Stockman, B. J., Westler, W. M., Mooberry, E. S., & Markley, J. L. (1988) *Biochemistry 27*, 136-142.
Torchia, D. A., Sparks, S. W., & Bax, A. (1988) *Biochemistry 27*, 5135-5141.
Torchia, D. A., Sparks, S. W., & Bax, A. (1989) *Biochemistry 28*, 5509-5524.
Wingfield, P. T., Payton, M., Tavernier, J., Barnes, M., Shaw, A., Rose, K., Simona, M. G., Demaczuk, S., Williamson, K., & Dayer, J.-M. (1986) *Eur. J. Biochem. 160*, 491-497.
Wingfield, P. T., Payton, M., Graber, P., Rose, K., Dayer, J.-M., Shaw, A., & Schmeissner, U. (1987a) *Eur. J. Biochem. 165*, 537-541.
Wingfield, P. T., Mattaliano, R. J., MacDonald, H. R., Craig, S., Clore, G. M., Gronenborn, A. M., & Schmeissner, U. (1987b) *Protein Eng. 1*, 413-417.
Wingfield, P. T., Graber, P., Movva, N. R., Clore, G. M., Gronenborn, A. M., & MacDonald, H. R. (1987c) *FEBS Lett 215*, 160-164.
Wingfield, P. T., Graber, P., Shaw, A. R., Gronenborn, A. M., Clore, G. M., & MacDonald, H. R. (1989) *Eur. J. Biochem. 179*, 565-571.
Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York.
Wüthrich, K., Billeter, M., & Braun, W. (1984) *J. Mol. Biol. 180*, 715-740.
Zuiderweg, E. R. P., & Fesik, S. W. (1989) *Biochemistry 28*, 2387-2391.

Accuracy of the *Eco*RI Restriction Endonuclease: Binding and Cleavage Studies with Oligodeoxynucleotide Substrates Containing Degenerate Recognition Sequences†

Vera Thielking, Jürgen Alves, Anja Fliess, Günter Maass, and Alfred Pingoud*
Zentrum Biochemie, Medizinische Hochschule Hannover, Konstanty-Gutschow-Strasse 8, D-3000 Hannover 61, West Germany
Received November 3, 1989; Revised Manuscript Received January 9, 1990

ABSTRACT: We have synthesized a series of 18 nonpalindromic oligodeoxynucleotides that carry all possible base changes within the recognition sequence of *Eco*RI. These single strands can be combined with their complementary single strands to obtain all possible *Eco*RI* sequences (left), or they can be combined with a single strand containing the canonical sequence to obtain double strands with all possible mismatches within the recognition sequence (right):

GCGC **A** A A T T C CGCG GCGC **A** A A T T C CGCG
CGCG **T** T T A A G GCGC CGCG C T T A A G GCGC

The rate of phosphodiester bond cleavage of these oligodeoxynucleotides by *Eco*RI was determined in single-turnover experiments under normal buffer conditions in order to find out to what extent the canonical recognition site can be distorted and yet serve as a substrate for *Eco*RI. Our results show that oligodeoxynucleotides containing mismatch base pairs are in general more readily attacked by *Eco*RI than oligodeoxynucleotides containing *Eco*RI* sites and that the rates of cleavage of the two complementary strands of degenerate oligodeoxynucleotides are quite different. We have also determined the affinities of these oligodeoxynucleotides to *Eco*RI. They are higher for oligodeoxynucleotides carrying a mismatch within the *Eco*RI recognition site than for oligodeoxynucleotides containing an *Eco*RI* site but otherwise do not correlate with the rate with which these oligodeoxynucleotides are cleaved by *Eco*RI. Our results allow details to be given for the probability of *Eco*RI making mistakes in cleaving DNA not only in its recognition sequence but also in sequences closely related to it. Due to the fact that the rates of cleavage in the two strands of a degenerate sequence generally are widely different, these mistakes are most likely not occuring in vivo, since nicked intermediates can be repaired by DNA ligase.

Type II restriction endonucleases are considered to be part of a defense system that protects bacterial cells against foreign DNA, in particular bacteriophage DNA (Smith, 1979). They operate by recognizing a defined DNA sequence and cleaving the DNA within this sequence, thereby inactivating it. They do not attack the chromosomal DNA, provided it is kept in a methylated state by a companion modification enzyme. Restriction endonucleases must be highly specific for their recognition sequence, since cleavage of the chromosomal DNA in other sequences not protected by methylation presumably is deleterious for the bacterial cell, as can be inferred from the observation that bacterial cells producing a restriction endonuclease are generally not viable in the absence of the corresponding modification enzyme [for review, cf. Wells et al. (1981), Malcolm (1981), Modrich and Roberts (1982), and Bennett and Halford (1989)].

Considerable progress has been made in understanding the molecular basis of the specificity of restriction enzymes, mainly due to chemical modification studies, in which the structural

† This work has been supported by grants from the Deutsche Forschungsgemeinschaft (Pi 122/2-3 and Ma 465/11-5) and the Fonds der Chemischen Industrie.
* To whom correspondence should be addressed.

0006-2960/90/0429-4682$02.50/0

elements of the DNA substrates needed for efficient cleavage were identified [reference nos. 5–64 in Fliess et al. (1988)]. These studies, however, did not answer the questions of how specific restriction endonucleases are for their canonical recognition sequences or how accurate these enzymes are in discriminating between the canonical recognition sequence and its most closely related sequences, i.e., those in which one base or one base pair is substituted for another one. The accuracy of restriction endonucleases can be determined by measuring the rates of cleavage of oligodeoxynucleotides of defined sequence in which each individual base or base pair within the recognition sequence is substituted by all alternative ones. For a restriction endonuclease that recognizes a palindromic hexanucleotide sequence, this implies that the rates of cleavage of 18 different oligodeoxynucleotides containing a single mismatch and of 9 oligodeoxynucleotides containing a single degenerate base pair have to be measured and compared with the rate of cleavage of the oligodeoxynucleotide with the canonical sequence.

This paper deals with the analysis of the *Eco*RI-catalyzed cleavage of oligodeoxynucleotides containing a degenerate but fully base paired recognition site or a recognition site with a single mismatch. We have chosen the *Eco*RI restriction endonuclease for this investigation, since it is the most thoroughly studied restriction enzyme, with respect to both its structure and its function. It is the only restriction endonuclease for which an X-ray structure analysis has been reported (McClarin et al., 1986), on the basis of which a detailed recognition mode was proposed. It is also one of the few restriction enzymes for which site-directed mutagenesis experiments (Wolfes et al., 1986; Yanofsky et al., 1987; Geiger et al., 1989; Alves et al., 1989a; Needels et al., 1989) and chemical modification studies (Kaplan & Nierlich, 1975; Berkner & Folk, 1977, 1979; Modrich & Rubin, 1977; Marchionni & Roufa, 1978; Petruska & Horn, 1980; Fliess et al., 1986; Brennan et al., 1986; Seela & Driller, 1986; McLaughlin et al., 1987; Seela & Kehne, 1987; Seela et al., 1988) were carried out to identify structural elements of the protein and the DNA, respectively, involved in specificity-determining interactions. We report here the preparation of 27 different oligodeoxynucleotides containing a single mismatch or a single degenerate base pair within the *Eco*RI recognition sequence. These oligodeoxynucleotides were analyzed as substrates for the *Eco*RI restriction endonuclease. Our results demonstrate that the *Eco*RI restriction endonuclease under normal buffer conditions discriminates very accurately between its canonical recognition sequence and sites differing in one *base pair*. It fails, however, to discriminate accurately between the canonical recognition site and sites that differ in one *base* from this sequence. Equilibrium binding and kinetic data demonstrate that the specificity of *Eco*RI is due to preferential binding (K^{assoc}) as well as processing (k_{cat}) of the canonical substrate as compared to degenerate substrates. Our data suggest that a consideration of the energetics of the enzyme–substrate complex and the transition-state complex presumably is not sufficient to explain the specificity of the *Eco*RI restriction endonuclease (or any other enzyme acting with high specificity on a large substrate). We argue instead that a major contribution to the specificity of this enzyme is due to the guidance afforded by the substrate to allow the enzyme to reach the transition state.

