Mg^{2+} Binding to the Active Site of *Eco*RV Endonuclease: A Crystallographic Study of Complexes with Substrate and Product DNA at 2 Å Resolution‡

Dirk Kostrewa and Fritz K. Winkler*

F. Hoffmann-LaRoche Ltd., Pharmaceutical Research—New Technologies, 4002 Basel, Switzerland

Received July 19, 1994; Revised Manuscript Received October 27, 1994®

ABSTRACT: The type II restriction endonuclease *Eco*RV was crystallized as a complex with the substrate DNA undecamer AAAGATATCTT (recognition sequence underlined). These crystals diffract to much better resolution (2 Å) than was the case for the previously reported complex with the decamer GGGATATCCC [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*, 1781−1795]. The crystal structure contains one dimer complex in the asymmetric unit and was solved by molecular replacement. The same kinked DNA conformation characteristic for enzyme-bound cognate DNA is observed. Crystals, soaked with Mg^{2+}, show the essential cofactor bound at only one active site of the dimer, and the DNA is not cleaved. The Mg^{2+} has one oxygen from the scissile phosphodiester group and two carboxylate oxygens, one from Asp74 and one from Asp90, in its octahedral ligand sphere. The scissile phosphodiester group is pulled by 1 Å toward the Mg^{2+}. After substrate cleavage in solution, isomorphous crystals containing the enzyme−product−Mg^{2+} complex were obtained. In this structure, each of the 5′-phosphate groups is bound to two Mg^{2+}. The kinked DNA conformation is essentially maintained, but the two central adenines, 3′ to the cleavage sites, form an unusual cross-strand base stacking. The structures have been refined to *R* factors of 0.16 at 2.1−2.0 Å resolution maintaining very good stereochemistry. On the basis of these structures and inspired by recent kinetic data [Vipond, I. B., & Halford, S. E. (1994) *Biochemistry* (second paper of three in this issue)], we have constructed a transition state model with two metals bound to the scissile phosphorane group.

*Eco*RV is one of the best studied type II restriction endonucleases, both functionally [for recent reviews, see Vipond and Halford (1993) and Halford et al. (1993)] and structurally (Winkler et al., 1993). In the presence of Mg^{2+}, the enzyme cleaves double-stranded DNA in the middle of its recognition sequence GATATC and produces blunt-ended fragments with 3′-hydroxyl and 5′-phosphate groups. *Eco*RV cleaves DNA at its recognition sequence at least 10^6 times faster than at any other DNA sequence (Taylor & Halford, 1989). In contrast to this remarkable cleavage site specificity, no specific binding of *Eco*RV to its cognate site is observed by gel-shift experiments in the absence of Mg^{2+} (Taylor et al., 1991). This led to the conclusion that the specificity of the enzyme is intimately linked to Mg^{2+} binding and the subsequent catalytic steps. Another observation supporting this view is that the apparent affinity of the enzyme for Mg^{2+} is much higher for cleavage at cognate sites than at noncognate sites (Taylor & Halford, 1989). Furthermore, specific binding in the presence of Mg^{2+} could be shown for the catalytically inactive mutant enzyme D90A (Thielking et al., 1992).

Crystal structure analysis of the free enzyme and of its complexes with cognate and noncognate DNA fragments has provided a structural basis for understanding the strong coupling between recognition and efficient catalysis (Winkleret al., 1993). In the complex with the cognate DNA fragment, the so-called recognition loops (R-loops), one from each subunit, reach into the major groove and make extensive contacts (including 12 hydrogen bonds) with the six recognition base pairs. Formation of this tight recognition interface forces the DNA into a strained conformation characterized by a 55° central kink. In the complex with noncognate DNA, the R-loops are disordered and the DNA maintains a B-like conformation (Winkler et al., 1993). As a consequence of the kinked conformation in the specific binding mode, the scissile phosphodiester group becomes positioned close to the acidic side chains of Asp74 and Asp90. The catalytically essential role of these residues was shown by mutagenesis studies (Selent et al., 1992), and it was postulated that these three negatively charged groups form the Mg^{2+} binding site (Winkler, 1992). Thus, the formation of the cofactor binding site and the correct positioning of the scissile group in the active site are structurally dependent on the kinked DNA conformation which, in turn, is stabilized by the recognition interactions. In the absence of Mg^{2+}, the energy cost of this strained conformation is approximately compensated by the favorable interactions of the specific binding mode. The resulting cofactor affinity appears appropriate to achieve saturation at physiological Mg^{2+} concentrations, thus making it very sensitive to any perturbation of the recognition interactions. This simple structural rationalization of the lack of binding specificity in the absence of Mg^{2+} and of the strong coupling between recognition and catalysis qualitatively explains the major functional characteristics of the *Eco*RV enzyme. In particular, the dramatic reduction of specificity observed with Mn^{2+} in place of Mg^{2+} is easily explained by the higher intrinsic binding energy of Mn^{2+} (Vermote & Halford, 1992).

‡ The refined coordinates have been deposited with the Protein Data Bank as entries 1RVA for the *Eco*RV−substrate complex without Mg^{2+}, 1RVB for the *Eco*RV−substrate complex with Mg^{2+}, and 1RVC for the *Eco*RV−product complex.

® Abstract published in *Advance ACS Abstracts,* December 1, 1994.

0006-2960/95/0434-683$09.00/0

Table 1: Data Statistics

	*Eco*RV complexed with substrate DNA	*Eco*RV complexed with substrate DNA and Mg^{2+}	*Eco*RV complexed with product DNA and Mg^{2+}
space group	*P*1	*P*1	*P*1
cell constants *a*, *b*, *c*, α, β, γ	49.2 Å, 50.4 Å, 64.1 Å, 96.4°, 108.9°, 108.5°	49.2 Å, 50.4 Å, 64.1 Å, 96.4°, 108.9°, 108.5°	49.3 Å, 50.3 Å, 63.9 Å, 96.7°, 108.8°, 108.4°
observed reflections	78653	40175	39654
unique reflections	33340	26271	26657
R_I[a]	5.1%	5.2%	4.3%
completeness	91.9% (25.0−2.0 Å) 70.2% (2.1−2.0 Å)	83.7% (25.0−2.1 Å) 65.0% (2.2−2.1 Å)	85.1% (25.0−2.1 Å) 57.4% (2.2−2.1 Å)

[a] $\sum_h\sum_i|I_i(h) - \langle I(h)\rangle|/\sum_h\sum_i I_i(h)$, where $I_i(h)$ and $\langle I(h)\rangle$ are the *i*th and mean measurement of the intensity of reflections *h*.

In this paper we present the crystal structures of *Eco*RV complexed with a different undecameric DNA fragment, AAAGATATCTT (recognition sequence underlined), in the absence and presence of Mg^{2+}, and that of the corresponding enzyme−product complex with Mg^{2+}. All three structures have been refined to high resolution; 2.0, 2.1, and 2.1 Å, respectively. Compared with the previously reported DNA−enzyme complexes refined to only 3.0 Å resolution, these new results permit a much more precise and detailed structure description, including the location of many well-defined water molecules in the protein−DNA interface. More importantly, however, this crystal form has allowed us to determine the location and coordination of the Mg^{2+} binding site(s) before and after cleavage. In the accompanying papers (Vipond & Halford, 1995; Baldwin & Halford, 1995), it is suggested that a second metal is involved in *Eco*RV catalysed DNA cleavage. We present evidence for its location and describe the construction and evaluation of a transition state model with two metals bound to a pentavalent phosphorane group.

EXPERIMENTAL PROCEDURES

Materials. *Eco*RV restriction endonuclease, prepared as described (Luke et al., 1987), was a gift from S. Halford. The oligodeoxynucleotide was synthesized by standard methods and was provided by the groups of A. Pingoud (University of Giessen) and W. Bannwarth (F. Hoffmann-LaRoche Ltd., Basel).

Crystallization. Crystals of the *Eco*RV restriction endonuclease complexed with the substrate DNA undecamer AAAGATATCTT were grown using the batch method. One volume, typically 6 μL, of *Eco*RV at 15 mg/mL in 10 mM phosphate buffer, pH 7.5, 250 mM NaCl, 1 mM EDTA,[1] and 0.1 mM DTT was mixed with one volume of oligodeoxynucleotide at 3.0 mg/mL in 10 mM cacodylate buffer, pH 6.0, and $^1/_3$ volume of PEG 4000, yielding final PEG concentrations of 0.8−1.4%. Crystals grew within a few days as triclinic prisms up to a size of 1.0 × 0.45 × 0.30 mm^3. After growth, crystals were kept in storage buffer containing 10 mM phosphate, pH 6.8, 110 mM NaCl, and 10% PEG 8000. For Mg^{2+} soaking experiments, crystals were put for 6 days in a buffer with the same ionic strength as the storage buffer and containing 10 mM phosphate, pH 6.8, 30 mM $MgCl_2$, 50 mM NaCl, and 10% PEG 8000.

The *Eco*RV−product complex was crystallized after incubation of 100 μL of a *Eco*RV/DNA mixture as described above in the presence of 10 mM Mg^{2+} at 4 °C for 24 h. The lower temperature was chosen to prevent melting of the short product DNA duplexes. PEG 4000 was then added to 12-μL drops in the cold to final concentrations of 1−2%. After a few days at 4 °C very thin platelets appeared which did not grow further in the cold. The crystallization setup was then brought to room temperature where all crystals dissolved and did not reappear when brought back into the cold. Finally, the crystallization setup was transferred back to room temperature where one single crystal grew to a size of 0.3 × 0.2 × 0.2 mm^3 within 2 months in one drop. As this was sufficient to solve the structure no further crystallization protocols were investigated.

Data Collection. Diffraction data were collected on a Siemens-Nicolet X100 multiwire area detector using monochromatized Cu Kα radiation produced by an Elliot GX21 rotating anode generator. Oscillation pictures were taken every 0.15°. All data were processed with XDS (Kabsch, 1988). The three crystals grew isomorphously in space group *P*1 with one dimer enzyme−substrate or enzyme−product complex in the asymmetric unit. Cell constants and data statistics are given in Table 1.

