Influence of Enzyme–Substrate Contacts Located outside the *Eco*RI Recognition Site on Cleavage of Duplex Oligodeoxyribonucleotide Substrates by *Eco*RI Endonuclease†

Mark D. Van Cleve‡ and Richard I. Gumport*,§

Department of Chemistry, University of Virginia, Charlottesville, Virginia 22901, and Department of Biochemistry and School of Basic Medical Sciences, University of Illinois at Urbana—Champaign, 506 South Matthews Avenue, Urbana, Illinois 61801

Received March 27, 1991; Revised Manuscript Received September 17, 1991

ABSTRACT: A complete understanding of the sequence-specific interaction between the *Eco*RI restriction endonuclease and its DNA substrate requires identification of all contacts between the enzyme and substrate, and evaluation of their significance. We have searched for possible contacts adjacent to the recognition site, GAATTC, by using a series of substrates with differing lengths of flanking sequence. Each substrate is a duplex of non-self-complementary oligodeoxyribonucleotides in which the recognition site is flanked by six base pairs on one side and from zero to three base pairs on the other. Steady-state kinetic values were determined for the cleavage of each strand of these duplexes. A series of substrates in which the length of flanking sequence was varied on both sides of the hexamer was also examined. The enzyme cleaved both strands of each of the substrates. Decreasing the flanking sequence to fewer than three base pairs on one side of the recognition site induced an asymmetry in the rates of cleavage of the two strands. The scissile bond nearest the shortening sequence was hydrolyzed with increasing rapidity as base pairs were successively removed. Taken together, the K_M and k_{cat} values obtained may be interpreted to indicate the relative importance of several likely enzyme–substrate contacts located outside the canonical hexameric recognition site.

DNA sequence recognition by enzymes and binding proteins is a basic biochemical process, giving the cell access to the information encoded in its genome. A convenient model system for the study of this phenomenon is the *Eco*RI restriction-modification system of *Escherichia coli* (Modrich & Roberts, 1982; Terry et al., 1987; Rosenberg et al., 1987), which comprises an endonuclease and a DNA methyltransferase. The endonuclease is a well-studied enzyme, having been cloned and overexpressed (Cheng et al., 1984; Botterman & Zabeau, 1985; Luke & Halford, 1985), purified to homogeneity (Luke & Halford, 1985; Modrich & Zabel, 1976), and characterized with respect to physical and kinetic properties (Modrich & Zabel, 1976) and its gene was sequenced (Greene et al., 1981; Newman et al., 1981). The method of recognition of the *Eco*RI site, duplex GAATTC, by the endonuclease has been studied using a variety of techniques (Terry et al., 1987; Rosenberg et al., 1987; Rosenberg, 1991) including substitution of nucleotide analogues into the recognition site [see Brennan et al. (1986) and references cited therein; Wells et al., 1981; McLaughlin et al., 1987; Jhon et al., 1988; Connolly et al., 1984; Lesser et al., 1990; Aiken et al., 1991) and protection–interference studies (Lu et al., 1981; Becker et al., 1988; Lesser et al., 1990), and models of structures of cocrystals of the enzyme with bound substrate have been derived from X-ray crystallographic studies (McClarin et al., 1986; Kim et al., 1990; Rosenberg, 1991). In addition, the catalytic mechanism of the enzyme has been extensively studied [see Bennett and Halford (1989) for a review].

In this study, we have searched for possible contacts adjacent to the canonical recognition hexamer. The existence of contacts between the enzyme and flanking sequences is suggested by several lines of evidence. Thomas and Davies (1975) reported that the five *Eco*RI sites in bacteriophage λ DNA are not cleaved at equal rates. The first cleavage occurs at the site nearest the right end of the DNA with a rate 10 times that at either of the two centrally located sites. They concluded that this phenomenon probably reflected local differences in the environments of the different *Eco*RI sites. Forsblom et al. (1976) saw rate disparities with *Eco*RI cleavage of adenovirus DNA. Halford et al. (1980) reported that the interactions with adjacent sequences responsible for the different reaction rates occurred during the transition state only. Rubin and Modrich (1978) suggested, from a study of *Eco*RI cleavage of plasmid DNAs, that the different cleavage rates were due to different rates of dissociation of enzyme from DNA cleaved in one strand.

Flanking-region contacts with *Eco*RI endonuclease have been shown directly in several studies. Lu et al. (1981), Becker et al. (1988), and Lesser et al. (1990) showed that covalent modification of either phosphates 3 or 4 (see below) with ethylnitrosourea interfered with binding of the enzyme to its target in the absence of Mg^{2+}. Slight but significant interference was also seen after alkylation of phosphate 2 (Lesser et al., 1990). The structure of the enzyme–substrate cocrystal reported by McClarin et al. (1986) indicated contacts between the enzyme and phosphodiester moieties 2, 3, and 4, and the current model also shows protein in this region of the DNA (Rosenberg, 1991).

