DNA duplexes with reactive dialdehyde groups as novel reagents for cross-linking to restriction–modification enzymes

M. G. Brevnov, O. M. Gritsenko, S. N. Mikhailov¹, E. V. Efimtseva¹, B. S. Ermolinsky¹, A. Van Aerschot², P. Herdewijn², A. V. Repyk³ and E. S. Gromova*

Department of Chemistry, Moscow State University, Moscow 119899, Russia, ¹Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, Vavilov Street 32, Moscow 117984, Russia, ²Rega Institute, Katholieke Universiteit Leuven, Minderbroedersstraat 10, B-3000 Leuven, Belgium and ³Institute of Biochemistry and Physiology of Microorganisms, Russian Academy of Sciences, 142292 Pushchino, Moscow Region, Russia

Received April 4, 1997; Revised and Accepted June 27, 1997

ABSTRACT

To create new, effective reagents for affinity modification of restriction–modification (R-M) enzymes, a regioselective method for reactive dialdehyde group incorporation into oligonucleotides, based on insertion of a 1-β-D-galactopyranosylthymine residue, has been developed. We synthesized DNA duplex analogs of the substrates of the EcoRI and MvaI R-M enzymes that contained a galactose or periodate-oxidized galactose residue as single substituents either in the center of the EcoRI (MvaI) recognition site or in the flanking nucleotide sequence. The dependence of binding, cleavage and methylation of these substrate analogs on the modified sugar location in the duplex was determined. Cross-linking of the reagents to the enzymes under different conditions was examined. M.EcoRI covalent attachment to periodate-oxidized substrate analogs proceeded in a specific way and to a large extent depended on the location of the reactive dialdehyde group in the substrate. The yield of covalent attachment to a DNA duplex with a dialdehyde group in the flanking sequence with EcoRI or MvaI methylases was 9–20% and did not exceed 4% for R.EcoRI.

INTRODUCTION

Type II restriction–modification (R-M) enzymes are attractive models for studying specific DNA–protein recognition. Recently the structures of several complexes of restriction endonucleases and DNA methyltransferases (Mtases) with substrates were solved and DNA–protein contacts were determined (1–3).

Affinity modification of the enzymes by modified substrates followed by proteolysis of a cross-linked product is an alternative promising way to determine the amino acid residues responsible both for specific recognition of the target DNA and for catalytic action. Previously, 5-bromouridine-containing substrates were photocross-linked to EcoRI and EcoRV restriction endonucleases (4). Thio analogs of thymidine and deoxyguanosine were introduced into synthetic substrates of the EcoRI and EcoRV R-M enzymes and these DNA duplexes were converted to photoaffinity reagents (5). DNA duplexes with an active monosubstituted pyrophosphate bond were used for affinity modification of several restriction endonucleases (6–8) and two DNA methyltransferases (6). However, these reagents are supposed to interact with any nucleophilic group of the proteins depending on its proximity to the modified internucleotide bond and formation of an unstable covalent enzyme–DNA bond is likely.

Periodate-oxidized nucleotides have been used extensively for affinity modification of enzymes involved in nucleic acid metabolism (9–10). This labeling technique involves formation of an unstable dihydroxymorpholine derivative, which may be reduced to stable morpholine compounds. Using a similar methodology and starting from 3′-modified oligonucleotides (ONs), conjugates with poly(lysine) were prepared (11).

We describe here a novel type of cross-linking reagent for R-M enzymes, DNA duplexes with reactive dialdehyde groups introduced at single defined positions of the ON strand. The advantages of such reagents are their specificity, mainly toward proximal lysine residues, and their ability to form stable covalent complexes upon reduction of the cross-linked products. To create such reagents, we developed a methodology for regioselective modification of ONs that enables introduction of reactive groups into the middle of an ON chain. Previously the reactive groups were introduced only at the 5′- (12) or 3′-ends (13). Introduction of additional cis diol groups may be achieved in several different ways, the simplest based on insertion of hexopyranosylnucleoside residues into an ON chain. These analogs have four hydroxyl groups, two of which may be used for incorporation into the ON chain and the other two for further modification. It should be noted that several ONs containing 2′,3′-dideoxyhexopyranosyl nucleosides have been prepared (14–16). It was shown that an ON with one or two such residues may form a stable duplex (14,16). It is noteworthy that such DNA duplexes with dialdehyde groups in the middle of the ON strand were not used for covalent attachment to DNA binding proteins. Ebralidse et al. (17) have reported mildly depurinated DNA cross-linking to histones.

In this work efficient synthesis of DNA duplexes with reactive dialdehyde groups is developed, the properties of such duplexes...
are described and cross-linking to EcoRII and MvaI R-M enzymes is investigated.

**MATERIALS AND METHODS**

Liquid secondary ion mass spectra (LSIMS) were obtained using a Kratos Concept 1H mass spectrometer. Column chromatography was performed on silica gel (0.06–0.20 mm), TLC was carried out on Kieselgel 260 F (Merck) with detection by UV light using: A, CHCl3; B, 98:2 CHCl3:EtOH; C, 95:5 CHCl3:EtOH. NMR spectra were recorded on Gemini 200 NMR and Bruker AMX 400 spectrometers at 22°C. Chemical shifts were measured relative to the solvent signals.

Buffers: A, 40 mM Tris–HCl, pH 7.0, 15 mM MgCl2; B, 1 M/3-aniline acetate, pH 4.5; C, 40 mM Tris–HCl, pH 7.6, 50 mM NaCl, 7 mM dithiothreitol (DTT); D, 40 mM Tris–HCl, pH 7.6, 50 mM NaCl, 5 mM MgCl2, 7 mM DTT; E, 40 mM Tris–HCl, pH 7.9, 5 mM DTT, 1 mM EDTA; F, 10 mM Tris–HCl, pH 7.5, 50 mM NaCl, 15 mM MgCl2, 0.1 mM DTT; G, 10 mM Tris–HCl, pH 8.5, 150 mM NaCl, 1 mM DTT, 100 µg/ml BSA; H, 50 mM Tris–HCl, pH 9.0, 20 mM NaCl, 1 mM DTT, 100 µg/ml BSA.