Structure Solution. The crystal structure of *Eco*RV complexed with substrate DNA was solved by molecular replacement using the program package MERLOT (Fitzgerald, 1988). As a search model we used one *Eco*RV subunit with one DNA strand (shortened to the central hexamer GATATC) from the refined crystal structure of the decamer complex (Winkler et al., 1993; PDB entry 4RVE). The cross-rotation function, using observed data between 8 and 4 Å, gave two clear maxima related by a noncrystallographic 2-fold symmetry axis. After application of the two rotation solutions to the search model, the translation function gave a unique solution for the intermolecular vector for these two half-sites. This molecular replacement solution was then applied to the *Eco*RV monomer and the complete decamer strand. The initial *R* factor was 0.47 for data between 10 and 3-Å resolution.

Refinement. Refinement was carried out with the program package X-PLOR (Bruenger et al., 1987) using parameters for ideal stereochemistry based on the Cambridge Crystallographic Database (Engh & Huber, 1991), modified to maintain better planarity and chirality. Electron density map interpretation and model building was done with the crystallographic computer graphics program O (Jones et al., 1991). The starting model from the molecular replacement solution was initially refined with rigid body rotational and translational parameters to an *R* factor of 0.43 and then with

[1] Abbreviations: EDTA, ethylenediaminetetraacetic acid; DTT, 1,4-dithio-DL-threitol; PEG; poly(ethylene glycol); *R* factor, $\sum|F_o - F_c|/\sum F_o$; F_o, observed structure factor amplitudes; F_c, calculated structure factor amplitudes; rms, root mean square; K_D, dissociation constant.

Table 2: Refinement Statistics

	*Eco*RV complexed with substrate DNA	*Eco*RV complexed with substrate DNA and Mg^{2+}	*Eco*RV complexed with product DNA and Mg^{2+}
R factor[a]	0.162	0.165	0.156
resolution range (Å)	20.0−2.0	20.0−2.1	20.0−2.1
number of reflections[b]	33333	26254	26645
protein atoms[c]	4046	4046	4046
DNA atoms[c]	446	446	448
water atoms[c]	266	185	219
metal atoms[c]		2 Mg^{2+}	4 Mg^{2+}
rms[d] deviation from ideal bond lengths (Å)	0.010	0.010	0.009
rms[d] deviation from ideal bond angles	1.86°	1.88°	1.83°
rms[d] deviation from ideal χ_1 dihedrals	13.5°	13.5°	13.4°
rms[d] deviation from ideal planarity (impropers)	0.80°	0.77°	0.78°
temperature factors (*B*)			
target restraints			
main chain 1−2/1−3	2.0 Å²/3.0 Å²	2.0 Å²/3.0 Å²	2.0 Å²/3.0 Å²
side chain 1−2/1−3	4.5 Å²/6.0 Å²	4.5 Å²/6.0 Å²	4.5 Å²/6.0 Å²
refined temperature factors			
B-factor range[e]	9.5−92.5 Å²	5.5−99.0 Å²	6.5−98.4 Å²
$\langle B \rangle \pm \sigma(B)$			
protein atoms	29.6 ± 14.3 Å²	34.6 ± 17.0 Å²	29.8 ± 15.2 Å²
DNA atoms	26.9 ± 12.7 Å²	31.3 ± 15.9 Å²	26.9 ± 12.9 Å²
water and Mg^{2+} atoms	34.7 ± 10.3 Å²	34.3 ± 10.4 Å²	32.3 ± 10.3 Å²

[a] $\Sigma|F_o - F_c|/\Sigma F_o$. [b] No amplitude cutoff was applied. [c] Only non-hydrogen atoms are analyzed. [d] Rms, root mean square. [e] Maximum temperature factor restricted to $B = 99$ Å².

stereochemically restrained atomic positional refinement to an *R* factor of 0.30 for data between 10 and 3 Å resolution. At this point, the DNA sequence of the model was corrected to that of the undecamer, and the extra 5′-adenine was fitted to a $2F_o - F_c$ electron density map. This model was then refined with the simulated annealing method using the slow cooling protocol (Bruenger et al., 1990) for the temperature range 4000−300 K against observed data between 6.0- and 2.0-Å resolution, followed by conventional restrained positional and restrained temperature factor refinement. The refinement converged to an *R* factor of 0.25 in the resolution range 6.0−2.0 Å. Several rounds of model building sessions led to a reinterpretation of flexible loop regions and addition of water molecules. Improved continuous electron density for the poorly ordered parts was observed after including all data in the resolution range 20.0−2.0 Å for positional refinement and for electron density calculation. The low-resolution data could be included after optimization of the parameters describing the disordered solvent in X-PLOR. A solvent probe radius of 0.25 Å was used together with a disordered solvent constant electron density term of 0.34 e/Å³ and a smoothing temperature factor of 50 Å². The optimization of these parameters will be described elsehwere (D.K., manuscript in preparation). Restrained individual temperature factors were refined against data in the resolution range 5.0−2.0 Å with the maximum temperature factor restricted to $B = 99$ Å². Complete models of both polypeptide chains could be fitted, although some flexible loops were still poorly defined.

For the substrate complex with Mg^{2+} and for the product complex, the following refinement protocol was applied, with the refined structure of the undecamer complex serving as a starting model. In the first step, an omit-refinement at 3 Å resolution was carried out in which all atoms lying within a sphere of 11 Å radius around the scissile phosphorus were excluded from the structure factor calculation. The resulting electron density difference maps were easily interpretable, and the omitted structure was rebuilt. In the second step, the structures were refined to 3 Å resolution with all atoms contributing to the structure factor calculation except for the water molecules in the active sites. Next, the structures were refined to the resolution limit of 2.1 Å. Finally, well-defined water molecules in the active sites were added to the model, and some corrections of protein side chain and DNA sugar−phosphate backbone torsion angles were applied as indicated in electron density difference maps.

In addition to this conventional omit refinement, a simulated annealing omit map (Hodel et al., 1992) at 2.1 Å resolution was calculated for the *Eco*RV−product complex using the same starting model as described above. This produced an unbiased electron density in the omitted region and unambiguously showed that the enzyme−product complex has been crystallized (Figure 9b). Refinement statistics for all three structures are given in Table 2.

Superpositions. All superpositions were calculated using the least-squares-fit procedure in O (Jones et al., 1991). The superpositions are based on the 181 Cα atoms of the DNA binding subdomain of *Eco*RV, comprising residue numbers 2−12, 38−140, 166−220, and 230−241.

Modeling. The goal of the modeling was to replace the scissile phosphodiester group with a pentavalent phosphorane maintaining acceptable stereochemistry. The phosphorane group should coordinate to two metal ions held in place by the essential amino acids Asp74, Asp90, and Glu45. In addition, the stereochemistry should be consistent with the result that in enzymatic assays using phosphorothioates, only the R_p diastereomer is a substrate for *Eco*RV (Grasby & Connolly, 1992). Thus, the O2P atom from the phosphodiester group must be coordinated to the two metal ions. To meet these requirements, the scissile phosphodiester group was placed between its position in the substrate complex and the position of the 5′-phosphate group in the product complex such that its O2P atom coordinated to the two Mg^{2+} from the substrate complex. Assuming an S_N2 cleavage mechanism, the obvious candidate for an attacking water was water 105 from the ligand sphere of Mg1. A covalent bond was formed between this water oxygen and the phosphorus. Small adjustments of the Mg^{2+} and water positions were followed

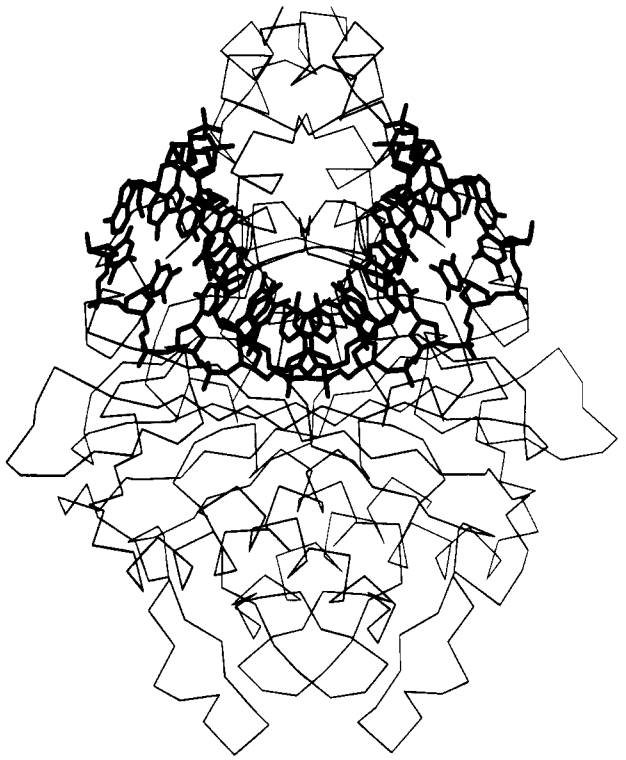

FIGURE 1: View of *Eco*RV complexed with the DNA undecamer AAAGATATCTT, showing the pronounced central kink in the DNA conformation. The protein is represented as a Cα trace in thin lines; the DNA is shown in thick lines.

by local energy minimization including neighboring protein and DNA residues using a united atom force field (Gerber & Müller, 1994).

RESULTS

We have determined the crystal structures of *Eco*RV complexed with, first, a cognate DNA undecamer without Mg^{2+} at 2.0 Å reolution, second, the same DNA undecamer with Mg^{2+} at 2.1 Å resolution, and third, the cleavage product of this undecamer with Mg^{2+} at 2.1 Å resolution. These three structures are very similar and major differences are seen only at the active sites. We first compare the structure of *Eco*RV complexed with the cognate DNA *undecamer* with the structure of *Eco*RV complexed with a cognate DNA *decamer* (Winkler et al., 1993). We then focus on the structures of the active sites of the three new complexes.

