Role of the 2-Amino Group of Deoxyguanosine in Sequence Recognition by EcoRI Restriction and Modification Enzymes*

(Received for publication, May 31, 1977)

PAUL MODRICH AND ROBERT A. RUBIN

From the Department of Biochemistry, Duke University Medical Center, Durham, North Carolina 27710

The dG residues within the EcoRI recognition sequence of ColEl DNA have been selectively replaced with dI. Methylation of the altered sequence by the EcoRI modification enzyme is extremely slow as compared with methyl transfer to the natural recognition site. Since the affinity of the modification enzyme for the dI-containing sequence is considerably less than that for the natural sequence, we have concluded that the 2-amino group of dG has an important role in DNA site recognition by this enzyme. In contrast, the altered site is subject to cleavage by EcoRI endonuclease at rates essentially identical with those observed with the natural sequence. These results strongly suggest that the two enzymes utilize different contacts within the EcoRI site and are consistent with our conclusion (Rubin, R. A., and Modrich, P. (1977) J. Biol. Chem. 252, 7265-7272) that the two proteins interact with their common recognition sequence in different ways.

As summarized in the previous paper (1), EcoRI endonuclease and methylase are proteins of quite distinct physical properties which interact with a common, 2-fold symmetric hexanucleotide sequence in duplex DNA (2, 3). However, the molecular mechanisms involved in specific sequence recognition by these proteins are not known. One approach to this problem is to study the effects of nucleotide analog substitution within the recognition sequence. Since a variety of nucleotide analogs are available which form good Watson-Crick base pairs (4), their utilization may permit identification of DNA functional groups involved in sequence recognition. This method has been employed previously in studies with T7 RNA polymerase (5). In this case, however, analog substitutions were extensive and thus may have resulted in significant alteration of general polynucleotide structure.

In this paper we describe the highly selective substitution of deoxyinosine (dI) for the deoxyguanosine (dG) residues present in the EcoRI recognition sequence of ColEl DNA, a change which results in loss of the 2-amino group of guanine exposed in the minor groove of the DNA helix. This work has demonstrated that although this amino group is not required for sequence recognition and cleavage by the endonuclease, it does have an important role in recognition by the methylase.

EXPERIMENTAL PROCEDURES

Materials

Enzymes—EcoRI endonuclease and methylase were homogeneous preparations described previously (1, 6). T4 DNA polymerase, isolated according to Goulian et al. (7), was further purified by gel filtration according to England (8), except that Bio-Gel P-150 was used. This step removed contaminating DNA ligase activity which interfered with the terminal labeling reaction catalyzed by the enzyme and yielded a preparation virtually free of contaminating exonuclease-like activities. T4 DNA ligase, Fraction VII of Weiss et al. (9), was generously provided by Dr. C. C. Richardson (Harvard University). Staphylococcus aureus nuclease and spleen phosphodiesterase were from Worthington.

Nucleotides—[γ-32P]dATP (Amersham) and [γ-32P]dATP, 3'-deoxythymidine 5'-monophosphate, and 3'-deoxyinosine 5'-monophosphate were used as purchased. All other nucleotides were obtained from Sigma Chemical Company. DNA was labeled with [32P]dATP by nick translation using T4 DNA ligase, Fraction VII (9). [32P]dGTP was isolated by paper chromatography in solvent I. All samples contained less than 1% of the added [32P]dATP, and the [32P]dGTP was isolated by paper chromatography in solvent I.

