Biochemistry

 Volume 35, Number 6 February 13, 1996

Articles

Random Mutagenesis Targeted to the Active Site of the *Eco*RV Restriction Endonuclease[†]

I. Barry Vipond and Stephen E. Halford*

Department of Biochemistry, Centre for Molecular Recognition, University of Bristol, Bristol BS8 1TD, U.K.

Received October 5, 1996; Revised Manuscript Received December 13, 1995[⊗]

ABSTRACT: Two segments of the gene for the *Eco*RV restriction endonuclease, each encoding 10 amino acids at the active site, were subjected to random mutagenesis with degenerate oligonucleotides. Mutations that abolished the activity of the *Eco*RV endonuclease were selected by viability in a strain of *Escherichia coli* that lacks the *Eco*RV methyltransferase, under conditions where the gene for the wild-type endonuclease is lethal to the cell. Sixty-five mutants were isolated and analyzed by DNA sequencing to identify the mutations. The collection of null mutants contained 49 with single amino acid substitutions, 15 with double substitutions, and one with a triple substitution. The single substitutions were located at many different positions within the two 10-amino acid segments, though several hot-spots gave rise to null mutants at high frequencies. Some hot-spots were readily explained by reference to the crystal structure of *Eco*RV since they were at the amino acids immediately adjacent to the scissile phosphodiester bond: for example, Asp90 and Lys92. These residues may be directly involved in the catalytic mechanism. Other hot-spots, such as Gln69, Tyr72, and Ala88, were at unexpected positions that appear to have no direct role in DNA binding or catalysis. At some of the unexpected hot-spots, the side chain of the amino acid lies distant from the DNA, yet the enzyme was still inactivated by conservative substitutions at these positions. The sensitivity of the *Eco*RV endonuclease to conservative substitutions may be due to its requirement to take up one particular conformation at the DNA−protein interface out of a large number of alternative conformations.

The *Eco*RV restriction endonuclease cleaves both strands of duplex DNA at a symmetrical sequence, GAT↓ATC (where ↓ marks the point of cleavage), in a reaction that requires only Mg^{2+} ions as a cofactor (Schildkraut et al., 1984; D'Arcy et al., 1985). The enzyme exists in solution as a dimer containing two identical subunits of 244 amino acids. In vivo, it functions alongside the *Eco*RV modification methyltransferase as part of a type II restriction−modification system (Bougueleret et al., 1984; Szczelkun & Connolly, 1995). The modification enzyme methylates the first adenine in the *Eco*RV recognition site (Nwosu et al., 1988) and thus protects the cellular DNA from cleavage by the restriction enzyme (Taylor et al., 1990).

In its DNA cleavage reaction, the *Eco*RV endonuclease shows a marked preference for its recognition site over any other DNA sequence. Noncognate sites that differ from the cognate site by just 1 bp[1] are cleaved at least a million times more slowly (Taylor & Halford, 1989; Waters & Connolly,

[†] This work was supported by the Biotechnology and Biological Sciences Research Council and by the Wellcome Trust.

* Correspondence to this author. Department of Biochemistry, School of Medical Sciences, University of Bristol, University Walk, Bristol BS8 1TD, U.K. Phone: +44 (0)11-928-7429. FAX: +44 (0)-117-928-8274.

[⊗] Abstract published in *Advance ACS Abstracts*, February 1, 1996.

[1] Abbreviations: Ap and Ap^r, ampicillin and resistance to Ap; BCIP, 5-bromo-4-chloro-3-indolyl phosphate; BME, β-mercaptoethanol; bp, base pair(s); BSA, bovine serum albumin; *ecoRVR* and *ecoRVM*, genes encoding the *Eco*RV restriction endonuclease and *Eco*RV modification methyltransferase, respectively; EDTA, ethylenedinitrilotetraacetic acid; LB, Luria broth; NBT, nitro blue tetrazolium chloride; Kn and Kn^r, kanamycin and resistance to Kn; PDB, (Brookhaven) Protein Data Base; SDM, site-directed mutagenesis; SDS−PAGE, polyacrylamide gel electrophoresis with sodium dodecyl sulfate; 2×YT, double-strength yeast-tryptone broth.

0006-2960/96/0435-1701$12.00/0

1994; Alves et al., 1995). Yet, in binding to DNA in the absence of Mg^{2+} ions, the *Eco*RV restriction enzyme shows no preference for its recognition site over other DNA sequences: it binds all sequences with equal affinity (Taylor et al., 1991). However, when the enzyme is located at its cognate site on DNA, it has a high affinity for Mg^{2+} ions but when located at a noncognate site, it has a much lower affinity for Mg^{2+}. Consequently, at all practicable concentrations of Mg^{2+}, only the cognate complex can proceed to cleave the DNA (Taylor & Halford, 1989). In contrast, the *Eco*RV endonuclease has a high affinity for Mn^{2+} ions when bound to either cognate or noncognate sites and, in the presence of Mn^{2+}, it cleaves both at similar rates (Vermote & Halford, 1992). The specificity of *Eco*RV for its recognition sequence is thus created in the ternary complex, containing enzyme, DNA, and divalent metal ions, rather than in the binary complex of enzyme and DNA (Baldwin et al., 1995). This view is consistent with DNA binding studies using either an inactive mutant enzyme with Mg^{2+} or the wild-type enzyme with an inactive cofactor, Ca^{2+}: in both cases, specific binding to the recognition site was observed (Thielking et al., 1992; Vipond & Halford, 1995).

This mechanism for DNA recognition can be related to the crystal structures of *Eco*RV bound to specific or nonspecific DNA (Winkler et al., 1993; Kostrewa & Winkler, 1995). In both complexes, the DNA is located in a deep cleft between the two subunits, with its minor groove facing the bottom of the cleft. Each subunit makes its main contacts to the DNA by two peptide loops. One loop, centered on Asn70, is known as the Q-loop since it carries two Gln residues, and this approaches the minor groove of the DNA. The other loop, centered on Asn185, is called the R (for Recognition)-loop since it carries the amino acids that make hydrogen bonds to the bases in the recognition sequence from the major groove. In the nonspecific complex, the R-loop lies outside the major groove and cannot interact with the bases. The specific and nonspecific complexes also contain very different DNA structures (Winkler et al., 1993). Nonspecific DNA bound to *Eco*RV retains a B-structure while the specific DNA is highly distorted. The energy for the distortion presumably comes from the contacts between the bases and the R-loop that are seen in the specific complex but not in the nonspecific complex. The distortion to the specific DNA places the relevant phosphodiester bond in the active site of the enzyme, while the nonspecific DNA lies distant from the active site. This positioning creates a binding site for Mg^{2+} in the specific complex, between the phosphodiester and two Asp residues (74 and 90), while no such site exists in the nonspecific complex (Winkler et al., 1993). The addition of Mg^{2+} to the crystals of *Eco*RV bound to cognate DNA leads to this 90/74 site being occupied by Mg^{2+}, but the DNA in the crystals remained intact (Kostrewa & Winkler, 1995). The reaction may require the binding of Mg^{2+} at a second site between Asp74 and Glu45 (Vipond et al., 1995; Baldwin et al., 1995). In the crystals of *Eco*RV bound to the product after a DNA cleavage reaction in solution, Mg^{2+} was observed at this 74/45 site and also at a second site adjacent to the main chain carbonyl group of Gln69 (Kostrewa & Winkler, 1995), though whether the Gln69 site has any catalytic significance has yet to be established.

Site-directed mutagenesis (SDM) of *Eco*RV has focused primarily on the residues that appear, from the crystal structures, to fulfill key roles: either in DNA recognition, in particular the R-loop (Thielking et al., 1991; Vermote et al., 1992; Wenz et al., 1994), or in phosphodiester hydrolysis (Selent et al., 1992; Thielking et al., 1992; Jeltsch et al., 1995). The latter had shown that the active site residues, Glu45, Asp74, Asp90, and Lys92, were all essential: nonconservative substitutions at each of these positions reduced *Eco*RV activity by factors of 10 000 or more. However, in contrast to SDM, random mutagenesis has the potential to identify key residues whose function is not apparent from the crystal structure. We describe here the random mutagenesis of two segments of the polypeptide chain of *Eco*RV. The two regions were targeted for mutagenesis by using "spiked" oligonucleotides in which the phosphoramidite for each stage in the synthesis was contaminated with small amounts of the other three phosphoramidites (Derbyshire et al., 1986; Hermes et al., 1989). Mutants with severely impaired activities were then isolated by the same genetic screen as in previous studies on the *Eco*RI, *Bam*HI, and *Nae*I endonucleases, viability in the absence of the companion methyltransferase (Yanofsky et al., 1987; Oelgeschläger et al., 1990; Xu & Schildkraut, 1991a; Holtz & Topal, 1994). The two regions of *Eco*RV selected for random mutagenesis covered amino acids 67−76 and 86−95 (Figure 1). The former includes the catalytic residue, Asp74, and the Q-loop. The loop makes several contacts to the sugar−phosphate backbone and to the edges of the bases facing the minor groove (Winkler et al., 1993), though it is doubtful whether the base contacts in the minor groove contribute toward specificity (Kostrewa & Winkler, 1995). The second region, positions 86−95, includes a β-strand that follows the sugar−phosphate backbone of the DNA in the cognate complex, making several contacts to the backbone, and also the catalytic residues, Asp90 and Lys92.

