4150–4162 *Nucleic Acids Research, 2015, Vol. 43, No. 8* *Published online 6 April 2015*
doi: 10.1093/nar/gkv219

Specificity of the ModA11, ModA12 and ModD1 epigenetic regulator N^6-adenine DNA methyltransferases of *Neisseria meningitidis*

Kate L. Seib[1,†], Freda E.-C. Jen[1,†], Aimee Tan[1], Adeana L. Scott[1], Ritesh Kumar[2], Peter M. Power[1], Li-Tzu Chen[3], Hsing-Ju Wu[4], Andrew H.-J. Wang[3], Dorothea M. C. Hill[5], Yvette A. Luyten[6], Richard D. Morgan[6], Richard J. Roberts[6], Martin C. J. Maiden[5], Matthew Boitano[7], Tyson A. Clark[7], Jonas Korlach[7], Desirazu N. Rao[2] and Michael P. Jennings[1,*]

[1]Institute for Glycomics, Griffith University, Gold Coast, Queensland 4222, Australia, [2]Department of Biochemistry, Indian Institute of Science, Bangalore, India, [3]Institute of Biological Chemistry, Academia Sinica, Taipei 115, Taiwan, [4]Center for Molecular Medicine, China Medical University Hospital, China Medical University, Taichung 40402, Taiwan, [5]Department of Zoology, University of Oxford, Oxford, UK, [6]New England Biolabs, Ipswich, MA 01938, USA and [7]Pacific Biosciences, Menlo Park, CA 94025, USA

Received January 14, 2015; Revised February 12, 2015; Accepted March 03, 2015

ABSTRACT

Phase variation (random ON/OFF switching) of gene expression is a common feature of host-adapted pathogenic bacteria. Phase variably expressed N^6-adenine DNA methyltransferases (Mod) alter global methylation patterns resulting in changes in gene expression. These systems constitute phase variable regulons called phasevarions. *Neisseria meningitidis* phasevarions regulate genes including virulence factors and vaccine candidates, and alter phenotypes including antibiotic resistance. The target site recognized by these Type III N^6-adenine DNA methyltransferases is not known. Single molecule, real-time (SMRT) methylome analysis was used to identify the recognition site for three key *N. meningitidis* methyltransferases: ModA11 (exemplified by M.NmeMC58I) (5′-CGY^{m6}AG-3′), ModA12 (exemplified by M.Nme77I, M.Nme18I and M.Nme579II) (5′-AC^{m6}ACC-3′) and ModD1 (exemplified by M.Nme579I) (5′-CC^{m6}AGC-3′). Restriction inhibition assays and mutagenesis confirmed the SMRT methylome analysis. The ModA11 site is complex and atypical and is dependent on the type of pyrimidine at the central position, in combination with the bases flanking the core recognition sequence 5′-CGY^{m6}AG-3′. The observed efficiency of methylation in the *modA11* strain (MC58) genome ranged from 4.6% at 5′-GCGC^{m6}AGG-3′ sites, to 100% at 5′-ACGT^{m6}AGG-3′ sites. Analysis of the distribution of modified sites in the respective genomes shows many cases of association with intergenic regions of genes with altered expression due to phasevarion switching.

INTRODUCTION

Phase variation, the high frequency ON/OFF switching of gene expression, is a common feature of host-adapted bacterial pathogens such as *Neisseria meningitidis* (1). In a series of recent studies we have reported that phase variation of expression can occur in N^6- adenine DNA methyltransferases (Mod) as a result of hypermutation of simple DNA repeats within the open reading frame (ORF), which leads to frame-shift mutations and the ON/OFF switching of Mod expression. The resulting changes in methylation of the genome lead to global changes in gene expression. These phase variable regulons, phasevarions (2), have been reported in a range of bacterial pathogens (reviewed in 3) such as *Moraxella catarrhalis* (4,5), *Haemophilus influenzae* (6), *Helicobacter pylori* (7) and the pathogenic *Neisseria,* where they have been demonstrated to control expression of surface antigens and virulence factors (8), and lead to altered phenotypes such as antibiotic resistance (9). Recent work has shown that similar randomly switching epigenetic regulation systems also occur in major Gram-positive pathogens, including *Streptococcus pneumoniae* (10), indicating that phasevarions are key regulators of virulence and

*To whom correspondence should be addressed. Tel: +61 755527050; Fax: +61 755528098; Email: m.jennings@griffith.edu.au
†These authors contributed equally to the paper as first authors.

immunoevasion in many of the bacterial pathogens that have a major impact on human health.

In *N. meningitidis*, two distinct Mod proteins have been studied in detail, ModA (exemplified by M.NmeMC58I and M.Nme77I) (8) and ModD (exemplified by M.Nme573I) (11). The genes encoding each of these Mod proteins have different alleles based on distinct amino acid sequences in their DNA recognition domain (12). Consequently, the Mod alleles defined by these distinct DNA recognition domains will likely methylate different DNA sequences, and distinct changes have been observed in expression profile analysis of the ModA11 (M.NmeMC58I), ModA12 (M.Nme77I) and ModD1 (M.Nme573I) phasevarions (8,11). DNA methyltransferases from different strains are classified as the same Mod allele if they have ≥95% identity in the DNA recognition domain. Particular ModA alleles have been shown to generate the same regulatory phenotype in different strains (8), therefore, for simplicity, the ModA or ModD allele name associated with a regulatory phenotype is used in this study. Furthermore, the ModX nomenclature is used, rather than the standard nomenclature (13), to distinguish these regulatory DNA methyltransferases from the non-regulatory (non-phase variable) methyltransferases that are associated with functional restriction-modification systems.

Investigations of the mechanism of epigenetic regulation of the genes in these phasevarions have been limited, as the target site for these Mods is unknown. Here we present the identification of the target specificities of these key enzymes from four meningococcal strains (Table 1) using single molecule, real-time (SMRT) methylome analysis (14).

MATERIALS AND METHODS

Bacterial strains and growth conditions

N. meningitidis strains MC58 (15), B6116/77 (16), FAM18 (17) and M0579 (18) are used in this study. MC58 *modA11* ON, B6116/77 *modA12* ON (8) and M0579 *modD1* ON (11) colonies were isolated, and the respective *modA11::kan*, *modA12::kan* and *modD1::kan* knockout mutants were generated previously (8,11). *N. meningitidis* strains were grown on Brain Heart Infusion (BHI, Oxoid) 1% agar and 10% Levinthal's Base medium at 37°C with 5% CO_2, with either kanamycin (kan) (100 μg/ml) or chloramphenicol (cm) (5 μg/ml) as required. *Escherichia coli* strains were cultured with Luria-Bertani medium (Difco), with ampicillin or kanamycin (100 μg/ml) as required.

Generation of the ModA11 locked ON (ON_1R) strain

The wild-type *modA11* gene (*nmb1375*) was amplified by polymerase chain reaction (PCR) from MC58 genomic DNA with 19 repeats (using primers ModA_F/ ModA_R, Table 2) and cloned into the pETSUMO expression vector (Life technologies, USA). Using site-directed mutagenesis, the pETSUMO*modA11*_1R construct was generated in which *modA11* contains only a single repeat unit (using primers ModA_1R_F / ModA_1R_R, Table 2). This construct was linearized with BstXI and transformed into MC58 *modA11::kan* by homologous recombination to generate strain MC58 *modA11* ON_1R. Colonies that lost kanamycin resistance were selected. Transformants containing one repeat in *modA11* were then confirmed by GeneScan fragment length analysis (Applied Biosystems, CA, USA) (8) and sequencing.

Generation of the *E. coli* ModA12 strain

The wild-type *modA12* gene (*nmc1310*) was PCR amplified from FAM18 genomic DNA (using primers M.Nme18orf1310exp_F and M.Nme18orf1310exp_R, Table 2) and cloned into the pRRS expression vector (NEB, Ipswich, MA, USA). The expression construct pRRS-*modA12* was transformed into the *E. coli* methylation deficient (dam-, dcm-) strain ER2796 (NEB; GenBank Accession: CP009644.1).

Analysis of Mod expression

The length of the *mod* repeats, and the percentage of each fragment length, were determined by GeneScan fragment length analysis as previously described (8,11). Primer pairs 6Fam-Him1 / Him3 and 6Fam-ModDRepF / ModDRepR were used to amplify the repeat regions of *modA* and *modD,* respectively (Table 2). Western blot analysis of Mod expression in whole cell lysates of meningococci strains was performed as previously described (19) using ModA (8) or ModD (11) antibodies.

Restriction inhibition assays

Chromosomal DNA was extracted from MC58 (*modA11* ON, *modA11::kan* and *modA11* ON_1R) and M0579 (*modD1* ON and *modD1::kan*) and digested overnight with AluI for ModA11, and HincII/AluI for ModD1. Digested fragments were separated on 1–1.8% agarose gels in Tris-borate-EDTA (TBE) buffer at 100 V for 1–2 h and visualized by UV illumination. Southern transfer was performed as described by (20), and hybridization and detection was performed using digoxigenin (DIG)-labeled (Roche, USA) PCR products as probes, according to the manufacturer's instructions. Probes were amplified using the primer pairs ModA11AluIF / ModA11AluIR for ModA11, and cutF / cutR for ModD1 (Table 2). Plasmid pCmGFP was extracted from C311#3 (21) and primers used for site-directed mutagenesis of pCmGFP are listed in Table 2. All restriction endonucleases were from New England Biolabs Inc. (NEB).

