Role of thymidine residues in DNA recognition by the EcoRI and EcoRV restriction endonucleases

Anja Fliess¹, Heiner Wolfes¹, Andre Rosenthal², Konrad Schweltnus³, Helmut Blöcker³, Ronald Frank³ and Alfred Pingoud¹

¹Abteilung Biophysikalische Chemie, Konstanty-Gutschow-Str. 8, D-3000 Hannover 61, FRG,

²Akademie der Wissenschaften der DDR, Zentralinstitut für Molekularbiologie, Abteilung für Zelldifferenzierung, Robert-Rössle-Str. 10, DDR-1115 Berlin-Buch, GDR and ³DNA-Synthese Gruppe, Gesellschaft für Biotechnologische Forschung mbH, Mascheroder Weg 1, D-3300 Braunschweig, FRG

Received 12 February 1986; Accepted 27 March 1986

ABSTRACT

We have synthesized a series of oligonucleotides containing the EcoRI (GAATTC) or EcoRV (GATATC) recognition site within which or adjacent to which thymidine was substituted by uridine or derivatives of uridine. The effects of these substitutions on the rate of the EcoRI and EcoRV catalyzed cleavage reaction were investigated. Our results show that most of the substitutions within the site are quite well tolerated by EcoRI, not, however, by EcoRV. We conclude that the thymine residues most likely are not directly involved in the recognition process of the EcoRI reaction. In contrast, they are major points of contact, between substrate and enzyme in the EcoRV reaction. The effects of substitutions in the position adjacent to the recognition site is also markedly different for EcoRI and EcoRV. Here, EcoRI seems to be considerably more selective than EcoRV.

INTRODUCTION

The high specificity with which restriction endonucleases attack their specific site within an excess of non-specific sites must reside in several unique contacts that can only be formed in the specific enzyme substrate complex, not, however, or only in part in non-specific complexes. It has been a subject of great interest for many years to elucidate the structural determinants on the DNA that are used by different restriction endonucleases for the discrimination between specific and non-specific DNA. The most widely used approach to define DNA contacts important for specific complex formation has been to analyze the effects of nucleotide substitutions within and adjacent to the recognition sequence on the cleavage reaction (for review cf. (1)).

For the EcoRI restriction endonuclease, the so far best studied restriction enzyme, it was shown, that the replacement of

cytosine in phage DNA with glucosylated hydroxymethylcytosine, not, however, with hydroxymethylcytosine, confers resistance to cleavage by EcoRI (2,3). Substitution of thymine in phage DNA by uracil does not affect the cleavage reaction, the substitution, by hydroxymethyluracil, however, significantly lowers the rate of cleavage (3,4). 5-Bromouracil-substituted DNA is also digested more slowly than the unsubstituted DNA (4-6). Replacement of guanine by hypoxanthine in phage DNA did not have any effect on the cleavage by EcoRI (7). Substitution of the adenine residues by 2-aminopurine in a synthetic oligodeoxynucleotide prevents double strand cleavage by EcoRI (8). A limitation of all these studies was the use of DNA or oligodeoxynucleotides in which all guanine residues, all adenine residues, all thymine, or all cytosine residues were substituted by the modified base. The effect of a given substitution at a particular position within or adjacent to the recognition site cannot be identified unequivocally by this method. We have, therefore, approached the problem of identifying the structural elements involved in the recognition process between EcoRI and its substrate DNA by synthesizing a set of oligodeoxynucleotides containing modified nucleotides at singular defined positions. In the present paper we describe the results of cleavage experiments with oligodeoxynucleotides in which thymidine residues were substituted by a variety of analogues within and adjacent to the recognition site of EcoRI

AAAGAATTCTT

For comparison, analogous experiments were carried out with EcoRV and oligodeoxynucleotides related to

AAAGATATCTT

The results obtained demonstrate that EcoRI, unlike EcoRV, tolerates most of the thymidine substitutions within the recognition sequence, indicating that the 5' positions of the thymine residues are not directly involved in the recognition process for the EcoRI reaction but are sites of contact in the EcoRV reaction. EcoRI and EcoRV are also not equally affected by substitutions at the position adjacent to the recognition site suggesting that these two enzymes also differ in their contacts with the flanking region.

