Sequence analysis of the *Ngo*PII methyltransferase gene from *Neisseria gonorrhoeae* P9: homologies with other enzymes recognizing the sequence 5'-GGCC-3'

Kevin M.Sullivan¹ and Jon R.Saunders

Department of Microbiology, University of Liverpool, Liverpool, L69 3BX, UK

Received February 22, 1988; Revised and Accepted March 4, 1988

Accession no.X06965

ABSTRACT

Recombinant plasmids harbouring the functional *M.Ngo*PII methyltransferase (specificity 5'-GGCC-3') were isolated from amplified gene libraries of gonococcal chromosomal DNA cloned in pBR322 and in *Escherichia coli* RR1. The *M.Ngo*PII gene was localized by sub-cloning and the nucleotide sequence of a cloned 1.6 kb segment of *Neisseria gonorrhoeae* DNA harbouring the methylase gene was determined. This data, coupled with sub-cloning experiments and *in vitro* transcription-translation studies, indicates a theoretical size of 38.5 kd for the methylase protein. The predicted amino acid sequence of the methylase contains significant regions of homology with the projected sequences of other cytosine-modifying methylases, upon which the activity of these enzymes is likely to depend.

INTRODUCTION

Relatively few restriction and modification genes have been sequenced, thereby limiting comparative studies of primary DNA sequence and of corresponding postulated protein sequences. Comparisons have been made between the *Escherichia coli* *dam* methylase sequence and those of *M.Dpn*II and the T4 *dam* methylases, all of which methylate the adenine residue in the sequence GATC (1,2). These studies suggest that the genes encoding these three enzymes have a common evolutionary origin and demonstrate the presence of three highly conserved regions that are likely to be involved in substrate recognition and enzyme activity.

Sequencing of the *M.Bsp*RI and *M.Bsu*RI genes of *Bacillus sphaericus* and *B.subtilis* and of the SPR methylase from the *B.subtilis* phage SPR has enabled a comparison of the deduced protein structures of these GGCC-modifying methylases (3,4). Substantial homology between *M.Bsp*RI and *M.Bsu*RI was revealed. However, homology was confined to two blocks of amino acids (aa) when SPR methylase, which is of dual specificity (GGCC and CCGG), was compared with the other two methylases. Further sequence comparisons have been made between SPR and two other *B.subtilis* methylases, *Phi*3T and *SPβ*

which are also of dual target specificity (GGCC and GCNGC)(5). Strong similarity between the two types of enzyme was apparent, apart from non-homologous inserts of 33 aa tracts at different positions in the proteins on which the unique specificity (GCNGC in *Phi3T* and *SP β* , and CCGG in *SPR*) of each methylase are postulated to depend. A further cytosine methylase *M.EcoRII* of specificity CC(A/T)GG, has also been sequenced (6). The postulated amino acid sequences of all the aforementioned cytosine methylases share several common domains, indicating involvement of these regions in cytosine methylation or in binding of S-adenosyl methionine (6).

At least seven sequence-specific type II methyltransferase activities have been characterized in strains of *Neisseria gonorrhoeae*, but only four cognate restriction endonucleases have been detected (7,8). This suggests that certain gonococcal methylases may not be part of a classical restriction-modification system. A number of bacterial methylases have been reported for which no corresponding endonuclease exists, for example the *dam* system of *E.coli* (9). However, the presence within a single strain of several methylases, as seen in *N.gonorrhoeae*, is uncommon. This may be attributed to problems created by the presence of excessively methylated DNA within the cell (10).

This paper describes the cloning in *E.coli* of the gene encoding the *NgoPII* methylase which recognises the nucleotide sequence GGCC. The deduced primary amino acid sequence of the *NgoPII* methylase is compared with other type II methyltransferases and functional domains postulated.

MATERIALS AND METHODS

Bacterial strains, bacteriophage, plasmids and media

N.gonorrhoeae P9-2 is a pilus antigenic variant of strain P9 (11). This strain was grown at 37°C for 18 h in the presence of 7% (v/v) CO₂ on CTA medium (1.125% [w/v] Proteose Peptone No.3 [Difco]; 1% [w/v] nutrient agar [Lab M No.2]; 0.1% [w/v] KH₂PO₄; 0.4% [w/v] K₂HPO₄; 0.5% [w/v] NaCl; 0.1% [w/v] soluble starch; and 1% [v/v] of a solution containing 40% [w/v] glucose; 0.5% [w/v] L-glutamine; 0.05% [w/v] Fe(NO₃)₃; 0.002% [w/v] thiamine pyrophosphate).

E.coli K-12 strain RR1 (*hsdS20*, *proA2*, *lacY1*, *rpsL20*, *ara-14*, *galK2*, *xyl-5*, *mtl-1*, *supE44*, λ^-) (12) was used for propagation of plasmid DNA. This was grown at 37°C in 2.5% (w/v) nutrient broth (Lab M No.2, London Analytical & Bacteriological Media Ltd.), or on nutrient agar (nutrient broth with agar [Lab M No.2] to 1.2% [w/v]). Where appropriate, liquid and solid media

contained ampicillin ($100 \mu\text{g} \cdot \text{ml}^{-1}$), or tetracycline hydrochloride ($12.5 \mu\text{g} \cdot \text{ml}^{-1}$). *E.coli* JM103 (lac, pro, thi, strA, supE, endA, sbcB15, hsdR4, F' traD36, proAB, lacIqZΔM15) was the host utilized for the propagation of all recombinant M13 bacteriophages (13). Phage stocks of λ vir were prepared by the plate lysis method (14).