In the present work, we attempt to more closely define the contacts outside the canonical sequence in terms of both their location and their importance to the recognition/cleavage process. Each substrate in this study is a duplex of nonpalindromic oligodeoxyribonucleotides so that hairpin formation will not complicate interpretation of the results. The sequence $p_1Ap_2Tp_3Cp_4Gp_5Ap_6Ap_7Tp_8Tp_9Cp_{10}Cp_{11}Gp_{12}Gp_{13}Cp_{14}Cp_{15}A$

† This work was supported in part by NIH Grant GM 25621.
* Author to whom correspondence should be addressed.
‡ University of Virginia.
§ University of Illinois at Urbana—Champaign.

0006-2960/92/0431-334$03.00/0

shows one strand of the longest duplex substrate (Table I, IA) with the recognition sequence in boldface. In one series of compounds, the recognition site is flanked by six base pairs on one side, and from zero to three base pairs at the other. This is referred to as the asymmetric series (Table I). Different-sized products are released by cleavage of the two strands, allowing their easy separation by homochromatography. Steady-state kinetic values were determined for the hydrolysis of each strand of these duplexes. The same was done for a series of substrates (symmetric series) in which the length of flanking sequence was varied simultaneously on both sides of the hexamer (Table II).

Experimental Procedures

Substrates. Oligonucleotides were synthesized on an Applied Biosystems Model 380B DNA synthesizer at the University of Illinois Genetic Engineering Center and purified by reversed-phase high-pressure liquid chromatography employing a 4.6 × 25 cm bonded-phase octylsilane column run at room temperature at a flow rate of 1 mL/min. The buffer was 50 mM potassium phosphate (pH 5.9) in H_2O (buffer A) and 25% methanol (buffer B). The peaks were separated by elution with a gradient of from 0 to 60% B over 20 min and then 60–100% B over 10 min followed by a wash at 100% B for 10 min. Purity was indicated by the appearance of a peak containing greater than 95% of the material absorbing at 260 nm upon re-injection of the purified oligomer. All oligonucleotides were phosphorylated by reaction with ATP and T4 polynucleotide kinase and then further purified by centrifugation through a 1-mL Sephadex G-25 column prior to use. [γ-^{32}P]ATP was used for labeling the 5′ end of oligonucleotides for which kinetics were being determined. The purity of the labeled oligomers was indicated by the appearance of a single spot containing greater than 95% of the total radioactivity on autoradiographs of thin-layer chromatography (TLC) plates developed as described under Assay 1. Oligomer concentrations were determined by measuring the absorbance at 260 nm in H_2O at room temperature and calculating masses by summing using the following extinction coefficients (all mM^{-1} cm^{-1}): adenosine, 9.4; guanosine, 10; cytosine, 5.0; thymidine, 8.0 (Brennan & Gumport, 1985).

*Enzymes. Eco*RI endonuclease and T4 polynucleotide kinase were gifts from P. Modrich and O. C. Uhlenbeck, respectively.

Other Materials. DEAE-cellulose HR/2-15 TLC plates were purchased from Brinkmann Instruments. [γ-^{32}P]ATP was purchased from New England Nuclear. Homomix 5 was prepared as described (Jay et al., 1974). The dinucleotide d(CpG) was from Boehringer-Mannheim Biochemicals.

EcoRI Endonuclease Assays. It was necessary to use two different assays in this study. Assay 1 is faster and more sensitive than assay 2, but requires that the substrates be labeled with [^{32}P]phosphate at the 5′ end. Assay 2 was used to determined turnover numbers for substrates lacking a 5′-phosphate.

(*A*) *Assay 1.* Oligonucleotide duplexes were tested as substrates for the endonuclease by combinging a radiolabeled oligonucleotide with its complement to give the desired final concentration of duplex. The specific activity of the labeled oligonucleotide was typically 2000 Ci/mmol in the assay. The complementary oligomers, in equimolar amounts, were dried with 5 μL of reaction buffer lacking bovine serum albumin (BSA) and dithiothreitol (DTT) in 1.5-mL Eppendorf tubes, redissolved in 5 μL of H_2O, heated to 65 °C, and cooled to 20 °C over 10–15 min. After this treatment, the tubes were centrifuged at 12500*g* for 30 s in an Eppendorf microfuge to collect the liquid. The final reaction mixture (20 μL) contained 100 mM NaCl, 50 mM Tris-HCl (pH 7.4), 10 mM $MgCl_2$, 50 μg/mL BSA, and 10 mM DTT. The BSA and DTT were added after the heating step as part of the cocktail described below.

Substrate was assayed at 20 °C, at concentrations ranging from 0.3 to 3 times the determined K_M value. Dimeric enzyme concentrations between 0.2 and 25 nM gave reaction rates of approximately 0.05–2% of substrate converted to product per minute, with less than 10% of the substrate being consumed during the course of any assay. Velocities of these reactions were linear for at least 10 min and are expressed as moles of phosphodiester bond cleaved per mole of endonuclease dimer. Reactions were started by addition of 15 μL of a cocktail containing an appropriate dilution of enzyme in reaction buffer to the 5 μL of redissolved substrate from the heating/cooling step, and four or five aliquots were withdrawn at 1-min intervals. The concentrations of DTT and BSA in the cocktail were such as to give final reaction concentrations of 10 mM and 50 μg/mL, respectively. Reaction aliquots were spotted on 20 × 20 cm DEAE TLC sheets that were developed in homomix 5 (Jay et al., 1974) at 65 °C after being dampened near the origin with water using a sprayer.