ModA11 protein expression and purification

The pETSUMO*modA11*_1R construct (see above) was sequenced and transformed into BL21(DE3) competent cells. After induction with 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG), the recombinant ModA11 protein with its N-terminal His6-SUMO tag was expressed at 20°C for 16 h and pelleted by centrifugation at 7000 rpm for 30 min. The cell pellet was resuspended in 100 ml lysis buffer (50 mM Tris–HCl, pH 8.0, 200 mM NaCl, 5 mM imidazole) and disrupted using a French press (30 000 psi; Sim-Aminco, USA). The total cell lysate was centrifuged at 20000 rpm for 40 min and the soluble recombinant ModA11 protein was purified by immobilized metal–ion chromatography with a Ni-NTA column (Amersham,

Table 1. Summary of DNA methylation by ModA11, ModA12 and ModD1

Mod allele[a]	Name[b]	GenBank accession (ORF) [strain: DRD coordinates]	Methyltransferase specificity
ModA11	M.NmeMC58I	NC_003112.2 (NMB1375) [MC58; 1400838–1401437]	5′-CGY^{m6}**A**G-3′
ModA12	M.Nme77I	CP007667 (B6116/77_01704) [B6116/77; 1732514–1733201]	5′-AC ^{m6}**A**CC-3′
	M.Nme18I	AM421808 (NMC1310) [FAM18; 1298475–1299159]	
	M.Nme579II	CP007668 (M0579_00017) [M0579; 17651–18336]	
ModD1	M.Nme579I	CP007668 (M0579_01779) [M0579; 1890047–1890626]	5′-CC ^{m6}**A**GC-3′

[a]DNA methyltransferases from different strains are classified as the same Mod allele if they have ≥95% identity in the DNA recognition domain. Particular ModA alleles have been shown to generate the same regulatory phenotype in different strains (8), therefore, for simplicity, the ModA allele name associated with this regulatory phenotype is used in this study.
[b]Each DNA methyltransferase identified from a particular *Neisseria spp.* strain is named according to standard nomenclature (13) and is deposited under this name in REBASE.

Table 2. Primers used in this study

Primers	Sequences (5′-3′)
Him3	CAAAAAAGCCGGTCAATTCATCAAA
6Fam-Him1	[6FAM]ATGGCGGACAAAGCACCGAAGG
ModA_F	ATGAAGACAGACATTCAAACCGAATTAACC
ModA_R	TTATTCGCCATCTTTTTTCTCCGCTTGATT
ModA_1R_F	CACCGAGGAAGGCGAAGAAATTTATTTTAAACGCAATAACAGCCAAAGACAAGAAATC TTTTTTAATCAAACCCTTGCTTTTGATG
ModA_1R_R	CATCAAAAGCAAGGGTTTGATTAAAAAAGATTTCTTGTCTTTGGCTGTTATTGCGTTT AAAATAAATTCTTCGCCTTCCTCGGTG
CmGFP_For	CTCGAGTGGAGCGCAGCAGCGCGACCTAAGCCGGCC
ModA11AluIF	CAGCTACCAAACGAGCCT
ModA11AluIR	GCAGCTGCAGATAAGGAA
ACGCAC_Rev	AGCTGCG**T**TACCGCTTCGCCTGCTGGCTTCGCT
GCGCAG_Rev	AGCTGCG**C**TACCGCTTCGCCTGCTGGCTTCGCT
CCGCAG_Rev	AGCTGCG**G**TACCGCTTCGCCTGCTGGCTTCGCT
TCGCAG_Rev	AGCTGCG**A**TACCGCTTCGCCTGCTGGCTTCGCT
T**T**GCAG_Rev	AGCTGC**A**ATACCGCTTCGCCTGCTGGCTTCGCT
TC**A**CAG_Rev	AGCTG**T**GATACCGCTTCGCCTGCTGGCTTCGCT
TCG**T**AG_Rev	AGCT**A**CGATACCGCTTCGCCTGCTGGCTTCGCT
CmGFP_SeqF	GGCTCATGTTGTATCTCGAAACCCCCG
CmGFP_SeqR	GATGAGTTGCTTTGTTCGCTTTTCGGC
6Fam-ModDRepF	[6FAM]GATGGAAGACGCAATTATGGC
ModDRepR	CGAAGTCTTTTGTGAAGACCA
cutF	ACGTCGGACGACTACATCGTTAC
cutR	GTTGACCATATATTCGATAG
pUC19 F1	ATCTGCGCTCTGCTGAAGCC
pUC19 F2	AAACAAACCACCGCTGGTAG
pUC19 F3	AAAAATCGACGCTCAAGTCA
pUC19 R1	GTTCCACTGAGCGTCAGACC
pUC19 R2	ACCAAATACTGTCCTTCTAG
M.Nme18orf1310 exp_F	TGCCTGCAGTTAAGGTTTAACATATGAAGGCAGACATTCAAACCGAATTAACC
M.Nme18orf1310 exp_R	TCTAGATCTTCCCCGGGGATCCTTATTATTCGCCATCTTTTTTCTCCGCTTGA

Bold underlined letters correspond to nucleotides that were mutated.

USA). The 50 ml solution containing the His6-SUMO-ModA11 protein was dialyzed against cleavage buffer (50 mM Tris–HCl, pH 8.0, 100 mM NaCl). Ten units of ULP-1 protease (Life technologies, USA) were added to remove His6-SUMO. The resulting ModA11 protein solution was re-loaded onto a Ni-NTA column to remove uncleaved proteins and the cleaved His6-SUMO tag. The flow-through was collected and dialyzed against storage buffer (30 mM HEPES, pH 7.0, 50 mM NaCl, 2 mM TCEP) and the protein (approx. 10 mg/ml) was stored at -80°C.

In vitro methylation

All methylation assays monitored the incorporation of tritiated methyl (^{3}H) groups into DNA by using a modified ion exchange filter-binding assay (22). Methylation assays were carried out in a reaction mixture (20 μl) containing DNA (pUC19 plasmid DNA, fragments of pUC19 plasmid DNA obtained by restriction digestion or PCR amplification, or duplex DNA (15 μM), with a single recognition sequence or modified recognition sequence (Supplementary Table S1)), [3H]Ado-Met (5 μM) (specific activity 66 Ci.mmol^{-1}) and purified protein (2 μM) in the reaction buffer (10 mM Tris/HCl, pH 8.0, 5 mM β-mercaptoethanol). After incubation at 37°C for 30 min, the reactions were stopped by snap-freezing in liquid nitrogen. Data were analyzed and compared to time 0 min and measurements after incubation in the absence of enzyme was subtracted. All methylation experiments were carried out at least in triplicate and the results were averaged.

Methylome determination by SMRT sequencing

Meningococcal strains were grown overnight on BHI agar, and genomic DNA extracted using the QIAGEN Genomic-tip 20/G kit as per the manufacturer's instructions. SMRTbell libraries were prepared as previously described (23) according to the manufacturer's instructions (PacBio, CA, USA). Briefly, genomic DNA was sheared to an average length of approximately 10 kb using g-TUBEs (Covaris; Woburn, MA, USA), treated with DNA damage repair mix, end repaired and ligated to hairpin adapters. Incompletely formed SMRTbell templates were digested using Exonuclease III (NEB) and Exonuclease VII (Affymetrix; Cleveland, OH, USA). Sequencing was carried out on the PacBio RS II (Menlo Park, CA, USA) using standard protocols for long insert libraries. Methylome data and evidence supporting the identification of the ModA11, ModA12 and ModD1 specificities have been reported to REBASE.

Bioinformatic analysis

The whole genome sequences determined by SMRT sequencing were annotated using the Prokka bacterial genome annotation tools (Prokaryotic Genome Annotation System, http://vicbioinformatics.com/) and the prepared file submitted to GenBank. The accession numbers for the annotated genomes are CP007667 for strain B6116/77 and CP007668 for strain M0579. The other strains investigated, MC58 (AE002098) and FAM18 (AM421808), were sequenced previously. The methylation sites were identified by SMRT sequencing and their location and relation to genome features were analyzed using perl, the R statistical package (24) and Artemis (25).

RESULTS

Phase variable DNA methyltransferases of *N. meningitidis*

In this study three key phase variable DNA methyltransferases of *N. meningitidis*, ModA11, ModA12 and ModD1 (Figure 1) are investigated to determine their DNA methylation sites to aid future mechanistic studies on the mechanism of epigenetic gene regulation. The *modA11* and *modA12* coding regions contain a tract of 5′-AGCC-3′ tetranucleotide repeats that mediate phase variation (8). These genes share a high degree of identity (94%) along the length of the ModA deduced amino acid sequence, excluding the central variable DNA recognition domain that shares only 18/33% identity/similarity (Figure 1A and B (i)). The DNA recognition domain, which binds DNA and dictates sequence specificity (26), defines the different Mod alleles. In previous studies ModA11 and ModA12 were found to regulate the expression of different sets of genes, consistent with differences in their DNA recognition domain (8). ModA11 and ModA12 are present in 17.8% and 78.5% of meningococcal clinical isolates examined, respectively (8). ModD1, on the other hand, contains a 5′-ACCGA-3′ pentanucleotide repeat tract in the *modD* coding sequence that mediates phase variation, and is only distantly related to ModA, sharing only ∼18/31% identity/similarity to the deduced amino acid sequence of ModA alleles in both the DNA recognition domain and the flanking regions. ModD1 is significantly associated with isolates of the hypervirulent clonal complex (cc) 41/44 and has been identified in 74% of cc41/44 clinical isolates examined (11) (Figure 1C (i)).