All clones were constructed with plasmids pBR322 (12), or pACYC184 (15). pSU19 (4) was kindly supplied by R.J.Roberts.

Enzymes and chemicals

All reagents were supplied by BDH unless stated otherwise. Ampicillin, chloramphenicol, tetracycline hydrochloride and bovine serum albumin (BSA) were from Sigma. Restriction endonucleases were obtained from Boehringer Mannheim, P&S Biochemicals and New England Biolabs. T4 DNA ligase, and calf intestinal alkaline phosphatase were from Boehringer Mannheim, and ribonuclease A was supplied by Sigma. Reaction conditions for each enzyme were essentially as described by Maniatis *et al.* (14).

Detection of restriction endonuclease activity

The organism to be screened for restriction endonuclease production was grown overnight on an appropriate medium: *N.gonorrhoeae* P9 was grown on 5 CTA plates, *E.coli* in 500 ml nutrient broth. After harvesting, *Neisseria* cell pastes were resuspended in 5 ml GC TES (100 mM NaCl, 20 mM Tris-HCl, pH 7.5), and lysed by adding 20 μg lysozyme, Triton X100 to 1.6% (v/v) and phenyl methyl sulphonyl fluoride (PMSF) to 0.1 mM, then incubating at 37°C with shaking. 1 g *E.coli* cells were resuspended in 3 ml sonication buffer (20 mM Tris-HCl, pH 7.4, 1mM EDTA, 0.1 mM PMSF, and 7 mM β -mercaptoethanol and sonicated in an ice/water bath for 10 bursts of 30 s, with 60 s cooling intervals. Cell debris was removed by centrifugation at $25,000 \times g$ for 10 min and the supernatant filtered through a 0.2 μm membrane. A 500 μl sample was applied to a 1 ml MonoQ anion-exchange column (Pharmacia) pre-equilibrated in running buffer (20 mM Tris-HCl, pH 7.4, 1 mM EDTA, 1 mM dithiothreitol (DTT), 10% [w/v] glycerol), and eluted with a linear 0-1 M KCl gradient in running buffer. 0.5 ml fractions were collected using a Pharmacia FPLC system (16) and to each was added 0.5 ml glycerol and 500mg bovine serum albumin (BSA), and the samples stored at -20°C. 2 μl of each fraction was assayed with 0.1 μg γ DNA in a total volume of 20 μl x 1 medium salt restriction buffer and then analysed by agarose gel electrophoresis.

DNA isolation and cloning procedures.

Large-scale plasmid isolations were from 1 litre cultures of *E.coli* by a combination of the cleared lysate method and CsCl isopycnic centrifugation

- (17). Small-scale plasmid isolation was by the method of Close and Rodriguez
(18). N.gonorrhoeae P9 total DNA was isolated by the method of Stern et al
(19) from bacterial lawns totalling 500 cm² in area.

Transformation of E.coli was by the CaCl₂-treatment method (20).
Restriction mapping, agarose gel electrophoresis and DNA sub-cloning
procedures were essentially as described by Maniatis et al. (14).

DNA sequencing

The dideoxy chain termination method was used for all sequence
determinations (21), as described previously (22). Clones were constructed in
M13mp18 and mp19 (23). [α -³⁵S]dATP and 17-mer universal primer, used in the
chain elongation reactions, were both supplied by Amersham International. DNA
sequences were compiled manually and analysed using the IBM-PC compatible
programs of Schwindinger & Warner (24) and Mount & Conrad (25). Protein
sequence comparisons were performed using the algorithm of M.Dayhoff.

Coupled transcription-translation analysis of recombinants

Analysis of polypeptides encoded from recombinant plasmid DNA by in
vitro coupled transcription-translation was by the method of Pratt (26) using
³⁵S-labelled methionine supplied by Amersham International.
SDS-polyacrylamide gel electrophoresis was performed on 10-17.5% gradient
Laemmli gels (27) and labelled proteins were visualized following
autoradiography (28).

RESULTS

Determination of the restriction endonuclease and methylase content of N.gonorrhoeae P9-2

Two endonuclease activities were partially purified by MonoQ FPLC from
other nucleases released during cell lysis. Multiple substrate digests and
comparisons with digestion patterns of corresponding isoschizomers indicated
that the enzymes recognised and cleaved the sequences 5'-GGCC-3' (Fig. 1) and
5'-CCGCGG-3' (Fig. 1) and were designated R.NgoPII and R.NgoPIII
respectively.

P9 chromosomal DNA is not cleaved by DpnI demonstrating the absence of
a dam-type methylase (Fig. 2). However this strain is refractory to; HaeII,
implying the presence of a PuG^mCGCPy-recognising methylase (M.NgoPI); HaeIII,
GG^mCC (M.NgoPII); SacII, CCGCGG (M.NgoPIII); NaeI GCCGGC (M.NgoPIV); BamHI,
NarI and NlaIV suggesting a GGNN^mCC methylation site (M.NgoPV)(29).