EXPERIMENTAL PROCEDURES

Strains and Plasmids. *Escherichia coli* CSH50 (Kunkel et al., 1991) was used as the host strain for all experiments, apart from preparations of ^{3}H-labeled pAT153 where the *rec*A strain, *E. coli* HB101 (Sambrook et al., 1989), was used instead. The plasmid pBSKSRV (Ap^r, *ecoRVR* $^-$) was constructed by cloning the *Eco*RI−*Hin*dIII fragment from pEMc3S (Vermote et al., 1992) in the polylinker site of a derivative of pBlueScript II KS^- (Ap^r, from Stratagene) that had been modified as described below. The *Eco*RI−*Hin*dIII fragment carries *ecoRVR* coupled to the λ P_L promoter at its optimal position for expression (Bougueleret et al., 1985), but the coding sequence is interrupted by a 30 bp "stuffer" fragment at the *Pst*I site close to its 5′-end (Vermote et al., 1992). Before the *Eco*RI−*Hin*dIII fragment was cloned, pBluescript had been modified so that the resultant construct, pBSKSRV, had no *Pst*I sites other than the pair flanking the stuffer fragment: the vector was first linearized with *Pst*I and then treated with T4 DNA polymerase to remove the 3′ single-strand overhangs, and finally recircularized with T4 DNA ligase. The *ecoRVR* gene in pBSKSRV was reactivated by using *Pst*I to excise the stuffer fragment, as detailed by Vermote et al. (1992), to yield pBSKSRVD. The plasmids p*cI*857 (Kn^r, λ*cI*857), pMetB (Kn^r, λ*cI*857, *ecoRVM*), and pAT153 have been described previously (Remaut et al., 1983; Vermote et al., 1992; Twigg & Sherratt, 1980).

DNA Manipulations. General procedures for bacterial cultures and DNA manipulations were carried out as in

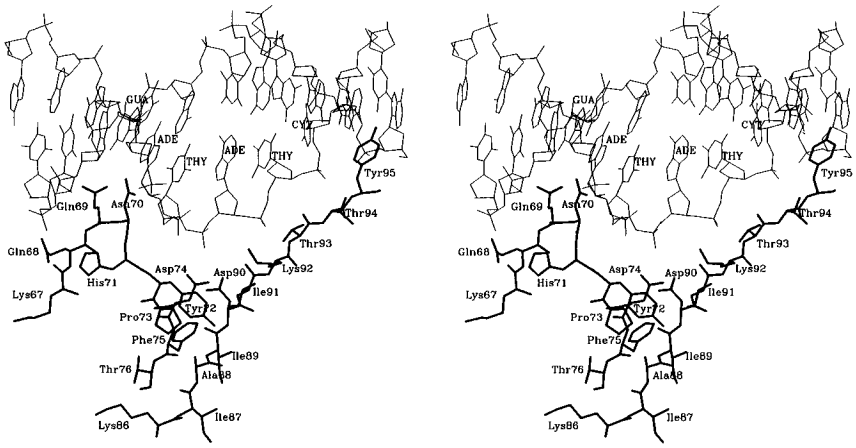

FIGURE 1: Active site region in *Eco*RV. The stereodiagram shows two segments of polypeptide chain (in bold type) from of the structure of the *Eco*RV restriction endonuclease bound to the duplex form of AAAGATATCTT (Kostrewa & Winkler, 1995). The two segments cover amino acids 67−76 and amino acids 86−95 from one of the two subunits of the dimeric protein: all of the amino acids are named in both segments. Both strands of the DNA are also shown in full, with the six bases in the recognition sequence named in one strand. Data taken from Brookhaven PDB file 1RVA.

Sambrook et al. (1989). Supercoiled pAT153 was labeled with ^{3}H-thymidine and purified as before (Halford & Goodall, 1988). Oligodeoxynucleotides were synthesized by L. Hall (this department) on a Millipore Expedite system with reagents from Cruachem (Glasgow, U.K.). The following were synthesized across the underlined regions by applying, at each cycle, equal volumes of two solutions of phosphoramidite: one solution contained the phosphoramidite for the base listed at the standard concentration; the other solution contained the phosphoramidites for all four DNA bases at 1/30th of the standard concentration.

oligo 67/76:

5′-TTTGT**ATATGTTGTTTTTATATCGATTGCAATTT**TTTTATT-3′

oligo 86/95:

5′-TAAAG**AGTAAAGTCAGGATAATGATTTTGTTGTT**TAGGTTC-3′

For the bases that are not underlined, the synthetic cycle employed only the phosphoramidite for the base listed.

Mutagenesis. The inactive copy of *ecoRVR* in pBSKSRV was mutagenized by the phosphorothioate method of Eckstein and colleagues (Taylor et al., 1985), using the mutagenesis kit from Amersham International plc., with either oligo 67/76 or oligo 86/95. Following mutagenesis, the mixture of plasmids was used to transform *E. coli* CSH50. The number of transformants was measured by plating a sample of the transformation mix (50 μL) onto LB agar containing Ap (100 μg/mL), and the colonies were counted after overnight growth at 37 °C. The rest of the transformation mix (950 μL) was diluted into 100 mL of LB containing Ap and grown overnight with shaking at 37 °C. A single preparation of plasmid DNA was made from the 100 mL culture and this was used as a gene pool of mutants. Aliquots of the gene pool DNA, typically 2 μg, were cleaved with *Pst*I to remove the stuffer fragment from the inactive copy of *ecoRVR* initially present in pBSKSRV. The cleaved DNA was then purified on a 1% agarose gel, recircularized with ligase, transformed into either *E. coli* CSH50 [pc*I*857] or CSH50 [pMetB], and the transformants were plated onto LB agar with Ap (100 μg/mL) and Kn (50 μg/mL). Transformants of CSH50 [pc*I*857] that gave rise to colonies after overnight growth at 28 °C potentially contained null mutants of the *Eco*RV endonuclease, while the overall efficiency of this process was assessed from the number of transformants obtained with CSH50 [pMetB].

Immunoblotting. Colonies containing potential null mutants of *Eco*RV were grown in 5 mL 2×YT containing Ap and Kn at 28 °C until the cultures reached an OD_{600} of about 0.3, typically after 2 h. At this point, *ecoRVR* was derepressed by the addition of 5 mL of 2×YT at 55 °C and the incubation continued for 3 h at 42 °C. The cells from 1.5 mL aliquots of these cultures were then harvested by centrifugation, and the pellets were resuspended in 200 μL of SDS gel-loading buffer (100 mM Tris-HCl, pH 6.8, 200 mM dithiothreitol, 40 mg of sodium dodecyl sulfate/mL, 2 mg of bromophenol blue/mL, 20% v/v glycerol). The samples were placed in a boiling water bath for 5 min and centrifuged, and aliquots (10 μL) from each supernatant were analyzed by SDS−PAGE. After electrophoresis, the proteins were transferred to nitrocellulose by electroblotting and probed using rabbit anti-*Eco*RV antisera as described previously (Vermote et al., 1992). The *Eco*RV band was visualized by binding to a goat anti-rabbit IgG coupled to alkaline phosphatase, followed by reaction with BCIP and NBT.

Sequencing. Plasmid DNA from viable transformants of *E. coli* CSH50 [pc*I*857] that were found by immunoblotting to express a full-size *Eco*RV-like protein were analyzed by DNA sequencing across the targeted region, typically 300 bp, to confirm the presence of a mutation and to identify the nature of the amino acid substitution(s). DNA sequencing was by the method of Sanger et al. (1977) using the sequenase kit from Amersham International. For a number of clones, the complete *ecoRVR* region was sequenced by the Genesis automated system (DuPont).

