

Full Paper

# Identification of the DNA methyltransferases establishing the methylome of the cyanobacterium *Synechocystis* sp. PCC 6803

Martin Hagemann<sup>1,\*</sup>, Katrin Gärtner<sup>1</sup>, Matthias Scharnagl<sup>1</sup>, Paul Bolay<sup>2</sup>, Steffen C. Lott<sup>2</sup>, Janina Fuss<sup>3</sup>, Bruno Huetzel<sup>3</sup>, Richard Reinhardt<sup>3</sup>, Stephan Klähn<sup>2,†</sup>, and Wolfgang R. Hess<sup>2,4</sup>

<sup>1</sup>Institute of Biosciences, Plant Physiology, University of Rostock, D-18059 Rostock, Germany, <sup>2</sup>Faculty of Biology, University of Freiburg, D-79104 Freiburg, Germany, <sup>3</sup>Max Planck-Genome-Centre Cologne, Max Planck Institute for Plant Breeding Research, D-50829 Köln, Germany, and <sup>4</sup>Freiburg Institute for Advanced Studies, University of Freiburg, D-79104, Freiburg, Germany

\*To whom correspondence should be addressed. Tel. +49 381 4986110. Fax. +49 381 4986112.  
Email: martin.hagemann@uni-rostock.de

†Present address: Helmholtz Centre for Environmental Research - UFZ, D-04318 Leipzig, Germany.

Edited by Prof. Takashi Ito

Received 22 September 2017; Editorial decision 23 January 2018; Accepted 24 January 2018

## Abstract

DNA methylation in bacteria is important for defense against foreign DNA, but is also involved in DNA repair, replication, chromosome partitioning, and regulatory processes. Thus, characterization of the underlying DNA methyltransferases in genetically tractable bacteria is of paramount importance. Here, we characterized the methylome and orphan methyltransferases in the model cyanobacterium *Synechocystis* sp. PCC 6803. Single molecule real-time (SMRT) sequencing revealed four DNA methylation recognition sequences in addition to the previously known motif <sup>m5</sup>CGATCG, which is recognized by M.Ssp6803I. For three of the new recognition sequences, we identified the responsible methyltransferases. M.Ssp6803II, encoded by the *slr0729* gene, modifies GG<sup>m4</sup>CC, M.Ssp6803III, encoded by *slr1803*, represents the cyanobacterial *dam*-like methyltransferase modifying G<sup>m6</sup>ATC, and M.Ssp6803V, encoded by *slr6095* on plasmid pSYSX, transfers methyl groups to the bipartite motif GG<sup>m6</sup>AN<sub>7</sub>TTGG/CCA<sup>m6</sup>AN<sub>7</sub>TCC. The remaining methylation recognition sequence GA<sup>m6</sup>AGGC is probably recognized by methyltransferase M.Ssp6803IV encoded by *slr6050*. M.Ssp6803III and M.Ssp6803IV were essential for the viability of *Synechocystis*, while the strains lacking M.Ssp6803I and M.Ssp6803V showed growth similar to the wild type. In contrast, growth was strongly diminished of the  $\Delta$ *slr0729* mutant lacking M.Ssp6803II. These data provide the basis for systematic studies on the molecular mechanisms impacted by these methyltransferases.

**Key words:** cyanobacteria, DNA methyltransferase, mutant, photosynthetic pigment, phylogenetics

## 1. Introduction

DNA serves as the universal carrier of information in living cells. In addition to the genetic information encoded in the nucleotide

sequence, an epigenetic level of information is encoded in DNA modifications, which play a fundamental role in the differentiation and development of eukaryotic cells, in cancer development and

prevention, aging, and acclimation to cellular and environmental stimuli.<sup>1,2</sup> DNA modification is carried out by methyltransferases, which transfer methyl groups from the universal substrate S-adenosyl-methionine (AdoMet) to their respective recognition sequences. Among bacteria, epigenetic modifications have been most commonly associated with the genome-wide methylation of specific DNA sequences, which are linked to restriction-modification (RM)-based defense mechanisms against foreign DNA, such as the DNA from phages.<sup>3,4</sup> C5-cytosine methylation is regarded as the most common DNA modification in eukaryotes,<sup>5</sup> but N6-adenine methylation has also been reported.<sup>6–8</sup> Prokaryotic DNA methyltransferases typically catalyze N6-adenine, N4- or C5-cytosine methylation.

In addition to the methyltransferases in RM systems, many prokaryotic genomes harbour orphan DNA methyltransferases that act independently of the RM systems. Most of these solitary or orphan methyltransferases are not well characterized, but functions in the regulation of gene expression, DNA replication, repair, and others have been suggested.<sup>9</sup> One of the most widespread and best characterized orphan methyltransferases is the *E. coli* Dam enzyme, an N6-adenine-specific methyltransferase modifying the target sequence GATC. Dam methylation plays an important role in DNA repair and replication (reviewed in [9]) but also in the regulation of gene expression<sup>10</sup> or phase variation of uropathogenic *E. coli*.<sup>11,12</sup> Dcm, which mediates cytosine DNA methylation, is another widespread orphan methyltransferase activity found in 162 strains of *E. coli*.<sup>13</sup> The molecular details of how DNA methylation regulates gene expression and subsequently the cell cycle have been elucidated in the bacterial model system *Caulobacter crescentus*.<sup>14</sup>

Single molecule real-time (SMRT) sequencing permits the parallel estimation of the methylation status of specific nucleotides. It was first used for the direct methylome profiling of *Mycoplasma pneumoniae*.<sup>15</sup> More recently, SMRT sequencing has been applied to characterize the methylomes of 230 prokaryotic strains, revealing that the majority of them contain extensive and variable DNA methylation patterns.<sup>16</sup> Cyanobacteria, which are the only prokaryotes that perform oxygenic photosynthesis and are increasingly used as cell factories in green biotechnology,<sup>17</sup> have been only scarcely characterized with regard to epigenetic modifications. In *Anabaena* sp. PCC 7120, four different orphan methyltransferases were detected in addition to the enzymes of the endogenous RM systems.<sup>18</sup> A high degree of adenine methylation was reported for the marine cyanobacterium *Trichodesmium* sp. NIBB1067.<sup>19</sup> In the present study, we analyzed the methylome of the model cyanobacterium *Synechocystis* sp. PCC 6803 (hereafter *Synechocystis*), which seems to be virtually free of endogenous restriction endonucleases.<sup>20</sup> Nevertheless, the chromosomal DNA of *Synechocystis* was found to be methylated, and the genome contains several orphan methyltransferase genes. Only the cytosine-specific orphan methyltransferase M.Ssp6803I, encoded by gene *str0214*, was previously analyzed.<sup>20</sup> The M.Ssp6803I-dependent modification of the <sup>m5</sup>CGATCG motif improved the integration efficiency of external DNA into the *Synechocystis* chromosome.<sup>21</sup> Here, we present the first methylome analysis of *Synechocystis*, which identified five DNA methylation recognition sequences and the corresponding methyltransferases.

## 2. Materials and methods

### 2.1. Strains and culture conditions

*Synechocystis* sp. 6803 substrain PCC-M<sup>22</sup> was used in all experiments. The axenic strain was maintained on agar plates supplemented with BG11 mineral medium<sup>23</sup> at 30°C under constant

illumination. Transformants were initially selected on media containing 10 µg ml<sup>-1</sup> kanamycin (Km; Sigma), while the segregation of clones and cultivation of mutants was performed at 50 µg ml<sup>-1</sup> Km. For physiological characterization, axenic cultures of the different strains (Supplementary Table S1) were grown photoautotrophically in BG11 medium, either under slight shaking in Erlenmeyer flasks at 50 µmol photons s<sup>-1</sup> m<sup>-2</sup>, or under bubbling with CO<sub>2</sub>-enriched air (5% [v/v]) in batch cultures at 29°C under continuous illumination of 180 µmol photons s<sup>-1</sup> m<sup>-2</sup> (warm light, Osram L58 W32/3). Contamination by heterotrophic bacteria was evaluated by microscopy or spreading of 0.2 mL culture on LB plates. The *E. coli* strains TG1, TOP10, and DH5α were used for routine DNA manipulations. *E. coli* was cultured in LB medium at 37°C. Growth was followed by measurements of the optical density at 750 nm (OD<sub>750</sub>) for *Synechocystis* and at 500 nm (OD<sub>500</sub>) for *E. coli*.