Cloning of the NgoPII methylase gene

0.4 μ g N.gonorrhoeae P9-2 chromosomal DNA partially digested with EcoRI

Figure 1. Comparison of digestion patterns generated by NgoPII and HaeIII on various DNA substrates. Lane 1, λ cut with HaeIII; 2, λ with NgoPII; 4, Ad2 with HaeIII; 5, Ad2 with NgoPII; 7, ϕ X174 with HaeIII; 8, ϕ X174 with NgoPII; 10, SV40 with HaeIII; 11, SV40 with NgoPII.

and ClaI, and ligated with 0.2 μ g of the large EcoRI-ClaI fragment of pBR322 (a gift from I.J.Nicolson), was transformed into E.coli RR1. The resultant c. 5,000 transformants (75% recombinants as determined by replica-plating on Ap and Tc-containing agar plates) were pooled and grown in 500 ml broth. Plasmid DNA constituting the gene bank was purified by bulk preparation on a CsCl gradient. A further gene library (a gift from I.J.Nicolson) was constructed by ligating 0.4 μ g P9-2 total DNA partially digested with EcoRI, with 0.2 μ g of suitably cleaved pBR322, and then transformed into E.coli. The resulting 48,000 transformants (85% recombinants) were pooled and the gene bank DNA

Figure 2. Incubation of P9-2 chromosomal DNA with various restriction enzymes. Digestions were at 37°C for 3 h with excess restriction enzyme. Electrophoresis was in a 0.7% agarose gel as follows: Lane 1, uncut P9-2 DNA; 2, P9-2 DNA plus DpnI; 3, MboI; 4, Sau3AI; 5, KpnI; 6, BamHI; 7, HaeII; 8, HaeIII; 9, SacII; 10, PstI; 11, PvuI.

isolated as described above. The plasmid DNA constituting the gene banks was digested with HaeIII (an NgoPII isoschizomer which is unable to digest DNA modified by NgoPII-specific methylation), enabling selection of clones carrying functional M.NgoPII genes.

Plasmid DNA constituting the P9-2 gene bank constructed from EcoRI-Sau3A digested chromosomal DNA exhibited a 2,500-fold reduction in the number of transformants following incubation with HaeIII, when compared with uncut DNA. Analysis of DNA isolated from individual clones revealed one

plasmid, designated pLV130, that was refractory to HaeIII cleavage, indicating the presence of a functional M.NgoPII methylase gene within the c. 5 kb gonococcal DNA insert. Similarly a plasmid designated pLV140, which was also HaeIII-resistant and contained a c. 8 kb gonococcal chromosomal DNA insert was isolated by screening of the EcoRI-ClaI gene bank.

Characterization of pLV130 and pLV140

Titration of λ vir on RRL (pLV130), RRL(pLV140) and RRL(pBR322) yielded identical e.o.p.s, indicating the absence of any restriction activity associated with the clones. This was confirmed by in vitro screening for endonuclease activity utilizing FPLC purification of cell lysates. Restriction digests enabled both clones to be restriction-mapped and indicated that the insert DNA of pLV130 mapped within that of pLV140 (Fig. 3). However, pLV140 lacks an internal EcoRI site correlating with the insert boundary of pLV130, suggesting the possibility that more than one copy of the methylase could be present in the genome. This was investigated by DNA hybridization analysis of chromosomal DNA digested with ClaI, and EcoRI with ClaI, using the internal 0.68 kb PvuII-HindIII fragment as a hybridization probe. Both digests yielded single hybridizing bands indicating the presence of a single copy of the methylase in the genome (data not shown). Thus, pLV140 is likely to have arisen through the ligation to pBR322 of an EcoRI-ClaI fragment and a non-contiguous ClaI-ClaI fragment of gonococcal DNA.

The M.NgoPII methylase gene was localized by sub-cloning fragments of pLV130 into pBR322 (Fig. 3). This localized the gene to the 3' end of the insert, within the 1.3 Kb HpaI-Sau3A (BamHI when ligated into the BamHI site of pBR322) fragment. The methylase was found to be expressed when sub-cloned in either orientation in the tet gene of pACYC184, strongly suggesting that the gonococcal methylase is capable of being expressed in E.coli from its own native promoter.

DNA sequencing strategy

The DNA sequence of the 1685 bp PvuII-Sau3A chromosomal DNA segment corresponding to the cloned gonococcal insert of pLV136 was determined. 100% of both strands was sequenced using fragments overlapping all restriction sites used in the cloning strategy apart from the HindIII site at nt position 1003 (Sau3A cleavage site = between nt 1 and 2). The sequencing strategy is detailed in Fig. 4 and part of the sequence is listed in Fig. 5. The strategy employed required the generation and cloning of particular HaeIII fragments of the NgoPII methylase gene, the activity of which protects DNA from

Figure 3. Restriction maps of M.NgoPII sub-clones. All sub-cloned fragments are in pBR322. Only gonococcal DNA inserts are shown.

cleavage by HaeIII endonuclease. This was achieved by first cloning a non-functional fragment of the methylase gene (1 kb HindIII-BamHI fragment of pLV136) into M13, re-isolating this fragment from a digestion of a bulk phage DNA preparation and then cutting with HaeIII endonuclease.

