M.SmaI is an N4-methylcytosine specific DNA-methylase

Saulius Klimašauskas⁺, Dana Steponavičienė, Zita Manelienė, Marytė Petrušytė, Viktoras Butkus and Arvydas Janulaitis*

Institute of Applied Enzymology, Fermentu 8, 232028 Vilnius, Lithuania, USSR

Received July 18, 1990; Revised and Accepted October 16, 1990

ABSTRACT

An enzymatic activity rendering DNA immune to the action of the SmaI restriction endonuclease in the presence of S-adenosyl-L-methionine has been detected in *Serratia marcescens* Sb. This methylase, M.SmaI, modifies the second cytosine residue of the substrate sequence CCCGGG yielding N4-methylcytosine.

INTRODUCTION

Serratia marcescens Sb is the source of the restriction endonuclease SmaI, which cuts the sequence CCC↓GGG as indicated by the arrow [1]. The endonuclease activity *in vitro* is inhibited by C5-methylation of the central CG dinucleotide [2,3] as well as by the presence of N4-methylcytosine (m4C) in either position of the recognition sequence [3]. *S.marcescens* DNA contains 5-methylcytosine (m5C) and N6-methyladenine (m6A), but no detectable m4C residues [3,4]. No enzymatic activity has been identified *in vitro* capable of protecting DNA from the restriction enzyme [1,5]. Recent cloning and sequence analysis of the genes coding for the SmaI restriction-modification system [5] suggested that the modification enzyme might be a DNA[cytosine-N4]methyltransferase [5,6].

We now report the isolation and initial characterization of the SmaI methylase.

MATERIALS AND METHODS

All chemicals were analytical reagent or higher grade commercial products. All enzymes and DNAs used are products of Fermentas (Vilnius) except for Nuclease P1 and venom phosphodiesterase which were purchased from Pharmacia and Boehringer Mannheim, respectively. *S.marcescens* strain Sb was from the collection of the Institute of Physiology and Biochemistry of Microorganisms Academy of Sciences, USSR. The dodecanucleotide dGGACCCGGGTCC, the methylated deoxynucleosides m4dCyd and m5dCyd were synthesized in our laboratory [7]. m6dAdo was purchased from Sigma.

Bacterial cells were grown in media containing yeast extract (5 g/l), peptone (5 g/l), glucose (2 g/l) and K-PO₄ (1 g/l) pH7.5. 150 g of the cell mass were resuspended in 300 ml of 10 mM K-PO₄ pH7.4, 1 mM EDTA, 7 mM 2-mercaptoethanol and

0.15 M KCl (Buffer A) and disrupted by sonication for 15 min. The solution was clarified by centrifugation at 20,000 rpm for 2 hours.

The crude extract was applied to a 2.5 cm×25 cm heparin sepharose column (Pharmacia) equilibrated with Buffer A. The column was developed by 1000 ml of a linear gradient from 0.15–0.8 M KCl at 60 ml/min in Buffer A. 10 ml fractions were collected and assayed for methylase activity. Active fractions (0.28–0.32 M KCl) were pooled and used as the source of methylase within several days.

Methylation reactions were performed by adding 5 µl of the crude methylase to 40 µl of the methylation buffer (20 mM Tris-HCl pH8.0, 50 mM NaCl, 7 mM 2-mercaptoethanol, 1 mM EDTA, 0.1 mg/ml bovine serum albumin) containing 0.1 mM S-adenosyl-L-methionine (SAM) and 2 µg of λ phage DNA and incubating at 37°C for 1 hour. The reaction mixture was then heated at 65°C for 15 min and made 10 mM MgCl₂ by adding 5 µl of a 250 mM solution followed by digestion with 20 units of R.SmaI at 37°C for 1 hour. The digests were analysed by electrophoresis on 0.7% agarose gel.

The dodecanucleotide dGGACCCGGGTCC (0.1 A₂₆₀ units) was annealed in 0.1 ml of the methylation buffer. 10 µl of [³H-methyl]-S-adenosyl-L-methionine ([³H-methyl]-SAM) (15 Ci/mmol, 1 mCi/ml, Amersham) and 8–15 µl of the methylase preparation were added. The reaction mixture was incubated overnight at 30° or 37°C, heated at 70°C for 5 min and then the oligo was isolated by gel-filtration on Sephadex G-25 (Pharmacia).

