Binding and Recognition of GATATC Target Sequences by the *Eco*RV Restriction Endonuclease: A Study Using Fluorescent Oligonucleotides and Fluorescence Polarization†

Sarah L. Reid, Damian Parry, Hsiao-Hui Liu, and Bernard A. Connolly*

Department of Biochemistry and Genetics, The University of Newcastle, Newcastle upon Tyne NE2 4HH, U.K.

Received August 16, 2000; Revised Manuscript Received December 14, 2000

ABSTRACT: Oligonucleotides labeled with hexachlorofluorescein (hex) have enabled the interaction of the restriction endonuclease *Eco*RV with DNA to be evaluated using fluorescence anisotropy. The sensitivity of hex allowed measurements at oligonucleotide concentrations as low as 1 nM, enabling K_D values in the low nanomolar range to be measured. Both direct titration, i.e., addition of increasing amounts of the endonuclease to hex-labeled oligonucleotides, and displacement titration, i.e., addition of unlabeled oligonucleotide to preformed hex-oligonucleotide/*Eco*RV endonuclease complexes, have been used for K_D determination. Displacement titration is the method of choice; artifacts due to any direct interaction of the enzyme with the dye are eliminated, and higher fluorescent-labeled oligonucleotide concentrations may be used, improving signal-to-noise ratio. Using this approach (with three different oligonucleotides) we found that the *Eco*RV restriction endonuclease showed a preference of between 1.5 and 6.5 for its GATATC target sequence at pH 7.5 and 100 mM NaCl, when the divalent cation Ca^{2+} is absent. As expected, both the presence of Ca^{2+} and a decrease in pH value stimulated the binding of specific sequences but had much less effect on nonspecific ones.

Restriction endonucleases which cut double-stranded DNA at defined sequences, typically four, six, or eight base pairs in length, are extensively used to digest DNA in recombinant DNA methodologies. Providing the appropriate buffer conditions are used, hydrolysis of DNA takes place almost exclusively at the recognition sequence of the restriction endonuclease, an observation which indicates an extremely high in vitro specificity. The in vivo role of these enzymes is to protect bacterial cells against foreign DNA invasion, with protection of the host genome being afforded by a partner DNA methyltransferase. It is thought that cutting at nontarget sites, which may not be protected by the methyltransferase, would be lethal to the host cells. As bacteria which contain restriction endonucleases/modification methylases are stable, a high in vivo specificity of the nuclease is also implied.

Although there seems little doubt that restriction endonucleases are extremely selective, both in vivo and in vitro, the exact mechanisms by which selectivity is achieved are controversial. Until comparatively recently, a division of restriction endonucleases into two families, first formally proposed by Barany and co-workers (*1*), has been suggested. One category achieves selectivity at both the binding and catalytic steps of their reaction cycle and the second solely at the catalytic steps. The well-studied enzymes R.*Eco*RI and R.*Eco*RV have provided paradigms for the two families. In the absence of Mg^{2+}, an essential cofactor for all restriction endonucleases, R.*Eco*RI clearly shows high binding selectivity for its target sequence, GAATTC (*2*−*5*), and this enzyme achieves selectivity at both binding and catalysis. In contrast, as first observed in Halford's group (*6*), R.*Eco*RV appears to bind with equal affinity to all sequences when Mg^{2+} is missing. This observation has been confirmed on numerous occasions (*7*−*11*), and it has been concluded that this enzyme generates selectivity solely at the catalytic step. In the presence of Ca^{2+}, a nonreactive analogue of Mg^{2+} that does not allow catalysis, R.*Eco*RV shows considerable preference for GATATC sites (*9*, *12*). Studies have placed R.*Bam*HI (*13*),[1] R.*Rsr*I (*14*), and R.*Sma*I (*15*) in the R.*Eco*RI family and R.*Taq*I (*16*), R.*Cfr*9I (*17*), R.*Cfr*10I (*18*), R.*Bcg*I (*19*), R.*Mun*I (*20*, *21*), and R.*Bgl*I (*22*) with R.*Eco*RV.

Recently, the nonspecific binding of R.*Eco*RV to DNA has been called into question by the observation that, under certain conditions, this protein demonstrates a marked preference for GATATC sequences (*23*). By consideration of parameters such as flanking sequences, ionic strength, and pH and the methodology used for K_D evaluation, it was possible to show pronounced binding selectivity in the absence of Mg^{2+}. In particular, pH appeared to play a crucial role; R.*Eco*RV bound to GATATC sequences about 10^4-fold more strongly than nonspecific sequences at pH 6.25.

† This work was supported by grants from the U.K. BBSRC, the European Community, and the U.K. Wellcome Trust to B.A.C. The work was carried out while B.A.C. was in receipt of a U.K. MRC Senior Research Leave Award. H.-H.L. was a Ph.D. student partially supported by a U.K. Universities ORS award. S.L.R. is an EPSRC supported Ph.D. student.

* To whom correspondence should be addressed: tel, +44-(0)191-222-7371; fax, +44-(0)191-222-7424; e-mail, b.a.connolly@ncl.ac.uk.

[1] Although R.*Bam*HI is usually quoted as belonging in the R.*Eco*RI class, recent work (Jen-Jacobson, University of Pittsburgh, personal communication) suggests that this enzyme actually shows little more binding selectivity than R.*Eco*RV.

10.1021/bi001956p CCC: $20.00
Published on Web 01/27/2001

Selectivity was much less at pH 7.5, about 10^2-fold. This study questioned the use of gel-shift analysis, most commonly used in earlier studies with R.*Eco*RV for K_D evaluation. The gels are routinely run in pH 8.3 buffers, and intermingling with the R.*Eco*RV/DNA sample (typically in a pH 7.5 buffer) might result in pH perturbations. As K_D seems very sensitive to pH, erroneous values could result. A more general point is that R.*Eco*RV binding to DNA at pH 7.5 is relatively weak and dissociation of DNA from the complex rapid. Therefore, dissociation could take place during the loading of the samples onto the gels, while they were running into the gels and during electrophoresis. Once again, this could give spurious K_D values. For these reasons it was concluded that gel-shift methods were inappropriate to the study of the binding of oligonucleotides to R.*Eco*RV at pH 7.5. A second restriction endonuclease R.*Mun*I (*20*, *21*) has also been shown to have high binding selectivity to CAATTG targets at pH 6.5 but little sequence preference at pH 8.3. Finally, the binding of R.*Eco*RI and R.*Bam*HI to their recognition sites is also weakened as the pH increases (Jen-Jacobson, personal communication). This suggests that a lowering of binding affinity for target sequences, and a consequent drop in selectivity, as pH rises might be a general feature of restriction enzymes.

