Simha, D., Palejwala, V. A., & Humayun, M. Z. (1991) *Biochemistry* (preceding paper in this issue).
Singer, B., Spengler, S. J., Chavez, F., & Kusmierek, J. T. (1987) *Carcinogenesis 8*, 745–747.
Walker, G. C. (1984) *Microbiol. Rev. 48*, 60–93.
Wells, R. D., Collier, D. A., Hanvey, J. C., Shimizu, M., & Wohlrab, F. (1988) *FASEB J. 2*, 2939–2949.
Witkin, E. M. (1976) *Bacteriol. Rev. 40*, 869–907.

*Eco*RV Restriction Endonuclease Binds All DNA Sequences with Equal Affinity†

John D. Taylor,‡ Ian G. Badcoe, Anthony R. Clarke, and Stephen E. Halford*
Department of Biochemistry, Centre for Molecular Recognition, University of Bristol, Bristol BS8 1TD, U.K.
Received April 19, 1991; Revised Manuscript Received June 18, 1991

ABSTRACT: In the presence of $MgCl_2$, the *Eco*RV restriction endonuclease cleaves its recognition sequence on DNA at least a million times more readily than any other sequence. In this study, the binding of the *Eco*RV restriction enzyme to DNA was examined in the absence of Mg^{2+}. With each DNA fragment tested, several DNA–protein complexes were detected by electrophoresis through polyacrylamide. No differences were observed between isogenic DNA molecules that either contained or lacked the *Eco*RV recognition site. The number of complexes with each fragment varied with the length of the DNA. Three complexes were formed with a DNA molecule of 55 base pairs, corresponding to the DNA bound to 1, 2, or 3 molecules of the protein, while >15 complexes were formed with a DNA of 381 base pairs. A new method was developed to analyze the binding of a protein to multiple sites on DNA. The method showed that the *Eco*RV enzyme binds to all DNA sequences, including the *Eco*RV recognition site, with the same equilibrium constant, though two molecules of the protein bind preferentially to adjacent sites on the DNA in a cooperative fashion. All of the complexes with a substrate that contained the *Eco*RV site dissociated upon addition of competitor DNA, but when the competitor was mixed with $MgCl_2$, a fraction of the substrate was cleaved at the *Eco*RV site. The fraction cleaved was due mainly to the translocation of the enzyme from nonspecific sites on the DNA to the specific site.

The occupancy of a given site on DNA by a protein is a function of not only the affinity of the protein for that site but also its affinity for the rest of the DNA molecule: the latter is governed by both its affinity for alternative sites on the DNA and the number of the alternatives (von Hippel & Berg, 1986). Many proteins fulfill their biological functions by simply binding to a specific site on DNA, and, in these cases, discrimination between DNA sequences can arise only from differences in binding energies. In other cases such as the restriction enzymes, the protein binds to a specific DNA sequence and catalyzes a reaction. For an enzyme, the discrimination between alternate substrates can stem from either the binding or the catalytic reaction, or both (Jencks, 1975). The type II restriction endonucleases recognize short DNA sequences, typically 4–8 bp,[1] and cleave both strands of the DNA at their recognition sites (Bennett & Halford, 1989; Rosenberg, 1991). These enzymes need Mg^{2+} as a cofactor; they have no catalytic activity in the absence of divalent cations. To date, the most extensively studied restriction enzyme is *Eco*RI. In the absence of Mg^{2+}, this protein binds its recognition sequence much more tightly than any other DNA sequence (Halford & Johnson, 1980; Terry et al., 1983). The difference in binding energies makes a major contribution to the specificity of *Eco*RI for DNA cleavage at its recognition site (Thielking et al., 1990; Lesser et al., 1990).

The *Eco*RV endonuclease is one of the few restriction enzymes apart from *Eco*RI to have been overproduced (Bougueleret et al., 1985) and crystallized (D'Arcy et al., 1985). High-resolution structures of the *Eco*RV nuclease have been solved by X-ray crystallography for the free protein in the absence of DNA, the protein bound to a duplex oligonucleotide containing the *Eco*RV recognition sequence, and the protein bound to nonspecific DNA [F. Winkler, personal communication; reviewed by Rosenberg (1991)]. *Eco*RV has no homology to *Eco*RI in either primary or tertiary structures (Bougueleret et al., 1984; Winkler et al., 1991). The *Eco*RV enzyme is a dimer of identical subunits, and it cleaves both strands of the DNA at the center of its recognition site, GATATC (Schildkraut et al., 1984; D'Arcy et al., 1985). The enzyme is extremely specific for this sequence. Under standard reaction conditions, k_{cat}/K_m for the reaction of *Eco*RV at its recognition site on pAT153 is 3×10^7 M^{-1} s^{-1}, while k_{cat}/K_m for double strand breaks at the next best site on this plasmid is 36 M^{-1} s^{-1}; the next best site is one of the locations on pAT153 that has the sequence GTTATC (Taylor & Halford, 1989). This 1 bp change in DNA sequence therefore produces a million-fold reduction in k_{cat}/K_m. The other sites on pAT153 that differ from the cognate site by 1 bp are cleaved even more slowly (Halford et al., 1986). In oligonucleotide substrates for *Eco*RV, most base analogues also cause large reductions in activity (Fliess et al., 1988; Mazzarelli et al., 1989; Newman et al., 1990).

When the *Eco*RV restriction enzyme is bound to its cognate site on DNA, the protein has a high affinity for Mg^{2+} ions, but when bound to noncognate sequences, it has a low affinity for Mg^{2+} (Taylor & Halford, 1989). This accounts for at least part of the difference in catalytic rates at cognate and noncognate sites. It also accounts for a difference in the mode of DNA cleavage: concerted double-strand scission at the recognition site; sequential single-strand breaks at other sites

† This work was supported by the Science and Engineering Research Council.
‡ Present address: Department of Biochemistry, Duke University Medical Center, Durham, NC 27710.

[1] Abbreviations: bp, base pair(s); BME, β-mercaptoethanol; BSA, bovine serum albumin; dNTP, deoxyribonucleotide triphosphate; EDTA, ethylenediaminetetraacetic acid: kb, 1000 bp; M_r, relative molecular mass; Tris, tris(hydroxymethyl)aminomethane.

0006-2960/91/0430-8743$02.50/0

(Halford & Goodall, 1988; Taylor & Halford, 1989). Presumably, both subunits of the protein bind Mg^{2+} during the reaction at the cognate site, but, at a noncognate site, only one subunit at a time is likely to bind Mg^{2+}. One consequence of the different modes of cleavage is that the effective specificity of *Eco*RV in vivo is even higher than that measured in vitro, due to proofreading by DNA ligase (Taylor et al., 1990).

This report describes equilibrium binding studies on the *Eco*RV endonuclease with DNA fragments that either contain or lack the *Eco*RV recognition site. The binding was carried out in the absence of Mg^{2+} to prevent DNA cleavage. In its catalytic reaction, the *Eco*RV enzyme has a compulsory-order mechanism, binding first to DNA and then to Mg^{2+} (Taylor & Halford, 1989). Hence, DNA–protein complexes formed in the absence of Mg^{2+} should be equivalent to intermediates in the catalytic reaction. Binding was monitored primarily by the gel shift method (Garner & Revzin, 1981; Fried & Crothers, 1981). For restriction enzymes, another method for monitoring the binding of the nuclease to its recognition site is the preferential cleavage assay (Jack et al., 1982), and this was also used.

Experimental Procedures

Proteins. The *Eco*RV restriction endonuclease was purified and its concentration determined as before (Luke et al., 1987; Halford & Goodall, 1988). The molarities cited are for the enzyme dimer of M_r 57 000 (D'Arcy et al., 1985; Luke et al., 1987). Apart from *Eco*RI (purified here by C. N. Parker), all other enzymes for DNA manipulation were obtained from either Gibco BRL or Boehringer Mannheim and were used as advised by the supplier. Nuclease-free solutions of BSA (Sigma) and proteinase K (Boehringer Mannheim) were prepared as described in Taylor and Halford (1989).

Plasmids. Three plasmids were used: pAT153 (Twigg & Sherratt, 1980), pAT153b (Taylor & Halford, 1989), and pACYC184/5 (Taylor, 1991). The sequence of pAT153b differs from that of pAT153 by the insertion of an 8 bp *Bgl*II linker at the *Eco*RV site. Plasmid pACYC184/5 is a 3.9 kb derivative of pACYC184 (Chang & Cohen, 1978) that has one site for *Eco*RV located 1.7 kb from the *Eco*RI site (Taylor, 1991). The plasmids were used to transform *Escherichia coli* HB101, and the DNA was purified as described previously (Halford & Goodall, 1988).

DNA Fragments. Fragments of DNA were obtained by either oligonucleotide synthesis (carried out by L. Hall, this department, using a Dupont Coder 300 DNA Synthesizer) or by restriction digests on either pAT153 or pAT153b followed by electrophoresis through polyacrylamide. The fragments were labeled at the 5′-ends by using the Klenow fragment of DNA polymerase I with [α-^{32}P]dATP (NEN Research Products, 3000 Ci/mmol) and, if needed, one or more additional dNTPs (Sambrook et al., 1989). After labeling, flush ends were generated by adding all four dNTPs to the Klenow reactions. The lengths of the DNA fragments given below are for the duplex DNA after end-filling, and the positions of restriction sites are numbered from the first base of the recognition sequence.

Specific 381-mer: The *Eco*RI–*Bam*HI fragment from position 0 to 375 on pAT153. This includes the *Eco*RV site at position 185.

Nonspecific 389-mer: An *Eco*RI–Ba*m*HI fragment equivalent to the above but from pAT153b. The 8 bp DNA inserted at the *Eco*RV site on this plasmid disrupts the sequence and neither cognate or noncognate sites for *Eco*RV were regenerated (Taylor & Halford, 1989).