For methylated base analysis the modified substrate was digested to nucleosides with nuclease P1 and bacterial alkaline phosphatase [3]. The resulting hydrolysate was made 100 mM K-PO₄ pH3.0 by adding 1 M stock solution and was mixed with m4dCyd, m5dCyd and m6dAdo standards. Samples were analysed on a Gilson gradient dual-wavelength HPLC system equipped with a NovaPak C-18 analytical column (Waters). The column was eluted with 25 mM K-PO₄ pH3.0 for 8 min followed by a linear acetonitrile concentration gradient to 30% in 17 min at 1 ml/min. Fractions corresponding to those of standard deoxynucleosides were collected and counted for ³H-radioactivity in toluene-triton scintillation fluid.

The procedures described previously were followed when analyzing the position of the methylated nucleotide [3] and

* To whom correspondence should be addressed

⁺ Present address: Cold Spring Harbor Laboratory, PO Box 100, Cold Spring Harbor, NY 11724 USA

comparing methylase sequences [6,8]. The sequences for the *SmaI* and *Cfr9I* methylases are from [5] and [6], respectively.

RESULTS AND DISCUSSION

During chromatography of a cellular extract from *S.marcescens* on heparin sepharose some fractions showed methylase activity that rendered DNA immune to the action of *R.SmaI* (Fig. 1). The activity required the presence of the methyl donor, S-adenosyl-L-methionine. Further purification proved impossible since after dialysis of the fractions from heparin sepharose the enzyme rapidly loses its activity. We have not yet been able to stabilize these fractions, although they were suitable for immediate use, in the experiments described below. This instability probably accounts for earlier failures [5] to detect activity.

Figure 1. Analysis of methylase activity in fractions during chromatography of a cellular extract from *S.marcescens* on heparin sepharose. The first 35 fractions out of 97 collected and analysed are shown. Fractions 17 through 23 render DNA immune to *R.SmaI*.

Figure 2. HPLC analysis of the nuclease P1 and alkaline phosphatase hydrolysis products of the [^3H -methyl]-modified dodecanucleotide with internal deoxynucleoside standards: A-m4C, B-m5C and C-m6A.

The methylase preparation was able to catalyse the transfer of methyl groups from [^3H -methyl]-SAM onto a double-stranded dodecanucleotide dGGACCCGGGTCC. This is an enzyme-specified modification of the substrate since ^3H -incorporation in the presence of the heat-inactivated methylase was lower by at least two orders of magnitude (not shown). The use of the *SmaI*-specific substrate eliminates possible effects of other methylating activities potentially present in the partially purified preparation. We can not exclude the possibility that the methylase itself recognizes a subset of the CCCGGG sequence, but that seems very unlikely in view of the extremely low abundance of *SmaI*-modified sites in the host DNA (see discussion below).

The modified oligonucleotide was further used to investigate the specificity of the methylase. An aliquot of the modified substrate was hydrolysed to deoxynucleosides and analysed by HPLC using corresponding m4C, m5C and m6A standards. A previously described elution system [3] was improved to increase the separation of the two methyldeoxycytidines (Fig.2). The peak of ^3H -radioactivity (31400 cpm, 98% of total counts) coincided with that of the m4C standard showing that *M.SmaI* forms N4-methylcytosine rather than 5-methylcytosine. This result is in accordance with the prediction from the primary structure of the methylase deduced from the DNA sequence [5]. The methylase sequence contains a TSPPY---(K,R) motif which is characteristic of the m4C-enzymes [6] but lacks most of the conserved patterns present in all m5C DNA-methylases [9]. The previous failure to detect m4C residues in *S.marcescens* chromosomal DNA [3,4] might be due to the underrepresentation of the recognition sequence in the host genome. From the major

Table 1. ^{32}P - and ^3H -radioactivity distribution in the products of partial venom phosphodiesterase digestion of pGGACCCGGGTCC methylated *in vitro* with *M.SmaI*.

Hydrolysis product	Radioactivity ^3H	Radioactivity ^{32}P	Ratio $^3\text{H}/^{32}\text{P}$
pGG	40	140	0.3
pGGA	50	100	0.5
pGGAC	90	250	0.4
pGGACC	150	120	1.3
pGGACCC	240	150	1.6