To overcome problems that might arise from the use of gel-shift assays, R.*Eco*RV binding to DNA has been investigated using fluorescence-based methods, a strictly equilibrium technique where net dissociation of the enzyme−substrate complex during the measurement is not possible (*24*). Two methods were used: first, fluorescence resonance energy transfer (FRET)[2] between protein tryptophans and a dansyl-labeled oligonucleotide; second, fluorescence anisotropy using oligonucleotides labeled with either dansyl or eosin groups. In all cases identical K_D values were obtained for oligonucleotides containing a GATATC sequence and controls which lacked this site. These experiments were conducted at pH 7.5, in the absence of Mg^{2+}, and appear to confirm the lack of binding selectivity under these conditions. However, difficulties arise due to the lack of luminosity of the fluorophores used. With dansyl labeling (for both FRET and fluorescence anisotropy) 500 nM concentrations were required, limiting the accuracy of K_D determination. Eosin was more sensitive and could be used at ≈100 nM levels. As the K_D values of the oligonucleotides used fell into this range, binding constants could be determined with a reasonable degree of accuracy. A recent investigation (*12*) concluded that R.*Eco*RV shows a small (5-fold) preference for its target site, over nonspecific sequences, at pH 7.5 and in the absence of calcium. Although this investigation used gel retardation for K_D evaluation, the buffers used for the diagnostic gels (especially with regard to pH and divalent metal ion content) were selected so as to disfavor protein−DNA dissociation during the measurement.

In this study we have used oligonucleotides labeled with hexachlorofluorescein (hex) in conjunction with fluorescence anisotropy to study the binding of the R.*Eco*RV to oligonucleotides. These fluorophores are much brighter than dansyl or eosin, allowing measurements at 1 nM levels. Thus, this approach may be generally applicable to the study of protein−DNA interactions, in cases where K_D values are about 1 nM or higher. At pH 7.5 a large selectivity for GATATC sites is seen, when Ca^{2+} is present. In the absence of this metal the preference for specific sequences is much reduced, to a maximum of 6.5-fold. The results are discussed with particular reference to R.*Eco*RV and also to restriction endonucleases in general.

EXPERIMENTAL PROCEDURES

Purification of R.EcoRV. R.*Eco*RV was purified from an overproducing strain (*25*) by a two-column procedure (phosphocellulose followed by gel filtration) as previously described (*10*, *26*). A Superdex-75 (30 × 1 cm) column (Pharmacia) and FPLC were used for the gel filtration step. Following the initial phosphocellulose column chromatography the fractions containing R.*Eco*RV were concentrated either by $(NH_4)_2SO_4$ (80%) precipitation (*27*) or using Centriprep-10 spin concentrators (Amicon). No differences were seen in the properties (steady-state hydrolysis rates and K_D values for oligonucleotides) of the enzyme concentrated by either of the two methods. After the second gel filtration column chromatography concentration was achieved using Centriprep-10 spin concentrators. The protein was stored in 30% (v/v) glycerol at −20 °C. Prior to use the R.*Eco*RV was dialyzed against the buffer used for fluorescence anisotropy experiments (10 mM Hepes−NaOH, pH 7.5, 100 mM NaCl, 1 mM DTT, and 1 mM EDTA) at 4 °C. Enzyme dialyzed in this way was stored at 4 °C and used within 2−3 days as it was unstable to storage for more than 1 week.

Determination of the Concentration of R.EcoRV by Binding Site Titration. The concentration of the dialyzed enzyme was determined by an active site fluorescence anisotropy titration. The oligonucleotide hex-GTCCGGATATCACCTA (hybridized to its nonlabeled complement), at a concentration of 5 nM, was incubated in 1 mL of 10 mM Hepes−NaOH, pH 7.5, 100 mM NaCl, 1 mM DTT, 1 mM EDTA, 6 mM $CaCl_2$, and 0.1 mg of acetylated bovine serum albumin and the fluorescence anisotropy measured (see K_D Determination by Fluorescence Anisotropy section). R.*Eco*RV was added in 1 nM aliquots [concentration determined by absorbance at 280 nm using an extinction coefficient of 1.04×10^5 M^{-1} cm^{-1} (*27*) for the active dimeric form] to a final concentration of 20 nM and the fluorescence anisotropy measured after each addition. In the presence of Ca^{2+}, the K_D for R.*Eco*RV binding to specific (GATATC containing) oligonucleotides is ≪5 nM such that each protein molecule added binds to the oligonucleotide (see Results section). It was routinely found that about 7 nM (concentration determined by absorbance) R.*Eco*RV was required to bind the 5 nM oligonucleotide under stoichiometric titration conditions; i.e., the enzyme has an activity of 71% when the binding site titration is compared with absorbance.

Preparation and Purification of Oligonucleotides. Oligonucleotides were synthesized on an Applied Biosynthesis 381 DNA synthesizer using standard reagents purchased from Cruachem. Hexachlorofluorescein (hex) was added to the 5′-ends of oligonucleotides, as an integral part of the solid-phase synthesis, using a hex phosphoramidite (Glen Re-

[2] Abbreviations: DTT, dithiothreitol; EDTA, ethylenediaminetetraacetic acid; FRET, fluorescence resonance energy transfer; Hepes, *N*-(2-hydroxyethyl)piperazine-*N*′-2-ethanesulfonic acid; hex, hexachlorofluorescein; HPLC, high-pressure liquid chromatography; M.*Eco*KI, *Eco*KI DNA methyltransferase; R.*Eco*RV, restriction endonuclease *Eco*RV (etc.).

search). Normal oligonucleotides (i.e., lacking fluorophores) were deprotected with ammonia in the normal fashion and purified twice by reverse-phase HPLC (initially dmt-on and then dmt-off) (*28*−*30*). With hex oligonucleotides the ammonia deprotection step was carried out at 50 °C for 4 h (rather than overnight as is usual) as recommended by Glen Research. Fluorescently labeled oligonucleotides were purified by reverse-phase chromatography on C-18 columns (*28*−*30*) using gradients prepared from 0.1 M triethylammonium acetate, pH 6.5, containing 5% CH_3CN (buffer A) or 65% CH_3CN (buffer B). A linear gradient of 5%−50% buffer B, over 30 min at a flow rate of 1 mL/min, was used. Columns were run at 50 °C. This step was usually sufficient to purify the oligonucleotide. Occasionally, further HPLC purification was required, using an anion-exchange column (NucleoPac PA-100, 4 × 250 mm; Dionex, Camberley, U.K.). The column was run at 50 °C and developed with a gradient consisting of 0−1.5 M ammonium acetate (containing 10% v/v CH_3CN), at 1 mL/min over 45 min. All oligonucleotides were desalted using disposable NAP-25 gel filtration columns (Pharmacia).