*Comparison with the Eco*RV*−decamer Complex.* The structure of the complex between *Eco*RV endonuclease and the DNA undecamer AAAGATATCTT without Mg^{2+} has been refined to 2 Å resolution which is of considerable improvement compared to the 3 Å of the previously reported complexes (Winkler et al., 1993). This has yielded much better defined conformations for many protein side chains and permitted us to locate 266 water molecules, 51 of which are part of the protein−DNA interface. As a whole, the structure of the complex with the undecamer is very similar to that of the complex with the decamer GGGATATCCC (Winkler et al., 1993) and shows the same kinked DNA conformation (Figure 1). The asymmetric unit of the crystals of the undecamer complex contains one dimer, while that of the decamer complex crystals contains three protein and three DNA monomers paired in one noncrystallographic and one crystallographic dimer. The previously disordered chain

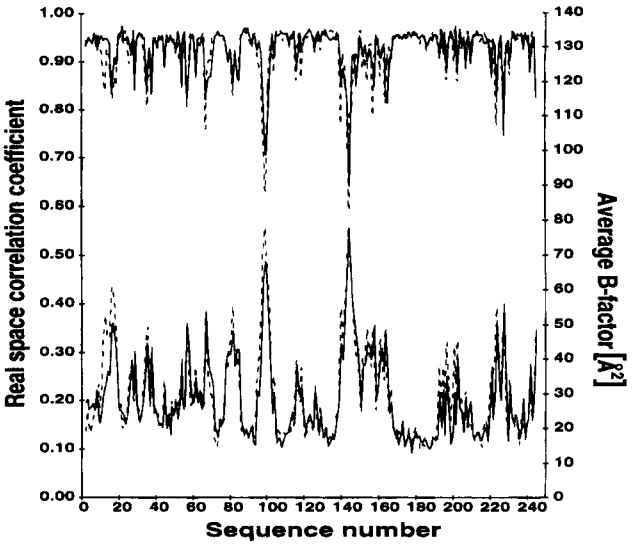

FIGURE 2: Real space correlation coefficient (top curves) and average temperature factor (bottom curves) for all protein atoms of the *Eco*RV−substrate complex against the residue sequence number. The real space correlation coefficient is calculated with O (Jones et al., 1991) using the coefficients $A_0 = 0.90$ and $C_0 = 0.91$ for the model atoms and the $2F_o - F_c$ electron density map of the refined structure. Results for subunit A are shown in solid lines; results for subunit B are shown in dashed lines.

segments 13−18, 99−100, 141−150, 221−229, and 242−245 could all be fitted in this new structure, but some are still poorly defined as indicated in Figure 2.

The DNA binding subdomains of the two crystallographically independent monomers of the undecamer complex superimpose very well onto the three crystallographically independent protein monomers of the decamer complex, with rms deviations of 0.66 ± 0.02 Å for 181 Cα atoms (see Experimental Procedures). Larger differences are restricted to more mobile surface loops and to flexible side chains. The superposition also reveals that the attached dimerization subdomains (Winkler et al., 1993) are in clearly different orientations in the different cognate complexes. When the DNA strands are superimposed based on this same least-squares transformation, a remarkable variation in the position of the DNA backbone with respect to the rather invariant protein can be seen (Figure 3). This variation correlates with conformational differences in the sugar−phosphate backbone of residue 0 (the DNA residues are numbered such that the residue carrying the scissile phosphodiester group on its 5′-side is number "0" with increasing numbers from the 5′-end to the 3′-end). In particular, the sugar conformation is 2′-*endo* in both strands of the undecamer complex but is 3′-*endo* in strands E and F of the decamer complex. As a result, the distances between the phosphorus atoms in positions 0 and +1 are 6.6−6.8 Å in the undecamer complex and 5.6−6.2 Å in the decamer complex. Thus, in the undecamer complex, the DNA strands are more B-DNA like at residue 0 than in the decamer complex (the backbone conformation of decamer strand D at residue 0 seems to be atypical).

All clear differences in the DNA−protein interface between the two cognate complexes occur either on the minor groove side of the bound DNA or along the sugar−phosphate backbone. Among the latter, we find a new interaction of the side chain of Arg226 with the phosphodiester groups in positions −5 and −4 of the undecamer strand. In the

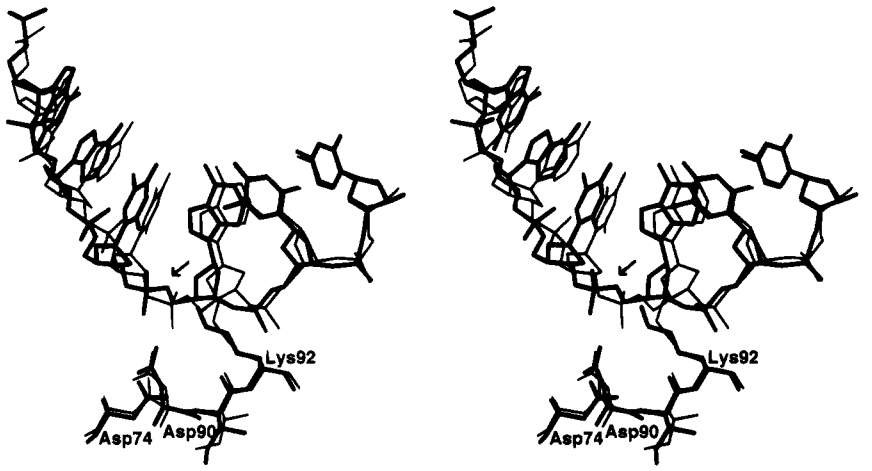

FIGURE 3: Stereopicture of the comparison of the relative DNA conformations at the active site of *Eco*RV for the undecamer and the decamer complex. The comparison is based on a superposition of the proteins DNA binding subdomains (see Experimental Procedures). Shown are the active sites of subunit B from the undecamer complex (thick lines) and subunit A from the decamer complex (thin lines). The relative DNA conformations of subunits B and C from the decamer complex lie between these two conformations. The scissile phosphodiester groups are marked with an arrow.

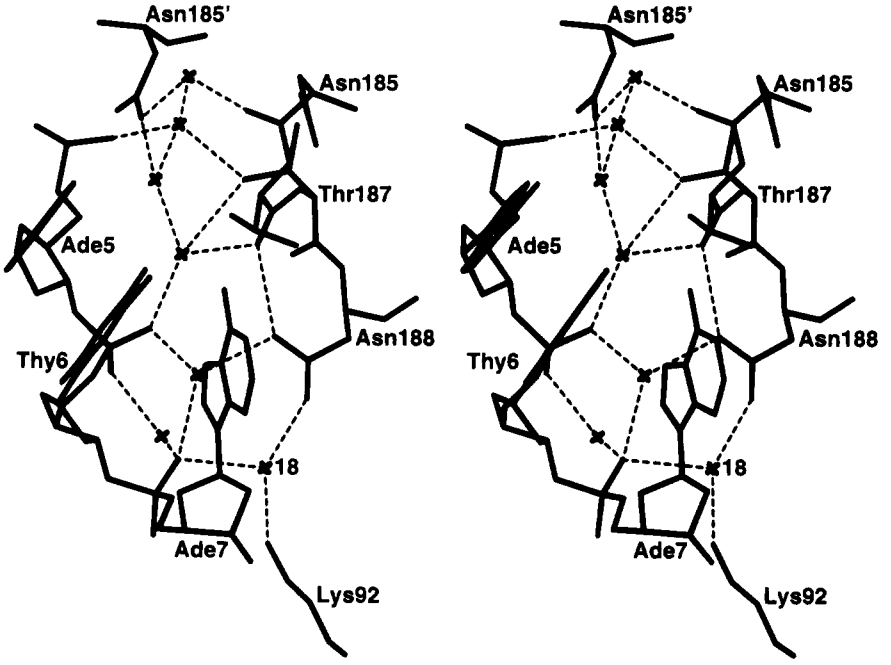

FIGURE 4: Stereopicture of the hydrogen bond network linking the recognition loop (182−186) to the phosphodiester backbone through a number of water molecules. Asn185′ comes from the other subunit of the dimer. Water 18 links the C-terminal extension of the recognition loop (Asn188) with the active site (Lys92).

decamer complex, the −5 phosphodiester group is missing and Arg226 is disordered. It is likely that this additional interaction contributes to the conformational stabilization of segment 221−229. An important set of water-mediated protein−DNA backbone interactions is shown in Figure 4. They link the phosphodiester groups of residues −2, −1, and 0 to the C-terminal part of the recognition loop (185−188). These water molecules are among the best defined and have low temperature factors (around 20 Å^2). In the decamer complex, solved at only 3 Å resolution, very restrictive criteria were applied for the acceptance of water molecules, and this intricate network was not identified, although it would be expected to be present. These indirect interactions between the DNA backbone and the recognition loop certainly allow for some variability in the relative position of these two structural elements which is further illustrated in two other cases. The first case is Ser112, where in the undecamer complex its Oγ makes a hydrogen bond to one of the nonesterified oxygens of the phosphodiester group in position −2. In the decamer complex, the distance is about 1 Å longer, and a water-mediated hydrogen bond is observed instead. The second case is Thr93, where in the decamer complex its Oγ1 makes a hydrogen bond to one of the nonesterified oxygens of the phosphodiester group in position +1. In the undecamer complex, this side chain is rotated by 120° around χ_1, and its hydrogen bond is mediated by a bridging water.