**Enzymes**

Restriction endonuclease MvaI (44 µg/ml, 20 µl) and T4 polynucleotide kinase were purchased from MBI Fermentas (Bmoons). T4 polynucleotide kinase was kept for 48 h at +2°C. Restriction endonuclease EcoRII (640 µg/ml, 80 µl) was purified by N.I.Matvienko (Institute of Protein, Russian Academy of Sciences). DNA methyltransferase EcoRII (840 µg/ml, 5–8 µl) was purified according to Buryanov et al. The solution was saturated with ammonium sulfate and filtered to evaporate the residues. The products were purified by column chromatography on a phenyl gel using solvents A and C to give as a foam (2.5 g, 76%).

**1-(2,3,4,6-O-Tetraacetyl-β-d-galactopyranosyl)thymine**

A suspension of dry thymine (9.45 g, 75 mmol) in hexamethyldisilazane (50 ml) and ammonium sulfate (0.5 g) was boiled under reflux in the absence of moisture until complete dissolution (10 h). The mixture was concentrated in vacuo to dryness and dry 1,2-dichloroethane (30 ml) was evaporated from the residue. A solution of 1,2,3,4,6-penta-O-acetyl-α-d-galactopyranose (27.4 g, 70.26 mmol) in dry 1,2-dichloroethane (200 ml) and SnCl2 (9.6 ml, 80 mmol) was added to the residue and the mixture was stirred for 16 h at 20°C. Chloroform (100 ml) and saturated aqueous sodium hydroxide solution (100 ml) were added, the mixture was stirred for 20 min at 20°C and then filtered through Hyflo SuperCel. The organic layer was separated, washed with aqueous sodium hydroxide solution (30 ml) and water, dried and evaporated. The products were purified by column chromatography on a silica gel column using solvent B to give a foam-like product (28.1 g, 88%).

**1-(β-d-Galactopyranosyl)-thymine (1)**

The solution of 1-(2,3,4,6-O-tetraacetyl-β-d-galactopyranosyl)-thymine (4.56 g, 10 mmol) in methanol (40 ml) semi-saturated with ammonia at 0°C was kept for 20 h at 20°C and then concentrated in vacuo to dryness. The residue was washed with acetone to yield a syrup (2.6 g, 90%). UV (λ = 170): [λ]max 267 nm (ε = 9500) (UV (λ = 170): [λ]max 267 nm (ε = 7000)).

1H NMR (D2O) (400 MHz): 7.68 q (1H, J = 1.2 Hz, H-6), 5.53 d (1H, J1,2 = 9.1 Hz, H-1), 4.02 dd (1H, J3,4 = 3.3 Hz, J4,5 = 1.0 Hz, H-4), 3.88 dd (1H, J3,2 = 9.6 Hz, H-2), 3.97 ddd (1H, J = 6.5 Hz, J3,6,5 = 5.6 Hz, H-6), 3.89 dd (1H, H-3′, 1H, J = 3.7 Hz, H-7′), 1.87 d (3H, Me-5). 1C NMR (D2O): 165.14 (C-6), 151.08 (C-2), 136.40 (C-6), 110.95 (C-5), 81.97 (C-1′), 76.98 (C-5′), 72.10 (C-3′), 67.75 (C-2′), 67.60 (C-4′), 59.88 (C-6′), 10.44 (Me-5).

1-(3,4-O-Isopropylidene-β-d-galactopyranosyl)thymine (2)

A suspension of nucleoside 1 (2.9 g, 10 mmol) in dry acetone (10 ml) and 2,2-dimethoxypropane (30 ml) in the presence of p-toluene sulfonic acid monohydrate (0.57 g, 3 mmol) was stirred at 20°C for 3 days. The mixture was neutralized with saturated aqueous sodium carbonate, filtered and filtrates were evaporated to dryness. The products were purified by column chromatography on a silica gel using solvents A and C to give as a foam (2.5 g, 76%).

**1-(β-d-Galactopyranosyl)-thymine-2′-O-phosphoryl-(2′-5′)-deoxyadenosine (3)**

Dinucleoside monophosphate 3 was prepared by condensation of 1-(3,4-O-isopropylidene-6-O-monomethoxytrityl-β-d-galactopyranosyl)thymine (0.71 mmol) and N′,3′-O-diacetyl-2′-deoxyadenosine 5′-phosphate (1 mmol) in dry pyridine (5 ml) in the presence of NaOCH3 (8 mmol) for 7 days at 20°C followed by removal of acetyl blocking groups (5 M ammonium in methanol for 48 h at 20°C) and the isopropylidene group (80% acetic acid for 16 h at 20°C) and separation on DEAE–cellulose using a NH4HCO3 concentration gradient (0.01–0.3 M). The pooled fractions (eluted at 0.05 M) were evaporated to dryness, co-evaporated with water (5 x 10 ml) and freeze dried. Yield 25%.

**Downloaded from https://academic.oup.com/nar/article/25/16/3302/1458086 by guest on 04 March 2022**
-11.4 Hz, J_{2a,b'} = 6.1 Hz, H-5′a), 3.91 ddd (1H, J_{5a,b'} = 5.1 Hz, H-5′b), 3.88 dt (1H, J_{5′a,b'} = 5.1 Hz, H-5′T), 3.81 dd (1H, H-3′T), 3.76 d (2H, H-6′a and H-6′b T), 2.70 ddd (1H, J_{2a,2b} = -14.0 Hz, H-2′a), 2.57 ddd (1H, H-2′b A), 1.71 q (3H, 5-Me T). The signals were assigned by the double resonance technique and two dimensional COSY method. 31P NMR (161.98 MHz) (D2O) chemical shift in p.p.m. from 80% phosphoric acid: 2.15.