Certain parts of the complete sequence appeared to be refractory to sub-cloning into M13. No fragments containing a functional methylase gene, i.e. spanning the 1.3 kb Sau3A-HpaI fragment were successfully cloned. E.coli strain JM103 contains the mcrB endonuclease which cleaves DNA with methylation pattern GG^mCC (30). Thus, it is evident that the M.NgoPII gene cannot be introduced into an mcrB⁺ host even when phage-borne. Furthermore, the blunt-ended AluI-AluI fragment spanning from nt position 158 to 511 could

Figure 4. Detailed restriction map and strategy for sequencing the *M.NgoPII* gene. Small arrows indicate origin and direction of restriction fragments cloned into M13. Bold arrow indicates the position of the methylase gene.

only be detectably cloned in one orientation in the *Sma*I site of M13mp19, as determined by T-tracking of DNA from 40 white plaques, all of which contained the insert in the same orientation. This problem was circumvented by isolating DNA from a bulk preparation of one of these M13mp19 clones, excising the cloned fragment by digestion with *Eco*RI and *Hind*III and then cloning into suitably cleaved M13mp18 DNA. Detection of clones containing the insert in one orientation only could be due to the presence of a promoter within the cloned fragment, capable of expressing, in the opposite orientation, the *lacZ* gene product as an in-frame fusion protein, which would result in a blue plaque. Further analysis subsequently demonstrated the presence of the putative methylase promoter within the *Alu*I-*Alu*I fragment under consideration.

Assignment of open reading frames to the modification gene.

The nucleotide sequence was analysed in all 6 possible reading frames for potential initiation and termination codons which resulted in open reading frames (orfs) of greater than 100 aa in length. The 1.329 kb *Sau*3A-*Hpa*I fragment contains the functional *NgoPII* methylase gene, as determined by sub-cloning experiments. Only one orf lies within this region and this was assigned to the methylation enzyme. The translational initiation codon (ATG) for this putative gene begins at nt 243 and terminates at 1266 (TGA), as shown in Fig. 5. The postulated promoter comprises a Pribnow box (TAAAT), at position -5 to -10 relative to the theoretical purine initiation nucleotide, but lacks an appropriate consensus -35 region. The -10 region shares homology with the consensus sequence (TATAAT) for *E.coli* promoters (31). In addition, a Shine-Dalgarno (SD) sequence (GAT) is situated 7 bp

```

      80
TGATCGCTCAACATAACTTATCTGAGACTACGGGGTATGTGGGGAATCGAAAAATCCATGCTTTGTCTTCAACTATATTT
      160
ACCAACGGAATATGGAAGTTTCAATTTTCATGGCCATTATCAACGAAGACAAATGGAACAGTTTAAACAATACAGAC
      240
AAGCTGCTTGCCATACAGGATTCCAAATCGGCCATTTCGGACATTA1333AAAAATCCCAACAACCGCGCGGATTAAAG
      320
AAATGCAAACTCATCACTACCACTTATAATCCTATGAAAATCATTAGTTTGTAGCGGTTGCGCGGTTTGGATTGTG
      400
MetGlnAsnSerSerProThrThrTyrAsnProMetLysIleIleSerLeuPheSerGlyCysGlyGlyLeuAspLeu
GGTTTCGAAAAAGCGGGATTGAAATCCCGCGGCCAACAATACGATAAAACCAATTGGGCAACCTTCAAGGCAAAACCA
      480
GlyPheGluLysAlaGlyPheGluIleProAlaAlaAsnGluTyrAspLysThrIleTrpAlaThrPheLysAlaAsnHis
TCCAAAGACCCATTGTATAGAAGCGGATATACGCAAGATTAAAGAAGAAGATTTCCCTGAAGAAATCGACGGGATTATCG
      560
ProLysThrHisLeuIleGluGlyAspIleArgLysIleLysGluGluAspPheProGluGluIleAspGlyIleIleGly
CGGCGCCACCTTCCAGCTCTTGGTCTGAAGCGGGAGCTTTGCGCGGCATCGACGATCGCGCGGACAGTTGTTTTTCGAC
      640
GlyProProCysGlnSerTrpSerGluAlaGlyAlaLeuArgGlyIleAspAspAlaArgGlyGlnLeuPhePheAsp
TACATCCGTATTTTGAAGCAACAGCCAAAAATCTTTTATAGCGGAAAACTGAGCGGAATGCTGGCAACCGCCACAA
      720
TyrIleArgIleLeuLysSerLysGlnProLysPhePheLeuAlaGluAsnValSerGlyMetLeuAlaAsnArgHisAsn
CGGACCGGTACAAAACCTGCTGAAAATGTTTACCGGATCGCGATACGAGTAACCTTGACTATGCCAACGCCAAAGACT
      800
GlyAlaValGlnAsnLeuLeuLysMetPheAspGlyCysGlyTyrAspValThrLeuThrMetAlaAsnAlaLysAspTyr
ACGGTGTAGCAGGAAACGCAAAAGGGTCTTCTACATCCGGTTCCGTAAGACTTGGAAATAAAAATTTCTTTTCCAAAA
      880
GlyValAlaGlnGluArgLysArgValPheTyrIleGlyPheArgLysAspLeuGluIleLysPheSerPheProLys
GGTTGCGCGGTGGAAGCAAGAGATTACATTGAAAGACGTTATTTGGGATTGCGAGGACACAGCGGTACCTTCGCG
      960
GlySerThrValGluAspLysAspLysIleThrLeuLysAspValIleTrpAspLeuGlnAspThrAlaValProSerAla
CCCGCAAAACAAGCAACACCCCGACGCGAGTCAACAACAACGAATATTTACCGGCAGTTTTCCTCTATTTTATGAGCC
      1040
ProGlnAsnLysThrAsnProAspAlaValAsnAsnAsnGluTyrPheThrGlySerPheSerProIlePheMetSerArg
GGAACCGAGTTAAAGCGTGGATGAACAGGGTTTACCGTCCAAGCTTCAGGCAGGAGTGCCAACTGCACCCGCAAGCC
      1120
AsnArgValLysAlaTrpAspGluGlnGlyPheThrValGlnAlaSerGlyArgGlnCysGlnLeuHisProGlnAla
CCAAAGATGGAAGACAGCGGGCAACGACTACCGTTTTCCTGCGCGCAAGAAACACTATATCGGAGGATGACGGTACG
      1200
ProLysMetGluLysHisGlyAlaAsnAspTyrArgPheAlaAlaGlyLysGluThrLeuTyrArgArgMetThrValArg
CGAAGTTGCAAGATCCAAAGCTTTCCGACAACTTCAAATTCATCTATCAAAATGTCAACGACGCATACAAAAATGATTG
      1280
GluValAlaArgIleGlnGlyPheProAspAsnPheLysPheIleTyrGlnAsnValAsnAspAlaTyrLysMetIleGly
GCAACGGCGTCCCGGTCAACCTTGCTACGAAATTCAGCGGCAATTA1333AAAAACCTAGAAAGGTGAACACCCCTTCCT
      1333
AsnAlaValProValAsnLeuAlaTyrGluIleAlaAlaAlaIleLysLysThrLeuGluArgEND
ACAGGCATCTGCCTTGTCTGTCATGCCGGAATAACGGTACTATTCCGAGTTAAC