Methods

Published in 1977 by Elsevier Science Inc. Printed in U.S.A.
tained 0.067 M Tris/HCl (pH 8.0), 0.033 M NaCl, 6.7 mM MgCl2, 6.7 mM 2-mercaptoethanol, 140 to 170 μg/ml of EcoRI-cleaved ColEl DNA, [α-32P]dITP or [α-32P]dGTP (3000 to 8000 cpm/pmol), prepared by dilution of α-32P-labeled triphosphates with an excess of the unlabeled component so that specific activities could be accurately determined, and 20 μg/ml of T4 DNA polymerase. After 30 min at 12°, reactions were terminated by addition of an equal volume of ice-cold 0.02 M Tris/HCl (pH 7.6), 0.12 M EDTA. After extraction twice with redistilled phenol (6), reactions were dialyzed versus 100 volumes of 0.02 M Tris/HCl (pH 7.6), 1 M NaCl, 1 mM EDTA (two changes, 16 h/change) followed by dialysis versus two changes of the same buffer containing only 0.05 M NaCl. Extent of terminal substitution was calculated on the basis of 32P present in dialyzed material. When carried through an identical procedure, control reactions lacking DNA polymerase contained less than 1% of the 32P present in complete reactions.

**Ligation Reactions**—Terminal substituted molecules were converted to the covalently closed circular form in reactions containing 0.066 M Tris/HCl (pH 7.6), 6.5 mM MgCl2, 55 μM ATP, 45 μg/ml of bovine serum albumin, 9 mM dithiothreitol, 1.5 to 2 μg/ml of DNA, and 0.04 unit/ml of T4 DNA ligase. After 20 to 40 h at 0° reactions were concentrated by step elution from small hydroxyapatite columns. Reactions (40 to 100 ml) were applied at 4° to 2-m columns of hydroxyapatite equilibrated with 0.05 M NaPO4 (pH 6.8). After washing with 4 ml of this buffer, columns were step-eluted at 23° with 0.5 M NaPO4 (pH 6.8). 32P-containing fractions were pooled and EDTA was added to 0.02 M. The covalently closed circular form was then isolated by CsCl/ethidium bromide equilibrium centrifugation (13). The resulting preparations were at least 95% covalently closed and were comprised essentially of monomer circles. Yields of closed circles in the ligation reaction were 40 to 50% as compared to 95% for unsubstituted linear molecules. Overall yields after equilibration centrifugation were 10 to 20%.

In some experiments, two independent preparations of terminally substituted, ligated molecules were employed. Preparation 1 was prepared using ColEl DNA uniformly labeled with 3H-thymidine (6), while Preparation 2 was derived from unlabeled DNA. Because the uniform 3H label of Preparation 1 obscured 3H transfer from S-adenosyl-L-[methyl-3H-jmethylene, reactions with EcoRI methylase employed only Preparation 2.

**Nearest Neighbor Analyses**—Digestion of DNA with S. aureus nuclease and spleen phosphodiesterase was performed according to JENSEN et al. (14). 3′-Mononucleotides were separated by descending chromatography on Whatman No. 1 using Solvent System II which was composed of 0.1 M NaPO4 (pH 6.8), 100 ml, (NH4)2SO4 (60 g), n-propanol (2 ml/100 ml of the NaPO4, (NH4)2SO4 solution). 3′-dNMP 32P cpm

### Table 1

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Measured stoichiometry</th>
<th>Nearest neighbor analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>dA 2.84</td>
<td>dG 1.520</td>
</tr>
<tr>
<td>2</td>
<td>dA 3.00</td>
<td>dG 1.700</td>
</tr>
</tbody>
</table>

in 3′-dAMP and 3′-dGMP, with label in the adenine nucleotide being 1.6 to 1.7 times that found in 3′-dGMP. Recovery of label in dGMP demonstrates that dG is in fact present as a penultimate nucleotide. Furthermore, the recovery of approximately twice as much 32P in 3′-dAMP as in 3′-dGMP indicates that dA is present on the 5′ side of dG at both ends of the linear molecule. The only sequence for ColEl EcoRI site consistent with these results is

\[
\begin{align*}
\text{Experiment 1} & : \text{d(A-G-G-A-A-T-T-C-T)} \\
\text{Experiment 2} & : \text{d(T-C-C-T-T-A-A-G-A)}
\end{align*}
\]

where sites of cleavage and methylation by EcoRI enzymes are indicated by arrows and asterisks, respectively (2, 3).