1-(6′-O-Dimethoxytrityl-3′,4′-O-isopropylidene-β-n-galacto-pyranosyl)-thymine

A standard procedure was used to prepare the title compound starting from 2. Yield 86% (foam). LSIMS (Thgly-NaOAc m/e) 653 (M+Na+, 5), 303 (DMTr, 100). 1H NMR (200 MHz) (CDCl3): 8.63 brs (1H, NH), 7.50–7.20 m (10H, arom-H, H-6), 6.84 s (4H, J = 8.8 Hz, arom-H), 5.65 d (1H, J_{1′,2′} = 8.5Hz, H-1′), 4.41–3.70 m (4H, H-2′, 3′, 4′, 5′), 3.81 s (6H, 2×MeO), 3.40 m (2H, J = 6.1 Hz, H-6′a 6′b), 1.90 s (3H, Me, Me), 1.57 s (3H, Me), 1.44 s (3H, Me). 13C NMR (CDCl3): 158.42, 149.66, 144.69, 135.98, 135.86, 135.46, 130.08, 128.15, 127.68, 126.70, 123.71 and 112.99 (Ph), 163.86 (C-4), 151.16 (C-2), 135.98 (C-6), 111.11 (MeO2C), 109.87 (C-5), 86.17 (PhC), 82.67 (C-1′), 78.65 (C-5′), 74.82 (C-3′), 73.41 (C-4′), 71.62 (C-2′), 62.22 (C-6′), 55.14 (2×MeO), 28.03 and 25.82 (Me2E), 12.37 (Me-5).

1-(6′-O-Dimethoxytrityl-3′,4′-O-isopropylidene-2′-O-(P-2-cyanoethyl-N,N-diisopropylaminophosphinyl)-β-n-galacto-pyranosyl)-thymine (4)

This was prepared by a standard procedure. Yield 79% (white powder). Rf (hexane:acetone:TEA-49:49:2): 0.58. LSIMS (NBA m/e) 831 (M+H+, 1), 303 (DMTr, 100). 1H NMR (200 MHz) (CDCl3): 8.85 brs (1H, NH), 7.50–7.15 m (10H, arom-H, H-6), 6.83 d (4H, J = 8.6 Hz, arom-H), 5.64 d (1H, J_{1′,2′} = 8.4Hz, H-1′), 4.41–3.15 m (10H, H-2′, 3′, 4′, 5′, 6′a, 6′b), CH in iPr, POCH2), 3.80 s (6H, 2×MeO), 2.81–2.50 m (2H, J = 6.2 Hz, CH2CN), 1.98 s and 1.95 s (3H, Me, 5′), 1.58 s and 1.56 s (3H, Me), 1.43 s and 1.42 s (3H, Me), 1.29–1.00 m (12H, J = 6.8 Hz, Me in iPr).

13C NMR (CDCl3): 158.32, 144.54, 153.72, 129.95, 128.02, 127.59, 126.63 and 112.89 (Ph), 163.27 and 163.10 (C-4), 150.42 and 150.25 (C-2), 135.05 (C-6), 117.57 (CN), 111.11 and 110.74 (Me2C), 109.72 and 109.63 (C-5), 86.10 (PhC), 81.64 and 81.48 (C-1′), 78.65 and 78.43 (C-3′), 74.78 (C-5′), 74.33 and 73.28 (d, J = 15Hz, C-2′), 73.17 (C-4′), 61.95 (C-6′), 58.09 and 57.72 (d, J = 18.0 Hz, POCH2), 55.09 (2×MeO), 43.27 and 42.90 (d, J = 12.7 Hz, NCH), 27.95 and 25.87 (Me2C), 24.7–24.0 (NCMe), 20.10, 19.98 (CH2CN), 12.29 (Me-5). 31P NMR (80.99 MHz) (CDCl3) (external reference = H3PO4 capil.): 152.6, 151.7.

Oligonucleotide synthesis

Oligonucleotide synthesis was performed on an ABI 392 at 1 μmol scale using commercial 2-cyanoethylphosphoramidites and standard methodology, except for a longer coupling time (80 s) and a higher concentration (0.15 M) for the amide 4 to ensure high coupling yields. The products were deprotected and removed from the solid support by concentrated ammonia (55°C, 16 h) and purified by ion exchange on a MonoQ column using a NaCl gradient (0.3–0.7 M) in 10 mM NaOH and reverse phase HPLC on a Nucleosil 100 C18, 5 μ column (4×250 mm) using a MeCN concentration gradient (10–20% MeCN over 25 min) in 0.1 M triethylammonium acetate, pH 6.9, with subsequent gel filtration of the products on a Toyopearl HW-40 column in water.

A solution of ON (3 OD260) in 80% acetic acid (0.1 ml) was kept at 20°C for 2 h and the mixture separated by HPLC under the above-mentioned conditions. Yields of ON containing Tgal IIa, IIb and IVa (the upper strands of duplexes II–IV; Table 1) were in the range 40–50% (the retention times were 17.3, 17.3 and 18.1 min respectively). The recovered connecting ONs with isopropylidene groups had retention times of 18.5, 18.7 and 23.3 min. 5′-32P-Labeling of the ONs were carried out using T4 polynucleotide kinase and [γ32P]ATP.

