Diekmann, S., & McLaughlin, L. W. (1988) *J. Mol. Biol. 202*, 823-834.
Diekmann, S., von Kitzing, E., McLaughlin, L., Ott, J., & Eckstein, F. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 8257-8261.
Froehler, B. C., & Matteucci, M. D. (1983) *Nucleic Acids Res. 11*, 8031-8036.
Griffith, J., Bleyman, M., Rauch, C. A., Kitchin, P. A., & Englund, P. T. (1986) *Cell (Cambridge, Mass.) 46*, 717-724.
Gupta, G., Sarma, M. H., & Sarma, R. H. (1988) *Biochemistry 27*, 7909-7919.
Hagerman, P. J. (1985) *Biochemistry 24*, 7033-7037.
Koo, H.-S., & Crothers, D. M. (1987) *Biochemistry 26*, 3745-3748.
Koo, H.-S., Wu, H.-M., & Crothers, D. M. (1986) *Nature (London) 321*, 501-506.
Letsinger, R. L., & Lunsford, W. B. (1976) *J. Am. Chem. Soc. 98*, 3655-3661.
Lilley, D. M. J. (1986) *Nature (London) 320*, 487-488.
Marini, J. C., Levene, S. D., Crothers, D. M., & Englund, P. T. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 7664-7668.
Matteucci, M. D., & Caruthers, M. H. (1981) *J. Am. Chem. Soc. 103*, 3185-3191.
McBride, L. J., Kierzek, R., Beaucage, S. L., & Caruthers, M. H. (1986) *J. Am. Chem. Soc. 108*, 2040-2048.
Nelson, H. C. M., Finch, J. T., Luisi, B. F., & Klug, A. (1987) *Nature (London) 330*, 221-226.
Seela, F., & Kehne, A. (1983) *Liebigs Ann. Chem.*, 876-884.
Seela, F., & Kehne, A. (1985) *Tetrahedron 41*, 5387-5392.
Seela, F., & Kehne, A. (1987) *Biochemistry 26*, 2232-2238.
Seela, F., & Driller, H. (1988) *Helv. Chim. Acta 71*, 1191-1198.
Sinha, N. D., Biernat, J., McManus, J., & Köster, H. (1984) *Nucleic Acids Res. 12*, 4539-4557.
Trifonov, E. N. (1985) *CRC Crit. Rev. Biochem. 19*, 89-106.
Ulanovsky, L., & Trifonov, E. N. (1987) *Nature (London) 326*, 720-722.
Ulanovsky, L., Bodner, M., Trifonov, E. N., & Chodev, M. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 862-866.
User Manual, 1986 Ed. (1986) Applied Biosystems, Weiterstadt, West Germany.
Zemlicka, J., & Holy, A. (1967) *Collect. Czech. Chem. Commun. 32*, 3159-3167.

Discrimination between DNA Sequences by the *Eco*RV Restriction Endonuclease†

John D. Taylor and Stephen E. Halford*
Department of Biochemistry, School of Medical Sciences, University of Bristol, Bristol BS8 1TD, U.K.
Received December 19, 1988; Revised Manuscript Received April 19, 1989

ABSTRACT: The *Eco*RV restriction endonuclease cleaves not only its recognition sequence on DNA, GATATC, but also, at vastly reduced rates, a number of alternative DNA sequences. The plasmid pAT153 contains 12 alternative sites, each of which differs from the recognition sequence by one base pair. The *Eco*RV nuclease showed a marked preference for one particular site from among these alternatives. This noncognate site was located at the sequence GTTATC, and the mechanism of action of *Eco*RV at this site was analyzed. The mechanism differed from that at the cognate site in three respects. First, the affinity of the enzyme for the noncognate site was lower than that for the cognate site, but, by itself, this cannot account for the specificity of *Eco*RV as measured from the values of k_{cat}/K_m. Second, the enzyme had a lower affinity for Mg^{2+} when it was bound to the noncognate site than when it was bound to its cognate site: this appears to be a key factor in limiting the rates of DNA cleavage at alternative sites. Third, the reaction pathway at the noncognate site differed from that at the cognate site. At the former, the *Eco*RV enzyme cleaved first one strand of the DNA and then the other while at the latter, both strands were cut in one concerted reaction. The difference in reaction pathway allows DNA ligase to proofread the activity of *Eco*RV by selective repair of single-strand breaks at noncognate sites, as opposed to double-strand breaks at the cognate site. The addition of DNA ligase to reactions with *Eco*RV made no difference to product formation at the cognate site, but products from reactions at noncognate sites were no longer detected.

The mechanism for the interaction of a protein with a specific DNA sequence must involve two processes, the recognition of the target sequence and the discrimination against alternative sequences. In principle, the recognition of the target sequence can be accounted for by the array of hydrogen-bonding functions along the edge of each base pair (bp)[1] (Seeman et al., 1976). Hydrogen bonds with the edges of the base pairs are now known to play major roles in the recognition of specific DNA sequences by many (though not all) DNA binding proteins (McClarin et al., 1986; Anderson et al., 1987; Otwinowski et al., 1988). However, it may be more difficult to account for the discrimination against alternative DNA sequences. If the alternative sequence differs from the recognition sequence by just 1 bp, it might be able to interact with

† This work was funded by the Science and Engineering Research Council.

[1] Abbreviations: AMP and ATP, adenosine 5′-phosphate and 5′-triphosphate, respectively; bp, base pair(s); BSA, bovine serum albumin; BME, β-mercaptoethanol; DMSO, dimethyl sulfoxide; DTT, dithiothreitol; EDTA, ethylenediaminetetraacetic acid; EtdBr, ethidium bromide; K_D, equilibrium dissociation constant; M_r, relative molecular mass; NAD, nicotinamide adenine dinucleotide; R/M, restriction/modification; Tris, tris(hydroxymethyl)aminomethane.

0006-2960/89/0428-6198$01.50/0

the protein with the loss of only one or two hydrogen bonds (Rosenberg & Greene, 1982; Rosenberg et al., 1987). Energies of hydrogen bonds between proteins and ligands vary substantially with the nature of the reacting partners and with their stereochemical alignment but typically fall in the range 2–15 kJ/mol (Fersht et al., 1985). The bottom of this range corresponds to a shift in an equilibrium constant by a factor of 2 and the top to a factor of 425. For some proteins such as the *cI* and *cro* repressors, the observed discrimination between alternative DNA sequences falls within this range (Ohlendorf et al., 1982; Hochschild et al., 1986). For other proteins, including that analyzed here, the discrimination factors are much larger and cannot readily be accounted for by one or two hydrogen bonds.

Type II restriction enzymes must discriminate effectively between their cognate recognition sequences and all alternative sequences on the DNA (Modrich, 1982; Bennett & Halford, 1989). In vivo, a cell carrying a R/M system would not remain viable if the nuclease had significant activity against noncognate DNA sequences. In vitro, all of the applications of restriction enzymes in the manipulation of DNA depend on this discrimination. However, the majority of the quantitative data currently available on the activity of restriction enzymes toward noncognate DNA sequences comes from studies with oligonucleotide substrates containing nucleoside analogues [for example, see Brennan et al. (1986), McLaughlin et al. (1987), and Fliess et al. (1988)]. Consequently, we report here on the activity of the *Eco*RV restriction enzyme both at its single recognition site on the plasmid pAT153 and at the noncognate DNA sequence that forms the next best site for *Eco*RV on this plasmid. These data yield a kinetic mechanism for the DNA sequence specificity of *Eco*RV.

At present, *Eco*RI forms the archetype for type II restriction enzymes (Modrich, 1982; Halford, 1983; McClarin et al., 1986). However, reactions of restriction enzymes at noncognate DNA sequences are often too slow to measure except at high concentrations of the enzyme, and the *Eco*RV enzyme is soluble at much higher protein concentrations than *Eco*RI (Luke & Halford, 1985). The *Eco*RV restriction enzyme is similar to *Eco*RI with respect to its requirement for Mg^{2+} as a cofactor for catalysis, its M_r value for the protein subunit (28 600), and its dimeric structure in solution (Luke et al., 1987). There is, however, no homology between its amino acid sequence and that of *Eco*RI (Bougueleret et al., 1984). Large amounts of the *Eco*RV restriction enzyme are available from an overproducing strain (Bougueleret et al., 1985), and it is one of the few restriction enzymes to have been crystallized (D'Arcy et al., 1985).

