Volume 15 Number 17 1987 Nucleic Acids Research

Cleavage of methylated CCCGGG sequences containing either N4-methylcytosine or 5-methylcytosine with MspI, HpaII, SmaI, XmaI and Cfr9I restriction endonucleases

V.Butkus, L.Petrauskienė, Z.Manelienė, S.Klimašauskas, V.Laučys and A.Janulaitis

ESP Fermentas, All Union Research Institute of Applied Enzymology, Fermentu 8, Vilnius 232028, Lithuanian SSR, USSR

Received June 5, 1987; Revised and Accepted August 5, 1987

ABSTRACT

The cleavage specificity of R.Cfr9I was determined to be C↓CCGGG whereas the methylation specificity of M.Cfr9I was C4mCCGGG. The action of MspI, HpaII, SmaI, XmaI and Cfr9I restriction endonucleases on an unmethylated parent d(GGACCCGGGTCC) dodecanucleotide duplex and a set of oligonucleotide duplexes, containing all possible substitutions of either 4mC or 5mC for C in the CCCGGG sequence, was investigated. It was found that 4mC methylation, in contrast to 5mC, renders the CCCGGG site resistant to practically all the investigated endonucleases. The cleavage of methylated substrates with restriction endonucleases is discussed.

INTRODUCTION

Research in bacterial restriction-modification (RM) received a new impetus with the discovery of a new type of cytosine methylation, i.e. with the discovery of DNA-methylases producing N4-methylcytosine (4mC) and thus preventing cleavage of host DNA [1]. At present 4mC is known to be present in the DNA of a large number of thermophilic and mesophilic bacteria [2,3]. The present report on methylase Cfr9I, producing the C4mCCGGG methylated sequence, is yet another addition to the list of methylases (M.BcnI - C4mC(C/G)GG [1]; M.MvaI - C4mC(A/T)GG [4]; M.PvuII and M.Cfr6I - CAG4mCTG [5]) of the 4mC type.

Besides complicating the accepted nomenclature of DNA methylases, the discovery of DNA(cytosine-N4)methyltransferases has brought up problems in restrictase sensitivity to the interchange of methylcytosines at the cognate methylation position of the recognition site, as well as in the classification of restriction endonucleases on the basis of their sensitivity to methylated substrates.

The principal task of the present study was to investigate cleavage of some methylated substrates with MspI, HpaII, SmaI and

XmaI restriction endonucleases of the CG type, which are commonly used in methylated DNA analysis [6,7]. These investigations were performed by the use of synthetic dodecadeoxynucleotide d(GGACCCGGGTCC) as well as a set of dodecanucleotides, containing all possible substitutions of either 4mC or 5mC for C in the CCCGGG sequence.

MATERIALS AND METHODS

Enzymes. Isolation procedures for R.Cfr9I and M.Cfr9I did not essentially differ from those published earlier [8]. Restriction endonucleases XmaI and HpaII were isolated according to the procedures described in [9,10]. Endonucleases MspI and SmaI and T4-polynucleotide kinase were commercial products of ESP Fermentas (Vilnius). All the endonucleases were checked for purity by prolonged incubation with 5'-^{32}P-labeled synthetic double- and single-stranded oligodeoxynucleotide substrates with excess of enzymes. Proteinase K and pancreatic DNAse were purchased from Serva, SPDE and VPDE from Merck, Nuclease P1 from Pharmacia and alkaline phosphatase from Sigma.

DNAs. Xanthomonas malvacearum and Serratia marcescens DNAs were isolated following procedure [11].

Materials. Chemicals were purchased from the following commercial sources: [γ-^{32}P]ATP from Isotope (Tashkent), Sephadex G-50f from Pharmacia. Homochromatography was performed on home-made TLC plates coated with a mixture (5:1) of microcrystalline cellulose MN 300 (Serva) and DEAE cellulose DE-41 (Whatman). All other reagents used were analytical grade commercial products.

Oligodeoxynucleotides The procedures for the synthesis, purification and characterization of oligodeoxynucleotides, including those containing N4-methylcytosine were described in [12].