```

Figure 5. Nucleotide sequence of the 1.33 kb *Sau*3A-*Hpa*I fragment containing the *M.Ngo*PII gene. The putative ribosome binding site is underlined and the putative -10 sequence is boxed.

upstream of the translational initiation codon. Three other types of gene from *N.gonorrhoeae* have been sequenced: pilin (32,33), opacity (protein PII) genes (34) and IgA protease (35). The pilin genes possess typical SD sequences and -10 regions but lack a recognisable -35 promoter sequence, whilst the opacity genes and IgA genes have all three motifs.

The base composition of the presumptive *Ngo*PII methylase gene is 52.9% A+T respectively. This value compares closely with an average gonococcal DNA A+T content of 49.5% (36). The putative *Ngo*PII methylase enzyme comprises 341 aa corresponding to a computed molecular weight of 38.49×10^6 .

Protein sequence comparisons

The aa sequence of the NgoPII modification enzyme was compared with published sequences of other GGCC-recognising methylases, all of which are from B.subtilis, B.sphaericus, or B.subtilis bacteriophages, and with the EcoRII methylase which recognises CC(A/T)GG (Fig. 6). Significant homology was detected between the NgoPII methylase and both the BsuRI and BspRI methylases, which, with the exception of the amino and carboxy termini, is localized in six blocks which extend over much of the protein (Fig. 7). Likewise, comparisons between SPR and both the NgoPII and BsuRI methylases demonstrated similar levels of homology in both comparisons, with sequence similarities localised in blocks. These regions of sequence are highly conserved in all GGCC-modifying methylases (Fig. 8), and correspond to blocks B, C, E and F previously defined in Fig. 7. Indeed, sub-sections of these highly conserved blocks may be common to all cytosine-modifying methylases, as demonstrated by comparison with the EcoRII methylase protein sequence (Fig. 8). Likewise, no significant homology was detectable between the M.NgoPII protein and the adenine-modifying methylases dam (37), HhaII (38), EcoRI (39) and EcoRV (40).

Computer alignment of proteins using Dayhoff's similarity scores between amino acids was also used to compare the methylase enzymes. The homology between M.NgoPII and the BsuRI, BspRI, SPR and EcoRII methylases was calculated to be 33.9%, 32.0%, 20.7% and 17.4% (matches/length) respectively. These compare with 64.0% for M.BsuRI/M.BspRI and 15.0% for M.BsuRI/M.SPR. By excluding the non-homologous N and C terminus regions, NgoPII-BsuRI methylase comparisons yielded >40% homology for the remainder of the sequence. This similarity is not reflected at the DNA sequence level, since DNA hybridizations between the M.NgoPII gene and the BsuRI R.M. system failed to demonstrate homology even under low stringency conditions (data not shown). This was confirmed by comparison of the methylase DNA sequences.

Coupled in vitro transcription-translation studies

A number of NgoPII methylase sub-clones were used to program an in vitro coupled transcription-translation (CTT) system. The α -³⁵S methionine-labelled products were visualised by autoradiography following SDS-PAGE (Fig. 9). pLV137, the smallest sub-clone to express a functional NgoPII methylase, produced a protein of calculated Mwt. 39.8×10^6 , compared with a theoretical Mwt. of 38.5×10^6 for the methylase, as predicted from the DNA sequence. A 44 kd protein band was also produced by pLV137. This is could be generated by read-through of the NgoPII methylase gene termination

Figure 6. Diagon comparisons between *M.NgoPII*, *M.BsuRI*, *M.BspRI*, *M.SPR* and *M.EcoRII*.

codon by the *E.coli* RNA polymerase, resulting in an extended protein. Both of these protein bands are also generated by pLV136.