Table 1. Properties of Tgal and Tgal*-containing DNA duplexes as substrates of M. EcorII, M. MvaI, R. EcorII and R. MvaI

<table>
<thead>
<tr>
<th>DNA duplex*</th>
<th>T_{mi} (°C)</th>
<th>Relative methylation (%)†</th>
<th>Relative cleavage (%)‡</th>
<th>Relative affinities‡</th>
<th>R. EcorII</th>
<th>M. MvaI</th>
<th>R. EcorII</th>
<th>M. MvaI</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>5′-GCCAACCCTGGCTCT</td>
<td>64</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>3′-CGTTTGGACCAGAGA</td>
<td>100</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>II</td>
<td>5′-GCCAACCGGGGTCTT</td>
<td>52</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3′-CGTTTGGACCAGAGA</td>
<td>0</td>
<td>30</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>III</td>
<td>5′-GCCAACCTGGCTCT</td>
<td>55</td>
<td>15</td>
<td>30</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3′-CGTTTGGACCAGAGA</td>
<td>0</td>
<td>90</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IV</td>
<td>5′-GCCAACCCTGGXCT</td>
<td>58</td>
<td>75</td>
<td>70</td>
<td>80</td>
<td>100</td>
<td>12</td>
<td>70</td>
</tr>
<tr>
<td>3′-CGTTTGGACCAGAGA</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>II</td>
<td>5′-GCCAACCGGYCTT</td>
<td>22</td>
<td>23</td>
<td>23</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3′-CGTTTGGACCAGAGA</td>
<td>0</td>
<td>30</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>III</td>
<td>5′-GCCACCGGGGTCTT</td>
<td>52</td>
<td>41</td>
<td>0</td>
<td>100</td>
<td>1</td>
<td>1</td>
<td>0.8</td>
</tr>
<tr>
<td>3′-CGTTTGGACCAGAGA</td>
<td>0</td>
<td>80</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Abb. Tgal; Y, Tgal*.

†Relative methylation was defined as the ratio of 3H radioactivity incorporation in DNA duplexes II–IV* to 3H radioactivity incorporation in DNA duplex I. For conditions of methylation see Materials and Methods.

‡For conditions see Materials and Methods. The data relate to cleavage of the upper and the lower strands of I–IV*.

§Relative affinities were defined as normalized values of the efficiencies of non-covalent complexes [M.EcorII(M.MvaI)–AdoHcy–DNA duplex(II*–IV*)] or [R.EcorII(R.MvaI)–DNA duplex (II*–IV*)] formation as described in Materials and Methods.
Enzymatic degradation of oligonucleotides

An oligonucleotide solution (0.2 OD) in 400 µl 50 mM Tris–HCl buffer, pH 8.6, 50 mM sodium chloride, 7 mM MgCl₂ was digested with 1.0 U snake venom phosphodiesterase (Pharmacia) at 37°C for 16 h. This digestion was followed by treatment with 1.0 U alkaline phosphatase (Boehringer Mannheim) for 4 h at 37°C. Aliquots of this solution (50 µl) were analyzed by HPLC on a Nucleosil 100 C18, 5 µ column (4 × 250 mm). Separation of the isopropylidene containing the ON digest was carried out in 0.1 M triethylammonium acetate, pH 6.9, 8% MeCN (flow rate 1 ml/min). The retention time was 7.6 min for 2 and 3,0, 3.9, 4.3, 6.5 min for dC, dG, T and dA.

Separation of unprotected ON digest was performed on the same column using a MeCN concentration gradient (0–10% of MeCN over 25 min, flow rate 1 ml/min) in 0.1 M triethylammonium acetate, pH 6.9. The retention time was 4.4 min for I and 7.2, 13.5, 14.3, 19.7 min for dC, dG, T and dA. All oligonucleotides showed a correct ratio of unmodified nucleosides over the modified one. The ON structure was also proved by Maxam–Gilbert sequencing (cleavage at Tgal residues proceeds more slowly than at thymidine).

Oxidation of oligonucleotides containing galactopyranose residues

Each [32P]ON (20–50 pmol) was dissolved in 40–100 µl 50 mM NaIO₄ and incubated at 37°C for 1.5–2.5 h. To stop the reaction, ON was precipitated by adding 100 µl 2 M LiClO₄ and 1 ml acetone (excess NaIO₄ and the products of its reduction were removed). The yield of oxidation was 65–75% as estimated by β-elimination in buffer B at 95°C for 15 min and subsequent electrophoresis on a 20% polyacrylamide gel containing 7 M urea.

Determination of Tₘ values

The thermal melting curves were obtained in buffer A using a Hitachi 150-20 spectrophotometer (Japan); concentration per duplex ([cₐ]) 1 µM. To determine the Tₘ values, the first derivative was calculated.

Gel mobility shift assay

M.EcoRII (3.16 µM), M.Mval (2.47 µM), R.EcoRII (1.45 µM) or R.Mval (0.5 µM) were incubated with [32P]-labeled duplexes ([cₐ] 0.35 µM) in 10 µl buffers E, H, C or G respectively, containing 8% glycerol and 0.1 mM AdoHcy in the case of Mtases, at room temperature for 5 min and at 0°C for 15 min. The reactions were run for 1.5–2 h at 120 V on native 8% polyacrylamide gels (the gels were pre-run for 1 h at 100 V). For autoradiography of the electrophoretic pattern, Kodak-XOMAT-S film was exposed with an intensifier screen at −20°C overnight. The efficiency of non-covalent complex formation (R) was determined as the ratio [c.p.m. bound DNA duplex/c.p.m. free DNA duplex]. To normalize data from different experiments, the ratios of R for modified substrates II*-IV* to R for canonical substrate I in each experimental data set were defined. The dependences of the non-covalent complex yields on the enzyme concentration were obtained using 0.23 µM DNA duplexes and 0.35–4.74 µM M.EcoRII, 0.35–4.95 µM M.Mval or 0.36–4.35 µM R.EcoRII. These data were normalized by dividing the yields of non-covalent complexes of canonical substrate I by the appropriate enzyme in each dataset [cₐ(I)] 0.23 µM; c(M.EcoRII) 4.95 µM; c(M.Mval) 4.35 µM).