The concentration of the oligonucleotides was determined by absorbance at 260 nm using extinction coefficients (units mM^{-1} cm^{-1}) of 14.7 (dA), 6 (dC), 11.8 (dG), 8.7 (T), and 32.2 (hex). The extinction coefficient of the oligonucleotide was obtained as the sum of the extinction coefficients of the bases plus the fluorophore (*29*, *30*). A correction was made for any hyperchromic effect by completely digesting a small sample with nuclease and determining the hyperchromicity factor (*28*). Oligonucleotides were hybridized by mixing complementary strands in the buffer used for fluorescence anisotropy measurements (10 mM Hepes−NaOH, pH 7.5, 100 mM NaCl, 1 mM EDTA, 1 mM DTT). The solution was heated to 90 °C and left to cool slowly to room temperature. For nonfluorescent oligonucleotides (i.e., used in competitive titrations) a 1:1 ratio of the two strands was used. For fluorescent oligonucleotides only one of the strands was hex-labeled. The ratio of fluorescent strand:nonfluorescent strand was 1:1.3, i.e., a slight excess of the nonfluorescent oligonucleotide.

K_D Determination by Fluorescence Anisotropy. Fluorescence anisotropy measurements were carried out at 25 °C using a SLM-Aminco 8100 fluorometer (SLM-Aminco, Urbana, IL). The excitation wavelength was 530 nm (excitation slit width 8 nm), and emission was detected through a 3 mm thick 570 nm cutoff filter (Schott OG-570; HV Skan Ltd., Solihull, U.K.). A 1 mL fluorescence cuvette, with excitation and emission path lengths each of 10 mm, was used. The initial volume was 1 mL containing 10 mM Hepes−NaOH, pH 7.5, 100 mM NaCl, 1 mM EDTA, 1 mM DTT, and 0.1 mg of acetylated bovine serum albumin (plus 6 mM $CaCl_2$ or at a pH of 7.2 when indicated). All buffers were prepared with Chelex-resin-treated water, and cuvettes were soaked in 10 mM EDTA solutions and washed with Chelex-resin-treated water prior to use. For direct titrations small volumes of R.*Eco*RV were added to the fluorescent oligonucleotide, and the anisotropy was measured after each addition. For competition titrations small volumes of oligonucleotide were added to a preformed R.*Eco*RV/fluorescent oligonucleotide complex, and the anisotropy was measured after each addition. Anisotropy was measured in the L-format using the time-trace function (total time of measurement = 315 s; integration time = 1 s; resolution = 9 s) to generate 35 data points which were averaged. A vertical excitation polarizer was used, and the emission polarizer was alternated between the vertical and horizontal positions to measure I_{VV} and I_{VH}, respectively. Anisotropy is defined by

$$\text{anisotropy} = (I_{VV} - I_{VH})/(I_{VV} + 2I_{VH})$$

where I_{VV} and I_{VH} are the intensities of the vertical and horizontal components of the emitted light using vertical polarized excitation. Differences in the response of the detector to vertical and horizontal polarized light were corrected (*G*-factor) automatically by the fluorometer. Correction was carried out "per measurement".

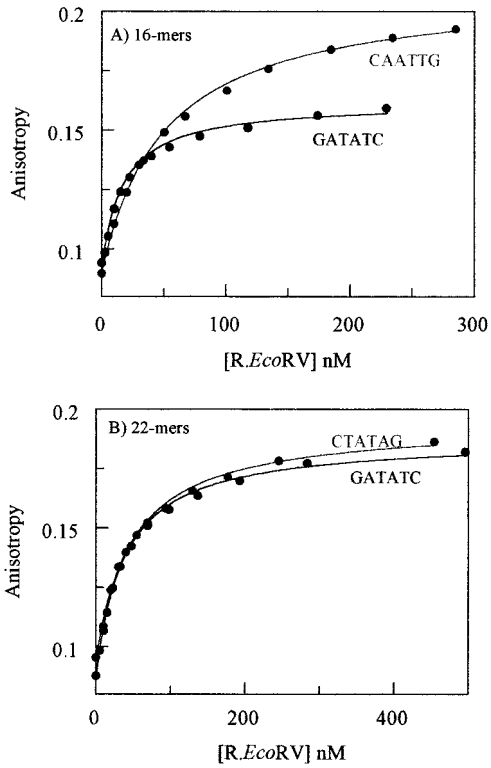

FIGURE 1: Direct titration of (A) the 16-mers hex-GTCCGATATCACCTA and hex-GTCCCAATTGACCTA and (B) the 22-mers hex-AATAGGTCGATATCGCGAATGG and hex-AATAGGTCCTATAGGCGAATGG with R.*Eco*RV (concentration determined by binding site titration) at pH 7.5 and in the absence of Ca^{2+}. In all cases the concentration of hex-labeled oligonucleotide was 5 nM. Data points were fitted to binding isotherms using GraFit (Experimental Procedures), and the K_D values found by averaging four experiments are summarized in Table 1.

Fluorescence Emission Intensity Measurements. When the effect on fluorescence emission intensity of adding R.*Eco*RV to hex-labeled oligonucleotides was under investigation, the setup was essentially as for anisotropy measurements. The time-trace function was used (total time of measurement = 305 s; integration time = 4 s; resolution = 5 s) to generate 61 data points which were averaged. "Magic angle" conditions were used to eliminate anisotropy effects on fluorescence intensity.