Protein Purifications. Mutant *Eco*RV proteins were purified from 1 L of 2×YT cultures (that had been derepressed for *ecoRVR* expression by temperature shock) by column chromatography on phosphocellulose (Whatman P11), essentially as described by Luke et al. (1987) but omitting the blue agarose column used for wild-type *Eco*RV. Fractions from the columns were analyzed by SDS−PAGE, and those containing the peak of *Eco*RV protein were pooled, precipitated by addition of ammonium sulfate to 75%

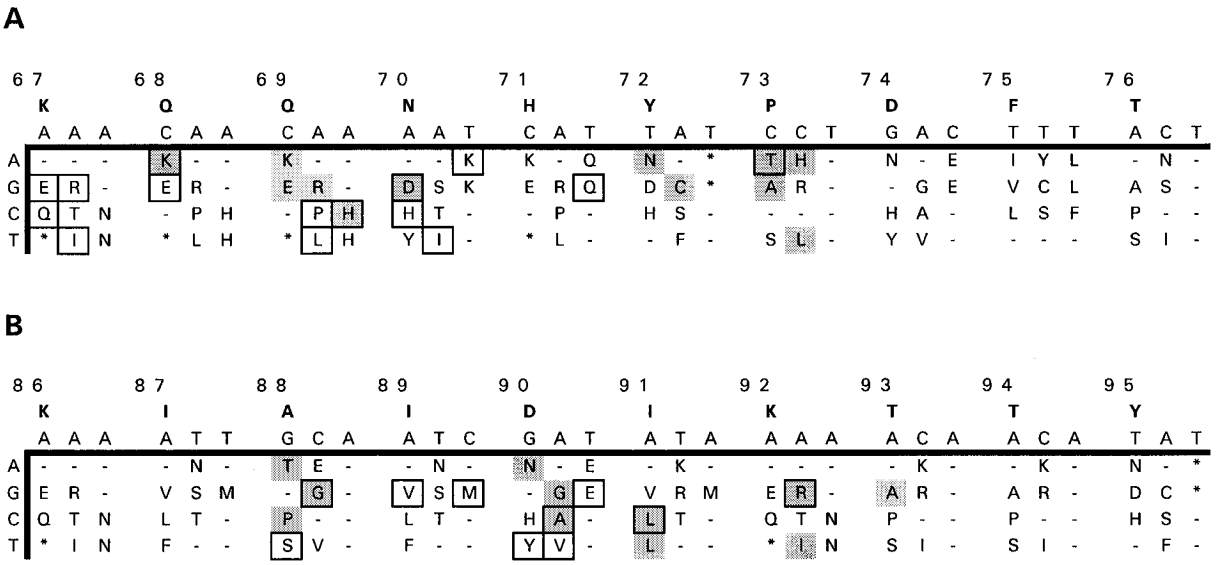

FIGURE 2: Possible mutations of *Eco*RV. The sequences above the lines list the single-letter amino acid sequences for the *Eco*RV endonuclease between residues 67−76 (in A) and between residues 86−95 (in B) and also the corresponding codons from the DNA sequence (Bougueleret et al., 1984). The amino acids below the lines indicate all possible substitutions that can be generated by changing the wild-type base at each position to the alternative base listed vertically on the left of the diagram. Dashes in the grid mark positions where the replacement of the base in the wild-type sequence by the alternative base makes no change to the amino acid from that codon. Base changes that generate stop codons are marked by stars. The shaded squares in the grid identify all of the single amino acid substitutions that gave rise to a null mutant of *Eco*RV isolated in this study: the corresponding single-base substitutions can thus be read off this grid. The boxed squares in the grid indicate all of the amino acid substitutions that gave rise to a null mutant of *Eco*RV as part of a multiple substitution in the targeted region. Shaded boxes represent replacements found in both single substitutions and as part of a multiple change. Three mutants were isolated that contained multiple base substitutions within a single codon and thus cannot be represented in the table: Y72K via TAT→AAG; A88R in the double mutant A88R+D90A via GCA→CGT; D90S via GAT→AGT on three occasions.

saturation, and stored at 4 °C as a slurry. Before use, an aliquot of the slurry was recovered by centrifugation and the pelleted enzyme dissolved in storage buffer (50 mM Tris-HCl, pH 7.5, 100 mM NaCl, 10 mM BME, 2 mM spermine, and 50% v/v glycerol). Protein concentrations were determined by the method of Bradford (1976) by reference to validated samples of *Eco*RV. All other enzymes were purchased from commercial suppliers and used as advised.

Enzyme Assays. Initial tests for DNA cleavage activity were carried out by serially diluting the resuspended enzyme in *Eco*RV dilution buffer (Halford & Goodall, 1988) and then adding 2 μL of each dilution to 18 μL of 10 nM ^{3}H-labeled pAT153 (>90% in its supercoiled form), in either MG buffer (50 mM Tris-HCl, pH 7.5, 100 mM NaCl, 10 mM BME, 100 μg of BSA/mL, 10 mM $MgCl_2$) or MN buffer (the same with 0.5 mM $MnCl_2$ instead of $MgCl_2$): the concentrations of *Eco*RV protein in the reaction mixtures spanned the range from 0.1 nM to 1.0 μM. After 1 h at 25 °C, the reactions were terminated by addition of 10 μL of stop mix (100 mM Tris-HCl, pH 8.0, 100 mM EDTA, 40% w/v sucrose, 100 μg of bromophenol blue/mL). The DNA was analyzed by agarose gel electrophoresis to separate the supercoiled, open circle, and linear forms of pAT153, and the amounts of each form were determined by scintillation counting (Halford & Goodall, 1988). Reaction rates were measured from 200 μL reactions, scaled up from above, by removing 15 μL aliquots of the mixture at various times and then analyzing each aliquot as above.

RESULTS

Mutagenesis of ecoRVR. Two 41-base oligodeoxynucleotides were synthesized, but both syntheses included 30 consecutive cycles in which the phosphoramidite for the correct base was "spiked" with low levels of the phosphoramidites for the other DNA bases. The spiked region in one oligonucleotide corresponded to amino acids 67−76 in the *Eco*RV endonuclease while the other covered amino acids 86−95 (Figure 1). In both oligonucleotides, the spiked regions were flanked by five or six bases where the synthesis employed only the cognate phosphoramidite, and these flanking regions were exactly complementary to the relevant loci in *ecoRVR*. The ratio of correct-to-incorrect phosphoramidites used here should result in each molecule of oligonucleotide carrying, on average, a single base change from the wild-type sequence, but this substitution may be at any position within the 30-base region across which the synthesis was spiked (Derbyshire et al., 1986; Hermes et al., 1989). The spiked oligonucleotides were used in a standard procedure for oligonucleotide-directed mutagenesis (Taylor et al., 1985) in order to generate random mutants of the *Eco*RV restriction enzyme, but in our procedure the mutations are targeted to preselected regions of the polypeptide chain. Single base changes over the segment of DNA encoding amino acids 67−76 have the potential to generate 61 different amino acid substitutions, while 63 substitutions are possible in the 86−95 region (Figure 2).

Plasmids carrying an active version of *ecoRVR* can only be maintained in *E. coli* strains that also express the *Eco*RV methyltransferase (Bougueleret et al., 1984). However, we observed previously that, even in the presence of the methyltransferase, plasmids of the type used here, with *ecoRVR* coupled to the λ P_L promoter (Bougueleret et al., 1985), were genetically unstable and rapidly lost *Eco*RV activity on account of spontaneous mutations that could occur at many different sites within the gene (Vermote et al., 1992). Consequently, in order to restrict the random mutations to the two selected segments of *Eco*RV, the mutagenesis was

carried out on a version of *ecoRVR* that had been inactivated by inserting a 30 bp DNA fragment at a unique *Pst*I site close to its 5′-end: the "stuffer" fragment contains stop codons in all three reading frames (Vermote et al., 1992). Following mutagenesis with either oligo 67/76 or oligo 86/95 (Experimental Procedures), the mutated DNA molecules were used to transform *E. coli* CSH50 and a single preparation of plasmid DNA was made from a bulk culture of the transformants. This preparation was used as a gene pool of random mutants. The gene pool DNA from mutagenesis with oligo 67/76 represented about 2000 transformants, while that from oligo 86/95 gave approximately 3400 transformants. The gene pools should contain many different plasmids, each carrying a mutation specified by an individual oligonucleotide in the mixture from the spiked synthesis, but it will also contain many plasmids on which the spiked oligonucleotides failed to generate a mutation. Since the mutagenesis and the subsequent amplification of the gene pool both employed an inactive form of *ecoRVR*, no selective advantage would have arisen from either the presence or the absence of a mutation from the spiked oligos.

