METHYLATED OLIGONUCLEOTIDES DERIVED FROM BACTERIOPHAGE fd RF-DNA MODIFIED *IN VITRO* BY *E. COLI* B MODIFICATION METHYLASE

H. van ORMONDT, J.A. LAUTENBERGER, S. LINN and A. de WAARD

Laboratory for Applied Enzymology, University of Leiden, Wassenaarseweg 62, Leiden, The Netherlands

and

Department of Biochemistry, University of California, Berkeley, Calif. 94720, USA

Received 2 May 1973

1. Introduction

E. coli B modification methylase [1, 2] converts four adenine residues (presumably two per strand) to *N*(6)-methylaminopurine in wild-type unmodified fd RF-DNA. When RF-DNA of the one-site mutant fd sB_1^0 sB_2 (strain 101) is the substrate, the DNA receives two methyl groups while fd sB_1^0 sB_2^0 RF-DNA has lost its susceptibility to the methylase by mutation [3]. This methylation is concomitant with an increase in infectivity on *E. coli* B spheroplasts [3, 4]. These results probably mean that wild-type RF-DNA has two recognizable and specific nucleotide sequences ("sites"), fd sB_1^0 sB_2 one, and fd sB_1^0 sB_2^0 none, which by becoming methylated confer resistance to the *E. coli* B restriction enzyme.

The present study was undertaken to elucidate the structure of those sites which bear "*E. coli* B specificity". To this end, unmodified RF-DNA of wild-type fd was methylated *in vitro* in the presence of purified modification methylase from *E. coli* B and methyl tritiated S-adenosylmethionine. The [^{3}H-methyl]RF-DNA was degraded to small fragments, and the tritiated oligonucleotides characterized as to their sequence. Although a definitive sequence was not obtained by this approach, an analysis of the sequenced oligonucleotides showed the 3′- and 5′-nearest neighbours of the methylated adenine residues to be A or C, G or C, respectively. This finding demonstrates that the sequence surrounding the modified base does not contain a twofold rotational symmetry.

Abbreviations:

RF-DNA: replicative form DNA; SAM: S-adenosylmethionine; spd: spleen phosphodiesterase; svd: snake venom phosphodiesterase; A; *N*-6-methyl-deoxyadenosine.

2. Materials and methods

Approximately 5 A_{260} units of unmodified fd RF-DNA were incubated for 10 hr with highly purified *E. coli* B modification methylase [1] and [^{3}H-methyl]SAM (Amersham; specific activity 8.5 Ci/mmole). The DNA was exhaustively dialysed to remove excess [^{3}H] SAM, and had incorporated about 600,000 cpm of [^{3}H]methyl groups. Duplex [^{3}H] DNA (40,000 cpm) was incubated in a siliconized tube in 80 μl containing 25 mM Tris-HCl, pH 8.9, 5 mM $MgCl_2$ and 40 μg of pancreatic DNAase. Degradation of duplex [^{3}H]DNA (40,000 cpm) with micrococcal nuclease was done as described previously [5]. The digests were lyophilized and redissolved in 10 μl of 0.1 M EDTA. About 90% of the [^{3}H] DNA digests was separated by two-dimensional electrophoresis, according to Sanger et al. [6], after mixing it with a digest of fd [^{32}P]DNA prepared with the same enzyme. The ratio of ^{3}H/^{32}P was roughly 1:1. The remaining 10% of the ^{3}H-digest was separated one-dimensionally on the side of the same DEAE paper, in order to provide a

