark 12 molecular weight marker; lane 2, flow through-fraction containing recombinant ParI after tag-removal. The theoretical calculated molecular weight is 54 kDa. Green arrow points to ParI in both figures.

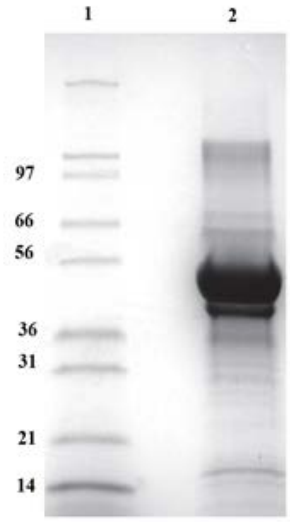


Figure 4. Gelfiltration of ParI. SDS-PAGE gel showing ParI after gelfiltration, 0,02 mg/ml of protein was loaded on gel.

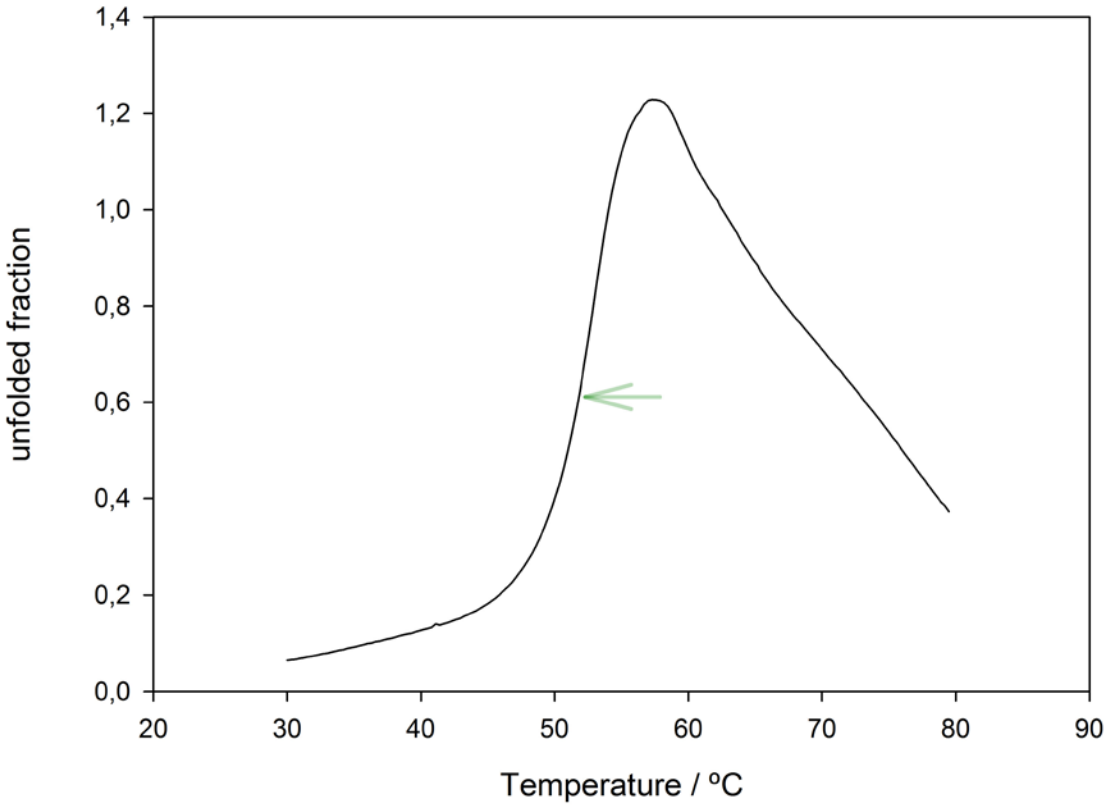


Figure 5. Thermo**gram showing temperature-induced unfolding of ParI.** The reaction was run across a temperature range of 5-95 ºC; here is shown the range 20-90 ºC. The green arrow indicates the calculated Tm of 53 ºC. Graph was made in SigmaPlot.