Selection of Null Mutants. Expression of wild-type *ecoRVR* from the λ P_L promoter is lethal to *E. coli* cells that lack *ecoRVM*, even when the expression of *ecoRVR* is repressed by the temperature-sensitive *cI*857 repressor at its permissive temperature, 28 °C (Taylor et al., 1990; Vermote et al., 1992). Incomplete repression by *cI*857 at 28 °C presumably results in a level of *Eco*RV endonuclease that is sufficient to kill cells lacking the methyltransferase, by DNA cleavage at unprotected *Eco*RV sites on the chromosome. This feature of the *Eco*RV restriction−modification system was exploited in a selection for "null" mutants of the endonuclease. The *Eco*RV genes in the gene pools were reactivated by excising the stuffer fragment with *Pst*I. The plasmids were then recircularized with DNA ligase and transformed into *E. coli* CSH50 [pc*I*857]. This strain carries the temperature-sensitive λ repressor, *cI*857. Hence, transformants that give rise to colonies at 28 °C carry either a mutated copy of *ecoRVR* expressing a variant of the *Eco*RV endonuclease with reduced activity or an unmutated form of *ecoRVR* that still contains the stuffer fragment. Transformants carrying either the gene for the wild-type *Eco*RV endonuclease or a mutant form of *ecoRVR* that codes for an active enzyme would not survive. In a typical selection, more than 100 viable colonies were obtained from the mutated DNA preparations, whereas transformations with the same amount of unmutated DNA yielded only one or two colonies. In contrast, transformations into *E. coli* CSH50 [pMetB], which carries both the *cI*857 repressor and the *Eco*RV methyltransferase, resulted in more than 1000 colonies from both mutated and unmutated DNA preparations. This would suggest that about 90% of the DNA molecules in the gene pools from the mutagenesis reactions were either unmutated or that the mutations failed to reduce *Eco*RV activity sufficiently to score as a null mutant.[2]

The surviving colonies may, however, contain *ecoRVR* plasmids that still possess the stuffer fragment or have undergone some gross rearrangement. Alternatively, random mutagenesis may have generated either stop codons in the targeted regions (Figure 2) or aberrant frame shifts in the reading frame for *ecoRVR*. Consequently, the clones selected as null mutants of *Eco*RV were first tested for the expression of a full-sized *Eco*RV-like protein by Western blotting with an anti-*Eco*RV serum. No *Eco*RV-like protein was detected in about 55% of the clones, while approximately 40% of the viable transformants produced a protein that cross-reacted with the serum and which had the same electrophoretic mobility as wild-type *Eco*RV. The remaining 5% of the clones made proteins that cross-reacted with the serum but had different electrophoretic mobilities (data not shown). The missized proteins may have been due to frame shift or nonsense mutations, but some could have been due to point mutations causing an increased sensitivity to proteolysis in vivo, as has been observed with some site-directed mutants of *Eco*RV (Taylor, 1991). The mutations giving rise to these missized proteins were not analyzed further.

Characterization of Mutants. Plasmid DNA was isolated from a number of the viable colonies that expressed a full sized *Eco*RV-like protein, and the DNA was sequenced across the regions that had been selected for mutagenesis (Figure 2). For the mutants generated with oligo 67/76, 31 clones were sequenced across the codons for amino acids 67−76, from a primer located 169 bp downstream of this locus. Of these 31 null mutants, 22 carried single amino acid changes while nine had double substitutions (Figure 3a): in all of the sequences read from this primer, typically 300 bp, no changes were found outside the targeted region. For the mutants produced by oligo 86/95, 34 clones were sequenced across the region for amino acids 86 to 95 from the same primer, now 105 bp downstream of this locus. The 34 null mutants in the 86−95 region contained 27 single amino acid substitutions, six double substitutions, and one triple substitution (Figure 3b): again, the 300 bp sequences from this primer revealed no changes outside the targeted region. Six mutants were sequenced over the entire gene to discover whether the mutagenesis had introduced secondary changes in other regions. The six mutants sequenced in full included two from the 67−76 region, Q69E and P73T, and four from the 86−95 region, D90N, I91L, K92R, and T93A. In all six cases, the only change from the sequence for wild-type *Eco*RV was the single base change shown in Figure 2. We cannot, however, exclude the possibility that the null phenotype of some of the other mutants is due to a secondary change elsewhere in the gene rather than to the change(s) identified in the targeted region.

For the six mutants that had been sequenced in full, together with three additional null mutants, N70D, Y72N and P73A, cultures of the clones were grown and then induced to overexpress *ecoRVR*. Following induction by temperature shock to inactivate the *cI*857 repressor, all nine cultures had approximately 10% of their protein as *Eco*RV endonuclease. The *Eco*RV proteins were then purified by chromatography on phosphocellulose. All of the mutant

[2] We attempted to refine this selection procedure by culturing the transformants of *E. coli* CSH50 [pc*I*857] at either 37 or 42 °C instead of at 28 °C. Higher temperatures inactivate the *cI*857 repressor and thus increase the level of expression of *ecoRVR* from the λ P_L promoter. Hence, the viable colonies at 28 °C might have encoded mutant *Eco*RV proteins with partially reduced activities while the viable colonies at higher temperatures could have encoded proteins with severely reduced activities. However, the analysis of several mutants isolated at 42 °C revealed exactly the same mutations that had been found previously among the colonies selected at 28 °C (data not shown). In addition, *Eco*RV activities were measured on a number of the mutants selected at 28 °C: the majority had activities that were too low to measure and the most active were 500 times worse than wild-type *Eco*RV (see below, Table 1).

A

Mutant	67 Lys	68 Gln	69 Gln	70 Asn	71 His	72 Tyr	73 Pro	74 Asp	75 Phe	76 Thr
1		Lys								
2			Glu							
3			Leu							
4			Lys							
5			Glu							
6			Glu							
7			Glu							
8			Arg							
9			His							
10				Asp						
11						Asn				
12						Cys				
13						Asn				
14						Lys				
15							Ala			
16							His			
17							Thr			
18							Leu			
19							Thr			
20							Thr			
21							Ala			
22							Thr			
23	Glu		Leu							
24	Arg				Gln					
25	Gln		Pro							
26	Ile			Lys						
27		Lys		His						
28		Glu	Pro							
29			His	Ile						
30			Pro		Gln					
31				Asp			Thr			

B

Mutant	86 Lys	87 Ile	88 Ala	89 Ile	90 Asp	91 Ile	92 Lys	93 Thr	94 Thr	95 Tyr
1			Thr							
2			Pro							
3			Thr							
4			Thr							
5			Gly							
6					Gly					
7					Asn					
8					Ser					
9					Ser					
10					Asn					
11					Ala					
12					Gly					
13					Ser					
14					Ala					
15					Asn					
16					Gly					
17					Gly					
18					Ala					
19					Gly					
20					Ala					
21						Leu				
22						Leu				
23						Leu				
24							Ile			
25							Arg			
26							Arg			
27								Ala		
28			Gly			Leu				
29			Gly			Leu				
30			Arg		Ala					
31			Ser		Glu					
32				Met	Tyr					
33				Val	Val					
34				Met	Tyr		Arg			

FIGURE 3: Null mutants of *Eco*RV. Random mutagenesis targeted to specified segments of the *Eco*RV endonuclease, followed first by a genetic selection for null mutants and then by DNA sequence analysis, identified the amino acid substitutions shown in panel A, for amino acids 67−76, and in panel B, for amino acids 86−95. In both panels, single amino acid substitutions are listed first, in sequence order, followed by double and triple substitutions, in sequence order with respect to the first substitution. The order in which mutations at a given position are listed simply reflects the random order in which they were isolated.

enzymes adhered to phosphocellulose and could be eluted from the column with an NaCl gradient, in the same manner as the wild-type *Eco*RV protein (Luke et al. 1987). This shows that none of the nine mutants was grossly misfolded: otherwise, their chromatography would have differed from wild-type (Vipond et al., 1996). After this one column, the proteins were about 90% pure, as judged by SDS−PAGE (data not shown).