Data Fitting. For direct titrations (addition of R.*Eco*RV to fluorescent oligonucleotides) binding isotherms were fitted by nonlinear least-squares regression using GraFit (*31*) to the standard equation describing the equilibrium D + E ↔ DE (D = oligonucleotide, E = enzyme, DE = oligonucle-

Table 1: Direct Titration of Hex-Labeled Oligonucleotides with R.*Eco*RV[a]

oligonucleotide	anisotropy change (%)	intensity change (%)	K_D (nM)
hex-GTCCGGATATCACCTA	+65 ± 6	+10 ± 2	18 ± 1
hex-GTCCGCAATTGACCTA	+132 ± 4	+16 ± 3	47 ± 5
hex-AATAGGTCGATATCGCGAATGG	+111 ± 3	−3 ± 0.3	38 ± 2
hex-AATAGGTCCTATAGGCGAATGG	+128 ± 30	+2 ± 0.3	53 ± 12

[a] Two pairs of hex-labeled oligonucleotides 16 and 22 base pairs in length were titrated with R.*Eco*RV at pH 7.5 in the absence of Ca^{2+}. One oligonucleotide, in each pair, contains a GATATC R.*Eco*RV recognition site, and the second is a control. The change in fluorescence anisotropy and intensity on going from free to enzyme-bound oligonucleotide and the K_D values seen for the interaction between the two macromolecules are given. K_D values were determined from binding isotherms as illustrated in Figure 1, and the conditions are given in this figure. All values are the averages of four measurements, and the errors represent 1 standard deviation.

otide−enzyme complex) (*32*, *33*):

$$A = A_{min} + [(D + E + K_D) - \{(D + E + K_D)^2 - (4DE)\}^{1/2}](A_{max} - A_{min})/(2D)$$

where A = measured anisotropy at a particular *total* concentration of R.*Eco*RV (E) and fluorescent oligonucleotide (D), A_{min} = minimum anisotropy (i.e., anisotropy of free oligonucleotide prior to addition of R.*Eco*RV), A_{max} = final anisotropy (i.e., anisotropy of the oligonucleotide bound to R.*Eco*RV), and K_D = dissociation constant. This equation does not rely on any simplifying assumptions, i.e., $[E]_{free} = [E]_{total}$.

Alternatively, the data were fitted, by nonlinear least-squares regression, using Scientist (*34*), which allows the system to be described as a set of interlinked equations that define each equilibrium, the relationship between the total and free concentration of each component, and how the observable signal arises (*35*). The Scientist software requires the construction of a model file describing the system, and the following was used:

IndVars: E_{tot}

DepVars: E_f, D_f, DE, A

Params: K_D, D_{tot}, YD, YDE

$$DE = D_f \times E_f/K_D$$

$$E_{tot} = E_f + DE$$

$$D_{tot} = D_f + DE$$

$$0 < E_f < E_{tot}$$

$$0 < D_f < D_{tot}$$

$$A = D_f \times YD + DE \times YDE$$

where E and D represent R.*Eco*RV and the fluorescent oligonucleotide, respectively. DE is the enzyme−oligonucleotide complex, K_D is the dissociation constant, and A the measured anisotropy at any particular point in the titration. The subscripts "f" and "tot" denote free and total concentrations, respectively. YD is the anisotropy of the uncomplexed oligonucleotide (per mole; i.e., anisotropy/concentration), and YDE is the anisotropy of the enzyme−oligonucleotide complex (per mole; i.e., anisotropy/concentration).

For competitive titrations data were fitted using Scientist (*34*) and the following model:

IndVars: C_{tot}

DepVars: E_f, D_f, C_f, DE, CE, A

Params: K_D, K_{D1}, D_{tot}, E_{tot}, YD, YDE

$$DE = D_f \times E_f/K_D$$

$$CE = C_f \times E_f/K_{D1}$$

$$E_{tot} = E_f + DE + CE$$

$$D_{tot} = D_f + DE$$

$$C_{tot} = C_f + CE$$

$$0 < E_f < E_{tot}$$

$$0 < D_f < D_{tot}$$

$$0 < C_f < C_{tot}$$

$$A = D_f \times YD + DE \times YDE$$

The symbols have the same meaning as above, with C being the nonfluorescent competitor oligonucleotide and K_{D1} the dissociation constant for the interaction of the nonfluorescent oligonucleotide with enzyme.

RESULTS

Direct Titration of Hex-Labeled Oligonucleotides with R.EcoRV in the Absence of Ca^{2+}. Initially the binding of R.*Eco*RV to oligonucleotides was measured in 10 mM Hepes−NaOH, pH 7.5, 100 mM NaCl, 1 mM EDTA, and 1 mM DTT, in the absence of Ca^{2+}, a cation known to promote the binding of specific sequences and to increase selectivity (*9*, *12*). As mentioned in the introduction, the degree to which R.*Eco*RV prefers GATATC sequences, when divalent cations are absent, is disputed. The pH value selected, 7.5, and the NaCl concentration of 100 mM have been used by many investigators in the field. Hepes was chosen as not only does it have a pK_a (7.55) almost identical to the pH used (7.5) but it seems to eliminate effects that may arise from a direct interaction between buffer components and nucleic acids (*36*). Two sets of hexachlorofluorescein-labeled oligonucleotides were used (Table 1): (1) the 16-mers, hex-GTCCGGATATCACCTA and hex-GTCCGCAATTGACCTA; (2)

the 22-mers, hex-AATAGGTCGATATCGCGAATTG and hex-AATAGGTCCTATAGGCGAATTG (in all cases hex-labeled oligonucleotides were hybridized to fully complementary strands lacking a fluorophore). In the first set a specific *Eco*RV site (GATATC) is placed at the center of the 16-mer, which in the nonspecific control is converted to a *Mun*I (CAATTG) site, retaining only two out of the six target bases. The second set has a GATATC site in the middle of a 22-mer and an "inverted" CATATG site in the control. For the highest accuracy in anisotropy-based titrations, the quantum yield of the free and bound oligonucleotide should be the same (*37*−*39*). Table 1 shows, with the 16-mer set, that there is a small increase in fluorescence intensity (+10% for the specific oligonucleotide and +16% for the nonspecific) following R.*Eco*RV binding. With the 22-mer set intensity changes were even lower. These relatively small changes were ignored in subsequent anisotropy titrations. Figure 1 gives binding isotherms observed when the hex-labeled oligonucleotides were titrated with increasing amounts of R.*Eco*RV, and the data are summarized in Table 1. The starting anisotropy for the free oligonucleotides varied between approximately 0.09 and 0.10. The absolute change in anisotropy on binding R.*Eco*RV is between +111% and 132% for three of the four oligonucleotides but is lower (+65%) for the GATATC-containing 16-mer. The K_D values (Table 1) show a slight preference for specific over nonspecific sequences: 2.6-fold in the case of the 16-mers and 1.4-fold with the 22-mers. Carrying out titrations in the presence of 5 mM (rather than the usual 1 mM) levels of EDTA did not result in a change in K_D values (data not shown). This eliminates the presence of trace metal ion contamination, in the standard buffer, obfuscating results.