We used *N. meningitidis* strains MC58, B6116/77 and M0579 in our studies to identify the DNA methylation sites of ModA11, ModA12 and ModD1, respectively. Isogenic mutants of each of these strains were generated by insertion of a kanamycin resistance cassette into the open reading frame of the *mod* genes (Figure 1A–C (i)). Western blot analysis confirmed the expression of Mod in the wild-type ON strains, and the absence of expression in the mutant strains (Figure 1A–C (ii)). Each of the wild-type strains were screened to isolate samples enriched for the respective *mod* gene phase varied ON, resulting in populations that were 86%, 82% and 91% Mod ON for *modA11, modA12* and *modD1*, respectively (Figure 1A–C (iii)). The Mod ON expression status of wild-type samples used in subsequent experiments was also verified by GeneScan fragment length analysis to determine the length of the phase variable repeat tract, and the percentage of the population containing each fragment length as previously described (8,11).

Identification of Mod target sites

Restriction inhibition analysis. Initial studies to determine the ModA11, ModA12 and ModD1 methylation target sites were based on inhibition of DNA restriction, as previously performed for ModA13 (8). Genomic or pCmGFP plasmid DNA was isolated from *N. meningitidis modA11,*

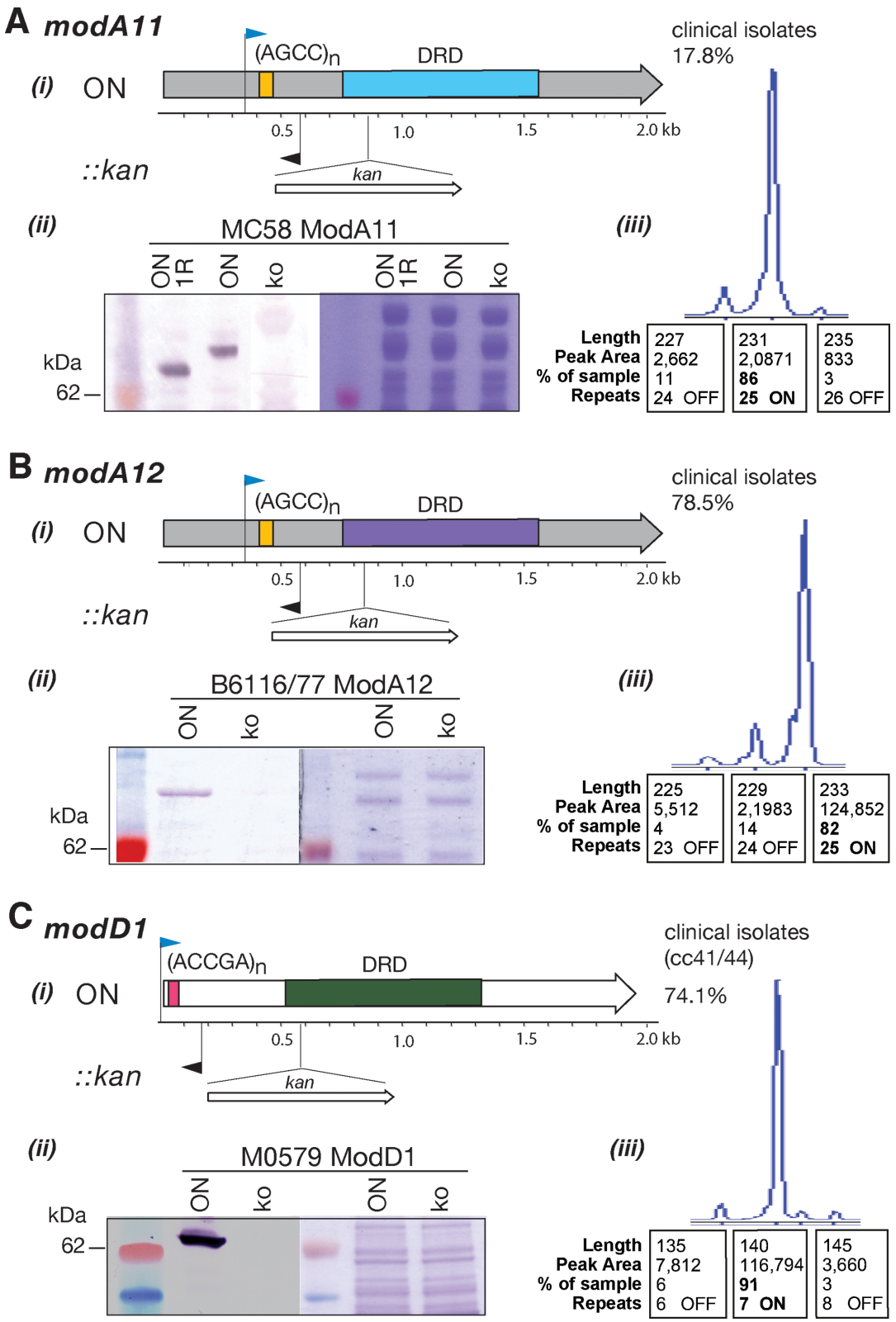


Figure 1. Diagrammatical representation of (**A**) *modA11,* (**B**) *modA12* and (**C**) *modD1* alleles of *N. meningitidis* and their expression status. (i) Schematic representation of the methyltransferase (*mod*) gene showing the DNA repeat tract responsible for phase variable expression, the variable DNA recognition domain (DRD), the PCR primers (arrowheads) used for GeneScan fragment length analysis and the site of insertion of the kanamycin resistance gene (kan) in the *mod::kan* knockout strains. The distribution of the allele in clinical isolates is shown on the right. (ii) Western blot analysis confirming Mod expression in whole-cell lysates of the natural *mod* ON strain (ON) and the locked ON strain with 1 repeat (ON_1R), and the absence of Mod expression in the *mod::kan* knockout (ko) strain. A Coomassie stained SDS-PAGE gel of the same samples is shown on the right. (iii). GeneScan fragment length analysis of samples used for SMRT analysis, showing the different fragment lengths (corresponding to different numbers of repeats) and their relative abundance in the sample.

modA12 and *modD1* ON strains and their corresponding *mod::kan* knockout mutants, and was digested with a range of restriction enzymes known to be inhibited by methylation of an adenine within their recognition sequence, as listed in REBASE (27). Differences in digestion patterns between DNA extracted from *mod* ON cells (Mod methylated DNA) and *mod::kan* cells (DNA not methylated by Mod) may indicate an overlap of the respective Mod methylated target and the restriction enzyme used. Unfortunately, no such fortuitous inhibition pairs were seen with the *modA11*, *modA12* and *modD1* strains (data not shown).

In vitro methyl ^{3}H labeling

The next approach to identify the ModA11 recognition site utilized an *in vitro* methylation assay to monitor the incorporation of tritiated methyl (^{3}H) groups into DNA. Recombinant ModA11 methyltransferase was expressed, purified and used to investigate the transfer of ^{3}H-labeled methyl groups from [methyl-^{3}H]AdoMet to adenine of a potential ModA11 recognition sequence in pUC19 plasmid DNA or fragments of pUC19 obtained using different combinations of restriction enzymes. A pUC19 fragment, obtained from EcoRI and BsrFI digestion, showed significant methylation, and this fragment was then used as a template to PCR amplify different sub-fragments (see Table 2 for primers sequence), and determine their methylation state. Based on the sequence of pUC19 sub-fragments showing methylation, a number of duplexes were designed for further investigation. Significant methylation of duplex MODA11–1 was observed, which contains the sequence 5′-GCGCAG-3′ that is common to the smallest sub-fragment of pUC19 that was methylated (product of primers pUC19 F2/ pUC19 R1). The remaining duplexes were designed based on MODA11–1, and ModA11 was found to methylate duplexes that contained both 5′-TCGCAG-3′ and 5′-GCGCAG-3′ sequences (Supplementary Table S1). However, replacement of the 5′-end 'T' (MODA11–9) with 'G' (MODA11–11) led to a 100-fold lower methylation activity. The kinetic parameters for ModA11 methylation are shown in Supplementary Table S2. This analysis indicated that ModA11 methylates the core site 5′-CGC**^{m6}**AG-3′.