### 2.2. Methylome analysis

For SMRT sequencing, high quality genomic DNA was isolated using the CTAB protocol.<sup>24</sup> Libraries were prepared according to the large SMRTbell gDNA protocol (Pacific Biosciences) with 10 kb insert size. Genomic DNA was sequenced with a PacBio RS II platform. Base modifications were analyzed using the program RS\_Modification\_Detection.1 from Pacific Biosciences (v. 2.3.0). For bisulfite sequencing, ~200 ng of DNA were bisulfite treated with the Zymo Gold kit (Zymo Research) and libraries constructed using the Ovation Ultra-Low Methyl-Seq library kit (NuGEN) following manufacturer's instructions, followed by sequencing on the Illumina HiSeq2500 system yielding 2 559 017 raw reads. The sequences were quality filtered and adapter trimmed using Trimmomatic v0.36<sup>25</sup>) and FastQC v0.67 (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/> (1 February 2018, date last accessed)) leaving 2 552 913 reads for further analysis. For mapping to the *Synechocystis* chromosome and quantitative evaluation we used Bismark (v0.17 with default options<sup>26</sup>) in conjunction with Bowtie 2.<sup>27</sup> All SMRT and Illumina sequencing raw data are available from the National Center for Biotechnology Information at <https://www.ncbi.nlm.nih.gov/biosample/8378604> (1 February 2018, date last accessed) (BioProject ID: PRJNA430784, BioSample: SAMN08378604, SRA: SRS2844079).

### 2.3. DNA manipulations

The isolation of total DNA from *Synechocystis* was performed as described previously.<sup>20</sup> All other DNA techniques, such as plasmid isolation, transformation of *E. coli*, ligations and restriction analysis (restriction enzymes were obtained from Promega and New England Biolabs) followed standard methods.<sup>28</sup> For the restriction analyses using chromosomal DNA from *Synechocystis*, the restriction endonucleases were used in a 10-fold excess and were incubated for at least 16 h at 37°C to ensure complete digestion. Synthetic primers were deduced from the complete genome sequence of *Synechocystis*<sup>29</sup> for the specific amplification of putative DNA-specific methyltransferase-coding genes (Supplementary Table S2). Interposon mutagenesis was used to generate mutants defective in these genes. For this purpose, DNA fragments containing their encoding sequences were amplified by PCR and cloned into pGEM-T (Promega). The *aphII* gene, conferring Km resistance from pUC4K (Pharmacia), was introduced into selected restriction sites (see Supplementary Table S1, Figs 2, 3, and 5), and verified constructs were transferred into *Synechocystis* as described.<sup>20</sup>

## 2.4. Generation of complementation strains

For the ectopic expression of *slI0729* and *ssl1378*, their coding sequences were fused to the regulatory sequences of the *ziaA* gene. The upstream region of *ziaA*, including *PziaA*, its 5'-UTR, and the upstream *ziaR*-gene, as well as *slI0729* or *ssl1378*, were amplified by PCR (for oligonucleotides, see [Supplementary Table S2](#)). The PCR products were digested with *ClaI*, followed by heat inactivation. The digested products were ligated and re-amplified with PCR, and the resulting *PziaA::slI0729* and *PziaA::ssl1378* fusions were cloned into the self-replicating, broad-host range vector pVZ322 via *XbaI/XhoI* cleavage sites. The resulting plasmids were introduced into  $\Delta slI0729$  via conjugal transfer, as described previously.<sup>30</sup>

## 2.5. Generation of recombinant proteins and methyltransferase assay

For overexpression and purification of the putative DNA-specific methyltransferases, the open reading frames (ORFs) were amplified from chromosomal DNA by PCR using primers for the directed in-frame cloning with the N-terminal His-tag into vectors of the pBAD/His A, B, C series (Invitrogen). Correct insertions were verified by sequencing. For overexpression, recombinant cells of the methyltransferase-defective *E. coli* strain TOP10 were cultured at 30°C in LB medium. The expression of the recombinant protein was induced at OD<sub>500</sub> = 1.0 by addition of arabinose (0.002% final concentration). The proteins were extracted from *E. coli* by sonication, and the fusion proteins were purified on a Ni-NTA matrix (ProBond™ Resin, novex, life technologies) after elution with an imidazole gradient. The resulting protein fractions were evaluated using SDS gels containing 12% acrylamide. The recombinant proteins were detected by immuno-blotting with an antibody specific for the N-terminal His-tag (Invitrogen).

The DNA-specific methyltransferase activity was assayed by incubation of non-methylated chromosomal DNA of *Micrococcus lysodeikticus* (Sigma) with [<sup>3</sup>H]-AdoMet (Amersham), as described.<sup>20</sup> The protein content was estimated according to reference.<sup>31</sup>

## 2.6. Phylogenetic analysis

The phylogenetic comparison included up to 20 of the most similar proteins found in BlastP searches<sup>32,33</sup> against Genbank with the amino acid sequence of M.Ssp6803II (SlI0729). In addition, the characterized DNA methyltransferases M.Ssp6803I (Slr0214), M.Ssp6803III (Slr1803) and some closely related proteins, as well as functionally characterized methyltransferases, were included. The protein sequences were aligned using ClustalW<sup>34</sup> embedded in the BioEdit software package. Extended sequence parts were manually removed from the N-terminal and C-terminal ends. The phylogenetic tree was calculated in MEGA5 using the neighbour joining method.<sup>35</sup>

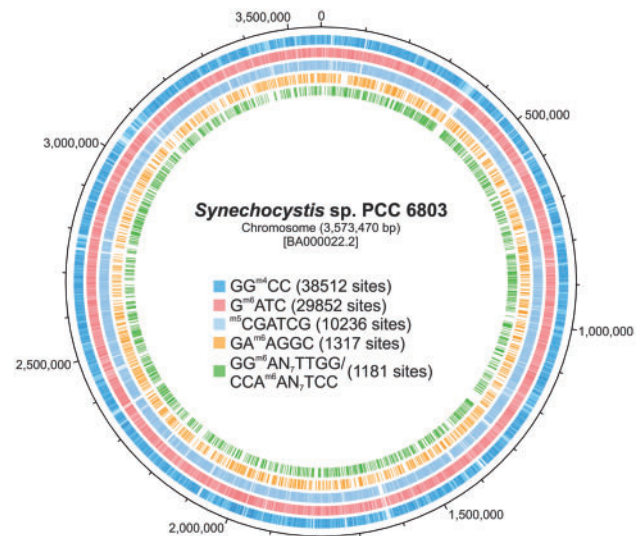
## 3. Results

### 3.1. The *Synechocystis* methylome

SMRT and bisulfite sequencing were used to analyze the *Synechocystis* methylome. The <sup>m</sup>CGATCG motif was previously characterized but the nature of the methylation could only be inferred based on protein similarity.<sup>20</sup> Bisulfite sequencing permits the direct detection of 5-methylcytosines, which was found for this motif, yielding <sup>m5</sup>CGATCG ([Supplementary Fig. S1](#)). In addition, SMRT sequencing indicated that the *Synechocystis* genome harbours

at least four further methylation recognition sequences, which were found on the chromosome with different frequencies: GG<sup>m</sup>CC, G<sup>m6</sup>ATC, GA<sup>m6</sup>AGGC, and GG<sup>m6</sup>AN<sub>7</sub>TTGG/CCA<sup>m6</sup>AN<sub>7</sub>TCC ([Fig. 1](#), [Table 1](#)). The SMRT sequencing results showed the methylation of GG<sup>m</sup>CC at the first C but failed to identify the precise modification. While SMRT sequencing can detect <sup>m6</sup>A with high accuracy and sensitivity, it is less sensitive towards <sup>m4</sup>C and performs badly on <sup>m5</sup>C methylation.<sup>36</sup> However, standard bisulfite sequencing protocols may also be used to map <sup>m4</sup>C, because <sup>m4</sup>C is partially resistant to bisulfite-mediated deamination.<sup>37</sup> Indeed, we observed a GG<sup>m</sup>CC methylation frequency of ~50% ([Supplementary Fig. S1](#)), matching previous records on the efficiency of bisulfite treatment for the detection of <sup>m4</sup>C.<sup>37</sup> In contrast, the methylation at C5 position of <sup>m5</sup>CGATCG motifs was detected to 100% by bisulfite sequencing ([Supplementary Fig. S1](#)) but was not observed by the SMRT method. The fact that the GGCC methylation was detected both by SMRT and by bisulfite sequencing suggests that it is at the N4 position (GG<sup>m4</sup>CC). The bisulfite sequencing data were used for a global analysis of the methylation of CGATCG and GGCC sites ([Supplementary Fig. S2](#)). This analysis revealed 80–85% methylated <sup>m5</sup>CGATCG sites. About 10% of the sites seem to be not completely or unmethylated ([Supplementary Fig. S2A](#)). In the case of the GGCC methylation site, bisulfite sequencing revealed 90% of all sites being GG<sup>m4</sup>CC ([Supplementary Fig. S2C](#)). The great majority of modifications of GGCC and CGATCG sites is consistent with previous restriction analysis, where DNA isolated from *Synechocystis* was resistant to treatment with methylation-sensitive restriction enzymes *HaeIII* (recognition sequence GGCC), *PvuI* (recognition sequence CGATCG), and *MboI* (recognition sequence GATC).<sup>20</sup> Similarly, the DNA of the marine cyanobacterium *Trichodesmium* sp. NIBB1067 was also modified at GATC sites leading to stimulation or inhibition of methylation-dependent restriction enzymes.<sup>19</sup>