Experimental Procedures

EcoRI Restriction Endonuclease Purification and Characterization. The *Eco*RI restriction endonuclease was isolated in homogeneous form from an overproducing strain as described recently (Geiger et al., 1989). It had a specific activity of approximately 3×10^6 units/mg of protein, as assayed with bacteriophage λ DNA. It was verified in single-turnover experiments, similarly as described by Alves et al. (1982), that all enzyme molecules participate in the cleavage reaction since maximum rates of cleavage were obtained at a 1:1 stoichiometry of enzyme and substrates.

Oligodeoxynucleotide Synthesis. Oligodeoxynucleotides were synthesized on solid phase with a Biosearch Model 8600 or Cruachem Model PS 200 DNA synthesizer. β-Cyanoethyl phosphoramidites were obtained from Cruachem, Pharmacia, and Milligen, tetrazole was from Cruachem, acetonitrile was from Baker, and all other chemicals were from Merck or Fluka. Synthesis, purification, and characterization followed the protocols described recently (Fliess et al., 1988).

Melting Curves. The thermal stabilities of double-stranded oligodeoxynucleotides were determined in a Zeiss DMR 10 spectrophotometer. For this purpose equimolar amounts of complementary single-stranded oligodeoxynucleotides were dissolved in cleavage buffer (20 mM Tris·HCl, pH 7.2, 50 mM NaCl, 10 mM $MgCl_2$), heated to 96 °C, and slowly cooled to room temperature. After annealing, the oligodeoxynucleotide concentration of the solutions was adjusted to A^{260nm}(1cm) = 0.9 with cleavage buffer. The melting curves were recorded between 10 and 80 °C. The temperature increase was 20 °C/h.

Cleavage Experiments. Cleavage experiments were performed with radioactively labeled substrates. For this purpose single-stranded oligodeoxynucleotides were phosphorylated at their 5′ end with [γ-^{32}P]ATP. Cleavage experiments were carried out with 1 μM double-stranded oligodeoxynucleotides composed of one radioactively labeled phosphorylated strand and one unlabeled unphosphorylated strand and 0.06–1.5 μM *Eco*RI at 20 °C in cleavage buffer. After defined time intervals aliquots were withdrawn 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 and quantitatively analyzed by liquid scintillation counting. Rates of cleavage (v_o) were determined from the initial linear portion of the reaction progress curve. k_{cat} values were derived from the rates of cleavage with the equation $v_o = k_{cat}$[ES], where [ES] is the concentration of the enzyme–substrate complex. The composition of the cleavage buffer used in this work allows rapid and accurate cleavage of DNA by *Eco*RI. A more alkaline pH and a lower ionic strength would lead to star activity, i.e., cleavage at degenerate recognition sites (Woodhead et al., 1981). However, buffers of a different composition, e.g., those recommended by Boehringer, BRL, New England Biolabs, Pharmacia, etc. that in general are more alkaline (pH 7.5–8.0) and have a higher ionic strength (I = 130–180 mM) than the one used by us (pH 7.2, I = 100 mM), support similarly rapid and accurate DNA cleavage and, therefore, could give results similar to those presented here.

Nitrocellulose Filter Binding Experiments. The affinity of double-stranded oligodeoxynucleotides containing degenerate recognition sequences to the *Eco*RI restriction endonuclease in the *absence* of Mg^{2+} ions was determined by nitrocellulose filtration experiments. A competition assay was used in which the binding of radioactively labeled double-stranded d(pTATAGAATTCTAT) in the presence of unlabeled oligodeoxynucleotides was measured. Binding experiments were carried out at 20 °C in 20 mM Tris·HCl, pH 7.2, 50 mM NaCl, 20 mM EDTA, 50 μM DTE, and 100 μg/mL bovine serum albumin with varying amounts of the competing oligodeoxynucleotide, at a fixed concentration of enzyme and

[^{32}P]d(pTATAGAATTCTAT). After incubation for 30 min, samples were filtered through nitrocellulose filters (Sartorius), washed, and dried. The filter-bound radioactivity was determined by liquid scintillation counting. The experimental data were fitted by a computer program using a multistep predictor/corrector method, which solves the cubic equation for the concentration of the enzyme–substrate complex in the presence of a competitor. The mass equations describing the two coupled equilibria are

$$K_s = (S_o - [ES])(E_o - [ES] - [EC])/[ES] \quad (1)$$

$$K_c = (C_o - [EC])(E_o - [ES] - [EC])/[ES] \quad (2)$$

where K_s and K_c are the equilibrium dissociation constants for the enzyme–substrate and the enzyme–competitor complexes, respectively. E_o, S_o and C_o denote the total concentrations of enzyme, substrate, and competitor, respectively, while [ES] and [EC] are the concentrations of the enzyme–substrate and enzyme–competitor complexes, respectively. Combining eqs 1 and 2, eliminating the term for the concentration of the enzyme–competitor complex, and replacing [ES] by the fractional saturation ($\theta = [ES]/S_T$), one obtains

$$(K_c - K_s)S_T\theta^3 + [K_s(E_o + S_o - C_o) - K_c(E_o - 2S_o) + K_s(K_s - K_c)]\theta^2 + [K_s(S_o - E_o) + K_c(2E_o + S_o) + K_sK_c]\theta - K_cE_o = 0 \quad (3)$$

Equation 3 was used for the fitting procedure to obtain the unknown binding constant K_c when K_s, E_o, S_o, and θ, which depend on the variable C_o, are known.

The equilibrium constant for the binding of the unmodified d(pGCGCGAATTCCGCG) to *Eco*RI in the absence of Mg^{2+} was determined directly as described recently (Alves et al., 1989a).

Results

Degenerate Oligodeoxynucleotide Substrates for EcoRI. We have synthesized 18 nonpalindromic oligodeoxynucleotides that carry all possible base changes within the recognition sequence for the *Eco*RI restriction endonuclease. These oligodeoxynucleotides were recombined with an oligodeoxynucleotide that contains the canonical *Eco*RI sequence as well as with their respective complementary oligodeoxynucleotides. As a result we have obtained 27 different double-stranded oligodeoxynucleotides, 18 of which have a single base and 9 a base pair substitution within the recognition site for *Eco*RI. These oligodeoxynucleotides can be considered as degenerate variants of the following fully base paired asymmetrical oligodeoxynucleotide, which contains the *Eco*RI recognition site in the center:

GCGC GAATTC CGCG

CGCG CTTAAG GCGC

An example of a degenerate but fully base paired oligodeoxynucleotide containing an *Eco*RI* site is

GCGC **A**AATTC GCGC

CGCG **T**TTAAG CGCG

Likewise, an example of a degenerate *Eco*RI substrate containing a mismatch within the *Eco*RI recognition site is

GCGC **A**AATTC GCGC

CGCG CTTAAG CGCG

Due to the GC base pairs, four at both sides, the oligodeoxynucleotide with the canonical *Eco*RI recognition site forms a stable duplex. Under the conditions given under Experimental Procedures, it dissociates reversibly into single strands with a T_M of 65 °C (Figure 1A). While the fully base paired degenerate variants can be assumed to have a similar stability, the oligodeoxynucleotides containing a single mismatch should be less stable. We have determined the T_M values of nine of these oligodeoxynucleotides to be in the range of 46–57 °C (Table I). A representative melting curve is shown in Figure 1B. The destabilization of the duplex

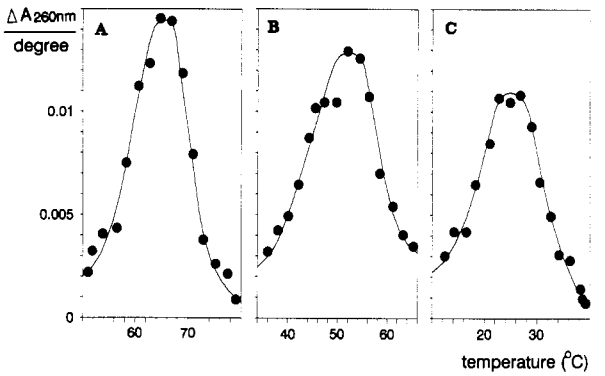

FIGURE 1: Melting curves of various oligonucleotides. Melting curves (in the differential form) of the following oligodeoxynucleotides were recorded in cleavage buffer:

(A) GCGC GAATTC CGCG

CGCG CTTAAG GCGC

(B) GCGC **C**AATTC CGCG containing a single mismatch

CGCG CTTAAG GCGC

(C) GCGC **C**AATTC CGCG containing a single mismatch

CGCG CTTAA A/G GCGC nicked in the canonical strand

Table I: Thermal Stabilities of Some of the Oligodeoxynucleotides Used as Substrates for the *Eco*RI Restriction Endonuclease

oligodeoxynucleotide	T_M[°C] [a]
fully base paired	
GCGC GAATTC CGCG CGCG CTTAAG GCGC	65°
mismatch	
GCGC **C**AATTC CGCG CGCG CTTAAG GCGC	52°
GCGC G**G**ATTC CGCG CGCG C TTAAG GCGC	53°
GCGC G**T**ATTC CGCG CGCG CTTAAG GCGC	48°
GCGC GA**C**TTC CGCG CGCG CTTAAG GCGC	46°
GCGC GA**T**TTC CGCG CGCG CTTAAG GCGC	48°
CGCG GAATA**C** GCGC GCGC CTTA AG CGCC	48°
CGCG GAAT**C**C GCGC GCGC CTTA AG CGCC	48°
CGCG GAAT**G**C GCGC GCGC CTTAAG CGCC	56°
CGCG GAATT**T** GCGC GCGC CTTAAG CGCC	57°
mismatch, nicked	
GCGC GAATT**T** CGCG CGCG CTTAAG GCGC	36°
GCGC **C**AATTC CGCG CGCG CTTAAG GCGC	26°

[a] Values are the mean of two determinations and are accurate within ±2 °C.

structure by the different mismatches is in accordance with published stability data (Patel et al., 1984; Aboul-ela et al., 1985; Werntges et al., 1986).

As will be shown later (vide infra) most degenerate oligodeoxynucleotides are cleaved by *Eco*RI with different rates in their two strands. The resulting nicked intermediate might dissociate into its constituent single-stranded oligodeoxynucleotides, which cannot serve as substrates for *Eco*RI [Greene et al., 1975; Goppelt et al., 1980; Alves et al., 1984; for an alternative view, cf. Bischofsberger et al. (1987)]. We have determined the stabilities of two of these nicked degenerate oligodeoxynucleotides (Figure 1C and Table I). As expected, they are considerably less stable than the intact duplex. The transition temperatures for the melting of these two oligodeoxynucleotides are 26 and 36 °C, respectively (Table I). The position of the mismatch in the oligodeoxynucleotides chosen for the T_M determinations is five base pairs away from the end. All oligodeoxynucleotides that have the mismatch further away from the end should be more stable, such that all oligodeoxynucleotides used here presumably have a duplex structure at 20 °C even after a nick has been introduced by *Eco*RI.

In order to monitor the cleavage of the individual strands of double-stranded oligodeoxynucleotides containing *Eco*RI* sites, we have radioactively labeled one strand for one set of experiments and the other strand for the other set of experiments, viz.

*pGCGC **A** A A T T C CGCG

CGCG **T** T T A A G GCGC

and

GCGC **A** A A T T C CGCG

CGCG **T** T T A A G GCGCp*

For substrates of this size the absence or presence of a phosphate group at the 5′ end has no detectable effect on the rate of cleavage (J. Alves, A. Fliess, and V. Thielking, unpublished results), in agreement with the finding that *Eco*RI interacts strongly with only two phosphate residues outside of the recognition sequence (Lu et al., 1981). The radioactive marker, therefore, can be considered to be an indifferent indicator for the progress of the cleavage reaction, which is monitored by homochromatography followed by autoradiography.

Cleavage of EcoRI Oligodeoxynucleotide Substrates by the EcoRI Restriction Endonuclease.* We have analyzed the *Eco*RI-catalyzed cleavage of nine oligonucleotides that contain all possible *Eco*RI* sites with a single degeneracy. These experiments, one of which is shown in Figure 2, were carried out with concentrations of substrate and enzyme in the micromolar range under single-turnover conditions in a *normal* cleavage buffer. The cleavage in the two strands of each double-stranded oligodeoxynucleotide was measured in two independent experiments. Table II presents a compilation of the results of these experiments, which can be summarized as follows:

(I) Oligodeoxynucleotides containing an *Eco*RI* sequence are cleaved considerably more slowly than the corresponding canonical substrate. One of them is not detectably attacked by *Eco*RI; with the detection limit of our assay, this means that this oligodeoxynucleotide is cleaved by at least a factor of approximately 400 000 more slowly than the canonical substrate.

(II) While in the canonical substrate both strands are attacked with similar rate constants, this is not the case with most of the *Eco*RI* substrates, whose two strands are cleaved with different rates. In one case, cleavage could only be detected in one strand, not however in the other one within the limits of detection (vide supra).

(III) The preference of the *Eco*RI endonuclease for double-strand cleavage in degenerate recognition sites containing

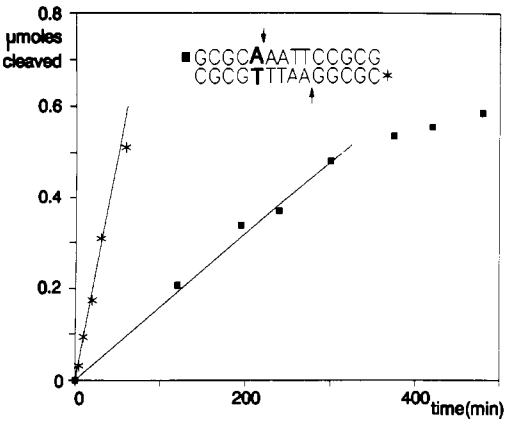

FIGURE 2: Kinetics of cleavage of an *Eco*RI* substrate by the *Eco*RI restriction endonuclease. The kinetics of cleavage of the oligodeoxynucleotide (1 μM) shown below by 1.5 μM *Eco*RI restriction endonuclease was measured at 20 °C in cleavage buffer. The upper trace shows the cleavage of the strand with the sequence -GAATTT- and the lower trace the cleavage of the complementary strand with the sequence -AAATTC-. The solid lines denote that part of the curve from which the initial rates were deduced.

GCGC **A** A A T T C CGCG

CGCG **T** T T A A G GCGC

Table II: Kinetic and Thermodynamic Parameters for the Cleavage of Oligodeoxynucleotides Containing *Eco*RI* Sequences

substrate	k_{cat}^{app} [min^{-1}] [a]	K^{Ass} [M^{-1}] [b]
canonical sequence		
GCGC GAATTC CGCG CGCG CTTAAG GCGC	4 ± 1 4.5 ± 1	(4 ± 1) x 10^{10}
EcoRI* sequences		
substrate	$k_{cat}/k_{cat(canonical)}$ [c]	$K^{Ass}/K^{Ass}_{(canonical)}$ [d]
Position 1/6 of the EcoRI sequence		
GCGC AAATTC CGCG CGCG TTTAAG GCGC	1/2000 1/400	1/7000
GCGC TAATTC CGCG CGCG ATTAAG GCGC	1/30000 1/400	1/4000
GCGC CAATTC CGCG CGCG GTTAAG GCGC	1/400000 1/23000	1/700
Position 2/5 of the EcoRI sequence		
GCGC GCATTC CGCG CGCG CGTAAG GCGC	1/130000 1/150000	1/5000
GCGC GGATTC CGCG CGCG CCTAAG GCGC	< 1/10^6 1/150000	1/5000
GCGC GTATTC CGCG CGCG CATAAG GCGC	1/130000 1/180000	1/8000
Position 3/4 of the EcoRI sequence		
GCGC GACTTC CGCG CGCG CTGAAG GCGC	1/1 1/700	1/2000
GCGC GAGTTC CGCG CGCG CTCAAG GCGC	1/7000 1/2000	1/400
GCGC GATTTC CGCG CGCG CTAAAG GCGC	< 1/10^6 < 1/10^6	1/7000