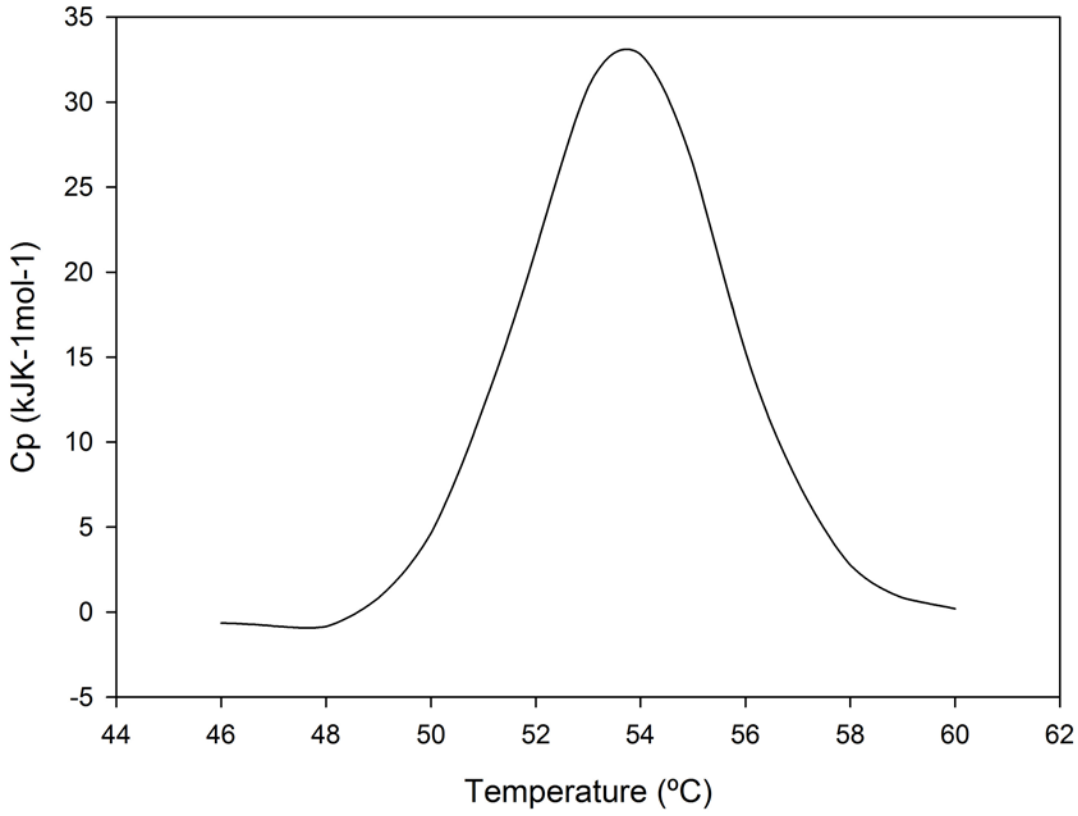


Figure 6. DSC thermogram for ParI. The protein was scanned with a scan rate 1 ºC /min in temperature range from 5-75 ºC. Here, only section where protein is unfolding (44-62 ºC) is shown. The melting temperature for ParI was 54 ºC, ΔH was 82.81 kcal/mol and ΔS was 0.2411 kcal/(molK). Graph was made in SigmaPlot.

