Changing the Hydrogen-Bonding Potential in the DNA Binding Site of *Eco*RI by Site-Directed Mutagenesis Drastically Reduces the Enzymatic Activity, Not, However, the Preference of This Restriction Endonuclease for Cleavage within the Site -GAATTC-†

Jürgen Alves, Thomas Rüter, Robert Geiger, Anja Fliess, Günter Maass, and Alfred Pingoud*

Zentrum Biochemie, Medizinische Hochschule Hannover, D-3000 Hannover, West Germany

Received June 14, 1988; Revised Manuscript Received October 18, 1988

ABSTRACT: According to the X-ray structure analysis of an *Eco*RI–oligodeoxynucleotide complex [McClarin et al. (1986) *Science 234*, 1526], sequence specificity is mediated by 12 hydrogen bonds, 6 from each of the two identical subunits of the dimeric enzyme to the recognition site -GAATTC-: Arg200 forms two hydrogen bonds with guanine, while Glu144 and Arg145 form four hydrogen bonds to adjacent adenine residues. Changing the hydrogen-bonding potential at the recognition site without perturbing the rest of the interface should lead to the recognition of degenerate sequences [Rosenberg et al. (1987) in *Protein Engineering* (Oxender, D. L., & Fox, C. F., Eds.) pp 237–250, Liss, New York]. We have shown previously that replacing Glu144 by Gln and Arg145 by Lys affects the activity of the enzyme, not, however, its specificity [Wolfes et al. (1986) *Nucleic Acids Res. 14*, 9063]. We show now that also the mutation of Arg200 to Lys, the double mutation Glu144Arg145 to GlnLys, and the triple mutation Glu144Arg145Arg200 to GlnLysLys do not lead to a detectable degeneracy of the specificity of cleavage by *Eco*RI but significantly impair the catalytic activity of this enzyme. A detailed analysis of the steady-state kinetics of cleavage of pUC8 DNA and a tridecadeoxynucleotide substrate demonstrates that the reduction in activity for all DNA binding site mutants investigated so far is mainly due to a decrease in k_{cat}, with the exception of the Arg200 to Lys mutant, which is only impaired in its K_M. This observation is confirmed by nitrocellulose filter experiments which show that this mutant has a lower affinity toward oligodeoxynucleotide substrates than the other DNA binding site mutants and a much lower affinity than wild-type *Eco*RI. While these and previous data clearly demonstrate that Glu144, Arg145, and Arg200 are essential for efficient substrate binding and catalysis, their involvement in the specific recognition is more intricate than proposed on the basis of the X-ray structure analysis. Our results also suggest that the specificity of *Eco*RI cannot easily be relaxed, and in particular not without affecting its activity.

Protein mutagenesis and kinetic analysis allow study of the involvement of individual amino acids in substrate binding and catalytic activity. Site-directed mutagenesis is a particularly fruitful approach when an X-ray structural analysis of the enzyme or, even better, of the enzyme–substrate complex at sufficient resolution exists to identify amino acid residues involved in substrate binding and/or catalysis.

*Eco*RI is the only restriction endonuclease for which an X-ray structural analysis of an enzyme–substrate complex has been published. A scheme for the recognition of the DNA substrate by the enzyme has been proposed on the basis of the X-ray data: specific binding is mediated by 12 hydrogen bonds, 6 from each of the two identical subunits of the dimeric enzyme to the recognition site -GAATTC-. Arg200 forms two hydrogen bonds with guanine; Glu144 and Arg145 form four hydrogen bonds to adjacent adenine residues (Figure 1). It has been suggested that the reduction of the hydrogen-bonding potential at the recognition site by site-directed mutagenesis should lead to recognition of degenerate sequences, provided that the mutation does not disrupt the protein structure at the substrate binding site. We have recently reported that the replacement in *Eco*RI of Glu144 by Gln and of Arg145 by Lys reduces the specific activity of the mutants by nearly 2 orders of magnitude but does not induce a detectable degeneracy of the specificity of cleavage. From these results it was concluded that specificity is established by a complex synergism of many individual interactions and that it cannot easily be relaxed by simply eliminating one or two of these interactions. It was of interest, therefore, to replace also Arg200 by Lys as well as to introduce double and triple mutations at the positions Glu144, Arg145, and Arg200 and to investigate whether these mutants still have enzymatic activity and, if so, whether they show a different specificity.

Experimental Procedures

Oligonucleotide Synthesis and Purification. Oligonucleotides used as substrates were synthesized on a 1-μmol scale as described in the accompanying paper (Geiger et al., 1989). The yield of the crude product was about 80 OD^{260}. The 5′-(dimethoxytrityl)-protected oligonucleotides were purified by preparative HPLC on a 2.5 × 25 cm LiChrosorb RP18 (Merck) column at a flow rate of 4 mL/min using a 90-min gradient from 15% to 50% acetonitrile in 0.1 M triethylammonium acetate (TEAA), pH 7.0. Purified oligonucleotides were lyophilized and deprotected in 80% acetic acid, lyophilized again, and extracted three times with diethyl

†This work has been supported by the Deutsche Forschungsgemeinschaft (Ma 465/11-5, Pi 122/2-3) and the Fonds der Chemischen Industrie, as well as a predoctoral fellowship of the BMFT to R.G.

*To whom correspondence should be addressed at the Abteilung Biophysikalische Chemie, Zentrum Biochemie, Medizinische Hochschule Hannover, Konstanty-Gutschow-Strasse 8, D-3000 Hannover, West Germany.