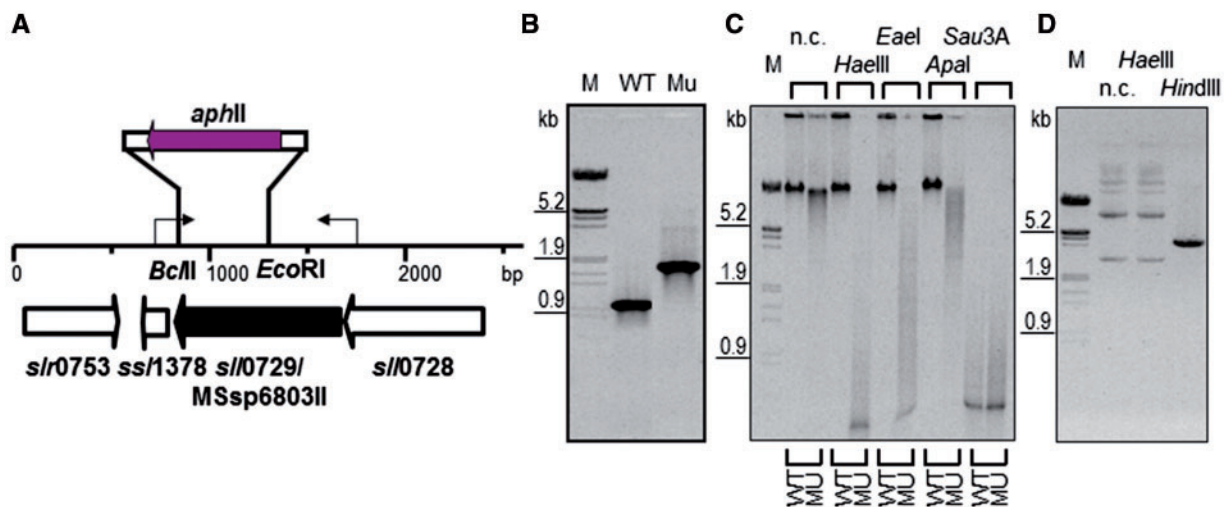
The Restriction Enzyme Database (REBASE<sup>®38</sup>) was searched for *Synechocystis* genes encoding putative DNA methyltransferases. In addition to *slr0214*, which encodes M.Ssp6803I modifying



**Figure 1.** Overview of the methylome of *Synechocystis* sp. PCC 6803. The methylome of *Synechocystis* sp. strain PCC 6803 comprises five different methylation motifs, which were detected using SMRT (Pacific Biosciences) and bisulfite sequencing. The genome plot shows the distribution of the recognition sequences on both strands of the chromosome and was made by using DNAPlotter.<sup>54</sup>

**Table 1.** Methyltransferases in *Synechocystis* sp. PCC 6803

Methyl-transferase	Encoding gene	Localized on	Recognition sequence	Sites on chromosome	Reference
M.Ssp6803I	<i>slr0214</i>	Chromosome	$m^5$ CGATCG	38 512	This study, Scharnagl <i>et al.</i> <sup>20</sup>
M.Ssp6803II	<i>sll0729</i>	Chromosome	GG $m^4$ CC	29 852	This study
M.Ssp6803III	<i>slr1803</i>	Chromosome	G $m^6$ ATC	10 236	This study
M.Ssp6803IV	<i>slr6050</i>	Plasmid pSYSX	GA $m^6$ AGGC	1317	This study, REBASE <sup>38</sup>
M.Ssp6803V	<i>slr6095</i>	Plasmid pSYSX	GG $m^6$ AN $_7$ TTGG/ CCA $m^6$ AN $_7$ TCC	1181	This study
M.Ssp6803VI	<i>ssl8010/sll8009</i>	Plasmid pSYSG	None		This study



**Figure 2.** Functional verification of M.Ssp6803II encoded by *sll0729*. (A) Construction strategy for the generation of the  $\Delta sll0729$  mutant defective in the ORF of M.Ssp6803II. Thin arrows indicate primer binding sites, which were used to verify the genotype of the mutants. (B) Characterization of  $\Delta sll0729$  mutant genotype by PCR using gene specific primers and as templates DNA from wild type (WT) and mutant cells (Mu). (C) Separation of fragments generated during a restriction analysis with GGCC-specific, methylation-sensitive restriction endonucleases (*HaellI*, *EaeI*, and *ApaI*) and GATC-specific enzyme (*Sau3A*) of chromosomal DNA of the wild type (WT) and  $\Delta sll0729$  mutant (Mu) by agarose gel electrophoresis. (n.c., uncut control DNA; M, DNA size marker  $\lambda$ -DNA cut by *HindIII* and *EcoRI*). Please note that GG $m^6$ C is cleaved by *HaellI*, whereas GG $m^6$ CC is resistant to cleavage (NEB). (D) Restriction analysis of plasmids, which were isolated from cells of the DNA-methylation negative *E. coli* strain TOP10 over-expressing M.Ssp6803II (*sll0729*), using methylation-sensitive enzyme (*HaellI*) and *HindIII* as control. (n.c., uncut control DNA; M, DNA size marker  $\lambda$ -DNA cut by *HindIII* and *EcoRI*).

$m^5$ CGATCG,<sup>20</sup> the database analysis revealed five further ORFs with significant similarities to characterized methyltransferase genes, which could be responsible for the observed methylation pattern. In particular, these are *sll0729* and *slr1803*, which are present on the chromosome, and *slr6050*, *slr6095* and *sll8009* (see Table 1 for an overview on methyltransferase nomenclature and corresponding genes), which are found on the plasmids pSYSX and pSYSG, respectively.

### 3.2. The *Synechocystis* DNA methyltransferases

#### 3.2.1. The GG $m^4$ CC motif is methylated by M.Ssp6803II, which is encoded by *sll0729*

According to the presence of a Dam and a D12 class N6-adenine-specific DNA methyltransferase domain (COG0338 and pfam02086), the gene *sll0729* likely encodes an adenine-specific methyltransferase. However, phylogenetic analyses showed separate clustering of this methyltransferase from characterized adenine-specific methyltransferases of heterotrophic bacteria such as Dam of *E. coli* (Supplementary Fig. S3). Using the BlastP algorithm<sup>32,33</sup> and

the NCBI nr database (January 2017), the closest homolog among cyanobacteria was found in the genome of *Fischerella* sp. JSC-11 (identity of 73%, BlastP *e* value of  $9e^{-148}$ ), and it should be mentioned that 234 other proteins of high similarity exist in cyanobacteria (*e*-value  $\leq 5e^{-15}$ ). The closest homolog beyond cyanobacteria exists in the genome of the archaeon *Methanosarcina mazei* (identity of 53%, BlastP *e* value of  $9e^{-112}$ ) and the bacterial strain *Spirochaetes bacterium* GWB\_1\_27\_13 (identity of 60%, BlastP *e* value of  $9e^{-111}$ ), respectively.

To study the function of *sll0729*, a mutant was generated in which the *aphII* gene, conferring Km resistance, was inserted into *sll0729*, leading to the deletion of a *BclI*-*EcoRI* fragment containing most of its coding sequence (Fig. 2A). Genotypic analysis revealed that a completely segregated  $\Delta sll0729$  mutant was obtained since only the fragment enlarged by the expected size of the inserted *aphII* gene was amplified by PCR, while a WT-sized fragment was not produced with  $\Delta sll0729$  mutant DNA as template (Fig. 2B). SMRT sequencing of the  $\Delta sll0729$  mutant revealed a lack of GG $m^4$ CC methylation but no other differences compared to the WT DNA methylation status, indicating that *sll0729* encodes for a

cytosine instead of an adenine-specific methyltransferase. Similarly, bisulfite sequencing showed no methylation of GGCC in the  $\Delta sll0729$  mutant, all methylation signals with mutant DNA were below the threshold differentiating valid signals from background and experimental error (Supplementary Figs S2B and C). To verify this observation experimentally, chromosomal DNA from  $\Delta sll0729$  mutant cells was incubated with various methylation-sensitive restriction enzymes to evaluate changes in the methylation pattern. Enzymes known to be unable to cut modified GG<sup>m</sup>4CC motifs, such as *Hae*III, *Eae*I, and *Apa*I, could cut mutant DNA but not WT DNA (Fig. 2C). This result was supported by over-expression of *sll0729* in *E. coli*. Plasmid DNA from *E. coli* clones expressing *sll0729* were protected against the action of *Hae*III (Fig. 2D). Moreover, recombinant Sll0729 protein was purified by affinity chromatography via a fused His-tag and was found to catalyze significant DNA-specific methyltransferase activity in an *in vitro* enzyme assay (Fig. 4).

Altogether, these results indicate that *sll0729* encodes a cytosine-specific DNA-methyltransferase responsible for modifying the core sequence 5'-GG<sup>m</sup>4CC-3'. Thus, following the established nomenclature of REBASE<sup>®</sup>, this DNA-specific *Synechocystis* methyltransferase was named M.Ssp6803II (Table 1).