DNA analysis. DNA (50 μg) was hydrolyzed to deoxy-nucleosides by the procedure of Gehrke et al. [13] with minor modifications. Deoxynucleoside composition was analyzed on Gilson model 303 gradient LC system with Spectroflow 773 absorbance detector (Kratos) equipped with chromatographic data analysis system Appligration II, (Dynamic Solutions Corporation) and NovaPak C-18 analytical column (Waters). Separation was performed

in 25 mM potassium phosphate (pH 4.7), with an acetonitrile gradient from 0.5% (for 0-8 min) to 30% (for 8-25 min) at a flow rate of 1 ml/min.

Enzyme specificity determination. The R.Cfr9I cleavage site was determined using the synthetic d(GGACCCGGGTCC) self-complementary dodecadeoxynucleotide, containing the CCCGGG recognition sequence. The 5'-^{32}P-labeled oligonucleotide duplex was cleaved with the enzyme, and the reaction products were analyzed by homochromatography [14].

The specificity of M.Cfr9I was determined using the same 5'-^{32}P-phosphorylated dodecanucleotide, which additionally was ^{3}H-methylated with M.Cfr9I in the presence of ^{3}H-SAM. The double-labeled oligonucleotide thus obtained was analyzed by the fingerprint technique [14], while ^{32}P-labeled areas were scraped off and counted for ^{3}H and ^{32}P radioactivity.

Restriction endonuclease assay. The oligonucleotides were tested as substrates for MspI, HpaII, XmaI, SmaI and Cfr9I endonucleases by combining 5'-^{32}P-labeled (approx. 4,000 cpm/µl) and unlabeled 5'-phosphorylated oligonucleotides in order to obtain the desired final concentration (1 µM in the reaction volume of 20 µl). The mixture of oligonucleotides was dissolved in a buffer for the corresponding endonuclease (5-15 µl), warmed to 75°C and slowly cooled to the required temperature. The reaction was initiated by adding 5-15 µl of endonuclease solution. Incubation time and temperatures were selected individually for each enzyme. Aliquots of the reaction mixture (5-10 µl) were spotted onto DEAE-cellulose TLC plate and chromatographed in homomix VI [14]. The amount of products formed was quantified by liquid scintillation counting of the excised areas containing radioactivity. The efficiency of substrate cleavage was calculated from the total amount of introduced radioactivity and expressed in per cent. At least three independant assays were performed for each enzyme-substrate combination.

RESULTS

Determination of Cfr9I endonuclease and methylase specificities. A preliminary study of R.Cfr9I and M.Cfr9I enzymes

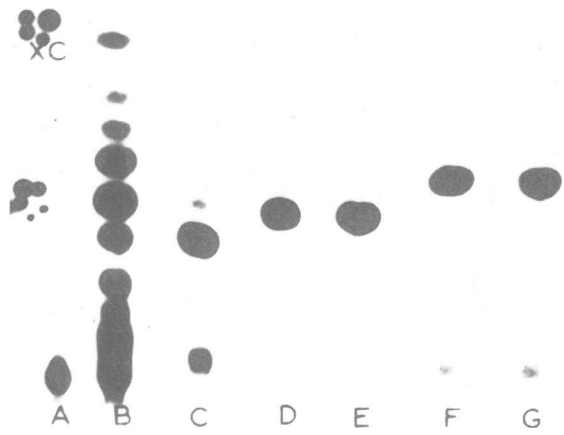

Figure 1. Cleavage of the 5'-^{32}P-labeled dodecanucleotide substrate d(GGACCCGGGTCC) with restriction endonucleases under investigation. Line (A) - initial dodecanucleotide; (B) - VPDE hydrolyzate; (C) - SmaI; (D) - HpaII; (E) - MspI; (F) - XmaI; (G) - Cfr9I. The reaction products were analyzed by TLC on DEAE-cellulose in homomix VI [14].

proved their ability to recognize the CCCGGG nucleotide sequence, in which the methylase produced N4-methylcytosine [8]. The present investigation shows the endonuclease to cleave the site after the first cytosine residue - C↓CCGGG (Fig.1). Thus, R.Cfr9I is an isoschizomer of R.XmaI, possessing the same cleavage specificity [9]. The Cfr9I methylase modifies the internal cytosine as follows: C4mCCGGG (Fig.2).