0006-2960/89/0428-2678$01.50/0

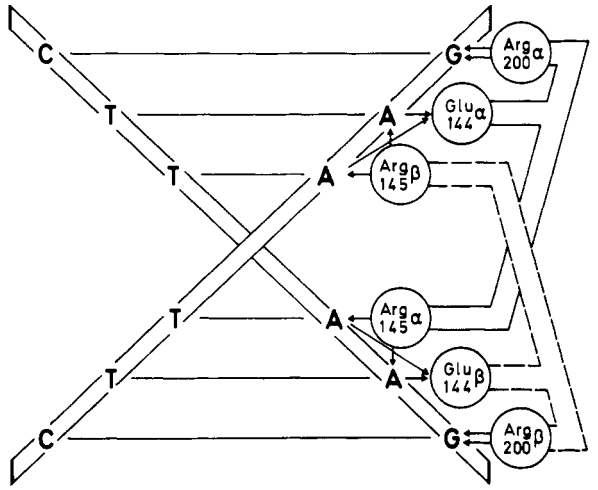

FIGURE 1: Scheme of the recognition of the DNA sequence -GAATTC- by *Eco*RI. This scheme is a modification of a schematic representation of the recognition interactions which according to McClarin et al. (1986) determine the specificity of *Eco*RI. It shows the symmetry of the recognition complex and indicates the overlapping interactions between the two identical subunits of *Eco*RI (α, β) and the two halves of the palindromic DNA sequence.

ether. The purity of the oligonucleotides was checked by HPLC on a wide-pore C_4 (Baker) column (0.46 × 25 cm) at a flow rate of 1 mL/min using a 25-min gradient from 5% to 17.5% acetonitrile in 0.1 M TEAA, pH 7.0. The yield of the purified oligodeoxynucleotide was about 10 OD^{260}.

EcoRI-Catalyzed Cleavage of Bacteriophage λ. Cleavage experiments with bacteriophage λ DNA were carried out to determine the specific activities of the mutant enzymes and to find out whether *Eco*RI mutants produce the same cleavage pattern as wild-type *Eco*RI. A total of 0.25 μg of bacteriophage λ DNA (λ cI 857 Sam 7 from Boehringer Mannheim) was incubated with wild-type or mutant *Eco*RI in 15 μL of 0.02 M Tris·HCl, pH 7.2, 10 mM $MgCl_2$, and 50 mM NaCl at 37 °C. The reaction was stopped after 15 min by addition of 5 μL of a solution containing 0.25 M EDTA, 0.2% (w/v) SDS, 25% (w/v) sucrose, and 0.1% (w/v) bromophenol blue to a 10-μL aliquot. Reaction products were separated on 1% (w/v) agarose gels. One unit is defined as the amount of enzyme required to produce a complete digest of 1 μg of λ DNA in 60 min. The wild-type *Eco*RI preparation which was used as a reference in this study was obtained by the same isolation procedure as the *Eco*RI mutants. It had a specific activity of 3×10^6 units/mg.

The specific activity was determined by varying the enzyme concentration over a wide range by diluting the stock solution by a factor of 10, 100, 1000, etc. in order to determine the approximate magnitude of the enzymatic activity. Within the range of interest the enzyme concentration was then varied by a factor of 2, 4, 6, and 8. The specific activity is, therefore, accurate within ~25%. Furthermore, all experiments were carried out with at least two batches of *Eco*RI and *Eco*RI mutants (different fermentations and preparations) and gave within this limit identical results.

The bacteriophage λ cleavage assay cannot be used to determine specific activities below $\sim 1 \times 10^2$ units/mg. The limited solubility of *Eco*RI and *Eco*RI mutants precludes quantitative estimates of enzymatic activities which are by a factor of 10^4–10^5 lower than *Eco*RI wild-type activity. This is true not only for the cleavage of bacteriophage λ DNA but also for other DNA substrates.

Determination of K_M and V_{max} Values from pUC8 Cleavage Data. K_M and V_{max} values for the cleavage of pUC8 were determined for the Lys200 mutant and the Gln144Lys145 double mutant. Concentrations of 0.18 μM Lys200 mutant and 0.036 μM Gln144Lys145 double mutant were incubated with 7.4, 14.8, 29.6, and 118.3 nM pUC8, respectively, at 37 °C in 0.02 M Tris·HCl, pH 7.2, 10 mM $MgCl_2$, and 0.05 M NaCl. Aliquots were withdrawn after 2, 4, 8, 20, 45, and 100 min, respectively, and the DNA was analyzed by electrophoresis on 2% (w/v) agarose gels. The ethidium bromide stained gels were photographed. The negatives were scanned with a LKB2202 Ultroscan laser densitometer, and the scans were digitized and numerically integrated. A calibration had shown that peak areas thus obtained are proportional to DNA concentration within the range used. The relative amount of superhelical, open circular, and linear DNA in each sample was normalized to the total DNA concentration present to correct for pipetting errors which had occurred during the loading of the electrophoresis gels. The resulting reaction progress curves, supercoiled pUC8 concentration as a function of time, were analyzed according to

$$\textit{Eco}\text{RI} + \text{sc pUC8} \rightleftharpoons \textit{Eco}\text{RI·sc pUC8} \rightarrow \textit{Eco}\text{RI} + \text{pUC8 (oc, li)}$$

(where sc, oc, and li denote the supercoiled, open circular, or linear form of the plasmid DNA) by a computer program that directly fits the K_M and V_{max} values to a set of cleavage experiments carried out at various concentrations. The reliability of the results of the numerical iteration procedure can be checked by carrying out simulations with K_M and V_{max} values which differ from the values obtained in the best fit and by comparing the root mean square deviations of the results of the iteration and the simulation. We have indicated in Tables II and III the ranges of the K_M and V_{max} values that give a ±10% difference of the root mean square deviation, since we estimate that this is the experimental error of each data point. An analogous procedure was used for the estimation of the experimental uncertainty in determining the K_{assoc} values given in Table IV.

Cleavage of Synthetic Oligonucleotides. Typically, 1–10 μM double-stranded d(TATAGAATTCTAT) was incubated with 0.5 μM *Eco*RI mutants in 0.02 M Tris·HCl, pH 7.2, 10 mM $MgCl_2$, and 50 mM NaCl at 25 °C. At defined time intervals the progress of the reaction was analyzed by HPLC on a 0.46 × 25 cm C_4 wide-pore column (Baker) with a 20-min gradient from 5% to 15% acetonitrile in 0.1 M TEAA, pH 7.0, using a Merck-Hitachi 655A-12 liquid chromatograph with an L-5000 LC controller, a 655A variable-wavelength UV monitor set at 260 nm, and a D-200 Chromato-Integrator.