**Construction and Characterization of dL Substituted Molecules**—The method employed for selective substitution of dL within the EcoRI site of ColEl DNA is summarized in Fig. 1. Control molecules were prepared in an identical fashion except [α-32P]dGTP replaced [α-32P]dITP in the T4 polymerase-catalyzed exchange reaction.

Since the penultimate dG on the 5′-side of one endonuclease cleavage point lies outside the minimal EcoRI recognition sequence, we wished to minimize exchange into this position. However, it was also necessary to ensure that essentially all termini participated in the exchange reaction. As shown in Table II, limitation of the exchange reaction to 30 min at 12° (see "Methods") resulted in preparations of control circles containing an average of 2.0 to 2.1 mol of substituted dG/mol of DNA. Furthermore, nearest neighbor analysis yielded 3′-dAMP in only 10 to 20% excess over 3′-dGMP demonstrating that exchange was limited primarily to the immediate 3′-terminal dG residues. The results of the nearest neighbor analyses can be used to calculate the average stoichiometry for molecules that participated in the exchange reaction. These predicted stoichiometries of 2.1 to 2.2 mol of dG/mol of DNA (Table II) are in good agreement with the values of 2.0 to 2.1 determined on the basis of nucleotide incorporation and indicate that 95% of the isolated circles had participated in the exchange reaction with dG.

However, when dL was the exchangeable nucleotide, significant exchange for the penultimate dG did occur in reaction times required to ensure maximum participation of the DNA. Measured stoichiometries of incorporation ranged from 2.2 to...
2.4 mol of dI incorporated/mol of DNA (Table II). Since maximal exchange would yield 3 mol of dI/mol of DNA, these results indicate 20 to 40% exchange of the penultimate dG residue. Again these measured stoichiometries were close to those predicted by nearest neighbor analysis. Comparison of these values indicates that at least 95% of the molecules of preparation 1 participated in the exchange reaction with dI, while 90% or more of the molecules of preparation 2 were substituted.

**Endonuclease Activity at EcoRI Sites Containing dI**—The activity of the endonuclease on dI and dG-substituted covalently closed circles was compared with that on natural ColEl DNA (Fig. 2). Under steady state conditions, the pattern of cleavage of dG-substituted molecules was very similar to that observed with the natural substrate. In particular, there was no accumulation of a Form II intermediate species (one single strand break in the EcoRI sequence) free in solution. As described previously (6), we interpret this to mean that the endonuclease cleaves both DNA strands within the EcoRI sequence prior to dissociation.

DNA circles containing the dI substitution were also subject to rapid and complete double strand cleavage by the endonuclease. However, accumulation of a Form II intermediate was observed with this substrate demonstrating dissociation of the enzyme after cleavage of only one strand of the helix. Despite this difference in mechanism of double strand cleavage, initial rates of phosphodiester bond hydrolysis with the analog containing dI were essentially identical with those observed with the dG control substrate (Table III). However, since initial rate measurements could be biased by the presence of a small fraction of molecules which did not participate in the T4 DNA polymerase exchange reaction, the times required for hydrolysis of 50% of the covalently closed circles were compared with that on found in 3'-dA and 3'-dG. Alternatively, the extent of substitution for the penultimate dG can be calculated as the ratio of label found in 3'-dI to that recovered in 3'-dG. For dI-substituted DNA this stoichiometry was estimated in two ways. In a manner analogous to that used for the dG substitution, label recovered in 3'-dA was compared with the sum of that found in 3'-dI and 3'-dG. Alternatively, the extent of substitution for the penultimate dG can be calculated as the ratio of label found in 3'-dI to the sum of that in 3'-dI plus 3'-dG. Both methods yielded comparable results, with values determined by the latter procedure shown in parentheses.