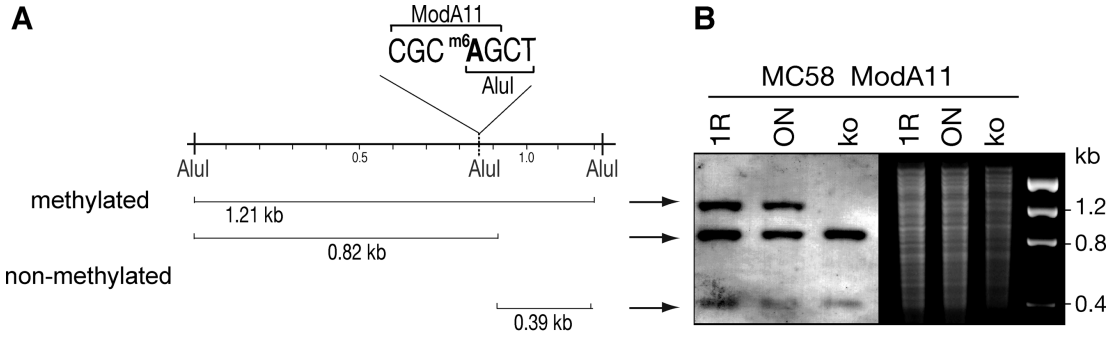
To confirm the ModA11 recognition sequences observed by *in vitro* methylation assays, a restriction inhibition assay was performed using the commercially available methyl-sensitive restriction enzyme AluI (recognizes and cleaves 5′-AGCT-3′, but not 5′-^{m6}AGCT-3′, sites), which overlaps with the 5′-GCG^{m6}AG-3′ ModA11 methylation sequence at several sites in the genome (Figure 2A). Southern blot analysis revealed that genomic DNA from MC58 *modA11* ON and MC58 *modA11* ON_1R (a strain constructed to contain 1 repeat in ModA11, which is locked ON and unable to phase vary), but not *modA11::kan*, were partially protected against AluI digestion, confirming that ModA11 methylates the proposed recognition sites 5′-CGC**^{m6}**AG-3′ (Figure 2B). To further investigate the flexibility of the methylation site that was seen in the *in vitro* methylation assays described above, the *Neisseria* spp. specific plasmid pCmGFP was used as a template to construct a panel of alternative versions of the proposed ModA11 recognition site with different bases either within, or at the flanking 5′ end of, the 5′-CGC**^{m6}**AG-3′ site (Figure 2C). Seven plasmids were constructed and transformed into MC58 *modA11* ON_1R and MC58 *modA11::kan* strains. AluI digestion of these constructs revealed that base alterations flanking the 5′ end of 5′-CGC**^{m6}**AG-3′ maintained ModA11 methylation, and DNA was partially protected against AluI digestion (Figure 2C, lanes 1–4). However, alterations to the central sequence 5′-CGCAG-3′ resulted in a decrease or loss of ModA11 methylation (Figure 2C, lanes 5–7) and confirmed this core, invariant component of the target sequence. Incomplete protection was observed in the construct containing 5′-CGTAG-3′ (lane 7) indicating that this site may be partially methylated.

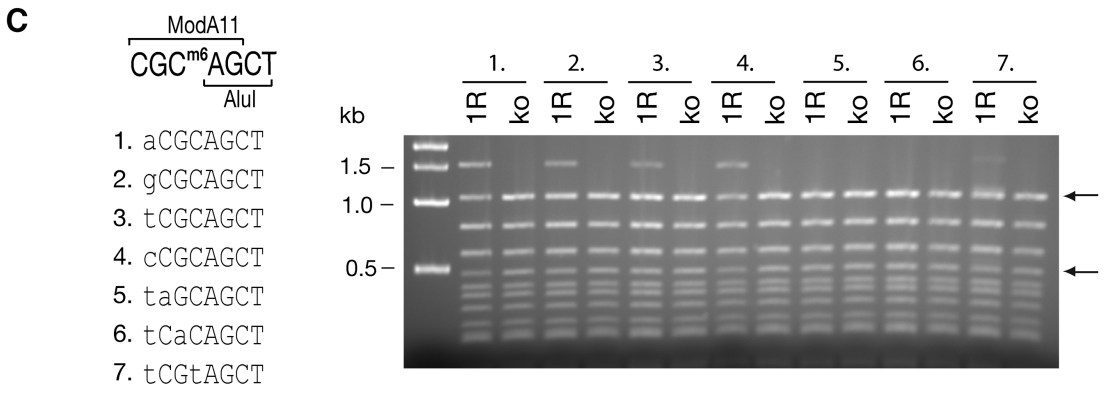
SMRT methylome analysis

At this stage in the study, SMRT methylome analysis was described as a rapid method to identify methylation sites in microbial genomes (23). We applied this method to all three strains of interest, sequencing the genomic DNA from both the *mod* ON and *mod::kan* knockout strains of MC58, B6166/77 and M0579 (Figure 1). This process involved complete *de novo* assembly of strains B6116/77 and M0579, and the genome sequences of these strains are deposited in GenBank (accession numbers CP007667 and CP007668, respectively). The N^{6}-methyladenine (^{m6}A) and N^{4}-methycytosine (^{m4}C) bases in these genomes were identified by their kinetic signatures and these were then aligned and clustered to identify the motifs that constitute the consensus recognition sequences for the active methyltransferases in these strains. Table 1, Figures 2 and 3 and Supplementary Figure S1 show a summary of results from SMRT sequencing, and details are given in the following sections.

ModA11 methylates 5′-CGY^{m6}AG-3′

The MC58 *modA11* ON and MC58 *modA11::kan* genomes were subjected to SMRT sequencing, which was mapped to the available MC58 genome (GenBank AE002098.2, which is actually the acapsulate mutant strain derived from MC58 called MC58¢3 (28)). Two distinct methylation patterns were detected in MC58 *modA11* ON; 5′-G**^{m6}**ACGC-3′ and 5′-CGY**^{m6}**AG-3′ (Figure 2D). The methyltransferase responsible for methylating the site 5′-G**^{m6}**ACGC-3′ was identified as M.NmeMC58II based on homology searches. Comparison of the *modA11* ON and the *modA11::kan* strains indicated that ModA11 (M.NmeMC58I) is responsible for methylation of ^{m6}A residues in the sequence 5′-CGY**^{m6}**AG-3′ (Figure 2D, Table 1), which is consistent with the site identified by *in vitro* methylation (above). However, only 8% (672/7957) of these sites in the genome were detected as methylated by ModA11 in the *modA11* ON strain. In order to determine whether this low rate of methylation was due to a low portion of the sample having ModA11 ON (original sample was 86% ON as determined by GeneScan fragment length analysis, Figure 1A), the MC58 *modA11* ON_1R strain in which the *modA11* is unable to phase vary was also investigated. SMRT sequencing and methylome analysis of the ModA11 ON_1R strain showed that 46% (3618/7957) of 5′-CGY^{m6}AG-3′ sites in the genome were





D

Methyltransferase Specificity	Modified Base	*modA11 ON*	*modA11 ON_1R*	*modA11 ::kan*	Name
5'-CGY^{m6}**A**G-3'	m6A	8% (672/7957)	46% (3618/7957)	0%	M.NmeMC58I (ModA11)
5'-G^{m6}**A**CGC-3'	m6A	98%	97%	89%	M.NmeMC58II

E

Methyltransferase Specificity	Modified Base	*modA12 ON*	*modA12 ::kan*	Name
5'-AC^{m6}**A**CC-3'	m6A	90.5% * (3591/3969)	0%	M.Nme77I (ModA12)
5'-G^{m6}**A**TC-3'	m6A	97.8%	100%	ND
5'-GA^{m6}**A**GG-3'	m6A	96.9%	99.9%	M.Nme77II

Figure 2. Identification of the (A–D) ModA11 5′-CGC^{m6}AG-3′ and (E) ModA12 5′-AC^{m6}ACC-3′ methylation sites. (**A**) Overview of the restriction inhibition assay used to confirm the ModA11 methylation site. The location of 3 methyl-sensitive AluI restriction enzyme sites within the probed region of the genome, and the size of fragments that are generated if the DNA is methylated at the middle AluI site or non-methylated, are shown. AluI recognizes the sequence 5′-AGCT-3′ but does not cut if the adenine is methylated in the overlapping ModA11 methylation site 5′-CGC^{m6}AG-3′. (**B**) Southern blot and DNA agarose gel of AluI digested genomic DNA from MC58 *modA11* ON, *modA11* ON_1R (1R) and *modA11::kan* (ko) strains. The ModA11 recognition site, 5′-CGCAG-3′, is methylated in the MC58 *modA11* ON and *modA11* ON_1R and is protected from digestion by AluI. (**C**) Seven constructs were made containing alterations to the 5′-CGCAG-3′ sequence overlapping of one of the AluI digestion sites in the pCmGFP plasmid (sites are shown on the left). The DNA gel on the right shows AluI digests of the pCmGFP constructs from *modA11* ON_1R (1R) and *modA11::kan* (ko) strains, where an undigested 1.561 kb fragment is seen when the site is methylated and partially protected from digestion. When unmethylated, bands of 1067 bp and 494 bp (shown by arrows on the right). (**D**) Methylated sequences identified by SMRT sequencing in the MC58 *modA11* ON, *modA11* ON_1R and *modA11::kan* strains. The % (number) of sites that are detected as methylated in the genome are shown. Assignment of the name is based on direct evidence [M.NmeMC58I (ModA11) encoded by *nmb1375*] or homology [M.NmeMC58II]. (**E**) Methylated sequences identified by SMRT sequencing in B6116/77 *modA12 ON* and *modA12::kan* strains. Underlined sequence represents methylation on opposite strand. Assignment of the name is based on direct evidence [M.Nme77I (ModA12) encoded by *B6116/77_01703/01704*] or homology [M.Nme77II]. ND, not determined; no methyltransferase could be reliably assigned as responsible for methylation of this site. * ModA12 site confirmed by expression in *E. coli* ER2796, with 99.78% (8,349/8,367) of 5′-AC^{m6}ACC-3′ sites detected as methylated in the genome.

