Volume 16 Number 14 1988 Nucleic Acids Research

Chimeric multispecific DNA methyltransferases with novel combinations of target recognition

Thomas A.Trautner, Tanjore S.Balganesh* and Brigitte Pawlek

Max-Planck-Institut für Molekulare Genetik, D-1000 Berlin 33, FRG

Received May 5, 1988

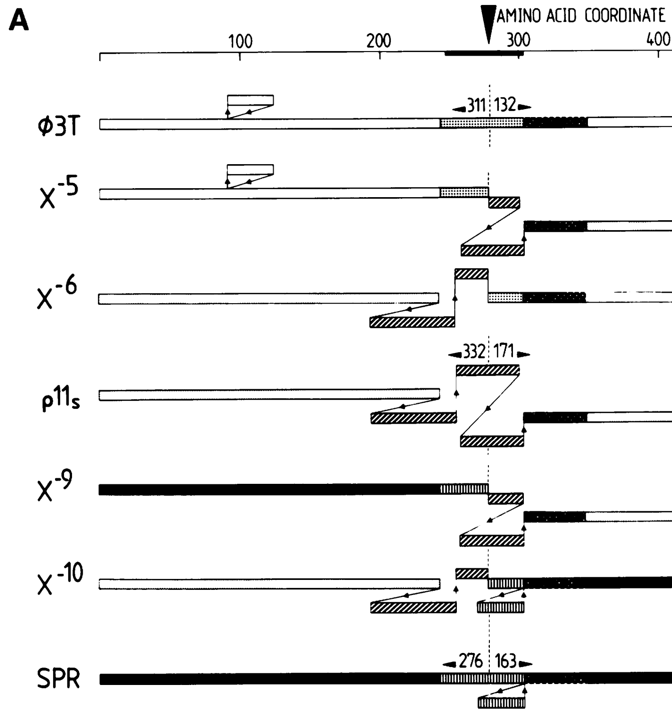
ABSTRACT

DNA target recognizing domains of different multispecific DNA-cytosine-methyltransferases can be rearranged through engineering of the corresponding genes to generate enzymes with novel combinations of target recognition.

INTRODUCTION

Prokaryotic DNA-cytosine5 methylating methyl-transferases (Mtases), whose amino acid sequences have been established, are closely related in their primary structures (1,2,3,4,5,6). They have two very similar regions of roughly 150 and 80 amino acids towards their NH_2- and COOH-terminal ends, respectively. These conserved regions are separated by a contiguous segment, in which the amino acids vary between different enzymes both in number and sequence. Our preceding studies on the multispecific Mtases encoded by the temperate *Bacillus subtilis* phages ϕ3T, SPR, and $\rho 11_s$ (Fig. 1) (1,2,5,7) showed that in these Mtases (and by analogy most likely also in the bacterial Mtases) the variable regions contain domains, which are responsible for the enzymes' DNA target recognition. The conserved NH_2- and COOH-terminal "core" sequences, on the other hand, carry domains required in general steps of the methylation reaction. In the multispecific phage Mtases, we could assign the enzymes' capacity to recognize individual targets to domains of about 50 amino acids each. These domains are sequentially arranged and nonoverlapping (8). A domain responsible for the recognition of the sequence GGCC (*Hae*III) is identical in the three phage Mtases and occupies the same relative position within the enzymes. Domains determining

© IRL Press Limited, Oxford, England.



B

SEQUENCE MODIFIED				
GGCC HaeIII	GCNGC Fnu4HI	G( )GC( )C Bsp1286	CCGG HpaII	CC A/T GG EcoRII
+	+	−		
+	−	−		
+	+	+		
+	−	+	−	−
+		−	−	−
+		+	+	−
+		−	+	+