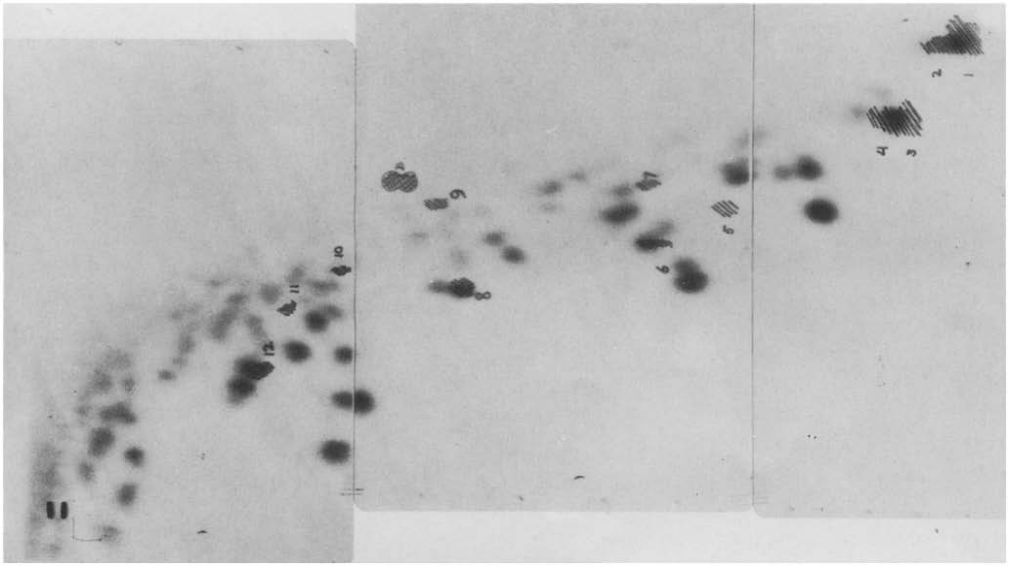

Fig. 1. Autoradiograph of micrococcal nuclease fingerprint of ^{32}P-labeled fd DNA. The shaded areas represent regions where micrococcal nuclease [^{3}H-methyl]oligonucleotides were localized. Numbers correspond with those of table 1. B = blue marker.

guide for locating the tritiated oligonucleotides in the fingerprint. This was performed by cutting the paper into 1 × 2 cm strips, which were suspended in scintillation fluid, and counted in a Tracerlab Corumatic apparatus. Beforehand, the ^{32}P spots had been localized by autoradiography.

The characterization of the oligonucleotides by digestion with alkaline phosphatase, snake venom phos-

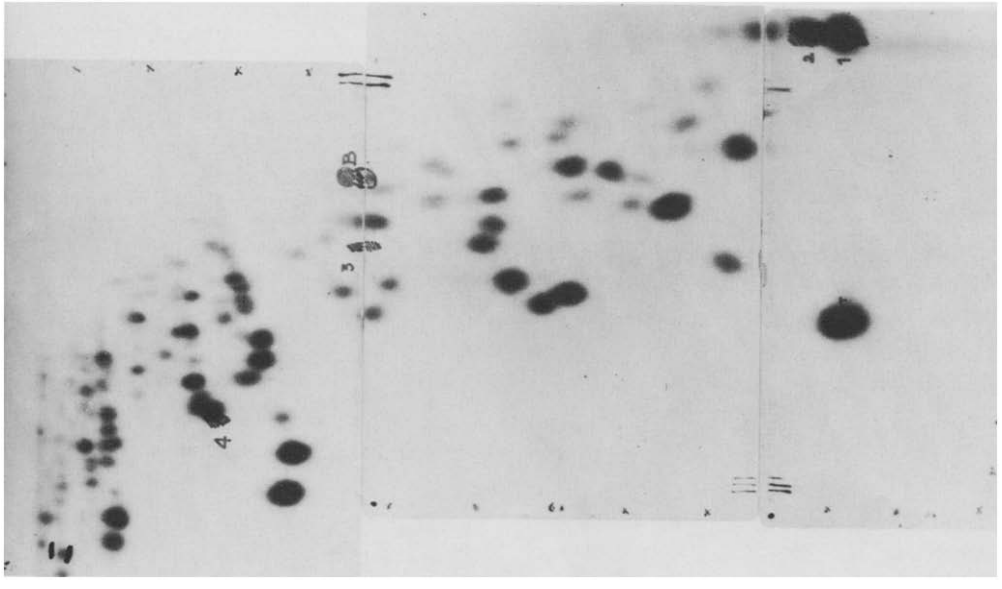

Fig. 2. Pancreatic DNAase fingerprint of ^{32}P-labeled fd DNA. The shaded areas represent regions where pancreatic DNAase [^{3}H-methyl]oligonucleotides were found. Numbers correspond with those of table 2. B = blue marker.

