An orphan DNA (cytosine-5-)-methyltransferase in *Vibrio cholerae*

Sanjib Banerjee and Rukhsana Chowdhury

Correspondence
Rukhsana Chowdhury
rukhsana@iicb.res.in

Biophysics Division, Indian Institute of Chemical Biology, 4 Raja S. C. Mullick Road, Calcutta 700 032, India

Received 24 October 2005
Revised 26 December 2005
Accepted 10 January 2006

5-Methyl cytosine (m5C) was detected in genomic DNA of the enteric pathogen *Vibrio cholerae* by HPLC analysis and immunoblotting with m5C-specific antibody. Although cleavage with the restriction endonuclease *EcoRII* revealed the absence of a Dcm homologue in *V. cholerae*, analysis of the genome sequence indicated the presence of a gene, designated in this study as *vchM*, which encodes a DNA (cytosine-5-)-methyltransferase (m5C-MTase) designated M.Vch. M.Vch is not associated with a restriction endonuclease or a mismatch very short patch repair (Vsr)-like endonuclease and is hence an 'orphan' or solitary MTase, although analysis of a phylogenetic tree indicated that related cytosine MTases are all components of restriction-modification systems. M.Vch recognizes and methylates the first 5' C in the degenerate sequence 5'-RCCGGY-3'. RT-PCR analysis suggested that *vchM* gene expression is increased during the stationary phase of growth. During stationary phase, the spontaneous mutation frequency in the *V. cholerae* wild-type strain was significantly higher than in the corresponding *vchM* mutant strain, suggesting that the presence of M.Vch and the absence of a very short patch (VSP) repair-like system imposes upon *V. cholerae* a mutator phenotype.

INTRODUCTION

DNA methylation is the most common epigenetic modification observed in the DNA of prokaryotes and eukaryotes (Jeltsch, 2002). Methylation occurs at the C-5 and N-4 positions of cytosine and at the N-6 position of adenine, and is catalysed by enzymes called DNA methyltransferases (MTases). DNA MTases can broadly be divided into those that methylate the exocyclic amino group of adenines and cytosines (N-MTases) and those that methylate the endocyclic C-5 of cytosine (m5C-MTases). In general, both types of MTases are two-domain proteins consisting of one large and one small domain with a DNA-binding cleft located in the domain interface. They contain amino acid motifs that differ between N- and m5C-MTases (Wilson, 1992). All m5C-MTases, irrespective of their phylogenetic origin, contain 10 conserved amino acid motifs separated by regions variable in size and amino acid sequence in different enzymes (Kumar *et al.*, 1994). The N-terminal region of the enzyme includes conserved motifs I–VIII containing the binding site for the universal methyl group donor S-adenosyl methionine (AdoMet) and the catalytic centre of the enzyme. Located beyond motif VIII is a variable region containing the target recognition domains, followed by the C-terminal motifs IX and X.

Abbreviations: AdoMet, S-adenosyl methionine; m5C, 5-methyl cytosine; MTase, methyltransferase; R-M, restriction-modification; VSP, very short patch (repair system); Vsr, very short patch repair (endonuclease).

In prokaryotes, a majority of DNA MTases are components of restriction-modification (R-M) systems, in conjunction with cognate restriction endonucleases. The sequence context of methylation serves to distinguish between self and non-self DNA and protects bacteria from 'foreign' invaders like bacteriophages, transposons and other genomic parasites. However, not all MTases are components of R-M systems. Certain adenine-MTases (Dam and CcrM) control gene expression and regulate cellular processes like post-replicative mismatch repair, DNA replication timing and cell cycle regulation (Messer & Noyer-Weidner, 1988; Reisenauer *et al.*, 1999). In some bacteria, cytosine-MTases (Dcm) are associated with very short patch repair (Vsr) endonucleases (Sohail *et al.*, 1990). Spontaneous hydrolytic deamination of m5C to form thymine can lead to the production of T:G mismatches. The mutagenic potential of T:G mismatches is counteracted by the Vsr endonuclease which recognizes and removes T from T:G mismatches in DNA.

Analysis of more than 300 bacterial and archaeal genomes in the REBASE database (Roberts *et al.*, 2005; <http://rebase.neb.com/rebase>) has revealed that about 90% of the genomes have at least one R-M system and about 80% contain multiple R-M systems, most of which have not been biochemically characterized. *In silico* genome analysis suggests that in many of the R-M systems, the restriction endonuclease component is inactive but the cognate MTase retains full activity. However, since very few such 'orphan' methylases have been characterized, their function is not

known and it is not clear why they are maintained in the genome. In this context, the single orphan MTase in *V. cholerae* has been characterized in this study.