Specific 235-mer: The *Eco*RI–*Nhe*I fragment from position 0 to 229 on pAT153.

Specific 100-mer: The *Mva*I–*Nhe*I fragment from position 131 to 229 on pAT153. All preparations of this fragment showed two bands of similar mobilities when analyzed by electrophoresis through polyacrylamide. Both contained the *Eco*RV site from pAT153. The stocks of *Mva*I were probably contaminated with an exonuclease.

Nonspecific 55-mer: Two self-complementary oligonucleotides of 51 bases were synthesized that, when annealed, generated a 47 bp duplex DNA with four-base 5′-extensions at both ends, one matching an *Eco*RI extension and the other *Nco*I. The duplex region had the same sequence as positions 1707–1753 on pAT153. This section contains the next best site for the *Eco*RV nuclease on pAT153, GTTATC at 1734 (Taylor & Halford, 1989), but no cognate site for *Eco*RV.

Specific 55-mer: Identical with that immediately above, except that the GTTATC sequence for the next best site was replaced with the recognition sequence for *Eco*RV, GATATC.

Gel Shift Assays. These were performed as described by Garner and Revzin (1981). The binding buffer was 50 mM Tris·HCl, 100 mM NaCl, 10 mM BME, 100 μg/mL BSA, and 0.1 mM EDTA, pH 7.5. The stock solutions of *Eco*RV were diluted to the appropriate concentration directly before use in the same buffer supplemented with 2 mM spermine. The reactions contained up to 0.1 nM ^{32}P-labeled DNA fragment and 0–300 nM *Eco*RV restriction endonuclease in 20 μL of binding buffer. (In the majority of experiments, the only DNA in the binding reactions was the ^{32}P-labeled substrate, but, for some experiments noted below, the reactions also contained various amounts of unlabeled pAT153b). The mixtures were incubated for 15 min at 20 °C before adding 8 μL of loading buffer (binding buffer with 40% (w/v) sucrose and 0.01% (w/v) bromophenol blue). Aliquots (15 μL) from these mixtures were immediately loaded on to 6% polyacrylamide gels and subjected to electrophoresis at room temperature. The voltage across the gels was 3 V/cm during loading and 13 V/cm both before and after loading: the gels were run for 2 h before use to ensure constant temperature. The electrophoresis buffer was TBE (Sambrook et al., 1989). [The following changes to the protocol made no difference to the results: leaving BSA out of the binding buffer; altering the incubation time of the DNA–protein mixtures from 0.5 to 30 min before adding the loading buffer; replacing the sucrose in the loading buffer with either 15% Ficoll 400 (Pharmacia LKB) or 33% glycerol; varying the voltage for electrophoresis or the percent polyacrylamide in the gels.] After electrophoresis, the gels were fixed in 1:1:8 acetic acid/methanol/H_2O, dried, and analyzed by autoradiography with preflashed Hyperfilm MP (Amersham International) at −70 °C. Band intensities on the autoradiographs were measured in a Joyce-Loebl densitometer (with software for data capture and analysis from A. Halestrap, this department). DNA concentrations were evaluated from the band intensities by reference to a linear scale established with known amounts of DNA. To ensure that the measured intensities fell within the linear range, the autoradiographs for each gel were generally taken at several different exposures (4–40 h).

Preferential Cleavage Assays. The assays were carried out on *Eco*RV essentially as in previous studies on the *Eco*RI restriction enzyme (Jack et al., 1982; Terry et al., 1983). The following DNA fragments were used as substrates: pACYC184/5 linearized with *Eco*RI; the nonspecific 389-mer; the specific 381- and 55-mers. They were labeled with [α-^{32}P]dATP and the ends filled in as above. Samples (20 μL) containing the substrate DNA ($\leq$0.1 nM) and the *Eco*RV enzyme (0–20 nM) in binding buffer at 20 °C were mixed with aliquots (5 μL) from a solution of binding buffer containing 100 nM unlabeled pAT153 and 50 mM $MgCl_2$. These were left for 30 s before terminating the reaction by adding 8 μL

of stop mix (0.1 M EDTA, 0.1 M Tris·HCl, 40% sucrose, 0.1% bromophenol blue, pH 8.0) followed by 1 μL of proteinase K (20 mg/mL). The solutions were incubated for 30 min, and the DNA was then analyzed by electrophoresis through either 6% polyacrylamide or 1% agarose. For the 381- and the 55-mers, the relative concentrations of the substrate DNA and the ^{32}P-labeled *Eco*RV product were determined by densitometry, as described above, and the agarose gels were used solely to check that not all of the pAT153 had been cleaved by *Eco*RV before the termination of the reaction: this was the case for all reactions reported here. For pACYC184/5, substrate and product concentrations were determined by scintillation counting on slices from the agarose gels (Halford & Goodall, 1988).

Computer Methods. Data fitting to single equilibrium constants was by nonlinear regression with ENZFITTER (Biosoft, Cambridge, U.K.). Multiple equilibrium constants were fitted to experimental data by numerical integration with FACSIMILE (Chance et al., 1977).

RESULTS

DNA Binding by EcoRV. The binding of the *Eco*RV restriction endonuclease to DNA in the absence of Mg^{2+} ions was examined by both the gel shift method (Garner & Revzin, 1981; Fried & Crothers, 1981) and by filter binding [the latter experiments followed the procedure used previously with the *Eco*RI enzyme (Halford & Johnson, 1980)]. DNA–protein complexes were readily detected by gel shift (Figure 1), but when the same mixtures of DNA and protein were applied to nitrocellulose filters, virtually none of the DNA was retained on the filters (data not shown). Perhaps the complexes have very low filter retention efficiencies, though filter binding studies on *Eco*RV have been reported previously (Fliess et al., 1988).

Gel shift experiments were carried out initially with a pair of isogenic DNA fragments: one of 381 bp that contains a single copy of the *Eco*RV recognition sequence and a second of 389 bp that differs from the first only in the disruption of the *Eco*RV site by an additional 8 bp DNA (Figure 1a,b). The gel shift method is based on DNA–protein complexes having different electrophoretic mobilities from the free DNA. It can be used to separate the free DNA from DNA–protein complexes, but complexes with different protein stoichiometries on the same DNA can also be separated from each other (Fried, 1989). Consequently, if the *Eco*RV restriction enzyme were to bind to DNA only at is recognition sequence, gel shift analysis would show a single retarded band of DNA with a specific 381-mer but no retarded bands with the 389-mer. However, if the protein were to bind to both the recognition site and to other DNA sequences, additional bands of retarded DNA would be observed with the 381-mer, and these additional bands would also be formed with the 389-mer. But, given the extreme specificity of the *Eco*RV restriction enzyme for DNA cleavage at its recognition site (Taylor & Halford, 1989), one might expect to observe additional bands only at very high concentrations of the protein. None of these expectations were met.

Following the electrophoresis of the specific 381-mer with the *Eco*RV enzyme, at least 15 discrete bands of DNA were counted on the autoradiographs (Figure 1a). The protein concentration needed to produce this number of bands was only slightly higher than that needed for the formation of the first shifted band. Both the free DNA and several retarded bands were found together in some samples. Electrophoresis of the 389-mer with the *Eco*RV enzyme gave a similar series of discrete bands (Figure 1b). For both the specific 381-mer and the nonspecific 389-mer, the decrease in the concentration of

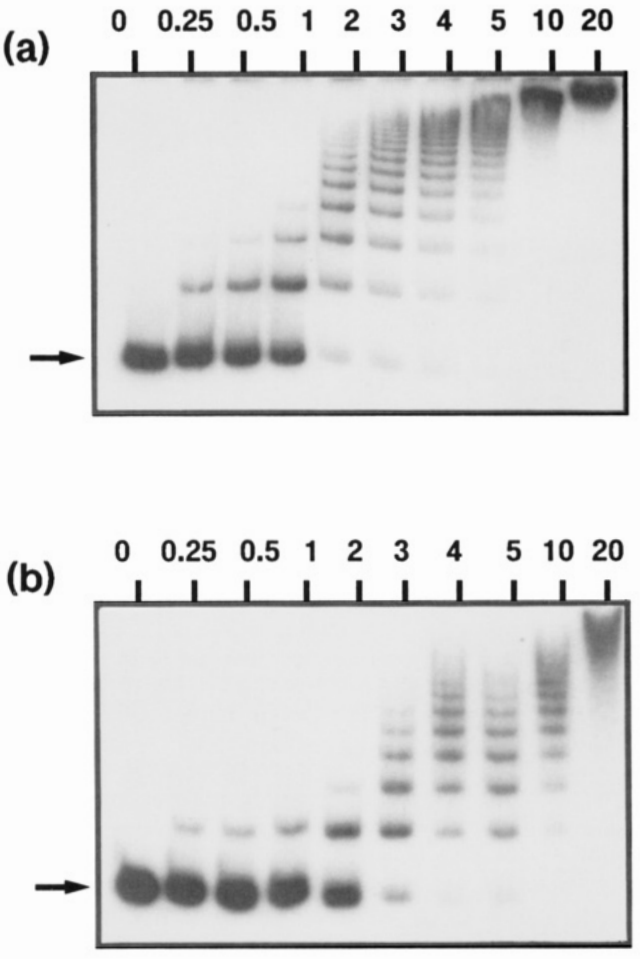

FIGURE 1: Gel shift assays on DNA with and without *Eco*RV sites. The reactions contained the *Eco*RV restriction endonuclease at the concentrations (nM) indicated above each lane and approximately 0.1 nM ^{32}P-labeled DNA in binding buffer. Lane 0 is in the absence of the enzyme. For panel a, the DNA was the specific 381-mer from pAT153. For panel b, the DNA was the nonspecific 389-mer from pAT153b. Aliquots from each mixture were subjected to electrophoresis through 6% polyacrylamide, and the gel was analyzed by autoradiography. The mobility of the free DNA is marked on the left of both autoradiographs.

the free (unretarded) DNA, caused by increasing concentrations of the enzyme, was measured by densitometry (Figure 2). No significant difference was found between the two DNA molecules.