MATERIALS AND METHODS

Oligonucleotides:

All oligonucleotides were synthesized via established phosphotriester chemistry on cellulose disks as segmental supports (9). A detailed description of the preparation of modified deoxyuridine building blocks, synthesis of the oligonucleotides and their characterization by solid phase chemical degradation sequencing (10) and fast atom bombardment mass spectrometry (11) is given elsewhere (manuscript in preparation). Melting curves of oligonucleotides were recorded at a concentration of 0.5 to 1.0 A₂₆₀ units in the buffer used for the cleavage experiments (vide infra) as described (14).

Restriction endonucleases:

EcoRI and EcoRV were isolated in homogeneous form from over-producing E.coli strains kindly provided by Dr. M. Zabeau (Gent) and Dr. F. Winkler (Basel) according to established procedures (12,13).

Cleavage experiments:

Cleavage experiments were carried out with oligonucleotides radioactively labelled at their 5'ends with γ -³²P-ATP (NEN) and T4 kinase (BRL). The EcoRI and EcoRV catalyzed cleavage of the purified phosphorylated oligonucleotides was carried out at 15°C in 20 mM Tris·HCl pH 7.6, 50 mM NaCl and 10 mM MgCl₂. For the EcoRI reaction 4 μ M oligonucleotide (double strand) and 0.019 to 0.19 μ M EcoRI (dimer) were employed, for the EcoRV reaction 2 μ M oligonucleotide and 0.012 to 0.12 μ M EcoRV (dimer). At defined time intervals aliquots were taken from the reaction mixture, spotted onto a DEAE-cellulose plate (Macherey-Nagel) and subjected to homochromatography. The separated radioactive substrates and products were located by autoradiography using a Kodak XAR 5 film and quantitatively analyzed by liquid scintillation counting.

RESULTS AND DISCUSSION

We have synthesized various undecanucleotides as potential substrates for EcoRI and EcoRV.resp.. These oligonucleotides have a selfcomplementary sequence comprising ten nucleotides which under appropriate conditions will pair to give a double

strand with a 5'-terminal unpaired nucleotide:

EcoRI: A A A₃G A₂A₁T¹T²C T³T
 T T³C T²T¹A A G A A A

EcoRV: A A A₃G A₂T¹A₁T²C T³T
 T T³C T²A T¹A G A A A

The oligonucleotides synthesized contain in the position for T¹, T² or T³ thymidine or analogues of thymidine, namely deoxyuridine (dU), bromodeoxyuridine (BdU), α -5-propyluridine (α propU), β -5-propyluridine (β propU) and β -D-arabinosyluracil (araU).

While it could be concluded from the results of several reports (14-17) that the unmodified oligonucleotides, as well as the dU, BdU and araU containing oligonucleotides would be double stranded under the conditions of our experiments, a prerequisite for their use as substrates for EcoRI and EcoRV, this was not the case for the α - and β propU containing oligonucleotides. We have, therefore, recorded melting curves for these compounds. Fig. 1 shows some of these experiments, the result of which is the following: the β propU containing oligonucleotides form duplex molecules with similar T_M -values as the unmodified oligonucleotides. The α propU containing oligonucleotides in which the modified nucleotide is located within the central tetranucleotide do not form a stable duplex.

Figure 1: Melting curves of d(AAAGAA β propUTCTT) (left) and d(AAAGAA α propUTCTT) (right). The melting curves were measured in the cleavage reaction buffer: 20 mM Tris-HCl pH 7.6, 50 mM NaCl and 10 mM MgCl_2 .