Nitrocellulose Filter Binding Experiments. The affinity of wild-type *Eco*RI and *Eco*RI mutants to specific and nonspecific oligonucleotides was determined in nitrocellulose filter binding experiments with [^{32}P]pTATAGAATTCTAT and [^{32}P]pTCGCGATATCGCG labeled to a specific activity of 3500 Ci/mol. These experiments were carried out at 25 °C in 0.02 M Tris·HCl, pH 7.2, 0.05 M NaCl, 0.02 M EDTA, 50 μM DTE, and 100 μg/mL bovine serum albumin, with Sartorius cellulose nitrate filters No. 11306 1251N. Filters were prewetted and rinsed with 0.02 M Tris·HCl, pH 7.2, 0.05 M NaCl, and 0.02 M EDTA. For a set of titrations the concentration of the oligonucleotide was varied in the range of the estimated K_{diss} as determined in preliminary experiments. The enzyme concentration in each titration ranged from one-fourth to 20 times the oligonucleotide concentration. The binding isotherms were analyzed by a computer program which can fit one binding constant to a set of different titrations in the same iteration process. It was not possible to obtain precise results for binding constants smaller than 1×10^4 M^{-1}, because

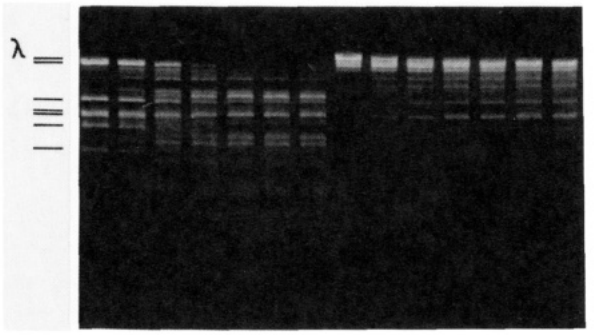

FIGURE 2: Bacteriophage λ DNA cleavage by wild-type and mutant (Lys200) *Eco*RI. λ DNA at 0.017 μg/μL was digested with 0.16 μM wild-type *Eco*RI (lanes 1–7) or 0.16 μM Lys200 mutant (lanes 8–14) for 15, 45, 90, 180, 240, 300, and 360 min, respectively, at 37 °C in 0.02 M Tris·HCl, pH 7.2, 10 mM $MgCl_2$, and 50 mM NaCl. At the high enzyme concentrations used wild-type *Eco*RI cleaves λ DNA not only within the six canonical sites but also at various degenerate sites.

Table I: Specific Activity of Wild-Type *Eco*RI and *Eco*RI Mutants

enzyme	units/mg (±25%)[a]
wild type	3×10^6 [b]
Gln144	6×10^4 [b]
Lys145	7×10^4 [b]
Lys200	$<1 \times 10^2$
Gln144Lys145	1×10^4
Gln144Lys145Lys200	nd

[a] For a definition of the specific activity, cf. Experimental Procedures. [b] Taken from Wolfes et al. (1986).

*Eco*RI and *Eco*RI mutants tend to aggregate at concentrations above 10 μM.

Results

Determination of the Activity and Specificity of Wild-Type EcoRI and EcoRI Mutants toward Cleavage of Bacteriophage λ DNA. The cleavage of bacteriophage λ DNA by *Eco*RI mutants was studied in order to determine the specific activity of these mutant enzymes and to find out whether sites other than -GAATTC- are cleaved by the mutant enzymes. Figure 2 shows the kinetics of cleavage of λ DNA by wild-type *Eco*RI and the Lys200 mutant, present in identical concentrations. While under these conditions, i.e., under normal buffer conditions but in the presence of 30000 units/mL *Eco*RI nonspecific cleavage ["star activity"; cf. Polisky et al. (1975)] is observed with wild-type *Eco*RI after several minutes (Pingoud, 1985), the reaction at the canonical site is not yet complete with the Lys200 mutant, even after 6 h. We estimate the Lys200 mutant to have an ~50000-fold lower specific activity than wild-type *Eco*RI (Table I). Furthermore, a detailed analysis of the cleavage pattern obtained with the Lys200 mutant gives no indication of cleavage at noncanonical sites.

Figure 3 shows the kinetics of cleavage of λ DNA by the Gln144Lys145 double mutant, and for comparison also that by wild-type *Eco*RI at an identical concentration. Under these conditions wild-type *Eco*RI has completed the cleavage at the canonical sites within several seconds and begins to show cleavage at noncanonical sites after several minutes, while the Gln144Lys145 mutant has cleaved λ DNA to completion only after 0.5 h. From these and similar experiments a specific activity of 1×10^4 was determined for this mutant (Table I). This value is by a factor of 300 smaller than the specific activity of wild-type *Eco*RI, but only by a factor of 6 and 7, respectively, smaller than those determined for the Gln144 and Lys145 single mutants. The Gln144Lys145 double mutant

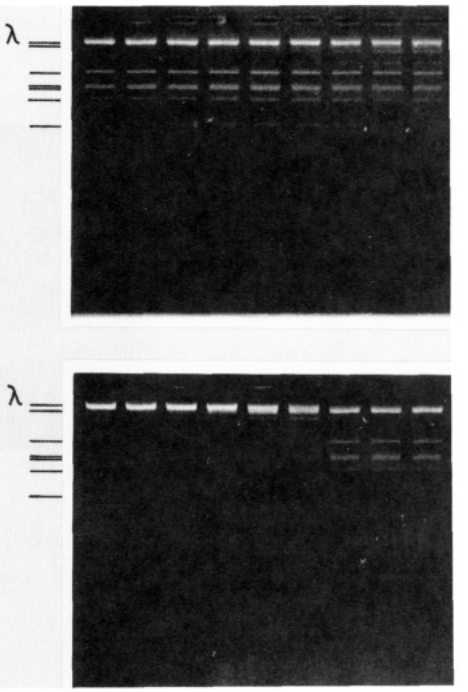

FIGURE 3: Bacteriophage λ DNA cleavage by wild-type and mutant (Gln144Lys145) *Eco*RI. λ DNA at 0.017 μg/μL was digested with 0.025 μM wild-type *Eco*RI (top) or 0.025 μM Gln144Lys145 (bottom) for 0.2, 1, 2, 4, 8, 15, 30, 60, and 120 min, respectively, at 37 °C in 0.02 M Tris·HCl, pH 7.2, 10 mM $MgCl_2$, and 50 mM NaCl. Under these conditions wild-type *Eco*RI begins to cleave λ DNA at degenerate sites after several minutes.