The simplest explanation for the multiple bands of DNA is that the least retarded band in the gel shifts (immediately above the free DNA) corresponds to the DNA bound to 1 molecule of the enzyme dimer while the second retarded band is the DNA complexed with 2 molecules of the enzyme and so forth. This explanation implies that the *Eco*RV restriction enzyme binds to DNA with little, if any, sequence specificity. The *Eco*RV recognition site cannot be the prime requirement for the association of this protein with DNA: only one of the two DNA molecules used in Figure 1 contains the site. Most sequences that differ from the *Eco*RV recognition site by 1 bp can also be cleaved by this enzyme, albeit slowly (Halford et al., 1986). However, the number of bands seen by gel shift (Figure 1) far exceeds not only the number of cognate sites for *Eco*RV but also the number of noncognate sites: the 381-mer has only one noncognate site (GAAATC at position 91). Moreover, if the *Eco*RV enzyme bound specifically to one site on these DNA molecules, the decrease in the concentration of free DNA with increasing levels of protein (Figure 2) should follow a rectangular hyperbola (Clore et al., 1982), but this was not observed (the deviations were especially

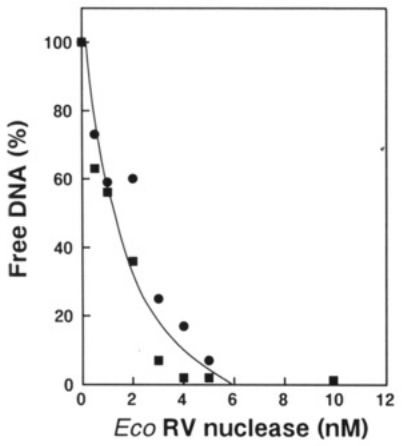

FIGURE 2: Binding curves for DNA with and without *Eco*RV sites. Mixtures of the *Eco*RV restriction enzyme and either the specific 381-mer (●) or the nonspecific 389-mer (■) were analyzed by electrophoresis. The percent of the input DNA found as free DNA was measured at each enzyme concentration. Each data point is an average from 3–5 separate experiments of the type shown in Figure 1. The line drawn is the theoretical line from the equation from Clore et al. (1982) for nonspecific binding of a protein to DNA without cooperativity (i.e., the same equilibrium constant for all potential sites on the DNA) fitted to the data assuming that $n = 25$. The optimal fit was with an equilibrium constant of 1.6×10^7 M^{-1}.

large as the level of free DNA approached zero). Instead, the binding curves could be reconciled to an equation from Clore et al. (1982) for nonspecific binding of a protein to multiple sites on the DNA (Figure 2).

Binding Buffers. The buffer used for the binding reactions was the same, except for the absence of $MgCl_2$, as the buffer in which the *Eco*RV nuclease shows both optimal activity for DNA cleavage at its recognition sequence and optimal discrimination against other DNA sequences (Halford & Goodall, 1988; Taylor & Halford, 1989). Binding reactions with the nonspecific 389-mer, which lacks the *Eco*RV site, were also carried out in the presence of 2 mM $MgCl_2$, and the products were analyzed by gel shift as above: for these samples, the EDTA in the electrophoresis buffer was also replaced by 2 mM $MgCl_2$. The gels showed the same series of bands as that in Figure 1b. The same number of DNA–protein complexes were formed in either the presence or absence of Mg^{2+} ions. However, in the presence of $MgCl_2$, the concentration of *Eco*RV enzyme needed to produce the retarded bands, and to reduce the amount of free DNA to zero (measured as in Figure 2), was about 3 times higher than that in the absence of $MgCl_2$ (data not shown).

When the concentration of NaCl in the binding buffer was raised above 100 mM, large numbers of discrete bands were again observed, but, as in the presence of $MgCl_2$, the products were formed only at higher protein concentrations than those used in Figure 1 (data not shown). Thus the affinity of the *Eco*RV enzyme for nonspecific DNA is affected by both $MgCl_2$ and NaCl. Part of the free energy for the association of a protein with DNA usually comes from the displacement of counterions such as Na^+ or Mg^{2+} (Record et al., 1978), and this probably applies to *Eco*RV.

DNA Length Dependency. Gel shift assays on the *Eco*RV restriction enzyme were also carried out with DNA fragments of different lengths (Figure 3). All three fragments used for Figure 3 contain the *Eco*RV recognition sequence. Two of these, the 235-mer and the 100-mer (Figure 3a,b), are restriction fragments from pAT153 that span the *Eco*RV site on this plasmid, just like the 381-mer. The third, the specific 55-mer (Figure 3c), is a synthetic DNA that carries the *Eco*RV site flanked by sequences from elsewhere in pAT153.

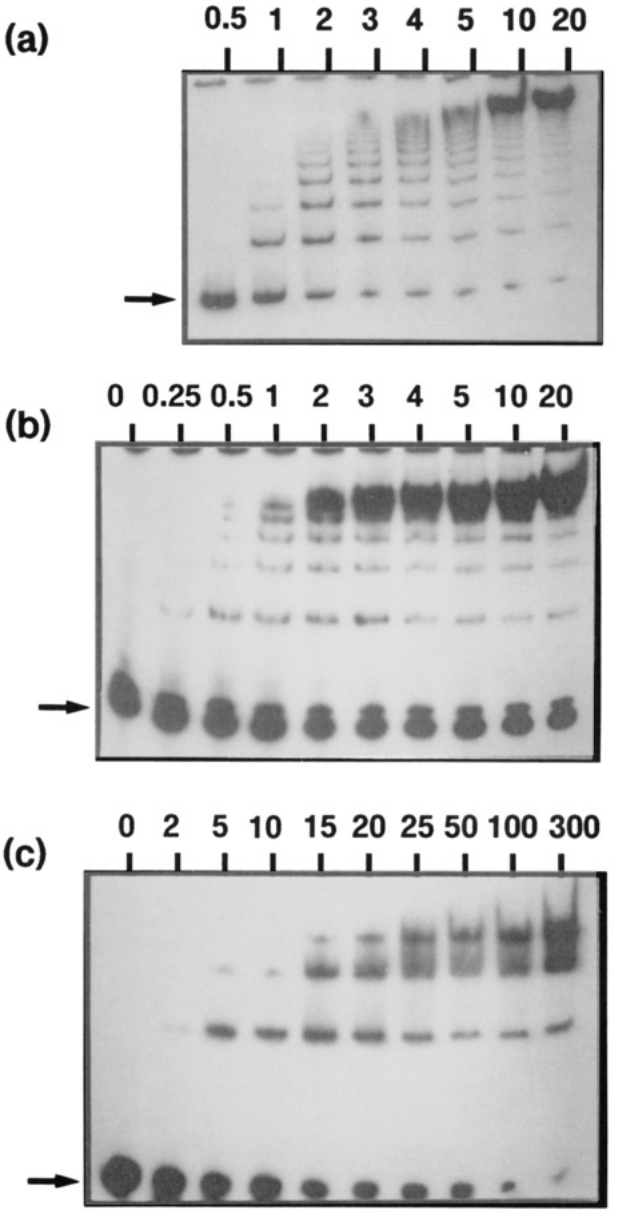

FIGURE 3: Gel shift assays on DNA fragments of different lengths. The reactions contained the *Eco*RV endonuclease at the concentrations (nM) indicated above each lane and approximately 0.1 nM ^{32}P-labeled DNA in binding buffer. Lane 0 is in the absence of the enzyme (not shown in panel a). Aliquots from each mixture were subjected to electrophoresis through 6% polyacrylamide, and the gels were analyzed by autoradiography. The mobility of the free DNA is marked on the left of all three autoradiographs. The DNA fragments were (a) the specific 235-mer, (b) the specific 100-mer, and (c) the specific 55-mer.

Further experiments were done with the nonspecific 55-mer, a DNA that is identical with the specific 55-mer except for 1 bp in the *Eco*RV site (see Experimental Procedures). Gel shift assays on the *Eco*RV enzyme with the nonspecific 55-mer gave results that were indistinguishable from those with the specific 55-mer (as in Figure 3c).