**Table II**

<table>
<thead>
<tr>
<th>Substituted nucleotide</th>
<th>Preparation</th>
<th>Nearest neighbor analysis</th>
<th>Predicted stoichiometry</th>
<th>Measured stoichiometry</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>3'-dNMP</td>
<td>=P cpm</td>
<td>mol nucleotide/mol ColEl DNA</td>
<td></td>
</tr>
<tr>
<td>dG</td>
<td>1</td>
<td>dA 378</td>
<td>2.17</td>
<td>2.12</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>dG 323</td>
<td>3.09</td>
<td>3.02</td>
</tr>
<tr>
<td></td>
<td></td>
<td>dT 4</td>
<td>2.17</td>
<td>2.12</td>
</tr>
<tr>
<td></td>
<td></td>
<td>dC 0</td>
<td>2.10</td>
<td>2.00</td>
</tr>
<tr>
<td>dI</td>
<td>1</td>
<td>dA 393</td>
<td>2.52</td>
<td>2.41</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>dG 136</td>
<td>2.52 (2.47)</td>
<td>2.41</td>
</tr>
<tr>
<td></td>
<td></td>
<td>dI 122</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>dT 5</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>dC 8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>dF</td>
<td>1</td>
<td>dA 321</td>
<td>2.49</td>
<td>2.16</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>dG 108</td>
<td>2.49 (2.50)</td>
<td>2.16</td>
</tr>
<tr>
<td></td>
<td></td>
<td>dI 108</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>dT 6</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>dC 6</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

A plausible, but unproven, explanation is that these rate differences are a consequence of differences in superhelical density. Since substituted molecules were ligated in vitro, their superhelical density would be much lower than that of natural ColEl DNA.

**Fig. 1.** Procedure for selective substitution of dI within the EcoRI sequence of ColEl DNA.

**Table III**

The characterization of substituted ColEl DNA circles

Determination of "measured" stoichiometries of nucleotide incorporation and nearest neighbor analyses were performed as described under "Methods." For determination of 32P, all samples were counted for at least 10 min, and activities shown were corrected for a background of 19 cpm. In the solvent system employed to separate deoxyribonucleoside monophosphates for nearest neighbor analysis (see "Methods"), dI ran just behind dT and occasionally was poorly resolved. This was the case with preparation 2 of dI-substituted molecules, in which case all label in this portion of the chromatogram was assumed to be present in 3'-dIMP. Predicted stoichiometries of substitution were calculated on the basis of nearest neighbor analysis and the known sequence about the ColEl EcoRI site. For dG-substituted molecules, this stoichiometry was based on the ratio of label recovered in 3'-dA to that recovered in 3'-dG. For dI-substituted DNA this stoichiometry was estimated in two ways. In a manner analogous to that used for the dG substitution, label recovered in 3'-dA was compared with the sum of that found in 3'-dI and 3'-dG. Alternatively, the extent of substitution for the penultimate dG can be calculated as the ratio of label found in 3'-dI to the sum of that in 3'-dI plus 3'-dG. Both methods yielded comparable results, with values determined by the latter procedure shown in parentheses.

1 A plausible, but unproven, explanation is that these rate differences are a consequence of differences in superhelical density. Since substituted molecules were ligated in vitro, their superhelical density would be much lower than that of natural ColEl DNA.
The presence of dl as the 3'-terminal nucleotide of the EcoRI-cleaved substrate analog was confirmed using T4 DNA polymerase. If \(^{32}\)PdIMP is present within the EcoRI site of the analog, then subsequent to cleavage with EcoRI endonuclease label should be released as 5'-dIMP upon incubation with the polymerase and dGTP via the terminal exchange reaction. Moreover, this release of label from DNA should not occur in the presence of dATP since this nucleotide would allow the polymerase to extend EcoRI-generated 3'-termini (Fig. 1) in a reaction that is much faster than exchange (8).

As shown in Table IV, release of \(^{32}\)P occurred only after prior treatment of dl-containing circles with the endonuclease and was suppressed by the presence of dATP but in the presence of dGTP was complete, yielding dIMP exclusively. Although not shown, analogous experiments with dG-substituted DNA yielded identical results except that \(^{32}\)P was recovered only in 5'-dGMP.