Caution is needed in the interpretation of the K_D values obtained by direct titration. Ideally, the fluorophore should be rigidly joined to the oligonucleotide (giving probe rotation perfectly correlated with that of the DNA) and also not interact directly with the protein. Hex has a negative charge, likely to be repelled by anionic nucleic acids, and is also attached to the oligonucleotide via a six-carbon spacer, properties unlikely to be compatible with rigid attachment. A recent study (*40*) has shown that this fluorophore does have a degree of independent motion, when linked to a nucleic acid, but there is sufficient correlation between the rotation of the two components to allow K_D determination. However, it is difficult to explain the difference in the anisotropy change seen when specific and nonspecific 16-mers are bound by R.*Eco*RV, although a similar effect has been observed with the *Eco*KI DNA methyltransferase (*40*). With M.*Eco*KI it was suggested that either a multimeric aggregate of the enzyme or more than one protein molecule binds to nonspecific sequences. Neither effect seems likely with R.*Eco*RV. On the basis of a number of crystallographic structures (*41*−*44*) a 16-mer, whether specific or nonspecific, is likely to accommodate only a single molecule of enzyme (as is a 22-mer). The amount of R.*Eco*RV used in the titration of the specific and nonspecific 16-mers was similar (Figure 1), making differences in the concentration-dependent aggregation state of the protein unlikely. Finally, specific and nonspecific complexes are similar in size (*44*); the anisotropy change difference cannot arise from the nonspecific complex being larger.

The simplest explanation for differences in absolute anisotropy change is a direct interaction between the protein and the dye. Under conditions where R.*Eco*RV strongly binds GATATC sequences (low pH values or the presence of Ca^{2+}), interactions are actually made to 12−14 base pairs, by contacting either flanking phosphates (directly or via H_2O) or flanking bases (via H_2O) (*41*, *42*, *44*, *45*). Therefore, even when specific binding to the GATATC sequence takes place, using 16-mers (and to a lesser extent 22-mers) would position the fluorescent probe near the protein. When R.*Eco*RV interacts with DNA nonspecifically (i.e., with controls lacking GATATC) or with little specificity (i.e., with GATATC at pH 7.5 in the absence of Ca^{2+}), 12−14 base pairs are bound mainly via protein−phosphate contacts (*44*). However, the enzyme can interact with any contiguous 12−14 bases in the 16-mer or 22-mer, necessarily giving rise to some binding modes that place the protein and fluorophore in close proximity. Therefore, with the fluorescent oligonucleotides used, protein−dye interactions cannot be rigorously excluded.

Competition Titration Using Hex-Labeled Oligonucleotides in the Absence of Ca^{2+}. To eliminate any artifacts that arise due to R.*Eco*RV interacting with the hexachlorofluorescein dye, competitive titrations have been used. Initially, complexes were formed between R.*Eco*RV and hex-labeled 16-mers (containing either GATATC or CAATTG sequences). In both cases unlabeled 16-mers (again with either GATATC or CAATTG sequences) were titrated into the mixture, and the decrease in anisotropy, as the hex-labeled oligonucleotide becomes displaced, was measured. The binding isotherms produced, using the hex-labeled nonspecific oligonucleotide as fluorescent probe (using the hex-labeled specific oligonucleotide gave similar graphs), are shown in Figure 2A and the results summarized in Table 2. Two separate estimates for the K_D values of the specific and nonspecific 16-mer competitors are obtained. With the specific 16-mer, GTCCGATATCACCTA, the two K_D values determined, 29 and 26 nM, are in good agreement and slightly higher than the 18 nM obtained by direct titration of hex-GTCCGATATCACCTA. With the nonspecific 16-mer competitor, GTCCGCAATTGACCTA, the values, 145 and 146 nM, are in excellent agreement and about 3-fold higher than was found for the analogous hex-labeled oligonucleotide by direct titration. The tighter binding observed, for both specific and nonspecific oligonucleotides, using direct as opposed to competitive titrations may be indicative of additional protein−dye interactions. Although we observe that the presence of the hex label may affect K_D values, we see little change in k_{cat} (single turnover cleavage rates) due to the dye (unpublised observations). To obtain authentic K_D values, the displacing oligonucleotide must be a true competitor of the hex-labeled oligonucleotide. The excellent fits seen in Figure 2A, which are to a model that assumes both oligonucleotides behave competitively, suggest this is the case. Furthermore, providing sufficient competitor is added (not shown), the anisotropy can be completely reduced to that of the free hex-labeled oligonucleotide, indicative of complete displacement, a requirement for competitive behavior. Therefore, we believe that the K_D values given in Table 2 represent the affinity of oligonucleotides for R.*Eco*RV, free from any dye influence. R.*Eco*RV prefers GATATC over nonspecific sequences by a factor of just over 5.