With all DNA fragments tested, multiple bands of DNA were observed after electrophoresis of the DNA–protein mixtures (Figure 3). However, the concentration of the *Eco*RV enzyme required to generate these products varied with the length of the DNA, more protein being needed with the shorter fragments. This is as expected for nonspecific binding of a protein to DNA (McGhee & von Hippel, 1974): a further analysis is given below. In addition, the number of bands that were resolved from the free DNA varied with the length of the DNA. The autoradiographs showed at least 12 discrete

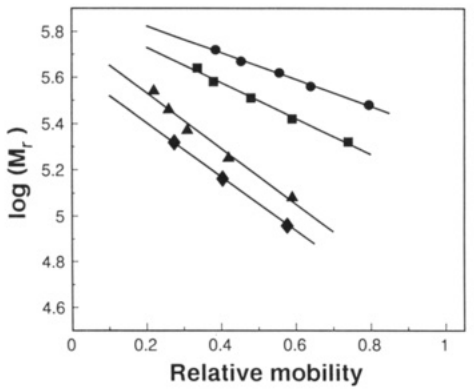

FIGURE 4: Electrophoretic mobilities from gel shifts. With each of the DNA fragments used in gel shift assays, the electrophoretic mobilities of the retarded bands were measured relative to the mobility of the free DNA, and these values were plotted against values of log M_r for the DNA–protein complexes. The values of log M_r were calculated by assuming that the first in a series of retarded bands from a given DNA denotes the DNA bound to one molecule of the *Eco*RV enzyme, the second the DNA bound to two molecules, and so forth: the M_r of the *Eco*RV nuclease (dimer) was taken as 57 000 and 1 bp as 662. Data from the following fragments were analyzed: the specific 381-mer (●); the specific 235-mer (■); the specific 100-mer (▲); and the nonspecific 55-mer (◆).

bands of retarded DNA with the 235-mer, 6 retarded bands with the 100-mer, but only 3 retarded bands with either the specific or nonspecific 55-mers. The variation in the number of bands is consistent with each series of bands coming from one DNA molecule bound to serially increasing numbers of protein molecules. If the resolved bands represent all possible protein–DNA stoichiometries on that DNA, then only three molecules of the protein dimer can bind to the 55-mer and six to the 100-mer. Hence, it appears that the *Eco*RV restriction enzyme covers about 15 bp when bound to DNA. If so, 15 bands should be visible with the 235-mer and 25 with the 381-mer. In both cases, the numbers of discrete bands that could be seen on the autoradiographs, 12 and 15, respectively (Figures 1a and 3a), were smaller than predicted. The deficit is probably due to inadequate separation of complexes containing large numbers of protein molecules at the top of the gels.

The estimate of 15 bp per protein is consistent with the crystallography on the *Eco*RV nuclease (F. Winkler, personal communication). The dimeric protein is in direct contact with 10 bp, but, given the overall shape of the protein, the center-to-center distance between two protein dimers cannot be much less than 15 bp.

Electrophoretic Mobilities. The multiple bands of DNA observed in gel shift assays with the *Eco*RV restriction endonuclease (Figures 1 and 3) are not necessarily due to DNA fragments being bound to 1, 2, or *n* molecules of the protein. An alternative is that the multiple bands correspond to configurational isomers of the DNA bound to one molecule of the protein (Fried, 1989): maybe the complexes differ from one another only in the location of the protein on the DNA, and each location for the protein causes a different electrophoretic mobility. Another possibility is that the bands correspond to conformational isomers of a single DNA–protein complex that could perhaps be generated by bending the DNA to differing extents (Gartenberg & Crothers, 1988).

To distinguish between these possibilities, M_r values were calculated for the postulated complexes (i.e., the DNA fragment and 1, 2, or more molecules of the *Eco*RV dimer), and these values were analyzed against the relative electrophoretic mobilities of the retarded bands (Figure 4). Linear relationships between log M_r and relative electrophoretic mobility have been established previously for several proteins that form

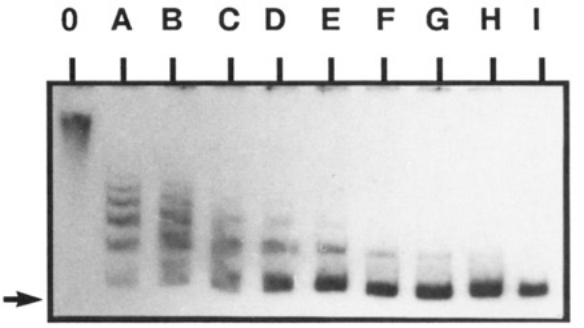

FIGURE 5: Competition assays. The *Eco*RV restriction enzyme and ^{32}P-labeled specific 381-mer were incubated in 20 μL binding buffer at 20 °C. Various amounts of pAT153b in 5 μL binding buffer were added and the samples were subsequently analyzed by electrophoresis through 6% polyacrylamide: an autoradiograph of the gel is shown with an arrow on the left to mark the mobility of the free DNA. The reactions contained 16 nM *Eco*RV enzyme, 0.05 nM 381-mer and either no pAT153b (lane O) or pAT153b at one of the following concentrations (nM); 0.12, 0.25, 0.5, 1, 2, 3.5, 7, 14 and 27 nM (lanes A–I respectively).

DNA–protein complexes with multiple stoichiometries and that give multiple bands in gel shift assays (Fried & Crothers, 1981; Bading, 1988; Fried, 1989). In each series of retarded bands produced by the *Eco*RV enzyme, the decrease in mobility from one band to the next was correlated to the change in M_r value expected from one additional molecule of the protein dimer (Figure 4).

Competition Assays. For DNA–protein complexes detected by gel shift, the distinction between specific and nonspecific complexes can be made by first mixing the protein with a radiolabeled DNA that carries the recognition sequence and then adding an unlabeled competitor DNA that lacks this sequence (Strauss & Varshavsky, 1984). Subsequent analysis by electrophoresis followed by autoradiography should show only the specific complexes. This method was tried with the *Eco*RV restriction enzyme. The *Eco*RV enzyme was first mixed with a ^{32}P-labeled DNA that contains the *Eco*RV recognition site; the reactions contained enough *Eco*RV enzyme to generate the multiply bound form of the DNA (Figure 5, lane 0). When these were challenged with increasing amounts of an unlabeled DNA that lacks the *Eco*RV site (pAT153b), none of the retarded bands survived (Figure 5, lanes A–I). If one of the retarded bands from DNA fragments carrying the *Eco*RV site had been due to preferential binding of the *Eco*RV nuclease to this site, the competitor DNA would have converted the multiple binding pattern to a single retarded band. This was not observed; instead, the competitor caused the dissociation of all of the complexes from the substrate DNA.

Preferential Cleavage Assays. For the *Eco*RI restriction enzyme, the fractional saturation of an *Eco*RI site on DNA can be measured by the preferential cleavage assay (Jack et al., 1982; Terry et al., 1983). In this assay, the restriction enzyme is added to a DNA substrate in the absence of Mg^{2+}, and the binding reaction is then mixed with a solution that contains both $MgCl_2$ and a competitor DNA. The competitor must contain the recognition site for the restriction enzyme so that both DNA molecules will be attacked in the presence of Mg^{2+}, but the hydrolytic reactions are terminated before the enzyme can catalyze repetitive turnovers. The concentration of the competitor DNA must be in excess of both the substrate DNA and the enzyme so that any free enzyme in the binding mixture will react with the competitor rather than the substrate. However, the enzyme in the binding mixture already on the substrate may be able to cleave that DNA before it dissociates from it (Jack et al., 1982). For preferential cleavage assays on *Eco*RV, pAT153 was used as the compe-

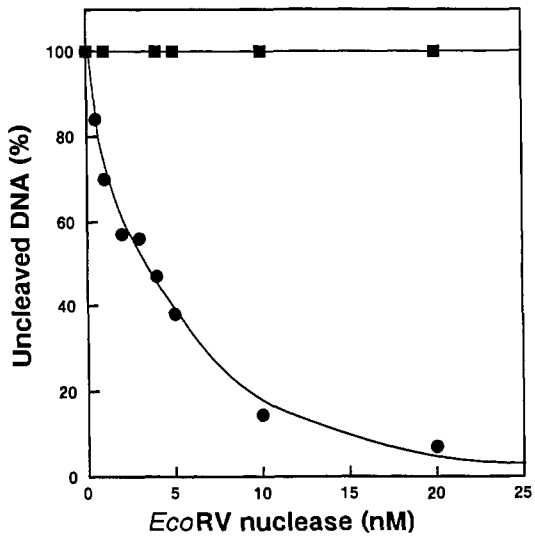

FIGURE 6: Preferential cleavage assays. Samples (20 μL) from solutions of binding buffer at 20 °C containing varied concentrations of the *Eco*RV restriction enzyme (as shown on the abscissa) and 0.1 nM ^{32}P-labeled DNA (given below) were mixed with aliquots (5 μL) of binding buffer containing 100 nM pAT153 and 50 mM $MgCl_2$. After 30 s, DNA cleavage was terminated as described under Experimental Procedures, and the ^{32}P-labeled DNA was analyzed by electrophoresis through 6% polyacrylamide followed by autoradiography. The fraction of the labeled DNA that had not been cleaved was assessed by densitometry and is given as a percent of the total. The labeled DNA was either the specific 381-mer (●) or the nonspecific 389-mer (■). The line drawn through the data with the 381-mer is the theoretical line for a rectangular hyperbola with K_{eff} at 2.5×10^8 M^{-1}.

titor DNA because it contains the *Eco*RV recognition site, in contrast to the competitor for the gel shifts in Figure 5 (pAT153b). Moreover, the concentration of competitor was set so that, at equilibrium in the absence of Mg^{2+}, all of the complexes on the substrate DNA would have dissociated (Figure 5). The time interval for the DNA cleavage reactions (30 s) was long enough for the *Eco*RV nuclease to make one double-strand break on DNA but not long enough for further reactions (Halford & Goodall, 1988).

Preferential cleavage assays on the *Eco*RV restriction endonuclease were carried out initially with the specific 381-mer and the nonspecific 389-mer. With the nonspecific 389-mer, the addition of $MgCl_2$ and the competitor lead to no detectable cleavage of this DNA (Figure 6). In the absence of Mg^{2+}, each molecule of this DNA binds several molecules of the *Eco*RV nuclease (Figure 1b). Presumably, all of the molecules of the enzyme dissociate from the nonspecific DNA before any cleavage reactions can occur. With the specific 381-mer, a fraction of the substrate was cleaved upon the addition of $MgCl_2$ and competitor DNA (Figure 6). The only product was that from cutting the *Eco*RV recognition site: products that could have arisen from reactions at other sites were not detected (data not shown). However, the fraction of the substrate DNA that was cleaved in this assay was much smaller than the fraction of the DNA bound to the *Eco*RV nuclease, as judged by gel shift. For example, at 4 nM *Eco*RV enzyme, ≥90% of the DNA molecules were bound to protein, and the majority of these carried >5 molecules of the *Eco*RV enzyme (Figures 1a and 2), yet only 50% of the DNA was cut in the preferential cleavage assay (Figure 6). The relationship between the fraction of the specific 381-mer that was cleaved by the *Eco*RV nuclease and the concentration of the enzyme was fitted to a rectangular hyperbola (Figure 6). The specific 381-mer yielded a value of 2.5×10^8 M^{-1} for K_{eff}, the effective equilibrium constant for binding the *Eco*RV restriction enzyme to its recognition sequence.