DISCUSSION

The primary DNA sequence of a 1.329 kb gonococcal chromosomal DNA fragment taken in conjunction with sub-cloning data enabled an orf to be attributed to the *NgoPII* methylase. The corresponding theoretical size of the protein (38.5 kd) was in good agreement with the observed polypeptide generated by *in vitro* transcription-translation experiments (39.8 kd) and *in vivo* when extracted from gonococcal lysates. The size of the methylase is typical of type II methyltransferases which are generally within the range 30-50 kd (41). The size of the active *NgoII* methylase has been reported atypically to be 25 kd following relatively crude purification of the enzyme (42). If the *NgoII* and *NgoPII* enzymes are the same, which seems to be highly likely, these results could be reconciled if the methylase(s) were to undergo post-translational processing. Alternatively and in our view more likely, the original estimate for the size of *NgoII* methylase could be erroneous.

The *M.NgoPII* gene showed no gross DNA homology with the *M.BsuRI* gene, despite sharing the same DNA sequence specificity. This implies extensive

divergent evolution of these two genes if they share a common ancestor. These results corroborate hybridization comparisons with other restriction and modification genes, such as those between the EcoRI, dam and T4 dam methylases which recognise similar sequences but display no detectable homology between their DNA and protein sequences (2). Previous comparisons of the BspRI, BsuRI and SPR methylase sequences have detected partial homologies, which were interpreted as an indication of evolutionary relatedness (3,4). Similarly SPR, 3T and SP β methylases share extensive homology (5). Comparisons of the NgoPII methylase protein sequence with the aforementioned enzymes also demonstrated homologies in the same regions of

1	M.NgoPII	Gly Gly Pro Pro Cys Gln Ser Trp Ser Glu Ala Gly	91
	M.BsuRI	Gly Gly Phe Pro Cys Pro Gly Phe Ser Glu Ala Gly	164
	M.BspRI	Gly Gly Phe Pro Cys Pro Gly Phe Ser Glu Ala Gly	163
	SPR	Gly Gly Ser Pro Cys Gln Ser Phe Ser Val Ala Gly	85
	Phi3T	Ser Gly Phe Pro Cys Pro Thr Phe Ser Val Ala Gly	85
	M.EcoRII	Ala Gly Phe Pro Cys Gln Pro Phe Ser Leu Ala Gly	193
2	M.NgoPII	Gln Pro Lys Phe Phe Leu Ala Glu Asn Val Ser Gly Met	127
	M.BsuRI	Gln Pro Glu Ile Phe Val Ala Glu Asn Val Lys Gly Met	199
	M.BspRI	Gln Pro Glu Ile Phe Val Ala Glu Asn Val Lys Gly Met	198
	SPR	Gln Pro Lys Phe Phe Val Phe Glu Asn Val Lys Gly Leu	121
	Phi3T	Lys Pro Lys Phe Val Ile Leu Glu Asn Val Lys Gly Leu	154
	M.EcoRII	Lys Pro Ala Ile Phe Val Leu Glu Asn Val Lys Gly Asn	153
3	M.NgoPII	Met Ala Asn Ala Lys Asp Tyr Gly Val Ala Gln Glu Arg Lys Arg Val Phe Tyr Ile Gly Phe Arg Lys Asp Leu	178
	M.BsuRI	Leu Leu Asn Ala Arg Asp Tyr Gly Val Pro Gln Leu Arg Glu Arg Val Ile Ile Glu Gly Val Arg Lys Asp Ile	250
	M.BspRI	Leu Leu Asn Ala Arg Asp Tyr Gly Val Pro Gln Ile Arg Glu Arg Val Ile Ile Val Gly Val Arg Asn Asp Ile	249
	SPR	Leu Leu Asn Ser Lys Phe Phe Asn Val Pro Gln Asn Arg Glu Arg Leu Tyr Ile Ile Gly Ile Arg Glu Asp Leu	172
	Phi3T	Leu Leu Asn Ser Lys Phe Phe Asn Val Pro Gln Asn Arg Glu Arg Val Tyr Ile Ile Gly Ile Arg Glu Asp Leu	205
	M.EcoRII	Lys Val Ile Asp Gly Lys His Phe Leu Pro Gln His Arg Glu Arg Ile Val Leu Val Gly Phe Arg Arg Asp Leu	297
4	M.NgoPII	Arg Arg Met Thr Val Arg Glu Val Ala Arg Ile Gln Gly Phe Pro Asp	303
	M.BsuRI	Arg Arg Leu Ser Val Lys Glu Ile Ala Arg Val Gln Thr Phe Pro Asp	361
	M.BspRI	Arg Arg Leu Ser Val Lys Glu Ile Lys Arg Ile Gln Thr Phe Pro Asp	360
	SPR	Arg Arg Leu Thr Pro Leu Glu Cys Phe Arg Leu Gln Ala Phe Asp Asp	393
	Phi3T	Arg Lys Leu Ser Pro Leu Glu Cys Trp Arg Leu Gln Ala Phe Asp Asp	397
	M.EcoRII	Arg Arg Leu Thr Pro Arg Glu Cys Ala Arg Leu Met Gly Phe Glu Lys	424
5	M.NgoPII	Tyr Lys Met Ile Gly Asn Ala Val Pro Val	325
	M.BsuRI	Tyr Lys Gln Ile Gly Asn Ala Val Pro Val	392
	M.BspRI	Tyr Lys Gln Ile Gly Asn Ala Val Pro Val	391
	SPR	Tyr Lys Gln Ala Gly Asn Ser Ile Thr Val	419
	Phi3T	Tyr Lys Gln Ala Gly Asn Ser Ile Thr Val	423
	M.EcoRII	Tyr Arg Gln Phe Gly Asn Ser Val Val Val	449

Figure 8. Amino acid sequences common to all characterized GGCC-recognising methylases. The region common to all such methylases is boxed. Blocks 1 and 2 lie within region B, block 3 corresponds to region C, block 4 with region E and block 5 with region F in Figure 7. Only part of block 3 is present in all GGCC-modifying enzymes and may be responsible for recognizing the sequence 5'-GGCC-3'.