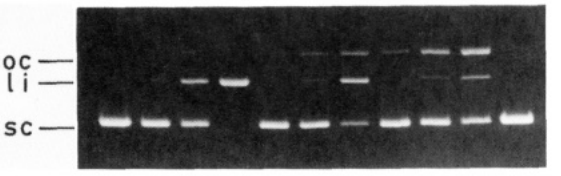

FIGURE 4: Cleavage of pUC8 DNA by wild-type *Eco*RI and *Eco*RI mutants. pUC8 at 10 nM was incubated with 0.06 nM wild-type *Eco*RI (lanes 2–4), 6 nM Gln144Lys145 double mutant (lanes 5–7), and 3 nM Lys200 mutant (lanes 8–10) for 1, 5, and 25 min, respectively. Product analysis was carried out by electrophoresis on 2% (w/v) agarose gels. Lanes 1 and 11 show the pUC8 reference.

does not cleave λ DNA at noncanonical sites, not even at 10-fold higher concentrations and after 6 h of incubation (data not shown).

Neither λ DNA nor supercoiled pUC8 DNA is cleaved detectably by the triple mutant Gln144Lys145Lys200. This means—considering the sensitivity of our assay—that the triple mutant is by more than a factor of 100 000 less active than the wild-type enzyme. It will be shown later that this mutant, however, cleaves oligodeoxynucleotides which contain the recognition site for *Eco*RI (vide infra).

Determination of Apparent K_M and V_{max} Values for the Cleavage of pUC8 DNA by Wild-Type EcoRI and EcoRI Mutants. The kinetics of cleavage of supercoiled pUC8 DNA by *EcoRI* mutants was determined at various substrate concentrations. This DNA contains only one *Eco*RI site and, therefore, is more suitable than λ DNA for a determination of K_m and V_{max} values. Cleavage was only detectable with the Lys200 mutant and the Gln144Lys145 double mutant, not, however, with the Gln144Lys145Lys200 triple mutant. Since only open circular and linear pUC8 DNA were obtained during the incubation of the DNA with these two mutants, it is assumed that only the single cognate site in pUC8 is attacked by the mutant enzymes.

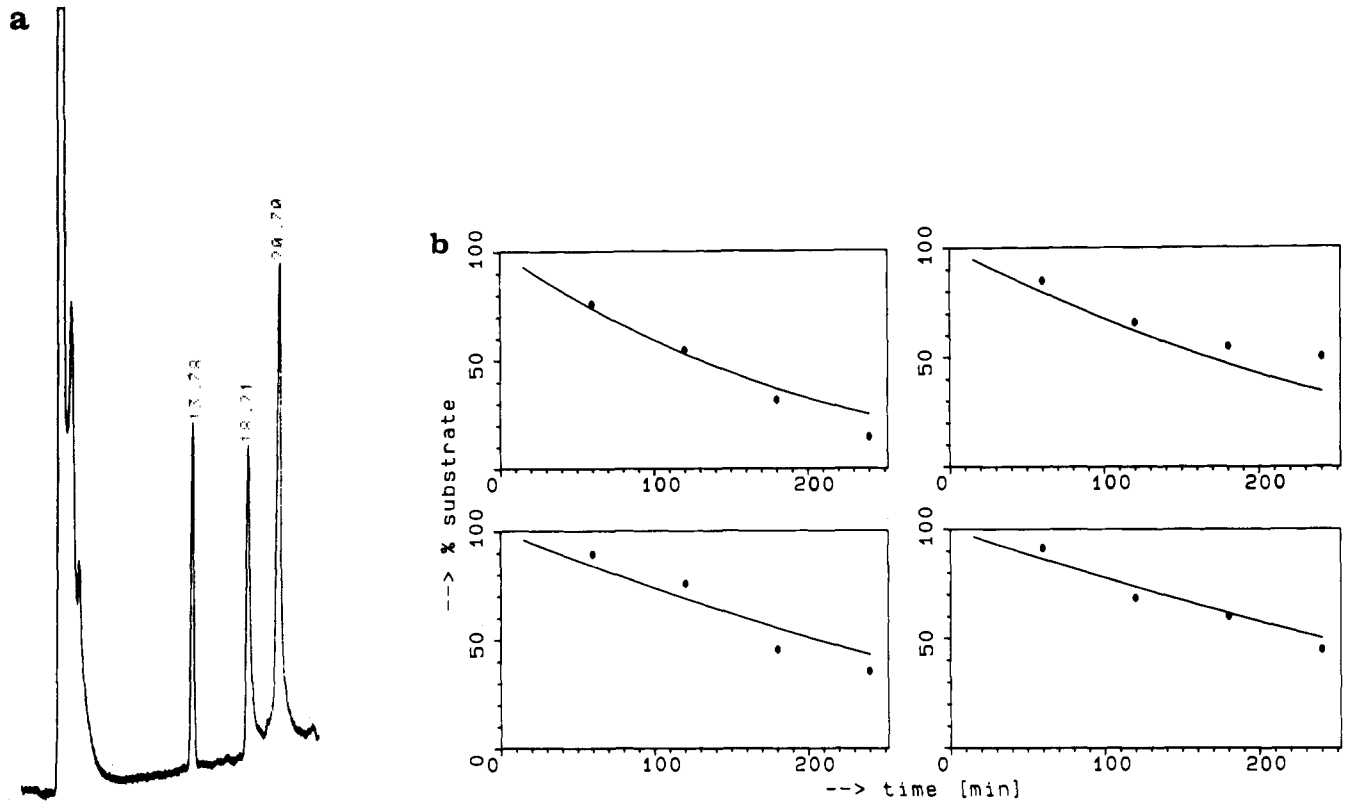

FIGURE 5: Cleavage of d(TATAGAATTCTAT) by mutant (Gln144Lys145Lys200) *Eco*RI. (a) Double-stranded d(TATAGAATTCTAT) (4 μM) was incubated at 25 °C for 4 h with 0.5 μM Gln144Lys145Lys200 triple mutant. A total of 20 μL of the reaction mixture was analyzed by HPLC as described under Experimental Procedures. The HPLC run was monitored at 260 nm with 0.1 absorbance unit full scale. The peak observed at 20.7 min represents uncleaved substrate, while the peaks at 13.78 and 18.21 min represent the products d(TATAG) and d(pAATTCTAT), respectively. (b) The kinetics of cleavage of d(TATAGAATTCTAT) by 0.5 μM Gln144Lys145Lys200 triple mutant were determined at 1 (upper left), 2 (upper right), 3 (lower left), and 4 μM (lower right) substrate concentrations and simulated according to a Michaelis–Menten scheme (cf. Experimental Procedures). The drawn out line is the theoretical curve given by a K_M of 1500 nM and a k_{cat} of 0.03 min^{-1}.