Preferential cleavage assays on the *Eco*RV nuclease were also carried out on two other DNA molecules that have the *Eco*RV site, a 3.9 kb plasmid and the specific 55-mer. Both yielded results that were similar to those with the specific 381-mer (Figure 6): the DNA was cleaved solely at the recognition site, but this was observed on the 3.9 kb DNA at lower concentrations of the enzyme than with the 381-mer, while the 55-mer needed higher concentrations (data not shown). In both cases, the increase in the fraction of cleaved DNA, with increasing concentrations of the enzyme, again followed a rectangular hyperbola and these gave values for K_{eff} at 1×10^9 M^{-1} with the 3.9 kb DNA and about 5×10^7 M^{-1} with the 55-mer. The three DNA molecules yield substantially different values for K_{eff}. This variation could have been due to the sequences flanking the *Eco*RV recognition sites, but exchanging these flanking sequences made no difference to *Eco*RV activity at the recognition site (Taylor, 1991). The value of K_{eff} thus depends on the length of the DNA molecule carrying the *Eco*RV site.

ANALYSIS[2]

Binding Parameters. If a DNA molecule has N bp and if one molecule of the protein covers l bp when bound to the DNA, the DNA contains $N - l + 1$ potential binding sites for the protein. But not all of the sites can be occupied simultaneously on account of the overlaps between sites (McGhee & von Hippel, 1974). The binding of a protein to multiple sites on a DNA molecule cannot be examined by a conventional Scatchard analysis. A Scatchard analysis fails because, after the binding of one or more molecules of the protein to the DNA, the number of free binding sites that remain depends on the location of the protein on the DNA. The only exceptions to this are when one site on the DNA has a much higher affinity for the protein than all other sites or, alternatively, when two or more proteins bind to nonoverlapping sites and the occupancy of each site can be measured [by, for example, gel shifts or DNase I footprints (Fried & Crothers, 1981; Brenowitz et al., 1986)]. These exceptions cannot apply to *Eco*RV as the enzyme bound equally well to isogenic DNA fragments that either possessed or lacked the *Eco*RV site, generating in both cases a large number of DNA–protein complexes (Figure 1). There is no reason to suggest that the binding sites for the *Eco*RV nuclease are nonoverlapping.

Analytical equations have been solved for multiple binding equilibria at overlapping sites on a homogeneous lattice (McGhee & von Hippel, 1974; Clore et al., 1982). However, these equations are not generally applicable to binding data obtained by the gel shift method of Garner and Revzin (1981) and Fried and Crothers (1981). First, the derivations were for long DNA molecules ($N > 1000$),[3] whereas the gel shift method is most effective with short molecules ($N < 500$). Second, the previous methods utilize only a fraction of the data

[2] The terms used in the analysis of binding equilibria are defined as follows: $K_{app,i}$, the apparent equilibrium constant for the binding of the ith molecule of the protein to the *molecule* of DNA; K_{eff}, the effective equilibrium constant for binding to the *recognition site* as measured by preferential cleavage; K_{int}, the intrinsic equilibrium constant for the binding of the protein to a *single isolated site* on the DNA; K_ω, the equilibrium constant for *protein–protein* associations (the binding constant to a site on the DNA adjacent to another molecule of the protein is therefore $K_{int} \times K_\omega$); l, length of DNA (bp) covered by 1 molecule of the protein; N, number of bp in the DNA molecule; n, maximum number of protein molecules that can bind to the DNA. One molecule of the *Eco*RV restriction enzyme is taken as the protein dimer.

[3] McGhee and von Hippel (1974) derived conditional probabilities in order to handle overlapping binding sites. The numerical solutions were obtained by summing the probabilities over infinitely long lattices. The same expressions were also used by Clore et al. (1982).

available from gel shifts. The experimental data needed for the equations of McGhee and von Hippel (1974) are values of v at different ligand concentrations, where v, the binding density, is defined as moles of ligand per mole of lattice residue. Though v may be a suitable parameter for binding data obtained by spectroscopic methods, it is ill-suited to gel shift assays. In principle, values of v could be calculated from the data in Figures 1 and 3 by first determining the concentrations of all of the separate DNA–protein complexes, then normalizing each against the number of proteins bound, and finally summing the total. However, if DNA molecules bound to 1, 2, or n molecules of the protein can be separated from each other by electrophoresis, then the gel shift data contain information about the individual binding reactions and only part of this information is utilized in the evaluation of v.

Multiple binding equilibria on DNA can also be followed by monitoring the concentration of the free DNA, the parameter measured in filter binding experiments (Clore et al., 1982). This approach was used to analyze the binding of *Eco*RV to the specific 381-mer and the nonspecific 389-mer (Figure 2), even though these fragments are shorter than ideal for this method. The data with both fragments were consistent with a model in which all of the potential sites on the DNA, regardless of their sequence, bind the *Eco*RV restriction enzyme with the same equilibrium constant and there is no cooperativity in binding to multiple sites (Figure 2). However, the more rigorous analysis given below revealed that the binding of *Eco*RV to DNA is cooperative, though the cooperativity constant is not large. The value of the equilibrium constant determined by the method from Clore et al. (1982), 1.6×10^7 M^{-1}, lies in between the values for binding to an isolated site on the DNA and for binding to a site adjacent to one already occupied by the protein (see below).

Statistical Analysis. Gel shifts can yield the concentrations of not only the free DNA (D) as a function of the protein concentration (P) but also the DP_1, the DP_2, and the DP_n complexes (the DNA bound to 1, 2, or n molecules of the protein; n is the maximum that can bind to the DNA). Each binding reaction can then be analyzed by a simple statistical relationship. The apparent equilibrium constant (K_{appl}) measured for the reaction

$$D + P \rightleftharpoons DP_1 \quad (1)$$

is related to the intrinsic equilibrium constant (K_{int}) for the binding of the protein to an individual site on the DNA by

$$K_{appl} = K_{int}[(N - l + 1)/1] \quad (2)$$

The term $[(N - l + 1)/1]$ is the statistical factor set by the number of possible states (i.e., configurational isomers) for the DP_1 complex divided by the number of possible states for the free DNA (Kelly et al., 1976). The thermodynamic basis for the statistical relationship between observed and intrinsic equilibrium constants is well known (Edsall & Gutfreund, 1983). The same relationship can be expressed more generally: for the binding of the ith molecule of protein to the DNA, the apparent equilibrium constant (K_{appi}) is given by

$$K_{appi} = K_{int}[X_i/X_{i-1}] \quad (3)$$

where X_i and X_{i-1} are the number of possible states for the DP_i and the DP_{i-1} complexes.

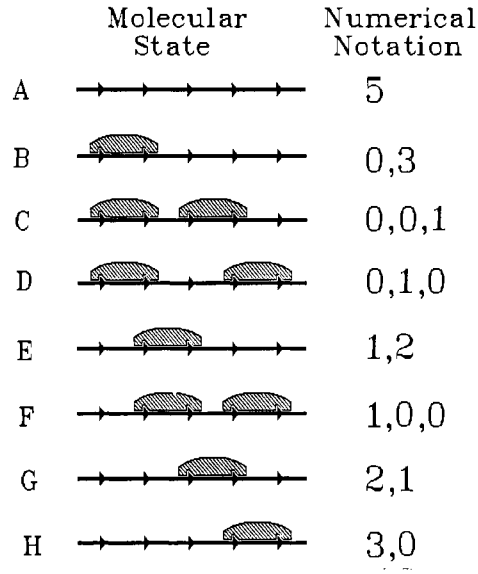

FIGURE 7: An algorithm to define all states in a DNA–protein interaction. To provide a systematic method of defining all possible complexes between a DNA molecule of N bp and a protein that covers l bp, a numerical notation was devised and is illustrated here for $N = 5$ and $l = 2$. The notation consists of a set of integers, each of which denotes a gap of ≥0 bp between two proteins or between a protein and the end of the DNA. State A is free DNA and has the notation "5". When the first protein is placed at the extreme left-hand site, the new structure (state B) is defined as "0,3". The next protein is placed immediately to the right of the first to give state C, which is noted as "0,0,1". At this point, no more proteins can be bound, and the second is then moved to the right in 1 bp steps until it reaches the end. In this example, only one more structure is generated, state D (notation "0,1,0"). The right-hand protein is then removed, and the remaining protein is moved 1 bp to its right to generate state E ("1,2"). Repetition of these processes covers all possible states. The numerical notation lists all states by starting with N, applying the following rules sequentially, and then ending when none of the rules can be applied: (i) If the last number is ≥l, subtract l and insert 0 before it (for example, "5" → "0,3" or "0,3" → "0,0,1"). (ii) If the last number is <l but >0, subtract 1 and add 1 to the number before it (for example, "0,0,1" → "0,1,0" or "2,1" → "3,0"). (iii) If the last number is 0, remove it, add $l - 1$ to the new last number, and add 1 to the new penultimate number (for example, "0,1,0" → "1,2" or "1,0,0" → "2,1"). The number of proteins bound in any state is given by the number of entries in the numerical notation minus 1 (i.e., "0,3" has one protein bound and "1,0,0" has two). The number of protein–protein contacts in each state is given by the number of zeroes in the notation, ignoring the first and last entries.