All oligonucleotides were phosphorylated at their 5' end (with very good yields, due to the presence of an unpaired nucleotide at the 5' end) and incubated with EcoRI and EcoRV, resp., at various concentrations as indicated in Materials and Methods. Representative experiments are shown in fig. 2. In table 1 the results of all cleavage experiments with EcoRI are given. The substitution of thymidine in either position of the recognition sequence -GAATTC- by deoxyuridine lowers the rate of cleavage by 30%.

This could be due to the fact that the 5-methyl group is missing which may be involved in a hydrophobic interaction in the enzyme substrate complex. The results of the analogous cleavage experiments with EcoRV argue against such a conclusion. On the other hand it is known from NMR studies, that the replacement of a dA-dT basepair by a dA-dU basepair in a selfcomplementary octanucleotide leads to a slight modification in the base orientation around the N-glycosidic bond of the substituted and adjacent basepairs (18), we argue, therefore, that conformational

Figure 2: Cleavage of d(AAAGATATCTT) by EcoRV. The cleavage reaction was carried out in 20 mM Tris.HCl pH 7.6, 50 mM NaCl and 10 mM MgCl₂ at 15°C with 2 μM of the phosphorylated double stranded oligonucleotide and 0.012 (□), 0.024 (○) and 0.12 (Δ) μM EcoRV dimer. The time course of the cleavage reaction (left) was evaluated in terms of initial rates, the concentration dependence of which is shown on the right.

Table 1: Cleavage of modified oligonucleotides by EcoRI

Oligonucleotide	Relative rate of cleavage
AAAGAATTCTT	100
AAAGAAUTCTT	70
AAAGAATUCTT	70
AAAGAA(BU)TCTT	50
AAAGAAT(BU)CTT	60
AAAGAA(α propU)TCTT	0
AAAGAAT(α propU)CTT	0
AAAGAA(β propU)TCTT	0
AAAGAAT(β propU)CTT	0
AAAGAA(araU)TCTT	0
AAAGAAT(araU)CTT	0
AAAGAATTCUT	22
AAAGAATTC(BU)T	42
AAAGAATTC(β propU)T	66
AAAGAATTC(araU)T	11

changes might well be responsible for the slightly reduced rate observed for the cleavage of the dU-substituted oligonucleotides. Substitution of thymidine by bromodeoxyuridine in either position of the recognition sequence reduces the rate of cleavage by 50 and 40% resp. These results are in very good agreement with the results of Berkner & Folk (3,4), and Petruska & Horn (6), who measured the effect of substitution of all thymidine residues in bacteriophage DNA by bromodeoxyuridine on the rate of cleavage by EcoRI. The covalent bond radius of the bromo-residue in bromodeoxyuridine is larger than that of the carbon atom of the methyl group in thymidine. Furthermore, from the comparison of the partition coefficients of toluene and bromobenzene in the octanol water system (19) it can be assumed that the bromo-residue in bromodeoxyuridine has also a similar if not larger hydrophobicity as the methyl group in thymidine (6). There is no information available as to the effect of the substitution of a Br atom for the CH₃ group in thymidine on the structure of DNA. In the absence of such information we conclude that the moderate reduction in the rate of the EcoRI catalyzed cleavage upon replacement of thymidine by bromodeoxyuridine in the oligonucleotide