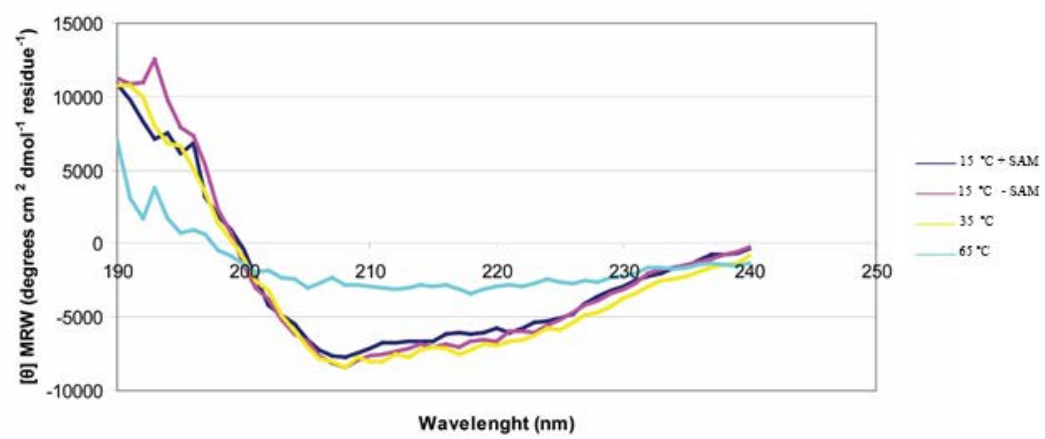


Figure 7. The converted output from CD measurements. Dark blue line presents secondary structure prediction at 15 ºC with SAM, purple is same only without SAM. Yellow line presents measurement done at 35 ºC and light blue at 65 ºC. Mean residue ellipticity (the measured molar ellipticity of the molecule divided by the number of monomer units (residues) in the molecule) is shown on y-axis and is shown in degrees cm^2 $dmol^{-1}$ per residue. On x-axis represents wavelength at which measurement was obtained.

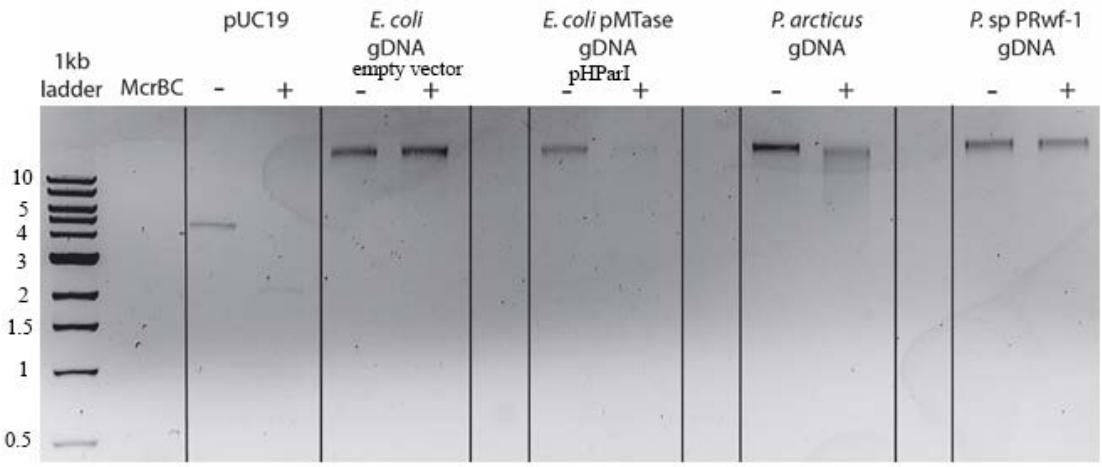
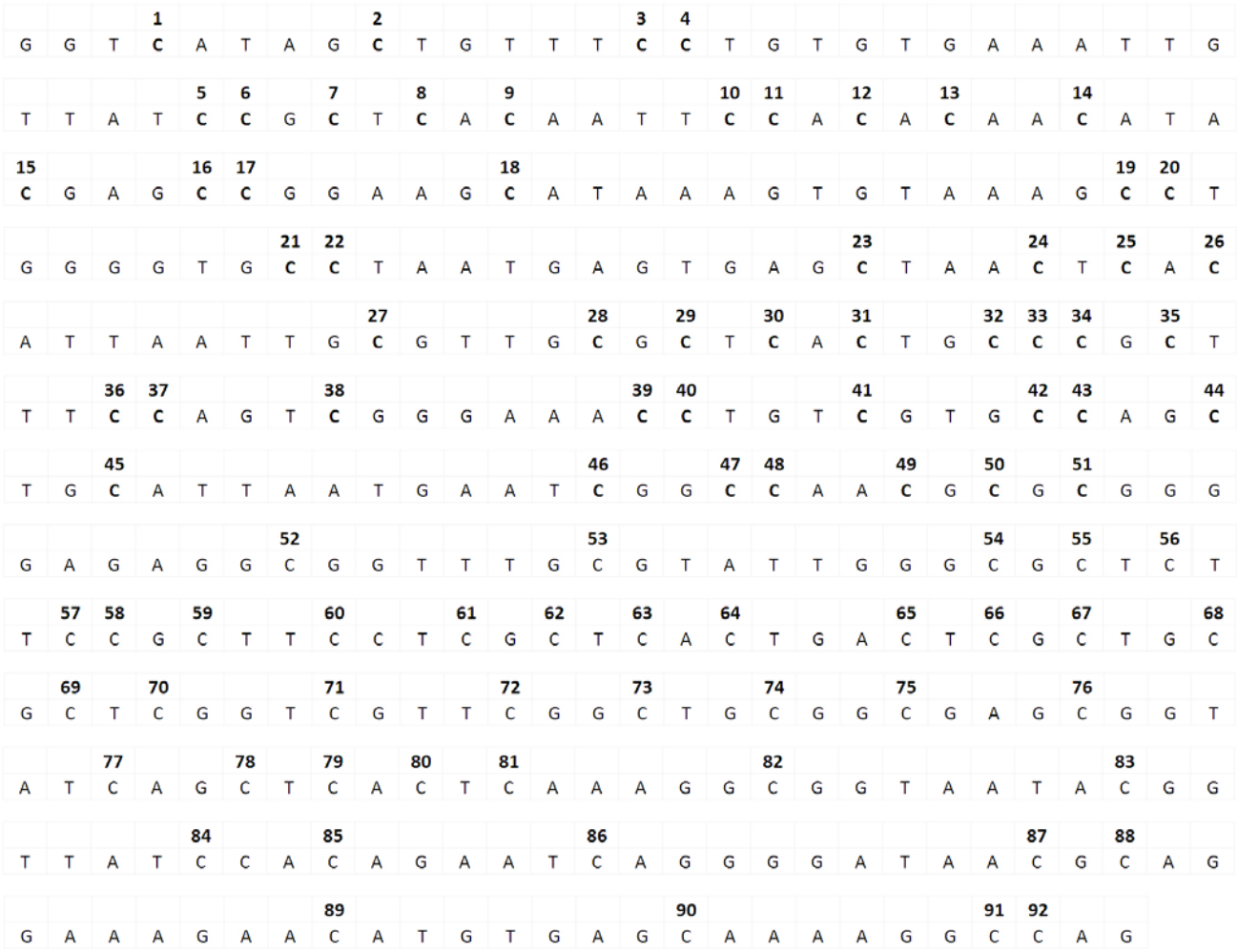