### 3.2.2. The G<sup>m</sup>6ATC motif is methylated by M.Ssp6803III, which is encoded by *slr1803*

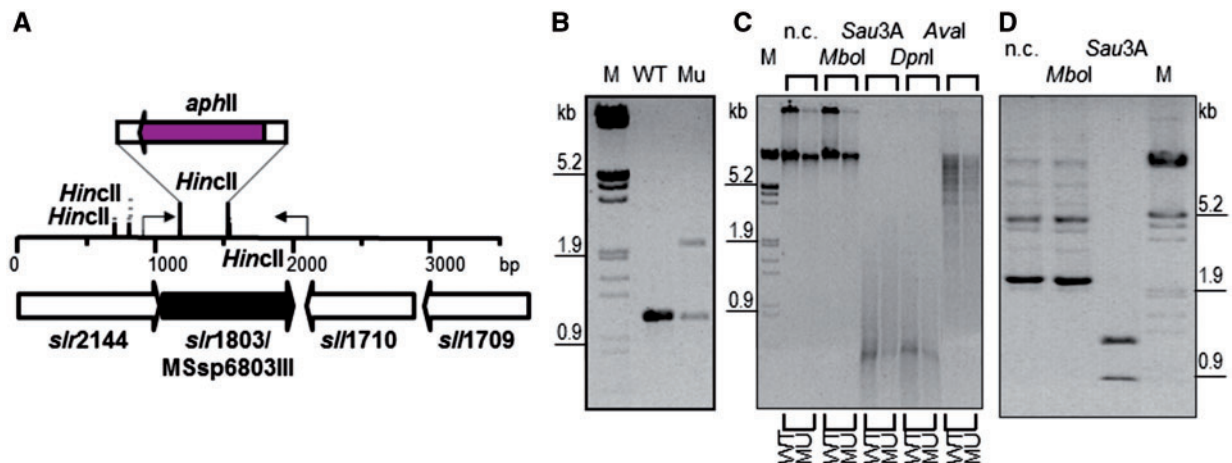
The protein sequence of the methyltransferase encoded by *slr1803* shows significant sequence similarities to Dam-like enzymes from many heterotrophic bacteria and forms one cluster with these enzymes (Supplementary Fig. S3). Furthermore, protein domain prediction at the NCBI Blast server showed that it also possesses a D12-class N6-adenine-specific DNA methyltransferase domain. The highest sequence similarity was found with the homolog from the cyanobacterium *Halotheca* sp. PCC 7418 (identity of 57%, BlastP *e*-value of  $1e^{-121}$ ). We found 290 highly similar proteins in other cyanobacteria (*e*-value  $\leq 5e^{-15}$ ). In addition, closely related proteins also exist in other Bacteria such as *Chloroflexi bacterium*

RBG\_16\_57\_9 (identity of 58%, BlastP *e*-value of  $5e^{-124}$ ) and in Archaea, such as *Methanobolus tindarius* DSM2278 (identity of 54%, BlastP *e*-value of  $2e^{-112}$ ). Among biochemically characterized enzymes, M.MboI from *Moraxella bovis*<sup>39</sup> was identified as the closest relative (identity of 41%, BlastP *e*-value of  $9e^{-63}$ ). A closely related enzyme was also reported from the filamentous cyanobacterium *Anabaena (Nostoc)* sp. PCC 7120.<sup>18</sup> These features strongly qualify Slr1803 as a candidate modifier of G<sup>m</sup>6ATC methylation motif, making *Synechocystis* DNA resistant against *Mbo*I restriction.

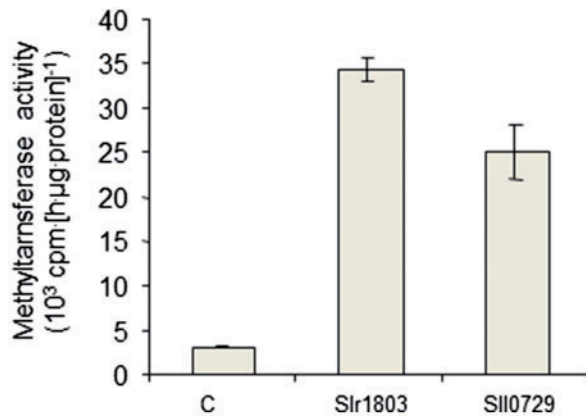
To verify the function of this putative N6-adenine methyltransferase, a deletion mutant of *slr1803* was generated; in this mutant, the internal *Hinc*II fragment of the ORF was replaced by an *aph*II gene (Fig. 3A). In addition to the mutated fragment, which is enlarged by the expected size of the inserted *aph*II gene, PCR analysis also detected the WT-sized PCR fragment from  $\Delta slr1803$  mutant DNA (Fig. 3B). The non-segregated status of the  $\Delta slr1803$  mutant could not be improved by cultivation at higher Km concentrations for many generations. This result indicates that the *slr1803* gene is essential for the viability of *Synechocystis* under our laboratory conditions. Consistent with the non-segregated genotype of the  $\Delta slr1803$  mutant, we observed no change in the methylation of the *Synechocystis* DNA, because *Mbo*I, which is Dam-methylation sensitive, did not cut DNA isolated from the mutant (Fig. 3C).

Since the methylation specificity of Slr1803 could not be verified using the  $\Delta slr1803$  mutant, the ORF was overexpressed in *E. coli*. Small amounts of recombinant protein of the expected size were found in crude extracts and could be isolated via the fused His-tag. Plasmid DNA from *E. coli* clones expressing *slr1803* was protected against the action of *Mbo*I (Fig. 3D) but could be still cut with *Sau*3A, which is not affected by N6-adenine methylation. Purified recombinant Slr1803 protein also showed significant DNA-specific methyltransferase activity in an *in vitro* enzyme assay (Fig. 4).

These results clearly indicate that *slr1803* encodes an N6-adenine-specific DNA methyltransferase modifying the core sequence 5'-G<sup>m</sup>6ATC-3'. Thus, this *Synechocystis* DNA-specific methyltransferase was named M.Ssp6803III (Table 1).



**Figure 3.** Functional verification of M.Ssp6803III encoded by *slr1803*. (A) Construction strategy for the generation of the  $\Delta slr1803$  mutant defective in the ORF of M.Ssp6803III. Thin arrows indicate primer binding sites, which were used to verify the genotype of the mutants. (B) Characterization of  $\Delta slr1803$  mutant genotype by PCR using gene-specific primers and as templates DNA from wild-type (WT) and mutant cells (Mu). (C) Separation of fragments generated during a restriction analysis with the N6-adenine methylation-sensitive restriction endonuclease (*Mbo*I) and enzymes (*Sau*3A, *Dpn*I, and *Ava*I), which are not sensitive to N6-adenine methylation, of chromosomal DNA of the wild-type (WT) and  $\Delta slr1803$  mutant (Mu) by agarose gel electrophoresis. (n.c., uncut control DNA; M, DNA size marker  $\lambda$ -DNA cut by *Hind*III and *Eco*RI). (D) Restriction analysis of plasmids, which were isolated from cells of the DNA-methylation negative *E. coli* strain TOP10 over-expressing M.Ssp6803III (Slr1803), using N6-adenine methylation-sensitive enzyme (*Mbo*I) and N6-adenine methylation-insensitive enzyme (*Sau*3A). (n.c., uncut control DNA; M, DNA size marker  $\lambda$ -DNA cut by *Hind*III and *Eco*RI).



**Figure 4.** *In vitro* methylation activities of M.Ssp6803II and M.Ssp6803III of *Synechocystis* sp. PCC 6803. Total DNA-specific methyltransferase activities towards non-methylated *Micrococcus* DNA was estimated with purified recombinant M.Ssp6803II (Sll0729) and M.Ssp6803III (Slr1803) using the NEB4-restriction buffer (C, control; incubation of DNA without added recombinant protein).

### 3.2.3. Two DNA methyltransferases are localized on the plasmid pSYSX

The methylome of *Synechocystis* includes two additional methylation motifs, GA<sup>m6</sup>AGGC and GG<sup>m6</sup>AN<sub>7</sub>TTGG/CCA<sup>m6</sup>AN<sub>7</sub>TCC that are not modified by the methyltransferases encoded on the chromosome. BlastP searches using proteins encoded on the *Synechocystis* plasmids revealed three additional genes for putative DNA-specific methyltransferases.