To obtain numerical values for the statistical factors, an algorithm was written to count the number of possible states at all possible DNA–protein stoichiometries (Figure 7).[4] The algorithm yields the number of isomers of the DP_1 complex (X_1 states, though X_1 must equal $N - l + 1$), the DP_2 complex (X_2 states), and likewise through to the DP_n complex (X_n states). It also counts the number of states for each complex in which two molecules of the protein are "touching" each other, i.e., where two proteins are bound to adjacent sites with no intervening base pairs. The algorithm can operate with any input values of N and l though the computer time needed to run the FORTRAN code for the algorithm increases to some extent as N is increased.[4]

The algorithm was used to analyze the binding of *Eco*RV to the specific 55-mer (Figure 3c). The *Eco*RV enzyme produced three retarded bands with this DNA, due to either 1, 2, or 3 molecules of protein binding to the DNA (Figure 4). With only three complexes, the relative concentrations of all three, and also the free DNA, could be measured at each protein concentration tested (see Figure 8). [The yield of each complex with the nonspecific 55-mer was no different from that with the specific 55-mer (data not shown).] Given the reaction mechanism

[4] All requests for the algorithm will be met. The request should include an E-mail number for the recipient and should be sent to I.G.B. at this address or by E-mail to I.BADCOE@UK.AC.BRISTOL. In a DEC MicroVAX 3900, the current code determines all states for the 55-mer, with l set at 15, in <1 s, but it requires about 100 s for the 381-mer.

Table I: Equilibrium Constants for the Binding of the *Eco*RV Restriction Enzyme to the Specific 55-mer

binding step[a]	K_{app} (M^{-1})[b]	statistical factor[c]	K_{int} (M^{-1})[d]
1	5.9×10^7	41/1	1.4×10^6
2	6.2×10^7	351/41	7.2×10^6
3	2.9×10^7	286/351	3.6×10^7

[a] The step refers to the binding of the first, second, and third molecule of the *Eco*RV enzyme to the specific 55-mer, corresponding to the three retarded bands in the gel shifts with this DNA (Figure 3c). [b] Values for the apparent equilibrium constants at each step were obtained by fitting the data to the scheme in eq 4. [c] The number of states for each complex was determined by the algorithm (Figure 7), and the statistical factors are ratios of these numbers. [d] The factors were used to convert the apparent constants to the intrinsic constants (as in eq 3).

$$D + P \rightleftharpoons DP_1 \rightleftharpoons DP_2 \rightleftharpoons DP_3 \qquad (4)$$

apparent values for the three successive equilibrium constants (K_{app1}, K_{app2}, and K_{app3}) were obtained by using FACSIMILE to fit to the concentrations of free protein (assumed to be equal to its total concentration; the protein was in large excess of the DNA), free DNA, and the complexes (Table I). In addition, the algorithm (Figure 7) was run with $N = 55$ and $l = 15$, and this gave the following values for the number of states for each species: 1 for D; 41 for DP_1 (the same as $N - l + 1$); 351 for DP_2; 286 for DP_3. The value for l came from the variation in the number of the complexes with the length of the DNA (Figure 3). The apparent equilibrium constants were then converted to the intrinsic constants by using the numbers from the algorithm (Table I).

If the *Eco*RV restriction enzyme had a higher affinity for the *Eco*RV site on the specific 55-mer compared to any other site, then K_{int} for binding the first molecule of the enzyme to the DNA ought to be larger than the values for binding the second and third enzymes. Alternatively, if the *Eco*RV nuclease showed no sequence specificity in binding to DNA, then all three equilibria in eq 4 should give the same value for K_{int}. Neither of these predictions were met. Instead, the values for K_{int} increased at each successive stage in the binding reaction (Table I).

Cooperativity. The progressive increments in the values of K_{int} indicate that the *Eco*RV enzyme binds cooperatively to DNA. The cooperativity can be analyzed by identifying how many of the 351 possible isomers of the DP_2 complex contain two molecules of the enzyme bound at adjacent sites and likewise how many of the 286 isomers of the DP_3 complex contain one or two sets of protein–protein contacts. This information is available from the algorithm (Figure 7). The mechanism in eq 4 was modified to

$$\begin{array}{l} \qquad\qquad\quad DP_2[0](325) \rightleftharpoons DP_3[0](165) \\ P + D \rightleftharpoons DP_1[0](41) \qquad\qquad\quad DP_3[1](110) \\ \qquad\qquad\quad DP_2[1](26) \rightleftharpoons DP_3[2](11) \end{array} \qquad (5)$$

Next to each intermediate, the number in square brackets specifies either zero protein–protein contacts, one set of contacts between two adjacent molecules of the enzyme, or two sets from three molecules and the number in parentheses denotes the number of possible isomers of that complex. The equilibria shown in eq 5 connect every complex to every other complex by one or more steps and are therefore sufficient to define the energetics of the system.

To test whether the scheme in eq 5 adequately describes the binding of the *Eco*RV enzyme to the 55-mer (Figure 3c), the concentrations of the free DNA, the DP_1, the DP_2, and the DP_3 complexes (measured at each concentration of the enzyme) were simultaneously fitted to the mechanism (Figure 8). The fit minimized the sum of the squares of the differences between model and data, varying only two equilibrium

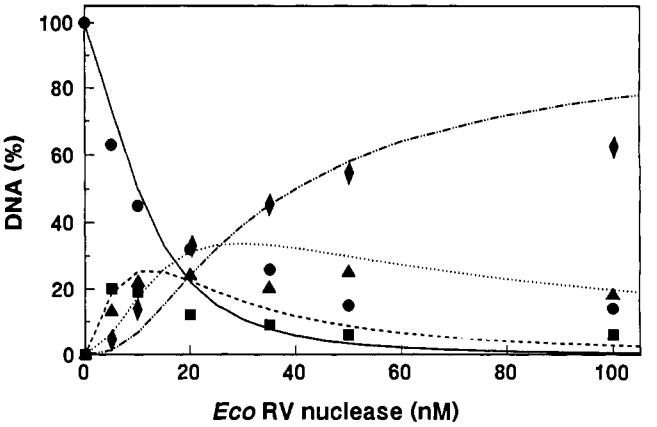

FIGURE 8: Analysis of the binding of *Eco*RV to DNA. Autoradiographs from gel shift assays on the binding of the *Eco*RV enzyme to the specific 55-mer were scanned by densitometry to determine the relative amounts of the DNA in its free form (●), in the first retarded band (■), in the second retarded band (▲), and in the third retarded band (◆). The amount of each form of the DNA is shown as a percent of the total. The values given are means from 5–9 repeat experiments at each concentration of the *Eco*RV nuclease: individual values were typically ±20% of the means. The lines drawn are theoretical lines from eq 5 with $K_{int} = 1.25 \times 10^6$ M^{-1} and $K_\omega = 75$. The lines are (continuous line) D; (dashes) DP_1; (dots) DP_2 (i.e., the sum of $DP_2[0]$ and $DP_2[1]$); (alternating dashes and dots) DP_3 (i.e., the sum of $DP_3[0]$, $DP_3[1]$, and $DP_3[2]$).

constants. The constants were K_{int}, the intrinsic binding constant to a single site on the DNA, and K_ω, the protein–protein interaction for two molecules bound immediately adjacent to each other. Each equilibration in eq 5 is then defined by either K_{int} alone or by $K_{int}K_\omega$, together with the appropriate statistical factor (eq 3). For example, the equilibration between $DP_2[0]$ and $DP_3[0]$ is given by $K_{int}(165/325)$ and that between $DP_2[0]$ and $DP_3[1]$ by $K_{int}K_\omega(110/325)$. The fitting yielded a value for K_{int} at 1.25×10^6 M^{-1} [5–95% confidence limits, $(1.0–1.5) \times 10^6$ M^{-1}], which is close to that determined above for the binding of the first molecule of the enzyme (1.4×10^6 M^{-1}; Table I). The fitted value for K_ω was 75 (5–95% limits, 57–101).

Considering that the mechanism in eq 5 contains a large number of binding reactions, the theoretical lines calculated with just two equilibrium constants were remarkably close to the experimental data points (Figure 8). Obviously, a closer fit could have been obtained by using multiple values for both K_{int} and K_ω, to account for the different sequences at each site on the DNA. In addition, the model neglected end effects. It is possible that the protein binds more strongly to the ends of linear DNA, as has been observed with other systems (Melançon et al., 1983). The algorithm (Figure 7) identifies the states where the protein is at the end of the DNA, and these could have given different values for K_{int}. Another possibility is that the protein on the end of the DNA overhangs the end so that it then covers less than l bp (B. Connolly, personal communication). Analysis with these more complicated models is not warranted at the present time.

Discussion

Binding Specificity. In the presence of Mg^{2+} ions, the *Eco*RV restriction enzyme cleaves DNA with an extremely high specificity for its recognition sequence, GATATC (Luke et al., 1987). A change of a single base pair in this sequence reduces k_{cat}/K_m by a factor of a million or more (Taylor & Halford, 1989). Yet, in equilibrium binding experiments on *Eco*RV in the absence of Mg^{2+}, every DNA fragment tested bound 1, 2, 3, or more molecules of the enzyme, giving rise to multiple bands in gel shifts (Figures 1, 3, and 4). The number of bands observed in each case was far larger than

the number of both cognate and noncognate sites for *Eco*RV so the protein must also bind to DNA sequences unrelated to the recognition sequence. Moreover, the enzyme bound equally well to isogenic DNA molecules that either contain or lack the *Eco*RV recognition sequence (Figure 2).