Experimental Procedures

Proteins. The *Eco*RV restricton enzyme was purified as described by Luke et al. (1987), except that the protein from the Blue Sepharose column was precipitated with ammonium sulfate (75% saturation). The slurry was stored at 4 °C. Prior to each series of experiments, the protein was collected from a sample of the slurry by centrifugation and resuspended in the requisite volume of buffer [10 mM K_2HPO_4, 100 mM NaCl, 5 mM BME, 0.5 mM EDTA, and 50% (v/v) glycerol, pH 7.0]: these aliquots were stored at −20 °C. The enzyme retained full activity throughout these stages. Protein concentrations were determined by amino acid analysis (Halford & Goodall, 1988), and molarities of the *Eco*RV restriction enzyme are given in terms of the dimeric protein of M_r 57 000.

DNA ligase from bacteriophage T4, that had been purified by the method of Murray et al. (1979), was a gift from L. R. Evans (this department). *Escherichia coli* DNA ligase was from Boehringer (Mannheim, FRG). All other enzymes for DNA manipulation were either from this supplier, Amersham International p.l.c. (Amersham, U.K.), or from Anglian Biotechnology Ltd. (Colchester, U.K.) and were used as advised by the supplier. BSA (Sigma) in 10 mM Tris was adjusted to pH 7.5 and heated for 16 h at 67 °C to inactivate contaminating nucleases.

DNA. A derivative of pAT153 (Twigg & Sherratt, 1980) was constructed by inserting a *Bgl*II linker at the *Eco*RV site: we refer to this molecule as pAT153b. The following procedure was carried out as in Maniatas et al. (1982). The octadeoxynucleotide CAGATCTG (Anglian Biotechnology Ltd.) was phosphorylated with polynucleotide kinase, ligated to pAT153 that had been cut previously with *Eco*RV, and then digested with *Bgl*II. After purification by electrophoresis through agarose, this DNA was treated successively with T4 DNA ligase and *Eco*RV and then used to transform *E. coli* HB101 (Boyer & Roulland-Dussoix, 1969). Transformants were identified by resistance to ampicillin and sensitivity to tetracycline (pAT153 confers resistance to both antibiotics, but its *Eco*RV site is within the gene for tetracycline resistance). These cells contained a plasmid that was shown by restriction analysis to be identical with pAT153 except for a single *Bgl*II site in place of the *Eco*RV site.

Transformants of *E. coli* HB101 with either pAT153 or pAT153b were used as sources of these plasmids. Procedures for labeling the DNA by the incorporation of [*methyl*-^{3}H]-thymidine in vivo and for plasmid purification were as in Halford and Goodall (1988). Analysis of the purified plasmids by electrophoresis through agarose showed that generally >90% of the DNA was the covalently closed form of the plasmid in its monomeric state, with <10% as either open-circle DNA or multimers of the plasmid. (The *rec*A mutation in HB101 reduces intermolecular recombination between plasmid molecules.)

The DNA sequence of pAT153 was taken from that of pBR322 (Sutcliffe, 1979; Peden, 1983) with adjustment for the deletion of nucleotides 1649–2353 from pBR322 (Twigg & Sherratt, 1980). All sequences given here are from the "top" strand and are numbered from the first base of the sequence. Computer analysis of this sequence was with the UWGCG programs (Devereux et al., 1984).

Kinetic Methods. For reactions at >50 nM *Eco*RV nuclease, the ammonium sulfate precipitate of the enzyme was resuspended as above and an aliquot of this (1–5 μL) added to a solution of ^{3}H-labeled pAT153 or pAT153b (195–199 μL) in buffer at 20 °C. Samples (10 or 15 μL) were removed from the reaction mixture at timed intervals and immediately vortexed with 8 μL of stop-mix [0.1 M EDTA, 0.1 M Tris, 40% (w/v) sucrose, and 100 μg/mL bromophenol blue, pH 8.0]. The samples were incubated with proteinase K (1 μL of 20 mg/mL) for 20 min at 20 °C, heated at 67 °C for 20 min, and then quenched on ice (otherwise, the high concentrations of *Eco*RV enzyme interfered with electrophoresis of the DNA). An aliquot (15 μL) from each sample was then analyzed by electrophoresis through 1.2% (w/v) agarose in Tris–acetate (Johnson & Grossman, 1977). Unless noted otherwise, the electrophoresis buffer also contained 0.5 μg/mL EtdBr. After electrophoresis, slices of the gel containing either substrate DNA or the product(s) were analyzed by scintillation counting (Halford & Goodall, 1988). Reactions at <50 nM *Eco*RV nuclease were carried out as above with prior dilution of the enzyme as in Halford and Goodall (1988).

For all reactions, zero time points were taken from the DNA solution before the addition of enzyme. Further samples of

the DNA were incubated at 20 °C for the duration of the reaction, in the same buffer as the reaction but in the absence of *Eco*RV enzyme. In both sets of samples, the fraction of the DNA in its covalently closed form was measured as above. For all experiments described here, these two values were within 2% of each other. Hence, the observed reactions must be due to the *Eco*RV restriction enzyme rather than any contaminant nuclease that might have been present in the buffer.

Rate constants (either zero or first order as appropriate) were evaluated from substrate or product concentrations as a function of time by regression analysis using ENZFITTER (Elsevier-Biosoft, Cambridge, U.K.). The same program was also used for the direct fitting of data to hyperbolic functions.

Buffers. The following buffers were used: buffer A, 50 mM Tris-HCl, 100 mM NaCl, 10 mM $MgCl_2$, 10 mM BME, and 100 μg/mL BSA, pH 7.5; buffer B, same as buffer A except at pH 8.5; buffer C, 50 mM Tris-HCl, 100 mM NaCl, 10 mM $MgCl_2$, 10 mM BME, 100 μg/mL BSA, and 10% (v/v) DMSO, pH 8.5; buffer D, same as buffer C but without NaCl.

Results and Discussion

Assay Systems. In the presence of Mg^{2+} ions, the *Eco*RV restriction endonuclease cleaves both strands of the DNA at the center of its recognition sequence, GATATC (Schildkraut et al., 1984; D'Arcy et al., 1985). The enzyme can also cleave, albeit very slowly, a number of alternative sites on the DNA (Halford et al., 1986). As with other restriction enzymes (Rosenberg & Greene, 1982; George & Chirikjian, 1982; Nasri & Thomas, 1986; Barany, 1988), the DNA sequence at each additional site differs from the recognition sequence by 1 bp. For *Eco*RV, the alternative sites include 14 out of the 18 possible sequences that show 5/6 homology to GATATC: sequences where the G was replaced by either C or T (or the symmetric equivalent, replacement of C by A or G) were cleaved too slowly to count as alternative sites. In addition, two of the noncognate sites for *Eco*RV, GAGATC and GATCTC, were only cleaved on DNA that lacked *dam* methylation: both contain the sequence GATC, the target for the *dam* methylase of *E. coli* (Hattman et al., 1978). We have used the plasmid pAT153 from a *dam*$^+$ strain of *E. coli* as a substrate for the reactions of *Eco*RV at both cognate and noncognate DNA sequences. This DNA molecule (3658 bp) contains one copy of the cognate recognition sequence for *Eco*RV. It also contains 12 further sites that can be cut by *Eco*RV (Halford et al., 1986), and these include representatives of 8 out of the 12 possible DNA sequences for noncognate *Eco*RV sites on *dam*$^+$ DNA.

At low concentrations of *Eco*RV nuclease (for example, 0.5 nM enzyme), the reactions over the time scales being considered here (<5 h) occur exclusively at the cognate site (Luke et al., 1987). These were monitored by analyzing samples from the reactions by electrophoresis, in order to determine the relative amounts of the supercoiled form of pAT153 (covalently closed DNA, intact in both strands), the open-circle form (cut in one strand), and the linear form (cut in both strands). At high concentrations of *Eco*RV nuclease (for example, 1 μM enzyme), pAT153 is cleaved first at the cognate site to yield the linear form of the plasmid, but this is subsequently cleaved at other sites to yield a series of smaller fragments (Luke et al., 1987). Samples from these reactions were also analyzed by electrophoresis in order to separate the full-length linear DNA from the smaller fragments: this detects product formation only when both strands of the DNA have been cleaved at a noncognate site. The rate of the reaction was measured from the decline in the concentration of linear pAT153. This rate could be due either to the reaction at the next best site on pAT153, if one noncognate site is cleaved preferentially to the others, or to multiple reactions at several sites. We show below that the former is the case.

Table I: Specificity of the *Eco*RV Restriction Endonuclease

buffer[a]	recognition site[b] k_{cat} (min^{-1})	recognition site[b] K_m (nM)	recognition site[b] k_{cat}/K_m (M^{-1} s^{-1})	next best site[b] k_{cat}/K_m (M^{-1} s^{-1})	ratio of k_{cat}/K_m values
A (Tris, NaCl, $MgCl_2$, pH 7.5)	0.9[c]	0.5[c]	3×10^7	36	1×10^6
B (same as A, pH 8.5)	0.8	[d]		180	
C (same as B + DMSO)	0.3	2	2.5×10^6	1700	1×10^3
D (same as C − NaCl)	0.4	[d]		2150	

[a] Buffer compositions are given in full under Experimental Procedures. [b] The recognition site and the next best site refer, respectively, to the DNA sequences GATATC and GTTATC at positions 185 and 1734 in pAT153: see text. [c] Halford and Goodall (1988). [d] Not determined.