The product and unconsumed substrate were visualized by autoradiography, and the amount of product was quantified by liquid scintillation counting of the excised radioactive segments of the chromatogram on a Beckman Model LS1801 liquid scintillation counter. The radioactivity of the substrate spots varied with substrate concentration and was typically between 5×10^3 and 2×10^5 cpm. The radioactivity of the product spots was between 2×10^2 and 2×10^4 cpm. A background of 80 cpm was subtracted from each value. The velocity of each reaction was calculated by linear-least-squares regression analysis and was determined from 2 to 8 times at each concentration. For each substrate, velocity was plotted as a function of substrate concentration using the program "Hyper", which executes an iterative least-squares fit to a Michaelis–Menten hyperbola and gives values for k_{cat} and apparent K_M and the respective standard errors of measurements (Cleland, 1979). In a typical experiment, the substrate was assayed in duplicate or triplicate at each of five or six duplex substrate concentrations. After the substrates were assayed in this way repeatedly (4–8 times), the standard errors were generally better than 20% for k_{cat} and 40% for K_M.

(*B*) *Assay 2.* Unlabeled oligonucleotide duplexes were tested as substrates for the endonuclease by mixing complementary oligonucleotides in reaction buffer as described above and adding 15 μL of the appropriate enzyme dilution. Aliquots were withdrawn and the reactions stopped by incubation at 95 °C for at least 5 min. The extent of reaction was quantified by separation of the reactants and products by HPLC as described above, and the area under the absorbance peaks was determined on an Altex Model CR1A integrator. For each oligomer assayed, the identity of peaks resulting from digestion was assigned by comparison of retention times with those of standards for the products dG and d(CpG) and undigested substrates. The sensitivity of the HPLC detector necessitates using substrate concentrations of at least 5 mM; thus, this assay cannot be used for accurate determination of K_M values below approximately 15 mM—a value 2–3 orders of magnitude higher than K_M values of the substrates determined in this study.

Results and Discussion

Oligonucleotide duplexes bearing the *Eco*RI recognition hexamer GAATTC and different lengths of flanking sequence

Table I: Asymmetric Substrates

	Substrate	k_{cat} (min^{-1})[a]	K_M (nM)[a]
IA	pATC **GAATTC** CGGCCA[b]	6.3 ± 0.6	22 ± 6
IB	TAG **CTTAAG** GCCGGTp	5.3 ± 0.2	28 ± 3
IIA	pTC **GAATTC** CGGCCA	14.0 ± 0.6	25 ± 3
IIB	AG **CTTAAG** GCCGGTp	4.7 ± 0.5	58 ± 15
IIIA	pC **GAATTC** CGGCCA	18.4 ± 1.1	61 ± 9
IIIB	G **CTTAAG** GCCGGTp	4.2 ± 0.6	111 ± 31
IVA	C **GAATTC** CGGCCA	< 0.5[c]	n.d[d]
IVB	G **CTTAAG** GCCGGTp	1.3 ± 0.1	n.d.
VA	p**GAATTC** CGGCCA	0.24 ± 0.03	120 ± 38
VB	**CTTAAG** GCCGGTp	0.35 ± 0.02	96 ± 13
VIA	**GAATTC** CGGCCA	< 0.5[c]	n.d.
VIB	**CTTAAG** GCCGGTp	0.02 ± 0.01	n.d.

[a] Results are given as means plus or minus standard errors. [b] The oligonucleotides are written as duplexes with the canonical recognition sequence in boldface and separated from the flanking sequences for clarity. Each strand is designated. [c] Determined by assay 2. All other values were determined by assay 1. [d] K_M values could not be determined by assay 2 and are given as not determined (n.d.).

were tested as substrates for the *Eco*RI endonuclease. The sequences of the substrates and the results are shown in Table I. The k_{cat} values obtained are consistent with those previously published for a short octameric duplex substrate, d(pGGAATTCC), of 9 min^{-1} per strand (Brennan et al., 1986) and for a plasmid DNA substrate of 4.6 min^{-1} for double-strand hydrolysis (Jack et al., 1981). The most rapid catalytic rate, 18.4 min^{-1}, is on the upper strand of substrate III. The K_M values are also generally consistent with previous reports: 180 nM for the short substrate d(pGGAATTCC) and 15 nM for the plasmid DNA.

General Considerations. Several points are relevant to the interpretation of these data. Although this study is an attempt to identify enzyme–substrate contacts lying outside the recognition hexamer, other factors may influence the results and complicate analysis of the data. The following discussion enumerates some of the potential ambiguities. See Aiken and Gumport (1991) for a complete discussion of the problems associated with using substrate analogues with restriction enzymes.