V. cholerae is a Gram-negative enteric pathogen that causes the diarrhoeal disease cholera (Kaper *et al.*, 1995). DNA adenine methylation has been reported in *V. cholerae*, and a role of the *V. cholerae* DNA adenine methyltransferase (Dam) in virulence (Julio *et al.*, 2001) and chromosome replication (Egan & Waldor, 2003) has been demonstrated. Although *V. cholerae* possesses an efficient Dam-directed mismatch repair pathway (Bera *et al.*, 1989), it lacks homologues of *Escherichia coli* Dcm and Vsr endonuclease (Bhakat *et al.*, 1999), known to be associated with the very short patch (VSP) repair system. An examination of the *V. cholerae* strain N16961 genome sequence database (www.tigr.org) revealed the presence of a single ORF annotated as a putative cytosine MTase. No gene encoding a restriction endonuclease or a Vsr-like endonuclease was detected in the vicinity of the cytosine MTase. In this report we present data demonstrating that *V. cholerae* strain O395 also contains an 'orphan' or solitary cytosine MTase that

methyates the first C in the sequence 5'-RCCGGY-3' and contributes to the generation of spontaneous mutants in a non-dividing cell population.

METHODS

Bacterial strains and plasmids. *V. cholerae* O1 strains O395 and CM319 ($\Delta vchM$) and other *V. cholerae* strains listed in Table 1 were used in the present study. The strains listed in Table 1 were obtained from the National Institute of Cholera and Enteric Diseases, Calcutta, India. All strains were grown in LB medium containing 1% tryptone (Difco), 0.5% yeast extract (Difco), 0.5% NaCl and stored at -70°C in 20% (v/v) glycerol. The suicide vector pCVD442, used for construction of the deletion mutant, was maintained in *E. coli* strain SM10 λ pir (Donnenberg & Kaper, 1991).

HPLC analysis. Chromosomal DNA (5 μg) from *V. cholerae* strain O395 or CM319 was digested with snake venom phosphodiesterase (50 U) followed by treatment with calf intestinal phosphatase (50 U). The samples were extracted with phenol/chloroform and analysed by HPLC using an RP-18 column equilibrated with 50 mM potassium phosphate buffer (pH 5.9). After injection of the samples, the column was washed with 5 ml of the same buffer and eluted with a linear gradient of 50 mM potassium phosphate buffer

Table 1. Presence of *vchM* gene in *V. cholerae* strains of different biotypes and serogroups

<i>V. cholerae</i> strain	Serogroup/biotype	Characteristic	<i>vchM</i> *
O395	O1/Classical	Hypertoxinogenic	+
569B	O1/Classical	Hypertoxinogenic	+
154	O1/Classical	Medium toxinogenic	+
E-82	O1/Classical	Non-toxinogenic environmental	—
N16961	O1/El Tor	Hypertoxinogenic	+
C6709	O1/El Tor	Hypertoxinogenic	+
GP-5	O1/El Tor	Non-toxinogenic	+
MAK757	O1/El Tor	Medium toxinogenic	+
E7946	O1/El Tor	Hypertoxinogenic	+
EW6	O1/El Tor	Low toxinogenic aquatic	+
CO457	Post-O139/El Tor	—	+
VC44	Pre-O139/El Tor	—	+
AP1	O139	Hypertoxinogenic clinical	+
AP2	O139	Hypertoxinogenic clinical	+
SG24	O139	Clinical	+
VO15	O139	Clinical	+
MO1	O139	Clinical	+
PO2	O139	Clinical	+
V94	Non O1	Clinical	—
VC09	Non O1	Clinical	+++
VCE22	Non O1	Environmental	++
VCE232	Non O1	Environmental	+
VCE299	Non O1	Environmental	++
VCE309	Non O1	Environmental	+++
VC042	Non O1	Clinical	+++

*The *vchM* gene was PCR-amplified from chromosomal DNA of different *V. cholerae* strains and examined by agarose gel electrophoresis. + denotes the presence of a single PCR product corresponding to the *vchM* gene; ++ denotes the presence of two PCR products; +++ denotes the presence of multiple PCR products.

(pH 5.9) and 50% (v/v) acetonitrile at a flow rate of 1 ml min⁻¹ (Eick *et al.*, 1983). 5-Methyl deoxycytidine and deoxycytidine (Sigma) were used as standard markers.

Immunoblotting. Chromosomal or plasmid DNA (5 µg) was denatured by boiling for 5 min and cooled rapidly on ice. The DNA was spotted onto nitrocellulose paper (Millipore) pre-soaked in 10 × SSC, dried at room temperature and baked for 2 h at 80 °C. The blot was blocked with 5% BSA and incubated overnight with mouse polyclonal antibody to m5C (1:10 000; Calbiochem), washed in Tris-buffered saline (TBS) and detected using HRP-conjugated rabbit anti-mouse antibody (1:5000; Jackson Laboratory) and enhanced chemiluminescence.

Expression and purification of His₆-M.Vch. The complete ORF of the *V. cholerae* *vchM* gene was PCR-amplified and inserted in-frame at the *Bam*HI/*Kpn*I sites of the expression vector pQE30 (Qiagen) to create an N-terminal His₆-tagged protein. The resulting plasmid, designated pQE*vchM*, was introduced into *E. coli* M15 and the His-tagged fusion protein was isolated following the manufacturer's instructions.