Unlike wild-type *Eco*RI but similar to the Gln144 and Lys145 mutants (Wolfes et al., 1986), the Gln144Lys145 double mutant and in particular the Lys200 mutant cleave supercoiled pUC8 such that the open circular intermediate accumulates (Figure 4). This may be due to a higher rate of dissociation of the nicked pUC8 DNA from the enzyme and/or to a lower rate of conversion to the linearized product. Both effects would favor the accumulation of the intermediate. For the Lys200 mutant which has a very high K_M for pUC8 DNA (vide infra), it is likely that the open circular DNA accumulates because of the high rate of dissociation of the intermediate from the enzyme.

For the determination of the K_M and V_{max} values for the cleavage of supercoiled pUC8, only the first phosphodiester bond cleavage which results in the open circular intermediate was considered. As can be seen from Table II the mutant enzymes are more affected in their V_{max} then in their K_M values. A similar result had been obtained for other *Eco*RI mutants with amino acid replacements in the DNA binding site (Wolfes et al., 1986).

It is interesting to note that the k_{cat}/K_M values for the cleavage of supercoiled pUC8 DNA by wild-type *Eco*RI and the various *Eco*RI mutants do not strictly correlate with the specific activities of these enzymes as determined with respect to cleavage of bacteriophage λ DNA. For example, the Gln144Lys145 double mutant cleaves pUC8 faster than the Gln144 and Lys145 mutants, but its activity toward λ DNA is smaller than those of the two single mutants. This is not a discrepancy but reflects the differences of the two assays: while with λ DNA the reaction rate of double-strand cleavage of six sites with different flanking sequences is measured by recording the time after the reaction is complete, with pUC8 DNA the initial rate of nicking of one particular site is determined.

Table II: Michaelis–Menten Parameters for the Cleavage of Supercoiled pUC8 DNA by Wild-Type *Eco*RI and *Eco*RI Mutants

enzyme	K_M (nM)	k_{cat} (min^{-1})	k_{cat}/K_M (nM^{-1} min^{-1})
wild type	50 ± 15[a]	19 ± 2[a]	0.38
Gln144	120 ± 20[a]	0.2 ± 0.02[a]	0.0017
Lys145	105 ± 20[a]	0.3 ± 0.02[a]	0.0029
Lys200	>2000[b]	>1[b]	0.00054
Gln144Lys145	55 ± 15	0.4 ± 0.1	0.0073
Gln144Lys145Lys200	nd	nd	nd

[a] Taken from Wolfes et al. (1986). [b] Saturation in the v_o vs c diagram was not reached such that only a lower limit can be given for the k_{cat} and K_M values; the k_{cat}/K_M value which is taken from the slope of the v_o vs c curve is accurate.

Determination of K_M and V_{max} Values for the Cleavage of d(TATAGAATTCTAT) by Wild-Type EcoRI and EcoRI Mutants. With a macromolecular substrate like pUC8 the apparent K_M value is a complex function of nonspecific and specific binding, as well as various other parameters which govern the ability of this enzyme to reach its recognition site by linear diffusion. In order to obtain Michaelis–Menten parameters which more directly reflect the interaction with the cognate site, we have chosen an oligonucleotide substrate for a further analysis.

We have included in this analysis also the Gln144Lys145Lys200 triple mutant, since this mutant failed to cleave λ DNA and pUC8 DNA in vitro but seemed to be

Table III: Michaelis–Menten Parameters for the Cleavage of d(TATAGAATTCTAT) by Wild-Type *Eco*RI and *Eco*RI Mutants

enzyme	K_M (nM)	k_{cat} (min^{-1})	k_{cat}/K_M (nM^{-1} min^{-1})
wild type	16 ± 3	18.0 ± 2	1.13
Gln144	400 ± 200	0.02 ± 0.002	0.00005
Lys145	400 ± 100	0.08 ± 0.005	0.0002
Lys200	1700 ± 300	39.0 ± 4	0.029
Gln144Lys145	50 ± 10	0.03 ± 0.001	0.0006
Gln144Lys145Lys200	1500 ± 300	0.03 ± 0.003	0.00002

active in an in vivo assay (see accompanying paper). As can be seen in Figure 5a, the triple mutant cleaves double-stranded d(TATAGAATTCTAT) at the correct position, albeit at a very low rate, due to an increase in K_M and a decrease in k_{cat} (Figure 2b) as compared to wild-type *Eco*RI.

In Table III the Michaelis–Menten parameters for the cleavage of this oligodeoxynucleotide substrate by wild-type *Eco*RI and several DNA binding site mutants of *Eco*RI are given. There is a good qualitative agreement between the K_M and k_{cat} values determined for the cleavage of pUC8 and the oligodeoxynucleotide. The results listed in Tables II and III demonstrate that the low activity of the Lys200 mutant is due solely to a large increase in K_M, while the reduced activity of the Gln144 and Lys145 mutants can be attributed to a moderate increase in K_M and a large decrease in k_{cat}. The Gln144Lys145 double mutant is slightly more active with pUC8 and double-stranded d(TATAGAATTCTAT) as substrates than the Gln144 and Lys145 single mutants. It seems as if the deleterious effect of one mutation is to some extent relieved by the other mutation. This is surprising, since the Gln144Lys145 double mutant has a decreased propensity to dimerize as compared to wild-type *Eco*RI and the Gln144 and the Lys145 mutants (see accompanying paper).