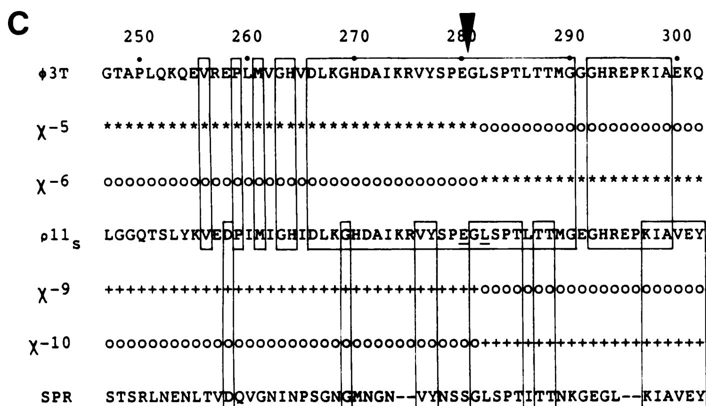


Fig. 1:
Alignment of parental and chimeric Mtases and their methylation potential. A: Schematic presentation of the alignment of Mtases based on the data presented in (5). Core regions of the three Mtases are distinguished by different shadowing. Unique target recognizing domains of the three enzymes are marked by (ϕ3T), (ρ11$_s$), and (SPR). The common GGCC target recognition domain is marked by . The location of the StuI site at the DNA equivalent of amino acid coordinate 281 is indicated by an arrow and a vertical interrupted line. In the parental molecules the numbers to the left and right of this location give the numbers of amino acids between this location and the NH_2- or COOH-termini. B: Methylation potential of parental and chimeric Mtases as determined in the experiment of Fig. 3. C: Amino acid sequences of parental and chimeric Mtases surrounding the joining point at amino acid coordinate 281 (arrow). The section shown is presented in the amino acid coordinate scale of Fig. 1A by a black bar. The ρ11$_s$ sequence is that of the StuI containing Mtase gene of plasmid pBB3 (Fig. 2). The numbers describe amino acid coordinates (5). In chimeras, *, o, and + represent amino acids of ϕ3T, ρ11$_s$, and SPR. Identical amino acids of ρ11$_s$ and ϕ3T, and of ρ11$_s$ and SPR are boxed.

recognition of the additional target sequences, which are distinct for each enzyme, are located NH_2-terminal to the common GGCC recognizing domain (Fig. 1A).

From this general building plan of the Mtases it appears that the variation in target recognition of different Mtases is essentially achieved by the association of a highly conserved core structure with a variety of single (bacterial enzymes), or multiple (phage enzymes) "modules" determining specificity. In such a situation one should be able to delete or add target

recognizing modules in the variable region, causing either a loss or the acquisition of novel methylating capacities. In support of this expectation, we had previously shown (8) that the domain in the SPR Mtase, determining CCGG methylation, was deletable without affecting the capacity of the mutant enzyme to methylate other target sequences. We had also reported (7) on the construction of chimeric Mtases with enzymatic activity. These experiments showed the interchangeability of the core structures of different phage enzymes.

Here we expand the latter studies. We report on the construction of chimeric Mtases, in which target recognizing domains from the Mtases of SPR and ϕ3T have become combined with those of $\rho 11_s$ to give enzymes with novel combinations of target recognition.

MATERIALS AND METHODS

Bacterial strains. *E. coli* strains used in the maintenance and construction of plasmids were methylation tolerant *rglB*$^-$ derivatives (9,10). The strains included HB101 (11), and GM1499 (kindly provided by M. Marinus), which is dcm$^-$ and which was used as a plasmid host to detect *Eco*RII (CC(A/T)GG) methylation.

Plasmids. All plasmids containing the entire or subfragments of the Mtase genes were derived from pBR328. Plasmid pKB131 (12) carrying the SPR Mtase gene is a derivative of pRB121 (13). pBN16 with the ϕ3T Mtase gene was described in (14). Plasmids with the $\rho 11_s$ Mtase gene were pSB11, derived from pBN52 (5) and pBB3 (Fig. 2), a new construct, which contained the Mtase gene, mutagenized to contain a *Stu*I site, within a *Pvu*II *Pst*I fragment of $\rho 11_s$ DNA. Plasmids pBB15, pBB16 and pBB9, pBB10 (Fig. 2), containing chimeric Mtase genes were constructed following conventional methods of engineering. All chimeric Mtase genes used, represent fusions of subgenic fragments at the unique *Stu*I site located at the DNA equivalent to amino acid 281 (5) (Fig. 1). Plasmids pSB11, pBB16 and pBB10 have an insert of λ DNA with two *Sac*I sites. In the absence of *Sac*I sites in the vector and the Mtase coding segment, the λ insert served to monitor *Sac*I methylation by the $\rho 11_s$ Mtase.