The conformation of base pairs in a DNA duplex is dependent on their distance from the end of the duplex. Connolly and Eckstein (1984) found in a study of duplex d(GGAATTCC) that phosphate residues in the interior of the octamer have a structure similar to that found in polymeric DNA whereas those near the termini more closely resemble dinucleoside monophosphates in conformation. The major deviation from DNA-like conformation occurred in the terminal three base pairs. Thus, in our study, the scissile bonds in the various strands may have different conformations, with shorter oligomers having a less DNA-like conformation at the scissile bond than the longer substrates. In addition, a terminal phosphate on an oligonucleotide bears a charge of approximately −2 at the pH value (7.4) used in this study, whereas phosphodiester bonds have a charge of −1. For this reason, removal of a terminal base changes not only the length of the substrate but also the charge density on the phosphate that becomes the new terminus.

The *Eco*RI substrate is duplex DNA (Greene et al., 1975), and since the melting points of oligomeric DNA duplexes change with their length and concentration, it is possible that the different substrates in these series will be present at different concentrations of duplex in the assay. The melting point (T_m) of the self-complementary octamer d(pGGAATTCC) is approximately 35 °C in the buffer and range of concentrations used in the present study (Brennan & Gumport, 1985). This duplex has the same overall base composition as substrate IX, the shortest duplex used in this study, and the sequence differs at only one position. The results of Breslauer et al. (1986) indicate that substrate IX would have a T_m greater than 35 °C, and since all rates in the present study were measured at 20 °C, the percentage of melted substrates is unlikely to be high enough to significantly affect our results.

The differences in rates between any two substrates may be specific for the identity of the base pair being removed, e.g., the difference between cleavage rates for strands IA and IIA might be some other value if IA had a 5′-terminal G rather than A. Since sequence-dependent conformation anomalies exist in B-form DNA (Patel et al., 1982), they are also likely in these flanking sequences. Finally, the rate-limiting step in the hydrolysis reaction may vary with the different substrates. Product release is the rate-limiting step for cleavage of plasmid DNA, but phosphodiester bond hydrolysis may become rate-limiting with shorter substrates (Terry et al., 1987; Brennan et al., 1986).

Asymmetric Series. (A) Measurements of k_{cat} Values. As the first and second base pairs are removed from the flanking region on the left side of the recognition hexamer, the cleavage rates in the upper strand of this series increase 2- and 3-fold, respectively (Table I). The first base is at the boundary of the 12 enveloped by the enzyme, and the phosphate (2) of the second base is also close to the protein because its ethylation introduces a modest interference with binding (Lesser et al., 1990). The rate on this strand then drops precipitously when phosphate 3 is removed. These individual strand-cleavage rates range from less than to twice that reported for the cleavage of duplex d(pGGAATTCC) (9 min^{-1} per strand) (Brennan et al., 1986). The catalytic rates for strands IVA and VIA were too low to be detected, while the rate of cleavage of strand VA was 0.24 ± 0.03 min^{-1}. Since strands IVA and VIA lack 5′-phosphates, assay 2 was used. The relative insensitivity of the HPLC assay makes it impossible to detect small amounts of cleavage; however, one can conclude that substrates lacking phosphates 3 and 4 are hydrolyzed much more slowly than those having them. According to the cocrystal structure (McClarin et al., 1986; Kim et al., 1990), these phosphates, the two immediately 5′ to GAATTC, are buried in the protein and make contact with basic or polar amino acid residues. These phosphates may be involved in correctly positioning the scissile bond for catalysis (Becker et al., 1988; Jen-Jacobson et al., 1991). When ethylated (Lu et al., 1981; Becker et al., 1988; Lesser et al., 1990), phosphates 3 or 4 drastically interfere with binding and are thought to be two of the three (the other being number 7 within the canonical site) that are necessary for recognition (Jen-Jacobson et al., 1991). Our results substantiate these findings and extend them by showing that the removal of these phosphates from one side of the target results in a dramatic drop in the rate of catalysis on both strands.

A number of explanations are possible for the increase in turnover numbers for strands of IIA and IIIA. Increased charge density on phosphates 2 and 3 as they are converted from diesters to monoesters could promote productive interactions with the enzyme. Another possibility relates to the "neo-R2 kink" in the enzyme–substrate complex described by McClarin et al. (1986). In this structure, the phosphate immediately 5′ to the G of the recognition hexamer (number 4) is in an unusual "kinked" conformation, which may (Nerdal

et al., 1989; Thomas et al., 1989) be accentuated by association of the substrate with the enzyme. Whether or not an analogous neo-R2 kink exists with substrates II and III, it is possible that they more easily attain a cleavable conformation than does the longer oligomer.

Cleavage rates in the "B" strands of substrates II and III do not show the increase observed with their complementary "A" strands; rather, they are hydrolyzed at nearly identical rates. There is a 3-fold drop in the rate, from 4.2 to 1.3 min^{-1}, upon removal of the terminal phosphate from IIIB to form IVB, another decrease of the same magnitude, to 0.35 min^{-1}, comparing strands IVB and VB, and then a 17-fold drop to 0.02 min^{-1} for strand VIB (about one cleavage per hour). Over the entire series, there is an approximately 250-fold drop in the rate of catalysis in the bottom strand, but the diminution is more gradual than the corresponding decreases in the upper strand cleavage rates.