Other water molecules, not seen in the decamer complex, mediate contacts between the recognition bases in the minor groove and residues 68−70 of the so-called Q-turn (Winkler et al., 1993). In contrast to the previously described interactions, the latter do not obey the noncrystallographic symmetry. This correlates with different conformations of the two Q-turns including a peptide-flip between Gln69 and

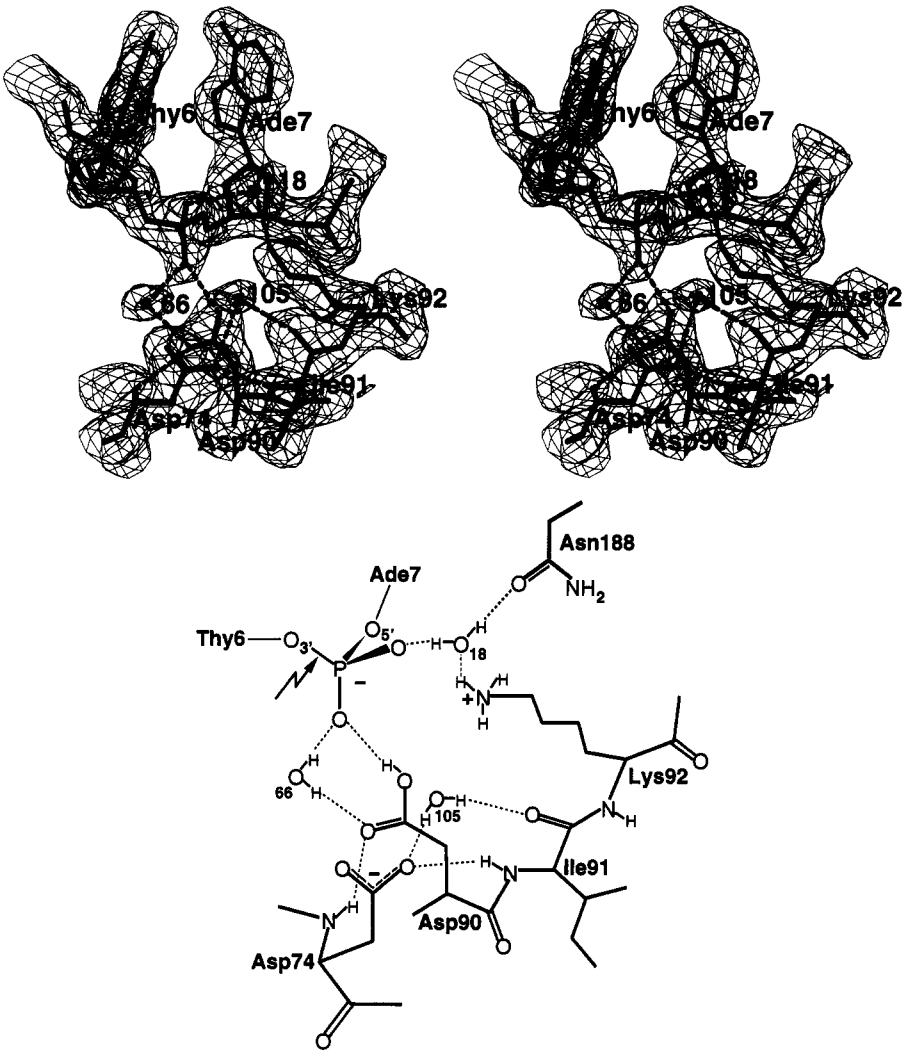

FIGURE 5: Active site of the refined *Eco*RV−substrate without Mg^{2+}. (a) Stereopicture of the active site at subunit B with its $2F_o - F_c$ electron density map at a contour level of 0.3 e/A^3 [$\approx 1\sigma(\varrho)$]. For clarity, only the electron density around the atoms is shown. (b) Schematic diagram illustrating the hydrogen bonding network. The arrow marks the scissile bond.

Asn70. The conformational flexibility of this turn has been noticed previously (Winkler et al., 1993) and enables the protein to smoothly adapt this important interface in the minor groove to variations in the DNA conformations.

The Active Site with Substrate DNA in the Absence of Mg^{2+}. The two crystallographically independent halves of the dimeric complex show essentially identical structures. Superposition of the two DNA binding subdomains yields an rms deviation of 0.30 Å for the 181 Cα atoms. Using this superposition, the amino acids Glu45, Asp74, Asp90, Ile91, and Lys92 of the two independent active sites show an rms deviation of 0.27 Å for all 42 atoms (0.14 Å if the side chain of the more flexible Lys92 is omitted beyond the Cβ atom). This deviation is not significant compared to the magnitude of the mean coordinate errors of 0.20 Å−0.25 Å (Luzzati, 1953). Similarly, no significant differences are observed for the two DNA strands. For the following description of the active site geometry we therefore focus on just one of the active sites.

Figure 5 shows the hydrogen bond network in the active site region (distances are listed in Table 3). The carboxylate group of the catalytically essential residue Asp74 forms two hydrogen bonds, one to the main chain nitrogen of Ile91, the other to water 105 which also donates a hydrogen bond to the main chain carbonyl oxygen of Ile91. The side chain carboxylate group of the other essential residue, Asp90, is ideally oriented and positioned to act as a donor in a hydrogen bond to one of the nonesterified oxygens of the scissile phosphodiester group. However, this implies that this carboxylate group is protonated at the slightly acidic pH (6.8) of the crystal mother liquor. In view of the clustering of negative charges in its immediate vicinity (phosphodiester group, Glu45, and Asp74) an upward shift of its pK_a appears plausible. In the decamer complex, crystallized at a similar pH, Asp90 is farther away from the scissile phosphodiester group and a direct interaction and therefore protonation of the carboxylate was not indicated (Winkler et al., 1993). The other carboxyl oxygen of Asp90 is linked through a water-mediated hydrogen bond (water 66) to the very same nonesterified oxygen of the scissile phosphodiester group, and it also makes a good hydrogen bond to the main chain nitrogen of Asp74. These favorable hydrogen bonding interactions contrast with an unfavorable torsion angle of −133° for χ_1. Interestingly, this strained side chain conformation is observed in all *Eco*RV structures thus far analyzed, whether bound to DNA or not, and it may be a consequence of the conserved intramolecular hydrogen bond to the main chain nitrogen of Asp74.

Table 3

(a) Comparison of Hydrogen Bond Distances at the Active Site of *Eco*RV Subunit B in Complexes with DNA−Substrate (S) and DNA−Product (P) with and without Mg^{2+} [a]

H-bond acceptor	H-bond donor	distance in complex *Eco*RV + S (Å)	distance in complex *Eco*RV + S + Mg^{2+} (Å)	distance in complex *Eco*RV + P + Mg^{2+} (Å)
Asp 74 Oδ2	Ile 91 N	2.9	2.9	2.9
Asp 74 Oδ2	HOH 105 O	3.3	2.7	
Asp 90 Oδ1	Asp 74N	3.1		3.0
Asp 90 Oδ1	HOH 66 O	2.9	2.8	2.6
Ade 7 O2P	Asp 90 Oδ2[b]	2.8	3.4	
Ile 91 O	HOH 105 O	3.3	2.8	3.3
Lys 92 Nζ	HOH 18 O	2.5	2.8	3.2
Ade 7 O1P	HOH 18 O	3.0	2.9	
Ade 7 O2P	HOH 66 O	2.8	2.8	
Ade 7 OXT	Thy6 O3′			2.6

(b) Comparison of Magnesium−Ligand Distances at the Active Site of *Eco*RV Subunit B in Complexes with DNA−Substrate (S) and DNA-Product (P)[c]

ligand atom	distance to magnesium ion in complex *Eco*RV + S + Mg^{2+} (Å)	distance to magnesium ion in complex *Eco*RV + P + Mg^{2+} (Å)
Asp 90 Oδ1	Mg1: 2.1	
HOH 66 O	Mg1: 2.1	
HOH 105 O	Mg1: 2.1	
HOH 107 O	Mg1: 2.1	
Ade 7 O2P	Mg1: 2.1	Mg3: 2.0
Asp 74 Oδ1	Mg1: 2.9	Mg3: 2.2
Glu 45 Oε2		Mg3: 2.0
Asp 74 Oδ2		Mg3: 2.3
HOH 291 O		Mg3: 2.1
HOH 292 O		Mg3: 2.0
Gln 69 O		Mg4: 2.1
Ade 7 O1P		Mg4: 1.9
HOH 126 O		Mg4: 2.3
HOH 168 O		Mg4: 2.3
HOH 254 O		Mg4: 1.9
HOH 290 O		Mg4: 2.2

[a] Distances greater than 3.5 Å are not considered. [b] From hydrogen bond geometry, it is assumed that the aspartic acid in the *Eco*RV−substrate complex is protonated. [c] The poorly defined Mg2 from the *Eco*RV−substrate complex is not considered.

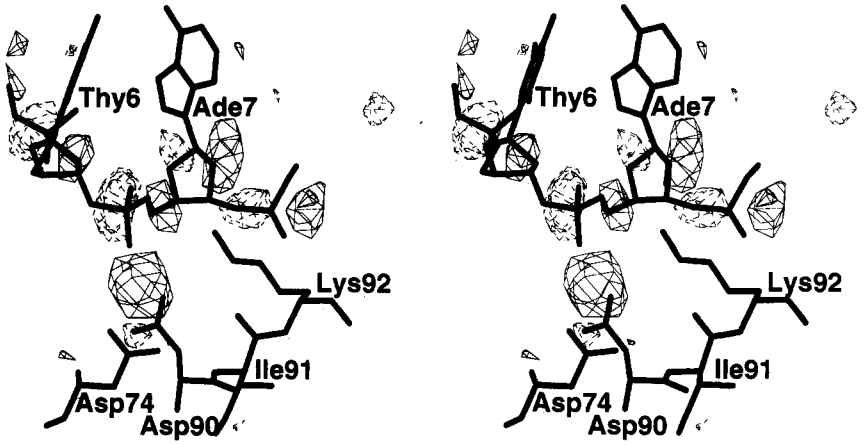

FIGURE 6: Stereopicture of the electron density difference map between the *Eco*RV−substrate complexes with and without Mg^{2+}. For the map calculation, the observed amplitude differences in the resolution range 20−3 Å, and phases from the refined *Eco*RV−substrate complex were used. Contour levels are +0.2 e/A^3 (solid lines) and −0.2 e/A^3 (dashed lines).

Water 18, one of the best defined water molecules, is hydrogen bonded to the amino group of Lys92, to one of the nonesterified oxygens of the scissile phosphodiester group, and to the carboxamide oxygen of Asn188. Hence, it provides a direct structural link between the C-terminal continuation of the recognition loop and the active site (see also Figure 4).