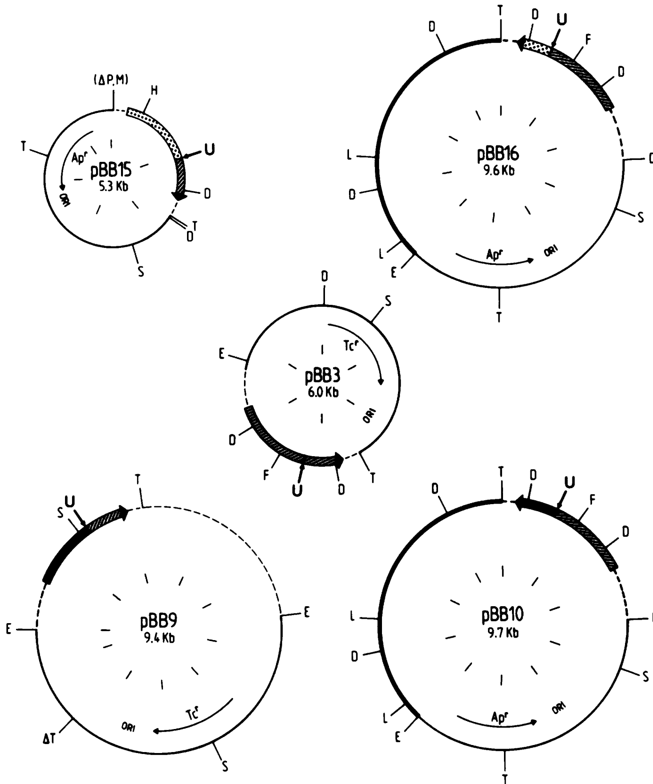


Fig. 2:
Plasmid constructs analyzed. The cloned Mtase genes (ϕ3T: [dotted], $\rho 11_s$: [hatched], SPR: [striped]) and their direction of transcription are given by double lined arrows. Vector DNA and insert DNA are shown as continuous or broken thin lines. λ DNA in plasmids pBB16 and pBB10 is represented by a heavy line. Some relevant restriction sites are: U: StuI, D: HindIII, E: EcoRI, F: AflII, H: HpaI, L: SacI, S: SalI, T: PstI. Δ describes the deletion of restriction site(s).

We only learned after the construction of these plasmids that the Mtase activity of $\rho 11_s$, which was determined by E. Hemphill (personal communication) to have SacI sites as targets, was actually directed against G(A,T,C)GC(T,A,G)C sequences (Bsp1286) of which the SacI sequence (GAGCTC) is a subset (P.A. Terschüren, unpublished). Bsp1286 sites are abundant in pBR328, such that the activity of the $\rho 11_s$ Mtase activity corresponding to Bsp1286 can be monitored also in plasmids like

pBB3 without the insert carrying *Sac*I sites. All plasmid constructs obtained could be stably maintained in *E. coli* strains provided they were permissive for methylated plasmid DNA (9,10).

Reagents and general techniques. Restriction endonucleases and other relevant enzymes were purchased from Boehringer (Mannheim, Germany). Standard DNA preparation, transformation and cloning techniques were used as described in (15).

Mutagenesis. To construct chimeric Mtases involving also the Mtase of phage $\rho 11_s$, we have introduced a unique StuI site (AGGCTT) at a position which is equivalent to the location of the StuI site previously introduced into the Mtase genes of ϕ3T and SPR (7). This is in the variable region and involves the DNA corresponding to amino acid coordinates 280-282 (Fig. 1A and C). Oligonucleotide directed mutagenesis followed the technique previously described (7). The combination of the 13mer oligonucleotide used (synthesized in an Applied Biosystems DNA synthesizer) and the relevant DNA sequence, contained in a 1189 bp *Sma*I/*Pst*I fragment of $\rho 11_s$ DNA cloned into M13 mp19 was the following (here and elsewhere DNA sequences are written 5′ → 3′ from left to right):

mutagenic primer		T	GAA	GGC	CTG	TCA	
$\rho 11_s$	...	CCT	GAT	GGG	GTG	TCA	...
amino acid	...	P	D(E)	G	V(L)	S	...
		279				283	

Base changes necessary to generate the *Stu*I site are indicated by asterisks in the primer sequence, its location by a bar in the $\rho 11_s$ sequence. Amino acids replacing those of the wild type gene after introduction of the *Stu*I site are shown in parentheses. Plasmid pBB3 (Fig. 2) was like plasmid pBN52 resistant to *Hae*III and *Bsp*1286 and sensitive to restriction by *Fnu*4HI (data not shown). Hence the conservative amino acid changes D → E and V → L encountered in introducing the *Stu*I site into the $\rho 11_s$ Mtase gene are silent with respect to the phenotpye analyzed here.