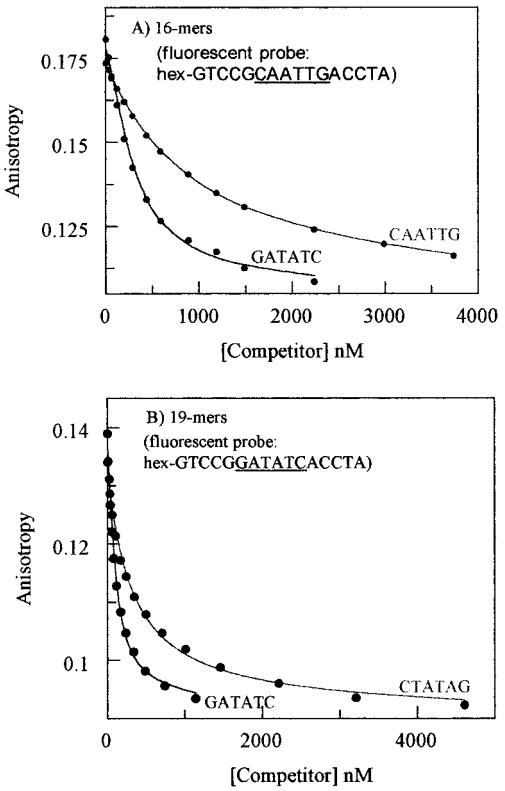

FIGURE 2: (A) Determination of the K_D for the 16-mers GTCCGATATCACCTA and GTCCCAATTGACCTA (pH 7.5 in the absence of Ca^{2+}) by competition titration using hex-GTCCCAATTGACCTA as the fluorescent probe (i.e., nonspecific fluorescent probe). The concentration of R.*Eco*RV (determined by binding site titration) was 240 nM and hex-GTCCCAATTGACCTA 150 nM (note that these concentrations were always used in competitive titrations with hex-GTCCCAATTGACCTA at pH 7.5 in the absence of Ca^{2+}). Data points were fitted to binding isotherms using Scientist (Experimental Procedures) with a K_D (dissociation constant for the R.*Eco*RV/hex-GTCCCAATTGACCTA interaction) of 47 nM. The K_D values found for the two 16-mer competitors (average from three experiments) are summarized in Table 2. (B) Determination of the K_D for the 19-mers GGGCCGATATCCGCGGGT and GGGCCCTATAGCGCGGGT (pH 7.5 in the absence of Ca^{2+}) using hex-GTCCGATATCACCTA as the fluorescent probe (i.e., specific fluorescent probe). The concentration of R.*Eco*RV (determined by binding site titration) was 75 nM and hex-GTCCCAATTGACCTA 50 nM (note that these concentrations were always used in competitive titrations with hex-GTCCGATATCACCTA at pH 7.5 in the absence of Ca^{2+}). Data points were fitted to binding isotherms using Scientist (Experimental Procedures) with a K_D (dissociation constant for the R.*Eco*RV/hex-GTCCGATATCACCTA interaction) of 18 nM. The K_D values found for the two 19-mer competitors (average from three experiments) are summarized in Table 2.

To confirm any preference for specific over nonspecific oligonucleotides, the K_D for two additional oligonucleotides has been determined by competition titration using hex-GTCCGATATCACCTA as the fluorescent oligonucleotide in each case. Initially the same 22-mers, used previously with hex labels in direct titrations, were evaluated. In this case, as summarized in Table 2, K_D values of 24 and 36 nM were obtained for the specific (GATATC) and nonspecific (CTATAG) oligonucleotides, respectively. This represents a 1.5-fold preference for specific over nonspecific sequences; however, given the errors reported in Table 2, differences as small as 1.5 approach the significance threshold of the method. The K_D values for both the specific and nonspecific 22-mers, determined by indirect titration, are both slightly lower than the comparable values found by the direct approach (Table 1). For the reasons detailed above we believe that the K_D values obtained by competition are more reliable. Finally, the K_D values for two 19-mers (GGGCCGATATCCCGCGGGT and GGGCCCTATAGCCGCGGGT) have been determined, 9 and 59 nM being found for the specific and nonspecific sequence, respectively (Figure 2B, Table 2). Therefore, with these oligonucleotides the specific is bound 6.5 times better than the nonspecific.

Titrations in the Presence of Ca^{2+}. Several investigators have shown that the binding of GATATC-containing sequences is strengthened in the presence of Ca^{2+}. Large factors are involved: 4.5×10^3 for a 382 base pair restriction fragment (*9*) and 10^4 with an oligonucleotide 16 base pairs long (*12*). Titration of 5 nM hex-GTCCGGATATCACCTA with the *Eco*RV endonuclease in the presence of 5 mM $CaCl_2$ gave the data illustrated in Figure 3A. The absolute change in anisotropy, +29%, was smaller than the value found when Ca^{2+} was absent (+65%, Table 1). Similar small increases in fluorescence intensity, +14% in the presence of Ca^{2+} and +10% in its absence (Table 1), were seen in both cases. Attempting to fit the data to a simple binding isotherm gave a K_D of 0.6 nM; however, the fit was poor with nonsystematic errors. The K_D determined for this oligonucleotide, in the absence of Ca^{2+}, was 18 nM (direct titration) or 27.5 nM (competition titration). If the presence of Ca^{2+} improved binding 10^3-10^4-fold (as expected), we would anticipate a K_D in the region 0.002−0.02 nM. Using 5 nM oligonucleotide would result in a stoichiometric titration (where every added molecule of enzyme bound to the nucleic acid) from which accurate K_D evaluation would not be possible. We believe, as shown in Figure 3A, that the data are best represented by a stoichiometric titration. As shown in this figure, 7 nM levels of the enzyme (concentration determined by UV absorbance) were required to completely bind 5 nM oligonucleotide, suggesting that only 71% of the protein is active. The protein purity is greater than 95% by SDS−PAGE, and we do not know if the reduced activity is due to about 30% denatured protein in the final preparation or to errors in the extinction coefficient used for concentration determination by absorbance spectroscopy. Throughout this publication we have used the concentration of endonuclease determined by binding site titration.

The nonspecific 16-mer, hex-GTCCGCAATTGACCTA, gave an anomalous direct titration with R.*Eco*RV in the presence of calcium.[3] Fortunately, hex-AATAGGTCCTATAGGCGAATTG gave a simple binding isotherm, when Ca^{2+} was present, which could be fitted to give a K_D of 18 nM (not shown). The absolute change in anisotropy was +130% (very similar to the +128% seen for this oligonucleotide in the absence of Ca^{2+}, Table 1), and a small decrease (−11%) in fluorescence intensity was also observed. Using a buffer that contains 6 mM $CaCl_2$ and 1 mM EDTA should give a

[3] When R.*Eco*RV was added to hex-GTCCGCAATTGACCTA in the presence of calcium, the binding isotherm describing the increase in anisotropy had a sigmoidal appearance, i.e., an initial slow increase followed by a more rapid increase and finally a tailing off. The change in fluorescence intensity, following enzyme addition, was also pronounced (about +60%) and much larger than seen with any other hex-labeled oligonucleotide. The complexities may arise because of two interaction modes, i.e., to the dye itself (resulting in the large intensity change) or to the oligonucleotide.