On the pSYSX plasmid, we found *slr6050* annotated to encode a hypothetical protein of 1100 amino acids. BlastP searches revealed that very similar proteins exist in many different bacteria but only in four other cyanobacteria (*Microcystis aeruginosa* spp. PCC 9701 and 9806, *Synechococcus* sp. PCC 73109, and *Prochlorothrix hollandica*). To verify the function of this putative N6-adenine methyltransferase, a deletion mutant of *slr6050* was generated; in this mutant, the internal *Cla*I fragment of the ORF was replaced by an *aphII* gene (Fig. 5A). Since the deleted *slr6050* fragment had approximately the same size as the inserted *aphII* gene, the PCR reaction using flanking primers (*slr6050*fw and *slr6050*rev, Supplementary Table S2) did not allow to judge whether or not the mutant was fully segregated. Alternatively, we used a primer pair (*slr6050\_i\_fw* and *slr6050\_i\_rev*, Supplementary Table S2), which binds to sequences located inside the deleted *Cla*I fragment. This PCR-detected DNA of same size using DNA from WT as well as the *slr6050* mutant (Fig. 5B), which clearly indicated that the mutant was not fully segregated. The *aphII* gene was detected with primers *aphII\_fw* and *aphII\_rev* only with DNA of the mutant (Fig. 5C). The non-segregated status of the  $\Delta$ *slr6050* mutant could not be improved by cultivation at higher Km concentrations for many generations. This result indicates that the *slr6050* gene is essential for the viability of *Synechocystis* under our laboratory conditions.

Comparison of Slr6050 against the Pfam database returned a single hit, the Eco57I bifunctional RM methylase, a type IV RM enzyme.<sup>40</sup> The Eco57I domain for AdoMet-dependent enzymes (pfam07669) is clearly present in the sequence of Slr6050 and weak similarity to the HSDR\_N restriction endonuclease domain could also be detected. Eco57I is sensitive to the methylation of CTGA<sup>m6</sup>AG or CTTC<sup>m6</sup>AG<sup>40</sup> making Slr6050 a likely candidate for

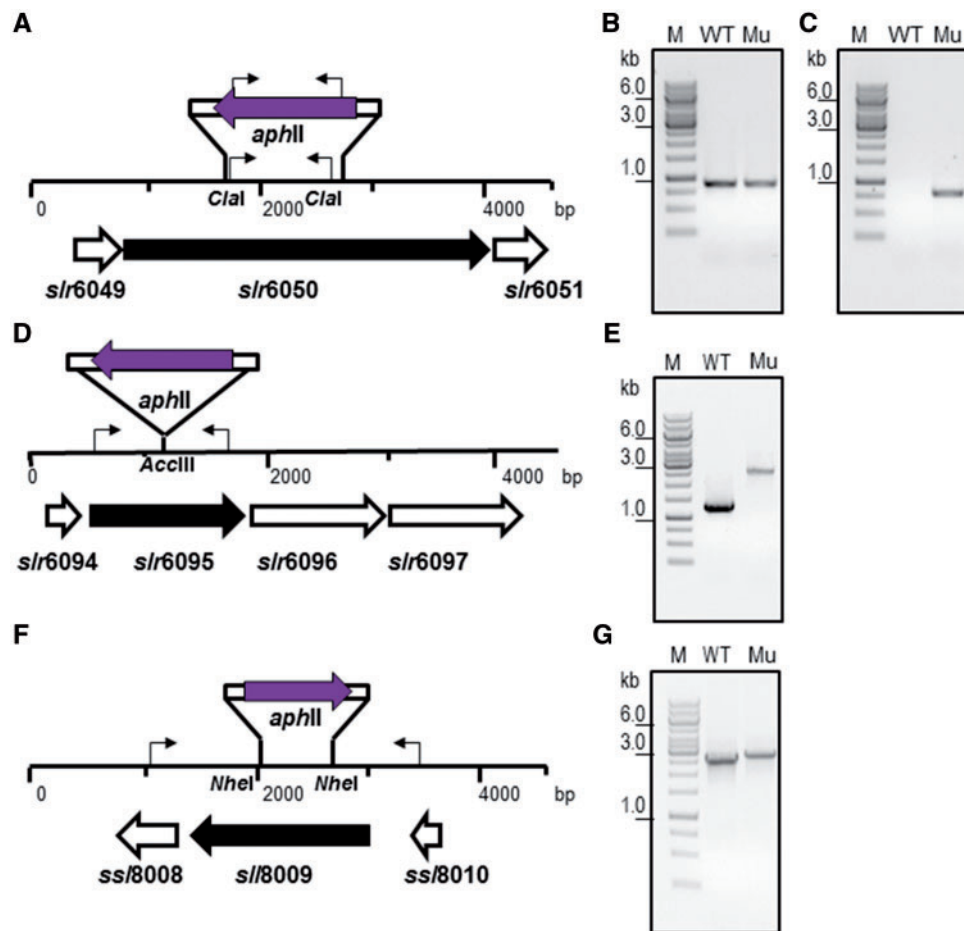
the modification of GA<sup>m6</sup>AGGC, which was detected by SMRT sequencing of *Synechocystis* DNA (Fig. 1). Unfortunately, all our attempts to obtain recombinant Slr6050 protein after expression of a codon-optimized *slr6050* gene in *E. coli* failed; hence, we could not directly verify its function as DNA-specific methyltransferase and the proposed specificity.

Therefore, we only tentatively annotate this putative DNA-specific methyltransferase to be responsible for the modification of GA<sup>m6</sup>AGGC site of *Synechocystis* and name it M.Ssp6803IV (Table 1).

Another methyltransferase candidate was found on plasmid pSYSX (*slr6095*), which shows similarities to DNA-specific methyltransferases of type I RM systems. Thus, this enzyme could be responsible for the modification of the remaining motif GG<sup>m6</sup>AN<sub>7</sub>TTGG/CCA<sup>m6</sup>AN<sub>7</sub>TCC. BlastP searches against the NCBI database revealed that 487 highly similar proteins are harbored in other cyanobacteria (*e*-value  $\leq 5e^{-15}$ ). To study the function of Slr6095, a mutant was generated in which the *aphII* gene, conferring Km resistance, was inserted into the coding sequence of *slr6095* at the single *Acc*III site (Fig. 5D). Genotypic analysis revealed that a completely segregated  $\Delta$ *slr6095* mutant was obtained since only the fragment enlarged by the expected size of the inserted *aphII* gene was amplified by PCR, while a WT-sized fragment was not produced with  $\Delta$ *slr6095* mutant DNA as template (Fig. 5E). SMRT sequencing of the  $\Delta$ *slr6095* mutant revealed a lack of GG<sup>m6</sup>AN<sub>7</sub>TTGG/CCA<sup>m6</sup>AN<sub>7</sub>TCC methylation but no other differences compared to the WT DNA methylation status, indicating that *slr6095* encodes the corresponding type I adenine-specific methyltransferase. Accordingly, this putative DNA-specific methyltransferase of *Synechocystis* was named M.Ssp6803V (Table 1).

### 3.2.4. The DNA methyltransferases on plasmid pSYSG is not active

Finally, the genes *sll8009* (M subunit), *ssl8010*, *sll8006* (S subunit), and *sll8049* (R subunit) on the pSYSG plasmid are annotated in CyanoBase<sup>41</sup> to encode all subunits required for a complete RM system. Similar proteins are encoded in about 50 other cyanobacterial genomes. However, after protein sequence comparisons we noticed that *sll8009* appears to encode a methyltransferase that is N-terminally truncated by 156 amino acid residues. The missing sequence is present in *ssl8010* and the sequence between *ssl8010* and *sll8009*. Because both genes are in the same reading frame, a point mutation leading to a TAA stop codon might have led to its inactivation, consistent with a frameshift mutation in *sll8049*. The entire region comprising *sll8009* and *ssl8010* was amplified and re-sequenced. This analysis confirmed the DNA sequence displayed in CyanoBase<sup>41</sup> and the truncated nature of Sll8009. Nevertheless, to study the function of the putative methyltransferase Sll8009, a mutant was generated in which the *aphII* gene, conferring Km resistance, was inserted into a deletion of an internal *Nbe*I fragment of the coding sequence of *sll8009* (Fig. 5F). Genotypic analysis revealed that a completely segregated  $\Delta$ *sll8009* mutant was obtained since only the fragment enlarged by the expected size of the inserted *aphII* gene was amplified by PCR, while a WT-sized fragment was not produced with  $\Delta$ *sll8009* mutant DNA as template (Fig. 5G). However, SMRT sequencing of the  $\Delta$ *sll8009* mutant revealed no differences compared to the WT DNA methylation status, indicating that *sll8009* is not involved in the methylation of *Synechocystis* DNA, because it most likely encodes an enzymatically non-active methyltransferase.