The nine mutant proteins that had been purified were tested for their ability to cleave the *Eco*RV recognition site on pAT153. Serial dilutions of the purified proteins were incubated with pAT153 for 1 h in reaction buffers containing either $MgCl_2$ or $MnCl_2$, and the amount of the supercoiled DNA substrate cleaved to either open-circle or linear forms was measured as described previously (Halford & Goodall, 1988). In the buffer containing Mg^{2+}, a reaction with 1 nM wild-type *Eco*RV would result in the complete cleavage of all of the substrate within 10 min (Halford & Goodall, 1988), while that with Mn^{2+} would be about 40% complete after 1 h (Vermote & Halford, 1992). However, with four of the nine mutants (N70D, P73T, D90N, and K92R), no DNA cleavage was detected after 1 h in the buffer containing Mg^{2+}, even when the enzyme concentration was increased to 1 μM (Table 1). In another three cases (Q69E, Y72N, and P73A), trace amounts of DNA cleavage were detected in the buffer with Mg^{2+} after 1 h with 1 μM enzyme, but the reaction rates were too slow to measure accurately: the Mg^{2+}-dependent activities of these three mutants is at least 10^4 times lower than wild-type *Eco*RV (Table 1). With the other two mutants (I91L and T93A), DNA cleavage was detected at a measurable rate in the presence of Mg^{2+}, but the rates were much lower than wild-type (Table 1).

Of the seven mutants that had failed to give a measurable rate for DNA cleavage with Mg^{2+} as the cofactor, four (P73A, P73T, D90N, and K92R) also gave rates that were too low to measure with Mn^{2+} while a fifth (Y72N) gave a very low rate with Mn^{2+} (Table 1). However, two of the seven mutants with near-zero activities in Mg^{2+}, Q69E and N70D, had activities with Mn^{2+} that were similar to the wild-type enzyme with Mn^{2+} as did one of the two mutants that had shown a low level of activity with Mg^{2+}, T93A (Table 1). A number of site-directed mutants of *Eco*RV have very low Mg^{2+}-dependent activities but are "rescued" by Mn^{2+} in that their Mn^{2+}-dependent activities approach the wild-type level with Mn^{2+} (Vermote et al., 1992; Selent et al. 1992). Similar behavior has also been observed with mutants of *Bam*HI (Xu & Schildkraut, 1991b). However, instead of being rescued to the wild-type level by Mn^{2+}, one of the mutants found here, I91L, was about 40 times more active than wild-type *Eco*RV in the presence of Mn^{2+} (Table 1).

Table 1: Activities of Mutant *Eco*RV Enzymes[a]

*Eco*RV protein	relative activities	
	MG buffer	MN buffer
wild-type	1.0	0.05
Q69E	$<1 \times 10^{-4}$	0.11
N70D	0	0.16
Y72N	$<1 \times 10^{-4}$	7×10^{-4}
P73A	$<1 \times 10^{-4}$	$<1 \times 10^{-4}$
P73T	0	0
D90N	0	0
I91L	1.6×10^{-3}	1.9
K92R	0	0
T93A	2.0×10^{-3}	0.22

[a] Rates of DNA cleavage were measured from reactions at 25 °C that contained 10 nM ^{3}H-labeled pAT153, in either MG buffer or MN buffer, with serial dilutions of each *Eco*RV protein as indicated. The reactions were initially carried with a range of concentrations of *Eco*RV protein, from 0.1 nM to 1.0 μM, and, where possible, the rates were then evaluated from reactions containing an appropriate enzyme concentration for DNA cleavage to be measured on a 1 h time scale. All reaction rates are given relative to a value of 1.0 for wild-type *Eco*RV in MG buffer. Relative rates given as $<1 \times 10^{-4}$ indicate that trace amounts of DNA cleavage were detected at the highest enzyme concentrations tested (1.0 μM), while values of 0 indicate that no DNA cleavage was detected with 1.0 μM enzyme.

Indeed, the Mn^{2+}-dependent activity of the I91L mutant was higher than the Mg^{2+}-dependent activity of wild-type *Eco*RV.

DISCUSSION

In this study, we used spiked oligos (Hermes et al., 1989) to generate random mutants within two selected regions of the *Eco*RV restriction endonuclease and then selected for null variants by viability in strains lacking the *Eco*RV methyltransferase. Similar strategies have been employed previously on the *Eco*RI, *Bam*HI, and *Nae*I endonucleases (Yanofsky et al., 1987; Xu & Schildkraut, 1991a; Holtz & Topal, 1994). In each of these cases, the entire gene for the endonuclease was mutated by hydroxylamine or nitrosoguanidine prior to the same selection for viable colonies, but null mutants were generated at relatively few sites in these proteins. For both *Eco*RI (Yanofsky et al., 1987) and *Bam*HI (Xu & Schildkraut, 1991a), the majority of the amino acids where substitutions produced the null phenotype have crucial functions in either DNA recognition or catalysis (Rosenberg, 1991; Newman et al., 1995). The other mutations causing the null phenotype, particularly in *Bam*HI (Xu & Schildkraut, 1991a) and *Nae*I (Holtz & Topal, 1994), were often substitutions at glycines, and these might have global effects on protein conformation. In contrast, over half of the 20 amino acids in *Eco*RV that were analyzed here could be mutated to yield the null phenotype (Figure 3).

The high frequency of null mutants in *Eco*RV, relative to either *Eco*RI or *Bam*HI, may be a reflection of the fact that, while *Eco*RI and *Bam*HI are similar to each other in terms of both structures and mechanisms (Aggarwal, 1995), *Eco*RV together with a number of other restriction enzymes such as *Taq*I and *Pvu*II belong to a distinct family of structures and/or mechanisms (Zebala et al., 1992; Cheng et al., 1994). For both *Eco*RV and *Taq*I, specific DNA recognition is achieved only in the catalytic reaction while sequence recognition by *Eco*RI and *Bam*HI also depends on the initial binding of the protein to the DNA and thus may be more robust. Alternatively, the increased frequency of null mutants with *Eco*RV, relative to the other restriction enzymes examined to date, could simply be a consequence of the efficiency of spiked-oligo mutagenesis (Hermes et al., 1989) compared to hydroxylamine mutagenesis, particularly as the process on *Eco*RV was targeted to its active site. However, the random substitutions of individual bases that are produced by spiked oligos will often result in either no change in the amino acid from a given codon or a conservative change, on account of the general nature of the genetic code (Figure 2). Hence, it might be envisaged that most substitutions of this kind would have little or no effect on enzyme activity, unless a critical functional group is replaced. In contrast to this expectation, spiked-oligo mutagenesis on *Eco*RV yielded a large number of null mutations in its active site, including many with conservative replacements (Figure 3).

Both X-ray crystallography and kinetic studies on *Eco*RV have shown that its DNA cleavage reaction must involve a series of conformational changes to both DNA and protein which result in a precise positioning of the catalytic functions, particularly the metal ions, against the scissile phosphodiester bond (Winkler et al., 1993; Kostrewa & Winkler, 1995; Baldwin et al., 1995). The sensitivity of *Eco*RV to mutations could therefore stem from the need to find one particular conformational state out of a large number of alternative states while maintaining the majority of the contacts between protein and DNA. If the protein−DNA interface is insufficiently flexible to maintain the contacts, the conformational rearrangements would be associated with high energy barriers (F. K. Winkler, personal communication). We attempt below to correlate the loss of activity from each of the mutants isolated here to the crystal structures of *Eco*RV.

Lys67−Gln68. No mutants were isolated with a single substitution at position 67, and only one, Gln→Lys, was isolated at position 68; however, several double substituents were found where either Lys67 or Gln68 was replaced alongside another change elsewhere in this segment (Figure 3A). However, the loss of activity in three of the four double mutants involving Lys67 can be ascribed to the second mutations at positions 69 and 70 (see below): the only one whose inactivity is likely to be due to the change at 67 is K67R+H71Q. In the high-resolution structure of *Eco*RV bound to its recognition sequence (Kostrewa & Winkler, 1995), Lys67 makes no direct contacts to the DNA though it interacts indirectly with Gln68 via a water molecule, while the amino group in the side chain of Gln68 interacts with a phosphate in the flanking DNA (Figure 4). Each subunit of the *Eco*RV protein makes the majority of its phosphate contacts to one strand of the DNA, but Gln68 contacts the opposite strand. Lys at position 68 might be able to contact this phosphate, but it would be precluded from the network of interactions involving Gln68.