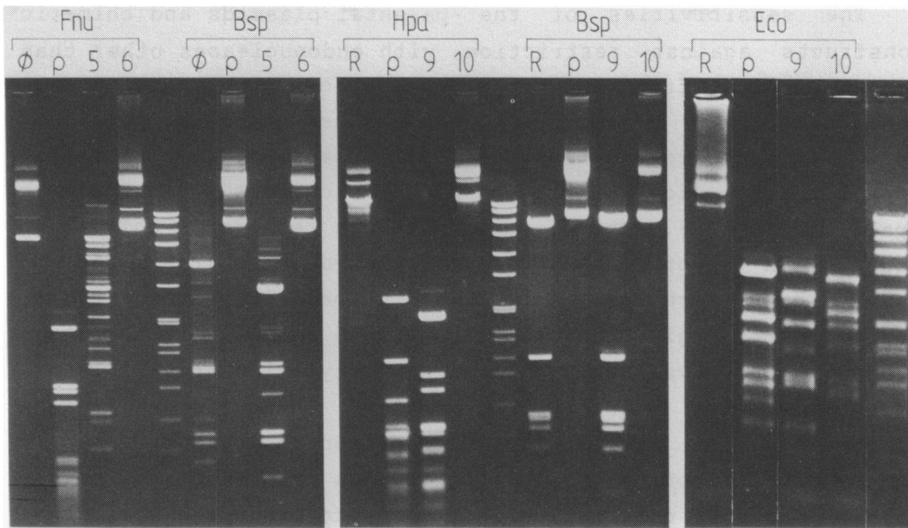


Fig. 3:
Electropherograms of digests with various restriction enzymes. Plasmids used were pBN16 (ϕ3T), pSB11 ($\rho 11_s$), pKB131 (SPR), pBB15 (chimera 5), pBB16 (chimera 6), pBB9 (chimera 9), pBB 10 (chimera 10). The plasmids were digested with Fnu4HI, Bsp1286, HpaII, and EcoRII as indicated. Plasmid and restriction enzyme denominations are abbreviated. EcoRI digested SPP1 DNA (16, tracks not marked) served as a molecular weight standard.

RESULTS AND DISCUSSION

To determine the methylation specificity of the Mtases encoded by the chimeric Mtase genes constructed, we have exposed plasmids pBB15/16 and pBB9/10 (Fig. 2) and plasmids containing the Mtase genes of ϕ3T, $\rho 11_s$, and SPR to relevant restriction endonucleases (Fig. 3, Fig. 1B). All plasmids analyzed were fully resistant to degradation by HaeIII (not shown). This demonstrates in the case of the chimeric genes reconstitution of a functional Mtase gene from the nonfunctional subgenomic StuI fragments. It also shows the interchangeability of the StuI generated subgenic fragments between $\rho 11_s$, ϕ3T and SPR. Furthermore, Mtase expression, as assayed here, is not measurably affected by the orientation of chimeric genes with respect to plasmid promoters in the various constructs of Fig. 2.

The sensitivities of the parental plasmids and chimeric constructs against restriction with endonucleases other than HaeIII follow from the experiment shown in Fig. 3. Here chimeras 5 and 9 had maintained neither of the methylations found in the corresponding parental plasmids, i.e. chimera 5 was sensitive to Fnu4HI and Bsp1286, chimera 9 sensitive to HpaII and Bsp1286. The plasmids with the reciprocal chimeras 6 and 10, on the other hand were methylated in their Bsp1268 and their Fnu4HI (chimera 6) or HpaII (chimera 10) sites. Hence, the in vitro engineered chimeric genes 6 and 10 encode active Mtases with novel combinations of methylation specificities, which have not been observed in vivo. Obviously the regions responsible for specificity recognition represent not only independently acting molecular domains as suggested before (7,8), but they can also be rearranged to produce new kinds of Mtases, in which the composite methylation capacities are compatible with each other.

The methylation specificities of the chimeric constructs also provide information about the domainal organization of target recognition sequences. (a) We can assign the Bsp1286 recognizing domain of $\rho 11_s$ to a region 5′ of the StuI joining site. The location of this domain had previously not been established. (b) The proficiency for HpaII methylation of chimera 10 is compatible with our previous assignment (7,8) of the CCGG recognizing domain to a region extending from a location at least 20 amino acids COOH-terminal to the StuI site. (c) With regard to the amino acid sequence requirements for recognition of the Fnu4HI or EcoRII sites by the Mtases, we had previously seen that mutations destroying the capacity for Fnu4HI or EcoRI methylation of ϕ3T and SPR were localized on both sides of the StuI site (8). No sequence recognizing domain could be assigned to the equivalent region in the $\rho 11_s$ gene. The absence of the SPR specific EcoRII methylating activity in chimeras 9 and 10 is readily understandable from the absence of amino acid identity in this region between the Mtases of SPR and $\rho 11_s$. Reciprocal fusions at the StuI site leading to chimeras 9 and 10 generate sequences strikingly different from either parental sequence. Corresponding results