The shortest DNA fragments used here were 55 bp long. If the *Eco*RV nuclease covers 15 bp when bound to DNA, these fragments contain 41 potential binding sites for the protein (from $N - l + 1$). By itself, the lack of preferential binding to the specific 55-mer, relative to the nonspecific 55-mer, shows that the affinity of the *Eco*RV restriction enzyme for its recognition site cannot be more than 40 times higher than that for any other sequence on this fragment. However, the intrinsic equilibrium constant determined from the binding of the first molecule of *Eco*RV to the specific 55-mer is smaller than those for binding the second and third molecules (Table I). Instead of indicating a 40-fold preference for the recognition site, the data for the binding of *Eco*RV to the 55-mer are consistent with a model in which there is no difference in affinities between the recognition site and any other DNA sequence (Figure 8): i.e., the value of K_{int} for the binding of *Eco*RV to an individual site on the DNA is the same for all sites including the recognition site. The model accommodated the successive increments in the binding constants by using two equilibrium constants to describe binding to a site immediately adjacent to one already occupied by the protein (eq 5): the same constant as above for the DNA–protein interaction and an additional constant for the protein–protein interaction (Figure 8).

The degree of cooperativity in the binding of the *Eco*RV restriction enzyme to DNA is unlikely to be significant in its reaction mechanism. The difference between $K_{int}K_{\omega}$ and K_{int} corresponds to a standard free energy of −2.9 kcal/mol. This value is comparable to that from transferring the side chain of tyrosine from aqueous to organic solvent, −2.3 kcal/mol (Creighton, 1984). Hence, the free energy change for the cooperativity could easily arise if two molecules of the *Eco*RV enzyme next to each other on the DNA bury a small area on the surface of the protein that would otherwise be accessible to solvent.

Site Location. The lack of specificity in DNA binding by the *Eco*RV restriction enzyme creates a problem for this enzyme in locating its recognition site. However, two results from the preferential cleavage assays demonstrate that the *Eco*RV enzyme can translocate from a nonspecific site to the specific site without dissociating from the DNA. Firstly, preferential cleavage yielded values for K_{eff}, the effective equilibrium constant for binding the recognition site, that increased steeply with the length of the DNA (from 5×10^7 M^{-1} with 55 bp to 1×10^9 M^{-1} with 3900 bp). If the enzyme encountered the recognition site directly, the value for K_{eff} would be independent of chain length (Berg et al., 1981). The variation with chain length shows that the encounter must be influenced by the remainder of the DNA molecule. Secondly, the intrinsic equilibrium constant for the binding of *Eco*RV to any individual site on the DNA (including the recognition site: $K_{int} = 1.25 \times 10^6$ M^{-1}) is smaller than any of the values for K_{eff}, even that from the shortest DNA fragment used here. The preferential cleavage assay on *Eco*RV therefore cannot be measuring only the occupancy of *Eco*RV recognition sites by the protein at equilibrium. It must also include the fraction of the sites that can be reached by the protein before it dissociates from the DNA.

The transfer of a protein from nonspecific to specific DNA can be achieved by several different mechanisms (Berg et al., 1981; von Hippel & Berg, 1989). However, for some of these mechanisms, two segments of the DNA that are separated from each other along the contour have to be brought into close physical proximity. It is highly unlikely that any mechanism of this type could operate on a DNA molecule of 55 bp: the fragment is shorter than the persistence length of DNA (Hagerman, 1988). Consequently, the most plausible mechanism for translocation by the *Eco*RV restriction enzyme is "sliding", the one-dimensional diffusion of the protein along the DNA (von Hippel & Berg, 1989). When bound to DNA, the *Eco*RV enzyme is probably in continual motion along the lattice, and any measurement of the occupancy of a given site presumably yields a time-averaged value.

Catalytic Specificity. The lack of discrimination between DNA sequences shown by the *Eco*RV restriction enzyme in binding to DNA is obviously in marked contrast to its specificity in cleaving DNA at the recognition sequence. Previously, it had been noted that base analogues, replacing individual bases in oligonucleotide substrates for *Eco*RV, caused major reductions in k_{cat} but only minor changes in K_m, and it was suggested that *Eco*RV discriminates between alternative substrates primarily at the transition state of the reaction rather than at the initial binding of the substrate (Newman et al., 1990). The data described here are consistent with this view. Our equilibrium binding studies eliminate the possibility that the discrimination could be due to the initial association of the protein with the DNA. Instead, it must be due to the catalytic process.

The *Eco*RV nuclease is not the only enzyme involved in DNA metabolism that discriminates against noncognate substrates mainly in its catalytic reaction rather than at substrate binding. One example is DNA polymerase I: with a suitable template/primer combination, the dNTP that is complementary to the template is bound by this protein with the same affinity as noncomplementary dNTPs (Eger et al., 1991). Another is the *dam* methyltransferase (Bergerat & Guschlbauer, 1990).

Magnesium Ions. The catalytic reaction of the *Eco*RV restriction enzyme requires Mg^{2+} ions [or a similar divalent cation (Luke et al., 1987)]. The binding of Mg^{2+} may determine the catalytic specificity of the nuclease. DNA cleavage by *Eco*RV follows a compulsory-order mechanism:

$$E + S + M \rightleftharpoons ES + M \rightleftharpoons ESM \rightarrow EPM \qquad (6)$$

where the enzyme (E) binds its target site on the DNA and then the Mg^{2+} ions (M) before catalyzing the hydrolysis of phosphodiester bonds. But the affinity of the ES complex for Mg^{2+} varies with the DNA sequence (Taylor & Halford, 1989). When bound to its recognition site, the *Eco*RV enzyme has a high affinity for Mg^{2+} (Halford & Goodall, 1988). The equilibration between ES and ESM in eq 6 then has a negative ΔG, so Mg^{2+} ions effectively trap the enzyme at the recognition site. In contrast, when the protein is bound at a noncognate site for *Eco*RV, it has a low affinity for Mg^{2+} (Taylor & Halford, 1989). For the protein at a noncognate site, the equilibration between ES and ESM has a positive ΔG at all practicable concentrations of $MgCl_2$. Thus the nuclease cannot be trapped at a noncognate site by Mg^{2+} ions, and, instead, it will continue to "slide" along the DNA.

In the absence of $MgCl_2$, competitor DNA caused the dissociation of all of the complexes formed between *Eco*RV and a DNA fragment carrying the *Eco*RV site (Figure 5). But in the presence of $MgCl_2$, competitor DNA failed to prevent the substrate from being cleaved at the recognition site though DNA lacking the recognition site was not cleaved (Figure 6). The protein at the recognition site must be able to bind Mg^{2+} and remain on the DNA for long enough to cut it while all other complexes dissociate before DNA cleavage. The inability of Mg^{2+} to trap the enzyme at noncognate sites was also demonstrated by the gel shift experiments with the nonspecific

389-mer in the presence of $MgCl_2$.

In binding buffer supplemented with sufficient $MgCl_2$ for maximal rates, the value of K_m for the reaction of the *Eco*RV endonuclease at its cognate site on pAT153 [0.5 nM (Halford & Goodall, 1988)] is much smaller than the reciprocal of the equilibrium constant for binding to an individual site on the DNA ($1/K_{int}$ = 800 nM). In contrast, the kinetics of DNA cleavage by *Eco*RV at a noncognate site on pAT153 yielded an equilibrium dissociation constant of 3 μM (albeit in a different buffer) (Taylor & Halford, 1989): the latter value is much closer to $1/K_{int}$ than the K_m for the recognition site. Moreover, in the presence of Mg^{2+}, the reaction at the recognition site on pAT153 is not inhibited by nonspecific DNA while that at the noncognate site is inhibited (A. J. Goodall, J.D.T., and S.E.H., unpublished). At least part of these differences between cognate and noncognate sites can be accounted for by Mg^{2+} ions locking the enzyme onto DNA specifically at the recognition sequence.

Binding Energies. In the crystal structure of the *Eco*RV restriction endonuclease bound to its cognate DNA sequence in the absence of Mg^{2+} [F. Winkler, personal communication; reviewed by Rosenberg (1991)], there are a large number of hydrogen bonds, of the type proposed by Seeman et al. (1976), between the protein and functional groups on the edges of the bases. These hydrogen bonds have no equivalents in the structure of *Eco*RV bound to nonspecific DNA. Given typical values for hydrogen-bond energies (Creighton, 1984), the difference in hydrogen-bonding patterns ought to result in a large difference in free energy for binding to specific and nonspecific DNA sequences. But this was not found: the same binding constant was observed for all DNA sequences, and thus the specific and nonspecific complexes are isoenergetic. However, the structure of the DNA in the specific complex is distorted radically from the regular structure for B-DNA, and the structure of the protein is also altered (F. Winkler, personal communication).

The binding energy available from the recognition of the cognate DNA sequence appears to be dissipated within the DNA–protein complex, presumably in the distortion of the DNA and in the alterations to the conformation of the protein [viz. Jencks (1975)]. The active site in the specific complex may be organized so that it is ready to bind Mg^{2+} ions, while the active site in the complex at a nonspecific sequence lacks the requisite organization. Alternatively, Mg^{2+} ions may play a direct role in recognizing the target sequence: perhaps the ion is coordinated between the protein and the DNA in the specific complex while this bridging interaction is missing in the nonspecific complex. With *Eco*RV, the specific and nonspecific complexes are isoenergetic but they are not isosteric.