Gln69. All possible amino acid substitutions that can be accessed by a single base change to the codon for Gln69 were found in the collection of null mutants (Figure 2). This is clearly a major hot-spot, but there is no obvious reason why this should be so. In the complexes of *Eco*RV bound to either specific or nonspecific DNA (Winkler et al., 1993), the methylene groups in the side chain of Gln69 lie parallel to the DNA backbone (Figure 4), making extensive van der Waals contacts, but no function can be ascribed to either the carbonyl or amino groups of the side chain. However, the main chain carbonyl of Gln69 interacts with a water molecule in the enzyme−substrate complexes with and without Mg^{2+}, while, in the enzyme−product complex, this

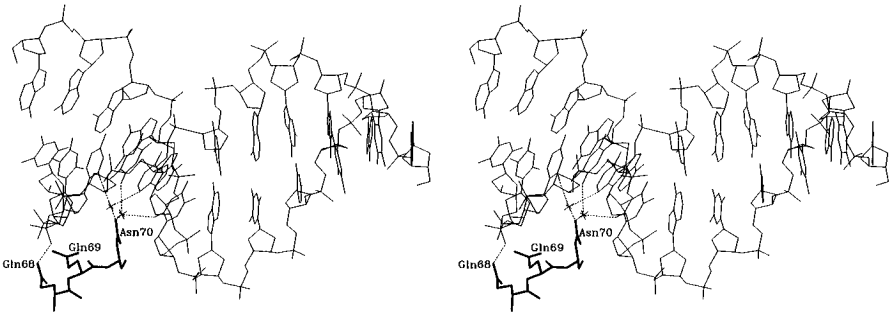

FIGURE 4: The Q-loop. The stereodiagram shows the positions of Gln68-Gln69-Asn70 in the A-subunit of the *Eco*RV restriction endonuclease in the structure of *Eco*RV bound to the duplex form of AAAGATATCTT (Kostrewa & Winkler, 1995). The DNA duplex is shown in full. Two water molecules are indicated by stars, and hydrogen bonds are indicated by dashed lines. The positions of the side chain of Asn70 and the water molecules in the B-subunit differ from that shown here. Data from PDB file 1RVA.

group is liganded to a Mg^{2+} ion which in turn is liganded to an oxygen on the 5′-phosphate of the cleaved DNA (Kostrewa & Winkler, 1995). One might have thought that the functions of both the methylene side chain and the main chain carbonyl on Gln69 could have been fulfilled by other amino acids, but this is not the case.

The principal role of Gln69 may be in determining the conformation and/or the flexibility of the Q-loop (Figure 4). Variations in both main chain and side chain conformations enable the Q-loop to mold itself around a variety of DNA structures while maintaining good backbone contacts in the minor groove, even with nonspecific DNA (Winkler et al., 1993). Alternatively, the binding of Mg^{2+} to the main chain carbonyl on Gln69 might have a direct role in catalysis, but this would require the Mg^{2+} to be bound to Gln69 at the transition state of the reaction rather than just in the enzyme−product complex as has been thought to date (Vipond et al., 1995). An additional possibility is that the carbonyl and/or amino groups in the side chain have a direct role at the transition state, even though no such role is apparent from the crystal structure of the ground state. Kinetic analysis of Q69E showed that the mutant was over 10^4 times less active than wild-type *Eco*RV in the presence of Mg^{2+}, but its Mn^{2+}-dependent activity was similar to wild-type (Table 1). Thus, the Gln69→Glu change cannot cause a gross misfolding of the protein but rather it disrupts the active center, interfering specifically with Mg^{2+}-supported catalysis. The fact that Mn^{2+} allows *Eco*RV to cleave noncognate DNA sequences better than Mg^{2+} (Vermote & Halford, 1992) suggests that Mg^{2+}-dependent reactions require a more precise alignment of the catalytic functions than Mn^{2+}-dependent reactions.

Asn70. Only one single substitution, Asn→Asp, was isolated at position 70 (Figure 3A). Asn70 is the only residue to interact with the bases in the minor groove (Figure 4). In the crystal structure of *Eco*RV bound to GGGATATCCC at 3.0 Å resolution, the interaction from Asn70 appears to be a direct hydrogen bond, but a similar bond was also observed in the structure of *Eco*RV bound to nonspecific DNA (Winkler et al., 1993). On the other hand, the 2.0 Å structure of *Eco*RV bound to AAAGATATCTT revealed that the interactions from Asn70 were mediated by water molecules, but these differed substantially between the two subunits in the dimer: the subunits are related by a non-crystallographic dyad (Kostrewa & Winkler, 1995). Hence, it seems unlikely that this minor groove contact plays a significant role in the recognition of the specific DNA sequence. Previously, a site-directed mutant with Gln at position 70 was reported to have an Mg^{2+}-dependent activity that was within a factor of 100 of the wild-type enzyme (Thielking et al., 1991). A 100-fold reduction might be insufficient to score as a null mutant: all of the mutants whose activities were measured were down by at least 500-fold with Mg^{2+} (Table 1). The reason why Asp at position 70 causes a larger reduction than Gln may be due to charge repulsion against the DNA phosphates.

His71. None of the null mutants carried a single substitution at position 71 (Figure 3A). This suggests that His71 lacks a significant role in either DNA recognition or catalysis. Many nucleases contain an essential histidine at their active sites that operates as an acid/base catalyst (Suck, 1992), but His71 in *Eco*RV is too far away from the scissile phosphodiester bond to fulfill such a function. However, the addition of $MnCl_2$ to the crystals of *Eco*RV bound to specific DNA results in a Mn^{2+} ion binding between His71 and a DNA phosphate outside the recognition sequence (F. K. Winkler, personal communication). Since this site is not filled by Mg^{2+}, on account of the low affinity of Mg^{2+} for nitrogen ligands, it is unlikely to have either recognition or catalytic functions.

Tyr72−Pro73. Both Tyr72 and Pro73 were hot-spots for null mutants (Figure 3A). Many restriction enzymes contain a sequence motif, P-D-...-(D/E)-X-K, where X is generally a long-chain aliphatic residue (Anderson, 1993), and, in both *Eco*RI and *Eco*RV, this motif is at the active site (Rosenberg, 1991; Winkler, 1992). However, *Eco*RV differs from *Eco*RI in that the relevant proline, at position 73, has the *cis* configuration while the analogous proline in *Eco*RI is the conventional *trans* isomer. The $\Delta G°$ for the *trans* → *cis* isomerization of proline is very much smaller than that for any other amino acid (Brandts et al., 1975). *cis*-Prolines thus occur in proteins at comparatively high frequencies, while *cis*-peptide bonds for all other amino acids are virtually unknown (MacArthur & Thornton, 1991). Hence, the replacement of Pro73 in *Eco*RV by any other amino acid will almost certainly result in a switch from *cis* to *trans* isomers and thus severely perturb the path of the peptide chain between the DNA recognition functions on the Q-loop and the catalytic residue, Asp74 (Figure 5). In an earlier study (Selent et al., 1992), SDM had been used to change Pro73 in *Eco*RV to either Ala or Gly: both had very low activities with Mg^{2+} as the cofactor but were reported to be only 3 times less active than wild-type with Mn^{2+}. In this study, the null mutants at position 73 encompassed four out of the six possibilities (Figure 2), including Ala and also His, Leu, and Thr (Figure 3A). In our hands, both P73A and P73T cleaved DNA too slowly to measure with either

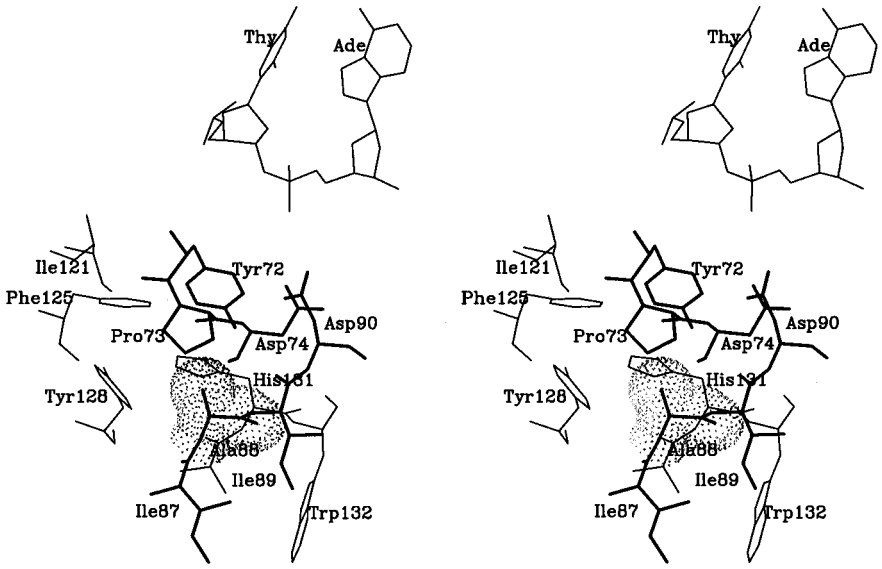

FIGURE 5: Tyr72, Pro73, and Ala88. The stereodiagram shows in bold type two segments of the polypeptide chain, Tyr72-Asp74 and Ile87-Asp90, from one subunit of the *Eco*RV restriction endonuclease, from the structure of *Eco*RV bound to the duplex form of AAAGATATCTT (Kostrewa & Winkler, 1995). The van der Waals surface of Ala88 is marked with dots. Also shown are a number of adjacent residues in the three-dimensional structure and the TpA step in the DNA that is cleaved by *Eco*RV. Data from PDB file 1RVA.