Figure 9. Coupled *in vitro* transcription-translation analysis of *M.NgoPII* subclones. α - ^{35}S methionine labelled CTT products were visualised by autoradiography following SDS-PAGE. The DNA substrates used to drive the reactions were track 1, pLV136; 2, pLV137; 4, pLV130 and 5, pBR322. Track 3 contains ^{14}C radiolabelled protein standards (Amersham International).

sequence as those conserved in methylases encoded by the *Bacillus* chromosome and *B.subtilis* phages. Thus, it could be concluded that the *NgoPII*, *BsuRI*, *BspRI* and *B.subtilis* phage methylases share a common evolutionary origin, and that the regions of homology common to all the enzymes (blocks 1 to 5, Fig. 8) are associated with the catalytic function of these GGCC-recognising enzymes. Furthermore, comparison with the *EcoRII* methylase, which recognises a dissimilar sequence, has highlighted homologous domains which may have a role in methylating cytosine, or in binding of S-adenosyl methionine (6). This in turn implies that the homologous region exclusive to GGCC-methylases (localized to block 3 in Fig. 8) may be required for their site-specific interaction with DNA.

A common evolutionary origin for the *M.NgoPII* gene of *N.gonorrhoeae* and corresponding genes of *B.subtilis* and *B.sphaericus* would have to predate divergence of Gram positive and Gram negative bacteria. A similar proposal

has been made for the relationship between the E.coli dam methylase and DpnII methylase of S.pneumoniae (1). These proteins share 30% homology, which was presumed to correlate with the expected amino acid sequence similarity for proteins evolving from such an ancient origin. The NgoPII methylase shares c. 34% and 32% homology with BsuRI and BspRI methylases respectively. Part of this overall homology is attributable to regions absent in the B.subtilis phage methylases (blocks A and D, Fig. 7) and are thus not essential for the catalytic activities of the bacterial enzymes. These non-essential homologies are unlikely to have arisen through convergent evolution from separate origins. Furthermore, greater overall homology between M.BsuRI and M.NgoPII compared with M.BsuRI and SPR implies that divergence of the B.subtilis and B.subtilis phage methylases predates even the divergence of Gram positive and Gram negative bacteria 10^9 years ago (43).

N.gonorrhoeae lacks detectable endogenous bacteriophages, yet many strains contain restriction-modification systems and the majority produce at least four methyltransferase activities (7,44). Postulated functions of these methylases are in the selection of gonococcal DNA from the environment during transformation of N.gonorrhoeae and in the protection of DNA in transit between cells during transformation (7). Evidence is accumulating for roles of methylation in the control of gene expression and in DNA recombination (44,45,46). Pilus antigenic variation in N.gonorrhoeae involves gene conversion-like events (33,47). M.NgoIV, which has a target site within the structural gene for pilin (the protein subunit of which pili are composed), could conceivably play a role in modulating the process of antigenic variation (47). Thus, the cloning of M.NgoPII and other gonococcal methylases should facilitate studies on the role of these activities in N.gonorrhoeae.

ACKNOWLEDGEMENTS

I.J. Nicolson and A.C.F. Perry are thanked for valuable discussions and S.Green for assistance with part of the computing. This work was supported by a Scholarship from Powell and Scholefield Ltd. to K. Sullivan and by grants from the Medical Research Council, Wellcome Trust and Wolfson Trust.

¹Present address: Home Office Central Research Establishment, Aldermaston, Reading, Berks RG7 4PN, UK

REFERENCES

1. Mannarelli, B.M., Balganes, T.S., Greenberg, W., Springhorn, S.S and Lacks, S.A. (1985) Proc. Natl. Acad. Sci. USA 82, 4468-4472.