Methylation assay

The efficiency of methylation was monitored by incorporation of radioactivity ([3H]Ado) into DNA duplexes I–IV*. Methylation reactions were carried out at 20°C for 30 min in 10 µl reaction mixtures containing buffer E (for M.EcoRII) or H (for M.Mval), 1 µCi [methyl-3H]AdoMet (15 Ci/mmol; Amersham) and 0.35 µM DNA duplexes I–IV*. Final concentrations of EcoRII and Mval Mtases were 1.58 and 0.82 µM respectively. Samples were spotted on DE 81 filters (2.5 cm; Whatman), washed with 50 mM KH₂PO₄ (5 × 150 ml), dried and counted in a liquid scintillation spectrometer. For each series blank values (without enzyme) were treated like the samples (with the complete washing procedure).

Cleavage assay

Enzymatic cleavage was performed by incubating 0.35 µM [32P]-labeled substrate analogs with R.EcoRII (1.45 µM) in 10 µl buffers D or with R.Mval (5.45 µM) enzyme in 10 µl buffer F at 20°C for 1 h. Enzymatic reactions were stopped by heating at 95°C for 2–4 min. The products of cleavage of [32P]-labeled DNA duplexes were analyzed in 20% polyacrylamide gels containing 7 M urea as described (19).

Cross-linking of EcoRII and Mval R-M enzymes to DNA duplexes

Cross-linking of M.EcoRII (0.35–4.74 µM), M.Mval (0.35–4.95 µM) or R.EcoRII (0.36–4.35 µM) to duplexes II*-IV* (0.23 µM) was performed in 10 µl buffers E, H and C respectively, containing 0.1 mM AdoHcy at room temperature for 5 min and at 0°C for 15 min. Aliquots of 6 µl 10 mM NaBH₄ was added and reaction mixtures were kept at 0°C for 40 min. Reactions were followed by 0.1% SDS–8% PAGE after heating the samples in 0.1% SDS at 95°C. The gels were analyzed by autoradiography and staining with Coomassie blue. Coexistence of DNA and protein in the same band proved formation of the DNA–enzyme crosslinker complex. The cross-linking yield was determined as the ratio of the covalent conjugate radioactivity to total radioactivity of the conjugate and unbound DNA. Competitive inhibition of M.EcoRII (3.16 µM) and M.Mval (3.3 µM) cross-linking to DNA duplexes II* and IV* (0.23 µM) was studied in the presence of increasing amounts of unlabeled DNA duplexes V (0, 0.46 and 4.6 µM) and VI (0 and 2.3 µM). To determine the influence of Mg²⁺ ions on R.EcoRII cross-linking to duplex IV*, MgCl₂ was added to final concentrations of 5, 15 and 25 mM.

RESULTS AND DISCUSSION

Design of substrate analogs for cross-linking to EcoRII and Mval R-M enzymes

EcoRII and Mval R-M enzymes recognize the DNA sequence 5’-CC[CG/CG]GGG-3’ (20). EcoRII and Mval restriction endonucleases (R.EcoRII and R.Mval) cleave DNA as shown by the arrows (|R.EcoRII; |, R.Mval). The cognate DNA Mtases (M.EcoRII and M.Mval) methylate the ‘inner’ cytosine residues at carbon C5 (M.EcoRII) or at nitrogen N4 (M.Mval). The enzymes require different cofactors. For Mtases it is S-adenosyl-L-methionine (AdoMet), which is converted to S-adenosyl-L-homocysteine (AdoHcy), whereas Mg²⁺ ions are necessary for the endonucleases.

To regiospecifically introduce the reactive dialdehyde groups into the EcoRII (Mval) recognition site, an ON thymidine residue
The presence of β-idene–O-An efficient preparative four step synthesis of 1-(3,4-·3′-end by one nucleotide residue (IV and IV*).

**Scheme 1.**

was replaced by 1-(β-D-galactopyranosyl)thymine (Tgal) as proposed in Scheme 1. This analog retains all the functions of a natural nucleoside crucial for binding to the enzymes. It also contains a cis diol group. Thus, after periodate oxidation one can obtain ONs with Tgal* (Scheme 1) where dialdehyde groups are in any desired position on the ON strand. 14mer DNA duplex I (Table 1) is a canonical substrate of EcoRII and Mval R-M enzymes (19). 14mer DNA duplexes with Tgal (II–IV) and their periodate-oxidized analogs with Tgal* (II*–IV*) are suggested as substrate analogs of the enzymes (Table 1). A modified sugar residue is introduced into the center of the EcoRII (Mval) recognition site (II and II*) or into the 5′-end flanking nucleotide sequence adjacent to the recognition site (III and III*) or separated from its 3′-end. 14mer DNA duplexes with Tgal* (Scheme 1) in high overall yield. We have examined several possibilities for the synthesis of 3,4-O-blocked 1-(β-D-galactopyranosyl)thymine and found the simplest and most effective to be that using the isopropylidene group.