EcoRV Compared to EcoRI. The results and the conclusions from experiments on *Eco*RV differ in many respects from previous studies on the *Eco*RI restriction enzyme. First, the *Eco*RI enzyme binds more strongly to a 40 000 bp DNA molecule that has an *Eco*RI site compared to the isogenic DNA that lacks the site (Halford & Johnson, 1980), while the *Eco*RV enzyme binds equally well to 55 bp fragments with and without *Eco*RV sites. Second, the number of retarded bands seen in gel shift experiments with *Eco*RI was equal to the number of *Eco*RI sites on the DNA (Terry et al., 1985), in contrast to the multiple bands with *Eco*RV (Figures 1 and 3). Third, the fraction of the DNA cut by *Eco*RI in preferential cleavage assays is the same as the fraction of the DNA molecules that have bound the enzyme (Jack et al., 1982; Terry et al., 1983): no such equivalence was seen with *Eco*RV (Figures 2 and 6). From these and other experiments, the binding affinities of *Eco*RI for specific and nonspecific DNA differ by a factor of about 1×10^5 (Clore et al., 1982; Terry et al., 1983; Lesser et al., 1990). An isoschizomer of *Eco*RI, *Rsr*I, behaves similarily (Aiken et al., 1991). However, the *Taq*I restriction enzyme binds readily to both specific and nonspecific DNA (F. Barany, personal communication). For *Eco*RV, the ratio of binding affinities is 1.

In other respects, there are similarities between the *Eco*RI and *Eco*RV restriction enzymes. With both, it appears that at least some of the binding energy available from interacting with their recognition sites is dissipated within the DNA–protein complex. For *Eco*RI, this was demonstrated by the isolation of a mutant (Gln 111) that has almost no activity but binds to the cognate sequence more tightly than the wild-type enzyme (Wright et al., 1989): the mutation alters the protein at the site for Mg^{2+} (Rosenberg, 1991). In addition, the two enzymes use similar pathways for finding their recognition sites on DNA. It had been shown previously that the *Eco*RI enzyme "slides" along DNA in both locating and leaving its recognition site (Jack et al., 1982; Ehbrecht et al., 1985; Terry et al., 1985). But at equilibrium in the absence of Mg^{2+} ions, a DNA molecule bound by one molecule of enzyme would have *Eco*RI on its recognition site while *Eco*RV could be anywhere on the DNA.

ACKNOWLEDGMENTS

We thank Bernard Connolly, Annette Goodall, Alfred Pingoud, Christian Vermote, Barry Vipond, and Fritz Winkler for materials, for discussions, and for unpublished data.

Registry No. *Eco*RV restriction endonuclease, 83589-02-0.

REFERENCES

Aiken, C. R., Fisher, E. W., & Gumport, R. I. (1991) *J. Biol. Chem.* (in press).

Bading, H. (1988) *Nucleic Acids Res. 16*, 5241–5248.

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

Berg, O. G., Winter, R. B., & von Hippel, P. H. (1981) *Biochemistry 20*, 6929–6948.

Bergerat, A., & Guschlbauer, W. (1990) *Nucleic Acids Res. 18*, 4369–4375.

Bougueleret, L., Schwartzstein, 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.

Brenowitz, M., Senear, D. F., Shea, M. A., & Ackers, G. K. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 8462–8466.

Chance, E. M., Curtis, A. R., Jones, I. P., & Kirby, C. R. (1977) FACSIMILE: *A Computer Program for Flow and Chemistry Simulation, and General Initial Value Problems*, H. M. Stationery Office, London, U.K.

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

Clore, G. M., Gronenborn, A. M., & Davies, R. W. (1982) *J. Mol. Biol. 155*, 447–466.

Creighton, T. E. (1984) *Proteins*, W. H. Freeman, New York.

D'Arcy, A., Brown, R. S., Zabeau, M., Van Resandt, R. W., & Winkler, F. K. (1985) *J. Biol. Chem. 252*, 1987–1990.

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

Eger, B. T., Kuchta, R. D., Carroll, S. S., Benkovic, P. A., Dahlberg, M. E., Joyce, C. M., & Benkovic, S. J. (1991) *Biochemistry 30*, 1441–1448.

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

Fliess, A., Wolfes, H., Seela, F., & Pingoud, A. (1988) *Nucleic*

Acids Res. 16, 11781-11793.
Fried, M. G. (1989) *Electrophoresis 10*, 366-376.
Fried, M. G., & Crothers, D. M. (1981) *Nucleic Acids Res. 9*, 6505-6525.
Garner, M., & Revzin, A. (1981) *Nucleic Acids Res. 9*, 3047-3060.
Gartenberg, M. R., & Crothers, D. M. (1988) *Nature (London) 333*, 824-829.
Hagerman, P. J. (1988) *Annu. Rev. Biophys. Biophys. Chem. 17*, 265-286.
Halford, S. E., & Johnson, N. P. (1980) *Biochem. J. 191*, 593-604.
Halford, S. E., & Goodall, A. J. (1988) *Biochemistry 27*, 1771-1777.
Halford, S. E., Lovelady, B. M., & McCallum, S. A. (1986) *Gene 41*, 173-181.
Jack, W. E., Terry, B. J., & Modrich, P. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 4010-4014.
Jencks, W. P. (1975) *Adv. Enzymol. Relat. Areas Mol. Biol. 43*, 219-410.
Kelly, R. C., Jensen, D. E., & von Hippel, P. H. (1976) *J. Biol. Chem. 251*, 7240-7250.
Lesser, D. R., Kurpiewski, M. R., & Jen-Jacobson, L. (1990) *Science 250*, 776-786.
Luke, P. A., McCallum, S. A., & Halford, S. E. (1987) *Gene Amplif. Anal. 5*, 183-205.
Mazzarelli, J., Scholtissek, S., & McLaughlin, L. W. (1989) *Biochemistry 28*, 4616-4622.
McGhee, J. D., & von Hippel, P. H. (1974) *J. Mol. Biol. 86*, 469-489.
Melançon, P., Burgess, R. R., & Record, M. T., Jr. (1983) *Biochemistry 22*, 5169-5176.
Newman, P. C., Williams, M. D., Cosstick, R., Seela, F., & Connolly, B. A. (1990) *Biochemistry 29*, 9902-9910.
Record, M. T., Jr., Anderson, C. F., & Lohman, T. M. (1978) *Q. Rev. Biophys. 11*, 103-178.
Rosenberg, J. M. (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.
Schildkraut, I., Banner, C. D., Rhodes, C. S., & Parekh, S. (1984) *Gene 27*, 327-329.
Seeman, N. C., Rosenberg, J. M., & Rich, A. (1976) *Proc. Natl. Acad. Sci. U.S.A. 73*, 804-808.
Strauss, F., & Varshavasky, A. (1984) *Cell 37*, 889-901.
Taylor, J. D. (1991) Ph.D. Thesis, University of Bristol.
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.
Terry, B. J., Jack, W. E., Rubin, R. A., & Modrich, P. (1983) *J. Biol. Chem. 258*, 9820-9825.
Terry, B. J., Jack, W. E., & Modrich, P. (1985) *J. Biol. Chem. 260*, 13130-13137.
Thielking, V., Alves, J., Fliess, A., Maass, G., & Pingoud, A. (1990) *Biochemistry 29*, 4682-4691.
Twigg, A. J., & Sherratt, D. J. (1980) *Nature (London) 283*, 216-218.
von Hippel, P. H., & Berg, O. G. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 1608-1612.
von Hippel, P. H., & Berg, O. G. (1989) *J. Biol. Chem. 264*, 675-678.
Winkler, F. K., D'Arcy, A., Blocker, H., Frank, R., & van Boom, J. H. (1991) *J. Mol. Biol. 217*, 235-238.
Wright, D. J., King, K., & Modrich, P. (1989) *J. Biol. Chem. 264*, 11816-11821.

Regulation of CapZ, an Actin Capping Protein of Chicken Muscle, by Anionic Phospholipids†

Steven G. Heiss and John A. Cooper*
Department of Cell Biology and Physiology, Washington University School of Medicine, St. Louis, Missouri 63110
Received April 24, 1991; Revised Manuscript Received June 24, 1991

ABSTRACT: Chicken muscle CapZ, a member of the capping protein family of actin-binding proteins, binds to the barbed end of actin filaments and nucleates actin polymerization. No regulation of the capping protein family has been described. We report that micelles of phosphatidylinositol 4,5-bisphosphate (PIP_2) bind to CapZ and completely inhibit its ability to affect actin polymerization as measured by several independent assays. Higher concentrations of other anionic phospholipids also completely inhibit the activity of CapZ. Neutral phospholipids have no effect. Mixed vesicles of PIP_2 with phosphatidylcholine or phosphatidylethanolamine also inhibit CapZ, but addition of Triton X-100 both prevents and reverses PIP_2's inhibition of CapZ.

CapZ is an actin-binding protein that is located at the Z line in skeletal muscle and binds to the barbed ends of actin filaments in vitro (Casella et al., 1986). It is a member of the capping protein family which are α/β heterodimers that bind the barbed end and do not require Ca^{2+} for activity (Pollard & Cooper, 1986). Capping proteins have been found in all eukaryotic cells examined to date and have been purified from *Acanthamoeba* (Isenberg et al., 1980; Cooper et al., 1984), *Dictyostelium* (Schleicher et al., 1984), bovine brain (Kilimann & Isenberg, 1982), chicken skeletal muscle (Casella et al., 1986; Caldwell et al., 1989a), *Xenopus laevis* (Ankenbauer et al., 1989), and *Saccharomyces cerevisiae*.[1] Capping protein genes (two α and one β) are expressed in all of many chicken

† This work was supported by grants from the NIH (GM38542) and the Lucille P. Markey Charitable Trust. S.G.H. is supported by the American Heart Association, and J.A.C. is a Lucille P. Markey Scholar.
* Address correspondence to this author at Box 8228, 660 S. Euclid Ave., St. Louis, MO 63110.

[1] J. F. Amatruda and J. A. Cooper, unpublished data (1991).

0006-2960/91/0430-8753$02.50/0