10	20	30	40	50	60	70
MPSKSSPLSVKLRSEPLEINGATLFEQDALSVLRRLPSGVSVCIVTSPFYWGLRDYDQIGLESSMTQFIN						
* * * * *						
MKKSHNNLDLFDQTECLESNLRSGKIIIVGDAREAVQGLDSEIFDCVVTSPFYWGLRDYDNGGGQGAEDNINDYIK						
10	20	30	40	50	60	70
80	90	100	110	120	130	140
RLVTIFSEAKRVLTDDGTLNVNIGDGYTSGNRYRAPDKKNPARAMAVRDRPEGPKPKDLIGIPWRLAFALQEDGW						
* * * * *						
DLVDLFRDVRTRLKDDGTLNINIGDYSYSGRTWRDKDKNKGKRAMSYRPTPEGLKPKDLIGIPWRLAFALQNDGW						
80	90	100	110	120	130	140
160	170	180	190	200	210	220
YLRSDIVNKNPNAMESVKDRPTRSHEFLFMTKSEKYYDWEAVREKDSGGF--RNKRTVMNVNTPKPFAGAHFAT						
* * * * *						
YLRDIIWNKPNQPSVDRPTRSHEYIFLLSGKYYDWEIKEPASDPKMDKKNRTVMNINTEPYGSHFAV						
160	170	180	190	200	210	220
230	240	250	260	270	280	290
FPTLIRPCILASTKPGDYVLDPPFGSGTGVGVVQDQYVGIENFETVDIAVNRQGEDTNVIRIAAA						
* * * * *						
FPRMARLCVLGRSPGGKVLDPFFGSGTGVGVVQDQYVGIENFETVDIAVNRQGEDTNVIRIAAA						
240	250	260	270	280	290	

Figure 3. Alignment of the *Cfr9I* (upper) and *SmaI* (lower) methylase sequences: * - identity, :- conservative substitution.

base composition of *S.marcescens* DNA, 40% A+T [4], methylation of the CCCGGG sequences would be expected to yield $0.3^6 \approx 0.07\text{mol}\%$ m4C in the genome if random distribution of bases is assumed. The actual level of m4C in this DNA lies beyond the sensitivity of detection by the HPLC technique used [3,4] and, therefore, might not exceed 0.01mol% [4], implying at least a sevenfold underrepresentation of the *Sma*I-modified sites.

The position of the methylated nucleotide within the target sequence has also been established. A portion of the modified oligonucleotide was ^{32}P -kinased and subjected to partial venom phosphodiesterase digestion. The hydrolysis products were separated by 2-dimensional electrophoresis-homochromatography (not shown) and analysed for ^3H -radioactivity [3]. The shortest fragment carrying both 5'- ^{32}P - and ^3H -methyl-labels was GGACC (Table 1), proving that the second cytosine residue of the recognition sequence is modified. ^3H -channel counts of the former three oligonucleotide spots are most probably due to natural background and/or the low energy fraction of the ^{32}P signals.

Thus, the specificity of the *Sma*I methylase is Cm4CCGGG, the same as that of *M.Cfr9I* [3]. This might have been expected given the high amino acid sequence homology of the methylases [8]: 159 residues are identical and 31 residues are conservative substitutions in the alignment of 288 amino acids (Fig.3). This structural and functional similarity suggests a close evolutionary relationship between the two methylases. In contrast, the corresponding restriction endonucleases are neoschizomers since *R.Sma*I cleaves in the middle of the sequence [1] whereas *R.Cfr9I* leaves tetranucleotide 5'-extensions [3]. No significant similarities have been found between their amino acid sequences (S.Klimasauskas, S.Menkevicius, A.Lubys, V.Butkus, A.Janulaitis, in preparation). This suggests that the modification and endonuclease counterparts of the *Sma*I and *Cfr9I* restriction-modification systems have evolved by separate evolutionary pathways.

ACKNOWLEDGEMENTS

The authors thank Dr.R.J.Roberts for helpful comments on the manuscript and J.Duffy and M.Ockler for artwork.

REFERENCES

1. Mulder,C., Greene,R. cited in Roberts,R.J. (1989) Nucleic Acids Res. **17**,r347-r387.
2. Youssofian,H., Mulder,C. (1981) J. Mol. Biol. **150**,133-136.
3. Butkus,V., Petrauskienė,L., Maneliene,Z., Klimasauskas,S., Laucys,V., Janulaitis,A. (1987) Nucleic Acids Res. **15**,7091-7102.
4. Ehrlich,M., Wilson,G.G., Kenneth,C.K., Gehrke,C.W. (1987) J. Bacteriol. **169**,939-943.
5. Heidmann,S., Seifert,W., Kessler,C., Domdey,H. (1989) Nucleic Acids Res. **17**,9783-9796.
6. Klimasauskas,S., Timinskas,A., Menkevicius,S., Butkiene,D., Butkus,V., Janulaitis,A. (1989) Nucleic Acids Res. **17**,9823-9832.
7. Petrauskienė,L., Klimasauskas,S., Butkus,V., Janulaitis,A. (1986) Bioorg. Khim. **12**,1597-1603.
8. Klimasauskas,S., Timinskas,A., Menkevicius,S., Butkiene,D., Butkus,V., Janulaitis,A. (1990) Eksperimentine Biologija (Vilnius) **1**,7-12.
9. Posfai,J., Bhagwat,A.S., Posfai,G., Roberts,R.J. (1989) Nucleic Acids Res. **17**,2421-2435.