DNA analysis. HPLC analysis of DNA hydrolyzates was performed using a chromatography system specially designed for unequivocal identification of 4mdCyd and 5mdCyd. For purposes of additional characterization, identification of respective methylated deoxycytidines based on absorbance ratios of peaks at two wavelengths was performed [13]. We found 5mC and 6mA to exist in the DNA of S.marcescens (not shown) and three methylated bases - 4mC, 5mC and 6mA in X.malvacearum DNA (Fig.3).

Cleavage of oligonucleotide duplexes with MspI, HpaII, XmaI, SmaI and Cfr9I endonucleases. The cleavage specificities of the investigated endonucleases and their cognate methylated sites are presented in Table 1. All endonucleases gave expected the cleavage products, verified by homochromatography alongside the VPDE hydrolysate of original 5'-labeled oligonucleotide.

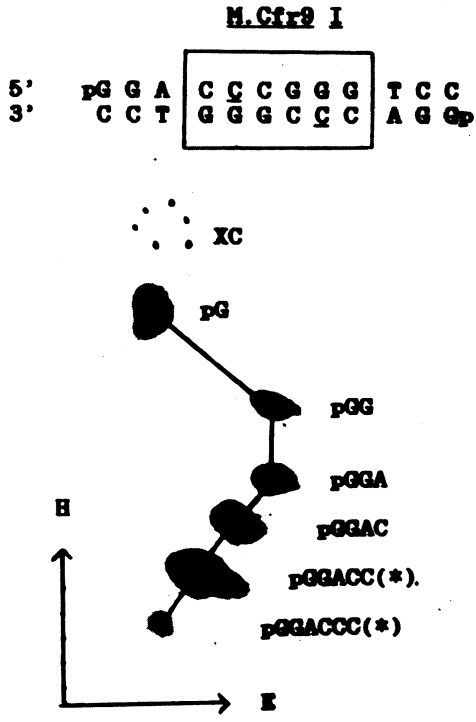

Figure 2. Nucleotide chart of the 5'- ^{32}P-labeled dodecanucleotide d(GGACCCGGGTCC) methylated with Cfr9I methylase in the presence of ^{3}H-SAM. Spots containing ^{3}H-radioactivity are marked by asterisk. The recognition site of M.Cfr9I is boxed and the determined methylated base is dashed.

Endonucleases released the following products: XmaI and Cfr9I tetranucleotide d(pGGAC), MspI and HpaII - pentanucleotide d(pGGACC) and SmaI - hexanucleotide d(pGGACCC) (Fig.1). Optimal assay temperatures, selected for each endonuclease separately, were as follows: 37°C for XmaI and Cfr9I, and 4°C for SmaI, HpaII and MspI. Cleavage of unmethylated substrates exceeded 90% in all studied endonucleases, with the exception of R.Cfr9I, which had approximate 50% cutting efficiency.

The cleavage of methylated oligonucleotide duplexes were also performed under the conditions found optimal for unmethylated substrates. However, with regard to some endonucleases, these particular methylated oligonucleotides demonstrated either very low or zero degree of cleavage (Table 2). The presence of inhibitors was tested by examining the