It is thus possible to change the *Eco*RI substrate to affect the kinetic constants for the cleavage of one strand, but not the other. This is not the first report of asymmetric activity during *Eco*RI cleavage of oligodeoxyribonucleotides, and the phenomenon has been observed with other endonucleases as well. In their 1979 study of cleavage by *Hpa*II and *Mno*I, Baumstark et al. (1979) found that, while the latter enzyme exhibited no strand preference, its isoschizomer *Hpa*II preferred the pyrimidine-rich strand of an oligomeric duplex substrate over the purine-rich strand by a factor of 3–4. Alves et al. (1984) found that, in the duplex substrate

5′-AAGAATTCCC-3′
3′-TTCTTAAGGG-5′

*Eco*RI cleaves the two strands at different rates, preferring dA·dT base pairs 5′ to the scissile bond. Rubin and Modrich (1978) noted that the less purine–pyrimidine 2-fold symmetry in the flanking sequences of plasmid DNA *Eco*RI sites, the more asymmetric the cleavage rates. In a study of the *Eco*RII endonuclease, Yolov et al. (1985) noted different cleavage rates of the two strands in an oligomer duplex substrate, and Gromova et al. (1987) reported markedly different strand-cleavage rates of a 30-mer duplex by *Mva*I.

In each of these cases, the difference between the two strands was in the sequences flanking the canonical recognition sites. In addition, changes within the recognition sequence itself can give rise to unequal rates of strand cleavage. The substitution of O^6-MedG for dG in the recognition sequence of some endonucleases induces asymmetric cleavage (Voight & Topal, 1990) as does the introduction of an incorrect base pair (*Eco*RI* site) or a mismatched base into the recognition site of *Eco*RI (Thielking et al., 1990). Jen-Jacobson et al. (1991) have, in addition, found that subtle perturbations within the canonical site, e.g., the stereospecific substitution of a phosphorothiodiester at position 7 in one strand, induced an asymmetry of cleavage. Structural asymmetry both within and outside the recognition site can thus cause asymmetry of cleavage. In natural DNA, structural asymmetry in flanking sequences could arise from sequence variability (Patel et al., 1982) and thereby contribute to the observed variabilities in cleavage rates at different *Eco*RI sites.

(*B*) *Measurements of K_M Values.* K_M values increase gradually over the asymmetric series, from 22 ± 6 nM in strand IA to 120 ± 38 nM in strand VA. Since the endonuclease is close to or in contact with its substrate in the flanking regions, this increase in the K_M is to be expected as possible contacts are lost. What may be surprising is that the loss of the two ionic bonds that have been proposed to exist between enzyme and substrate in this region (McClarin et al., 1986) results in only a 6-fold increase in the K_M value, corresponding to the loss of about 1 kcal/mol of binding energy. These contacts may not be important for substrate binding, or the endonuclease may adjust for their loss by forming new contacts. The relatively greater decreases in hydrolysis rates, compared to Michaelis constants, with the loss of phosphates 3 and 4 indicate that contact with these phosphates may serve to stabilize the transition state of the reaction.

Table II: Symmetric Substrates

	Substrate	k_{cat} (min^{-1})[a]	K_M (nM)[a]
VIIA	pATC **GAATTC** CGG[b]	4.5 ± 0.2	45 ± 7
VIIB	TAG **CTTAAG** GCCp	5.8 ± 1.6	69 ± 36
VIIIA	pTC **GAATTC** CG	5.8 ± 0.8	12 ± 8
VIIIB	AG **CTTAAG** GCp	3.3 ± 0.4	19 ± 9
IXA	pC **GAATTC** C	4.1 ± 0.8	225 ± 75
IXB	G **CTTAAG** Gp	5.8 ± 0.6	270 ± 66

[a] Results are given as means plus or minus standard errors. [b] The oligonucleotides are written as duplexes with the canonical recognition sequence in boldface and separated from the flanking sequences for clarity.

Different Michaelis constants were measured for the two strands in some of these duplexes, with the more slowly cleaved strand generally having the higher apparent K_M. An explanation for this result is that each strand may competitively inhibit hydrolysis of the other. *Eco*RI may or may not dissociate from singly cut substrate, depending on the nature of the substrate (Rubin & Modrich, 1978). When the two strands of a particular duplex are viewed as different substrates A and B, the following could occur: (1) duplex AB binds to the active site; (2) substrate A is hydrolyzed; (3) the complex dissociates with strand B intact. Substrate B experiences this sequence of events as nonproductive binding, i.e., as occupying an active site that does not lead to catalysis. As in any case of competitive inhibition, the result will be an unchanged k_{cat} and an elevated apparent K_M. The extent to which this effect manifests itself depends on whether cleavage of the first strand leads to dissociation of the enzyme–duplex complex, which may be different for the two strands, since their cleavage products are of different lengths. Cleavage of the lower strand yields longer products that may not dissociate from the cleaved complex so readily.