A

Methyltransferase Specificity	Modified Base	*modD1 ON*	*modD1 ::kan*	Name
5'-CC^{m6}AGC-3'	m6A	97.4% (4914/5043)	0%	M.Nme579I (ModD1)
5'-AC^{m6}ACC-3'	m6A	16.9%#	92.4%	M.Nme579II (ModA12)
5'-AGTA^{m4}CT-3'	m4C	57.6%	46.7%	M.Nme579III
5'-G^{m6}ACGC-3'	m6A	99.5%	98.6%	M2.Nme579IV

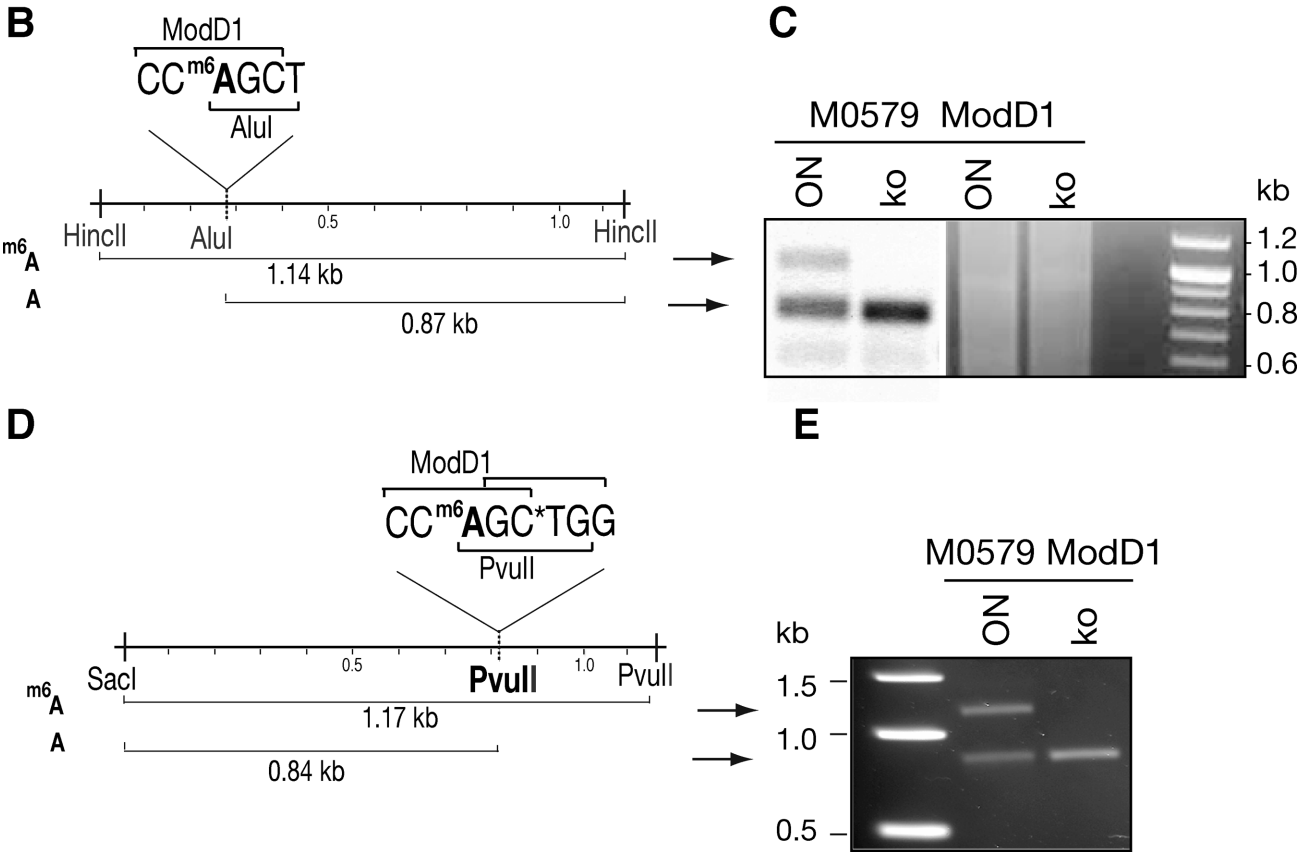


Figure 3. Identification of the ModD1 5′-CC^{m6}AGC-3′ methylation site. (**A**) Methylated sequences identified by SMRT sequencing in M0579 *modD1* ON and *modD1::kan* strains. The % (number) of sites that are detected as methylated in the genome are shown. '#'; the reduced level of ModA12 methylation is due to a lower % of ModA12 ON in the *modD1* ON versus *modD1::kan* strain. Assignment of the name is based on direct evidence [M.NmeM579I (ModD1) encoded by *M0579_01779*] or homology [M.Nme579II (ModA12), M.Nme579III and M2.Nme579IV]. (**B**) Overview of the restriction inhibition assay used. Location of a methyl-sensitive AluI restriction enzyme site between 2 HincII sites in the genome, and the size of fragments that are generated using these 2 restriction enzymes if the DNA is methylated at the AluI site (^{m6}A) or is not methylated (A), are shown. AluI recognizes the sequence 5′-AGCT-3′ but does not cut if the adenine is methylated in the overlapping ModD1 methylation site 5′-CC^{m6}AGC-3′. (**C**) Southern blot and DNA agarose gel of AluI/HincII digested genomic DNA from M0579 ModD1 ON and Δ*modD1::kan* (ko) strains. ModD1 recognition site, 5′-CCAGC-3′, is methylated in the M0579 ON strain and is protected from digestion by AluI. (**D**) Overview of the pCm-GFP plasmid restriction inhibition assay used. Location of a methyl-sensitive PvuII restriction enzyme site between a SacI and PvuII sites in the pCm-GFP plasmid, and the sizes generated using these two restriction enzymes if the DNA is methylated at the central PvuII site (m6A) or is not methylated (A), are shown. The overlapping PvuII-ModD1 methylation site is symmetrical, and '*' indicates methylation on the opposite strand. (**E**) DNA agarose gel of PvuII/SacI digested pCm-GFP DNA from M0579 ModD1 ON and *modD1::kan* (ko) strains. ModD1 recognition site, 5′-CCAGC-3′, is methylated in the M0579 ON strain and is protected from digestion by PvuII.

detected as methylated (Figure 2D), with 5′-CGT^{m6}AG-3′ (58.2%) having a slightly higher methylation rate than 5′-CGC^{m6}AG-3′ (41.3%) (Table 3).

More detailed analysis of the methylome data, taking into account all possible 5′ and 3′ flanking nucleotides and sub-sequences, revealed a complex and variable 7 base pair target sequence for ModA11. All sites with the sequence 5′-nCGY^{m6}AGn-3′ were detected as methylated but the efficiency of methylation varied from ∼5 to 100% and was dependent on which of the central C or T residues was present in combination with the flanking residues (where n = A, C, G, or T, with variable methylation efficacy based on the combination of bases present). For example, 4.6% of 5′-GCGTAGG-3′ sites were detected as methylated compared to 100% of 5′-ACGTAGG-3′ sites. Adenine at the 5′ end of 5′-nCGYAGn-3′ resulted in the highest level (85.1%) of ModA11 methylation, while guanine had the lowest level (18.6%). Thymine at the 3′ end of 5′-nCGYAGn-3′ resulted in the highest methylation level (70.5%) compared to the other three bases (29.4–50.8%) (Table 3).

ModA12 methylates 5′-AC^{m6}ACC-3′

SMRT sequencing and *de novo* assembly of B6116/77 *modA12* ON and B6116/77 *modA12::kan* resulted in a closed genome for each strain (2,187,672 and 2,188,817 bp, respectively). Three distinct methylation patterns were detected in B6116/77 *modA12* ON; 5′-G^{m6}ATC-3′, 5′-GA^{m6}AGG-3′ and 5′-AC^{m6}ACC-3′ (Figure 2E, underlining indicates methylated base on opposite DNA strand). The site 5′-G^{m6}ATC-3′ is methylated by the DNA adenine methyltransferase (Dam) (29) but could not be attributed to a specific methyltransferase. The methyltransferase responsible for methylating the site 5′-GA^{m6}AGG-3′ is predicted to be encoded by M.Nme77II based on similarities in REBASE. From comparison of the *modA12* ON and the *modA12::kan* strains, it is clear that ModA12 (M.Nme77I) is active and modifies the second A residue in the sequence 5′-AC^{m6}ACC-3′ (Figure 2E, Table 1). SMRT sequencing of another *N. meningitidis* strain containing ModA12, FAM18 detected the same three distinct methylation patterns as detected in B6116/77 *modA12* ON (Supplementary Table S3). Unfortunately, no restriction enzyme is commercially available that overlaps with the ModA12 methylation site, so methylation inhibition experiments to confirm the ModA12 site determined by SMRT sequencing could not be performed. Therefore, the ModA12 methylation site was confirmed by expressing ModA12 in the *E. coli* methylation deficient strain ER2796. Only one methylation pattern was detected by SMRT sequencing of genomic DNA isolated from ER2796-*modA12*; 99.78% (8,349/8,367) of 5′-AC^{m6}ACC-3′ sites were detected as methylated in the genome.

ModD1 methylates 5′-CC^{m6}AGC-3′

The M0579 m*odD1* ON and the *modD1::kan* genomes were subjected to SMRT sequencing, and *de novo* assembly resulted in a closed genome for each strain (2,324,822 bp and 2,326,237 bp, respectively). Four distinct methylation patterns were detected in M0579 *modD1* ON; 5′-AC^{m6}ACC-3′, 5′-AGTA^{m4}CT-3′, 5′-G^{m6}ACGC-3′ and 5′-CC^{m6}AGC-3′ (Figure 3A). ModA12 (M.Nme579II) is responsible for methylating the sequence 5′-AC^{m6}ACC-3′ (as described above), the site 5′-AGTA ^{m4}CT-3′ is methylated by M.Nme579III (based on homology to M.NmeSI (30)), and 5′-G^{m6}ACGC-3′ by M2.Nme579IV (based on homology).