The Active Site with Substrate DNA in the Presence of Mg^{2+}. The most striking result from the analysis of the data of a crystal soaked with Mg^{2+} is that significant changes occur in only one of the two active sites of the noncrystallographic dimer and that no DNA cleavage is observed. The other active site shows no structural changes within the magnitude of coordinate errors. An electron density map computed with the differences between the observed structure factor amplitudes of the *Eco*RV−substrate complex with and without Mg^{2+} and with phases calculated from the refined structure without Mg^{2+} is shown in Figure 6. The strongest difference density peak with a peak height of 9σ (Table 4) is seen within short distance of the scissile phosphodiester

Table 4: Metal Binding at the Active Sites of the *Eco*RV−Substrate Complex[a]

*Eco*RV−substrate complex soaked with:	metal binding and peak heights[b] at the active site of subunit A	DNA shift at the active site of subunit A[c]	metal binding and peak heights[b] at the active site of subunit B	DNA shift at the active site of subunit B[c]	maximum peak height[b]
30 mM Mg^{2+}[d]	90/74: −	none	90/74: 9.0σ	large	9.0σ
	74/45: −		74/45: 4.6σ		
30 mM Mn^{2+}	90/74: −	none	90/74: 14.6σ	large	14.6σ
	74/45: −		74/45: −		
30 mM Co^{2+}	(8.4σ)[e]	none	90/74: 13.1σ	large	13.1σ
			74/45: 9.7σ		
30 mM Ca^{2+}	90/74: 8.9σ	none	90/74: 11.9σ	small	11.9σ
	74/45: −		74/45: −		
15 mM Mg^{2+} and 15 mM Ca^{2+}	90/74: 8.9σ	none	90/74: 13.3σ	none	13.3σ
	74/45: −		74/45: −		
1 mM Mn^{2+} and 20 mM Ca^{2+}	90/74: 7.4σ	none	90/74: 11.8σ	small	11.8σ
	74/45: 5.2σ		74/45: 8.0σ		

[a] The results shown are based on electron density maps calculated using the observed amplitude differences after metal binding and the phases from the refined model without metal ions in the resolution range 20−3 Å. [b] Peak heights are given as multiples of the rms density. [c] Indicated by alternating positive and negative electron density difference peaks; large means about 1 Å. [d] Soaking with 5 mM Mg^{2+} gave essentially the same binding sites. [e] The two peaks are not resolved; the maximum of the two overlapping peaks is given.

group and of the carboxylate oxygens of Asp74 and Asp90. Its position is thus very consistent with a Mg^{2+} binding site. In addition, pairs of positive and negative difference densities are seen along the adjacent DNA backbone. They indicate that this part of the DNA molecule has moved significantly in response to Mg^{2+} binding. Initial positional refinement at 3 Å resolution produced a considerable shift in the position of the DNA backbone and confirmed this interpretation. In order to prove that the observed DNA shift was induced by Mg^{2+} binding and not present in the crystal by accident, we repeated the soaking experiment. One crystal was cut into two halves, one of which was soaked in the Mg^{2+} buffer, while the other was kept in storage buffer without Mg^{2+}. For both halves of the crystal, data sets were collected and difference electron densities calculated as described above. Indeed, only the half crystal exposed to Mg^{2+} showed the strong positive electron density difference peak for the Mg^{2+} site and the alternating positive and negative electron density difference peaks along the DNA, thus confirming the correlation between Mg^{2+} binding and the DNA shift.

Refinement at 2.1 Å resolution, as described under Experimental Procedures, yielded the active site structure illustrated in Figure 7. The Mg^{2+} is octahedrally coordinated by one of the nonesterified oxygens of the scissile phosphodiester group, by two carboxylate oxygens, one from Asp74 and one from Asp90, and by three water oxygens. Compared to its position in the absence of Mg^{2+}, the scissile phosphodiester group is pulled about 1 Å toward the bound Mg^{2+} (Figure 8). The magnitude of this shift in the DNA backbone is largest at the active site and decreases quickly toward the 3′-end but slowly toward the 5′-end. Five of the six Mg^{2+}−O distances are close to the observed average value of 2.08 Å (Glusker, 1991) (very weak restraints were used during refinement for these distances). The distance to the carboxylate oxygen of Asp74 is significantly longer (2.9 Å). This points to an interesting difference between Asp74 and Asp90 in their response to Mg^{2+} binding. Asp74 retains its side chain conformation and the hydrogen bonds to the main chain nitrogen of Ile91 and water 105, which has become a metal ligand. In contrast, Asp90 changes its side chain conformation such that it ideally fits into the ligand sphere of Mg^{2+}. However, this is only achieved by disrupting the hydrogen bond to the main chain nitrogen of Asp74. Water 18 has shifted by 0.6 Å but maintains its three hydrogen bonding partners (Figure 7b) with only small changes in distances. Hydrogen bonding and metal−oxygen distances are listed in Table 3.

A two metal ion mechanism is proposed for the *Eco*RV catalyzed cleavage reaction in the accompanying paper by Vipond and Halford (1995). We have therefore carefully examined electron density difference maps for additional metal binding sites in the active site region. Only one significant uninterpreted electron density peak occurs, and it is located between the carboxylate groups of Glu45 and Asp74 (the 74/45 site). At a short distance (2.1 Å) from it, there are two water molecules which are not observed in the other active site. This could be regarded as suggestive of a second Mg^{2+} binding site. To clarify the nature of this peak, we have undertaken additional soaking experiments with Mn^{2+}, Co^{2+}, Ca^{2+}, and binary mixtures of Mg^{2+} or Mn^{2+} with Ca^{2+} at similar concentrations as were used with Mg^{2+}. Data sets were collected and the corresponding electron density difference maps calculated and inspected as described. The preliminary results show that Mn^{2+} binds to the same 74/90 site as Mg^{2+} and induces about the same structural changes in the corresponding DNA strand. Co^{2+}, in contrast, binds in each active site of the dimer at both the 74/90 and the 74/45 sites. The same binding behavior was observed with a mixture of Mn^{2+} with Ca^{2+}. However, in both cases a DNA shift is only indicated on the side where Mg^{2+} binding was observed. Finally, Ca^{2+} (with or without Mg^{2+}) appears to bind to both 74/90 sites of the dimer, only, with virtually no shift in either DNA strand (Table 4).

The Active Site with Product DNA in the Presence of Mg^{2+}. As no product formation could be obtained by adding Mg^{2+} to the crystalline complex, crystallization experiments were carried out with a preincubated mixture of enzyme, DNA, and Mg^{2+}. Surprisingly, isomorphous crystals were obtained and inspection of an electron density difference map indeed confirmed that the enzyme−product complex was present. Except for large changes in the immediate vicinity of the cleavage site, no other structural changes were indicated. During refinement as described under Experimental Procedures it became apparent that two Mg^{2+} ions were bound to each of the enzymatically produced 5′-phosphate groups. The noncrystallographic symmetry is again very well obeyed in this structure and one of the two nearly identical active sites is shown in Figure 9. The 5′-phosphate group is 4.6 Å away

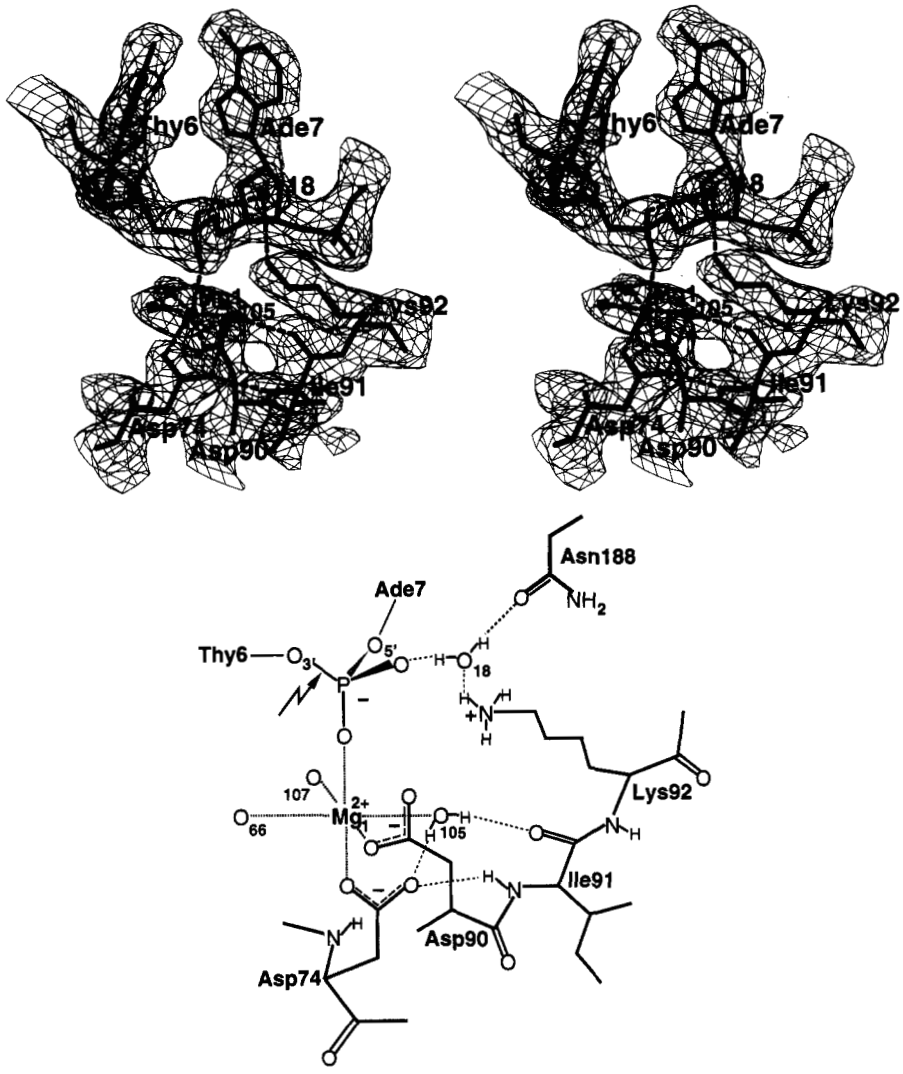

FIGURE 7: Active site of the refined *Eco*RV−substrate complex with Mg^{2+}. (a) Stereopicture of the active site at subunit B where Mg^{2+} has bound with its $2F_o - F_c$ electron density map at a contour level of 0.3 e/A^3 [$\approx 1\sigma(\varrho)$]. For clarity, only the electron density around the atoms is shown. (b) Schematic diagram illustrating the hydrogen bonding network and the metal coordination. The arrow marks the scissile bond.