Figure 8. ParI methylates DNA and makes it McrBC sensitive. The *in vivo* methylation assay was used to analyze the restriction of genomic DNA from *E. coli* transformed with the pParI17 vector (*E. coli* MTase) and *E. coli* transformed with empty vector (*E. coli* gDNA), *P. arcticus* and *P.* sp. Prwf-1. DNA was, after purification, treated with McrBC that cleaves methylated DNA in the form $(G/A)^{m}C$ and loaded on the agarose gel. As positive control pUC19 vector pre-methylated with M.SssI was used as a substrate for McrBC treatment.

A.


1 2 3 4
G G T C A T A G C T G T T T C C T G T G T G A A A T T G
5 6 7 8 9 10 11 12 13 14
T T A T C C G C T C A C A A T T C C A C A C A A C A T A
15 16 17 18 19 20
C G A G C C G G A A G C A T A A A G T G T A A A G C C T
21 22 23 24 25 26
G G G G T G C C T A A T G A G T G A G C T A A C T C A C
27 28 29 30 31 32 33 34 35
A T T A A T T G C G T T G C G C T C A C T G C C C G C T
36 37 38 39 40 41 42 43 44
T T C C A G T C G G G A A A C C T G T C G T G C C C A G C
45 46 47 48 49 50 51
T G C A T T A A T G A A T C G G C C A A C G C G C G G G
52 53 54 55 56
G A G A G G C G G T T T G C G T A T T G G G C G C T C T
57 58 59 60 61 62 63 64 65 66 67 68
T C C G C T T C C T C G C T C A C T G A C T C G C T G C
69 70 71 72 73 74 75 76
G C T C G G T C G T T C G G C T G C G G C G A G C G G T
77 78 79 80 81 82 83
A T C A G C T C A C T C A A A G G C G G T A A T A C G G
84 85 86 87 88
T T A T C C A C A G A A T C A G G G G A T A A C G C A G
89 90 91 92
G A A A G A A C A T G T G A G C A A A A G G C C A G

B.