Thus, EcoRI endonuclease can clearly recognize and cleave within a recognition sequence containing dl. Furthermore, since the endonuclease reactions described here were performed at subsaturating DNA concentrations (15% of the \(K_m\) concentration determined previously for ColEl DNA (6)), the kinetic parameters for cleavage within a dl-containing sequence cannot be less favorable than those for cleavage within the natural site. Therefore, the 2-amino group of deoxyguanosine does not have an important role in recognition and catalysis by this enzyme.

**Methylase Activity at EcoRI Sites Containing dl**  
In contrast to results obtained with the endonuclease, the presence of dl within the EcoRI recognition sequence resulted in a marked reduction in the rate of methylation (Fig. 3). Under steady state conditions, the rate of methylation of dl-containing DNA circles was only 4% of that observed with dG-substituted control circles or with natural ColEl DNA. To determine whether this low level of activity reflected methylation of sequences containing dl as opposed to methylation of a small fraction of residual sites with dG, methyl transfer was examined as a function of enzyme concentration (Fig. 4).

### Table III  
**Rates of hydrolysis of substituted ColEl DNA**

Data shown in the first three rows were obtained from experiments shown in Fig. 2, while that in the last three were obtained from an identical experiment using independent preparations of substituted molecules. Initial rates are expressed in terms of phosphodiester bond hydrolysis which was calculated as twice the rate of appearance of Form III DNA plus the rate of appearance of Form II molecules. The latter term was significant only with dl-substituted DNA. The time required for cleavage of 50% of the Form I DNA initially present is designated \(T_{50}\).

<table>
<thead>
<tr>
<th>Substrate</th>
<th>Preparation</th>
<th>Initial rate of hydrolysis</th>
<th>(T_{50})</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>pmol min(^{-1}) ml(^{-1})</td>
<td>min</td>
</tr>
<tr>
<td>dG-substituted</td>
<td>1</td>
<td>0.34</td>
<td>4.6</td>
</tr>
<tr>
<td>dl-substituted</td>
<td>1</td>
<td>0.34</td>
<td>4.2</td>
</tr>
<tr>
<td>Natural ColEl</td>
<td>0.14</td>
<td>8.2</td>
<td></td>
</tr>
<tr>
<td>dG-substituted</td>
<td>2</td>
<td>0.23</td>
<td>7.2</td>
</tr>
<tr>
<td>dl-substituted</td>
<td>2</td>
<td>0.30</td>
<td>5.4</td>
</tr>
<tr>
<td>Natural ColEl</td>
<td>0.10</td>
<td>N.D.(^a)</td>
<td></td>
</tr>
</tbody>
</table>

\(^a\) Not determined.

### Table IV  
**Presence of dl at 3' termini of cleaved molecules**

First incubation reactions (0.03 ml) were performed as described in the legend to Fig. 2 except DNA concentration was 3.4 \(\mu\)M and, when present, the endonuclease concentration was 0.8 \(\mu\)M. After 30 min at 37°, reactions were heated at 65° for 3 min. After cooling to 0°, 2-mercaptoethanol was added to 6 \(\mu\)M dGTP or dATP to 57 \(\mu\)M, and T4 DNA polymerase to 20 \(\mu\)g/ml to yield a reaction volume of 0.035 ml. After 45 min at 12°, 2 \(\mu\)l of 0.5 \(\mu\)M EDTA were added, and reactions were subjected to chromatography on Whatman No. 1 in system II (see “Methods”) in the presence of deoxyribonucleoside 5'-monophosphate markers. \(^{32}\)P was determined as described in the legend to Table II. Label associated with DNA remains at the origin.