[a] k_{cat} values are the result of four experiments using oligodeoxynucleotides phosphorylated in one strand. [b] The K^{assoc} value is the result of five titrations using the oligodeoxynucleotide phosphorylated in both strands. [c] k_{cat} values are the result of a minimum of two experiments using oligodeoxynucleotides phosphorylated in one strand and are accurate within at least ±30%. [d] K^{assoc} values are the result of a minimum of three titrations using the unphosphorylated oligodeoxynucleotide and are accurate within at least ±50%.

a single noncognate base pair corresponds to the following hierarchy in which the six positions of the recognition sequence are considered separately:

position 1: G >>> A > T > C

position 2: A >>>> G ≈ C ≈T

position 3: A >> C > G >> T

position 4: T >> G > C >> A

position 5: T >>>> C ≈ G ≈ A

position 6: C >>> T > A > G

(≈ indicates the same order of magnitude; >, >>, >>>, and >>>> correspond to a difference by 1, 2, 3, and 4 orders of magnitude, respectively). It is evident from this rank order that the *Eco*RI restriction endonuclease at position 1 strongly prefers a GC over any other base pair. An AT or TA, not however a CG, base pair is marginally acceptable in this position. Position 2 is the most sensitive one to a base pair substitution: there is an almost absolute requirement for an AT base pair. In contrast, in position 3 an AT base pair is preferred over all other ones, but a CG or GC base pair is also accepted. This result is in agreement with chemical modification studies (McLaughlin et al., 1987), which have shown that the same base substitution in position 2 has a more drastic effect than in position 3. It is noteworthy that substitution of an AT base pair in position 3 of the recognition sequence by a CG base pair leads to the best *Eco*RI* substrate studied here while substitution by a TA base pair results in an oligodeoxynucleotide completely resistant to cleavage. The reason for this may be that the exocyclic amino group of the cytosine residue mimics the exocyclic amino group of the adenine residue normally present in this position and that the methyl group of the thymine residue has an effect similar to that of the methyl group at the exocyclic amino group of the adenine residue, introduced by the *Eco*RI methyltransferase to prevent cleavage at *Eco*RI sites. The effects of substitutions in positions 4–6 are expected to be complementary to those at positions 3–1, due to the symmetry relationship in the palindromic recognition sequence, and have, therefore, not been investigated. Small differences, however, have to be expected, since the recognition sequence in the *Eco*RI* oligodeoxynucleotide substrates is embedded in a nonsymmetrical environment. The influence of different flanking sequences is reflected in the slightly different rates with which the two strands of the canonical oligodeoxynucleotide substrate are cleaved by the *Eco*RI restriction endonuclease [Table II and Alves et al. (1984)].

From the data presented so far it is not clear whether the low efficiency with which *Eco*RI* oligodeoxynucleotides are cleaved by the *Eco*RI restriction endonuclease is due to a drastic decrease in affinity (K^{assoc}) and/or a reduction in the intrinsic rate of phosphodiester bond cleavage (k_{cat}). We have carried out nitrocellulose filter binding experiments in order to determine the affinities of *Eco*RI for the different *Eco*RI* oligodeoxynucleotide substrates. These experiments were carried out with unphosphorylated oligodeoxynucleotides in the absence of Mg^{2+} ions to prevent cleavage of the substrate. This means that the K^{assoc} values were not obtained under exactly the same conditions as the k_{cat} values. While we know that the absence of a single terminal phosphate group has only a small effect on the affinity of the tetradekadeoxynucleotides under consideration for the *Eco*RI restriction endonuclease, $\Delta\Delta G°$ being approximately 0.7 kcal mol^{-1} (J. Alves, unpublished results), we can only extrapolate from stopped-flow

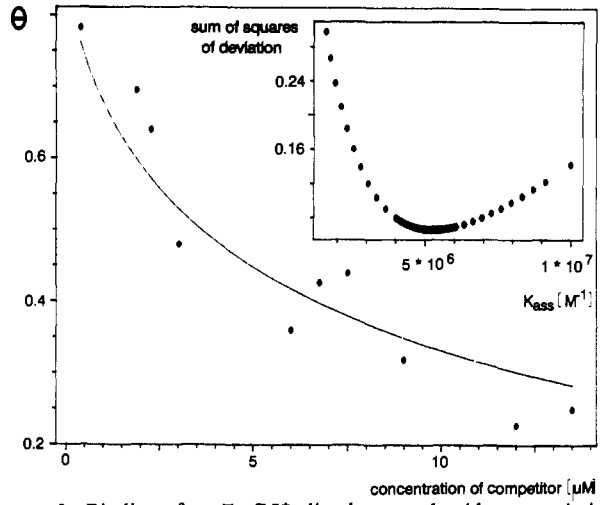

FIGURE 3: Binding of an *Eco*RI* oligodeoxynucleotide to restriction endonuclease *Eco*RI. The binding of the oligodeoxynucleotide shown below to the *Eco*RI restriction endonuclease was determined in competition experiments at 20 °C in cleavage buffer in which 10 mM $MgCl_2$ had been replaced by 20 mM EDTA to prevent cleavage. For this purpose 20 nM radioactively labeled d(pTATAGAATTCTAT) and varying amounts of the unlabeled *Eco*RI* oligodeoxynucleotide were incubated together with 20 nM *Eco*RI restriction endonuclease. θ is a measure for enzyme–d(pTATAGAATTCTAT) complex formation. It was determined by nitrocellulose filter binding experiments and analyzed to give the relative affinities of the *Eco*RI* oligodeoxynucleotide and the canonical substrate for the *Eco*RI restriction endonuclease (cf. Experimental Procedures). The solid line represents the theoretical curve with $K^{assoc} = 2 \times 10^9$ M^{-1} for d(pTATAGAATTCTAT) and 5×10^6 M^{-1} for the *Eco*RI* oligodeoxynucleotide. The insert illustrates the confidence limit for this determination: an increase or decrease of the fitted binding constant by more than 30% leads to an unacceptable large deviation of measured and calculated data points.

GCGC G A G T T C CGCG

CGCG C T C A A G GCGC

experiments with d(pTATAGAATTCTAT) (Alves et al., 1989b) that in the presence of Mg^{2+} the affinity of oligodeoxynucleotide substrates is nearly as high as or higher than that in the absence of Mg^{2+}. The results of the binding experiments, one of which is shown in Figure 3, are given in Table II. They demonstrate that oligodeoxynucleotides containing an *Eco*RI* site are bound on average by 3–4 orders of magnitude more weakly than the corresponding oligodeoxynucleotide with the canonical *Eco*RI recognition site. This means that the specificity of the *Eco*RI restriction endonuclease is at least partially due to preferential binding of its canonical substrate. At high concentrations of enzyme and substrate, however, when complex formation is nearly complete, as approximated in our cleavage experiments, the relative rate of cleavage [v_o/v_o(canonical)] is dominated by the difference in k_{cat}. Hence, it can be concluded that the accuracy of *Eco*RI in discriminating *Eco*RI from *Eco*RI* sites is due to preferential binding as well as turnover of the canonical substrate, which means that thermodynamic *and* kinetic parameters determine the specificity of *Eco*RI. The results presented in Table II also demonstrate that there is no correlation between the rate of cleavage of *Eco*RI* oligodeoxynucleotides and the affinity of these oligodeoxynucleotides for the *Eco*RI restriction endonuclease. This indicates that the absence of favorable contacts or the presence of unfavorable contacts during the interaction between *Eco*RI* oligodeoxynucleotide substrates and the *Eco*RI restriction endonuclease has different effects on the stability of the enzyme–substrate and the transition-state complexes, respectively.