In vitro methylation assay. Plasmid pBR322 DNA isolated from *E. coli* strain GM31 (*dcm*⁻; Sohail *et al.*, 1990) was linearized with *Eco*RI and used as substrate for an *in vitro* methylation reaction (Marks *et al.*, 2003) with *V. cholerae* M.Vch. Briefly, 250 ng linearized pBR322 DNA was treated with M.Vch in the presence of 200 µM AdoMet for 2 h. Plasmid DNA was purified from the reaction mixture, cleaved with *Bsr*FI (5 U, 3 h) and electrophoresed on an 8% polyacrylamide gel. Ethidium bromide stained gels were visualized using a Gel Doc 1000 system (Bio-Rad).

Construction of *V. cholerae* *vchM* mutant. The *V. cholerae* *vchM* gene was cloned as a 1523 bp fragment in vector pBluescript KS to give plasmid pSB319. An internal 935 bp region was deleted from the cloned gene by digestion with *Nde*I and *Nsi*I followed by religation. The resulting plasmid was digested with *Sal*I and *Sac*I, and a 651 bp fragment was obtained that contained -105 bp to +177 bp of the *vchM* ORF fused to a 306 bp region downstream of the ORF. This fragment was cloned into the positive selection suicide vector pCVD442 (Ap^r, *sacB*) and transformed into *E. coli* SM10 λpir. The plasmid was conjugally transferred into *V. cholerae* O395 (streptomycin-resistant, Sm^r) and ampicillin-resistant (Ap^r) Sm^r colonies arising due to a single recombination event were selected and spread on plates containing 10% sucrose. Sm^r Ap^s and sucrose-resistant colonies were selected and Southern hybridization and nucleotide sequencing was used to confirm that deletion had occurred in the *vchM* gene.

RNA isolation and RT-PCR. For isolation of RNA, *V. cholerae* strains O395 or CM319 were grown to exponential phase (3 h) or overnight in LB at 37 °C. Total RNA was extracted and purified using guanidinium isothiocyanate (Ausubel *et al.*, 1989). The RNA was treated with RNase-free DNase I (1 U µg⁻¹, amplification grade; Invitrogen) in the presence of an RNase inhibitor (RNasin; Gibco-BRL) and RT-PCR was performed using a single tube RT-PCR kit (Gibco-BRL). DNase-treated RNA (200 ng) was used in all reactions. Samples were removed after 20, 25, 30 and 35 cycles (94 °C, 30 s; 55 °C, 30 s; 72 °C, 30 s; followed by 7 min extension at 72 °C), electrophoresed on 2% agarose gels with ethidium bromide and analysed using a Gel Doc 1000 system (Bio-Rad). The data obtained in the linear range of PCR amplification were considered for analysis. Genomic DNA served as a positive control and RNA that had been treated with DNase, but not reverse transcribed, was used as negative control. Each set of experiments was performed at least thrice.

Determination of spontaneous mutation frequency. Overnight cultures of *V. cholerae* strains O395 and CM319 were diluted 1:10⁶

in fresh LB medium (approx. 500 c.f.u. ml⁻¹) and grown at 37 °C. Samples were removed in the exponential phase of growth, the number of c.f.u. ml⁻¹ was assayed and 6 ml culture was concentrated by centrifugation and plated on LB agar plates containing 10 µg rifampicin (Rif) ml⁻¹ or 3 µg novobiocin (Nov) ml⁻¹. Similarly, at 24 h, the number of c.f.u. ml⁻¹ was determined and 1 ml culture was plated on antibiotic plates. Spontaneous mutation frequency was determined as the ratio of the number of colonies obtained on antibiotic plates to the total number of cells plated.

RESULTS AND DISCUSSION

Presence of m5C in *V. cholerae* DNA

It has previously been reported that *V. cholerae* does not contain a Dcm-like m5C-MTase (Bhakat *et al.*, 1999). However, plasmid pBR322 isolated from *V. cholerae* could be transformed into *E. coli* strains carrying functional McrAB at a very low frequency (less than 10⁻⁹), but could be transformed into an *E. coli* *mcrA*⁻ *B*⁻ strain, XL1Blue MRF', at frequencies greater than 10⁻⁷. Since the *E. coli* McrAB system restricts DNA containing m5C (Raleigh & Wilson, 1986), this result indicated that pBR322 isolated from *V. cholerae* contains m5C. Estimation of m5C in *V. cholerae* genomic DNA by HPLC analysis of nucleotides released by extensive digestion of genomic DNA with nucleases indicated the presence of m5C (Fig. 1a). The presence of m5C in *V. cholerae* chromosomal DNA and plasmids isolated from *V. cholerae* was also demonstrated by immunoblotting using anti-m5C antibodies (Fig. 1b). The results obtained indicated that, although *V. cholerae* O395 genomic DNA and plasmid pBR322 DNA isolated from strain O395 contained m5C, the proportion is much lower than in calf thymus

Fig. 1. (a) Genomic DNA from *V. cholerae* O395 (top panel) and *V. cholerae* CM319 (middle panel) were digested with snake venom phosphodiesterase and analysed by HPLC for detection of m5C. The positions of m5C and C used as markers are shown in the lower panel. (b) Immunological detection of m5C in genomic DNA (top panel) and plasmid DNA (lower panel) from *V. cholerae* strains O395 (I) and CM319 (II). (III) *E. coli* DNA (top panel) and calf thymus DNA (lower panel) were used as controls.