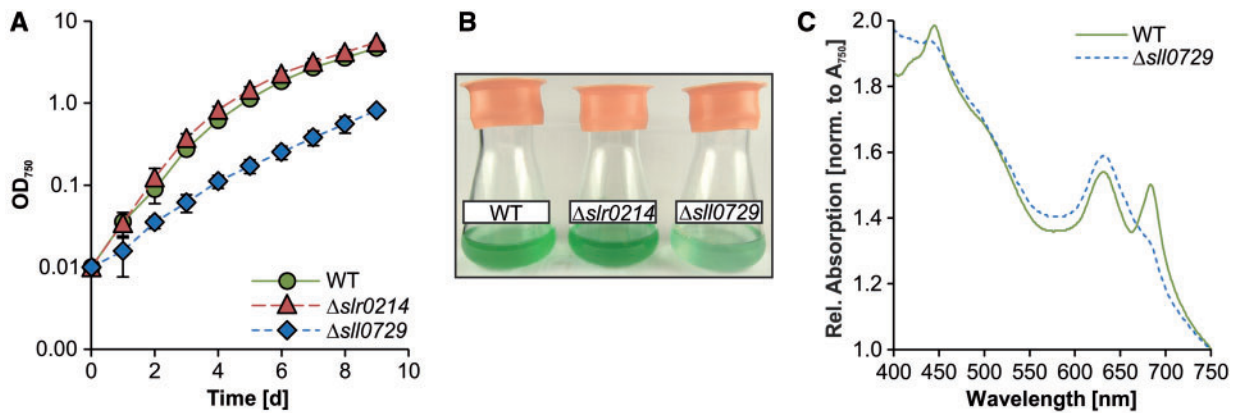
**Figure 5.** Mutation of plasmid-located methyltransferase genes. (A) Construction strategy for the generation of the  $\Delta$ *slr6050* mutant defective in the ORF of M.Ssp6803IV. Thin arrows indicate primer binding sites, which were used to verify the genotype of the mutants. (B) Separation of PCR fragments using primers *slr6050\_i\_fw* and *Slr6050\_i\_rev* and DNA from wild-type (WT) or mutant cells (Mu) verifying the non-segregated genotype. (C) Separation of PCR fragments using primers *aphII\_fw* and *aphII\_rev* verifying the insertion of the Km-resistance cartridge. (D) Construction strategy for the generation of the  $\Delta$ *slr6095* mutant defective in the ORF of M.Ssp6803V. Thin arrows indicate primer binding sites, which were used to verify the genotype of the mutants. (E) Characterization of the fully segregated  $\Delta$ *slr6095* mutant genotype by PCR using gene-specific primers and as template DNA from WT or Mu cells. (F) Construction strategy for the generation of the  $\Delta$ *ssl8009* mutant. Thin arrows indicate primer binding sites used for genotype verification. (G) Characterization of the  $\Delta$ *ssl8009* mutant genotype by PCR using gene-specific primers and WT or Mu cell template DNA.

### 3.3. Physiological characterization of mutants defective in DNA methyltransferases

To gain insight into possible physiological functions of DNA methylation in *Synechocystis*, mutants defective in these methyltransferase-encoding genes were studied. No clear phenotypical alterations, in comparison to WT, were observed for the  $\Delta$ *slr0214*,  $\Delta$ *slr6095* and the partially segregated  $\Delta$ *slr1803* and  $\Delta$ *slr6050* mutants. Because completely segregated  $\Delta$ *slr1803* and  $\Delta$ *slr6050* mutants could not be obtained (Figs 3B and 5), M.Ssp6803III and M.Ssp6803IV play essential roles for cell viability. Interestingly, cultivation under identical conditions indicated a very severe growth deficiency of mutant  $\Delta$ *ssl0729*, whereas the mutant  $\Delta$ *slr0214* grew like WT under these conditions. Moreover, a bluish appearance due to changed pigment composition was characteristic for this mutant compared to WT and mutant  $\Delta$ *slr0214* over the entire cultivation time (Fig. 6). Whole-cell absorbance spectra clearly indicated that the  $\Delta$ *ssl0729* cells contained a reduced amount of chlorophyll a (represented by the peaks at 440 and 680nm), while the content of phycocyanin (represented by the

peak at 630 nm) was not changed (Fig. 6C). The decreased chlorophyll content most probably reflects reduced photosynthetic capacity, which corresponds to the diminished growth of mutant  $\Delta$ *ssl0729*.

According to transcriptomic data available for *Synechocystis*,<sup>42</sup> the *ssl0729* gene potentially comprises an operon with two adjacent genes (Fig. 2A), the upstream located *ssl0728* (*accA*) gene encoding the acetyl-CoA carboxylase alpha subunit and the downstream-located *ssl1378* gene encoding a small hypothetical protein. To rule out polar effects on the expression of the downstream gene, we generated complementation strains in which *ssl0729* or *ssl1378* were ectopically expressed. The expression of intact *ssl0729* fully reversed the phenotype back to WT-like growth and pigmentation (Supplementary Fig. S4), whereas  $\Delta$ *ssl0729* mutant cells expressing *ssl1378* did not change the phenotype compared to the original  $\Delta$ *ssl0729* mutant. These experiments clearly show that defects in the DNA-specific methyltransferase M.Ssp6803II encoded by *ssl0729* result in strong physiological defects; hence, this enzyme seems to play an important role in *Synechocystis*.



**Figure 6.** Phenotype profiling of mutants  $\Delta slr0214$  and  $\Delta slI0729$  defective in M.Ssp6803I and M.Ssp6803II. (A) Growth curves of the WT and the mutants  $\Delta slr0214$  and  $\Delta slI0729$ . (B) Phenotypic appearance of liquid cultures. (C) Whole cell absorbance spectra of the WT and the mutant  $\Delta slI0729$ .

#### 4. Discussion

*Synechocystis* possesses at least five different methylation activities toward specific DNA sequences, which were detected using SMRT and bisulfite sequencing. The three DNA methyltransferases encoded on the chromosome seem to belong to the type II methyltransferase group since they modify bases in short palindromic sequences (M.Ssp6803I-III), whereas M.Ssp6803V is a type I enzyme that modifies a larger bi-partite motif (see Fig. 1). The M.Ssp6803IV is not modifying a palindromic sequence and, based on its similarity to Eco571 RM enzymes, qualifies as a type IV enzyme.<sup>40</sup> The occurrence of five methylated motifs and five methyltransferase-encoding genes is similar to other bacteria. A recent study of the epigenetic landscape of prokaryotes revealed that only a few genomes are not methylated, whereas others contain multiple different motifs, up to 19.<sup>16</sup> This study included the cyanobacteria *Leptolyngbya* sp. PCC 6406 and *Mastigocladopsis repens* PCC 10914, with 12 and 2 modified motifs, respectively. A survey of palindromic sequences and their putative modifying methyltransferases identified several types among cyanobacteria.<sup>43</sup> Particularly widespread among cyanobacteria are the palindromic sequences GCGATCGC (highly iterated palindrome 1, HIP1),<sup>44,45</sup> GGCC, and GATC. These three sequences contain the motifs for the DNA methyltransferases M.Ssp6803I-III, which are encoded on the chromosome of *Synechocystis* (this work and [20]), as well as related enzymes in *Anabaena* (*Nostoc*) sp. PCC 7120.<sup>43</sup> It is very likely that the proteins of high similarities encoded in many other cyanobacterial genomes show identical methylation specificities.

The previously identified C5-cytosine-specific enzyme M.Ssp6803I modifies the core motif within the HIP1 sequences.<sup>20</sup> Despite the frequent occurrence of methylated HIP1 sequences, their methylation by M.Ssp6803I seems to have no significant impact on the physiology of *Synechocystis* under laboratory conditions, because in previous work only slight growth retardation was observed for the corresponding *Synechocystis* mutant<sup>20</sup> and the mutant defective in the ortholog of *Anabaena* (*Nostoc*) sp. PCC 7120.<sup>18</sup> Similarly, the expression of the gene for M.Ssp6803I in tobacco chloroplasts led to the methylation of the plastome DNA, but the transplastomic lines showed no alterations in plastid gene expression and were phenotypically indistinguishable from wild-type plants.<sup>46</sup> The minor effects of the mutation of the gene for M.Ssp6803I on cell viability and gene expression contrast its widespread occurrence among cyanobacteria. However, the close correlation between the presence of this methyltransferase and the occurrence of HIP1 sequences has led to a model

wherein M.Ssp6803I could be involved in the methylation-directed mismatch repair of DNA, which is potentially of high importance for cyanobacteria exposed to strong light intensities, including UV.<sup>43</sup>

The enzyme M.Ssp6803III modifies adenine in the sequence 5'-GATC-3', which is an internal part of the HIP1 sequence but often also occurs separately. Correspondingly, genes for methyltransferases modifying G<sup>m5</sup>CGATCGC and G<sup>m6</sup>ATC are usually co-occurring in the genomes of cyanobacteria.<sup>43</sup> The activity of the Dam-like enzyme M.ssp6803III was clearly proven by the inhibition of *MboI* cleavage in *E. coli* cells expressing this *Synechocystis* gene. Its activity is also sufficiently high to modify virtually all GATC sites in the *Synechocystis* DNA, which is completely resistant against *MboI* treatments. Similar results were obtained when *slr1803* was expressed in tobacco plastids.<sup>46</sup> The M.Ssp6803III (Slr1803) appears homologous to related Dam-like enzymes from many bacteria (Supplementary Fig. S3). It is well known that Dam methylation has many physiological functions such as in the initiation of DNA replication, nucleoid segregation, post-replicative DNA mismatch repair, and gene expression regulation.<sup>9,47</sup> The Dam-like protein M.Ssp6803III plays an essential role in *Synechocystis* since our attempts to generate a null mutant were not successful and led only to a partial gene replacement. Similar results were reported for the ortholog in *Anabaena* (*Nostoc*) sp. PCC 7120.<sup>18</sup> In contrast, *dam* mutants were obtained for *E. coli* that remained viable under standard conditions but showed an increased mutation rate (reviewed in [47]), whereas Dam methylation is essential for viability in *Vibrio cholerae*.<sup>48</sup> The different dependence of these bacteria on Dam could be explained by differences in the mode of chromosomal replication. In chromosome II of *V. cholerae*, the origin for DNA replication (*oriC*) is different from that of *E. coli*. It replicates in a DnaA-independent manner but was found to strictly depend on the methylation by Dam.<sup>48-50</sup> The molecular details of DNA replication in *Synechocystis* are less well understood, but it has been shown that DnaA is not essential for the initiation of DNA replication.<sup>51,52</sup> Thus, it might be possible that the Dam-dependent DNA methylation is essential for the mode of DNA replication in *Synechocystis* similar to the case of *V. cholerae*.