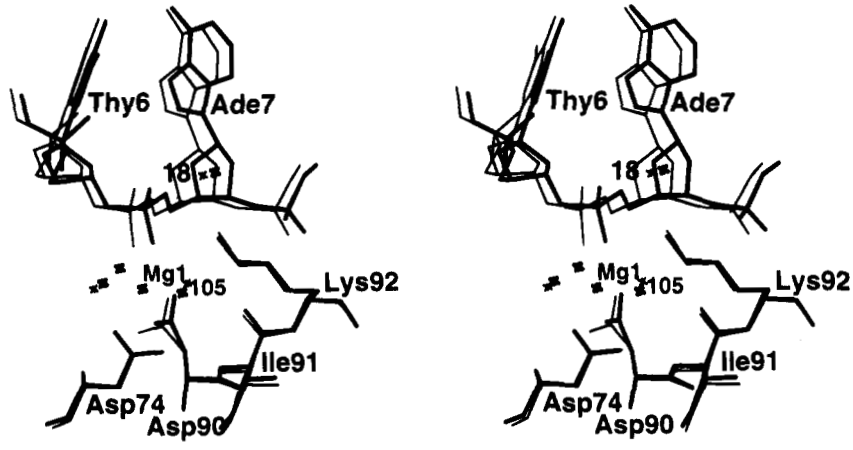

FIGURE 8: Stereopicture of the superposition of the refined *Eco*RV−substrate complex with Mg^{2+} (thick lines) and without Mg^{2+} (thin lines), illustrating the relative DNA displacement at the active site where Mg^{2+} has bound.

from its position before cleavage. It makes a hydrogen bond to the 3′-hydroxyl group of Thy6 to which it was covalently linked before cleavage, and it forms a salt bridge to the ammonium nitrogen of Lys38 coming from the other subunit (not shown in Figure 9). In addition, the 5′-phosphate group is bound to two Mg^{2+}. None of these two Mg^{2+} positions is directly comparable to the Mg^{2+} position in the *Eco*RV−substrate−Mg^{2+} complex. However, if we compare the coordination sphere of Mg1 in the substrate−cofactor complex (Figure 7 and Table 3) with the coordination sphere

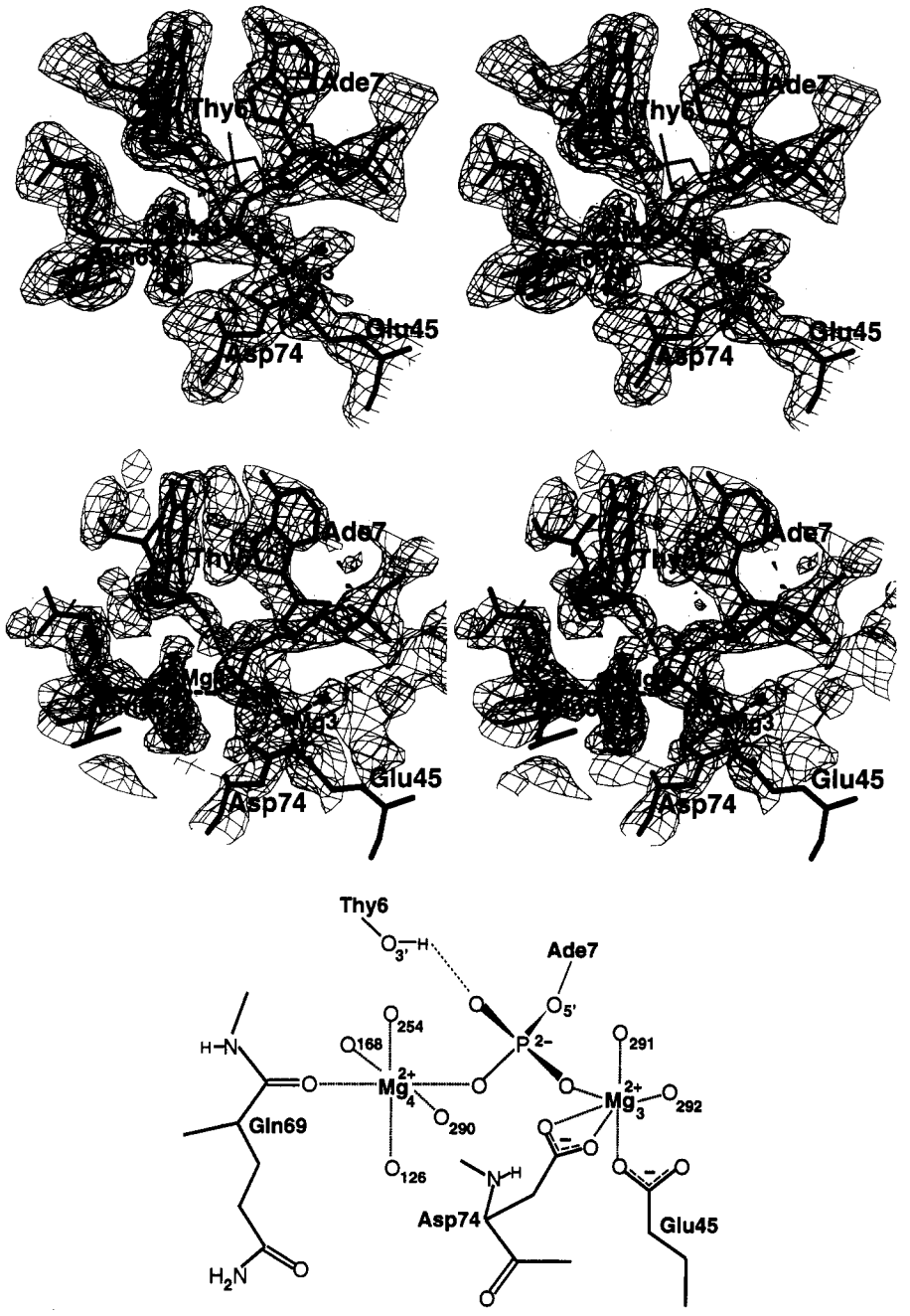

FIGURE 9: Active site of the refined *Eco*RV−product complex with Mg^{2+}. (a) Stereo picture of the active site at subunit B with its $2F_o - F_c$ electron density map at a contour level of 0.3 e/A^3 [$\approx 1\sigma(\varrho)$]. For clarity, only the electron density around the atoms is shown. For comparison, the DNA of the *Eco*RV−substrate complex is shown in thin lines. (b) Same view as in panel a, but showing the simulated annealing $F_o - F_c$ omit map within a box at a contour level of 0.15 e/A^3. The same atoms as in panel a are shown. (c) Schematic diagram illustrating the hydrogen bonding network and the metal coordination.

of Mg3 in this complex (Figure 9 and Table 3), we find in both one nonesterified oxygen from the scissile phosphodiester group and one oxygen from the carboxylate group of Asp74 as ligands. In contrast to the binding to Mg1 in the substrate complex, Asp74 binds with both carboxylate oxygens in a bidentate way to Mg3 in the product complex. This does not appear to be an optimal arrangement for Mg^{2+} coordination (Glusker, 1991). However, the local environment of Mg3 does not permit for an optimal octahedral ligand sphere without disrupting the structurally conserved hydrogen bond between the Asp74 carboxylate group and the main chain nitrogen of Ile91. The ligand sphere of Mg3 is completed by one carboxylate oxygen from Glu45 and two oxygens from water molecules (Figure 9 and Table 3). Mg4 is octahedrally coordinated by one of the phosphate oxygens, by the main chain carbonyl oxygen of Gln69 and by four oxygens from water molecules (Figure 9 and Table 3). In this product complex, the electron density for the Glu45 side chain is very well defined, in contrast to the other structures where it shows a variable extent of disorder. Thus, its ligation to Mg^{2+} stabilizes its otherwise flexible conformation.

The large motion of the 5′-phosphate group is accompanied by a shift of the corresponding sugar moiety and a reorientation of the attached adenine base. Instead of stacking with the thymine bases on their 3′-side, the two adenine bases are involved in a cross-strand stacking with each other (Figure 10). This correlates with a change of the glycosyl

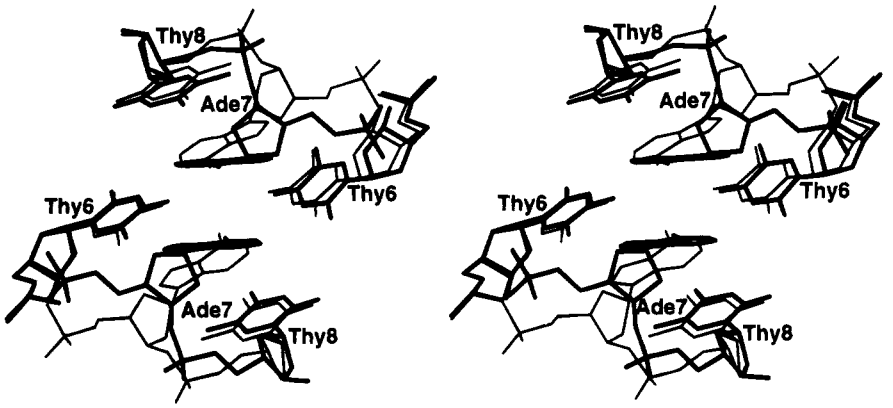

FIGURE 10: Cross-strand base stacking of the two central adenines after cleavage in the *Eco*RV−product complex (thick lines). For comparison, the uncleaved DNA strands of the *Eco*RV−substrate complex are shown (thin lines).

bond torsion angle χ from −125° in the substrate complex to +170° in the product complex. Furthermore, the Watson−Crick base pairing of the two central A-T base pairs at the kink is severely distorted with a buckle angle of about 30°.