Binding of Wild-Type EcoRI and EcoRI Mutants to Various Oligodeoxynucleotides in the Absence of Mg^{2+} Ions. Although a preferential binding to the canonical site is not an absolute requirement for the specific cleavage of DNA by *Eco*RI (Goppelt et al., 1980; Langowski et al., 1980), it is undoubtedly beneficial for the catalytic efficiency, because it prevents *Eco*RI from being trapped in nonspecific binding. Nevertheless, nonspecific binding can be advantageous for locating the target site by linear diffusion on a macromolecular substrate (Jack et al., 1982; Ehbrecht et al., 1985; Terry et al., 1985). We have wondered, therefore, whether the very low activity of some of the mutants in particular the Lys200 mutant toward macromolecular substrates is due to a decrease in affinity toward the canonical site and/or an increase in affinity toward noncanonical sites. To answer this question, binding experiments were carried out in the absence of Mg^{2+} ions in order to prevent cleavage of the substrate. We have evidence from stopped-flow kinetic data that binding constants determined for wild-type *Eco*RI and d(TATAGAATTCTAT) in the absence of Mg^{2+} approximate those determined in its presence [J. Alves, C. Urbanke, A. Fliess, G. Maass, and A. Pingoud, unpublished results). Since the rate of cleavage of DNA by *Eco*RI depends on the sequences flanking the recognition site (Thomas & Davis, 1975; Goldstein et al., 1975; Forsblom et al., 1976; Halford et al., 1980; Berkner & Folk, 1983; Alves et al., 1984), we have chosen two oligodeoxynucleotides each containing the canonical site but having different flanking sequences, viz., d(TATAGAATTCTAT) and d(TCGCGAATTCGCG). For comparison an oligodeoxynucleotide which does not contain the *Eco*RI recognition site, viz., d(TCGCGATATCGCG), was included in this study. Figure 6 shows a set of nitrocellulose filter binding experiments

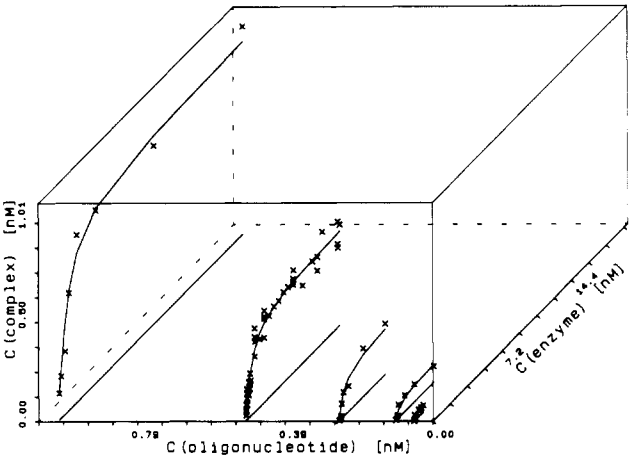

FIGURE 6: Composite diagram of the results of nitrocellulose filter binding experiments with wild-type *Eco*RI and d(pTATAGAATTCTAT). Binding experiments with wild-type *Eco*RI and double-stranded d(pTATAGAATTCTAT) were carried out as described under Experimental Procedures. Eight titrations obtained at different oligonucleotide concentrations (50, 100, 250, 500, and 1000 pM) were evaluated simultaneously by a nonlinear least-squares procedure. The titration at 500 pM was carried out four times, with two different enzyme preparations, in order to give an impression of the reproducibility of the measurements and the consistency of the lot to lot quality of the enzyme. Experimental data are indicated by (×); the theoretical curve for $n = 1$, $K_{assoc} = 2.5 \times 10^9$ M^{-1} is given by the drawn out line.

Table IV: Equilibrium Constants (M^{-1}) for the Binding of Various Oligodeoxynucleotides to Wild-Type *Eco*RI and *Eco*RI Mutants in the Absence of Mg^{2+}

	ligand		
enzyme	d(TATAGA-ATTCTAT)	d(TCGCGA-ATTCGCG)	d(TCGCGA-TATCGCG)
wild type	2 (±1) × 10^9	4 (±1) × 10^9	<1 × 10^4
Gln144	nd	1 (±0.8) × 10^7	nd
Lys145	nd	3 (±2) × 10^6	nd
Lys200	1 (±0.5) × 10^4	nd	5 (±2) × 10^5
Gln144Lys145	6 (±4) × 10^5	nd	2 (±1) × 10^5
Gln144Lys145Lys-200	4 (±3) × 10^5	nd	8 (±2) × 10^4

carried out to determine the affinity of wild-type *Eco*RI to d(TATAGAATTCTAT). All titrations were evaluated by a nonlinear least-squares fitting procedure as described under experimental procedures. Similar experiments were carried out with the other two oligodeoxynucleotides and with the *Eco*RI mutants, respectively. The results are summarized in Table IV. Wild-type *Eco*RI has a very high affinity to tridecadeoxynucleotides containing the *Eco*RI recognition site. Tridecadeoxynucleotides containing the *Eco*RV recognition site which differs in two base pairs from the *Eco*RI site are bound by 5 orders of magnitude more weakly. With smaller oligodeoxynucleotides, e.g., octadeoxynucleotides, the preference is considerably less pronounced (Goppelt et al., 1980). DNA binding site mutants of *Eco*RI all show a strongly diminished affinity toward their oligodeoxynucleotide substrate. Of particular interest is the Lys200 mutant: it has a higher affinity toward d(TCGCGATATCGCG), which it does not cleave, than to d(TATAGAATTCTAT), which it cleaves readily. These results demonstrate that specificity is not necessarily associated with preferential binding. They can also explain why the Lys200 mutant is so inefficient in cleaving high molecular weight substrates: it might indeed get trapped in nonspecific binding. We cannot offer a convincing argument why the Gln144Lys145Lys200 triple mutant shows a slight preference in binding to the specific oligodeoxynucleotide, while

the Lys200 mutant does not. The triple mutant behaves in this respect more like the Gln144Lys145 double mutant.