have been observed with ϕ3T/SPR chimeras (7). Different from the $\rho11_s$ and SPR situation, amino acid homology is extensive between $\rho11_s$ and ϕ3T in the region concerned (Fig. 1C). From the proficiency of chimera 6 to methylate Fnu4HI sites we conclude that the ϕ3T and $\rho11_s$ sequences NH_2-terminal to the StuI site are interchangeable with respect to providing this function. Sequence differences COOH-terminal to the StuI site must therefore be responsible for the absence of Fnu4HI methylating activity both in $\rho11_s$ and in chimera 5. Differences between the two genes are apparent at amino acid coordinate 291, where G in ϕ3T is represented by E in $\rho11_s$. Also COOH-terminal to the A at coordinate 299 amino acid homology is absent. Experiments are presently performed to determine whether the $\rho11_s$ Mtase can be endowed with Fnu4HI methylating potential by site directed mutagenesis at these locations.

We have shown here that target recognizing domains can be newly combined. Our analysis, however, was limited to the phage Mtases with very similar core structures. At this time we are also attempting to interchange variable, non-core sequences between phage and bacterial Mtases. Provided active Mtases were formed also in such combinations, this could facilitate the definition of target recognizing domains also of monospecific bacterial Mtases. Furthermore, such experiments would allow one to determine to what extent individual core structures and target recognizing domains are adjusted to each other.

ACKNOWLEDGEMENT

We thank Sebastian Rausch, Peter-Andreas Terschüren, Mario Noyer-Weidner, and Roland Lauster for communication of unpublished results, advice and critical comments.

*Present address: Astra Research Centre, PO Box 359, Bangalore 560 003, India

REFERENCES

1. Buhk,H.J., Behrens,B., Tailor,R., Wilke,K., Prada,J.J., Günthert,U., Noyer-Weidner,M., Jentsch,S. and Trautner,T.A. (1984) Gene, 29, 51-61.

2. Tran-Betcke,A., Behrens,B., Noyer-Weidner,M. and Trautner, T.A. (1986) Gene, 42, 89-96.
3. Posfai,G., Baldauf,F., Erdei,S., Posfai,J., Venetianer,P. und Kiss,A. (1984) Nucl. Acids. Res. 12, 9039-9049.
4. Sznyter,L.A., Slatko,B., Moran,L., O'Donnell,K.H. und Brooks, J.E. (1987) Nucl. Acids. Res. 15, 8249-8266.
5. Behrens,B., Noyer-Weidner,M., Pawlek,B., Lauster,R., Balganesh,T.S. and Trautner,T.A. (1987) EMBO J., 6, 1137-1142.
6. Lauster, R. (1987) Doctoral Thesis, Freie Universität Berlin.
7. Balganesh,T.S., Reiners,L., Lauster,R., Noyer-Weidner,M., Wilke,K. and Trautner,T.A. (1987) EMBO J., 6, 3543-3549.
8. Wilke,K., Rauhut,E., Noyer-Weidner,M., Lauster,L., Pawlek, B., Behrens,B. and Trautner,T.A. (1988) EMBO J. in press.
9. Noyer-Weidner,M., Diaz,R. and Reiners,L. (1986) Mol. Gen. Genet., 205, 469-475.
10. Raleigh,E.A. and Wilson,G. (1986) Proc. Natl. Acad. USA, 83, 9070-9074.
11. Boyer,H.W. and Roulland-Dussoix,D. (1969) J. Mol. Biol., 41, 459-472.
12. Wilke,K. (1987) Doctoral Thesis, Freie Universität Berlin.
13. Günthert,U., Reiners,L. and Lauster,R. (1986) Gene 41, 261-270.
14. Noyer-Weidner,M., Jentsch,S., Kupsch,J., Bergbauer,M. and Trautner,T.A. (1985) Gene 35, 143-150.
15. Maniatis,T., Fritsch,F.E. and Sambrook,F. (1982) Molecular Cloning. A Laboratory Manual. Cold Spring Harbor Laboratory, New York, N.Y.
16. Ratcliff,S.W., Luh,J., Ganesan,A.T., Behrens,B., Thompson, R., Montenegro,M.A., Morelli,G. and Trautner,T.A. (1979) Mol. Gen. Genet., 168, 165-172.