DNA and only slightly lower than in *E. coli* K-12 DNA used as controls in these experiments.

***V. cholerae* cytosine methyltransferase (*vchM*) gene**

Examination of the *V. cholerae* strain N16961 genome sequence database revealed the presence of a single gene in chromosome 2 (accession no. Q9KMW8, TIGR locus VCA0198; www.tigr.org) annotated as a putative cytosine-MTase. To discover if the gene was also present in the *V. cholerae* strain O395, primers flanking the gene were designed based on the genome sequence of strain N16961 and used to PCR-amplify homologous sequences, if any, from *V. cholerae* O395 genomic DNA. A 1.5 kb PCR product was obtained, the nucleotide sequence of which was found to be identical to that of the putative cytosine-MTase in strain N16961 (TIGR locus VCA0198). The *V. cholerae* O395 putative cytosine-MTase was designated *vchM*. The predicted amino acid sequence of the gene product contains all 10 motifs characteristic of m5C-MTases and the order of the conserved and variable sequences is also conserved. The *vchM* gene product, designated M.Vch, shares high sequence homology with m5C-MTases from *Neisseria meningitidis* (58 % identity, 74 % positive), *Citrobacter freundii* (56 % identity, 71 % positive) and *Nostoc punctiforme* (54 % identity, 71 % positive), but no significant homology with N4-cytosine and N6-adenine MTases.

Most bacterial cytosine MTases reported are associated with cognate restriction endonucleases and some Dcm-like cytosine MTases are associated with VSP repair nicking endonucleases like Vsr. An unidentified ORF is present adjacent to the putative cytosine MTase gene of strain N16961 (TIGR locus VCA0199; www.tigr.org) and also the *vchM* cytosine MTase gene of strain O395 (see below), but it

bears no homology with *vsr*. Indeed, *dcm-vsr* homologues have been reported to be absent in *V. cholerae* (Bhakat *et al.*, 1999). Also, the unidentified ORF is unlikely to encode a restriction enzyme since an in-frame deletion mutation in the *vchM* cytosine methylase gene could be constructed (described below). Thus, the *V. cholerae* M.Vch is an 'orphan' or solitary cytosine MTase. However, a phylogenetic tree for homologues of *V. cholerae* M.Vch (Fig. 2) indicates that many homologues of M.Vch are components of R-M systems.

What could be the origin of the M.Vch MTase? The fact that the *vchM* gene region has a lower (G+C) content (39.16 mol%) than the remainder of the genome (47.2 mol%) indicates that this gene may have been acquired horizontally. Indeed, a transposase gene (IS1004 transposase) is located in the vicinity of the cytosine MTase gene in strain N16961 (www.tigr.org).

Cloning of *V. cholerae* *vchM* gene and construction of mutant

The *vchM* gene from *V. cholerae* strain O395 was PCR-amplified as a 1523 bp fragment and cloned into plasmid pBluescript KS to give plasmid pSB319. The plasmid could be transformed into *E. coli* strain XL1Blue MRF' (*mcrAB*⁻) but no transformants were obtained using *E. coli* strains XL1Blue or DH5 α . Nucleotide sequencing of the cloned fragment revealed that it contained a 1152 bp complete ORF identical to the cytosine-MTase gene of N16961 (locus VCA0198) and the 5' segment of an ORF identical to a hypothetical gene of strain N16961 (locus VCA0199) located immediately downstream. Thus, both the gene sequence and genomic organization of the *vchM* gene of strain O395 are identical to the cytosine-MTase encoding

Fig. 2. Phylogenetic tree for M.Vch and related m5C-MTases. A neighbour-joining tree from the amino acid sequences is shown with percentage bootstrap values at the branches. The name of each MTase is followed by the recognition sequence, name of the organism and gene organization. Based on data taken from REBASE and NCBI databases (<http://rebase.neb.com/rebase>; www.ncbi.nlm.nih.gov).

Fig. 3. (a) RT-PCR was performed with RNA isolated from strains O395 (lane 1) and CM319 (lane 2) grown to the exponential (lanes 1 and 2) or stationary (lanes 3 and 4) phase for estimation of *vchM* expression. (b) RT-PCR for estimation of *dapA* gene expression was performed with RNA from *V. cholerae* O395 grown to exponential (lane 1) or stationary (lane 2) phase.

gene of strain N16961 (VCA0198) in the *V. cholerae* genome sequence database (www.tigr.org).