Discussion

In the present and a previous paper (Wolfes et al., 1986) we have studied mutants of the *Eco*RI endonuclease, viz., Glu144 → Gln, Arg145 → Lys, Arg200 → Lys, Glu144Arg145 → GlnLys, and Glu144Arg145Arg200 → GlnLysLys, with respect to their reactivity toward DNA. These five mutants carry mutations in the DNA binding site: amino acid residues that on the basis of the X-ray structure analysis are directly involved via hydrogen bonds in the specific interaction with the DNA were substituted such that the hydrogen-bond potential was decreased by 2, 4, or 6 out of 12 interactions per dimer.

It has been speculated that the reduction of the hydrogen-bonding potential at the recognition site by site-directed mutagenesis of *Eco*RI should lead to recognition of degenerate sequences, i.e., sequences other than -GAATTC-. For example, replacing Glu144 by Gln would substitute a hydrogen-bond acceptor which forms a hydrogen bond to adenine N6 (cf. Figure 1) by a hydrogen-bond donor which might form a hydrogen bond to guanine O6. This could lead to recognition of -GGGCCC- (Rosenberg et al., 1987). According to Rosenberg et al. (1987) such a prediction requires that the respective mutation does not disrupt the DNA binding site or interfere with conformational transitions involved in the catalytic process. Results presented here, in the accompanying paper, and in a previous publication (Wolfes et al., 1986) demonstrate that the Glu144 → Gln mutation while decreasing the activity of the enzyme does not change its specificity and does not seem to affect the conformational integrity of the enzyme. Similar results were obtained for two other *Eco*RI mutants in which the other two amino acid residues postulated to be involved in DNA recognition were replaced: the Arg145 → Lys and the Arg200 → Lys mutants are strongly affected in their catalytic activity but are specific in cleaving DNA only within the site -GAATTC-. Their overall structure is unaltered. Further reduction of the hydrogen-bonding potential in the recognition site by substituting two and even three amino acid residues at a time, as in the Glu144Arg145 → GlnLys double mutant and in the Glu144Arg145Arg200 → GlnLysLys triple mutant, also does not lead to a noticeable reduction in specificity, only to a strong impairment of the catalytic activity. In the case of the triple mutant the endonucleolytic activity can only be detected with an oligodeoxynucleotide substrate. The low activities of the double and triple mutants may in part be due to their decreased propensity to form the catalytically active dimer (see accompanying paper).

The results presented above show clearly that Glu144, Arg145, and Arg200 are involved in substrate binding and/or catalysis; they cannot, however, prove or disprove the involvement of these residues in the specific recognition process. Indeed, our finding that some of the mutants we have investigated are impaired mainly in their k_{cat}, like the Glu144 → Gln, the Arg145 → Lys, and the Glu144Arg145 → GlnLys mutants, another one only in its K_M, like the Arg200 → Lys mutant, and again another one equally in its K_M and k_{cat}, like the Glu144Arg145Arg200 → GlnLysLys mutant, demonstrates that the contribution of the individual residues to the catalytic process must be quite different. Arg200 probably is involved mainly in binding to specific as well as to nonspecific DNA, since its replacement by lysine does not decrease k_{cat} but increases K_M and K_{diss}. Glu144 and Arg145 are presumably mainly involved in stabilizing the transition state, because their conservative replacement has a greater effect on k_{cat} than on K_M and the effect on K_{diss} is much smaller than that with the Arg200 → Lys mutant.

The crystallographic results concerning the *Eco*RI–oligodeoxynucleotide complex have been reported to a resolution of 3.0 Å. At this resolution, it is generally possible to trace the polypeptide backbone and to identify some of the amino acid side chains. It is not possible, however, to delineate unequivocally the location of each atom in the backbone or in the side chains. Assignments of interactions between individual atoms of the protein and the ligand, therefore, must be considered as tentative, and even more so, of course, any recognition model derived from such tentatively assigned interactions. Furthermore, it has to be emphasized that the *Eco*RI–oligodeoxynucleotide complex was crystallized in the absence of Mg^{2+} ions, the essential cofactor for catalysis. It constitutes an inactive conformer which isomerizes to an active enzyme upon addition of Mg^{2+} (McClarin et al., 1986). Even if it is assumed that the addition of Mg^{2+} does not alter the structure of the interface between enzyme and substrate, it is doubtful whether one can understand the specificity of *Eco*RI just by considering the structure of the *Eco*RI–oligodeoxynucleotide complex. The approach to deduce the structural basis of specificity by analyzing the structure of the protein–DNA complex has been very successful with DNA binding proteins [e.g., Harrison et al. (1988) and Otwinowski et al. (1988)]. While repressors, however, exert their function by binding, restriction enzymes work by binding *and* catalysis. Specificity, therefore, is not solely dependent on the formation of a specific complex but is at least in part achieved by additional interactions which become accessible during catalysis and are responsible for the stabilization of the transition state. This implies that only the combination of structural and kinetic investigations on wild-type *Eco*RI and *Eco*RI mutants will eventually lead to a detailed description of how this restriction enzyme achieves its high specificity. The power of this combined approach has been amply documented for several enzymes [e.g., Fersht (1987), Knowles (1987), Benkovic et al. (1988), and Carter and Wells (1988)].

Acknowledgments

We are obliged to K. Zander and R. Mull for expert technical assistance as well as E. Schuchardt for typing the manuscript.

Registry No. *Eco*RI, 80498-17-5; L-Arg, 74-79-3; L-Glu, 56-86-0; L-Gln, 56-85-9; L-Lys, 56-87-1; d(TATAGAATTCTAT), 118977-31-4; d(TCGCGAATTCGCG), 118977-29-0; d(TCGCGATATCGCG), 118977-30-3.

References

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.

Benkovic, S. J., Fierke, C. A., & Naylor, A. M. (1988) *Science 239*, 1105–1110.

Berkner, K. L., & Folk, W. R. (1983) *Anal. Biochem. 129*, 446–456.

Carter, P., & Wells, J. A. (1988) *Nature 332*, 564–568.

Ehbrecht, H.-J., Pingoud, A., Urbanke, C., Maass, G., & Gualerzi, C. (1985) *J. Biol. Chem. 260*, 6160–6166.

Fersht, A. (1987) *Biochemistry 26*, 8031–8037.