Thus, reaction of 1 with 2,2-dimethoxypropane in acetone in the presence of p-toluenesulfonic acid for 3 days at 20 °C gave 1-3,4-O-isopropylidene-β-D-galactopyranosyl)thymine (2) in 50% yield. To investigate the periodate oxidation reaction in detail, a model dinucleoside phosphate, namely 1-(β-D-galactopyranosyl)-thymine-2′-O-phosphoryl-(2′-5′)-2′-deoxyadenosine (3) was prepared. Oxidation of the cis diol groups using a 10-fold excess of sodium periodate was complete in 1.5 h at 20 °C. The reaction was monitored by HPLC on a Nucleosil C18 column using a concentration gradient of MeCN (5–15%) in 0.1 M triethylammonium acetate, pH 6.9. The oxidized product was stable in this buffer for at least 1 day at room temperature. Successful synthesis of this dimer and the results of its oxidation with periodate encouraged us to use phosphoramidite 4 for ON synthesis. Assembly of ONs was in an ABI 392 at 1 µmol scale using standard methodology, except that a longer coupling time (80 s) and a higher concentration of amidite 4 (0.15 M) were used to ensure high coupling yields. The isopropylidene group was removed with 80% acetic acid (2 h at room temperature). During this time 40–50% of isopropylidene groups were cleaved. The products and starting ONs may be cleanly separated by HPLC and ONs with an isopropylidene group may be easily recovered and used for further deblocking. Prolonged treatment resulted in formation of a complex mixture of products, evidently due to partial depurination.

To be sure that no random depurination took place during removal of the isopropylidene blocking group, unprotected ONs were incubated in buffer B for 15 min at 95 °C. These conditions for RNA sequence determination favor selective cleavage of internucleotide bonds at abasic sites (22). We show that 32P-labeled model ONs were stable after such treatment (Fig. 1). The galactose-containing ONs were also stable under conditions of 10% piperidine treatment at 95 °C for 25 min. A set of ONs containing Tgal IIa–IVa (the upper strands of duplexes II–IV; Table 1) were prepared. Oxidation of the ON (Scheme 1) was carried out in 20–200 mM NaIO₄ solution for 1.5–2.5 h at 25 and 37 °C. To estimate the yields of oxidation products, they were cleaved in 10% piperidine at 95 °C for 25 min at the phosphodiester bond, 5′-terminal relative to the oxidized galactose residues. The yields were 65–75%. The structures of Tgal*-containing ONs IIa*–IVa* (the upper strands of duplexes II*–IV*; Table 1) were confirmed by selective cleavage at the phosphodiester bond neighboring the oxidized galactose residues (Fig. 1). The reaction was carried out

**Figure 1.** Stability analysis of the Tgal- and Tgal*-containing oligonucleotides (ON). (a) Lanes 1 and 2, ON IIa; lanes 3 and 4, ON IIIa; lanes 5 and 6, ON IVa; lanes 1, 3 and 5, incubation in buffer B, 95 °C, 15 min; lanes 2, 4 and 6, incubation in buffer E, 95 °C, 25 min. (b) Lanes 1, ON III*a; lane 2, ON II*a; lane 3, ON IV*a; lanes 1, 2 and 3, incubation in buffer B, 95 °C, 15 min.

**Synthesis of oligonucleotides with Tgal and Tgal* and their properties**

An efficient preparative four step synthesis of 1-(3,4-O-isopropylidene-β-D-galactopyranosyl)thymine was developed starting from β-D-galactopyranose. Fully acetylated β-D-galactopyranose was condensed with bis-trimethylsilylthymine under V orbruggen’s conditions (21) to give, after deacetylation, 1-((β-D-galactopyranosyl)thymine (Tgal) in 50% yield. To investigate the periodate oxidation reaction in detail, a model dinucleoside phosphate, namely 1-(β-D-galactopyranosyl)thymine-2′-O-phosphoryl-(2′-5′)-2′-deoxyadenosine (3) was prepared. Oxidation of the cis diol groups using a 10-fold excess of sodium periodate was complete in 1.5 h at 20 °C. The reaction was monitored by HPLC on a Nucleosil C18 column using a concentration gradient of MeCN (5–15%) in 0.1 M triethylammonium acetate, pH 6.9. The oxidized product was stable in this buffer for at least 1 day at room temperature. Successful synthesis of this dimer and the results of its oxidation with periodate encouraged us to use phosphoramidite 4 for ON synthesis. Assembly of ONs was in an ABI 392 at 1 µmol scale using standard methodology, except that a longer coupling time (80 s) and a higher concentration of amidite 4 (0.15 M) were used to ensure high coupling yields. The isopropylidene group was removed with 80% acetic acid (2 h at room temperature). During this time 40–50% of isopropylidene groups were cleaved. The products and starting ONs may be cleanly separated by HPLC and ONs with an isopropylidene group may be easily recovered and used for further deblocking. Prolonged treatment resulted in formation of a complex mixture of products, evidently due to partial depurination.

To be sure that no random depurination took place during removal of the isopropylidene blocking group, unprotected ONs were incubated in buffer B for 15 min at 95 °C. These conditions for RNA sequence determination favor selective cleavage of internucleotide bonds at abasic sites (22). We show that 32P-labeled model ONs were stable after such treatment (Fig. 1). The galactose-containing ONs were also stable under conditions of 10% piperidine treatment at 95 °C for 25 min. A set of ONs containing Tgal IIa–IVa (the upper strands of duplexes II–IV; Table 1) were prepared. Oxidation of the ON (Scheme 1) was carried out in 20–200 mM NaIO₄ solution for 1.5–2.5 h at 25 and 37 °C. To estimate the yields of oxidation products, they were cleaved in 10% piperidine at 95 °C for 25 min at the phosphodiester bond, 5′-terminal relative to the oxidized galactose residues. The yields were 65–75%. The structures of Tgal*-containing ONs IIa*–IVa* (the upper strands of duplexes II*–IV*; Table 1) were confirmed by selective cleavage at the phosphodiester bond neighboring the oxidized galactose residues (Fig. 1). The reaction was carried out
in buffer B for 15 min at 95°C or in 10% piperidine for 25 min at 95°C. In contrast to Tgal compounds, their cleavage produced corresponding 32P-labeled peptidyl nucleotides, tetranucleotides and undecanucleotides with 3'-phosphate groups (Fig. 1). In control experiments we observed slight decomposition of the periodate-oxidized oligonucleotides during gel electrophoresis.