<table>
<thead>
<tr>
<th>First incubation</th>
<th>Second incubation</th>
<th>Origin</th>
<th>dA</th>
<th>dG</th>
<th>dl</th>
<th>dT</th>
<th>dC</th>
</tr>
</thead>
<tbody>
<tr>
<td>No enzyme</td>
<td>T4 DNA polymerase</td>
<td>dGTP</td>
<td>270</td>
<td>7</td>
<td>11</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>EcoRI endonuclease</td>
<td>T4 DNA polymerase</td>
<td>dGTP</td>
<td>11</td>
<td>6</td>
<td>9</td>
<td>282</td>
<td>12</td>
</tr>
<tr>
<td>EcoRI endonuclease</td>
<td>T4 DNA polymerase</td>
<td>dATP</td>
<td>269</td>
<td>8</td>
<td>8</td>
<td>27</td>
<td>12</td>
</tr>
</tbody>
</table>
FIG. 3. Reduced rate of methylation of dI-substituted ColEl DNA. Reactions (0.1 ml) contained 0.1 M Tris/HCl (pH 8.0), 5 mM EDTA, 2.5 mM dithiothreitol, 1.1 μM S-adenosyl-L-[methyl-3H]methionine (9500 cpm/pmol), 500 μg/ml of bovine serum albumin, 3.3 nM DNA (as molecules), and 0.2 nM EcoRI methylase. Incubations were at 37°C. Samples (4 μl) were removed at indicated times and extent of methyl transfer determined as described (1). O—O, Natural ColEl DNA; □—□, dG-substituted ColEl DNA; △—△, dI-substituted ColEl DNA.

As can be seen, extensive methylation of molecules containing dI could be observed, but only when enzyme was present in molar excess over DNA. Hence, sequences containing dI can be methylated, but at a much reduced rate. This defect in methylation could be located at the level of sequence recognition or, alternatively, recognition and binding may be normal with the block being at the level of catalysis. If the latter possibility is correct, then methylation of the natural sequence should be inhibited in the presence of the dI-substituted analog. As shown in Table V, the rate of methyl transfer to ColEl DNA was not affected by the presence of a molar excess of dI-containing circles. This indicates that the affinity of the methylase for the sequence containing dI is considerably less than that for the natural recognition site, although additional defects at the level of catalysis have not been ruled out. We have, therefore, concluded that the 2-amino group of guanine has an important role in binding and specific sequence recognition by EcoRI methylase.

DISCUSSION

The experiments described here demonstrate that the 2-amino group of guanine within the EcoRI sequence is required for normal recognition by the methylase but not by the endonuclease. This observation lends additional support to our conclusion that the two proteins interact with their recognition site in different ways (1).

Furthermore, since the extent of analog substitution employed was extremely limited, it seems unlikely that these results reflect a subtle conformational transition in the substrate which blocks sequence recognition only by the methylase. In the absence of such conformational perturbations, our results would implicate the 2-amino group of guanine as an important contact utilized in the process of recognition by the methylase, but not by the endonuclease. Since this amino group is exposed in the minor groove of the helix while the 6-amino group of adenine which is methylated by the enzyme (3) is exposed only in the major groove, this would indicate that the methylase makes contacts with nucleotide determinants that are exposed in both grooves of B form DNA. (This is not meant to imply, however, that the methylase necessarily interacts with the B form structure.)

Inasmuch as cleavage occurs immediately adjacent to dG residues within the EcoRI sequence (2), it may seem anomalous that the dI substitution is without major effect on the endonuclease reaction. However, Polisky et al. (18) have demonstrated that at elevated pH and low ionic strength the requirement for dG.dC base pairs is suppressed, and endonuclease recognition specificity is reduced to the tetranucleotide sequence,

$$\text{d}(\text{T-T-A-A})$$

Although cleavage under such conditions is extremely slow, the fact that cleavage can occur in the absence of the terminal dG.dC base pairs is suppressed, and endonuclease recognition specificity is reduced to the tetranucleotide sequence,

$$\text{d}(\text{A-A-T-T})$$

However, it should be noted that the dI substitution is without major effect on the endonuclease reaction.

REFERENCES

Natl. Acad. Sci. U. S. A. 69, 3448-3452
Role of the 2-amino group of deoxyguanosine in sequence recognition by EcoRI restriction and modification enzymes.

P Modrich and R A Rubin


Access the most updated version of this article at http://www.jbc.org/content/252/20/7273

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 0 references, 0 of which can be accessed free at http://www.jbc.org/content/252/20/7273.full.html#ref-list-1