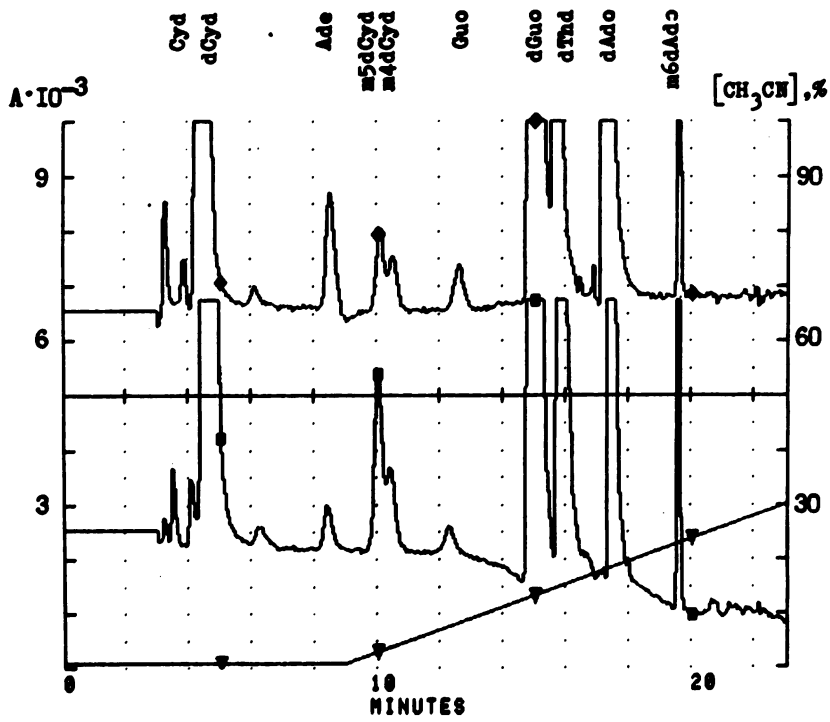

Figure 3. HPLC analysis of X.malvacearum DNA hydrolysate. The elution profiles marked with ■ and ◆ symbols correspond to UV absorbtion at 280 and 254 nm, respectively. Chromatographic conditions are described in the Materials and Methods section.

ability of the methylated oligonucleotides to inhibit cleavage of the unmethylated substrate. In order to avoid heteroduplex formation, the methylated and unmethylated duplex solutions were mixed at 0°C, and then allowed to reach the required temperature. In all cases, the unmethylated (control) substrate was hydrolyzed with an identical efficiency both individually

Table 1. Cleavage and cognate methylation specificity of the investigated endonucleases

Endonuclease	Cleavage	Cognate methylation
MspI	C↓CGG	5mCCGG
HpaII	C↓CGG	C5mCGG
SmaI	CCC↓GGG	?
XmaI	C↓CCGGG	?
Cfr9I	C↓CCGGG	C4mCCGGG

Table 2. Cleavage of methylated substrates with restriction endonucleases

Site	Restriction endonuclease				
	MspI	HpaII	SmaI	XmaI	Cfr9I
CCCGGG	97	94	92	91	51
5mCCCGGG	96	92	5	1	5
C5mCCGGG	0	20	70	10	8
CC5mCGGG	50	0	0	23	46
4mCCCGGG	89	80	0	0	0
C4mCCGGG	31	0	0	0	0
CC4mCGGG	16	0	0	0	0
Reaction conditions	4°C, 20h, 1 u/pmol	4°C, 22h, 2 u/pmol	4°C, 5h, 0.5 u/pmol	37°C, 18h, 3 u/pmol	37°C, 20h, 3 u/pmol

The u/pmol is unit of endonuclease per 1 pmol of substrate (single strands). 1 u is the amount of endonuclease required for complete digest of 1 μg of lambda DNA within 1 hour at 37°C.

and in a mixture with the methylated dodecanucleotide, which allowed us to exclude the possibility of adventitious enzyme inhibitors.

As was expected, neither 4mC nor 5mC neighboring the CCGG recognition sequence, inhibited cleavage with R.MspI and R.HpaII; both endonucleases also were unable to cleave their cognate methylated sites - 5mCCGG and C5mCGG respectively (Table 2). Cleavage of non-cognate methylated sites (5mCCGG with R.HpaII and C5mCGG with R.MspI) proceeded with a lower degree of efficiency in comparison with unmethylated substrates (20 and 50% respectively). Endonuclease HpaII did not cleave any site with N4-methylcytosine, while R.MspI hydrolyzed 4mCCGG and C4mCGG sites to 30 and 15% correspondingly.