**Thermal stability of modified duplexes**

Melting temperatures \(T_m\) of DNA duplexes II–IV with a single galactopyranose residue were determined in buffer A (Table 1). Insertion of the modified sugar moiety resulted in a destabilization of 6–12°C compared with the regular duplex I. We could not determine the \(T_m\) of DNA duplexes II*–IV* due to their rapid decomposition. Incubation of duplexes with Tgal* under melting conditions (a temperature change from 15 to 80°C, buffer A) resulted in ~50% breaking of the upper strands at the modified sugar residue. It was shown previously that in the case of a 14mer DNA duplex with the dT residue of the recognition site replaced by Tgal the \(T_m\) was 14°C below that of a regular duplex. Hence, one can expect a similar destabilization effect for duplexes II*–IV*. Also, drastic double helix distortions are unlikely, since modified DNAs are still substrates of highly sequence-specific enzymes (see below). All cross-linking experiments were done under conditions favorable for thermodynamic stability of the periodate-oxidized duplexes.

**Interaction of EcoRII and MvaI Mtases with substrate analogs containing Tgal and Tgal***

**Specific binding.** The use of DNA duplexes with reactive dialdehyde groups as reagents for cross-linking to R-M enzymes requires evaluation of both their ability to specifically bind to the enzymes and their substrate properties. To obtain correct complexes and to eliminate methylation, EcoRII and MvaI binding to the reagents was studied in the presence of the cofactor analog AdoHcy. In addition, we took into account that complex formation between M.EcoRII and the 14mer regular substrate was substantially reduced if no cofactor was added (24).

-Ternary complex (Mtase–reagent–AdoHcy) formation was monitored in gel mobility shift assays. An increase in binding is observed when the enzyme concentration is increased (data not shown). For both Mtases very poor binding occurred with reagent III* (Table 1). M.EcoRII and M.MvaI had high and approximately equal affinity for DNA duplexes II* (Tgal* is in the center of the recognition site) and IV* (Tgal* is in the flanking nucleotide sequence) (Table 1).

**Methylation.** Insertion of Tgal into DNA duplex I reduced or even eliminated the activity of EcoRII and MvaI Mtases. For both enzymes the substrate properties of Tgal-containing DNA duplexes strongly depended on the location of the galactopyranose residue (Table 1). DNA duplexes with the modified sugar residue in the 5'-end flanking nucleotide sequence adjacent to the recognition site (III) or separated from its 3'-end by one nucleotide residue (IV) retain their ability to be methylated. DNA duplex II with Tgal in the center of the recognition site was not methylated.

-Methylation of DNA duplexes with periodate-oxidized galactose residues (II*–IV*) proceeded similarly to galactose-containing duplexes II–IV (Table 1). We used all three reagents II*–IV* for affinity modification of EcoRII and MvaI Mtases, taking into account that duplex II* with Tgal in the center of the recognition site was able to bind to the enzymes but was not methylated.

**Cross-linking.** The reagents with dialdehyde groups likely react primarily with proximal lysine residues according to Scheme 2.

-Formation of either the Schiff base or a cyclic derivative is possible. We believe that reaction pathway (1) is less probable since modified DNAs are still substrates of highly sequence-specific enzymes (see below). All cross-linking experiments were done under conditions favorable for thermodynamic stability of the periodate-oxidized duplexes.
Covalent complexes were separated from the unreacted reagent by SDS–8% PAGE.

To find optimal conditions for cross-linking, the interaction of M.EcoRII with DNA duplex IV* was studied in detail. To increase the yield of the covalent complex, the M.EcoRII concentration was varied (Fig. 2). Ability of the enzyme to form a covalent complex correlated with its binding capacity. Excess enzyme over reagent was used in all further experiments. We found that the optimal concentration of NaBH₄ was 3.75 mM. Varying the reduction time from 20 to 40 min resulted in an increased yield of the covalent complex. A further increase in the reduction time to 2 h had no effect on yield. No cross-linking occurred with the non-oxidized DNA duplexes. The cross-linking yield was extremely low in the absence of AdoHcy at the concentrations of M.EcoRII and duplex IV* indicated in Materials and Methods. In the absence of NaBH₄, cross-linking products were easily destroyed (data not shown).

Table 2 summarizes the M.EcoRII cross-linking results under optimal conditions. There is a remarkable difference in the cross-linking behavior of reagents II*–IV* depending on the reactive dialdehyde group location. The highest yield was observed with substrate analog IV*. DNA duplex II*, which is able to bind to M.EcoRII but is not methylated, was successfully cross-linked to M.EcoRII, though somewhat less well than duplex IV*. No covalent attachment of the enzyme to duplex III* was detected (data not shown).

As to M.MvaI, covalent attachment of DNA duplexes II* and IV* was also observed under conditions of M.EcoRII cross-linking, but in a different incubation buffer (Table 2). One can see that the cross-linking yields were lower than those for M.EcoRII and no difference in reactivity was found between reagents II* and IV*. The concentration of M.MvaI was varied as well. The same correlation between binding affinity and cross-linking efficiency was observed as for M.EcoRII.

The specificity of M.EcoRII and M.MvaI cross-linking to DNA duplexes II* and IV* was examined in competition experiments with 14mer unmodified cognate (V) and non-cognate (VI) DNA duplexes (Table 2). The 20-fold excess of substrate V inhibited cross-linking of reagents II* and IV* to both enzymes. The 10-fold excess of DNA duplex VI, lacking the recognition site, did not alter the ability of duplexes II* or IV* to modify M.EcoRII and resulted in an ~2-fold decrease in the yields of M.MvaI covalent complexes with these reagents. These data imply a specific covalent attachment of M.EcoRII to DNA duplexes II* and IV*. Taking into account that II* failed to be methylated, one may suppose that structural rearrangement of the enzyme–substrate complex has to occur before the methylation step and such an altered conformation is forbidden in the case of the complex of M.EcoRII with reagent II*.