The DNA-methyltransferase M.Ssp6803II modifies the *HaeIII* recognition sequence 5'-GGCC-3'. Its recognition site was verified by screening *HaeIII*-resistant plasmids in a *Synechocystis* gene library, by mutation and overexpression of the ORF *slI0729*. Moreover, bisulfite sequencing revealed that M.Ssp6803II is specific for N4-cytosine leading to 5'-GG<sup>m4</sup>CC-3' and modifies at least 90% of the



recognition sequence. However, the protein shows structural features that are different from the well-conserved C5-cytosine-specific DNA methyltransferases, including M.HaeIII. Instead of cytosine-specific enzymes, the most similar proteins all belong to the group of N6-adenine-specific DNA methyltransferases, which is documented by the phylogenetic analysis (Supplementary Fig. S3). Structural and sequence comparisons of cytosine- and adenine-specific enzymes revealed that most of the conserved motifs are shared by both enzyme classes, only the organization of these conserved motifs and minor sequence differences within seem to determine whether an enzyme is specific for cytosine or adenine.<sup>53</sup> Correspondingly, a sequence alignment of M.Ssp6803II with several previously characterized Dam-like sequences revealed distinct sequence differences between the cytosine-specific and the adenine-specific enzymes (Supplementary Fig. S5).

The DmtB enzyme in *Anabaena* (*Nostoc*) sp. PCC 7120 shows similar functional and structural features to M.Ssp6803II, which included the N4-methylation of the first cytosine leading to the inhibition of *HaeIII* restriction activity.<sup>18</sup> The deletion of M.Ssp6803II ( $\Delta$ *slI0729* mutant) led to a strong phenotype and the mutant could only be maintained at conditions permitting slow growth. Hence, the modification of the *HaeIII* recognition sequence is important for the performance of *Synechocystis* under conditions promoting high growth rates. However, further experiments are needed to identify the primary cause of this strong phenotypic alteration. We hypothesize that the absence of GGCC methylation could either have a broad impact on gene expression or the coordination of DNA replication with cell propagation.

Moreover, we analyzed three additional DNA methyltransferases in *Synechocystis*, two of which modify sequence motifs that have not been previously detected among cyanobacteria. Albeit lacking genetic and biochemical evidence, we conclude that M.Ssp6803IV, which is encoded by the *slr6050* gene on the plasmid pSYSX, is likely responsible for the modification of GA<sup>m6</sup>AGGC motifs. This assumption is supported by the REBASE database since a search using this recognition sequence revealed the Slr6050 protein as M.Ssp6803IV and by its similarity to Eco57I with a closely related recognition sequence CTGA<sup>m6</sup>AG.<sup>40</sup> Moreover, it is also supported by the elimination of other possible candidate enzymes, since with M.Ssp6803V we identified the enzyme modifying the only remaining, bipartite motif GGAN<sub>7</sub>TTGG/CCAAN<sub>7</sub>TCC. M.Ssp6803V is encoded by the *slr6095* gene on plasmid pSYSX. It is annotated as part of a type I RM system.<sup>41</sup> Nevertheless, the restriction system appears not to be active in *Synechocystis*, since we obtained a mutant lacking the respective methylation. Furthermore, we characterized the *slI8009* gene, which is also annotated to encode a methyltransferase of a type I RM system.<sup>41</sup> However, a closer inspection of this DNA locus indicated that the gene encodes a truncated, inactive protein. Correspondingly, the mutation of *slI8009* has no impact on DNA methylation of *Synechocystis*. Taking this into account, M.Ssp6803IV is the only remaining DNA methyltransferase for GA<sup>m6</sup>AGGC modification.

Altogether, this study provides the first comprehensive methylome analysis of the cyanobacterial model strain *Synechocystis* sp. PCC 6803. Moreover, it can be regarded as the groundwork for systematic analyses of the possible impact and molecular mechanisms linking methyltransferase activities and particular phenotypes in cyanobacteria. Particularly, the detected ~10% unmethylated CGATCG and GGCC sites in the bisulfite sequencing analyses constitute a solid basis for further detailed analyses of their functional relevance.

## Acknowledgements

We thank Richard J. Roberts (New England Biolabs) for helpful discussion. Klaudia Michl and Viktoria Reimann are acknowledged for technical assistance.

## Accession numbers

All bisulfite and SMRT sequencing raw data were uploaded to the databases at the National Center for Biotechnology Information (BioProject ID: PRJNA430784, BioSample: SAMN08378604, and SRA: SRS2844079).

## Funding

This study was funded by the German Research Foundation (Deutsche Forschungsgemeinschaft) via a joint grant to MH (HA2002/17-1) and WRH (HE 2544/10-1).

## Conflict of interest

None declared.

## Supplementary data

Supplementary data are available at DNARES online.

## References

1. D'Urso, A. and Brickner, J. H. 2014, Mechanisms of epigenetic memory, *Trends Genet.*, **30**, 230–6.
2. Vidalis, A., Živković, D., Wardenaar, R., Roquis, D., Tellier, A. and Johannes, F. 2016, Methylome evolution in plants, *Genome Biol.*, **17**, 264.
3. Noyer-Weidner, M. and Trautner, T. A. 1993, *Methylation of DNA in prokaryotes*, Birkhäuser: Basel.
4. Tock, M. R. and Dryden, D. T. 2005, The biology of restriction and anti-restriction, *Curr. Opin. Microbiol.*, **8**, 466–72.
5. Vanyushin, B. F., Tkacheva, S. G. and Belozersky, A. N. 1970, Rare bases in animal DNA, *Nature*, **225**, 948–9.
6. Greer, E. L., Blanco, M. A. and Gu, L. 2015, DNA methylation on N6-adenine in *C. elegans*, *Cell*, **161**, 868–78.
7. Fu, Y., Luo, G. Z., Chen, K., et al. 2015, N6-methyldeoxyadenosine marks active transcription start sites in *Chlamydomonas*, *Cell*, **161**, 879–92.
8. Zhang, G., Huang, H., Liu, D., et al. 2015, N6-methyladenine DNA modification in *Drosophila*, *Cell*, **161**, 893–906.
9. Casadesús, J. 2016, Bacterial DNA methylation and methylomes. In: Jeltsch, A., Jurkowska, R. Z. (eds) *DNA Methyltransferases – Role and Function*, *Adv. Exp. Med. Biol.*, **945**, 35–61. Vol. Springer International, Switzerland, p.
10. Erova, T. E., Kosykh, V. G., Sha, J. and Chopra, A. K. 2012, DNA adenine methyltransferase (Dam) controls the expression of the cytotoxic enterotoxin (*act*) gene of *Aeromonas hydrophila* via tRNA modifying enzyme-glucose-inhibited division protein (GidA), *Gene*, **498**, 280–7.
11. Braaten, B. A., Nou, X., Kaltenbach, L. S. and Low, D. A. 1994, Methylation patterns in *pap* regulatory DNA control pyelonephritis-associated pili phase variation in *E. coli*, *Cell*, **76**, 577–88.
12. Nou, X., Braaten, B., Kaltenbach, L. and Low, D. A. 1995, Differential binding of Lrp to two sets of *pap* DNA binding sites mediated by Pap I regulates Pap phase variation in *Escherichia coli*, *EMBO J.*, **14**, 5785–97.
13. Militello, K. T., Simon, R. D., Qureshi, M., et al. 2012, Conservation of Dcm-mediated cytosine DNA methylation in *Escherichia coli*, *FEMS Microbiol. Lett.*, **328**, 78–85.