Table 2: Competitive Titrations Comprising the Addition of Increasing Amounts of Unlabeled Oligonucleotide to a Preformed Complex of R.*Eco*RV and a Hex-Labeled Oligonucleotide[a]

oligonucleotide (fluorescent probe)	oligonucleotide (competitor)	K_D (nM) (competitor)
hex-GTCCGGATATCACCTA	GTCCGGATATCACCTA (16-mer)	29 ± 2
hex-GTCCGCAATTGACCTA	GTCCGGATATCACCTA (16-mer)	26 ± 2
hex-GTCCGGATATCACCTA	GTCCGCAATTGACCTA (16-mer)	145 ± 19
hex-GTCCGCAATTGACCTA	GTCCGCAATTGACCTA (16-mer)	146 ± 20
hex-GTCCGGATATCACCTA	AATAGGTCGATATCGCGAATGG (22-mer)	24 ± 4
hex-GTCCGGATATCACCTA	AATAGGTCCTATAGGCGAATGG (22-mer)	36 ± 5
hex-GTCCGGATATCACCTA	GGGCCGATATCCCGCGGGT[b] (19-mer)	9 ± 3
hex-GTCCGGATATCACCTA	GGGCCCTATAGCCGCGGGT[b] (19-mer)	59 ± 7
hex-AATAGGTCCTATAGGCGAATGG	GTCCGGATATCACCTA (16-mer) (5 mM Ca^{2+})	≪1
hex-AATAGGTCCTATAGGCGAATGG	GTCCGCAATTGACCTA (16-mer) (5 mM Ca^{2+})	40 ± 5
hex-GTCCGCAATTGACCTA	GTCCGGATATCACCTA (16-mer) (pH 7.2)	13 ± 1.5
hex-GTCCGCAATTGACCTA	GTCCGCAATTGACCTA (16-mer) (pH 7.2)	151 ± 16

[a] The K_D values for 16-mers (specific, GATATC; nonspecific CAATTG), 22-mers (specific, GATATC; nonspecific CTATAG), and 19-mers (specific, GATATC; nonspecific CTATAG) (competitors) were determined by displacement of hex-labeled oligonucleotides (fluorescent probe) from R.*Eco*RV. Conditions: pH 7.5 in the absence of Ca^{2+} unless (last four rows) otherwise stated. K_D values were determined from binding isotherms, representative examples of which are illustrated in Figures 2, 3B, and 4. The R.*Eco*RV and hex-labeled oligonucleotide concentrations used, as well as the K_D that describes the interaction of the hex-labeled oligonucleotide with the enzyme, are given in the legends to these figures. The K_D values are the means of three experiments, and the errors represent 1 standard deviation. [b] The complementary strands used with the 19-mers contained a three-base dangling extension (TTT) at their 3′-ends.

free Ca^{2+} concentration of at least 5 mM. This level has been found to be saturating by other investigators (*23*), and we observed (data not shown) that increasing the $CaCl_2$ level to 10 mM too did increase the binding affinity of the oligonucleotide. The successful determination of a binding constant with hex-AATAGGTCCTATAGGCGAATTG, in the presence of Ca^{2+}, has enabled us to carry out competition titrations in the presence of this cation. As shown in Figure 3B, the K_D obtained for the nonspecific 16-mer (GTCCGCAATTGACCTA), using this approach, was 40 nM. This represents an increase in binding affinity of just under 4-fold, when compared to competition titrations using the same oligonucleotide in the absence of Ca^{2+} (Figure 2A, Table 2). The specific 16-mer (GTCCGGATATCACCTA) displaced the fluorescent oligonucleotide at much lower concentrations than the nonspecific 16-mer (Figure 3B), clearly indicating tighter binding for GATATC-containing sequences when Ca^{2+} is present. Although the data found using GTCCGGATATCACCTA could be fitted to give a K_D of about 1 nM, inspection of Figure 3B shows a linear decrease in anisotropy as each aliquot of oligonucleotide was added. Additionally, the titration finished when only a slight excess of oligonucleotide [130 nM over enzyme (100 nM)] had been added. As above, this indicates tight, near stoichiometric, binding where each molecule of competitor added binds to the *Eco*RV endonuclease. Under this condition an accurate K_D cannot be obtained, and we believe the real K_D is much lower than 1 nM.

Binding at pH 7.2. The Jen-Jacobson group (*23*) have shown that a decrease in pH (from 7.4 to 6.25) increases the binding of specific GATATC sequences but has minimal consequences to nonspecific binding. However, no experiments were carried out at pH 7.5, although K_D values were determined at pH 7.2. Therefore, to allow a better comparison of our work with that of Jen-Jacobson, we have also carried out competition titrations at pH 7.2. Preliminary direct titrations (not shown) gave a K_D of 47.5 ± 5 nM ($n = 3$) for hex-GTCCGCAATTGACCTA at pH 7.2. This value is identical to that of 47 nM seen at pH 7.5 (Figure 1, Table 1). The direct titration also revealed that the absolute changes in both fluorescence anisotropy and intensity were essentially identical to the values determined at pH 7.5 and given in Table 1. Competition titrations (Figure 4), using this hex-labeled oligonucleotide, revealed a K_D of 13 nM for the specific 16-mer GTCCGGATATCACCTA, a 2-fold improvement over the value seen at pH 7.5 (Table 2). In contrast, the K_D found for the nonspecific 16-mer, GTCCGCAATTGACCTA, at pH 7.2 was 151 nM, essentially unchanged when compared to the pH 7.5 value (Table 2). Thus we find a preference of 11.6 for specific over nonspecific sequences at pH 7.2, a clear increase over the ratio of 5.3 observed at pH 7.5.