Table III: Kinetic and Thermodynamic Parameters for the Cleavage of Oligodeoxynucleotides with an *Eco*RI Sequence Containing a Mismatch Base Pair

substrate	k_{cat}^{app} [min^{-1}] [a]	K^{Ass} [M^{-1}] [b]
canonical sequence		
GCGC GAATTC CGCG CGCG CTTAAG GCGC	4 ± 1 4.5 ± 1	(4 ± 1) x 10^{10}
mismatch containing sequences		
substrate	k_{cat}/k_{cat}(canonical) [c]	K^{Ass}/K^{Ass}(canonical) [d]
Position 1 of the EcoRI sequence		
GCGC **A**AATTC CGCG CGCG CTTAAG GCGC	1/40 1/1	1/70
GCGC **C**AATTC CGCG CGCG CTTAAG GCGC	1/130 1/1.5	1/80
GCGC **T**AATTC CGCG CGCG CTTAAG GCGC	1/180 1/2	1/50
Position 2 of the EcoRI sequence		
GCGC G**C**ATTC CGCG CGCG CTTAAG GCGC	1/170 2/1	1/20
GCGC G**G**ATTC CGCG CGCG CTTAAG GCGC	1/1600 1/1.5	1/30
GCGC G**T**ATTC CGCG CGCG CTTAAG GCGC	1/1000 1/4	1/30
Position 3 of the EcoRI sequence		
GCGC GA**C**TTC CGCG CGCG CTTAAG GCGC	1/45 3/1	1/30
GCGC GA**G**TTC CGCG CGCG CTTAAG GCGC	1/20 1/5	1/10
GCGC GA**T**TTC CGCG CGCG CTTAAG GCGC	1/5000 1.3/1	1/20
Position 4 of the EcoRI sequence		
CGCG GAA**A**TC GCGC GCGC CTT AAG CGCG	1/600 1/70	1/100
CGCG GAA**C**TC GCGC GCGC CTT AAG CGCG	1/1800 1/95	1/50
CGCG GAA**G**TC GCGC GCGC CTT AAG CGCG	1/25 1/30	1/60
Position 5 of the EcoRI sequence		
CGCG GAAT**A**C GCGC GCGC CTTA AG CGCG	1/6500 1/10	1/60
CGCG GAAT**C**C GCGC GCGC CTTA AG CGCG	1/7500 1/7	1/30
CGCG GAAT**G**C GCGC GCGC CTTA AG CGCG	1/65 1/70	1/70
Position 6 of the EcoRI sequence		
CGCG GAATT**A** GCGC GCGC CTTAAG CGCG	1/110 1/70	1/70
CGCG GAATT**G** GCGC GCGC CTTAAG CGCG	1/150 1/55	1/50
CGCG GAATT**T** GCGC GCGC CTTAAG CGCG	1/130 1/50	1/130

[a] k_{cat} values are the result of four experiments using oligodeoxynucleotides phosphorylated in one strand. [b] The K^{assoc} value is the result of five titrations using the oligodeoxynucleotide phosphorylated in both strands. [c] Values are the result of a minimum of two kinetic experiments and are accurate within at least ±25%. [d] Values are the result of a minimum of three titrations and are accurate within at least ±30%.

Cleavage of EcoRI Oligodeoxynucleotides Substrates Containing a Single Mismatch Base Pair. Oligodeoxynucleotides containing a single mismatch base pair within the *Eco*RI recognition site form stable double strands and as such should be accepted as substrates by the *Eco*RI restriction endonuclease, provided the distortion introduced by the mismatch is not too large. We have analyzed the *Eco*RI-catalyzed cleavage of 18 oligodeoxynucleotides that contain all possible mismatches within the *Eco*RI recognition site. These 18 double-stranded oligodeoxynucleotides belong to two classes, which differ in the arrangement of the sequences flanking the recognition site (cf. Table III). They were obtained by

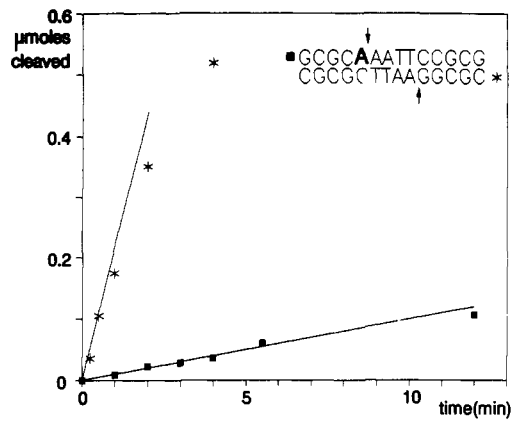

FIGURE 4: Kinetics of cleavage of a mismatch oligodeoxynucleotide by the *Eco*RI restriction endonuclease. The kinetics of cleavage of the oligodeoxynucleotide (1 μM) shown below by 1.5 μM *Eco*RI restriction endonuclease was measured in cleavage buffer. The upper trace shows the cleavage of the canonical strand with the sequence -GAATTC- and the lower trace the cleavage of the strand containing the mismatch -AAATTC-. The solid lines denote that part of the curve from which the initial rates were deduced.

GCGC A A A T T C CGCG

CGCG C T T A A G GCGC

combining two single strands containing a canonical *Eco*RI sequence with the same 18 single-stranded oligodeoxynucleotides as used in the experiments described above to produce *Eco*RI* substrates. Since a GC base pair next to an *Eco*RI site has a similar effect on the rate of cleavage by *Eco*RI as a CG base pair (Alves et al., 1984; this paper), this procedure, which allowed us to synthesize 18 rather than 36 oligodeoxynucleotides, seemed justified. Cleavage experiments were carried out with the same protocol as described for the *Eco*RI* oligodeoxynucleotides (vide supra), i.e., with concentrations of substrate and enzyme in the micromolar range under single-turnover conditions. The cleavage in the canonical strand and the strand containing the noncanonical base was measured in two independent experiments. An example is shown in Figure 4. The results of all cleavage experiments are given in Table III. They can be summarized as follows:

(I) *Double-strand cleavage* occurs more slowly in oligodeoxynucleotides containing a mismatch than in the normal substrate. In general, however, mismatch oligodeoxynucleotides are much better substrates than *Eco*RI* oligodeoxynucleotides; some of them are cleaved nearly as quickly as the corresponding canonical oligodeoxynucleotide.

(II) In the majority of cases the two strands are cleaved at considerably different rates: usually, the strand containing the noncanonical base is cleaved more slowly than the normal strand.

(III) The effect of a mismatch on the rate of nicking, i.e., cleavage of the first strand, depends on the position of the mismatch relative to the phosphodiester bond cleaved: the rate of nicking of the normal strand is nearly as high as or even higher than the rate of nicking of the canonical substrate when the mismatch is in the distal half of the recognition sequence and not as high when in the proximal half. A similar relationship is not observed for the cleavage of the strand containing the noncanonical base.

The most surprising result of the cleavage experiments with oligodeoxynucleotides containing a mismatch within the *Eco*RI recognition site is the fact that in general they are rather good substrates for the *Eco*RI restriction endonuclease. This may be rationalized by two not mutually exclusive explanations. A mismatch within a double-stranded DNA destabilizes the

duplex structure but does not necessarily lead to major distortions of the backbone, presumably because tautomerization (Topal & Fresco, 1976a), wobble interactions (Crick, 1966), anti–syn isomerization (Topal & Fresco, 1976a,b; Traub & Sussman, 1982), and ionization (Topal & Fresco, 1976b) allow formation of noncanonical hydrogen-bonded base pairs. This implies that the strand which contains the canonical sequence will have slight and localized alterations in conformation and will be more flexible compared to the same strand in the fully base paired canonical substrate. This higher flexibility is a consequence of a decrease in stacking interactions, which leads to a lower duplex stability. This may explain why several mismatch substrates are nicked very efficiently in the canonical strand by the *Eco*RI restriction endonuclease, particularly when the mismatch is far away from the site of phosphodiester bond cleavage, which is also the major site of *base* recognition according to X-ray structure analysis of an *Eco*RI–DNA complex (McClarin et al., 1986) viz.