An in-frame deletion in the *vchM* gene was constructed in the positive selection vector pCVD442 as described in Methods. The resulting plasmid was used for allelic exchange in strain O395 to generate strain CM319 ($\Delta vchM$). The mutation was confirmed by Southern hybridization, nucleotide sequencing and by PCR using primers flanking the *vchM* gene (data not shown). RT-PCR analysis also indicated that *vchM*-specific transcripts were not present in the mutant strain CM319 (Fig. 3a, lanes 2 and 4). Furthermore, m5C could not be detected in genomic DNA from the mutant strain CM319 by HPLC (Fig. 1a), and immunoblotting with m5C-specific antisera did not give a signal with chromosomal and plasmid DNA isolated from strain

CM319, although a specific signal was obtained with DNA from the wild-type parent strain (Fig. 1b). These results indicated that M.Vch is the only m5C-MTase in *V. cholerae* strain O395.

In view of the fact that *vchM* has been annotated as a 'putative' MTase in the *V. cholerae* genome sequence database (www.tigr.org), RT-PCR was used to examine whether the gene is expressed in *V. cholerae* strain O395. RNA was isolated from *V. cholerae* O395 grown in LB medium at 37 °C, and *vchM*-specific transcripts were estimated by RT-PCR. The results obtained indicated that, although significant *vchM* expression was detected in the exponential phase of growth, the expression was two- to threefold higher in stationary phase cells (Fig. 3a, lanes 1 and 3). Expression of the housekeeping gene *dapA* decreased in the stationary phase (Fig. 3b, lanes 1 and 2).

Identification of the M.Vch methylation site

The recognition sequence of M.Vch was identified by analysis of cytosine methylation sensitive restriction endonuclease digestion patterns of plasmid and chromosomal DNA from the *V. cholerae* strains O395 and CM319 ($\Delta vchM$). *EcoRII*, which recognizes unmethylated 5'-CCWGG-3' sites, the target for Dcm methylation, completely cleaves chromosomal DNA from both strains O395 and CM319 (Fig. 4e), indicating that 5'-CCWGG-3' sites are not methylated in *V. cholerae* and suggesting that the M.Vch MTase recognizes a different target. The restriction endonucleases *MspI* and *HpaII* both recognize the sequence 5'-CCGG-3'. While *HpaII* will not cleave if either C residue in the recognition sequence is methylated, *MspI* cleavage is blocked by methylation of the 5' C (Roberts *et al.*, 2005). All *HpaII*/*MspI* recognition sites in plasmid pBR322 were cleaved by both *HpaII* and *MspI* in pBR322 isolated from

Fig. 4. Restriction digestion pattern analysis. Plasmid pBR322 DNA from *E. coli* (lane 1), *V. cholerae* O395 (lane 2) and *V. cholerae* CM319 (lane 3) digested with *MspI* (a), *HpaII* (b) and *BsrF1* (c). The arrows denote fragments missing from and arrowheads denote new fragments appearing in the *HpaII*/*MspI*-digested pBR322 DNA isolated from *V. cholerae* O395 (see text). (a and b) The 631 bp new fragment in lane 2 could not be distinguished from the 622 bp fragment in lanes 1 and 3. One of the two 160 bp fragments present in lanes 1 and 3 was missing in lane 2. (c) A 9 bp fragment could not be detected. Lane M, 100 bp DNA ladder. (d and e) Digestion of genomic DNA from *V. cholerae* strains O395 (lane 1) and CM319 (lane 2) with *BsrF1* (d) and *EcoRII* (e).

E. coli DH5 α , yielding fragments of expected sizes (Figs 4a and b, lanes 1). However, when pBR322 isolated from strain O395 was digested with the enzymes *Hpa*II or *Msp*I, nine fragments of 622, 404, 201, 160, 123, 90, 76, 34 and 26 bp could not be detected; instead additional fragments of sizes 631, 438, 326, 227 and 147 bp were obtained (Figs 4a and b, lanes 2). Analysis of the molecular masses of the additional fragments obtained indicated that they represent fused fragments. The presence of the fused fragments suggested that certain 5'-CCGG-3' sites in pBR322 from strain O395 were resistant to cleavage by *Hpa*II and *Msp*I. Further analysis of these refractory sites reveal that they fall within the sequence 5'-RCCGGY-3'. Failure of both *Hpa*II and *Msp*I to cleave at 5'-RCCGGY-3' sites suggested that the first 5' C within the sequence 5'-RCCGGY-3' is methylated by M.Vch, since methylation of only this cytosine blocks both *Msp*I and *Hpa*II cleavage (Kwiatek *et al.*, 2004; Roberts *et al.*, 2005). When plasmid pBR322 isolated from strain CM319 ($\Delta vchM$) was digested with *Hpa*II or *Msp*I, all potential sites were completely cleaved, yielding a pattern similar to that obtained in *Hpa*II and *Msp*I digests of pBR322 DNA from *E. coli* (Figs 4a and b, lanes 1 and 3), indicating absence of m5C in the recognition sites.