probably reflects a steric hindrance of the Br atom in the enzyme substrate complex formation (6). This conclusion is reinforced by the result of cleavage experiments with oligonucleotides containing within the recognition site β -5-propyluridine instead of thymidine. The substitution of the methyl group by the considerably larger propyl group makes these oligonucleotides resistant to cleavage. The fact that the α -5-propyluridine containing oligonucleotides are not cleaved is not surprising, since they are not double stranded under the conditions of our experiments. The oligonucleotides in which β -D-arabinosyl uracil replaces a thymidine residue within the recognition sequence is also not cleaved by EcoRI. It is known, that arabinosyl nucleotides prefer the 3'endo over the 2'endo sugar puckering (20). Whether this tendency is also manifest within an oligonucleotide, is not known. A reduction in the rate of the EcoRI catalyzed cleavage of the octanucleotide d(GGAATTCC) in which the first deoxyadenosine residue was replaced by β -D-arabinosyladenine were also discussed in terms of an altered sugar puckering (16). The fact that pyrimidine nucleosides have a higher propensity to adopt a 3'endo conformation than purine nucleosides (21), may be the reason, why the substitution of thymidine by arabinosyluracil has a much larger effect than the replacement of deoxyadenosine by arabinosyladenine. It is evident from stereochemical considerations that the introduction of the hydroxyl group in the 2'position, if it led to a change of the normal sugar pucker, would concomitantly produce a distortion of the regular geometry of the phosphodiester backbone which would render the DNA resistant to cleavage.

The effect of flanking sequences on the rate of cleavage of DNA by restriction endonucleases is now well documented for several of these enzymes (14, 17, 22, 23). Preferential cleavage at particular sites is most likely due to the influence which flanking sequences exert on the conformation of the recognition sequence or its propensity to undergo a conformational change necessary for cleavage to occur (24). It might also be that differences in the rate of cleavage are caused by contacts between the restriction enzyme and the regions adjacent to the recognition site. We have been interested to find out, whether and to

Table 2: Cleavage of modified oligonucleotides by EcoRV

Oligonucleotide	Relative rate of cleavage
AAAGATATCTT	100
AAAGAUATCTT	0
AAAGATAUCTT	0
AAAGA(BU)ATCTT	20
AAAGATA(BU)CTT	20
AAAGA(α propU)ATCTT	0
AAAGA(β propU)ATCTT	0
AAAGATA(α propU)CTT	0
AAAGATA(β propU)CTT	0.5
AAAGA(araU)ATCTT	0
AAAGATA(araU)CTT	0
AAAGATATCUT	120
AAAGATATC(BU)T	106
AAAGATATC(β propU)T	118
AAAGATATC(araU)T	0

what extent modification of the kind introduced within the cleavage site will affect the rate of cleavage when located adjacent to the recognition site. As can be seen in table 1, the replacement of thymidine by uridine and bromodeoxyuridine in the position adjacent to the recognition site has a similar if not larger effect than within the site, indicating that the nucleotide flanking the 3' end of the recognition sequence is at least as much involved in the interaction between enzyme and substrate as the two thymidine residues within the recognition sequence. The contact, however, is probably not as close, since the bulky β -5-propyluridine substitution is better tolerated if located adjacent to rather than within the cleavage site. We conclude from this set of results that flanking sequences influence the conformation of the recognition sequence.

The slow rate with which the arabinosyluracil substituted oligonucleotide with the substitution in the position flanking the recognition site is cleaved indicates that also in this position the distortion of the DNA backbone by the presumed 3' endo conformation of the arabinose residue impairs the correct jux-

taposition of enzym and substrate.

Taken together our results suggest that in forming the specific complex EcoRI does not interact with the methyl group of the thymidine residues within the recognition sequence. Its approach to the major groove in this region is sufficiently close for steric hindrance to occur, when the methyl group of thymidine is replaced by larger substituents. EcoRI is sensitive to slight modifications in the base adjacent to the recognition sequence demonstrating the involvement of flanking sequences in complex formation. Furthermore the conformation of the phosphodiester backbone at the thymidine residues within the recognition sequence and adjacent to it seems to be very critical for efficient cleavage to occur. After completion of the work described here the 3 Å structure of an oligodeoxynucleotide EcoRI restriction endonuclease complex was published (24), from which it was concluded that the direct sequence specific interactions between EcoRI and its substrate consist of twelve hydrogen bonds between arginine and glutamic acid side chains of the protein and the adenine and guanine bases of the DNA. Pertinent to our results and in good agreement is the apparent absence of direct contacts between EcoRI and the thymine residues of the recognition sequence. It has to be pointed out, however, that our experiments were carried out in the presence of Mg^{2+} ions, while the X-ray analysis was performed on co-crystals obtained in the absence of Mg^{2+} .