From comparison of the *modD1* ON and the *modD1::kan* strains, it is clear that ModD1 (M.NmeM579I) is active and forms ^{m6}A residues in the sequence 5′-CC^{m6}AGC-3′ (Figure 3, Table 1). To confirm the ModD1 DNA methylation site, Southern blot analysis was performed on genomic DNA, isolated from M0579 *modD1* ON and *modD1::kan* and digested with the methyl-sensitive restriction enzyme AluI (as described above for ModA11). Some AluI sites in the genome overlap with the 5′-CC^{m6}AGC-3′ ModD1 methylation site (Figure 3B). Digestion of one such region with HincII and AluI results in 872 and 252 bp fragments when unmethylated, or a 1,130 bp fragment if ModD1 methylation interferes with AluI cleavage. Partial inhibition of DNA digested by AluI at the site investigated is seen in *modD1* ON (Figure 3C), confirming the ModD1 methylation site. This ModD1 site was also confirmed using pCmGFP plasmid DNA isolated from M0579 *modD1* ON and *modD1::kan* strains that was digested with the methyl-sensitive restriction enzyme PvuII, which also overlaps the ModD1 methylation site (Figure 3D and E).

Distribution of ModA and ModD methylation sites in the *N. meningitidis* genomes

Bioinformatic analysis of the recognition sequences methylated by ModA11, ModA12 and ModD1 shows that they are distributed widely throughout the respective genomes, as may be expected for 5 base pair recognition sequences with high %GC content. They are present both within ORFs and intergenic regions, and there is no clear bias for their location. For example, in the case of ModA12, 90.5% (3,591/3,969) of the 5′-ACAGC-3′ sites present in the genome were detected as methylated. Of the methylated sites, 11.7% are in intergenic regions (419/3,591), and intergenic regions comprise 18.3% of the genome. The intergenic regions have a lower %GC than coding regions (44% versus 51.66%), which may account for lower frequency of 5′-ACAGC-3′ sites (60% GC) in these locations. For ModD1, 97.4% (4,914/5,043) of the 5′-CCAGC-3′ sites present in the genome were detected as methylated. Of the methylated sites, 4.7% are in intergenic regions (238/4,914), and intergenic regions comprise 18% of the genome (80% GC).

In several cases, methylated sites are present in the intergenic regions of genes that have been shown to be differentially expressed in the respective ModA11, ModA12 (7) and ModD1 (11), phasevarions (Supplementary Figures S2–S4). These data provide the basis for future studies to determine how changes in DNA methylation regulate expression of these genes.

DISCUSSION

Random switching of gene expression (phase variation) in host adapted bacterial pathogens is an established mechanism of altering expression of individual genes and in most

Table 3. Distribution and methylation state of ModA11 5′-nCGY^{m6}AGn-3′ recognition sites in the genome of *N. meningitidis* MC58

Methyltransferase Specificity	Sites Detected as Modified	Sites in Genome	% Detected Modified	IPD Ratio	Mean Score	Mean Coverage
5′-nCGT**A**Gn-3′	1190	2045	58.2			
5′-aCGTAGn-3′	216	224	96.4			
5′-aCGT**A**Ga-3′	49	50	98	6.1	122.9	100.6
5′-aCGT**A**Gc-3′	54	60	90	6.5	113.7	98.4
5′-aCGT**A**Gg-3′	61	61	100	6	118.9	103.5
5′-aCGT**A**Gt-3′	52	53	98.1	6	123.3	96.4
5′-cCGT**A**Gn-3′	466	694	67.1			
5′-cCGT**A**Ga-3′	91	126	72.2	4.6	98.1	103.8
5′-cCGT**A**Gc-3′	140	251	55.8	4.9	89.4	98.9
5′-cCGT**A**Gg-3′	123	192	64.1	4.5	92.5	104.4
5′-cCGT**A**Gt-3′	112	125	89.6	5	102.3	99.1
5′-gCGTAGn-3′	247	612	40.4			
5′-gCGT**A**Ga-3′	53	115	46.1	4.4	88.7	100.8
5′-gCGT**A**Gc-3′	56	170	32.9	5.1	90.3	102.9
5′-gCGT**A**Gg-3′	73	209	34.9	4.2	90.3	105.6
5′-gCGT**A**Gt-3′	65	118	55.1	4.2	88.3	102.2
5′-tCGTAGn-3′	261	515	50.7			
5′-tCGT**A**Ga-3′	55	104	52.9	4.4	91.5	102.3
5′-tCGT**A**Gc-3′	80	176	45.5	4.7	88.6	103
5′-tCGT**A**Gg-3′	46	137	33.6	4.3	88.2	102.5
5′-tCGT**A**Gt-3′	80	98	81.6	5.1	108.6	104.4
5′-nCGC**A**Gn-3′	2439	5912	41.3			
5′-aCGCAGn-3′	913	1102	82.8			
5′-aCGC**A**Ga-3′	146	160	91.3	5.3	110.5	98
5′-aCGC**A**Gc-3′	305	353	86.4	6.6	108.8	98
5′-aCGC**A**Gg-3′	307	421	72.9	4.7	97.3	100.6
5′-aCGC**A**Gt-3′	155	168	92.3	5.7	120.5	98.7
5′-cCGCAGn-3′	948	1931	49.1			
5′-cCGC**A**Ga-3′	178	486	36.6	4.3	90.8	98.1
5′-cCGC**A**Gc-3′	453	744	60.9	5.2	91.4	100.9
5′-cCGC**A**Gg-3′	128	476	26.9	4.1	85.4	105
5′-cCGC**A**Gt-3′	189	225	84	5	106	99.2
5′-gCGCAGn-3′	302	2341	12.9			
5′-gCGC**A**Ga-3′	43	284	15.1	4.2	89.6	100.8
5′-gCGC**A**Gc-3′	84	514	16.3	4.9	90.5	100.6
5′-gCGC**A**Gg-3′	57	1226	4.6	3.9	85	102.6
5′-gCGC**A**Gt-3′	118	317	37.2	4.3	87.4	96.8
5′-tCGCAGn-3′	276	538	51.3			
5′-tCGC**A**Ga-3′	59	106	55.7	4.6	98.3	101.5
5′-tCGC**A**Gc-3′	93	156	59.6	5.5	91.1	99.9
5′-tCGC**A**Gg-3′	63	200	31.5	4.2	89	107.4
5′-tCGC**A**Gt-3′	61	76	80.3	5.5	113.1	100.5
5′-aCGY**A**Gn-3′	1129	1326	85.1			
5′-cCGY**A**Gn-3′	1414	2625	53.9			
5′-gCGY**A**Gn-3′	549	2953	18.6			
5′-tCGY**A**Gn-3′	368	482	76.3			
5′-nCGY**A**Ga-3′	674	1431	47.1			
5′-nCGY**A**Gc-3′	1231	2424	50.8			
5′-nCGY**A**Gg-3′	858	2922	29.4			
5′-nCGY**A**Gt-3′	832	1180	70.5			

IPD, Interpulse duration.

cases results in the ON/OFF switching of expression of a single protein with a limited phenotypic effect (1). The combination of many such loci can generate a diverse population of individuals, however the accumulation of substantially distinct phenotypes requires many stepwise changes in the hypermutable loci that control the expression of individual genes. However, it has been discovered that a wide range of host adapted Gram-negative bacterial pathogens contain DNA methyltransferases (Mod) that are also subject to phase variation. The finding that global changes in methylation patterns that result from Mod switching can change global patterns of gene expression has elevated the status of these phase variable methyltransferases to global regulators (2,5,7–8,11). The key to examining the role of DNA methylation in the gene regulation events that underpin virulence and immunoevasion phenotypes is the identification of the sequence and location of the sites that are modified by these DNA methyltransferases, the majority of which are unknown.

The identification of the ModA11, ModA12 and ModD1 methylation sites by SMRT methylome analysis will enable in-depth investigations of the mechanisms and outcome of epigenetic gene regulation in *N. meningitidis.* These Mods are N^6-adenine DNA methyltransferases that are associated with Type III Restriction-Modification (R-M) systems. R-M systems are classified into four types based on their subunit organization, sequence recognition, the position of DNA hydrolysis/modification, cofactor requirements and substrate specificity (31,32). Type III R-M systems consist of Mod and Res subunits, and Mod is responsible for recognition and methylation of short non-palindromic sequences. Res, when in complex with Mod, hydrolyzes unmethylated DNA approximately 25 bp from the recognition site (reviewed in 33). The ModA12 and ModD1 recognition sequences are typical Type III Mod recognition sequences, in which one strand of a 5 bp non-palindromic sequence is methylated. As with similar SMRT studies (14,23,34), greater than 90% of the deduced target sites were detected as methylated in the native hosts. However, the ModA11 recognition sequence is a complex and variable 7 bp sequence. Essentially, a core 5 bp motif is present, with variable levels of methylation seen depending on the pyrimidine present at the central position (e.g. 58% methylation of 5′-nCGT^{m6}**A**Gn-3′ versus 41% of 5′-nCGC^{m6}**A**Gn-3′sites) as well as the combination of 5′ and 3′ flanking bases present (e.g. methylation ranges from 4.6% of 5′-gCGC^{m6}**A**Gg-3′ to 100% of 5′-aCGT^{m6}**A**Gg-3′ sites). Typically, both the Mod and Res of R-M systems are highly site-specific (31). While some studies have demonstrated promiscuous nonspecific or 'off-target' activity, this has been seen under non-native conditions and their biological relevance is unclear (e.g. during overexpression in *E. coli* (23,34–35); the use of high concentrations of recombinant proteins to characterize 'star activity' (36–38 and references therein); or direct attempts to engineer promiscuous activity (39–41)). To the best of our knowledge, this is the first time that this type of promiscuity has been characterized for a DNA methyltransferase in its natural host, and has interesting implications in terms of its evolution and function.