The restriction endonuclease *Bsr*F1 recognizes the sequence 5'-RCCGGY-3' and does not cleave substrate DNA containing a methylated C residue. Chromosomal DNA from the wild-type strain O395 was resistant to cleavage by *Bsr*F1, while that from strain CM319 ($\Delta vchM$) was completely cleaved by *Bsr*F1 (Fig. 4d), indicating that *Bsr*F1 sites were methylated by M.Vch in the wild-type *V. cholerae* DNA. To further examine the substrate specificity of M.Vch, plasmid pBR322, which contains seven *Bsr*F1 recognition sites, was isolated from *V. cholerae* O395 and the *vchM* mutant strain CM319, and digested with *Bsr*F1. Plasmid pBR322 isolated from strain O395 was not cleaved by *Bsr*F1, indicating the presence of methylation in the recognition sequence (Fig. 4c, lane 2). However, pBR322 isolated from strain CM319 ($\Delta vchM$) was completely digested by *Bsr*F1 (Fig. 4c, lane 3), yielding a pattern similar to that obtained in the *Bsr*F1 digest of pBR322 isolated from *E. coli* DH5 α (Fig. 4c, lane 1).

Taken together, the *Bsr*F1, *Hpa*II and *Msp*I digestion patterns suggest that *V. cholerae* M.Vch recognizes the sequence 5'-RCCGGY-3' and methylates the first 5' C residue within the sequence.

***In vitro* methylation by M.Vch**

To examine if M.Vch directly methylates the 5'-RCCGGY-3' sites in plasmid pBR322, an *in vitro* experiment was carried out by incubating pBR322 with M.Vch followed by digestion with the restriction endonuclease *Bsr*F1. The results obtained indicate that, in the presence of AdoMet, M.Vch can methylate the 5'-RCCGGY-3' sites in pBR322, rendering them refractory to cleavage by *Bsr*F1 (Fig. 5 lane 1). pBR322 incubated without M.Vch was cleaved by *Bsr*F1 (Fig. 5, lane 2).

Fig. 5. *In vitro* methylation assay. pBR322 DNA was incubated with (lane 1) or without (lane 2) M.Vch as described in Methods and digested with *Bsr*F1. Lane M, 100 bp DNA ladder.

Presence of *vchM* in other *V. cholerae* strains

A number of *V. cholerae* strains of different serogroups and biotypes were examined for the presence of the *vchM* gene by PCR amplification using *vchM* internal primers corresponding to the highly conserved DNA binding and AdoMet binding domains. Of the 12 *V. cholerae* serogroup O1 strains of the classical and El Tor biotypes examined, 11 contained the *vchM* gene (Table 1). The gene was present in all six strains belonging to *V. cholerae* serogroup O139 examined and in six of the seven non-O1 non-O139 strains examined (Table 1). Southern blot analysis was used to confirm the absence of *vchM* in strains E82 and V94 (data not shown). In some non-O1 strains, two or more bands were obtained (Table 1), suggesting that they may contain other cytosine MTases in addition to *vchM*. To examine if the *vchM* gene is expressed in all strains containing the *vchM* gene, RT-PCR analysis was performed with randomly selected strains of both O1 and non-O1 serogroups. *vchM* expression was detected in all strains carrying the *vchM* gene, but not in strains in which the gene is absent (Table 1, Fig. 6).

Fig. 6. RT-PCR analysis of *vchM* expression in *V. cholerae* strains N16961 (lane 1), C6709 (lane 2), E82 (lane 3), VCE232 (lane 4) and V94 (lane 5). Lane M, 100 bp DNA ladder.

Spontaneous mutation frequency

Spontaneous deamination of m5C results in the formation of T:G mismatches which, if not repaired, lead to C→T mutations. In the exponential phase of growth, T:G mismatches are repaired, together with a variety of other mismatches, by the general sequence non-specific methyl-directed mismatch repair (MMR) pathway, while in the stationary phase T:G mismatches are specifically repaired by the VSP repair system (Bhagwat & Lieb, 2002). The spontaneous mutation frequencies of *V. cholerae* strains O395 and CM319 to rifampicin resistance and novobiocin resistance were measured in the exponential and stationary phases of growth (Table 2). Nucleotide sequencing has revealed that the sequences of the genes *rpoB* and *gyrB*, mutations in which confer resistance to rifampicin and novobiocin respectively, in strain O395 are identical to those