Experimental Conditions. The ability of restriction enzymes to discriminate between DNA sequences varies with the reaction conditions. This was discovered first with the *Eco*RI enzyme, relaxed specificity being observed at high pH and low salt concentrations and also in the presence of water-miscible organic solvents (Polisky et al., 1975; Tikchonenko et al., 1978). These or similar reaction conditions have since been shown to relax the specificities of many other restriction enzymes including *Eco*RV (Halford et al., 1986; Bennett & Halford, 1989). The activities of the *Eco*RV restriction enzyme against both its recognition site on pAT153 and its next best site on this DNA molecule, under a variety of reaction conditions, are given in Table I. (The experiments that yielded the values in Table I are described below.) These data define the specificity of the *Eco*RV nuclease, the specificity of an enzyme being given by the ratio of the values for k_{cat}/K_m with cognate and noncognate substrates (Fersht, 1985).

In buffer A (Tris, NaCl, and $MgCl_2$ at pH 7.5; see Experimental Procedures), the standard conditions used previously for *Eco*RV reactions (Halford & Goodall, 1988), this restriction enzyme discriminates very effectively against noncognate DNA sequences. Under these conditions, the value for k_{cat}/K_m at the next best site on pAT153 was lower than that for the cognate site by a factor 1×10^6 (Table I). However, successive alterations to buffer A, by first raising the pH, second adding DMSO, and third omitting NaCl (buffers B–D), each resulted in increased values for k_{cat}/K_m at the noncognate site (Table I). In contrast, the activity of *Eco*RV at the cognate site underwent only minor alterations with these changes in reaction conditions. Consequently, in buffer C, the value for k_{cat}/K_m at the next best site on pAT153 was lower than that for the recognition site by a factor of only 1×10^3. The change in reaction conditions between buffer A and buffer C thus caused a 1000-fold reduction in the specificity of *Eco*RV. This reduction is similar to that obtained from relative rates for the *Taq*I nuclease (Barany, 1988).

The rate at which the *Eco*RV enzyme cleaved noncognate DNA sequences in buffer A was so slow (viz., Figure 6b) that it was difficult to accumulate kinetic data under these conditions. Hence, the majority of these experiments were carried out in buffer C, where the enhanced reaction rates at noncognate sites facilitated the kinetic analysis. (Rates at noncognate sites were even higher in buffer D, but the absence of NaCl in this buffer meant that any change in the concentration of $MgCl_2$ also caused a general ionic strength effect.) However, given enough time, the *Eco*RV nuclease cleaved the same series of noncognate sites on pAT153 in buffer A (not shown), buffer C (this study), and buffer D (Halford et al.,

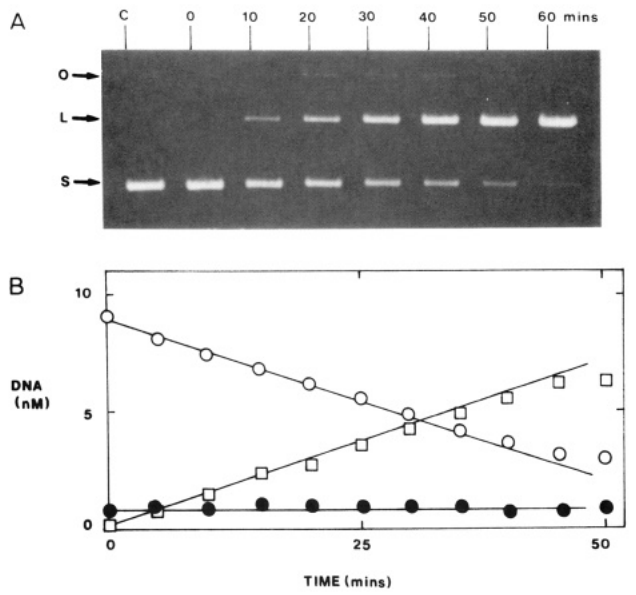

FIGURE 1: Reaction at the cognate site on pAT153. The reaction contained 0.5 nM *Eco*RV restriction endonuclease and 10 nM pAT153 in buffer C at 20 °C. (A) Samples were withdrawn from the reaction at the times indicated above the gel, quenched, and subsequently analyzed by electrophoresis through agarose. Only a representative selection of the samples that were taken from this reaction was run on the gel shown. The electrophoretic mobilities of the supercoiled, open-circle, and linear forms of pAT153 are marked as S, O, and L, respectively. (B) The concentrations of the supercoiled (○), the open-circle (●), and the linear (□) forms of pAT153 were determined relative to the total DNA at each time point during the reaction.

1986). The same processes occur in each of these buffers.

Reactions at the Cognate Site. In buffer A, the *Eco*RV enzyme simultaneously cleaves both strands of the DNA at its recognition site on pAT153 in one concerted reaction (Luke et al., 1987; Halford & Goodall, 1988). However, several restriction enzymes are known to follow different reaction pathways under different reaction conditions (Rubin & Modrich, 1978; Maxwell & Halford, 1982), and these include *Eco*RV (Halford & Goodall, 1988). Hence, comparisons of cognate and noncognate sites required data for the reaction at the cognate site in buffer C.

In buffer C, as in buffer A, the steady-state reaction of *Eco*RV at its recognition site on pAT153 resulted in the supercoiled DNA being converted directly to linear DNA: none of the open-circle DNA was formed during the reaction (Figure 1). Initial rates were measured from both the decline in the concentration of supercoiled DNA and the increase in the concentration of linear DNA with time. The rates were determined across a range of concentrations of the DNA substrate (1–20 nM), steady-state conditions being maintained with $[E_0] \ll [S_0]$ for all reactions: the concentration of $MgCl_2$ was held constant at 10 mM. Direct fitting of these data to the Michaelis–Menten equation yielded values (for this concentration of $MgCl_2$) of 0.3 min^{-1} for k_{cat} and 2 nM for K_m. Hence, in buffer C, the value for k_{cat} is slightly lower and the value for K_m slightly higher than in buffer A, with the result that k_{cat}/K_m is reduced 10-fold (Table I).

Steady-state velocities were also measured from the reactions of *Eco*RV at its cognate site on pAT153 in buffer C across a range of concentrations of $MgCl_2$ (1–10 mM) at a fixed DNA concentration: no systematic variation was observed (Figure 5b below). Hence, this system must be fully saturated with Mg^{2+} at the lowest concentration of $MgCl_2$ tested.

For the equivalent reactions in buffer A, Halford and Goodall (1988) had used a scheme involving the following forms of the DNA:

$$S{\cdot}S \overset{k_1}{\rightleftharpoons} (E{\cdot}E)\ [S{\cdot}S \xrightarrow{k_2} S{\cdot}P \xrightarrow{k_3} P{\cdot}P] \xrightarrow{k_4} P{\cdot}P \quad (1)$$

where S·S is the DNA duplex with an intact *Eco*RV site and S·P and P·P are DNA molecules cut at the *Eco*RV site in one or both strands; forms within the brackets are bound to (E·E), the protein dimer. However, the coupling between the two hydrolytic reactions (i.e., $k_3 \gg k_2$) was dependent on both subunits of the enzyme binding Mg^{2+}. Otherwise, the two reactions were no longer coupled, and k_2 was then twice as large as k_3 (Halford & Goodall, 1988). The rate-limiting step was the final dissociation of P.P., the DNA cut in both strands. Since the reactions in buffer C yielded the same profile as those in buffer A, this mechanism for cleaving the cognate site also applies to the buffer conditions used here.