As for M.MvaI, the pronounced inhibition produced by substrate V also confirmed that the reaction was specific. A low inhibitory effect of duplex VI, which was not an MvaI substrate, was probably due to the appreciable affinity of M.MvaI for non-cognate DNA.

### Interaction of EcoRII and MvaI endonucleases with substrate analogs containing Tgal and Tgal*

Specific binding. Binding of periodate-oxidized substrate analogs II*–IV* to EcoRII and MvaI endonucleases was measured in the absence of Mg²⁺ ions (Table 1). At the concentrations of endonucleases and reagents indicated in Materials and Methods no complexes with reagent II* were detected; M.MvaI does not bind to III* either. Tgal* substitution in the flanking nucleotide sequence one nucleotide away from the recognition site (duplex IV*) had little effect on binding to R.EcoRII or R.MvaI. The R.MvaI complexes migrated faster than those of R.EcoRII, as was expected from the Mr of the proteins (28 000 and 44 000 respectively) (data not shown).

Cleavage. The substrate properties of DNA duplexes II–IV and II*–IV* are markedly dependent on the enzyme and the modified sugar residue (Table 1). Only DNA duplex IV was a substrate of R.EcoRII. It is consistent with our previous observations that R.EcoRII is very sensitive to different types of modifications in the substrate (19). It is also evident that R.EcoRII cannot cleave the 2',5'-phosphodiester bond in III or III*.

### Table 2. Cross-linking of M.EcoRII and M.MvaI to DNA duplexes II* and IV*: competition inhibition of cross-linking by cognate and non-cognate DNA duplexes

<table>
<thead>
<tr>
<th>DNA duplex</th>
<th>M.EcoRII</th>
<th>M.MvaI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cross-linking yield (%)</td>
<td>Cross-linking in presence of inhibitor</td>
<td>Cross-linking in presence of inhibitor</td>
</tr>
<tr>
<td></td>
<td>Inhibitor</td>
<td>Ratio [inhibitor]/ [reagent]</td>
</tr>
<tr>
<td>II*</td>
<td>14</td>
<td>V</td>
</tr>
<tr>
<td></td>
<td></td>
<td>VI</td>
</tr>
<tr>
<td>IV*</td>
<td>20</td>
<td>V</td>
</tr>
<tr>
<td></td>
<td></td>
<td>VI</td>
</tr>
</tbody>
</table>

*Cross-linking conditions and the determination of the product yields are described in Materials and Methods.

**5'-ACCTAAGGAGTGGT (V), 5'-GCCAACCGGCTCTA (VI) 3'-TGGATACCCGACCA 3'-CGGTTGCGCAGAT**

Relative cross-linking yields were determined as the ratio of the cross-linking yields in the presence of different amounts of the inhibitor to the cross-linking yield in the absence of the inhibitor.
All galactopyranose and periodate-oxidized galactopyranose-containing DNA duplexes were substrates for R.MvaI (Table 1). Both strands of DNA duplexes IV and IV* were cleaved. DNA duplexes II, II*, III and III* were cleaved by R.MvaI only in the non-modified strand, which was characteristic also of the other substrate analogs tested with R.MvaI (19).

It is noteworthy that in parallel with the enzyme digestion, partial destruction of the reagents bearing the dialdehyde group may occur in the cleavage buffers (see above). As a result, traces of 5′-terminal 32P-labeled fragments of modified strands of II*–IV* were seen on the gels.

Cross-linking. DNA duplex IV*, which can bind specifically to EcoRII-R enzymes, covalently joins to this endonuclease in the absence of Mg2+ (Fig. 3). As the enzyme concentration was increased from 0.36 to 2.9 µM, the yield of covalent complex also increased. The maximum yield of covalent complex did not exceed 4%, which was considerably less than in the case of the Mtases.

Elevation of the Mg2+ concentration from 0 to 5 × 10–3 M had no effect on reagent IV* covalent joining to EcoRII. One may suggest that reagent IV* does not join to the catalytic site of the enzyme.

Preliminary data on R.MvaI showed that no cross-linking of R.MvaI to duplexes II* and III* occurred under the conditions indicated in Materials and Methods.

CONCLUSIONS

Periodate-oxidized DNA duplex analogs of substrates of type II restriction endonucleases and DNA methyltransferases that contain an active dialdehyde group located as a single substituent were tested. The activity of these enzymes, at the level of the Mtases, was considerably less than in the case of the Mtases. The efficiency of obtained proved to be efficient for affinity modification of these enzymes, especially in the case of the Mtases. The efficiency of cross-linking was dependent on the active group location and correlated with enzyme binding affinity for the reagents. An advantage of the reagents was their ability to specifically react, mainly with lysines of the enzymes, under conditions of lysine proximity to the reactive dialdehyde groups to form a stable covalent bond between the ON and the protein. To understand the functional meaning of the cross-linking dependence on the reactive group location, it is necessary to determine the amino acid residues responsible for binding to the reagent. Such investigations are now in progress. These data will be important for the determination of the active site structure and the mechanism of catalysis of EcoRII and MvaI R-M enzymes.

ACKNOWLEDGEMENTS

This research was supported by an award from the Howard Hughes Medical Institute (grant HHMI 75195-54501), INTAS (project 93-1500), NATO (grant HTECH.LG 961324) and the Russian Foundation for Basic Research (projects 96-04-49178 and 95-04-12283A). A.V. A. is a research associate of the Belgian NSF.

REFERENCES