Analogous cleavage experiments as carried out with EcoRI were done with EcoRV. The results of these experiments are shown in table 2. The replacement of either thymidine residue within the recognition sequence by uridine leads to a complete resistance against cleavage by EcoRV. Obviously, hydrophobic contacts between the enzyme and its substrate at both positions within the recognition site are essential for catalysis to occur. Substitution of thymidine by bromodeoxyuridine and β -5-propyluridine reduces the rate of cleavage by 80% and almost to zero, resp., an indication of steric hindrance. The α -5-propyluridine substituted oligonucleotides, with the substitution within the recognition site which are single stranded under the conditions of the experiment were not expected to be

cleaved. Similar as with EcoRI, the distortion of the phosphodiester backbone by introducing arabinosyluracil in the place of either thymidine within or adjacent to the recognition site is not tolerated by EcoRV. Different from the EcoRI reaction, uridine, bromouridine or 8-5-propyluridine in the position flanking the recognition site do not impair the rate of oligonucleotide cleavage in the EcoRV reaction.

Taken together these results suggest that EcoRV in forming the specific complex strongly relies on hydrophobic contacts to the methyl groups of the thymidine residues within the recognition sequence and that it approaches its substrate very closely in this region such that bulky substituents in place of these methyl groups cause effective steric hindrance. It is interesting to note that the strict dependence of EcoRV on hydrophobic contacts in forming the recognition complex under normal buffer conditions must be relieved in the presence of dimethyl sulphoxide and alkaline pH, since under these conditions EcoRV* sites are cleaved in which thymidine is replaced by other nucleosides (25). The configuration of the phosphodiester backbone both within and adjacent to the recognition site is critical for efficient cleavage to take place. The influence of the nucleotides adjacent to the 3' side of the recognition site is small.

CONCLUSION

We have studied the process of recognition of specific sites on DNA by the two restriction endonucleases EcoRI and EcoRV. Our experimental approach has been to modify residues of interest and to study the effect of these modifications in the rate of cleavage. We have concentrated in the present work on the role of thymidine in the recognition process and in particular on a possible involvement of the 5-methyl group of thymidine in a hydrophobic interaction in the enzyme substrate recognition complex. Our results demonstrate that such interactions are not relevant for the EcoRI catalyzed reaction. In contrast, EcoRV relies in its catalytic action strongly on hydrophobic contacts to both thymidine residues of its recognition sequence. A similar, albeit not as pronounced an effect, was observed recently for EcoRII (26). The confirmation of our results regarding the EcoRI endonuclease by X-ray crystallography stu-

dies lends credit to the predictive value of our analogue studies. Since structural studies using X-ray crystallography and electron microscopy of the EcoRV endonuclease and of its complexes with short DNA duplices are well under way (Winkler, F., personal communication), it will be possible to verify our results for the EcoRV endonuclease in the near future.

ACKNOWLEDGEMENTS:

We thank Dr. M. Zabeau (Gent) and Dr. F. Winkler (Basel) for the EcoRI and EcoRV overproducing *E. coli* strains. We are obliged to Dr. G. Maass, Dr. C. Urbanke and Dr. J. Alves for stimulating discussions. We are grateful to Dr. F. Winkler and Dr. S. Halford (Bristol) for making available to us results prior to publication. Thanks are due to M. Becker, C. Giesa, W. Heikens and H. Krause for their help in oligonucleotide synthesis, R. Mull for drawings and photographs, E. Schuchardt and M. Buchholz for typing the manuscript. This work was supported by grants from the Deutsche Forschungsgemeinschaft (Pi 122, 2-2) and the Fonds der Chemischen Industrie.