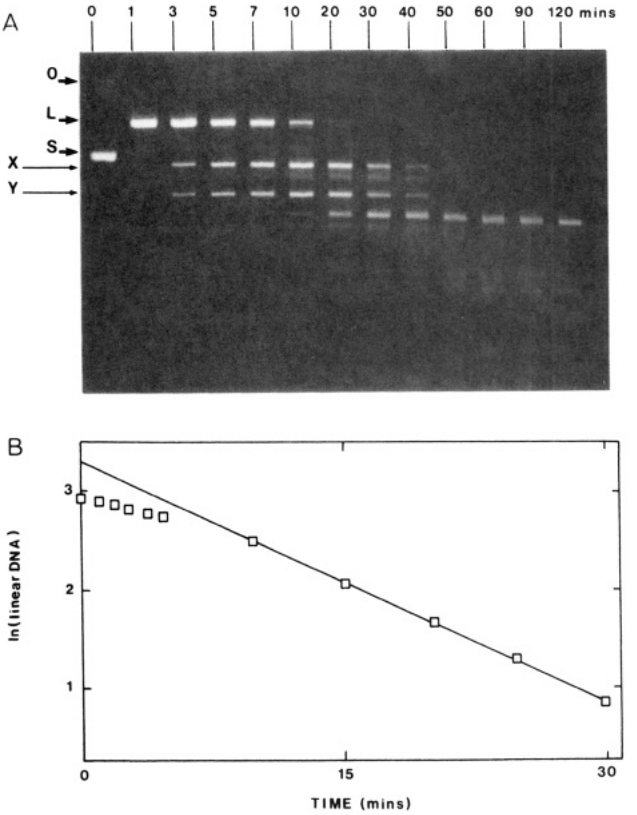

FIGURE 2: Reactions at noncognate sites on pAT153. The reaction contained 1.2 μM *Eco*RV restriction endonuclease and 10 nM pAT153 in buffer C at 20 °C. (A) Samples withdrawn from the reaction at the indicated times were analyzed as in Figure 1A. Two DNA fragments produced during this reaction are marked X and Y. (B) The concentration of the full-length linear form [(□) marked above as L] of pAT153 was determined at each time point during the reaction and is shown in semi-logarithmic format. (At zero time, the concentration given is the sum of the supercoiled and open-circle forms of the DNA.)

Reactions at Noncognate Sites. Figure 2 shows an identical reaction with that in Figure 1, except for a 2400-fold increase in the concentration of *Eco*RV restriction endonuclease. At this enzyme concentration, the recognition site for *Eco*RV on pAT153 was cleaved in both strands to give the linear form of the DNA within 1 min. The full-length linear form of pAT153 was then cleaved at one further site to give two products of approximately 2100 and 1550 bp: these are labeled in Figure 2A as X and Y. These two DNA fragments were subsequently cleaved by *Eco*RV at other sites to yield a series of smaller fragments of DNA (Figure 2A). As pAT153

contains only one copy of the cognate recognition sequence for *Eco*RV, both the conversion of the linear form of pAT153 to X and Y and the subsequent fragmentation of the DNA must be due to *Eco*RV reactions at noncognate sites. At low enzyme concentrations (Figure 1), these must have been too slow to detect.

In Figure 2A, the initial cleavage of the linear forms of pAT153 into two defined fragments, X and Y, indicates that 1 out of the 12 noncognate sites for *Eco*RV is more susceptible to the enzyme than the others. (If all 12 noncognate sites had been cleaved at similar rates, the linear form of pAT153 would have given a complicated mixture of fragments.) We refer to this as the next best site, after the recognition site, for the *Eco*RV nuclease on pAT153. The location of this site was mapped relative to the *Sal*I and *Pst*I sites on pAT153. This placed the preferred noncognate site at about position 1700 in the sequence of this plasmid. Inspection of the DNA sequence in this vicinity revealed three sequences that form noncognate sites for *Eco*RV: GGTATC at position 1704, GTTATC at position 1734, and GATAAC at position 1753. Further mapping, relative to the *Nsp*HI site at position 1769 on pAT153, identified the middle of these as the next best site for *Eco*RV: band X (Figure 2A) yielded a DNA fragment of about 40 bp when digested with *Nsp*HI (data not shown).

At present, we do not know why the DNA sequence GTTATC, at position 1734 on pAT153, is more susceptible to *Eco*RV than any other site on this molecule apart from the cognate recognition sequence, GATATC. It cannot be the sequence itself: GTTATC occurs at two other sites on the plasmid, neither of which is particularly susceptible to *Eco*RV. Moreover, the *Eco*RV modification enzyme methylates the first A within GATATC (Nwosu et al., 1988), and hence this base is likely to be critical for recognition by the *Eco*RV nuclease. However, with DNA substrates that contain several copies of the recognition sequence for a restriction enzyme, the reactivity at each site can be different (Thomas & Davis, 1975), and this could also apply to noncognate sites. With other plasmids as substrates for *Eco*RV, such as pACYC184 (Chang & Cohen, 1978) and pA03 (Oka et al., 1979), there was no one noncognate site that was preferred over the others on these DNA molecules (data not shown). With pAT153, the existence of a preferred noncognate site means that the kinetics of the reaction at an individual noncognate site can now be measured (Figure 2B).

Kinetics at the Noncognate Site. Given the following conditions, the decline in the concentration of substrate during an enzyme-catalyzed reaction should follow a single exponential: this yields a first-order rate constant that is equal to $(k_{cat}/K_m)[E_0]$. The conditions are that the enzyme obeys Michaelis–Menten kinetics and that the enzyme concentration is higher than that of the substrate but lower than the K_m value (Halford & Johnson, 1981). These apply to the reaction shown in Figure 2. However, one exponential is insufficient to describe the decline in the concentration of the substrate for the reaction at the noncognate site (the full-length linear form of pAT153 generated during the first minute of the reaction). The semi-log plot in Figure 2B shows a lag phase preceding the exponential phase. The simplest explanation for the lag phase, confirmed below, is that it corresponds to the cleavage of one strand of the DNA at the noncognate site: this yields a nicked linear DNA molecule that has the same electrophoretic mobility as the linear DNA substrate. The exponential phase corresponds to the cleavage of the second strand of the DNA at this noncognate site.

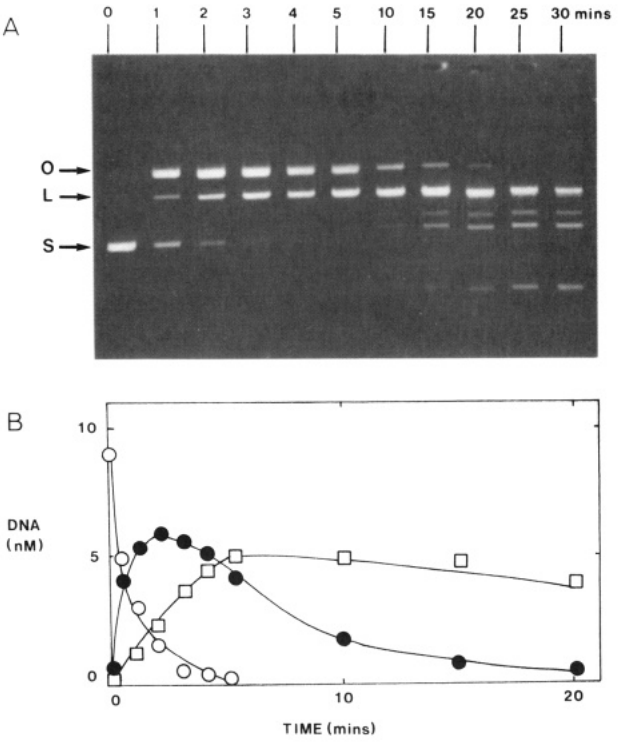

FIGURE 3: Reactions at noncognate sites on pAT153b. The reaction contained 1.2 μM *Eco*RV restriction endonuclease and 10 nM pAT153b in buffer C at 20 °C. (A) Samples withdrawn from the reaction were analyzed as in Figure 1A. (B) The concentrations of the supercoiled (○), open-circle (●), and full-length linear (□) forms of pAT153b were determined at each time point during the reaction: the electrophoretic mobilities of these forms are marked above as S, O, and L, respectively.

The first-order rate constant from the exponential phase in Figure 2B, when normalized against the enzyme concentration, is therefore equal to k_{cat}/K_m for double-strand breaks at this particular site. The values for k_{cat}/K_m for the *Eco*RV enzyme at its next best site on pAT153 (Table I) were determined by this method from reactions in buffers A–D. In buffers A and B, the reactions were monitored over a longer time base (typically 5 h). In each case, the values were checked by repeating the measurements at several different enzyme concentrations. The kinetic data in Table I thus refer to the formation of product cleaved by *Eco*RV in both strands of the DNA, at either the single recognition site on pAT153 (as in Figure 1) or the noncognate site at position 1734 in the DNA sequence (as in Figure 2), this being preferred to the other noncognate sites under all conditions tested (data not shown). The other noncognate sites must yield k_{cat}/K_m values that are even lower than those reported in Table I for the preferred noncognate site.

Mechanism at the Noncognate Site. In order to analyze the mechanism of action of the *Eco*RV nuclease at its preferred noncognate site, we constructed a derivative of pAT153 by inserting 8 bp of DNA in the middle of the cognate recognition sequence for *Eco*RV (Experimental Procedures). This derivative, pAT153b, no longer contains the *Eco*RV recognition sequence, and no new noncognate sites were generated by the insertion. Hence, any cleavage of pAT153b by the *Eco*RV restriction enzyme must be at one or more of the noncognate sites present on the parent plasmid.