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

Geiger, R., Rüter, T., Alves, J., Fliess, A., Wolfes, H., Pingoud, V., Urbanke, C., Maass, G., Pingoud, A., Düsterhöft, A.,

& Kröger, M. (1989) *Biochemistry* (preceding paper in this issue).
Goldstein, K., Thomas, M., & Davis, R. W. (1975) *Virology 66*, 420-427.
Goppelt, M., Pingoud, A., Maass, G., Mayer, H., Köster, H., & Frank, R. (1980) *Eur. J. Biochem. 104*, 101-107.
Halford, S. E., Johnson, N. P., & Grinsted, J. (1980) *Biochem. J. 191*, 581-592.
Harrison, S. C., Anderson, J. E., Koudelka, G. B., Mondragon, A., Subbiah, S., Wharton, R. P., Wolberger, C., & Ptashne, M. (1988) *Biophys. Chem. 29*, 31-37.
Jack, W. E., Terry, B. J., & Modrich, P. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 4010-4014.
Knowles, J. R. (1987) *Science 236*, 1252-1258.
McClarin, J. A., Frederick, C. A., Wang, B.-C., Greene, P. J., Boyer, H. W., Grable, J., & Rosenberg, J. M. (1986) *Science 234*, 1526-1541.
Otwinowski, Z., Schevitz, R. W., Zhang, R.-G., Lawson, C. L., Joachimiak, A., Marmorstein, R. Q., Luisi, B. F., & Sigler, P. B. (1988) *Nature 335*, 321-329.
Pingoud, A. (1985) *Eur. J. Biochem. 147*, 105-109.
Polisky, B., Greene, P., Garfie, D. E., McCarthy, B. J., Goodman, H. M., & Boyer, H. W. (1975) *Proc. Natl. Acad. Sci. U.S.A. 72*, 3310-3314.
Rosenberg, J. M., Wang, B.-C., Frederick, C. A., Reich, N., Greene, P. J., Grable, J., & McClarin, J. (1987) in *Protein Engineering* (Oxender, D. L., & Fox, C. F., Eds.) pp 237-250, Liss, New York.
Terry, B. J., Jack, W. E., Rubin, R. A., & Modrich, P. (1985) *J. Biol. Chem. 260*, 13130-13137.
Thomas, M., & Davis, R. W. (1975) *J. Mol. Biol. 91*, 315-328.
Wolfes, H., Alves, J., Fliess, A., Geiger, R., & Pingoud, A. (1986) *Nucleic Acids Res. 14*, 9063-9080.

Random and Site-Directed Mutagenesis of Bacterial Luciferase: Investigation of the Aldehyde Binding Site†

Lorenzo H. Chen and Thomas O. Baldwin*
Department of Biochemistry and Biophysics, Texas A&M University, and Texas Agricultural Experiment Station, College Station, Texas 77843

Received September 8, 1988; Revised Manuscript Received November 18, 1988

ABSTRACT: Numerous luciferase structural gene mutants of *Vibrio harveyi* have been generated by random mutagenesis and phenotypically characterized [Cline, T. W., & Hastings, J. W. (1972) *Biochemistry 11*, 3359-3370]. All mutants selected by Cline and Hastings for altered kinetics in the bioluminescence reaction had lesions in the α subunit. One of these mutants, AK-20, has normal or slightly enhanced thermal stability and enhanced $FMNH_2$ binding affinity but a much-reduced quantum yield of bioluminescence and dramatically altered stability of the aldehyde–C4a-peroxydihydroflavin–luciferase intermediate (IIA), with a different aldehyde chain length dependence from that of the wild-type luciferase. To better understand the structural aspects of the aldehyde binding site in bacterial luciferase, we have cloned the *lux*AB genes from the *V. harveyi* mutant AK-20, determined the nucleotide sequence of the entire *lux*A gene, and determined the mutation to be TCT → TTT, resulting in a change of serine → phenylalanine at position 227 of the α subunit. To confirm that this alteration caused the altered kinetic properties of AK-20, we reverted the AK-20 *lux*A gene by oligonucleotide-directed site-specific mutagenesis to the wild-type sequence and found that the resulting enzyme is indistinguishable from the wild-type luciferase with respect to quantum yield, $FMNH_2$ binding affinity, and intermediate IIA decay rates with 1-octanal, 1-decanal, and 1-dodecanal. To investigate the cause of the AK-20 phenotype, i.e., whether the phenotype is due to loss of the seryl residue or to the properties of the phenylalanyl residue, we have constructed mutants with alanine, tyrosine, and tryptophan at α227. The luciferases with α227 Tyr and α227 Trp exhibited kinetic and bioluminescence properties comparable to those of the AK-20 luciferase (α227 Phe) and distinctly different from those of the wild-type luciferase (α227 Ser), whereas luciferase with α227 Ala was similar to the wild-type enzyme in kinetics and quantum yield, suggesting that the properties of AK-20 must be due more to the introduction of the bulky aromatic group at α227 than to the loss of the serine hydroxyl group at that position. The results obtained here do not support a role of αSer_{227} in catalysis, since the αAla_{227} mutant has nearly wild-type activity.

Bacterial luciferase is a flavin monooxygenase that catalyzes the oxidation of a long-chain aliphatic aldehyde to the corresponding carboxylic acid (Scheme I). During the course of this reaction, an enzyme-bound excited state of the flavin is generated that, upon decay, emits blue-green light (λ_{max} = 490 nm). The enzyme is a heterodimer with a single active center thought to reside primarily on the α subunit, although the β subunit is absolutely required for activity [see Ziegler and Baldwin (1981) for a review].

Investigation of the mechanism of action of many enzymes has been enhanced substantially by analysis of mutant forms of the proteins. Although the structural details of proteins provided by high-resolution X-ray crystallographic data are extremely valuable, the data are seldom sufficient to allow one to draw conclusions regarding the mechanisms by which enzymes carry out their functions. To date, we do not have a

† This research was supported in part by grants from the National Science Foundation (NSF DMB 85-10784) and the Robert A. Welch Foundation (A865). The support of the Texas Agricultural Experiment Station is gratefully acknowledged.

* Address correspondence to this author at the Department of Biochemistry and Biophysics, Texas A&M University.

0006-2960/89/0428-2684$01.50/0