Mg^{2+} or Mn^{2+} (Table 1). Mutations at the analogous proline in *Eco*RI have little effect on its activity (Grabowski et al., 1995).

In protein structures, the probability of a proline taking up the *cis* configuration varies markedly with the nature of the adjacent residues: the highest frequency is when the Pro is immediately preceded by Tyr (MacArthur & Thornton, 1991), as is the case in *Eco*RV. The preference for *cis*-Pro in Tyr-Pro sequences can be accounted for by the *syn* configuration of this linkage allowing the rings of Tyr and Pro to stack against each other (MacArthur & Thornton, 1991). Exactly this stacking arrangement is seen in the *Eco*RV endonuclease at Tyr72-Pro73 (Figure 5). Hence, the reason why Tyr72 is a hot-spot for null mutants is not necessarily a direct effect of the substitution. Instead, Asn, Cys, or Lys at position 72 (Figure 3A) might cause Pro73 to revert to the energetically favored *trans* configuration.

Asp74-Phe75-Thr76. No mutants were recovered at these three positions (Figure 3A). *Eco*RV might have tolerated substitutions at Phe75 and Thr76, since both are located in the hydrophobic core of the protein (Figure 1). Proteins can often accommodate alternative packing arrangements in their core regions (Lim & Sauer, 1989). But tolerance to substitutions cannot explain the absence of null mutants at Asp74 since this residue coordinates one of the Mg^{2+} ions required for the catalytic reaction (Kostrewa & Winkler, 1995). Site-directed mutants of *Eco*RV replacing Asp74 by Asn or Ala had no cleavage activity at all, though Glu at this position gave 2% of wild-type activity (Selent et al., 1992). While D74E mutants might have escaped the selection for null mutants, the absence of any other single substituents at this position, and also all double substituents, cannot be reconciled to the known function of Asp74. However, the two collections of mutants, generated either with oligo 67/76 (Figure 3A) or with oligo 86/95 (Figure 3B), both lacked any representatives with mutations from close to the 3′-end of the spiked region. Hence, the most likely explanation for the absence of mutations at Asp74 is that spiked oligos with mismatches to the template strand close to their 3′-ends either failed to anneal properly to the template or failed to prime the chain extension reaction that forms part of the mutagenic procedure. The spiked oligos carried five or six canonical bases either side of the spiked region, but these had to be primarily A or T bases, on account of the very high AT content of *ecoRVR* (Bougueleret et al., 1984).

Ile87-Ala88-Ile89. No single replacements of Ile87 or Ile89 were isolated though three double mutants contained one of their changes at Ile89. In all three cases, the loss of activity in these double mutants can be assigned to the additional change at Asp90 (see below). Both Ile87 and Ile89 are buried deep within the hydrophobic core, and, since single base changes to the codons for Ile87 and Ile89 give primarily hydrophobic amino acids (Figure 2), substitutions at these positions might well be tolerated. Ala88 is also in the hydrophobic core but, in contrast to the above, several null mutants were recovered with single substitutions at this position. In this case, disruption of the close packing around Ala88 by any amino acid substitution will probably have severe consequences on the organization of the active site. In particular, the side chain of Ala88 is virtually wedged between the two stacked rings of Tyr72 and Pro73 (Figure 5) so that the replacement of Ala by either larger residues (Thr, Pro, or Ser) or by a smaller residue (Gly) must alter the packing of the Tyr-Pro element against position 88. This in turn will affect the spacing between Asp74 and Asp90, two residues that coordinate an Mg^{2+} ion at the active site (Kostrewa & Winkler, 1995).

Asp90. More null mutants contained substitutions at Asp90 than at any other position in this segment (Figure 3B). As expected for a key residue involved in binding a catalytic metal ion, its substitution by either Ala, Asn, Gly, Ser, Tyr, or Val destroyed activity. The D90N mutant had no detectable activity in the presence of either Mg^{2+} or Mn^{2+} (Table 1). These mutations extend earlier studies by SDM where complete loss of activity was caused by replacing Asp90 with either Asn or Ala, though the conservative substitution with Glu left wild-type activity (Selent et al., 1992). One of our double mutants contained Glu at position 90, but this also carried Ser in place of Ala at position 88 so

the loss of activity in this case can be assigned to the change at position 88. However, the D90A mutant of *Eco*RV retains the ability to bind Mg^{2+} (Thielking et al., 1992; Jeltsch et al., 1995), so the function of Asp90 appears to be in the positioning of the Mg^{2+} ion against the scissile phosphodiester bond rather than in binding Mg^{2+} *per se*.

Ile91. In contrast to the situations at Ile87 and Ile89, three null mutants of *Eco*RV were recovered with single substitutions at Ile91 and another two double mutants also had a change at this residue (Figure 3B). All five isolates had the same amino acid substitution, Ile→Leu, but, since this was generated by two different base changes (Figure 2), they cannot all be duplications of the same mutation during the amplification of the gene pool. This is a very conservative substitution for a null mutant, but the kinetic properties of the I91L protein were unusual (Table 1). The null phenotype observed in vivo is obviously a reflection of its low DNA cleavage activity with Mg^{2+} as the cofactor, but its Mn^{2+}-dependent activity was 40 times higher than that for wild-type *Eco*RV (Table 1). The properties of I91L are described more fully in the following paper (Vipond et al., 1996). The following paper also explains, at least in part, why all of the null mutants at position 91 carry Leu in place of Ile: mutants with Ala or Val at position 91 have activities that are similar to wild-type.

Lys92. In *Eco*RV, Lys92 is part of the P-D-...-(D/E)-X-K sequence motif found in many restriction enzymes (Anderson, 1993), but the function of this residue has yet to be defined. The assignment of a function to this Lys is complicated by the fact that the analogous position in *Bam*HI is filled by Glu and that the Glu→Lys mutation in *Bam*HI inactivates the enzyme (Dorner & Schildkraut, 1994). Site-directed mutants of *Eco*RV with Ala or Gln at position 92 were severely disabled for DNA cleavage under all conditions tested, while Glu at this position disabled activity in the presence of Mg^{2+} but allowed for nearly full activity with Mn^{2+} (Selent et al., 1992). This study yielded two new mutations at this position that prevent enzyme activity, K92I and K92R (Figure 3B). Even the conservative Lys→Arg change abolished DNA cleavage activity with either Mg^{2+} or Mn^{2+} (Table 1). Hence, it appears that only Lys can fulfill the function required at this position in *Eco*RV. Lys92 is located close to the scissile phosphodiester bond (Figure 1) and it has been suggested that its function is to neutralize the additional negative charge on the phosphorous at the transition state (Winkler et al., 1993). If so, one might have expected that Arg at this position would have been an adequate substitute, retaining at least partial activity, but this was not so. However, the interaction between Lys92 and the phosphate at the scissile bond is mediated by a water molecule, and this water also interacts with Asn188 on the R-loop in *Eco*RV [see Figure 5 in Kostrewa and Winkler (1995)]. Lys92 thus provides a physical link between the catalytic functions at the active site and the DNA recognition functions on the R-loop. Hence, it may not be directly involved in the chemistry of the *Eco*RV reaction and, instead, its principal function may be to organize the water at the active site.

Thr93-Thr94-Tyr95. These residues lie on a β-strand that runs parallel to the sugar−phosphate backbone in the complex of *Eco*RV with specific DNA (Figure 1). The consecutive amino acids in this region interact with consecutive phosphates in the DNA, either directly or via water. Thr93 also makes water-mediated contacts with other amino acids in this region that make further contacts to the phosphates (Kostrewa & Winkler, 1995). In the complex with nonspecific DNA, where the DNA retains a B-structure instead of being highly distorted, the DNA is clearly separated from this section of the polypeptide chain (Winkler et al., 1993). The phosphate contacts from Thr93, Thr94, and Tyr95 might therefore contribute to specificity by anchoring the preceding catalytic residues onto the DNA. However, this study yielded only one mutant with a substitution in these three residues, Thr→Ala at position 93. The absence of mutants at Thr94 and Tyr95 may be due to the same technical reason suggested above to account for the absence of mutants at Asp74.