GCGC **A** A A T T C CGCG

CGCG C T T A A/G GCGC

Furthermore, the decrease in stacking interaction in the immediate vicinity of the mismatch may be favorable for the introduction of the type I and II neokinks in the center of and adjacent to the recognition site, which according to Frederick et al. (1984) may facilitate access by the *Eco*RI restriction endonuclease to the major groove of the DNA. The other strand that contains the noncanonical base will have a slightly altered conformation at and around the position of the mismatch as well as an altered functional group pattern as compared to the canonical strand of the mismatch oligodeoxynucleotide. Since both conformation *and* structural elements involved in recognition are different in the noncanonical strand, it is in general less readily cleaved by the *Eco*RI restriction endonuclease than the canonical strand. However, there is no apparent correlation between rate of cleavage on one hand and type and stability of the mismatch within the sequence on the other hand. It is interesting to compare this result obtained for a sequence-specific restriction enzyme with the results of a study in which an attempt was made to correlate thermodynamic and structural parameters of mismatches in oligodeoxynucleotides with repair efficiencies (Werntges et al., 1986). It was shown in this study that there is no correlation between T_M and the biological repair efficiency but that in general mismatches which give rise to wobble base pairs (T/G, G/G, C/A, A/A, A/G) are efficiently repaired while all others (T/T, C/T, T/C, C/C) which are more likely to stay unpaired show very low repair efficiencies. Thus, the repair system does not check the DNA for local instabilities but generally works more effectively on some types of mismatches and less so on others. Such a generalization is not possible for the *Eco*RI restriction endonuclease.

We have also determined the affinities with which the mismatch oligodeoxynucleotides are bound by the *Eco*RI restriction endonuclease. These experiments were carried out with unphosphorylated oligodeoxynucleotides in the absence of Mg^{2+} ions in order to prevent cleavage (Figure 5). As shown in Table III, oligodeoxynucleotides containing a single mismatch within the *Eco*RI recognition sequence are bound only 1–2 orders of magnitude more weakly than the canonical substrate. Within the class of mismatch oligodeoxynucleotides there is no correlation between affinity to the *Eco*RI restriction endonuclease and the rate with which these oligodeoxynucleotides are cleaved. It must be emphasized that most of the *Eco*RI* oligodeoxynucleotides are bound more weakly and

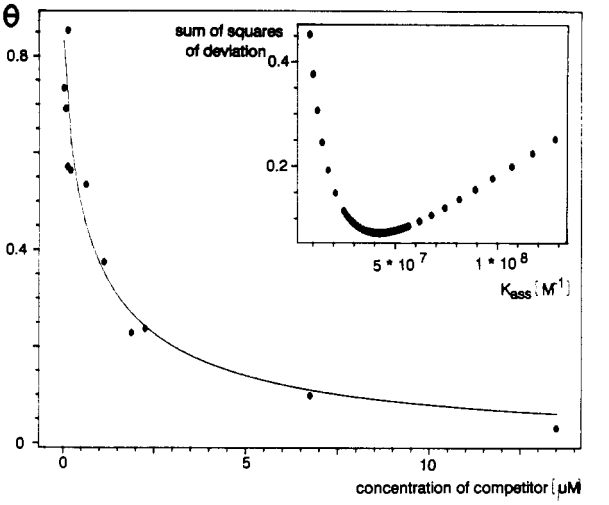

FIGURE 5: Binding of a mismatch oligodeoxynucleotide to the *Eco*RI restriction endonuclease. The binding of the oligodeoxynucleotide shown below to the *Eco*RI restriction endonuclease was determined in competition experiments as described in the legend to Figure 3 and under Experimental Procedures. In this experiment the binding of 20 nM d(pTATAGAATTCTAT) to 20 nM *Eco*RI restriction endonuclease was determined in the presence of varying amounts of the mismatch oligodeoxynucleotide. The solid line represents the theoretical curve with $K^{assoc} = 2 \times 10^9$ M^{-1} for d(pTATAGAATTCTAT) and $K^{assoc} = 4 \times 10^7$ M^{-1} for the mismatch oligodeoxynucleotide. The insert illustrates the confidence limits for this determination: an increase or decrease of the fitted binding constant by more than 30% leads to an unacceptably large deviation of measured and calculated data points.

GCGC G A A T T **G** CGCG

CGCG C T T A A G CGCG

cleaved more slowly than mismatch oligodeoxynucleotides (cf. Tables II and III). As noted above, however, within the class of *Eco*RI* oligodeoxynucleotides there also is no correlation between affinity and rate of cleavage.

Discussion

In the study presented here, which deals with the mechanism of DNA recognition and cleavage by the *Eco*RI restriction endonuclease, nonpalindromic oligodeoxynucleotides were used in which a single base or base pair within the *Eco*RI recognition sequence was exchanged for the three other bases or base pairs that are present in DNA. This study, therefore, differs in two respects from previous studies in which the effects of functional group changes in the *Eco*RI recognition sequence on the cleavage reaction catalyzed by the endonuclease were studied (Fliess et al., 1986; Brennan et al., 1986; McLaughlin et al., 1987; Seela & Kehne, 1987): (I) modifications were introduced only in half of the palindromic recognition sequence; (II) modifications consisted of a substitution by one base or base pair from the set of naturally occurring bases or base pairs. The advantage of this approach is that the effect of a *single* replacement for the cleavage of the first and second strand by the *Eco*RI endonuclease could be studied separately. Since only substrates and substrate analogues containing the natural four bases were employed, our results allow conclusions to be made as to how accurately the *Eco*RI restriction endonuclease works. Type II restriction endonucleases display extremely high specificities toward their target sites on double-helical DNA. Cleavage at noncanonical sites can be detected only at very high enzyme concentrations after extended periods of incubation. It has been shown, for example, that the *Eco*RI restriction endonuclease cleaves not only its recognition sequence -GAATTC- but also at very much reduced

rates a number of alternative sequences that differ in one base pair from the canonical sequence: in pUC8 DNA the sites -TAATTC- and -GAGTTC- are cleaved at 3000 and 9000, respectively, times lower rates than the canonical site (J. Alves and A. Pingoud, unpublished results). Likewise, the *Eco*RV restriction endonuclease in pBR322 DNA attacks not only its canonical site, -GATATC-, but several other sites, one of which, -GAAATC-, is cleaved 15000 times more slowly than the canonical site (Landgraf, 1987). More recently, Taylor and Halford (1989) have shown that in pAT153 DNA the *Eco*RV restriction endonuclease cleaves its canonical site by a factor of 10^6 faster than the next best site, -GTTATC-.

So far, no systematic study of the accuracy of any restriction enzyme was carried out. In this paper we present such a study. We have analyzed the interaction of the *Eco*RI restriction endonuclease with oligodeoxynucleotides containing the canonical recognition site or all possible closely related alternative sites, i.e., those which differ in one base (mismatch oligodeoxynucleotides) or one base pair (*Eco*RI* oligodeoxynucleotides) from the canonical site. They can be considered as the naturally occurring competitors for a canonical *Eco*RI site. Our analysis of the interaction of these oligodeoxynucleotides with the *Eco*RI restriction endonuclease consisted of determining K^{assoc} and k_{cat} values. As expected from the few data obtained with plasmid DNA (vide supra), the *Eco*RI restriction endonuclease discriminates very efficiently between its cognate recognition sequence and naturally occurring base paired alternative sequences, i.e., *Eco*RI* sequences. The discrimination is due to a general decrease in the affinity of the enzyme for these degenerate substrates and to a variable decrease in the rate constant for cleavage. This variability determines a rank order in the hydrolysis rates of the various *Eco*RI substrates, which intriguingly is identical with the one proposed by Goodman et al. (1977) and Rosenberg and Greene (1982) for the cleavage of *Eco*RI* sites in natural DNA by the *Eco*RI restriction endonuclease *under EcoRI* conditions*. This means that this enzyme makes the same mistakes, but less frequently, under optimum than under suboptimum conditions. *Eco*RI* sites with a single noncanonical base pair occur approximately 10 times as frequently as an *Eco*RI site. During the course of evolution the *Eco*RI restriction endonuclease must have been developed such that cleavage in *Eco*RI* sites became very rare, since cleavage in such sites on the chromosomal DNA of the *Escherichia coli* cell (approximately 10000 per *E. coli* chromosome) would be deleterious. One way to achieve this goal was to increase the accuracy; another one was to make the rates of cleavage in the two strands different. In vivo, this would allow the DNA ligase to repair the nicked intermediate. It was indeed shown for the *Eco*RV restriction endonuclease that double-strand cleavage at an *Eco*RV* site in pAT153 was effectively suppressed by DNA ligase, which closes the nicks introduced by the *Eco*RV restriction endonuclease (Taylor & Halford, 1989). Similarly, Heitman et al. (1989) have shown that temperature-sensitive mutants of the *Eco*RI restriction endonuclease are more toxic for *E. coli* cells deficient in DNA ligase than for those with normal DNA ligase levels. Our results demonstrate that the *Eco*RI restriction endonuclease not only effectively discriminates canonical from *Eco*RI* sites (Table II) but in the majority of cases attacks the two strands of an *Eco*RI* site with different rates. Due to the intrinsic accuracy of *Eco*RI and the "proofreading" activity of the DNA ligase, as well as the compartmentation of the *Eco*RI restriction endonuclease in the periplasm [Kohring et al., 1985; for an alternative view, cf. Didier et al. (1988)] and the nonspecific protection of the chromosomal DNA by histone-like proteins (Pingoud et al., 1984), the integrity of the *E. coli* chromosome is therefore not likely to be endangered by the action of the *Eco*RI restriction endonuclease at *Eco*RI* sites in vivo.