2. Hattman, S., Wilkinson, J., Swinton, D., Schlagman, S., Macdonald, P.M. and Mosig, G. (1985) *J. Bacteriol.* 164, 932-937.
3. Posfai, G., Baldauf, F., Erdei, S., Posfai, J., Venetianer, P. and Kiss, A. (1984) *Nucl. Acids Res.* 12, 9039-9049.
4. Kiss, A., Posfai, G., Keller C.C., Venetianer, P. and Roberts, R.J. (1985) *Nucl. Acids Res.* 13, 6403-6421.
5. Tran-Beckte, A., Behrens, B., Noyer-Weidner, M. and Trautner, T.A. (1986) *Gene* 42, 89-96.
6. Som, S., Bhagwhat, A.S. and Friedman, S. (1987) *Nucl. Acids Res.* 15, 313-332.
7. Korch, C, Hagblom, P. and Normark, S. (1983) *J. Bacteriol.* 155, 1324-1332.
8. Stein, D., Gregoire, S. and Piekarowicz, A. (1988) *Infect. Immun.* 56, 112-116.
9. Marinus, M.G. and Morris, N.R. (1973) *J. Bacteriol.* 114, 1143-1150.
10. Adams, R.L.P. and Burdon, R.H. (1985) *Molecular Biology of DNA methylation.* Springer-Verlag, New York.
11. Nicolson, I.J., Perry, A.C.F., Heckels, J.E. and Saunders, J.R. (1987) *J. Gen. Micro.* 133, 553-561.
12. Bolivar, R., Rodriguez, R.L. Greene, P.J., Betlach, M.C., Heynecker, H.L. and Boyer, H.W. (1977) *Gene* 2, 95-113.
13. Messing, J., Crea, R. and Seeburg, P.H. (1981) *Nucl. Acids Res.* 9, 309-321.
14. Maniatis, T., Frisch, E.F. and Sambrook, J. (1982) *Molecular Cloning. A Laboratory Manual.* Cold Spring Harbor, NY.
15. Chang, A.C.Y. and Cohen, S.N. (1978) *J. Bacteriol.* 134, 1141-1156.
16. Walker, J.M.B., Dean, P.D.G and Saunders, J.R. (1986) *Nucl. Acids Res.* 14, 1293-1301.
17. Humphreys, G.O., Willshaw, G.A. and Anderson, E.S. (1975) *Biochim. Biophys. Acta* 383, 457-463.
18. Close, J.C. and Rodriguez, R.L. (1982) *Gene* 20, 305-316.
19. Stern, A., Nickel, P., Meyer, T.F. and So, M. (1984) *Cell* 37, 447-456.
20. Brown, M.G.M., Weston, A., Saunders, J.R. and Humphreys, G.O. (1979). *FEMS Microbiol. Lett.* 5, 219-222.
21. Sanger, F., Nicklen, S. and Coulson, A.R. (1977) *Proc. Natl. Acad. Sci. USA* 74, 5463-5467.
22. Nicolson, I.J., Perry, A.C.F., Virji, M., Heckels, J.E. and Saunders, J.R. (1987) *J. Gen. Micro.* 133, 825-833.
23. Norrander, J., Kempe, T. and Messing, J. (1983) *Gene* 26, 101-106.
24. Schwindinger, W.F. and Warner, J.R. (1984) *Nucl. Acids Res.* 12, 601-604.
25. Mount, D.W. and Conrad, B. (1984) *Nucl. Acids Res.* 12, 811-818.
26. Pratt, J.M. (1984) In Hames, B.D. and Higgins, S.J. (Eds.) *Transcription and Translation: a practical approach.* IRL Press, Oxford, pp. 179-209.
27. Laemmli, U.K. (1970) *Nature* 227, 680-685.
28. Hames, B.D. (1982) In Hames, B.D. and Rickwood, D. (Eds.), *Gel Electrophoresis of Proteins.* IRL Press, Oxford, pp. 1-91.
29. Sullivan, K.M., Macdonald, H.J. and Saunders, J.R. (1987) *FEMS Microbiol. Lett.* 44, 389-393.
30. Raleigh, E.A. and Wilson, G. (1986) *Proc. Natl. Acad. Sci. USA* 83, 9070-9074.
31. Rosenberg, M. and Court, D. (1979) *Ann. Rev. Genet.* 13, 319-353.
32. Perry, A.C.F., Nicolson, I.J. and Saunders, J.R. (1988) *Gene* 60, 85-92.
33. Meyer, T.F., Billyard, E., Haas, R., Storzbach, S. and So, M. *Proc. Natl. Acad. Sci. USA* 81 (1984) 6110- 6114.
34. Stern, A., Brown, M., Nickel, P. and T.F. Meyer. (1986) *Cell* 47, 61-71.
35. Pohlner, J., Halter, R., Beyreuther, K. and Meyer, T.F. (1987). *Nature* 325, 458-462.

36. Belozersky, A.N. and Spirin, A.B. (1960). In *The Nucleic Acids*, Vol. 3. Chargaff, E. and Davidson, J.N. (Eds.) Academic Press, New York pp. 147-185.
37. Brooks, J.E., Blumenthal, R.B. and Gingeras T.R. (1983) *Nucl. Acids Res.* 11, 837-851.
38. Schoner, B., Kelly, S. and Smith, H.O. (1983) *Gene* 24, 227-236.
39. Greene, P.J., Gupta, M. and Boyer, H.W. (1981) *J. Biol. Chem.* 256, 2143-2153.
40. Bougueleret, L., Schwartzstein, M., Tsugita, A. and Zabeau, M. (1984) *Nucl. Acids Res.* 12, 3659-3676.
41. Smith, H.O. and Kelly, S.V. (1984). In *DNA methylation: biochemistry and biological significance*. Razin, A., Cedar, H and Riggs, A.D. (Eds.), Springer-Verlag, New York, pp 1-37.
42. Clanton, D.J., Riggsby, W. and Miller, R.V. (1979). *J. Bacteriol.* 137, 1299-1307.
43. Schwartz, R.M. and Dayhoff, M.O. (1978). *Science* 199, 395-403.
44. Radman, M. and Wagner, R. (1984). *Curr. Top. Microbiol. Immunol.* 108, 23-23.
45. Roberts, D., Hoopes, B.C., McClure, W.R. and Kleckner, N. (1985). *Cell* 43, 117-130.
46. Sternberg, N., Sauer, B., Hoess, R. and Abremski, K. (1986). *J. Mol. Biol.* 187, 197-212.
47. Segal, E., Billyard, E., So, M., Storzbach, S. and Meyer, T.F. (1985). *Cell* 40, 293-300.