of strain N16961, the genome of which has been completely sequenced (www.tigr.org), and contain two and four M.Vch methylation sites, respectively. In the exponential phase, the *V. cholerae* wild-type and *vchM* mutant strains had comparable spontaneous mutation frequencies (Table 2), indicating that T:G mismatches arising from deamination of m5C in the parent strain were efficiently repaired. This result was consistent with reports that an efficient MMR repair system is present in *V. cholerae* (Bera *et al.*, 1989). In the stationary phase of growth, the spontaneous mutation frequency of the parent strain was significantly higher than that of the corresponding *vchM* mutant strain CM319 (Table 2). One explanation for this difference may be that, since the Vsr-like endonuclease, which corrects T:G mismatches in non-dividing cells, might be absent in *V. cholerae* (Bhakat *et al.*), the T:G mismatches arising from deamination of m5C in the parent strain are not efficiently repaired, giving rise to C→T mutations within the M.Vch methylation sites (5'-RCCGGY-3') in the *rpoB* and *gyrB* genes. Since the T:G mismatches do not arise in the *vchM* mutant strain, the spontaneous mutation frequency of the mutant is relatively lower. To test this hypothesis, fragments of the *rpoB* and *gyrB* genes (carrying the 5'-RCCGGY-3' sites), from Rif^R and Nov^R mutants of strain O395 and CM319, were PCR-amplified and sequenced. One of the three Rif^R mutants of strain O395 examined carried a C→T mutation at the first C residue in the first 5'-RCCGGY-3' site within the *rpoB* gene, while two of five Nov^R mutants of strain O395 examined contained a C→T change at the first C residue of the first and fourth 5'-RCCGGY-3' sites within the *gyrB* gene. An equal number of Rif^R and Nov^R mutants of strain CM319 were examined, but none carried a C→T change within the 5'-RCCGGY-3' sites of *rpoB* or *gyrB*. Thus, the presence of M.Vch and the absence of a VSP repair-like system impose upon *V. cholerae* a mutator phenotype. It

Table 2. Spontaneous mutation frequency in *V. cholerae* strains

<i>V. cholerae</i> strain	Growth phase	10 ⁹ × Mutation frequency*		Relative frequency†	
		Rif ^R	Nov ^R	Rif ^R	Nov ^R
O395	Exponential	4.0 ± 0.33	4.6 ± 0.4	1.00	1.00
	Stationary	18.0 ± 0.22	19.0 ± 0.25	4.55	4.18
CM319	Exponential	3.5 ± 0.45	3.6 ± 0.8	1.00	1.00
	Stationary	3.0 ± 0.35	2.6 ± 0.4	0.81	0.72
N16961	Exponential	2.6 ± 0.15	2.3 ± 0.25	1.00	1.00
	Stationary	7.4 ± 0.42	5.6 ± 0.3	2.85	2.44
C6709	Exponential	2.3 ± 0.18	1.5 ± 0.12	1.00	1.00
	Stationary	5.2 ± 0.22	5.9 ± 0.2	2.25	3.86
E82	Exponential	3 ± 0.12	1.3 ± 0.12	1.00	1.00
	Stationary	3.3 ± 0.27	1.7 ± 0.32	1.12	1.32
V94	Exponential	6.2 ± 0.18	5.0 ± 0.12	1.00	1.00
	Stationary	5.3 ± 0.22	4.4 ± 0.2	0.85	0.88

*The ratio of the c.f.u. on plates with antibiotics to the total c.f.u. of cells plated. The results represent the mean ± SD of five independent experiments.

†The mutation frequency in the exponential and stationary phases of growth of each strain was compared.

may be mentioned in this context that the *V. cholerae* genome (4.03 Mb) contains only 2172 M.Vch methylation sites (5'-RCCGGY-3') while 5906 such sites are present in the *E. coli* genome (4.64 Mb).

The spontaneous mutation frequency in the exponential and stationary phases of growth was also determined for *V. cholerae* strains N16961 and C6709, which carry a functional *vchM* gene, and strains E82 and V94, which lack the *vchM* gene (Table 1). The results obtained indicate that in the strains N16961 and C6709 the spontaneous mutation frequency in stationary-phase cells was higher than in cells grown to the exponential phase (Table 2). However, the spontaneous mutation frequency of strains E82 and V94 was similar in the exponential and stationary phases of growth.

What could be the biological role of the M.Vch MTase? It is known that mutations occur in non-dividing microbial populations and this phenomenon provides a source of genetic variants from which individuals with a variety of advantages may be selected in evolution. It is attractive to hypothesize that in *V. cholerae* the presence of M.Vch and absence of Vsr may contribute to the generation of such mutants in non-growing conditions, under which *V. cholerae* may be presumed to spend most of the time in its natural aquatic habitats. Furthermore, there is evidence that R-M systems may behave selfishly in that their loss from a cell may lead to cell killing through restriction of the genome. It has been suggested that, in bacteria containing orphan MTases, prior methylation of a specific sequence in the genome may afford protection against parasitism by R-M systems recognizing the same sequences (Takahashi *et al.*, 2002). Whether the M.Vch MTase, which methylates the sequence 5'-RCCGGY-3', has a similar role in conferring protection against parasitism by R-M complexes needs to be investigated.