A reaction of the *Eco*RV nuclease at the same high enzyme concentration as in Figure 2, with pAT153b as the substrate, is shown in Figure 3. In this reaction, the supercoiled form of the plasmid was converted initially to its open-circle form

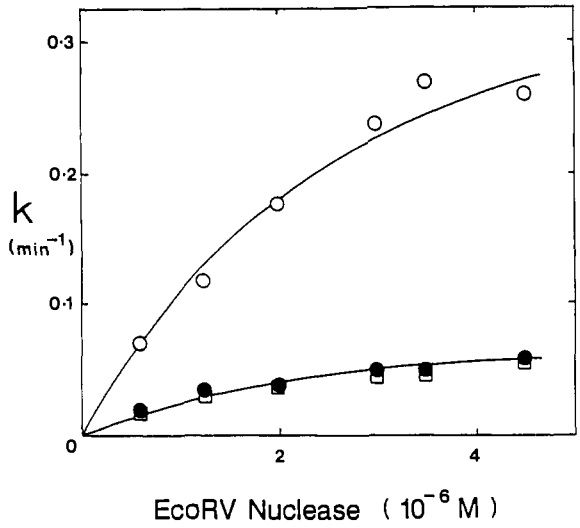

FIGURE 4: Enzyme concentration dependence. Reactions at 20 °C contained *Eco*RV restriction endonuclease (at the concentration indicated) and 10 nM DNA (either pAT153 or pAT153b) in a buffer identical with buffer C except the concentration of $MgCl_2$ was 2 mM. With pAT153 as the substrate, apparent rate constants (□) were evaluated from the exponential phase in the decline in the concentration of the full-length linear form of the DNA. With pAT153b as the substrate, values for the apparent rate constants k_a (○) and k_b (●) were determined from the decline in the concentrations of the supercoiled and the open-circle forms of the DNA, respectively. The lines drawn are theoretical lines from eq 3 with the following parameters: K_s at 3 μM, $K_c \gg$ [M], and k_h[M]/K_c at either 0.45 min^{-1} (upper line) or 0.09 min^{-1} (lower line).

by nicking one strand of the DNA, then to its full-length linear form by cleaving both strands at the same site, and finally to a series of smaller DNA fragments. The sites on pAT153b that had been cleaved in one or both strands, to yield the open-circle and linear forms, respectively, were identified by restriction mapping: both reactions had occurred at the same position as the preferred noncognate site on pAT153 (data not shown). For mapping the position of the nick on the open-circle DNA, we used alkaline agarose gels where the DNA runs in its single-stranded form (McDonnell et al., 1977): we did not determine whether one strand was cleaved preferentially to the other.

The reaction in Figure 3B was fitted to two first-order rate constants: one for the decline in the concentration of the supercoiled form of pAT153b (starting from zero time) and a second for the decline in the concentration of the open-circle DNA (starting from a time point subsequent to where the open-circle form was at its maximum during the reaction). This yields two apparent rate constants, k_a and k_b, respectively, for the *Eco*RV reactions that cleave the first and the second strands of the DNA at its preferred noncognate site. In the reaction shown, the value for k_a (0.56 min^{-1}) was 4 times larger than that for k_b (0.14 min^{-1}). The reactions of the *Eco*RV enzyme on pAT153b were also carried out at a series of different enzyme concentrations (at a fixed concentration of $MgCl_2$) and at a series of different $MgCl_2$ concentrations (at a fixed concentration of enzyme). All of these reactions yielded profiles similar to that in Figure 3. The values for k_a and k_b varied considerably with the concentrations of $MgCl_2$ and enzyme (Figures 4 and 5a), yet the ratio of k_a/k_b was always about 4/1. With the parent pAT153 as the substrate, we also measured the apparent rate constants for double-strand breaks at the preferred noncognate site as in Figure 2B: over the range of enzyme and $MgCl_2$ concentrations tested, these agreed closely with the values for k_b from equivalent reactions with pAT153b (Figures 4 and 5a).

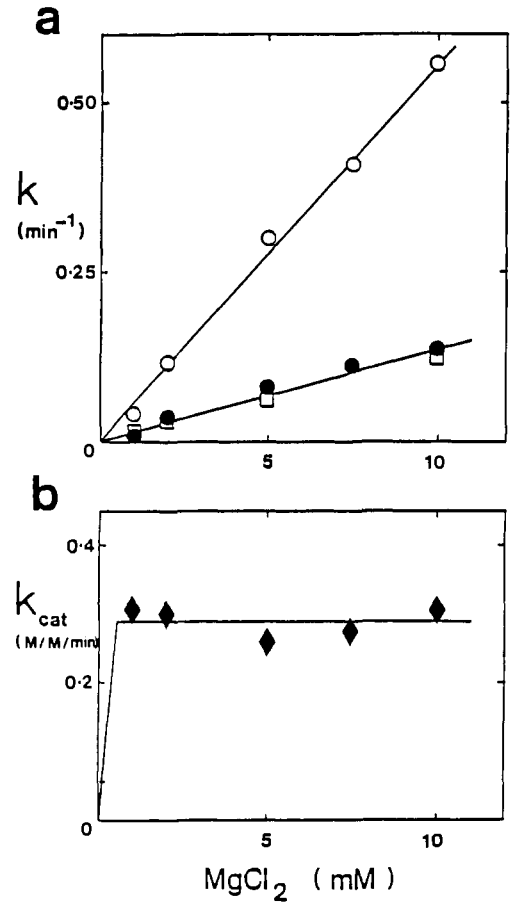

FIGURE 5: [$MgCl_2$] dependence at noncognate (a) and cognate (b) sites. (a) Reactions at 20 °C contained 1.2 μM *Eco*RV restriction endonuclease and 10 nM DNA (either pAT153 or pAT153b) in a buffer identical with buffer C except for variations in the concentration of $MgCl_2$ as indicated. With pAT153 as the substrate, apparent rate constants (□) were evaluated from the exponential phase in the decline in the concentration of the full-length linear form of the DNA. With pAT153b as the substrate, values for the apparent rate constants k_a (○) and k_b (●) were determined from the decline in the concentrations of the supercoiled and the open-circle forms of the DNA, respectively. (b) Reactions contained 0.5 nM *Eco*RV restriction endonuclease and 10 nM pAT153 but were otherwise identical with (a). Values for k_{cat} (♦) were determined from the steady-state reaction velocities.

The data in Figures 4 and 5a are consistent with a compulsory order mechanism:

$$E + S + M \overset{K_s}{\rightleftharpoons} E{\cdot}S \overset{K_c}{\rightleftharpoons} E{\cdot}S{\cdot}M \xrightarrow{k_h} E{\cdot}P \qquad (2)$$

where S is the target site on the DNA and M is Mg^{2+}: K_s and K_c are the K_D values for binding the DNA and the magnesium cation, respectively, and k_h is the rate constant for the hydrolytic reaction. If the bimolecular steps are faster than the hydrolytic reaction and if the initial concentration of substrate is lower than that of the enzyme (as in our reactions), the apparent rate constant for product formation (either k_a or k_b, depending on whether eq 2 refers to the cleavage of the first or the second strand) is then given by

$$k_a = \frac{k_h[E][M]}{[E]([M] + K_c) + K_sK_c} \qquad (3)$$

[The data in Figures 4 and 5a eliminate an alternative mechanism in which the enzyme binds Mg^{2+} before the DNA. For this alternative, an equation analogous to eq 3 was derived, but it predicted concentration dependencies that were inconsistent with the experimental data (not shown).]

For the dependence of either k_a or k_b on the concentration of the enzyme, at a fixed concentration of $MgCl_2$ (Figure 4), eq 3 predicts a hyperbolic relationship, where the maximum value of k_a (or k_b) is equal to $k_h[M]/(K_c + [M])$ and the enzyme concentration that gives half of this maximum is equal to $K_sK_c/(K_c + [M])$. When the value of K_c is much larger than the concentration of $MgCl_2$ (which is the case for these reactions, see below), these two terms simplify to $k_h([M]/K_c)$ and K_s, respectively. The dependence of the rates on the concentration of *Eco*RV fitted closely to a hyperbola (Figure 4), with maximum values for k_a and k_b of 0.45 and 0.09 min^{-1}, and respectively, and an enzyme concentration for the half-maximal rate (from either k_a or k_b) at 3 μM. We equate the latter with K_s, the equilibrium dissociation constant of the *Eco*RV nuclease from its preferred noncognate site on pAT-153.