Symmetric Series. (*A*) *Measurements of k_{cat} Values.* There is little change in the turnover numbers for the members of this series (Table II). The values for both strands of substrate VII are similar to those for substrate I, suggesting that base pairs 13 through 15 contribute little to the rates of formation of product. When base pairs 11 and 12 are then successively removed (VIII and IX), the rate of cleavage on either strand is relatively unchanged with respect to the control (VII). These results with VIII and IX contrast with those of their analogues (II and III) in the asymmetric series. When two base pairs were removed from the left-side flanking sequence in the asymmetric substrates, the "A" strand-cleavage rate increased to 3 times that of the "B" strand. The relatively constant hydrolysis rates for both strands of the symmetric series substrates argue against duplex-end proximity, kink formation, or changing charge density on terminal phosphates as being responsible for the increases in "A" strand cleavage rates observed with the asymmetric series. Because the scissile bonds in the "A" strands of substrates I through III and VII through IX, respectively, are equally close to the ends of the duplexes, their conformations are likely similar in the analogous cases. Thus, the asymmetries of cleavage rates in substrates II and

III are likely due to an induced asymmetry in the *Eco*RI endonuclease homodimer caused by interaction with the asymmetric substrates. This possibility is difficult to reconcile with the finding of several investigations (Lu et al., 1981; Becker et al., 1988; McClarin et al., 1986) that base pairs 13, 14, and 15, the source of the asymmetry, are not contacted by the enzyme.

The octamer substrate IX is similar to a substrate used in a previous study (Brennan et al., 1986), differing only in the sequence at one position. That symmetric substrate, d-(pGGAATTCC), was cleaved at the rate of 9.0 min^{-1} per strand. Substrate IXB, d(pGGAATTCG), is hydrolyzed at 5.8 min^{-1} and IXA, d(pCGAATTCC), at 4.1 min^{-1}. The significance of the approximately 2-fold difference in hydrolysis rates in the two studies is unclear. It is possible that the enzyme used in the present study was less active than that used earlier or that the slightly different substrates are actually hydrolyzed at different rates. The failure to observe significant asymmetry of cleavage in this series (VII through IX), the members of which lack a longer flanking sequence on the 3′ side of the recognition site, supports the view that these base pairs must be responsible for an induced asymmetry in the enzyme that causes the two strands to be cleaved at different rates in the asymmetric series. These oligonucleotide substrates show that base pairs beyond those usually thought to be in contact with the enzyme can influence its activity.

(*B*) *Measurements of K_M Values.* The K_M values for substrate VII are only slightly greater than those for substrate I, indicating that interactions three base pairs beyond the canonical site do not contribute significantly to this parameter. The slight K_M value increments for substrates, VII and IX in comparison to I and III, respectively, may be due to conformational differences at the duplex termini or to the loss of interactions between the enzyme and the extra bases available on the longer substrates. We have no satisfactory explanation for the anomalously low K_M values for substrate VIII. The values for this substrate were the lowest observed, although less flanking sequence is available for interaction with the enzyme than on some other substrates with higher K_M values. The terminal phosphate (number 2) of substrate VIII is contacted by the enzyme (Lu et al., 1981; McClarin et al., 1986; Lesser et al., 1990). Although removal of phosphate 2 has only a small effect on k_{cat} in the asymmetric substrate, and no effect in the symmetric compounds, it has a marked effect on the K_M values in both series. Thus, contact with this phosphate appears to stabilize the enzyme–substrate complex but to have little effect on enzyme binding to the transition state. The increased negative charge on the phosphomonoester at position 2 in substrate VIII may cause the better binding. A related possibility is that the enzyme is better able to compensate for other lost contacts with this substrate than with the others. Replacement of lost contacts by formation of new ones has been observed in interactions of *Eco*RI with analogue-containing substrates (Jen-Jacobson et al., 1991) and in some *Lac* repressor–operator interactions (Mossing et al., 1985).

Conclusions

Varying flanking-sequence length has delineated the importance of several phosphate contacts in terms of their contributions to the hydrolysis reaction. In addition, we have shown that asymmetry in the substrate arising from unequal flanking-sequence lengths introduces asymmetry into the relative rates of cleavage of the two strands of the duplex. This asymmetry manifests itself even when it is caused by base pairs beyond those thought to be in contact with the enzyme. The relatively greater effects on k_{cat} compared to K_M upon removing the two phosphate contacts immediately adjacent to the recognition site indicate that these interactions may serve more to stabilize binding to the transition state than to the substrate itself.

Acknowledgments

We thank Paul Modrich and Olke Uhlenbeck for their gifts of enzymes and Linda Jen-Jacobson and John Rosenberg for sharing results prior to their publication.

References

Aiken, C. R., & Gumport, R. I. (1991) *Methods Enzymol. 208*, 433–457.

Aiken, C. R., McLaughlin, L. W., & Gumport, R. I. (1991) *J. Biol. Chem. 266*, 19070–19078.

Alves, J., Pingoud, A., Haupt, W., Langowski, J., Peters, F., Maass, G., & Wolff, C. (1984) *Eur. J. Biochem. 140*, 83–92.