ACKNOWLEDGEMENTS

We thank all members of the Biophysics Division for cooperation, encouragement and helpful discussions during the study, and I. Guha Thakurta and P. Majumdar for excellent technical support. We are grateful to K. E. Klose, University of Texas Health Science Center, San Antonio, Texas, USA, for the generous gift of the plasmid pCVD442. S.B. is grateful to the Council of Scientific and Industrial Research, Government of India, for a research fellowship. The work was supported by Network Program SMM003.

REFERENCES

Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A. & Struhl, K. (editors) (1989). *Current Protocols in Molecular Biology*. New York: Wiley.

Bera, T. K., Ghosh, S. K. & Das, J. (1989). Cloning and characterization of *mutL* and *mutS* genes of *Vibrio cholerae*: nucleotide sequence of the *mutL* gene. *Nucleic Acids Res* **17**, 6241–6251.

Bhagwat, A. S. & Lieb, M. (2002). Cooperation and competition in mismatch repair: very short-patch repair and methyl-directed mismatch repair in *Escherichia coli*. *Mol Microbiol* **44**, 1421–1428.

Bhakat, K. K., Sharma, S. & Das, J. (1999). The *mutK* gene of *Vibrio cholerae*: a new gene involved in DNA mismatch repair. *J Bacteriol* **181**, 879–883.

Donnenberg, M. S. & Kaper, J. B. (1991). Construction of an *eae* deletion mutant of enteropathogenic *Escherichia coli* by using a positive selection suicide vector. *Infect Immun* **59**, 4310–4317.

Egan, E. S. & Waldor, M. K. (2003). Distinct replication requirements for the two *Vibrio cholerae* chromosomes. *Cell* **114**, 521–530.

Eick, D., Fritz, H. & Doerfler, W. (1983). Quantitative determination of 5-methyl cytosine in DNA by reverse-phase high-performance liquid chromatography. *Anal Biochem* **135**, 165–171.

Jeltsch, A. (2002). Beyond Watson and Crick: DNA methylation and molecular enzymology of DNA methyltransferases. *ChemBiochem* **3**, 274–293.

Julio, S. M., Heithoff, D. M., Provenzano, D., Klose, K. E., Sinsheimer, R. L., Low, D. A. & Mahan, M. J. (2001). DNA adenine methylase is essential for viability and plays a role in the pathogenesis of *Yersinia pseudotuberculosis* and *Vibrio cholerae*. *Infect Immun* **69**, 7610–7615.

Kaper, J. B., Morris, J. G. & Levine, M. M. (1995). Cholera. *Clin Microbiol Rev* **8**, 48–86.

Kumar, S., Cheng, X., Klimasauskas, S., Mi, S., Posfai, J., Roberts, R. J. & Wilson, G. G. (1994). The DNA (cytosine-5) methyltransferases. *Nucleic Acids Res* **22**, 1–10.

Kwiatkiewicz, A., Kobes, M., Olejnik, K. & Piekarczyk, A. (2004). DNA methyltransferases from *Neisseria meningitidis* and *Neisseria gonorrhoeae* FA1090 associated with mismatch nicking endonucleases. *Microbiology* **150**, 1713–1722.

Marks, P., McGeehan, J., Wilson, G., Errington, N. & Kneale, G. (2003). Purification and characterization of a novel DNA methyltransferase, M.Ahd1. *Nucleic Acids Res* **31**, 2803–2810.

Messer, W. & Noyer-Weidner, M. (1988). Timing and targeting the biological functions of Dam methylation in *E. coli*. *Cell* **54**, 735–737.

Raleigh, E. A. & Wilson, G. (1986). *Escherichia coli* K-12 restricts DNA containing 5-methylcytosine. *Proc Natl Acad Sci U S A* **83**, 9070–9074.

Reisenauer, A., Kahng, L. S., McCollum, S. & Shapiro, L. (1999). Bacterial DNA methylation: a cell cycle regulator. *J Bacteriol* **181**, 5135–5139.

Roberts, R. J., Vincze, T., Posfai, J. & Macelis, D. (2005). REBASE – restriction enzymes and DNA methyltransferases. *Nucleic Acids Res* **33**, D230–D232.

Sohail, A., Lieb, M., Dar, M. & Bhagwat, A. S. (1990). A gene required for very short patch repair in *Escherichia coli* is adjacent to the DNA cytosine methylase gene. *J Bacteriol* **172**, 4214–4221.

Takahashi, N., Naito, Y., Handa, N. & Kobayashi, I. (2002). A DNA methyltransferase can protect the genome from post-disturbance attack by a restriction-modification gene complex. *J Bacteriol* **184**, 6100–6108.

Wilson, G. G. (1992). Amino acid sequence arrangements of DNA methyltransferases. *Methods Enzymol* **216**, 259.