For the dependence of either k_a or k_b on the concentration of $MgCl_2$, at a fixed concentration of the *Eco*RV enzyme (Figure 5a), eq 3 again predicts a hyperbolic relationship. In this case, the maximum values for the rates will be equal to k_h, and the concentration of $MgCl_2$ for the half-maximal rate equal to $K_c(1 + K_s/[E])$. However, the values of both k_a and k_b increased linearly with increasing concentrations of $MgCl_2$ over the range tested (up to 10 mM) (Figure 5a). Hence, for these reactions, the parameter $K_c(1 + K_s/[E])$ must be very much larger than 10 mM. Given the value for K_s determined from Figure 4 (3 μM), and the fixed enzyme concentration used for the experiments in Figure 5a (1.2 μM), it then follows that K_c itself must be very much larger than 3 mM. It was impossible to confirm this estimate for K_c by extending these measurements to higher concentrations of $MgCl_2$ because the reaction rates were then influenced by general ionic strength effects. However, we cannot find any other pair of values, other than those given here (K_s at about 3 μM, $K_c \gg 3$ mM), that are internally consistent with the data in both Figure 4 and Figure 5a.

We can now compare the dependencies on $MgCl_2$ concentration for *Eco*RV reactions at either its preferred noncognate site or its cognate site on pAT153 (Figure 5a and Figure 5b, respectively). Across a range of concentrations of $MgCl_2$ where the rates of the reaction at the noncognate site *increased linearly* with increasing [$MgCl_2$] (Figure 5a), the rate of the reaction at the cognate site *showed no increase* at all (Figure 5b). The latter is given as k_{cat} values, measured from steady-state kinetics: eq 5 of Halford and Johnson (1981) gives the dependence of k_{cat} on [$MgCl_2$]. Hence, for the reaction at the cognate site, the lowest concentration of $MgCl_2$ tested is in large excess of K_c. Across the range 1–10 mM $MgCl_2$, the *Eco*RV enzyme is fully saturated with Mg^{2+} during its reaction at the cognate site, yet, under the same conditions, it has a very low degree of saturation with Mg^{2+} during the reaction at the noncognate site.

Differences between Cognate and Noncognate Sites. The reaction of the *Eco*RV restriction enzyme at its preferred noncognate site on pAT153 differs in three respects from that at its recognition site. First, the reactions proceed by different pathways. At the cognate site, both strands of the DNA were cleaved in a single concerted reaction (Figure 1; Halford & Goodall, 1988). At the noncognate site, the two strands were cleaved in sequential reactions with separate kinetics (Figure 3). All of our data on reactions at noncognate sites came from experiments where the enzyme was in large molar excess over the plasmid, so it is possible that the DNA nicked at the noncognate site remains associated with the enzyme (as in eq 1). However, the dependence of both k_a and k_b on the enzyme concentration indicates otherwise (Figure 4), and we have confirmed that the nicked DNA dissociates from the enzyme between the two hydrolytic reactions (see below). Barany (1987) had reported previously a similar difference for the *Taq*I restriction enzyme.

The second difference is in the affinities of the enzyme for cognate and noncognate sites. In buffer C, the cognate site yielded a K_m value of 2 nM (Table I), and though K_m values are not a direct measure of the affinity of an enzyme for its substrate, it remains far below the value of 3 μM for K_s at the noncognate site (Figure 4). However, this may be insufficient to account for the large difference in k_{cat}/K_m between cognate and noncognate sites (Table I). The fact that product dissociation is rate limiting for the cognate site (Halford & Goodall, 1988) creates an opportunity for compensation between k_{cat} and K_m (Fersht, 1985). The product formed at a noncognate DNA sequence might dissociate from the enzyme more rapidly than that formed at the cognate site. If product dissociation remained rate limiting, the reaction at the noncognate site could have a higher k_{cat} than that at the cognate site.

The third difference is in the affinity of the enzyme for Mg^{2+}, depending on whether the protein is bound to its cognate site or to a noncognate site on the DNA. The reaction of the *Eco*RV nuclease at its recognition site on pAT153 yields a much lower value for K_c than that at the next best site (Figure 5). This appears to be a critical factor in determining the DNA sequence specificity of the *Eco*RV restriction enzyme. Even when the enzyme is bound to a noncognate site on the DNA, it still cannot cleave the DNA at that site unless it has also bound Mg^{2+} ions. The specificity of the *Eco*RI restriction enzyme is thought to involve a process called "allosteric activation" (McClarin et al., 1986; P. Modrich, personal communication): the recognition of the correct DNA sequence may trigger a conformational change within the DNA–protein complex that aligns the catalytic functions of the system, whereas an incorrect DNA sequence fails to promote the organization of the catalytic center. Our data on *Eco*RV are fully consistent with this proposal. The different affinities for Mg^{2+} indicate that the organization of the catalytic center is dependent on the DNA sequence that is bound to the enzyme.

Cleavage of the noncognate site by two separate reactions, as opposed to a single concerted reaction at the cognate site, is a direct consequence of the different affinities for Mg^{2+}. The reaction at the cognate site also involves two separate reactions whenever the system is not fully saturated with Mg^{2+} (Luke et al., 1987; Halford & Goodall, 1988), but in this case, k_a was twice k_b instead of the 4-fold difference seen at the noncognate site (Figure 3). The change in this ratio can be accounted for by different statistical factors that arise from the binding of a ligand to a dimeric protein. If each subunit of the dimer binds the ligand with the same intrinsic K_D, then the apparent K_D for binding one molecule of the ligand to either subunit will be half the intrinsic K_D, that for binding one molecule to a specified subunit equal to the intrinsic K_D, and that for binding the second molecule twice the intrinsic K_D (Edsall & Gutfreund, 1983). This model predicts that the *Eco*RV enzyme requires Mg^{2+} to be bound to both subunits of the dimer in order to cleave the second strand at a noncognate site, while the equivalent reaction at the cognate site requires Mg^{2+} on one specified subunit.

Proofreading by DNA Ligase. Nicked DNA is formed as an intermediate during *Eco*RV reactions at noncognate sites, but it is not formed in equivalent reactions at the cognate site (Figures 1 and 3). This raises the possibility that DNA ligase might be able to edit the specificity of the *Eco*RV restriction

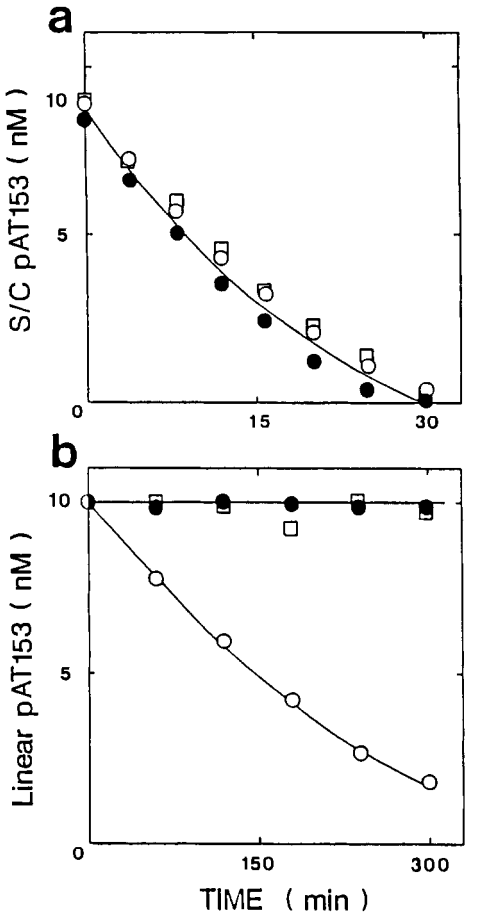

FIGURE 6: Proofreading by DNA ligase. Reactions at 20 °C in 50 mM Tris, 10 mM $MgCl_2$, 10 mM DTT, 1.3 mM ATP, and 100 μg/mL BSA, pH 7.5, contained 10 nM pAT153, *Eco*RV restriction endonuclease at either 0.5 nM (a) or 3 μM (b), and either no DNA ligase (○) or 10 nM DNA ligase from phage T4 (□) or 10 nM *E. coli* DNA ligase (●). For reactions containing *E. coli* DNA ligase, the ATP was replaced by 1.3 mM NAD. (a) With 0.5 nM *Eco*RV nuclease, the concentration of the supercoiled form of pAT153 (the substrate) was measured over the time scale given. (b) With 3 μM *Eco*RV nuclease, the concentration of the full-length linear form of pAT153 (the substrate generated during the first minute of this reaction) was measured at timed intervals thereafter. For this reaction, the concentration of DNA substrate given at zero time is the total concentration of DNA in the reaction.

enzyme by selectively repairing DNA cut at noncognate sites in preference to DNA cut at the cognate site. DNA ligase catalyzes the synthesis of phosphodiester bonds, either to repair nicks in DNA duplexes or to join separate termini (Engler & Richardson, 1982). Hence, in the presence of DNA ligase, the forms of the DNA generated during *Eco*RV reactions can be interconverted as shown in eq 4 (for the *Eco*RV recognition site) and eq 5 (for a noncognate site). Forms of the DNA

$$\text{S·S} \underset{w_1}{\overset{v_1}{\rightleftharpoons}} \text{P·P} \tag{4}$$

$$\text{S·S} \underset{w_2}{\overset{v_2}{\rightleftharpoons}} \text{S·P} \underset{w_3}{\overset{v_3}{\rightleftharpoons}} \text{P·P} \tag{5}$$

are defined as in eq 1: v_1, v_2, and v_3 and w_1, w_2, and w_3 are the reaction velocities determined by the concentrations of *Eco*RV nuclease and of DNA ligase, respectively, that are present in the mixture. For *Eco*RV, v_1 will be much greater than v_2 which in turn is 4 times greater than v_3. DNA ligase has a much higher activity at repairing nicks than at joining ends (Sugino et al., 1977), so w_2 will be much greater than either w_1 or w_3. Efficient proofreading by ligase demands that $v_1 > w_1$ and that $v_2 < w_2$.