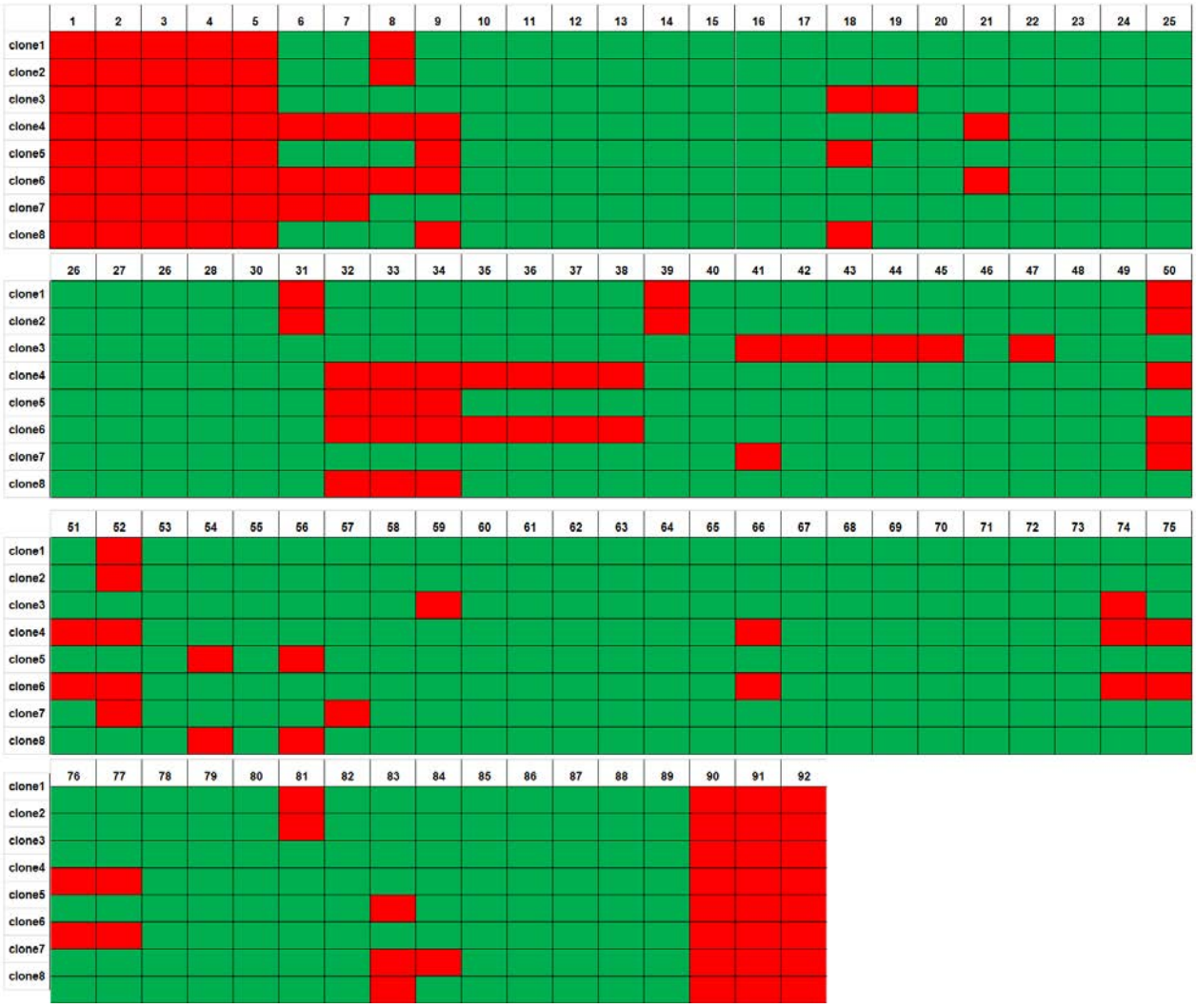


Figure 9. Analysis of the ParI methylation site A. Sequence alignment of template used for bisulfite conversion. All cytosines are numbered. **B**. **Presentation of sequences obtained from bisulfite conversion**. Eight samples that were analysed are presented. All methylated cytosines are represented by green colour, while non methylated cytosines are presented by red colour. Numbers above each column are same as number of cytosines in sequence.