DISCUSSION

High-resolution structure analysis of *Eco*RV−DNA substrate and product complexes has given us new insights into the structural changes accompanying recognition and catalysis. The higher resolution of the complex with the undecamer duplex AAAGATATCTT (2.0 Å) has considerably improved the accuracy to which structural details are defined, as compared to the previously reported complex with the cognate DNA decamer GGGATATCCC (3.0 Å; Winkler et al., 1993). This is of particular interest with respect to the interactions in the protein−DNA interface where many well-defined water molecules mediating a number of protein−DNA contacts could be identified. Furthermore, most of the residues that were poorly ordered in the structure of the decamer complex are now much better defined. In the case of segment 221−229 this is certainly helped by the interaction of Arg226 with the additional phosphodiester group present at the 5′-end of the undecamer as compared to the decamer. As described (Winkler et al., 1993), each DNA strand of a bound duplex (S1 and S2) interacts with both protein subunits (E1 and E2) of the dimeric complex. However, each strand interacts more strongly with the subunit responsible for its cleavage, and we have defined this as the E1S1 interface (equivalent to the E2S2 interface). In the crystal structures of the decamer and undecamer complexes, we have five crystallographically independent DNA strands (three in the decamer and two in the undecamer complex). Detailed comparison of the five E1S1-type interfaces has revealed some clear differences accompanied by a number of small local adaptations. Most strikingly, the backbone of the two undecamer strands appears considerably shifted compared to that of the three decamer strands after superposition of the rather invariant protein subunits (Figure 3). The relative displacement is largest at the scissile phosphodiester group but is significant all along the strand 5′ to the scissile bond. This shift correlates with a different backbone conformation of residue 0 which results in different phosphorous(0) to phosphorous(+1) distances. In the undecamer complex, the corresponding ribose sugar is in a 2′-*endo*-like conformation resulting in a standard B-DNA P(0)−P(+1) distance of 6.8 Å. In the decamer complex, the distances are between 5.6 and 6.2 Å, and the sugar conformation is 3′-*endo*-like in two of the three cases. All hydrogen bonding contacts between the protein and the phosphodiester groups 5′ to the scissile phosphodiester group are mediated through mobile water molecules or involve adaptable side chains (Lys, Arg, Thr, Ser). This permits the *Eco*RV−DNA complex to maintain these favorable contacts through local adaptations to the changes in the relative position of the DNA backbone. A different kind of adaptation is seen with the Q-turn (68−70), which makes extensive contacts with the sugar−phosphate backbone on the minor groove side 5′ to the scissile phosphodiester group (Winkler et al., 1993). Through the combined variation of its backbone and side chain conformation, this turn manages to form similarly good contacts with a variety of local DNA structures, including that observed in the even more different complex with a noncognate DNA fragment (Winkler et al., 1993). The observed variation in the structure of the DNA backbone has a negligible effect on the positioning of the recognition base pairs on the major groove side and the recognition interfaces are essentially identical. Thus, this comparison of the different cognate complexes indicates that considerable flexibility is built into the DNA−backbone protein interface at the reactive phosphodiester group and on its 5′-side but that the recognition interface is rather invariant.

Contrary to our expectations, soaking experiments with Mg^{2+} did not lead to cleavage in the crystalline undecamer complex even after heating to 37 °C for 1 week. In addition, Mg^{2+} binding is seen in just one of the two crystallographically independent active sites despite the absence of significant structural differences between the two. Mg^{2+} binding is accompanied by a shift of only that DNA strand which is directly involved in metal binding (Figure 8). The shift again propagates primarily on the 5′-side and is another illustration of the variability in this part of the DNA−protein interface described above. To explain the asymmetric binding, we are forced to postulate that the shift is needed for a sufficiently strong Mg^{2+} binding but that it cannot occur on both sides because of crystal packing forces. However, we have not been able to identify the structural origin of this packing constraint. Analysis of the packing of a modeled dimer with both strands symmetrically displaced shows unfavorable crystal contacts, but it is not obvious why these cannot be relaxed at a small energy cost. The only crystal

contacts involving the DNA duplex are formed betwen the end 5′ to the Mg^{2+} binding site at strand D and subunit B of a neighboring *Eco*RV dimer. No crystal contacts occur between DNA molecules. Crystal packing forces must also be responsible for the prevention of cleavage on the side where Mg^{2+} does bind. We have excluded the possibility that an inactive enzyme has been crystallized, by measuring the enzyme activity of redissolved crystals (data not shown). Furthermore, Mg^{2+} soaking of the decamer crystals does result in cleavage (C. Chéne and F. K. Winkler, unpublished). As there are no crystal contacts in the vicinity of the active site, we think that there must be some quaternary structure changes along the reaction pathway that are in conflict with the crystal contacts. The observed variation in the DNA conformation suggests that the postulated structural change has its origin primarily in a conformational change of the substrate DNA during catalysis. As described above, the structural perturbation at the scissile phosphodiester group induces a shift of the DNA backbone mainly on the 5′-side. This shift propagates all the way to the surface of the complex. Certainly, it will be accompanied by other structural adaptations, including some variation of the quaternary structure. Given the invariant backbone structure of the active site region of the enzyme in the various structures (with and without DNA), we consider it unlikely that this backbone undergoes larger structural changes during catalysis. However, some side chain rearrangements may occur.

Our new structural results, including the metal binding studies and the analysis of a product complex, have important implications for mechanistic considerations and hypotheses. A brief summary of the mechanistically most relevant experimental results may therefore be helpful. Using oligonucleotide substrates carrying a phosphorothioate rather than a phosphodiester group at the reactive site, it was shown that cleavage occurs with inversion of configuration at the phosphorus (Grasby & Connolly, 1992). Thus, the simplest and generally favored mechanism is a direct in-line attack by water with a trigonal bipyramidal transition state. In the same work it was shown that, of the two phosphorothioate diastereomers, only the R_P isomer was a substrate. In this isomer, the sulfur atom replaces the nonesterified oxygen which is not coordinated to the Mg^{2+} ion (Figure 7). Given the large preference of Mg^{2+} for oxygen against sulfur coordination (Eckstein, 1985), the observed Mg^{2+} binding mode nicely explains the discrimination of the two diastereomers. Both aspartic acid residues which ligate to the Mg^{2+} ion (74 and 90) have been shown by site-directed mutagenesis experiments to be essential for catalysis (Selent et al., 1992) and can only be substituted by glutamic acid. These essential acidic functions are conserved in three other type II restriction enzymes, *Eco*RI (Rosenberg, 1991), *Bam*HI (Newman et al., 1994), and *Pvu*II (Kokkinidis, personal communication), which share a common active site architecture with *Eco*RV. Substitution of another important active site residue in *Eco*RV, Lys92, by Ala or Gln yields an essentially inactive enzyme (Selent et al., 1992). The change to Glu still reduces activity about 1000-fold with Mg^{2+} as cofactor but is very well tolerated with Mn^{2+} (Selent et al., 1992). Somewhat surprisingly, this lysine residue is not conserved in *Bam*HI where a glutamic acid residue is found. Its replacement in *Bam*HI by lysine produces an inactive enzyme maintaining specific DNA binding (Dorner & Schildkraut, 1994). The role of Lys92 in *Eco*RV in catalysis is currently not clear, but two functions have been considered. First, the positive charge may be needed to balance the crowding of negatively charged groups in the active site region and stabilize the doubly charged pentavalent transition state. A glutamic acid substitution, when combined with an additional metal binding site, could then fulfil a similar function (Selent et al., 1992), and the same explanation is possible for *Bam*HI. However, there is no experimental evidence for such an additional metal binding site. Alternatively, Lys92 could function as the general base needed to deprotonate the attacking water molecule. Structurally, this appears plausible, as Lys92 forms a hydrogen bond to water 18 (Figure 7) which is close to a position from where a water molecule is expected to attack. Chemically, this is much less plausible as there is no obvious reason to justify the necessary lowering of the p*K* of this lysine. In addition, the observed hydrogen bonding of water 18 to the scissile phosphodiester group makes water 18 a poor candidate for nucleophilic attack. Some structural rearrangement, including the breakage of this hydrogen bond, would have to occur first. Glu113 of *Bam*HI is certainly a better candidate to act as a general base, which would imply, however, that this enzyme functions differently from the other three. For *Eco*RI and *Eco*RV it has been suggested that the group activating the attacking water could be one of the phosphoryl oxygens of the +1 phosphodiester group (Jeltsch et al., 1993). Evidence for such a substrate-assisted catalysis was provided by showing that modified substrates lacking either this negatively charged oxygen or the +1 phosphodiester group altogether were cleaved with dramatically reduced rates. The experimental evidence clearly indicates that a negatively charged group at the position of this *pro-R*p phosphoryl oxygen is important. Jeltsch et al. (1993) argue that the proposed water activation mechanism also appears structurally reasonable because the backbone conformation between the two phosphodiesters in the complex of *Eco*RV with the cognate decamer (Winkler et al., 1993) is such that the P(0)−P(+1) distance is about 1 Å shorter than in normal B-DNA, and the *pro-R*p phosphoryl oxygen is suitably oriented for water activation. As discussed here, this shorter P(0)−P(+1) distance is not observed in the undecamer complex even after Mg^{2+} binding. Nevertheless, the *pro-R*p phosphoryl oxygen is still similarly oriented. Hence, it could be argued that it is just this change in the DNA backbone conformation, bringing the two phosphodiester groups closer together, that is blocked in this crystal form. On the other hand, such a conformational change appears too small to explain why cleavage should be completely blocked. Of more concern, however, is the difficulty to rationalize the large p*K* shift that is required to make the phosphodiester group, whose normal p*K* is around 1 (Guthrie, 1976), a suitable base. Obviously, there is as yet no convincing candidate base for water activation.