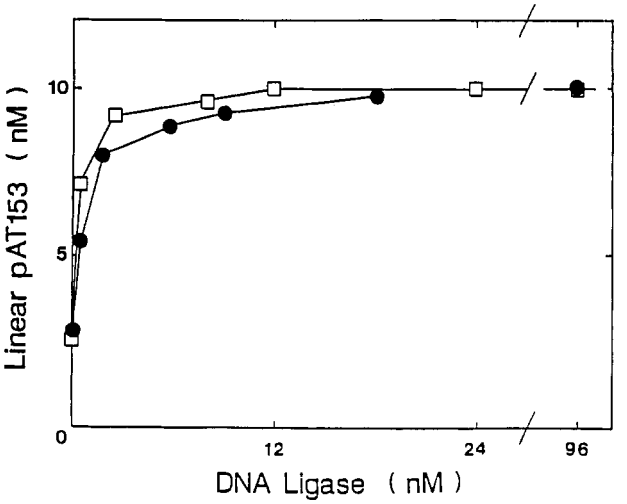

FIGURE 7: Repair at noncognate sites. Reactions contained 3 μM *Eco*RV restriction endonuclease and 10 nM pAT153 in 50 mM Tris, 10 mM $MgCl_2$, 10 mM DTT, and 100 μg/mL BSA, pH 7.5, with either T4 DNA ligase and 1.3 mM ATP (□) or *E. coli* DNA ligase and 1.3 mM NAD (●). The concentrations of DNA ligase are given on the abscissa. After 5 h at 20 °C, the concentration of the full-length linear form of pAT153 was measured in each reaction.

To test this model, we analyzed the reactions of the *Eco*RV nuclease, at both its cognate and its noncognate sites on pAT153, in the presence of DNA ligase. Two ligases were used, that encoded by phage T4 and that from *E. coli*. The two enzymes differ in their cofactor requirement, T4 ligase using ATP and *E. coli* ligase NAD (Engler & Richardson, 1982): neither ATP nor NAD, at the concentrations used here, had any effect on the activities of *Eco*RV at either cognate or noncognate sites (data not shown). The two ligases also differ in their relative abilities at joining blunt-ended DNA fragments: the DNA ligase from phage T4 has a much higher activity in this reaction than the *E. coli* enzyme (Sugino et al., 1977; Zimmerman & Pheiffer, 1983).

At a low concentration of the *Eco*RV nuclease, the supercoiled form of pAT153 is cleaved only at the recognition site for *Eco*RV (Figure 1). The rate at which the *Eco*RV enzyme cleaved its cognate site was unaffected by the addition of either T4 DNA ligase or *E. coli* DNA ligase (Figure 6a). The same result was also obtained when the experiments in Figure 6a were repeated with a fixed concentration of *Eco*RV nuclease (0.5 nM) but across a range of concentrations of DNA ligase (0–600 nM for the T4 enzyme, 0–450 nM for the *E. coli* enzyme) (data not shown). In these experiments, the only species of DNA seen on the gels were the supercoiled, open-circle, and linear forms of pAT153: products that might have arisen by intermolecular ligations were not detected. These results were as expected for *E. coli* DNA ligase, as this enzyme would have virtually no activity at joining the blunt ends formed by *Eco*RV reactions at the cognate site. However, the rate at which T4 ligase catalyzes blunt-end ligations must also be much lower than the rate at which *Eco*RV cleaves both strands of the DNA at its cognate site, even when this ligase is present at a 1200-fold excess over *Eco*RV.

At high concentrations of the *Eco*RV enzyme, the supercoiled form of pAT153 is converted rapidly to its linear form to yield a substrate for further reactions at noncognate sites (Figure 2). Under these conditions, in the absence of ligase, the linear DNA was cleaved by *Eco*RV into smaller fragments, but in the presence of either T4 or *E. coli* ligase, the DNA remained as its full-length linear form (Figure 6b). The amount of DNA ligase that was required to prevent the formation of any products from *Eco*RV reactions at noncognate

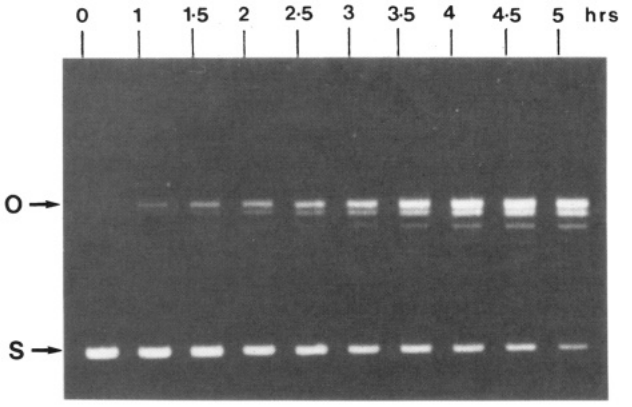

FIGURE 8: Relaxation with *Eco*RV and ligase. The reaction contained 60 nM *Eco*RV restriction endonuclease, 10 nM T4 DNA ligase, and 10 nM pAT153b in 50 mM Tris, 2 mM $MgCl_2$, 10 mM DTT, 1.3 mM ATP, 100 μg/mL BSA, and 10% DMSO, pH 8.5, at 20 °C. Samples were withdrawn from the reactions and subsequently analyzed by electrophoresis through agarose in the *absence* of EtdBr: the time of withdrawal is noted above each slot of the gel. The electrophoretic mobilities of the supercoiled and the open-circle forms of pAT153b are marked S and O by the left of the gel.

sites was also determined (Figure 7). More than 90% of the DNA remained in its full-length linear form when either ligase was present at 0.1% of the level of the *Eco*RV nuclease. These experiments were carried out under conditions similar to buffer A (modified for the requirements of DNA ligase), where the *Eco*RV restriction enzyme discriminates between cognate and noncognate DNA sequences by a factor of 1×10^6 (Table I). In the presence of ligase, *Eco*RV reactions at noncognate DNA sequences are no longer 1×10^6 times less frequent than at cognate sites but are now too infrequent to detect.

There remains the possibility that ligase may have enhanced the specificity of *Eco*RV by somehow inhibiting reactions at noncognate sites rather than by repairing nicks. However, in reactions of the *Eco*RV enzyme on pAT153b (that lacks *Eco*RV sites) in the presence of T4 DNA ligase, the supercoiled form of the plasmid retained its covalently closed structure but became topologically relaxed (Figure 8). Electrophoresis of the DNA from this reaction through agarose in the absence of EtdBr separated the product into a series of topoisomers with mobilities similar to that of the open-circle form of the DNA (Figure 8). The relaxed covalently closed structure of this DNA was confirmed by electrophoresis in the presence of chloroquine (Shure et al., 1977). The relaxation of pAT153b was not due to the topoisomerase activity of DNA ligase (Modrich et al., 1972), as no relaxation was observed in the absence of *Eco*RV (not shown): ligase functions as a topoisomerase only in the presence of AMP (Modrich et al., 1972). Hence, the simplest explanation for the observed relaxation is that, after *Eco*RV has nicked the plasmid, one strand of the DNA is free to rotate around the other before ligase reseals the nick. This mechanism demands that the DNA, nicked at the noncognate site, dissociates from the *Eco*RV enzyme.