*Eco*RI* sites occur more frequently in natural DNA than canonical sites, whereas degenerate *Eco*RI sites with a single mismatch produced as the result of replication errors and recombination events will be encountered very rarely in DNA from natural resources. The *Eco*RI restriction endonuclease, therefore, probably did not have to be developed during evolution to avoid DNA cleavage in such sites. It is nevertheless interesting to know to what extent a highly specific enzyme like the *Eco*RI restriction endonuclease will process a slightly variant substrate. Our results show that the *Eco*RI restriction endonuclease attacks oligodeoxynucleotides containing a mismatch within the recognition site much more readily than *Eco*RI* oligodeoxynucleotide substrates, which indicates that a single mismatch induces only a small perturbation in the structure of the DNA, since large distortions would not be tolerated by an enzyme as accurate as the *Eco*RI restriction endonuclease. This conclusion is in agreement with published data on the structure of oligodeoxynucleotides containing a mismatch, which demonstrate that a mismatch does not disrupt the double helix and has in most cases only localized effects on backbone conformation [for review, cf. Patel et al. (1987), Modrich (1987), and van de Ven and Hilbers (1988)]. The observed differences in the rate of cleavage of mismatch oligodeoxynucleotides by the *Eco*RI restriction endonuclease, nevertheless, are not only the consequence of an altered functional group pattern of the bases of the recognition sequence but may also reflect small distortions of the substrate at the mismatch and its immediate vicinity. At present, it is impossible to separate these effects.

The two identical subunits of the *Eco*RI restriction endonuclease were shown to cooperate in the binding and cleavage of the palindromic substrate (Alves et al., 1982). The X-ray structure analysis of an *Eco*RI oligodeoxynucleotide complex (McClarin et al., 1986) has provided detailed information that allows understanding of the molecular basis of the cooperation in the binding process: the purine nucleotide residues of each half of the recognition sequence are in contact with *both* subunits of the protein via the same set of interactions. With degenerate sites as studied here a particular problem arises: *Eco*RI* and mismatch oligodeoxynucleotides comprise an unmodified and a modified "half-site" (note: half-site is defined by the set of interactions from one subunit to the substrate and does not necessarily coincide with one strand or the left or right half of the recognition sequence). While contacts from the enzyme to the former presumably are close to normal, those to the latter are not. The data presented here demonstrate that cleavage in degenerate sites occurs with preference in the unmodified part of the recognition sequence, e.g.

GCGC G A T T T C CGCG

CGCG C T T A A G GCGC

↑

According to the X-ray structure analysis the subunit that is responsible for this cleavage forms only two hydrogen bonds to the adenine bases of the unmodified half-site, while the other subunit is involved in four hydrogen bonds to the adenine and guanine residues of this site. This means that the subunit which is responsible for the cleavage in the half-site forms fewer specific contacts to this site than the other one. Therefore, it must be triggered by the other subunit to initiate the catalytic reaction. Our data suggest that the cooperation

of the two subunits in the cleavage process is based on a molecular interaction between the recognition module of one subunit with the catalytic module of the other. The demonstration that the amino acid residue Glu111 of the *Eco*RI restriction endonuclease, which is far away from the DNA binding site but close to the subunit/subunit interface of the enzyme, fulfills an essential role in the catalytic process (Wolfes et al., 1986; King et al., 1989; Wright et al., 1989) may be taken as an indication that the catalytic center is activated upon specific recognition via this residue.

Mismatch substrates are cleaved more readily and bound more tightly by the *Eco*RI restriction endonuclease than *Eco*RI* substrates, but otherwise, there is no apparent correlation between K^{assoc} and k_{cat} (Tables II and III): a more firmly bound substrate is not necessarily processed more quickly. This can be rationalized by assuming that a particular oligodeoxynucleotide substrate has a high affinity for the restriction endonuclease but that the enzyme–substrate complex is unproductive, in as much as the interactions that stabilize the ground-state complex destabilize the transition-state complex. Furthermore, it should be considered that reaching the transition state involves a series of more or less random conformational fluctuations whose efficacy for the approach to the transition state depends on the covalent structure of the substrate or substrate analogue. Some structural elements of the substrate or its analogue may be helpful to approach the transition state by favoring conformational fluctuations that eventually lead to the transition state; others may trap the enzyme–substrate complex in unproductive conformational states. Since the substrate of the *Eco*RI restriction endonuclease is large, its dimensions being roughly defined by the numbers of nucleotide residues in contact with the protein, there are multiple possibilities for conformational states and substates, which may or may not lead to the transition state. This means that even when the free enthalpies of an enzyme–substrate complex and an enzyme–substrate analogue complex are the same and the respective transition-state complexes also have the same higher free enthalpy, there might be a difference in the k_{cat} values with which substrate and substrate analogue are processed, because the substrate provides proper "guidance" for the enzyme in the approach to the transition state while the substrate analogue does not. Thus, relatively minor, low-energy conformational changes in the course of enzyme catalysis determine not only the efficiency of an enzyme as discussed in detail by Kraut (1988) but also its specificity. A major conclusion of our work, therefore, is that the structural analysis of the enzyme–substrate complex alone cannot resolve the question how the *Eco*RI restriction endonuclease recognizes its substrate with high specificity. This question can only be answered by a very detailed comparison of the structure and conformational options of canonical and noncanonical enzyme–substrate complexes and of their rate of conversion to enzyme and products.

ACKNOWLEDGMENTS

We thank Kornelia Zander and Rolf Mull-Grotefend for expert technical assistance and Edeltrud Schuchardt and Andrea Meyer for typing the manuscript. We are grateful to Dr. Stephen Halford for making available to us manuscripts prior to publication.

REFERENCES

Aboul-ela, F., Koh, D., & Tinoco, I. (1985) *Nucleic Acids Res. 13*, 4811–4824.

Alves, J., Pingoud, A., Langowski, J., Urbanke, C., & Maass, G. (1982) *Eur. J. Biochem. 124*, 139–142.

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

Alves, J., Rüter, T., Geiger, R., Fliess, A., Maass, G., & Pingoud, A. (1989a) *Biochemistry 28*, 2678–2684.

Alves, J., Urbanke, C., Fliess, A., Maass, G., & Pingoud, A. (1989b) *Biochemistry 28*, 7879–7886.

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

Berkner, K. L., & Folk, W. R. (1977) *J. Biol. Chem. 252*, 3185–3193.

Berkner, K. L., & Folk, W. R. (1979) *J. Biol. Chem. 254*, 2551–2560.

Bischofsberger, N., Ng, P. N., Webb, T. R., & Matteucci, M. D. (1987) *Nucleic Acids Res. 15*, 709–718.

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

Crick, F. H. C. (1966) *J. Mol. Biol. 19*, 548–555.

Didier, S., Lazzaroni, J. C., & Portalier, R. (1988) *Appl. Microbiol. Biotechnol. 28*, 468–470.