The evolution of R-M systems and novel or promiscuous R-M recognition sequences has recently been reviewed, and in general it is assumed that for most types of R-M systems the evolution of novel specificities must involve the coevolution of both components, since Res are toxic in the absence of their cognate Mod (32). However, there is likely to be less selective pressure on Type III R-M systems, such as ModA11, where DNA recognition by Res is dependent on the Mod. Furthermore, the Res associated with ModA11 contains a frameshift mutation and is non functional. Similarly, inactivating mutations have been identified within the *res* gene associated with *modA* in multiple strains of *H. influenzae* (6), and it appears that the restriction function has been lost and that Mod is dedicated to a gene regulation function.

The broad and variable specificity of ModA11 provides an added layer of depth and complexity to epigenetic regulation in *N. meningitidis*, and other organisms that may express similar DNA methylases. Given that meningococcal strains may contain up to three phase variable Mods (ModA, ModB and ModD) that switch ON/OFF independently, there are numerous potential combinations of *mod* gene expression, and as such of expression of genes within phasevarions. For example, in addition to the ModA and ModD proteins described herein, both the MC58 and M0579 genomes contain the *modB1* gene, whose product is known to methylate 5′-CCACC-3′ (42), while strain B6116/77 contains the *modB2*, which has an unknown target site. However, analysis of the SMRT genome sequences indicated that the *modB* alleles in these strains are phase varied OFF and not expressed. Furthermore, different levels of expression of each Mod are possible (i.e.% ON versus OFF), which may alter the methylation level in the genome. This ability to vary epigenetic signals greatly increases the adaptability of these organisms.

While epigenetic regulation in bacteria is well described, the specific mechanisms by which these phase variable DNA methylases mediate differential gene transcription are unknown. DNA methylation can directly regulate transcription by altering interaction of regulatory proteins with DNA-binding sites (25), and in each of the three Mods characterized here, modified sites were identified upstream of genes within the respective phasevarions (Supplemental Figure S2–S4). This information will guide future investigations of the molecular mechanism of regulation. However, epigenetic regulation may also be indirect and multifactorial, with expression of a regulator being affected, or expression of a protein in a tightly regulated system being affected resulting in altered expression of others protein in the same regulon. Indeed, methylation sites were only found near or within a subset of the genes regulated by the prophage encoded M.EcoGIII Mod in a pathogenic *E. coli* strain, suggesting that in many cases methylation affects gene expression indirectly (34).

Traditionally, it has been difficult, or impossible, to identify and study numerous DNA base modifications that exist in nature, including N^6-adenine DNA methylation, which is the most common modification within bacterial genomes (43). To date, the majority of studies have focused on 5-methylcytosine methylation in eukaryotes using indirect detection methods (44), or on trial and error approaches based on inhibition of DNA restriction to identify N^6-adenine methylation in prokaryotes (e.g. 8). How-

ever, the ongoing development of PacBio's SMRT DNA sequencing method has recently opened the field of bacterial epigenomics with the possibility of genome-wide mapping of methylated sites (45). SMRT methylome analysis detects the rate of DNA base incorporation during sequencing of native unamplified double-stranded input DNA, enabling the identification of a broad spectrum of DNA base modifications, including strand specific N^6-methyladenines (^{m6}A) and N^4-methylcytosines (^{m4}C) that are associated with reliable, robust kinetic signatures (44). For both ModA11 and ModD1, the 6mA sites identified by SMRT sequencing were validated by traditional methods including *in vitro* methylation and restriction enzyme inhibition assays. However, these traditional methods of DNA methylation detection would not have been sufficient to enable the characterization of the flexible ModA11 site. SMRT sequencing is increasingly being used to rapidly identify methylation sites in bacterial genomes (5,10,14,23,46,47), and has paved the way for an increased understanding of bacterial epigenetic gene regulation and its role in pathogenesis.

ACCESSION NUMBERS

The sequences reported in this paper have been deposited in the GenBank database (accession nos. CP007667 and CP007668).

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We would like to thank The Centers for Disease Control and Prevention (CDC) Atlanta, USA for providing the strain M0579.

FUNDING

Australian National Health and Medical Research Council (NHMRC) [Project Grant 2163530 and Career Development Fellowship to K.L.S., Program Grant 565526 to M.P.J.]; Australian Research Council (ARC) [Discovery Project Grant 130103141 to M.P.J]. D.M.C.H was supported by The Meningitis Research Foundation. M.C.J.M was supported by the Wellcome Trust. Y.A.L., R.D.M. and R.J.R. were supported by New England Biolabs (NEB). M.B., T.A.C. and J.K. were supported by Pacific Biosciences. Funding for open access charge: Australian National Health and Medical Research Council Program Grant 565526.
Conflict of interest statement. None declared.

REFERENCES

1. Moxon,R., Bayliss,C. and Hood,D. (2006) Bacterial contingency loci: the role of simple sequence DNA repeats in bacterial adaptation. *Annu. Rev. Genet.*, **40**, 307–333.
2. Srikhanta,Y.N., Maguire,T.L., Stacey,K.J., Grimmond,S.M. and Jennings,M.P. (2005) The phasevarion: a genetic system controlling coordinated, random switching of expression of multiple genes. *Proc. Natl. Acad.Sci. U.S.A.*, **102**, 5547–5551.
3. Srikhanta,Y.N., Fox,K.L. and Jennings,M.P. (2010) The phasevarion: phase variation of type III DNA methyltransferases controls coordinated switching in multiple genes. *Nat. Rev. Microbiol.*, **8**, 196–206.
4. Seib,K.L., Peak,I.R. and Jennings,M.P. (2002) Phase variable restriction-modification systems in Moraxella catarrhalis. *FEMS Immunol. Med. Microbiol.*, **32**, 159–165.
5. Blakeway,L.V., Power,P.M., Jen,F.E., Worboys,S.R., Boitano,M., Clark,T.A., Korlach,J., Bakaletz,L.O., Jennings,M.P., Peak,I.R. *et al.* (2014)ModM DNA methyltransferase methylome analysis reveals a potential role for Moraxella catarrhalis phasevarions in otitis media.FASEB J.,**28**,5197–5207.
6. Fox,K.L., Dowideit,S.J., Erwin,A.L., Srikhanta,Y.N., Smith,A.L. and Jennings,M.P. (2007) Haemophilus influenzae phasevarions have evolved from type III DNA restriction systems into epigenetic regulators of gene expression. *Nucleic Acids Res.*, **35**, 5242–5252.
7. Srikhanta,Y.N., Gorrell,R.J., Steen,J.A., Gawthorne,J.A., Kwok,T., Grimmond,S.M., Robins-Browne,R.M. and Jennings,M.P. (2011) Phasevarion mediated epigenetic gene regulation in Helicobacter pylori. *PLoS One*, **6**, e27569.
8. Srikhanta,Y.N., Dowideit,S.J., Edwards,J.L., Falsetta,M.L., Wu,H.J., Harrison,O.B., Fox,K.L., Seib,K.L., Maguire,T.L., Wang,A.H. *et al.* (2009) Phasevarions mediate random switching of gene expression in pathogenic Neisseria. *PLoS Pathog*, **5**, e1000400.
9. Jen,F.E., Seib,K.L. and Jennings,M.P. (2014) Phasevarions mediate epigenetic regulation of antimicrobial susceptibility in Neisseria meningitidis. *Antimicrob. Agents Chemother.*, **58**, 4219–4221.
10. Manso,A.S., Chai,M.H., Atack,J.M., Furi,L., De Ste Croix,M., Haigh,R., Trappetti,C., Ogunniyi,A.D., Shewell,L.K., Boitano,M. *et al.* (2014) A random six-phase switch regulates pneumococcal virulence via global epigenetic changes. *Nat. Commun.*, **5**, 5055.
11. Seib,K.L., Pigozzi,E., Muzzi,A., Gawthorne,J.A., Delany,I., Jennings,M.P. and Rappuoli,R. (2011) A novel epigenetic regulator associated with the hypervirulent Neisseria meningitidis clonal complex 41/44. *FASEB J.*, **25**, 3622–3633.
12. Gawthorne,J.A., Beatson,S.A., Srikhanta,Y.N., Fox,K.L. and Jennings,M.P. (2012) Origin of the diversity in DNA recognition domains in phasevarion associated modA genes of pathogenic Neisseria and Haemophilus influenzae. *PLoS One*, **7**, e32337.
13. Roberts,R.J., Belfort,M., Bestor,T., Bhagwat,A.S., Bickle,T.A., Bitinaite,J., Blumenthal,R.M., Degtyarev,S., Dryden,D.T., Dybvig,K. *et al.* (2003) A nomenclature for restriction enzymes, DNA methyltransferases, homing endonucleases and their genes. *Nucleic Acids Res.*, **31**, 1805–1812.
14. Murray,I.A., Clark,T.A., Morgan,R.D., Boitano,M., Anton,B.P., Luong,K., Fomenkov,A., Turner,S.W., Korlach,J. and Roberts,R.J. (2012) The methylomes of six bacteria. *Nucleic Acids Res.*, **40**, 11450–11462.
15. McGuinness,B.T., Clarke,I.N., Lambden,P.R., Barlow,A.K., Poolman,J.T., Jones,D.M. and Heckels,J.E. (1991) Point mutation in meningococcal por A gene associated with increased endemic disease. *Lancet*, **337**, 514–517.
16. Maiden,M.C., Bygraves,J.A., Feil,E., Morelli,G., Russell,J.E., Urwin,R., Zhang,Q., Zhou,J., Zurth,K., Caugant,D.A. *et al.* (1998) Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms. *Proc. Natl. Acad.Sci. U.S.A.*, **95**, 3140–3145.
17. Bentley,S.D., Vernikos,G.S., Snyder,L.A., Churcher,C., Arrowsmith,C., Chillingworth,T., Cronin,A., Davis,P.H., Holroyd,N.E., Jagels,K. *et al.* (2007) Meningococcal genetic variation mechanisms viewed through comparative analysis of serogroup C strain FAM18. *PLoS Genet.*, **3**, e23.
18. Budroni,S., Siena,E., Hotopp,J.C., Seib,K.L., Serruto,D., Nofroni,C., Comanducci,M., Riley,D.R., Daugherty,S.C., Angiuoli,S.V. *et al.* (2011) Neisseria meningitidis is structured in clades associated with restriction modification systems that modulate homologous recombination. *Proc. Natl. Acad.Sci. U.S.A.*, **108**, 4494–4499.
19. Seib,K.L., Oriente,F., Adu-Bobie,J., Montanari,P., Ferlicca,F., Giuliani,M.M., Rappuoli,R., Pizza,M. and Delany,I. (2010) Influence of serogroup B meningococcal vaccine antigens on growth and survival of the meningococcus in vitro and in ex vivo and in vivo models of infection. *Vaccine*, **28**, 2416–2427.

20. Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd edn. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
21. Jen,F.E., Warren,M.J., Schulz,B.L., Power,P.M., Swords,W.E., Weiser,J.N., Apicella,M.A., Edwards,J.L. and Jennings,M.P. (2013) Dual pili post-translational modifications synergize to mediate meningococcal adherence to platelet activating factor receptor on human airway cells. *PLoS Pathog*, **9**, e1003377.
22. Rubin,R.A. and Modrich,P. (1977) EcoRI methylase. Physical and catalytic properties of the homogeneous enzyme. *J Biol Chem*, **252**, 7265–7272.
23. Clark,T.A., Murray,I.A., Morgan,R.D., Kislyuk,A.O., Spittle,K.E., Boitano,M., Fomenkov,A., Roberts,R.J. and Korlach,J. (2012) Characterization of DNA methyltransferase specificities using single-molecule, real-time DNA sequencing. *Nucleic Acids Res.*, **40**, e29.
24. R-Development-Core-Team.(2008)R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, http://www.R-project.org.
25. Rutherford,K., Parkhill,J., Crook,J., Horsnell,T., Rice,P., Rajandream,M.A. and Barrell,B. (2000) Artemis: sequence visualization and annotation. *Bioinformatics*, **16**, 944–945.
26. Humbelin,M., Suri,B., Rao,D.N., Hornby,D.P., Eberle,H., Pripfl,T., Kenel,S. and Bickle,T.A. (1988) Type III DNA restriction and modification systems EcoP1 and EcoP15. Nucleotide sequence of the EcoP1 operon, the EcoP15 mod gene and some EcoP1 mod mutants. *J. Mol. Biol.*, **200**, 23–29.
27. Roberts,R.J., Vincze,T., Posfai,J. and Macelis,D. (2010) REBASE–a database for DNA restriction and modification: enzymes, genes and genomes. *Nucleic Acids Res.*, **38**, D234–D236.
28. Virji,M., Makepeace,K., Peak,I.R., Ferguson,D.J., Jennings,M.P. and Moxon,E.R. (1995) Opc- and pilus-dependent interactions of meningococci with human endothelial cells: molecular mechanisms and modulation by surface polysaccharides. *Mol. Microbiol.*, **18**, 741–754.
29. Martin,P., Sun,L., Hood,D.W. and Moxon,E.R. (2004) Involvement of genes of genome maintenance in the regulation of phase variation frequencies in Neisseria meningitidis. *Microbiology*, **150**, 3001–3012.
30. Bart,A., Pannekoek,Y., Dankert,J. and van der Ende,A. (2001) NmeSI restriction-modification system identified by representational difference analysis of a hypervirulent Neisseria meningitidis strain. *Infect. Immun.*, **69**, 1816–1820.
31. Bickle,T.A. and Kruger,D.H. (1993) Biology of DNA restriction. *Microbiol. Rev.*, **57**, 434–450.
32. Vasu,K. and Nagaraja,V. (2013) Diverse functions of restriction-modification systems in addition to cellular defense. *Microbiol. Mol. Biol. Rev.*, **77**, 53–72.
33. Rao,D.N., Dryden,D.T. and Bheemanaik,S. (2014) Type III restriction-modification enzymes: a historical perspective. *Nucleic Acids Res.*, **42**, 45–55.
34. Fang,G., Munera,D., Friedman,D.I., Mandlik,A., Chao,M.C., Banerjee,O., Feng,Z., Losic,B., Mahajan,M.C., Jabado,O.J. *et al.* (2012) Genome-wide mapping of methylated adenine residues in pathogenic Escherichia coli using single-molecule real-time sequencing. *Nat. Biotechnol.*, **30**, 1232–1239.
35. Drozdz,M., Piekarowicz,A., Bujnicki,J.M. and Radlinska,M. (2012) Novel non-specific DNA adenine methyltransferases. *Nucleic Acids Res.*, **40**, 2119–2130.
36. Kamps-Hughes,N., Quimby,A., Zhu,Z. and Johnson,E.A. (2013) Massively parallel characterization of restriction endonucleases. *Nucleic Acids Res.*, **41**, e119.
37. Wei,H., Therrien,C., Blanchard,A., Guan,S. and Zhu,Z. (2008) The Fidelity Index provides a systematic quantitation of star activity of DNA restriction endonucleases. *Nucleic Acids Res.*, **36**, e50.
38. Borgaro,J.G., Benner,N. and Zhu,Z. (2013) Fidelity index determination of DNA methyltransferases. *PLoS One*, **8**, e63866.
39. Cohen,H.M., Tawfik,D.S. and Griffiths,A.D. (2004) Altering the sequence specificity of HaeIII methyltransferase by directed evolution using in vitro compartmentalization. *Protein Eng. Des. Sel.*, **17**, 3–11.
40. Rockah-Shmuel,L. and Tawfik,D.S. (2012) Evolutionary transitions to new DNA methyltransferases through target site expansion and shrinkage. *Nucleic Acids Res.*, **40**, 11627–11637.
41. Timar,E., Groma,G., Kiss,A. and Venetianer,P. (2004) Changing the recognition specificity of a DNA-methyltransferase by in vitro evolution. *Nucleic Acids Res.*, **32**, 3898–3903.
42. Adamczyk-Poplawska,M., Lower,M. and Piekarowicz,A. (2009) Characterization of the NgoAXP: phase-variable type III restriction-modification system in Neisseria gonorrhoeae. *FEMS Microbiol. Lett.*, **300**, 25–35.
43. Wion,D. and Casadesus,J. (2006) N6-methyl-adenine: an epigenetic signal for DNA-protein interactions. *Nat. Rev. Microbiol.*, **4**, 183–192.
44. Korlach,J. and Turner,S.W. (2012) Going beyond five bases in DNA sequencing. *Curr. Opin. Struct. Biol.*, **22**, 251–261.
45. Davis,B.M., Chao,M.C. and Waldor,M.K. (2013) Entering the era of bacterial epigenomics with single molecule real time DNA sequencing. *Current Opin. Microbiol.*, **16**, 192–198.
46. Krebes,J., Morgan,R.D., Bunk,B., Sproer,C., Luong,K., Parusel,R., Anton,B.P., Konig,C., Josenhans,C., Overmann,J. *et al.* (2014) The complex methylome of the human gastric pathogen Helicobacter pylori. *Nucleic Acids Res.*, **42**, 2415–2432.
47. Lluch-Senar,M., Luong,K., Llorens-Rico,V., Delgado,J., Fang,G., Spittle,K., Clark,T.A., Schadt,E., Turner,S.W., Korlach,J. *et al.* (2013) Comprehensive methylome characterization of Mycoplasma genitalium and Mycoplasma pneumoniae at single-base resolution. *PLoS Genet.*, **9**, e1003191.

Author/s:
Seib, KL; Jen, FE-C; Tan, A; Scott, AL; Kumar, R; Power, PM; Chen, L-T; Wu, H-J; Wang, AH-J; Hill, DMC; Luyten, YA; Morgan, RD; Roberts, RJ; Maiden, MCJ; Boitano, M; Clark, TA; Korlach, J; Rao, DN; Jennings, MP

Title:
Specificity of the ModA11, ModA12 and ModD1 epigenetic regulator N-6-adenine DNA methyltransferases of Neisseria meningitidis

Date:
2015-04-30

Citation:
Seib, K. L., Jen, F. E. -C., Tan, A., Scott, A. L., Kumar, R., Power, P. M., Chen, L. -T., Wu, H. -J., Wang, A. H. -J., Hill, D. M. C., Luyten, Y. A., Morgan, R. D., Roberts, R. J., Maiden, M. C. J., Boitano, M., Clark, T. A., Korlach, J., Rao, D. N. & Jennings, M. P. (2015). Specificity of the ModA11, ModA12 and ModD1 epigenetic regulator N-6-adenine DNA methyltransferases of Neisseria meningitidis. NUCLEIC ACIDS RESEARCH, 43 (8), pp.4150-4162. https://doi.org/10.1093/nar/gkv219.

Persistent Link:
http://hdl.handle.net/11343/249575

File Description:
published version