Baumstark, B., Roberts, R., & RajBhandary, U. (1979) *J. Biol. Chem. 254*, 8943–8950.

Becker, M., Lesser, D., Kurpiewski, M., Baranger, A., & Jen-Jacobson, L. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 6247–6251.

Bennett, S. P., & Halford, S. E. (1989) *Curr. Top. Cell. Regul. 30*, 57–104.

Botterman, J., & Zabeau, M. (1985) *Gene 37*, 229–239.

Brennan, C. A., & Gumport, R. I. (1985) *Nucleic Acids Res. 13*, 8665–8684.

Brennan, C. A., Van Cleve, M. D., & Gumport, R. I. (1986) *J. Biol. Chem. 261*, 7270–7278.

Breslauer, K. J., Frank, R., Blocker, H., & Marky, L. A. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 3746–3750.

Cheng, S.-C., Kim, R., King, K., Kim, S.-H., & Modrich, P. (1984) *J. Biol. Chem. 259*, 11571–11575.

Cleland, W. W. (1979) *Methods Enzymol. 63*, 103–138.

Connolly, B. A., & Eckstein, F. (1984) *Biochemistry 23*, 5523–5527.

Connolly, B. A., Eckstein, F., & Pingoud, A. (1984) *J. Biol. Chem. 259*, 10760–10763.

Forsblom, S., Rigler, R., Ehrenberg, M., Pettersson, U., & Philipson, L. (1976) *Nucleic Acids Res. 3*, 3255–3269.

Greene, P. J., Poonian, M. S., Nussbaum, A. L., Tobias, L., Garfin, D. E., Boyer, H. W., & Goodman, H. M. (1975) *J. Mol. Biol. 99*, 237–261.

Greene, P. J., Gupta, M., Boyer, H., Brown, W. E., & Rosenberg, J. M. (1981) *J. Biol. Chem. 256*, 2143–2152.

Gromova, E., Kubareva, E., Pain, K.-D., Oretskaia, T., Shabarova, Z., Tsekh, D., & Prokof'ev, M. (1987) *Dokl. Acad. Nauk. SSSR 295*, 1493–1497.

Halford, S. E., Johnson, N. P., & Grinsted, J. (1980) *Biochem. J. 191*, 581–592.

Jack, W. E., Rubin, R. A., Newman, A., & Modrich, P. (1981) in *Gene Amplification and Analysis* (Chirikjian, J. G., & Pappas, T. S., Eds.) Vol. 1, pp 165–179, Elsevier/North-Holland, New York.

Jay, E., Bambara, R., Padmanabhan, R., & Wu, R. (1974) *Nucleic Acids Res. 1*, 331–353.

Jen-Jacobson, L., Lesser, D. R., & Kurpiewski, M. R. (1991) *Nucleic Acids and Molecular Biology* (Eckstein, F., & Lilley, D. M. J., Eds.) Vol. 5, Springer-Verlag, Berlin (in press).

Jhon, N.-I., Casas-Finet, J. R., Maki, A., & Modrich, P. (1988) *Biochim. Biophys. Acta 949*, 189–194.

Kim, Y., Grable, J. C., Love, R., Greene, P. J., & Rosenberg, J. M. (1990) *Science 249*, 1307–1309.

Lesser, D. R., Kurpiewski, M. R., & Jen-Jacobson, L. (1990) *Science 250*, 776-786.
Lu, A.-L., Jack, W., & Modrich, P. (1981) *J. Biol. Chem. 256*, 13200-13206.
Luke, P. A., & Halford, S. E. (1985) *Gene 37*, 241-246.
McClarin, J. A., Frederick, C. A., Wang, B.-C., Greene, P., Boyer, H. W., Grable, J., & Rosenberg, J. M. (1986) *Science 234*, 1526-1541.
McLaughlin, L. W., Benseler, F., Graeser, E., Piel, N., & Scholtissek, S. (1987) *Biochemistry 26*, 7238-7245.
Modrich, P., & Zabel, D. (1976) *J. Biol. Chem. 251*, 5866-5874.
Modrich, P., & Roberts, R. J. (1982) in *Nucleases* (Linn, S. M., & Roberts, R. J., Eds.) pp 109-154, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
Mossing, M. C., & Record, M. T. (1985) *J. Mol. Biol. 186*, 295-305.
Nerdal, W., Hare, D. R., & Reid, B. R. (1989) *Biochemistry 28*, 10008-10021.
Newman, A. K., Rubin, R. A., Kim, S.-H., & Modrich, P. (1981) *J. Biol. Chem. 256*, 2131-2139.
Patel, D., Pardi, A., & Itakura, K. (1982) *Science 216*, 581-590.
Rosenberg, J. M. (1991) *Curr. Opin. Struct. Biol. 1*, 104-113.
Rosenberg, J. M., McClarin, J. A., Frederick, C. A., Grable, J., & Boyer, H. W. (1987) in *Gene Amplification and Analysis*, (Chirikjian, J., Ed.) Vol. 5, pp 119-145, Elsevier/North-Holland, New York.
Rubin, R. A., & Modrich, P. (1978) *Nucleic Acids Res. 5*, 2991-2997.
Terry, B., Jack, W., & Modrich, P. (1987) in *Gene Amplification and Analysis* (Chirikjian, J., Ed.) Vol. 5, pp 103-118, Elsevier/North-Holland, New York.
Theilking, U., Alves, J., Fliess, A., Maass, G., & Pingoud, A. (1990) *Biochemistry 29*, 4682-4691.
Thomas, M., & Davis, R. (1975) *J. Mol. Biol. 91*, 315-328.
Thomas, G. A., Kubasek, W. L., Peticolas, W. L., Greene, P., Grable, J., & Rosenberg, J. M. (1989) *Biochemistry 28*, 2001-2009.
Voight, J. M., & Topal, M. O. (1990) *Biochemistry 29*, 1632-1637.
Wells, R. D., Klein, R. D., & Singleton, C. K. (1981) *Enzymes (3rd Ed.) 14*, 177-182.
Yolov, A., Gromova, E., Kubareva, E., Potapov, V., & Shabarova, Z. (1985) *Nucleic Acids Res. 13*, 8969-8981.