For SmaI, XmaI and Cfr9I endonucleases the 5mCCCGGG sequence was not a satisfactory substrate, since it was hydrolyzed to 14, 2 and 5% respectively. In the same order the three endonucleases hydrolyzed the C5mCCGGG site to 70, 30 and 8%. 5-Methylcytosine in the rightmost position of the site (CC5mCGGG) made it immune to R.SmaI, but R.XmaI and R.Cfr9I cleaved the site to 23 and 46%. All the three endonucleases indicated could hydrolyze neither of sites with N4-methylcytosine.

Cleavage of hemimethylated substrates. The methylated oligonucleotide duplexes poorly cut by a particular endonuclease, were tested in hemimethylated duplexes. The probes were prepared by

Table 3. Cleavage of hemimethylated substrates with restriction endonucleases

Site	Restriction endonuclease				
	MspI	HpaII	SmaI	XmaI	Cfr9I
5' 5mCCCGGG 3' GGGCCC	[+++] [+++]	[+++] [+++]	+ ++	- -	- -
5' C5mCCGGG 3' G GGCCC	[++] [-]	++ ++	[+++] [+++]	++ ++	+ ++
5' CC5mCGGG 3' GG GCCC	[+++] [+++]	[--] [--]	[---] [---]	+++ ++	++ ++
5' 4mCCCGGG 3' GGGCCC	[+++] [+++]	[+++] [+++]	- -	- -	- -
5' C4mCCGGG 3' G GGCCC	[++] [++]	+ ++	- +	-- +	[--] [-]
5' CC4mCGGG 3' GG GCCC	+++ ++	- ++	- -	- -	- -

The results of cognate hemimethylated site cleavage are boxed. Dashed box denotes predicted cognate site.
(+++) - denotes strand cleavage higher than 50%; (++) - 20-50%; (+) - 5-20%; (-) - 1-5% and (--) - absence of cleavage. Results in the squared brackets are taken from Table 2.

annealing corresponding methylated and unmethylated dodecanucleotides. In order to differentiate cleavage of methylated and unmethylated strands, two kinds of probes were prepared. Firstly, the methylated dodecanucleotide was 5'-labeled and then annealed with a 50-fold amount of unmethylated oligonucleotide. The second probe was prepared by combining a labeled unmethylated oligonucleotide with excess of non-radioactive methylated component.

Results of the hemimethylation assay are presented in Table 3. In general, the hemimethylated duplexes were cleaved more efficiently than the corresponding methylated homoduplexes. Further, in the majority of cases the methylated strands were cut with either identical or lower efficiency, as compared to the non-methylated strands. The cognate methylated strands in hemimethylated duplexes were not cleaved with HpaII and Cfr9I restrictases. Endonuclease MspI proved an exception, cleaving the cognate hemimethylated site; besides, the methylated strand (5mCCGG) was more readily cleaved as compared to the unmethylated strand (Table 3).

DISCUSSION

The data presented above clearly indicate that non-cognate methylation, of the recognition sequence reveals a wide range of cleavage inhibition. For example, the introduction of N4-methylcytosine blocked the action of the majority of endonucleases under investigation (HpaII, SmaI, XmaI and Cfr9I), with the inhibitory effect of 5-methylcytosine significantly lower. Although both methyl groups are exposed in the major groove of B-DNA, the endonucleases strictly differentiated between these methylations, thus displaying different sensitivity to the substrates. Our assumption is that besides sterically obstructing the interaction with enzymes (similar to C5-methyl group) N4-methylation of cytosine destroys the essential C:G base pair recognition contact. Thus interpreted 4mC, i.e. the "blocked cytosine", could be included in the list of already used modified oligonucleotide substrates, designed to investigate DNA-protein interaction.

Our experiments demonstrated the different cleavage abilities of MspI endonuclease. Among them, the introduction of 4mC in any position of the recognition site did not prevent cleavage (Table 2). Most likely R.MspI, in contrast to most investigated endonucleases, has predominant contacts in the minor groove of B-DNA. A similar mode of action in the case of BspRI endonuclease recognizing the GGCC sequence has been reported earlier [15]. Furthermore, R.MspI cleaved the methylated strand better than the unmethylated one in cognate (5mCCGG:CCGG) and non-cognate (4mCCGG:CCGG) hemimethylated duplexes (Table 3). A more direct explanation of this result is that the endonuclease binds the unmethylated strand and cleaves the phosphodiester bond of the methylated strand. Such mode of interaction with the substrates was earlier postulated for some endonucleases recognizing six-base sequences [16].