In its native strain of *E. coli* (Bougueleret et al., 1984), the *Eco*RV restriction enzyme forms about 0.01% of the total protein, which corresponds to a concentration in vivo of about 200 nM (i.e., about 100 molecules/cell). The chromosome of *E. coli* will contain about 12000 noncognate sites for *Eco*RV, given their statistical frequency on random DNA. Given these concentrations together with the lowest value of k_{cat}/K_m for noncognate sites in Table I (that in buffer A), we can calculate that the chromosome of *E. coli* will be subjected to about 100 double strand breaks from *Eco*RV reactions at noncognate sites within each cell cycle of 30 min. [No reactions will occur at cognate sites due to the *Eco*RV methylase (Nwosu et al., 1988).] Hence, a bacterial cell that carries the R/M system will remain viable only if one or more of the following pertain: the restriction enzyme is compartmentalized away from the chromosome; the intrinsic activity of the enzyme toward noncognate DNA sequences is at least 100 times lower in vivo than in vitro; a proofreading system exists to repair DNA cut at noncognate sites. *E. coli* cells contain about 300 molecules of DNA ligase (Modrich et al., 1973). Hence, if we extrapolate from the data in vitro (Figures 6 and 7), the ratio of DNA ligase to *Eco*RV restriction enzyme in vivo would result in no repair of double-strand breaks at the cognate site but complete repair of single-strand breaks at noncognate sites: i.e., noncognate sites would never be cleaved in both strands. However, the relative activities of these enzymes in vivo are not known.

ACKNOWLEDGMENTS

We thank Annette Goodall for the construction of pAT153b, Marc Zabeau and Christian Vermote for strains and discussions, and Nigel Brown for advice on the manuscript. We also thank Brenda Fowler for the preparation of the manuscript.

Registry No. d(GATATC), 90937-02-3; d(GTTATC), 121288-62-8; Mg, 7439-95-4; restriction endonuclease *Eco*RV, 83589-02-0.

REFERENCES

Anderson, J. E., Ptashne, M., & Harrison, S. C. (1987) *Nature (London) 326*, 846–852.

Barany, F. (1987) *Gene 56*, 13–27.

Barany, F. (1988) *Gene 65*, 144–165.

Bennett, S. P., & Halford, S. E. (1989) *Curr. Top. Cell. Regul.* (in press).

Bougueleret, L., Schwartzstein, M., Tsugita, A., & Zabeau, M. (1984) *Nucleic Acids Res. 12*, 3654–3676.

Bougueleret, L., Tenchini, M. L., Botterman, J., & Zabeau, M. (1985) *Nucleic Acids Res. 13*, 3823–3839.

Boyer, H. W., & Roulland-Dussoix, D. (1969) *J. Mol. Biol. 41*, 459–472.

Brennan, C. A., Van Cleve, M. D., & Gumport, R. I. (1986) *J. Biol. Chem. 265*, 7270–7278.

Chang, A. C. Y., & Cohen, S. N. (1978) *J. Bacteriol. 134*, 1141–1156.

D'Arcy, A., Brown, R. S., Zabeau, M., Van Resandt, R. W., & Winkler, F. K. (1985) *J. Biol. Chem. 260*, 1987–1990.

Devereaux, J., Haeberli, P., & Smithies, O. (1984) *Nucleic Acids Res. 12*, 387–395.

Edsall, J. T., & Gutfreund, H. (1983) *Biothermodynamics*, Wiley, New York.

Engler, M. J., & Richardson, C. C. (1982) *Enzymes (3rd Ed.) 15*, 3–29.

Fersht, A. R. (1985) *Enzyme Structure and Mechanism*, 2nd ed., W. H. Freeman, New York.

Fersht, A. R., Shi, J., Knill-Jones, J., Lowe, D. M., Wilkinson, A. J., Blow, D. M., Brick, P., Carter, P., Waye, M. M. Y., & Winter, G. (1985) *Nature (London) 314*, 235–238.

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

George, J., & Chirikjian, J. G. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 2432–2436.

Halford, S. E. (1983) *Trends Biochem. Sci. (Pers. Ed.) 8*, 455–460.

Halford, S. E., & Johnson, N. P. (1981) *Biochem. J. 199*, 767–777.

Halford, S. E., & Goodall, A. J. (1988) *Biochemistry 27*, 1771–1777.

Halford, S. E., Lovelady, B. M., & McCallum, S. A. (1986) *Gene 41*, 173–181.
Hattman, S., Brooks, J. E., & Masureker, M. (1978) *J. Mol. Biol. 126*, 367–380.
Hochschild, A., Douhan, J., & Ptashne, M. (1986) *Cell 47*, 807–816.
Johnson, P. H., & Grossman, L. I. (1977) *Biochemistry 16*, 4217–4225.
Luke, P. A., & Halford, S. E. (1985) *Gene 37*, 241–246.
Luke, P. A., McCallum, S. A., & Halford, S. E. (1987) *Gene Amplif. Anal. 5*, 183–205.
Maniatas, T., Fritsch, E. E., & Sambrook, J. (1982) *Molecular Cloning, a Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
Maxwell, A., & Halford, S. E. (1982) *Biochem. J. 203*, 85–92.
McClarin, J. A., Frederick, C. A., Wang, B.-C., Greene, P. J., Boyer, H. W., Grable, J., & Rosenberg, J. M. (1986) *Science (Washington, D.C.) 234*, 1526–1541.
McDonell, M. W., Simon, M. N., & Studier, F. W. (1977) *J. Mol. Biol. 110*, 119–146.
McLaughlin, L. W., Bensler, E., Graesser, E., Piel, N., & Scholtissek, S. (1987) *Biochemistry 26*, 7238–7245.
Modrich, P. (1982) *CRC Crit. Rev. Biochem. 13*, 287–323.
Modrich, P., Lehman, I. R., & Wang, J. C. (1972) *J. Biol. Chem. 247*, 6370–6372.
Modrich, P., Anraku, Y., & Lehman, I. R. (1973) *J. Biol. Chem. 248*, 7495–7502.
Murray, N. E., Bruce, S. A., & Murray, K. (1979) *J. Mol. Biol. 132*, 493–505.
Nasri, M., & Thomas, D. (1986) *Nucleic Acids Res. 14*, 811–821.
Nwosu, V. U., Connolly, B. A., Halford, S. E., & Garnett, J. (1988) *Nucleic Acids Res. 16*, 3705–3720.
Ohlendorf, D. H., Anderson, W. E., Fisher, R. G., Takeda, Y., & Matthews, B. W. (1982) *Nature (London) 298*, 718–723.
Oka, A., Nomura, N., Morita, M., Sugisaka, H., Sugimoto, K., & Takanami, K. (1979) *Mol. Gen. Genet. 172*, 151–159.
Otwinowski, Z., Schevitz, R. W., Zhang, R.-G., Joachimiak, A., Marmorstein, R. Q., Luisi, B. F., & Sigler, P. B. (1988) *Nature (London) 335*, 321–329.
Peden, K. W. C. (1983) *Gene 22*, 277–280.
Polisky, B., Greene, P., Garfin, D., McCarthy, B., Goodman, H. M., & Boyer, H. W. (1975) *Proc. Natl. Acad. Sci. U.S.A. 72*, 3310–3314.
Rosenberg, J., & Greene, P. (1982) *DNA 1*, 117–124.
Rosenberg, J. M., McClarin, J. A., Frederick, C. A., Grable, J., Boyer, H. W., & Greene, P. J. (1987) *Gene Amplif. Anal. 5*, 119–145.
Rubin, R. A., & Modrich, P. (1978) *Nucleic Acids Res. 5*, 2991–2997.
Schildkraut, I., Banner, C. D., Rhodes, C. S., & Parekh, S. (1984) *Gene 27*, 327–329.
Seeman, N. C., Rosenberg, J. M., & Rich, A. (1976) *Proc. Natl. Acad. Sci. U.S.A. 73*, 804–808.
Shure, M. P., Pulleyblank, D. E., & Vinograd, J. (1977) *Nucleic Acids Res. 4*, 1183–1205.
Sugino, A., Goodman, H. M., Heyneker, H. L., Shine, J., Boyer, H. W., & Cozzarelli, N. R. (1977) *J. Biol. Chem. 252*, 3987–3994.
Sutcliffe, J. G. (1979) *Cold Spring Harbor Symp. Quant. Biol. 43*, 77–90.
Terry, B. J., Jack, W. E., & Modrich, P. (1987) *Gene Amplif. Anal. 5*, 103–118.
Thomas, M., & Davis, R. W. (1975) *J. Mol. Biol. 91*, 315–328.
Tikchonenko, T. I., Karamov, E. V., Zavizon, B. A., & Naroditsky, B. S. (1978) *Gene 4*, 195–212.
Twigg, A. J., & Sherratt, D. J. (1980) *Nature (London) 283*, 216–218.
Zimmerman, S. B., & Pheiffer, B. H. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 5852–5856.