REFERENCES:

- 1 Modrich, P. & Roberts, R.J. (1982) in *Nucleases* (Lin, S. & Roberts, R.J., eds.) Cold Spring Harbor Laboratory, Cold Spring Harbor 1982
- 2 Kaplan, D. & Nierlich, P. (1975) *J. Biol. Chem.* 250, 2395-2397
- 3 Berkner, K.L. & Folk, W.R. (1977) *J. Biol. Chem.* 252, 3185-3193
- 4 Berkner, K.L. & Folk, W.R. (1979) *J. Biol. Chem.* 254, 2551-2559
- 5 Marchionni, M.A. & Roufa, D.J. (1978) *J. Biol. Chem.* 253, 9075-9081
- 6 Petruska, J. & Horn, D. (1983) *Nucleic Acids Res.* 11, 2495-2510
- 7 Modrich, P. & Rubin, R.A. (1977) *J. Biol. Chem.* 252, 7273-7278
- 8 Haupt, W. (1983) Dissertation Universität Hannover
- 9 Frank, R., Heikens, W., Heisterberg-Moutsis, G. & Blöcker R. (1983) *Nucleic Acids Res.* 11, 4365-4377
- 10 Rosenthal, A., Schwertner, S., Hahn, V. & Hunger, H.-D. (1985) *Nucleic Acids Res.* 13, 1173-1184
- 11 Grotjahn, L., Frank R. & Blöcker, H. (1982) *Nucl. Acids Res.* 10, 4671-4677
- 12 Scholtissek, S., Pingoud, A., Maass, G. & Zabeau, M. (1986) *J. Biol. Chem.* in press
- 13 d'Arcy, A., Brown, R.S., Zabeau, M., van Resandt, R.W. & Winter, F. (1985) *J. Biol. Chem.* 260, 1987-1990
- 14 Alves, J., Pingoud, A., Haupt, W., Langowski, J., Peters, F., Maass, G. & Wolff, C. (1984) *Eur. J. Biochem.* 140, 83-92
- 15 Goppelt, M., Pingoud, A., Maass, G., Mayer, H., Köster, H., & Frank, R. (1980) *Eur. J. Biochem.* 104, 101-107
- 16 Ohtsuka, E., Ishino, Y., Ibaraki, K. & Ikehara, M. (1984) *Eur. J. Biochem.* 139, 447-450

- 17 Wolfes, H., Fliess, A. & Pingoud, A. (1985) *Eur.J.Biochem.* 150, 105-110
- 18 Delort, A.-M., Neumann, J.M., Molko, D., Herve, M., Teoule, R. & Tran Dinh, S. (1985) *Nucl.Acids Res.* 13, 3343-3355
- 19 Rekker, R.F. (1977) *The Hydrophobic Fragmental Constant*, p. 26, Elsevier Scientific Publishing Company, Amsterdam
- 20 Remin, M., Darzyukiewicz, E., Ekiel, I. & Shugar, J. (1976) *Biochim.Biophys.Acta* 435, 405-416
- 21 Saenger, W. (1984) *Principles of Nucleic Acid Structure*, Springer Verlag, New York
- 22 Armstrong, K.A. & Bauer, W.R. (1982) *Nucleic Acids Res.* 10, 993-1007
- 23 Armstrong, K.A. & Bauer, W.R. (1983) *Nucleic Acids Res.* 11, 4109-4116
- 24 Rosenberg, J.M. McClarin, J.A. Frederick, C.A., Wang, B.-C., Boyer, H.W. & Greene, P. (1985) *Proc. 14th Aharon Katzir Conference*, 18-19
- 25 Halford, E.S. Lovelady, B.M. & McCallum, S.A.(1986), *Gene*, in press
- 26 Yolov, A.A., Vinogradova, M.N., Gromova, E.S., Rosenthal, A., Cech, D., Veiko, V.P., Metelev, V.G., Kosykh, V.G. Buryanov, A.A. & Shabarova, Z.A. (1985) *Nucleic Acids Res.* 13, 8983-8998