Table 1
^{3}H-containing oligonucleotides in micrococcal nuclease digests (see fig. 1).

Spot no.	^{3}H (cpm)	Sequence
M1	250	$\underline{\overset{*}{A}p}$
M2	80	$\overset{*}{A}$pCp
M3	350	ApGp$\overset{*}{A}$pCp
M4	450	$\underline{TpGp\overset{*}{A}p}$

Clearly established sequences are underlined.

phodiesterase (svd) and spleen phosphodiesterase (spd) was performed as described previously [5]. The ^{32}P-oligonucleotides in the relevant areas of the pancreatic DNAase fingerprint were sequenced in a similar way.

3. Results

Figs. 1 and 2 represent fingerprints of micrococcal nuclease and pancreatic DNAase digests of fd [^{32}P]DNA, respectively. The shaded areas are the sites where tritium counts were found in the paper. The methylated oligonucleotides coincided partially or completely with their non-methylated analogues.

In table 1 the results for the micrococcal nuclease digestion are summarized. Spot M1, coinciding with [^{32}P]Ap, yielded 6-Me-deoxyadenosine on phosphatase treatment. Spot M2 coincided with [^{32}P]ApCp. Since [^{32}P]CpAp was not found in micrococcal nuclease digests of either single-stranded or duplex DNA, spot M2 is surmised to be $\overset{*}{A}$pCp. Spot M3 coincided with [^{32}P]ApGpApCp. Partial spd hydrolysis showed that the 5′-nucleotide was non-methylated A, while one of the products had the mobility of $\overset{*}{A}$pCp. Therefore, M3 was surmised to be ApGp$\overset{*}{A}$pCp, since the other possibility ApGpCp$\overset{*}{A}$p (containing the AGC sequence) has a lower mobility than ApGpApCp [5]. Spot M4 was found in the [^{32}P]TpGpAp region. Its sequence was confirmed to be TpGp$\overset{*}{A}$p by digestion with phosphatase, and then cleavage of the dephosphorylated product with spd to yield Gp$\overset{*}{A}$ and $\overset{*}{A}$.

The scarcity of small labeled oligonucleotides formed by micrococcal nuclease forced us to use another, less specific DNAase. More tritiated oligonucleotides were found in a pancreatic DNAase fingerprint, as shown in fig. 2. The sequence results are summarized in table 2. Spot P1, coinciding with [^{32}P]p(A,C) was shown to be a mixture of p$\overset{*}{A}$pC and pCp$\overset{*}{A}$ by svd and spd hydrolysis after phosphatase digestion. The mobility of P2, before or after phosphatase digestion indicated that its composition is p($\overset{*}{A}C_2$) and degradation with svd and spd showed that $\overset{*}{A}$ was in the middle position. Spot P3 coincided with [^{32}P]pApA and after dephosphorylation, yielded methyladenosine on svd digestion, so that its structure was deduced to be p$\overset{*}{A}$pA. P4 is distinct from spot P3, and carries an extra C, as judged from its mobility after phosphatase digestion. Spot P5 coincided with [^{32}P]TpGpA, a dephosphorylated oligonucleotide produced by a phosphomonoesterase impurity in the pancreatic DNAase preparation we used. P5 has the same mobility as [^{32}P]TpGpA at pH 3.5 and contains a 3′-terminal $\overset{*}{A}$. Spot P6 coincides with [^{32}P]pGpA; consecutive phosphatase and spd digests yield products travelling as [^{32}P]GpA and $\overset{*}{A}$, respectively. Spot P7 coincides with [^{32}P]pGpCpA or pCpGpA and carries the methylated adenine in the 3′-terminal position.

Table 2
^{3}H-containing oligonucleotides in pancreatic DNAase digests (see fig. 2).