14. Fioravanti, A., Fumeaux, C., Mohapatra, S. S., et al. 2013, DNA binding of the cell cycle transcriptional regulator GcrA depends on N6-adenosine methylation in *Caulobacter crescentus* and other alphaproteobacteria, *PLoS Genet.*, **9**, e1003541.
15. Lluch-Senar, M., Luong, K., Lloréns-Rico, V., et al. 2012, Comprehensive methylome characterization of *Mycoplasma genitalium* and *Mycoplasma pneumoniae* at single-base resolution, *PLoS Genet.*, **9**, e1003191.
16. Blow, M. J., Clark, T. A., Daum, C. G., et al. 2016, The epigenomic landscape of prokaryotes, *PLoS Genet.*, **12**, e1005854.
17. Hagemann, M. and Hess, W. R. 2018, Systems and synthetic biology for the biotechnological application of cyanobacteria, *Curr. Opin. Biotechnol.*, **49**, 94–9.
18. Matveyev, A. V., Young, K. T., Meng, A. and Elhai, J. 2001, DNA methyltransferases of the cyanobacterium *Anabaena* PCC 7120, *Nucleic Acids Res.*, **29**, 1491–506.
19. Zehr, J. P., Ohki, K., Fujita, Y. and Landry, D. 1991, Unique modification of adenine in genomic DNA of the marine cyanobacterium *Trichodesmium* sp. strain NIBB 1067, *J. Bacteriol.*, **173**, 7059–62.
20. Scharnagl, M., Richter, S. and Hagemann, M. 1998, The cyanobacterium *Synechocystis* sp. strain PCC 6803 expresses a DNA methyltransferase specific for the recognition sequence of the restriction endonuclease *PvuII*, *J. Bacteriol.*, **180**, 4116–22.
21. Wang, B., Yu, J., Zhang, W. and Meldrum, D. R. 2015, Premethylation of foreign DNA improves integrative transformation efficiency in *Synechocystis* sp. strain PCC 6803, *Appl. Environ. Microbiol.*, **81**, 8500–6.
22. Trautmann, D., Voss, B., Wilde, A., Al-Babili, S. and Hess, W. R. 2012, Microevolution in cyanobacteria: re-sequencing a motile substrain of *Synechocystis* sp. PCC 6803, *DNA Res.*, **19**, 435–48.
23. Rippka, R., Deruelles, J., Waterbury, J. B., Herdman, M. and Stanier, R. Y. 1979, Generic assignments, strain histories and properties of pure cultures of cyanobacteria, *J. Gen. Microbiol.*, **111**, 1–16.
24. Wilson, K. 2001, Preparation of genomic DNA from bacteria, *Curr. Prot. Molec. Biol.*, 00:1:2.4:2.4.1-2.4.5; <http://dx.doi.org/10.1002/0471142727.mb0204s56> (1 February 2018, date last accessed).
25. Bolger, A. M., Lohse, M. and Usadel, B. 2014, Trimmomatic: a flexible trimmer for Illumina sequence data, *Bioinformatics*, **30**, 2114–20.
26. Krueger, F. and Andrews, S. R. 2011, Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. Krueger F1, Andrews SR, *Bioinformatics*, **27**, 1571–2.
27. Langmead, B. and Salzberg, S. 2012, Fast gapped-read alignment with Bowtie 2, *Nat. Methods*, **9**, 357–9.
28. Sambrook, J., Fritsch, E. F. and Maniatis, T. 1989, *Molecular cloning*. Cold Spring Harbour Laboratory Press: Cold Spring Harbor, NY.
29. Kaneko, T., Sato, S., Kotani, H., et al. 1996, Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp. strain PCC 6803. II. Sequence determination of the entire genome and assignment of potential protein-coding regions, *DNA Res.*, **3**, 109–36.
30. Zinchenko, V. V., Piven, I. V., Melnik, V. A. and Shestakov, S. V. 1999, Vectors for the complementation analysis of cyanobacterial mutants, *Russ. J. Genet.*, **35**, 228–32.
31. Lowry, O. H., Rosenbrough, N. J., Farr, A. L. and Randall, R. J. 1951, Protein measurement with the folin phenol reagent, *J. Biol. Chem.*, **193**, 265–75.
32. Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., et al. 1997, Gapped BLAST and PSI-BLAST: a new generation of protein database search programs, *Nucl. Acid. Res.*, **25**, 3389–402.
33. Altschul, S. F., Wootton, J. C., Gertz, E. M., et al. 2005, Protein database searches using compositionally adjusted substitution matrices, *FEBS J.*, **272**, 5101–9.
34. Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., et al. 2007, ClustalW and ClustalX version 2, *Bioinformatics*, **23**, 2947–8.
35. Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M. and Kumar, S. 2011, MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods, *Mol. Biol. Evol.*, **28**, 2731–9.
36. Murray, I. A., Clark, T. A., Morgan, R. D., et al. 2012, The methylomes of six bacteria, *Nucleic Acids Res.*, **40**, 11450–114362.
37. Vilkaitis, G. and Klimašauskas, S. 1999, Bisulfite sequencing protocol displays both 5-methylcytosine and N4-methylcytosine, *Anal. Biochem.*, **271**, 116–9.
38. Roberts, R. J., Vincze, T., Posfai, J. and Macelis, D. 2010, REBASE—a database for DNA restriction and modification: enzymes, genes and genomes, *Nucl. Acid. Res.*, **38** (Database issue), D234–6.
39. Ueno, T., Ito, H., Kimizuka, F., Kotani, H. and Nakajima, K. 1993, Gene structure and expression of *MboI* restriction-modification system, *Nucleic Acids Res.*, **21**, 2309–13.
40. Janulaitis, A., Vaisvila, R., Timinskas, A., Klimasauskas, S. and Butkus, V. 1992, Cloning and sequence analysis of the genes coding for Eco57I type IV restriction-modification enzymes, *Nucl. Acid. Res.*, **20**, 6051–6.
41. Nakao, M., Okamoto, S., Kohara, M. and Fujishiro, T. 2010, CyanoBase: the cyanobacteria genome database update 2010, *Nucl. Acid. Res.*, **38** (Database issue), D379–81.
42. Kopf, M., Klähn, S., Scholz, I., Matthiessen, J. K., Hess, W. R. and Voß, B. 2014, Comparative analysis of the primary transcriptome of *Synechocystis* sp. PCC 6803, *DNA Res.*, **21**, 527–39.
43. Elhai, J. 2015, Highly iterated palindromic sequences (HIPs) and their relationship to DNA methyltransferases, *Life (Basel)*, **5**, 921–48.
44. Robinson, N. J., Robinson, P. J., Gupta, A., Bleasby, A. J., Whitton, B. A. and Morby, A. P. 1995, Singular over-representation of an octameric palindrome, HIP1, in DNA from many cyanobacteria, *Nucl. Acid. Res.*, **23**, 729–35.
45. Robinson, P. J., Cranenburgh, R. M., Head, I. M. and Robinson, N. J. 1997, HIP1 propagates in cyanobacterial DNA via nucleotide substitutions but promotes excision at similar frequencies in *Escherichia coli* and *Synechococcus* PCC 7942, *Mol. Microbiol.*, **24**, 181–9.
46. Ahlert, D., Stegemann, S., Kahlau, S., Ruf, S. and Bock, R. 2009, Insensitivity of chloroplast gene expression to DNA methylation, *Mol. Genet. Genomics*, **282**, 17–24.
47. Adhikari, S. and Curtis, P. D. 2016, DNA methyltransferases and epigenetic regulation in bacteria, *FEMS Microbiol. Rev.*, **40**, 575–91.
48. Demarre, G., Chattoraj, D. K. and Burkholder, W. F. 2010, DNA adenine methylation is required to replicate both *Vibrio cholerae* chromosomes once per cell cycle, *PLoS Genet.*, **6**, e1000939.
49. Egan, E. S. and Waldor, M. K. 2003, Distinct replication requirements for the two *Vibrio cholerae* chromosomes, *Cell*, **114**, 521–30.
50. Val, M.-E., Kennedy, S. P., Soler-Bistué, A. J., et al. 2014, Fuse or die: how to survive the loss of Dam in *Vibrio cholerae*, *Mol. Microbiol.*, **91**, 665–72.
51. Richter, S., Hagemann, M. and Messer, W. 1998, Transcriptional analysis and mutation of a *dnaA*-like gene in *Synechocystis* sp. strain PCC 6803, *J. Bacteriol.*, **180**, 4946–9.
52. Ohbayashi, R., Watanabe, S., Ehira, S., Kanesaki, Y., Chibazakura, T. and Yoshikawa, H. 2016, Diversification of DnaA dependency for DNA replication in cyanobacterial evolution, *Isme J.*, **10**, 1113–21.
53. Malone, T., Blumenthal, R. M. and Cheng, X. 1995, Structure-guided analyses reveals nine sequence motifs conserved among DNA amino-methyltransferases, and suggests a catalytic mechanism for these enzymes, *J. Mol. Biol.*, **253**, 618–32.
54. Carver, T., Thomson, N., Bleasby, A., Berriman, M. and Parkhill, J. 2009, DNAPlotter: circular and linear interactive genome visualization, *Bioinformatics*, **25**, 119–20.