Proton NMR Studies of [*N*-MeCys^3,*N*-MeCys^7]TANDEM Binding to DNA Oligonucleotides: Sequence-Specific Binding at the TpA Site†

Kenneth J. Addess,‡ Dara E. Gilbert,‡ Richard K. Olsen,§ and Juli Feigon*,‡
Department of Chemistry and Biochemistry and Molecular Biology Institute, University of California, Los Angeles, California 90024, and Department of Chemistry and Biochemistry, Utah State University, Logan, Utah 84322
Received July 24, 1991; Revised Manuscript Received October 3, 1991

ABSTRACT: [*N*-MeCys^3,*N*-MeCys^7]TANDEM, an undermethylated analogue of Triostin A, contains two *N*-methyl groups on the cysteine residues only. Footprinting results showed that [*N*-MeCys^3,*N*-MeCys^7]TANDEM binds strongly to DNA rich in A·T residues [Low, C. M. L., Fox, K. R., Olsen, R. K., & Waring, M. J. (1986) *Nucleic Acids Res. 14*, 2015-2033]. However, it was not known whether specific binding of [*N*-MeCys^3,*N*-MeCys^7]TANDEM requires a TpA step or an ApT step. In 1:1 saturated complexes with the octamers $[d(GGATATCC)]_2$ and $[d(GGTTAACC)]_2$, [*N*-MeCys^3,*N*-MeCys^7]TANDEM binds to each octamer as a bis-intercalator bracketing the TpA step. The octadepsipeptide ring binds in the minor groove of the DNA. Analysis of sugar coupling constants from the phase-sensitive COSY data indicates that the sugar of the thymine in the TpA binding site adopts predominantly an N-type sugar conformation, while the remaining sugars on the DNA adopt an S-type conformation, as has been observed in other Triostin A and echinomycin complexes. The drug does not bind to the octamer $[d(GGAATTCC)]_2$ as a bis-intercalator. Only weak nonintercalative binding is observed to this DNA octamer. These results show unambiguously that [*N*-MeCys^3,*N*-MeCys^7]TANDEM binds sequence specifically at TpA sites in DNA. The factors underlying the sequence specificity of [*N*-MeCys^3,*N*-MeCys^7]TANDEM binding to DNA are discussed.

[*N*-MeCys^3,*N*-MeCys^7]TANDEM[1] is a cyclic octadepsipeptide antibiotic that contains a disulfide cross bridge and two quinoxaline rings attached to two D-Ser residues (Chart I). [*N*-MeCys^3,*N*-MeCys^7]TANDEM (hereafter referred to as CysMeTANDEM) is a methylated analogue of the undermethylated TANDEM (des-*N*-tetramethyl-Triostin A). CysMeTANDEM is also an undermethylated analogue of the naturally occurring antibiotic Triostin A, which contains *N*-methyl substituents on the Cys and Val residues. Triostin A and its two synthetic analogues belong to a class of antibiotics called triostins. Echinomycin, which also contains

† This work was supported by grants from the NIH (R01 GM 37254-01) and by a NSF Presidential Young Investigator Award (DMB 89-58280) with matching funds from AmGen Inc., Du Pont/Merck Pharmaceuticals, Monsanto Co., and Sterling Drug Inc. D.E.G. was supported in part by an NIH Predoctoral Cell and Molecular Biology training grant (GM 07185).
‡ University of California, Los Angeles.
§ Utah State University.

[1] Abbreviations: 2D NMR, two-dimensional nuclear magnetic resonance spectroscopy; [*N*-MeCys^3,*N*-MeCys^7]TANDEM; CysMeTANDEM; COSY, correlation spectroscopy; HOHAHA, homonuclear Hartmann-Hahn spectroscopy; *N*-MeCys, *N*-methylcysteine; NOESY, nuclear Overhauser effect spectroscopy; P.E.COSY, primitive exclusive COSY; P.COSY, purged COSY.

0006-2960/92/0431-339$03.00/0