Spot no.	^{3}H (cpm)	Sequence
P 1	800	$\underline{p\overset{*}{A}pC}$ + $\underline{pCp\overset{*}{A}}$
P 2	250	pCp$\overset{*}{A}$pC
P 3	550	p$\overset{*}{A}$pA
P 4	120	p($\overset{*}{A}$,A,C)
P 5	100	TpGp$\overset{*}{A}$
P 6	250	$\underline{pGp\overset{*}{A}}$
P 7	150	p(C,G)$\overset{*}{A}$
P 8	200	$\underline{p\overset{*}{A}pApT}$
P 9	150	p(A,G,C)$\overset{*}{A}$
P10	60	p$\overset{*}{A}$pApTpA ?
P11	100	?
P12	800	$\underline{pTpGp\overset{*}{A}}$

Clearly established sequences are underlined.

Spot P8 partially coincides with [^{32}P]pApApT. After simultaneous dephosphorylation and partial svd digestion it yielded products with the mobility of [^{32}P]ApApT and ApA and the dinucleotide had methyladenosine away from the 3′-terminus. The sequence of P8 was therefore deduced to be p$\overset{*}{\text{A}}$pApT. Spot P9 coincides with [^{32}P]nucleotides of composition p(A_2GC). The methylated adenine is in the 3′-position; presumably, it has the structure p(pA, pG, pC)p$\overset{*}{\text{A}}$. Spot P10 contains very few tritium counts. It partially coincides with [^{32}P]pApTpApA, and the 3′-nucleotide is not methylated. In view of this evidence, and of the sequence of spot P8, the primary structure of P10 might be p$\overset{*}{\text{A}}$pApTpA. Spot P11 was not characterized. Spot P12 coincides with [^{32}P]pTpGpA; phosphatase treatment and then spd digestion yielded products travelling like [^{32}P]TpGpA, GpA and 6-methyl-deoxyadenosine; its sequence, therefore, should be pTpGp$\overset{*}{\text{A}}$.

3. Discussion

The 3′- and 5′-terminal sequences of the DNA fragments produced by scission with restriction endonucleases R (from *Hemophilus influenzae* Rd) R_I and R_{II} (coded for by resistance factors in *E. coli*) have been determined [7–9]. The recognition sites deduced from these data show a twofold rotational symmetry. Attempts to sequence the restriction sites formed by the restriction enzymes of *E. coli* B or K have been unsuccessful due to the resistance of the termini formed by these enzymes to phosphorylation by polynucleotide kinase. Moreover, a report by Horiuchi and Zinder [10] suggests that the breaks might not be introduced into recognition sites, but elsewhere in the molecule. Our method using the methylase is technically more difficult and has not yielded enough information to obtain a unique sequence. However, our results (table 1 and 2) show that the methyladenine residues formed by the *E. coli* B modification enzyme have two different 5′-neighbours, and two 3′-neighbours. Given these "nearest neighbours" it is not possible to construct a double-stranded sequence with twofold rotational symmetry; hence, the sequence must be of another type than that which applies to the sites specific for the *Hemophilus influenzae* Rd [7], R_I [8] and R_{II} [9] enzymes.

References

[1] J.A. Lautenberger and S. Linn, J. Biol. Chem. 19 (1972) 6176.
[2] U. Kühnlein, S. Linn and W. Arber, Proc. Natl. Acad. Sci. U.S. 63 (1969) 556.
[3] U. Kühnlein and W. Arber, J. Mol. Biol. 63 (1972) 9.
[4] A. De Waard, Proc. Kon. Ned. Acad. Wet. (A'dam) C74 (1971) 183.
[5] H. Van Ormondt and J.W. Maagdenberg, Biochim. Biophys. Acta 272 (1972) 141.
[6] F. Sanger, G.G. Brownlee and B.G. Barrell, J. Mol. Biol. 13 (1965) 373.
[7] T.J. Kelly and H.O. Smith, J. Mol. Biol. 51 (1970) 393.
[8] J. Hedgpeth, H.M. Goodman and H.W. Boyer, Proc. Natl. Acad. Sci. U.S. 69 (1972) 3448.
[9] K. Murray, personal communication.
[10] K. Horiuchi and N.D. Zinder, Proc. Natl. Acad. Sci. U.S. 69 (1972) 3220.