Fliess, A., Wolfes, H., Rosenthal, A., Schwellnus, K., Blöcker, H., Frank, R., & Pingoud, A. (1986) *Nucleic Acids Res. 14*, 3463–3474.

Fliess, A., Wolfes, H., Seela, F., & Pingoud, A. (1988) *Nucleic Acids Res. 16*, 11781–11793.

Frederick, C. A., Grable, J., Melia, M., Samudzi, C., Jen-Jacobson, L., Wang, B., Greene, P., Boyer, H. W., & Rosenberg, J. M. (1984) *Nature 309*, 327–331.

Geiger, R., Rüter, T., Alves, J. Fliess, A., Wolfes, H., Pingoud, V., Urbanke, C., Maass, G., Pingoud, A., Düsterhöft, A., & Kröger, M. (1989) *Biochemistry 28*, 2667–2677.

Goodman, H. M., Greene, P. J., Garfin, D. E., & Boyer, H. W. (1977) in *Nucleic Acids Protein Recognition* (Vogel, H. J., Ed.) pp 239–259, Academic Press, New York.

Goppelt, M., Pingoud, A., Maass, G., Mayer, H., Köster, H., & Frank, R. (1980) *Eur. J. Biochem. 104*, 101–107.

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.

Heitman, J., Zinder, N. D., & Model, P. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 2281–2285.

Kaplan, D. A., & Nierlich, D. P. (1975) *J. Biol. Chem. 250*, 2395–2397.

King, K., Benkovic, S. J., & Modrich, P. (1989) *J. Biol. Chem. 264*, 11807–11815.

Kohring, G.-W., & Mayer, F. (1985) *Eur. J. Cell Biol. 37*, 1–6.

Kraut, J. (1988) *Science 242*, 533–540.

Landgraf, A. (1987) Diploma Thesis, Universität Hannover.

Lu, A.-L., Jack, W. E., & Modrich, P. (1981) *J. Biol. Chem. 256*, 13200–13206.

Malcolm, A. D. B. (1981) in *Genetic Engineering 2* (Williamson, R., Ed.) pp 129–173, Academic Press, New York.

Marchionni, M. A., & Roufa, D. J. (1978) *J. Biol. Chem. 253*, 9075–9081.

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. (1987) *Annu. Rev. Biochem. 56*, 435–466.

Modrich, P., & Rubin, R. A. (1977) *J. Biol. Chem. 252*, 7273–7278.

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.
Needels, M. C., Fried, S. R., Love, R., Rosenberg, J. M., Boyer, H. W., & Greene, P. J. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 3579–3583.
Patel, D. J., Kozlowski, S. A., Ikuta, S., & Itakura, K. (1984) *Fed. Proc., Fed. Am. Soc. Exp. Biol. 43*, 2663–2670.
Patel, D. J., Shapiro, L., & Hare, D. (1987) *Annu. Rev. Biophys. Biophys. Chem. 16*, 423–454.
Petruska, J., & Horn, D. (1980) *Biochem. Biophys. Res. Commun. 96*, 1317–1325.
Pingoud, A., Urbanke, C., Alves, J., Ehbrecht, H.-J., Zabeau, M., & Gualerzi, C. (1984) *Biochemistry 23*, 5697–5703.
Rosenberg, J. M., & Greene, P. J. (1982) *DNA 1*, 117–124.
Seela, F., & Driller, H.-J. (1986) *Nucleic Acids Res. 14*, 2319–2332.
Seela, F., & Kehne, A. (1987) *Biochemistry 26*, 2232–2238.
Seela, F., Driller, H.-J., Herdering, W., & DeClercq, E. (1988) *Nucleosides Nucleotides 7*, 347–363.
Smith, H. O. (1979) *Science 205*, 455–462.
Taylor, J. D., & Halford, S. E. (1989) *Biochemistry 28*, 6198–6207.
Topal, M. D., & Fresco, J. R. (1976a) *Nature 263*, 285–289.
Topal, M. D., & Fresco, J. R. (1976b) *Nature 263*, 289–293.
Traub, W., & Sussman, J. L. (1982) *Nucleic Acids Res. 10*, 2701–2708.
van de Ven, F. J. M., & Hilbers, C. W. (1988) *Eur. J. Biochem. 178*, 1–38.
Wells, R. D., Klein, R. D., & Singleton, C. K. (1981) *Enzymes (3rd Ed.) 13*, 157–191.
Werntges, H., Steger, G., Riesner, D., & Fritz, H.-J. (1986) *Nucleic Acids Res. 14*, 3773–3790.
Wolfes, H., Alves, J., Fliess, A., Geiger, R., & Pingoud, A. (1986) *Nucleic Acids Res. 14*, 9063–9080.
Woodhead, J. L., Bhave, N., & Malcolm, A. D. B. (1981) *Eur. J. Biochem. 115*, 293–296.
Wright, D. J., King, K., & Modrich, P. (1989) *J. Biol. Chem. 264*, 11816–11821.
Yanofsky, S. D., Love, R., McClarin, J. A., Rosenberg, J. M., Boyer, H. W., & Greene, P. J. (1987) *Proteins 2*, 273–282.

Molecular Cloning, Primary Structure, and Orientation of the Vertebrate Photoreceptor Cell Protein Peripherin in the Rod Outer Segment Disk Membrane[†,‡]

Gregory J. Connell and Robert S. Molday*
Department of Biochemistry, Faculty of Medicine, University of British Columbia, 2146 Health Sciences Mall, Vancouver, B.C., Canada V6T 1W5

Received June 28, 1989; Revised Manuscript Received January 16, 1990

ABSTRACT: Peripherin, a 39-kDa membrane protein, has been previously localized to the rim region of the vertebrate rod photoreceptor disk membrane by use of monoclonal antibodies and immunocytochemical labeling techniques. As an initial step in determining the structure and function of this protein, we have cloned and sequenced cDNA containing its complete coding sequence. A bovine retinal λgt11 expression library was screened with the antibodies, and a 583 base pair clone was initially isolated. The remaining part of the coding sequence was obtained from subsequent rescreenings of the same library and an independent λgt10 library. A C-terminal CNBr fragment of peripherin was purified by immunoaffinity chromatography and reverse-phase high-performance liquid chromatography. The amino acid sequence of the isolated C-terminal peptide and the N-terminal sequence analysis of immunoaffinity-purified peripherin are in agreement with the cDNA sequence. The cDNA sequence predicts that there are possibly four transmembrane domains. On the basis of immunocytochemical studies and sequence analysis, the hydrophilic C-terminal segment containing the antigenic sites for the antiperipherin monoclonal antibodies has been localized on the cytoplasmic side of the disk membrane. There are three consensus sequences for asparagine-linked glycosylation. Deglycosylation studies have indicated that at least one of these sites is utilized. The possible function of peripherin in relation to its primary structure is discussed.

Rod outer segments (ROS)[1] of vertebrate rod photoreceptor cells are specialized organelles that function in the transduction of light into electrical signals as part of the visual excitation process. These organelles contain hundreds of closely stacked membrane disks that are discontinuous with a surrounding plasma membrane throughout most of the ROS (Rosenkranz, 1977). The photoactive protein, rhodopsin, is the major constituent of both the plasma membrane and the disk membrane. The C-terminus of rhodopsin and the F_1–F_2 loop that joins the fifth and sixth membrane-spanning helices are orientated on the cytoplasmic face of the membranes where they are able to interact with other components of the visual cascade such as rhodopsin kinase and transducin or G-protein (Kuhn,

[†] This research was supported by grants from the National Institutes of Health (EY 02422) and the Medical Research Council of Canada (MT 5822).
[‡] The nucleic acid sequence in this paper has been submitted to GenBank under Accession Number J02884.

[1] Abbreviations: ROS, rod outer segment; SDS, sodium dodecyl sulfate; Ig, immunoglobulin; DEAE, diethylaminoethyl; HPLC, high-performance liquid chromatography; Endo H, endo-β-*N*-acetylglucosaminidase H; bp, base pair; Tris, tris(hydroxymethyl)aminomethane.

0006-2960/90/0429-4691$02.50/0