Among the investigated endonucleases, only MspI [17] and HpaII [10] cognate methylation types were known. The determination of C4mCCGGG specificity for M.Cfr9I (this report) raises a question about the cognate methylation type of other

endonucleases (XmaI and SmaI), recognizing the same nucleotide sequence. However, our attempts to isolate corresponding methylases with the aim of determining their specificity, were unsuccessful. Nevertheless, new data was obtained following analysis of genomic DNAs isolated from the strains producing the corresponding enzymes. The results obtained showed the occurence of two methylated bases, i.e. 5mC and 6mA in the DNA of S.marcescens. Since the SmaI site consists of only cytosine and guanine nucleotides, one might suppose that M.SmaI is a methylase with 5mC specificity. The presence of 6mA is most likely related to dam-like methylation identified in this strain [18,19]. Three methylated bases - 4mC, 5mC and 6mA were detected in the DNA of X.malvacearum (fig.3). Besides, three RM systems were detected in this bacteria [20], yet E.coli-like dam methylation was not found [18,19]. Since only the XmaII site (CTGCAG) contains the adenine base, N6-methyladenine most probably is a product of the associated methylase. Methylase XmaIII was investigated earlier and found to be 5mC (CGG5mCCG) specific [21]. It is likely that XmaI RM system uses N4-methylcytosine for its DNA protection.

Additional information on possible cognate methylations was obtained after the cleavage of methylated substrates. As expected, the respective cognate methylated sites were strictly immune to MspI, HpaII and Cfr9I endonucleases (Table 2). This feature is common in most class II RM enzymes, being predetermined by their biological function. Among 5mC containing sites, only CC5mCGGG was immume to the SmaI endonuclease cleavage (Table 2). This, in junction with the absence of 4mC in S.malvacearum DNA, most likely indicates that the above site represents M.SmaI cognate methylation. Supporting the presumption about the 4mC type of M.XmaI, the associated endonuclease cleaved all possible variants of 5mC containing sites. None of the 4mC containing sites, however, were cleaved with the endonuclease, which prevented the determination of the modified base position.

The presented results were obtained using comparatively short oligonucleotide duplexes as substrates, which may be characterized by a slightly different interaction with enzymes as compared to native DNA [22]. Nevertheless, we assume that our

results reflect general inhibitory effects of the methylation in respect to investigated endonucleases.

The conclusion of the present work is that the cleavage efficiency of non-cognate methylated sites may be rather different as compared to the unmethylated site. In short, a particular methylated site can not be unequivocally attributed to those which are cleaved or those that are not cleaved with the respective endonuclease. At this point we see a drawback in currently available tables [23,24] of restriction enzyme sensitivity to methylated substrates. Currently these tables include contradicting data; some of them are related to the endonucleases under investigation. For example, it is not yet clear whether R.HpaII cleaves the 5mCCGG sequence or not; analogously, the CC5mCGGG site was reported to be cleavable with R.XmaI in DNA of eukaryotes and not cleavable in the DNA of H.parainfluenzae [23,24]. Some of the discrepancies concerning R.MspI sensitivity to methylated sites were elucidated after establishing the inhibitory role of flanking sequences [25,26].

On the basis of the data obtained we can explain the discrepancy in reports on R.HpaII and R.XmaI sensitivity to 5mCCGG and CC5mCGGG sites. As seen from Table 2 the cleavage of the discussed sites is strongly inhibited by C5-methylation. The cleavage efficiency of the sites was five-fold lower as compared to unmethylated substrates. To conclude, we may suppose that enzyme-substrate combinations with the cleavage efficiency ranging from 10 to 30% (Table 2), are potential candidates for erroneous interpretation. In these cases either the absence or presence of cleavage should be dependant on the experimental conditions (incubation time, excess of enzyme, etc.). This should be taken into account when using restriction endonucleases in the analysis of methylated DNA.