Conclusions. Random mutagenesis, followed by genetic selection for null mutants, has demonstrated the sensitivity of the *Eco*RV endonuclease to amino acid substitutions at and around its active site. A number of the random mutants were at positions that had been analyzed previously by SDM (Thielking et al., 1991; Selent et al., 1992), but the collection of random mutants has generally extended the range of amino acid replacements tested at each position. In addition, the nature of the genetic code ensures that random substitutions of single bases results in many of these mutants containing conservative amino acid substitutions, such as Ile→Leu or Lys→Arg. If a conservative substitution of this type generates a null mutant, then the amino acid in question must be fulfilling a key function in the protein. The amino acids identified here as fulfilling key functions include several that might not have been anticipated from inspection of the crystal structures of *Eco*RV. In particular, the interactions between Tyr72, Pro73, and Ala88, in the hydrophobic core just underneath the active site (Figure 5), appear to play a vital role in the positioning of the active site residues. In addition, Gln69 clearly fulfills a crucial function though, in this case, the function is not apparent from the current crystal structures of *Eco*RV. X-ray crystallography on this and on some the other mutants isolated here is in progress (R. L. Brady and F. K. Winkler, personal communication) and the new structures may reveal why these mutations inactivate the *Eco*RV endonuclease.

ACKNOWLEDGMENT

We thank Geoff Baldwin, Bernard Connolly, Symon Erskine, Dirk Kostrewa, Tim Nobbs, Mark Oram, Alfred Pingoud, Mark Szczelkun, Lois Wentzell, and Fritz Winkler for discussions and unpublished data.

REFERENCES

Aggarwal, A. K. (1995) *Curr. Opin. Struct. Biol. 5*, 11−19.

Alves, J., Selent, U., & Wolfes, H. (1995) *Biochemistry 34*, 11191−11197.

Anderson, J. E. (1993) *Curr. Opin. Struct. Biol. 3*, 24−30.

Baldwin, G, S., Vipond, I. B, & Halford, S. E. (1995) *Biochemistry 34*, 705−714.

Bougueleret, L., Schwarzstein, M., Tsugita, A., & Zabeau, M. (1984) *Nucleic Acids Res. 12*, 3659−3676.

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

Bradford, M. (1976) *Anal. Biochem. 72*, 248−254.

Brandts, J. F., Halvorson, H. R., & Brennan, M. (1975) *Biochemistry 14*, 4953−4963.

Cheng, X., Balendiran, K., Schildkraut, I., & Anderson, J. E. (1994) *EMBO J. 13*, 3927−3935.

D'Arcy, A., Brown, R. S., Zabeau, M., van Resandt, R. W., & Winkler, F. K. (1985) *J. Biol. Chem. 260*, 1987−1990.
Derbyshire, K. M., Salvo, J. J., & Grindley, N. D. F. (1986) *Gene 46*, 145−152.
Dorner, L. F., & Schildkraut, I. (1994) *Nucleic Acids Res. 22*, 1068−1074.
Grabowski, G., Jeltsch, A., Wolfes, H., Maass, G., & Alves, J. (1995) *Gene 157*, 113−118.
Halford, S. E., & Goodall, A. J. (1988) *Biochemistry 27*, 1771−1777.
Hermes, J. D., Parekh, S. M., Blacklow, S. C., Koster, H., & Knowles, J. R. (1989) *Gene 84*, 143−151.
Holtz, J. K., & Topal, M. D. (1994) *J. Biol. Chem. 269*, 27286−27290.
Jeltsch, A., Maschke, H., Selent, U., Wenz, C., Köhler, E., Connolly, B. A., Thorogood, H., & Pingoud, A. (1995) *Biochemistry 34*, 6239−6246.
Kostrewa, D., & Winkler, F. K. (1995) *Biochemistry 34*, 683−696.
Kunkel, T. A., Bebenek, K., & McClary, J. (1991) *Methods Enzymol. 204*, 125−139.
Lim, W. A., & Sauer, R. T. (1989) *Nature 339*, 31−36.
Luke, P. A., McCallum, S. A., & Halford, S. E. (1987) *Gene Amplif. Anal. 5*, 185−207.
MacArthur, M. W., & Thornton, J. M. (1991) *J. Mol. Biol. 218*, 397−412
Newman, M., Strzelecka, T., Dorner, L. F., Schildkraut, I., & Aggarwal, A. K. (1995) *Science 269*, 656−663.
Nwosu, V. U., Connolly, B. A., Halford, S. E., & Garnett, J. (1988)- *Nucleic Acids Res. 16*, 3705−3720.
Oelgeschläger, T., Geiger, R., Rüter, R., Alves, J., Fliess, A., & Pingoud, A. (1990) *Gene 89*, 19−27.
Remaut, E., Tsao, H., & Fiers, W. (1983) *Gene 22*, 103−113.2
Rosenberg, J. (1991) *Curr. Opin. Struct. Biol. 1*, 104−113.
Sambrook, J., Fritsch, E. F., & Maniatis, T. (1989) *Molecular Cloning, a Laboratory Manual*, 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Sanger, F., Nicklen, S., & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5463−5467.
Schildkraut, I., Banner, C. D., Rhodes, C. S., & Parekh, S. (1984) *Gene 27*, 327−329.
Selent, U., Rüter, T., Köhler, E., Liedtke, M., Thielking, V., Alves, J., Oelgeschläger, T., Wolfes, H., Peters, F., & Pingoud, A. (1992) *Biochemistry 31*, 4808−4815.
Suck, D. (1992) *Curr. Opin. Struct. Biol. 2*, 84−92.
Szczelkun, M. D., & Connolly, B. A. (1995) *Biochemisty, 34*, 10724−10733.
Taylor, J. D. (1991). Ph.D. Thesis, University of Bristol, U.K.
Taylor, J. D., & Halford, S. E. (1989) *Biochemistry 28*, 6198−6207.
Taylor, J. D., Goodall, A. J., Vermote, C. L., & Halford, S. E. (1990) *Biochemistry 29*, 10727−10733.
Taylor, J. D., Badcoe, I. G., Clarke, A. R., & Halford, S. E. (1991) *Biochemistry 30*, 8743−8753.
Taylor, J. W., Ott, J., & Eckstein. F. (1985) *Nucleic Acids Res. 13*, 8765−8785.
Thielking, V., Selent, U., Köhler, E., Wolfes, H., Pieper, U., Geiger, R., Urbanke, C., Winkler, F. K., & Pingoud, A. (1991) *Biochemistry 30*, 6416−6422.
Thielking, V., Selent, U., Köhler, E., Landgraf, Z., Wolfes, H., Alves, J., & Pingoud, A. (1992) *Biochemistry 31*, 3727−3732.
Twigg, A. J., & Sherratt, D. J. (1980) *Nature (London) 283*, 216−218.
Vermote, C. L. M., & Halford, S. E. (1992) *Biochemistry 31*, 6082−6089.
Vermote, C. L. M., Vipond, I. B., & Halford, S. E. (1992) *Biochemistry 31*, 6089−6097.
Vipond I. B., & Halford, S. E. (1995) *Biochemistry 34*, 1113−1119.
Vipond, I. B., Baldwin, G. S., & Halford, S. E. (1995) *Biochemistry 34*, 697−704.
Vipond, I. B., Moon, B. Y., & Halford, S. E. (1996) *Biochemistry 35*, 1712−1721.
Waters, T. R., & Connolly, B. A. (1994) *Biochemistry 33*, 1812−1819.
Wenz, C., Selent, U., Wende, W., Jeltsch, A., Wolfes, H, & Pingoud, A. (1994) *Biochim. Biophys. Acta. 1219*, 73−80.
Winkler, F. K. (1992) *Curr. Opin. Struct. Biol. 2*, 93−99.
Winkler, F. K., Banner, D. W., Oefner, C., Tsernoglou, D., Brown, R. S., Heathman, S. P., Bryan, R. K., Martin, P. D., Petratos, K., & Wilson, K. S. (1993) *EMBO J. 12*, 1781−17945.
Xu, S.-Y., & Schildkraut, I. (1991a) *J. Biol. Chem. 266*, 4425−4429.
Xu, S.-Y., & Schildkraut, I. (1991b) *J. Bacteriol. 173*, 5030−5035.
Yanofsky, S. D., Love, R., McClarin, J. A., Rosenberg, J. M., Boyer, H. W., & Greene, P. J. (1987) *Proteins 2*, 273−282.
Zebala, J. A., Choi, J., Trainor, G., & Barany, F. (1992) *J. Biol. Chem. 267*, 8106−8116.

BI952391D