Inspired by the two-metal ion mechanism that seems well established for the 3′-to-5′ exonuclease activity of DNA polymerase (Beese & Steitz, 1991; Steitz, 1993), an analogous mechanism has also been considered for *Eco*RV endonuclease (Winkler et al., 1993). In particular, it was pointed out that Glu45 together with Asp74 formed the lead heavy atom binding site in the structure of the free enzyme and in the noncognate complex, and that the large rate reductions observed for a number of Glu45 mutants (Selent

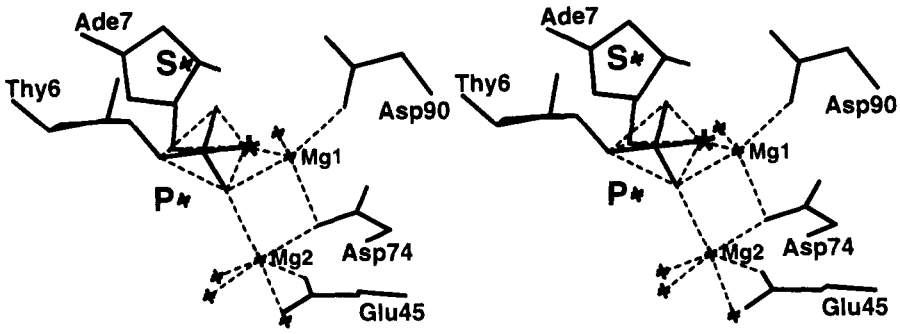

FIGURE 11: Transition state model with two metals bound to the pentavalent phosphorus. The attacking water (105) after bond formation is marked with an asterisk. For comparison, the positions of the phosphorus atoms in the substrate complex and in the product complex are shown with small crosses and marked with S and P, respectively.

et al., 1992) indicated an important role in catalysis. It was further noticed that this side chain occurs in several conformations and shows signs of disorder in all structures. More recent soaking experiments of crystals of the free enzyme with Mn^{2+} (Winkler, unpublished) have shown that this metal also binds to the 74/45 site. A two metal ion mechanism was, however, considered unlikely by Jeltsch et al. (1993) on the basis of the hyperbolic (and not sigmoidal) dependence of the *Eco*RV cleavage rate on Mg^{2+} concentration. They concluded that only one weakly bound metal ion (K_D > 1 mM) could be involved in catalysis but could not rule out the participation of another more strongly bound metal ion. New evidence for a two-metal mechanism in the case of *Eco*RV is now presented in the two accompanying papers (Vipond & Halford, 1995; Baldwin & Halford, 1995). In the first paper, it is shown that binary mixtures of Mg^{2+} or Mn^{2+} with Ca^{2+} show effects that are difficult to reconcile with a single metal ion bound in the active site. This holds true only for *Eco*RV and not for *Eco*RI, and it has been previously pointed out that Glu45 has no counterpart in *Eco*RI (Winkler et al., 1993). In the second paper, changes in the intrinsic tryptophan fluorescence of *Eco*RV endonuclease were used to monitor the rates of structural changes during binding and cleavage of an oligonucleotide substrate. Two processes with a different dependence on Mg^{2+} concentration were identified and tentatively associated with binding of one metal to the 90/74 site and a second to the 74/45 site. An intermediate in which metal ions are bound simultaneously to both these sites was postulated to be responsible for phosphodiester hydrolysis, in analogy to the two-metal mechanism discussed by Steitz (1993).

These results have provoked us to consider the structural possibilities for such a two-metal mechanism. In our previous thinking, we had assumed that the pentavalent transition state would be more or less within the range of DNA conformations observed in the two different cognate complexes. This seemed to be confirmed by the result of the Mg^{2+} soaking study, showing that metal binding does not involve a major structural change. At the same time, the lack of cleavage, despite seemingly productive cofactor binding, was puzzling. A second metal can indeed bind between Asp74 and Glu45 without further conformational changes, as shown by soaking the crystals with Co^{2+} or Ca^{2+}/Mn^{2+} mixtures (Table 4). In the case of Mg^{2+} there is just a weak hint of low occupancy metal binding at this site. While this assigns a clear function to Glu45, the role of this second metal site during catalysis is not obvious. With a transition state close to the observed DNA conformation there is no possibility for a direct interaction of this metal with the scissile phosphodiester group or for its participation in activating the attacking water.

Thus far we have ignored the structure of the product complex in our mechanistic considerations. As shown in Figures 9 and 10, this complex has a rather different structure of the DNA backbone in the active site region and has two metal ions bound to the newly generated 5′-phosphate group. One of them (Mg3) is also coordinated by Asp74 and Glu45, and, as suggested by Baldwin and Halford (1995), it is tempting to postulate that this metal site has a function before product formation. We therefore examined whether the structure of the transition state could be more toward that of the product complex, and whether a structurally acceptable model for the pentavalent transition state with two metals bound to the phosphorane group could be modeled. The result of some modeling studies, including energy minimization within the active site area, is shown in Figure 11. The pentavalent phosphorane group occupies an intermediate position between the phosphodiester group observed in the substrate complex and the 5′-phosphate group observed in the product complex. Its O2P atom originating from the phosphodiester coordinates to two metals. This is consistent with the fact, that only the *R*p phosphorothioate diastereomer is a substrate for *Eco*RV (Grasby & Connolly, 1992). The two metals are within 1.5 Å of the positions found in the soaking experiments. Only negligible adjustments were made to the amino acid side chains of the active site. One important result of this modeling exercise is that the architecture of the active site region permits substantial conformational changes of the DNA backbone and that conformations with the scissile phosphodiester group coordinated to two metals with 74/90 and 74/45 as additional acidic ligand pairs appear possible. Such a substantial conformational change of the DNA backbone on the way to the transition state certainly makes it easier to rationalize why cleavage is prevented by crystal packing forces. A second important result is that, in the case of *Eco*RV, the stereochemistry of such a transition state cannot be the same as in the two-metal mechanisms discussed by Steitz (1993). In the latter cases, the two metals are always on a line parallel to the axial direction of the trigonal bipyramid at the phosphorus. In our case, such an arrangement is virtually impossible: given the disposition of the three acidic residues with respect to the overall direction of the DNA backbone, the two metal sites have to be more or less in the equatorial plane of the bipyramid. In the model shown in Figure 11, the coordination of the two metals is not completely

satisfactory. We would have preferred to bring the carboxylate group of Asp74 into a bridging position where each oxygen is a ligand to one of the two metals. Although not impossible, this could not be done in a very satisfactory way. On the other hand, in the arrangement shown it is conceivable that a water (105) from the first coordination shell of Mg1, activated by Asp74, becomes the attacking nucleophile. In view of the speculative nature of this model, a discussion of such detailed mechanistic aspects is, however, premature.

Our new structural studies have not resolved all of the open mechanistic questions but have provided experimental evidence for two metal binding sites involving the acidic residues 90/74 and 74/45, respectively. In addition, new possibilities for the DNA conformation at the transition state of the reaction have emerged. The simple picture that recognition is structurally coupled to catalysis primarily by the special nature of the productive substrate conformation can be maintained, but the productive conformation may be more different from that observed in the crystals with cognate DNA fragments than previously thought. This has important consequences for how one tries to understand the characteristics of different metal cofactors or their combinations at the structural level. Finally, one wonders how the conserved and nonconserved features of the active site structure and residues in *Eco*RV, *Eco*RI, *Bam*HI, and *Pvu*II relate to common and different aspects in their mechanism of phosphodiester cleavage.

ACKNOWLEDGMENT

We thank U. Selent and A. Pingoud (University of Giessen) and W. Bannwarth (F. Hoffmann-LaRoche, Ltd., Basel) for providing oligonucleotide samples and S. Halford (University of Bristol) for providing purified enzyme. We are grateful to S. Halford for critical comments on the manuscript and for providing results prior to publication.

REFERENCES

Baldwin, G. S., & Halford, S. E. (1995) *Biochemistry* (third paper of three in this issue).

Beese, L. S., & Steitz, T. A. (1991) *EMBO J. 10*, 25−33.

Bruenger, A. T., Kuriyan, J., & Karplus, M. (1987) *Science 235*, 458−460.

Bruenger, A. T., Krukowski, A., & Erickson, J. (1990) *Acta Crystallogr. A46*, 585−593.

Dorner, L. F., & Schildkraut, I. (1994) *Nucleic Acids Res. 22*, 1068−1074.

Eckstein, F. (1985) *Annu. Rev. Biochem. 54*, 367−402.

Engh, R. A., & Huber, R. (1991) *Acta Crystallogr. A47*, 392−400.

Fitzgerald, P. M. D. (1988) *J. Appl. Crystallogr. 21*, 273−278.

Gerper, P., & Müller, K. (1994) *J. Comput.-Aided Mol. Des.* (manuscript submitted).

Glusker, J. P. (1991) *Adv. Protein Chem. 42*, 1−76.

Grasby, J. A., & Connolly, B. A. (1992) *Biochemistry 31*, 7855−7861.

Guthrie, J. P. (1976) *J. Am. Chem. Soc. 99*, 3991−4001.

Halford, S. E., Taylor, J. D., Vermote, C. L. M., & Vipond, I. B. (1993) in *Nucleic Acids and Molecular Biology* (Eckstein, F., & Lilley, D. M. J., Eds.) Vol. 7, pp 47−69, Springer-Verlag, Berlin, Heidelberg.

Hodel, A., Kim, S.-H., & Bruenger, A. T. (1992) *Acta Crystallogr. A48*, 851−858.

Jeltsch, A., Alves, J., Wolfes, H., Maass, G., & Pingoud, A. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 8499−8503.

Jones, T. A., Zou, J.-Y., Cowan, S. W., & Kjeldgaard, M. (1991) *Acta Crystallogr. A47*, 110−119.

Kabsch, W. (1988) *J. Appl. Crystallogr. 21*, 916−924.

Luke, P. A., McCallum, S. A., & Halford, S. E. (1987) *Gene Amplif. Anal. 5*, 183−205.

Luzzati, V. (1953) *Acta Crystallogr. 6*, 142−152.

Newman, M., Strzelecka, T., Dorner, L. F., Schildkraut, I., & Aggarwal, A. K. (1994) *Nature 368*, 660−664.

Rosenberg, J. M. (1991) *Curr. Opin. Struct. Biol. 1*, 104−113.

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.

Steitz, T. A. (1993) *Curr. Opin. Struct. Biol. 3*, 31−38.

Taylor, J. D., & Halford, S. E. (1989) *Biochemistry 28*, 6198−6207.

Taylor, J. D., Badcoe, I. G., Clarke, A. R., & Halford, S. E. (1991) *Biochemistry 30*, 8743−8753.

Thielking, V., Selent, U., Köhler, E., Landgraf, A., Wolfes, H., Alves, J., & Pingoud, A. (1992) *Biochemistry 31*, 3727−3732.

Vermote, C. L. M., & Halford, S. E. (1992) *Biochemistry 31*, 6082−6089.

Vipond, I. B., & Halford, S. E. (1993) *Mol. Microbiol. 9*, 225−231.

Vipond, I. B., & Halford, S. E. (1995) *Biochemistry* (second paper of three in this issue).

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*, 1781−1795.

BI9416263