REFERENCES

1. Janulaitis, A., Klimašauskas, S., Petrušytė, M. and Butkus, V. (1983) FEBS Lett., 161, 131-134.
2. Ehrlich, M., Gama-Sosa, M.A., Carreira, L.H., Ljungdahl, L.G., Kuo, K.C. and Gehrke, C.W. (1985) Nucleic Acids Res., 13, 1399-1413.
3. Ehrlich, M., Wilson, G.G., Kuo, K.C. and Gehrke, C.W. (1987) J. Bacteriol., 169, 939-943.
4. Butkus, V., Klimašauskas, S., Keršulytė, D., Vaitkevičius, D.,

Lebionka, A. and Janulaitis, A. (1985) Nucleic Acids Res., 13, 5727-5746.
5. Butkus, V., Klimašauskas, S., Petrauskienė, L., Manelienė, Z., Lebionka, A. and Janulaitis, A. (1987) Biochim. Biophys. Acta, in press.
6. Waalwijk, C. and Flavell, R.A. (1978) Nucleic Acids Res., 5, 3231-3236.
7. Bird, A.P., Taggart, M.H. and Smith, B.A. (1979) Cell, 17, 889-901.
8. Janulaitis, A.A., Stakėnas, P.S., Petrušytė, M.P., Bitinaitė, J.B., Klimašauskas, S.J. and Butkus, V.V. (1984) Molekulyarnaja Biologiya (Moscow), 18, 115-129.
9. Endow, S.A. and Roberts, R.J. (1977) J. Mol. Biol., 112, 521-529.
10. Mann, M.B. and Smith, H.O. (1977) Nucleic Acids Res., 4, 4211-4221.
11. Saito, H. and Miura, K.J. (1963) Biochem. Biophys. Acta, 72, 619-629.
12. Petrauskienė, L.J., Klimašauskas, S.J., Butkus, V.V. and Janulaitis, A.A. (1986) Bioorg. Khim.(Moscow) 12, 1597-1603.
13. Gehrke, C.W., McCune, R.A., Gama-Sosa, M.A., Ehrlich, M., Kuo, K.C. (1984) J. Chromatogr., 301, 199-219.
14. Jay, E., Bambara, R., Padmanabhan, R. and Wu, R. (1974) Nucleic Acids Res. 1, 331-353.
15. Wolfes, H., Fliess, A. and Pingoud, A. (1985) Eur. J. Biochem., 150, 105-110.
16. Ono, A., Sato, M. Ohtani, Y. and Ueda, T. (1984) Nucleic Acids Res., 12, 8939-8949.
17. Walder, R.Y., Langtimm, C.J., Chatterjee, R. and Walder, J.A. (1983) J. Biol. Chem., 258, 1235-1241.
18. Barbeyron, T., Kean, K. and Forterre, P. (1984) J. Bacteriol., 160, 586-590.
19. Gingeras, T.R., Blumenthal, R.M., Roberts, R.J. and Brooks, J.E. (1982) In: Metabolism and Enzymology of Nucleic Acids, eds. Zelinka J. and Balan, J., 4, 329-340, Bratislava.
20. Kunkel, L.M., Silberklang, M. and McCarthy, B.J. (1979) J. Mol. Biol., 132, 133-139.
21. Trautner T.A. In: Methylation of DNA, ed. Trautner T.A., Current Topics in Microbiology and Immunology 108, 11-22.
22. Kita, K., Hiraoka, N., Kimizuka, F., Obayashi, A., Kojima, H., Takahashi, H. and Saito, H. (1985) Nucleic Acids Res. 13, 7015-7024.
23. McClelland, M. and Nelson, M. (1985) Nucleic Acids Res., 13, Supplement, r201-r207.
24. Kessler, C., Neumaier, T.S. and Wolf, W. (1985) Gene, 33, 1-102.
25. Busslinger, M., de Boer, E., Wright, S., Grosveld, F.G. and Flavell, R.A. (1983) Nucleic Acids Res., 11, 3559-3569.
26. Keshet, E. and Cedar, H. (1983) Nucleic Acids Res., 11, 3571-3580.