DISCUSSION

The publication has two main purposes. The first is to explore the use of hex-labeled oligonucleotides to study protein−DNA interactions by fluorescence anisotropy. Although fluorescence anisotropy has been widely used for this purpose (*37*, *46*, *47*), there are only a few cases in which hexachlorofluorescein has been used as the fluorophore. One study with hex-labeled DNA (in which the hex was attached to the DNA via the same six-carbon spacer used in this publication) and the *Eco*KI methyltransferase (*40*) showed clear changes in anisotropy following protein binding which, coupled with the sensitivity of hex, enabled K_D values in the low nanomolar range to be evaluated. This publication also determined the fluorescence lifetime (2.87 ns) and rotational correlation time (1.7−2.5 ns) for a hex-labeled 21 base pair duplex. While the hex fluorophore was not attached to the DNA in a completely rigid manner (probably due to the six-carbon spacer), there was sufficient correlation between the motion of the probe and the nucleic acid to ensure a change in anisotropy following protein binding. In this publication we have extended the method to a second enzyme, the *Eco*RV restriction endonuclease. Direct titrations (Figure 1), at 5 nM concentrations of hex oligonucleotide, lead to straightforward measurement of K_D values that varied between 18 and 53 nM. A slight concern is a degree of direct protein−dye interaction (perhaps reflected in the variations of absolute anisotropy and intensity changes observed with

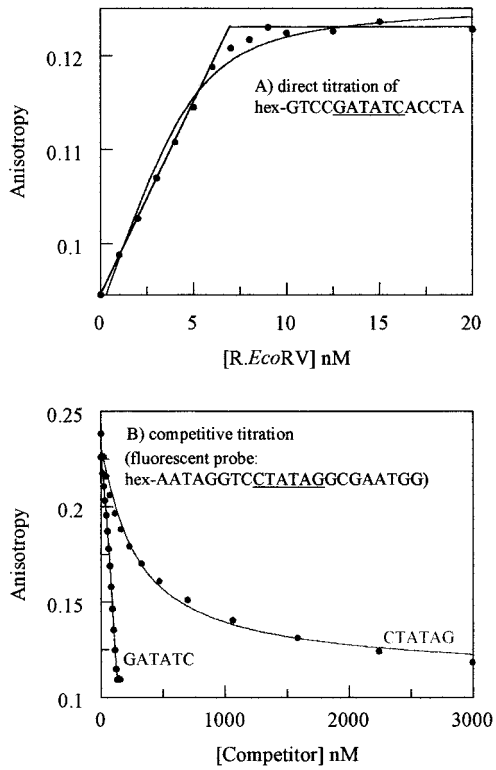

FIGURE 3: (A) Direct titration of hex-GTCCGATATCACCTA (5 nM) with R.*Eco*RV (concentration determined by UV absorbance spectroscopy) in the presence of 5 mM $CaCl_2$. The curved line represents a best fit to a simple binding isotherm using GraFit (Experimental Procedures) and gives a K_D of 0.66 nM. The two straight lines (with the linear increase defined by the first six points and the end point by the last five) represent a stoichiometric titration (i.e., $K_D \ll 5$ nM, the oligonucleotide concentration used) in which every protein molecule added binds to the fluorescent probe. As mentioned in the text, we believe the stoichiometric titration is the better description of the data. The two straight lines intersect at an R.*Eco*RV concentration of 7 nM, which given the 5 nM levels of oligonucleotide suggests that 71% of the enzyme is active. (B) Displacement of hex-AATAGGTCCTATAGGCGAATGG (60 nM) from R.*Eco*RV (100 nM, concentration determined by binding site titration) by GTCCGATATCACCTA and GTCCCAATTGACCTA in the presence of 5 mM $CaCl_2$. The curved line is a fit to a binding isotherm using Scientist (Experimental Procedures, K_D for the interaction between R.*Eco*RV and hex-AATAGGTCCTATAGGCGAATGG fixed at 18 nM). Three such titrations gave a K_D of 40 ± 5 nM for the nonspecific competitor GTCCCAATTGACCTA (Table 2). The straight line shows an essentially linear decrease in anisotropy on addition of each aliquot of GTCCGATATCACCTA. The last three data points, which occur at 130, 140, and 150 nM concentrations, have the same final anisotropy value. Although fits gave a K_D of about 1 nM, we believe the true K_D is much below this and that the binding of each molecule of competitor added to the enzyme prevents accurate evaluation of K_D values (see text).

different oligonucleotides) that may compromise K_D values, particularly for oligonucleotides which have similar affinity for the enzyme. This difficulty is most simply overcome using competitive titrations (Figures 2−4), which allow simple and direct comparisons between different oligonucleotides free from any influence of the fluorophore. Competition methods have the added benefit of using higher hex oligonucleotide concentrations, giving better signal-to-noise ratios. The advantages of hexachlorofluorescein include (1) commercial availability as a phosphoramidite enabling straightforward incorporation at the 5′-end of chemically synthesized nucleic acids and (2) high luminescence; we estimate hexachlorofluorescein is about an order of magnitude brighter than the more commonly used fluorescein (unpublished observations). Although, in this study, we have routinely used 5 nM concentrations in direct titrations, we have easily been able to carry out measurements at 1 nM levels (unpublished observations), a concentration previously used with *Eco*KI methyltransferase (*40*). Therefore, we anticipate that anisotropy methods using hex-labeled oligonucleotides, especially competitive titrations, will be widely useful for the study of protein−DNA interactions. The main disadvantage is that it is still not possible to routinely measure K_D values <1 nM, either by direct or competition methods (see below). However, protein−DNA interactions are weakened by an increase in ionic strength. Therefore, K_D values that cannot be determined at 0.1 M NaCl may be measurable at higher salt concentrations. Nevertheless, for very tight binding gel-shift analysis or filter binding with ^{32}P-labeled nucleic acids is still the method of choice.

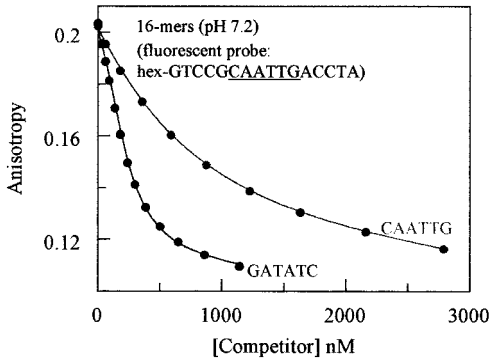

FIGURE 4: Determination of the K_D for the 16-mers GTCCGATATCACCTA and GTCCCAATTGACCTA, at pH 7.2, by competition titration using hex-GTCCCAATTGACCTA as the fluorescent probe (i.e., nonspecific fluorescent probe). The concentration of R.*Eco*RV (determined by binding site titration) was 240 nM and hex-GTCCCAATTGACCTA 150 nM. Data points were fitted to binding isotherms using Scientist (Experimental Procedures) using a K_D (dissociation constant for the R.*Eco*RV/hex-GTCCCAATTGACCTA interaction at pH 7.2) of 47.5 nM. The K_D values found for the competitors (average of three experiments) are summarized in Table 2.

The second, and main, aim of the publication was to investigate the selectivity of R.*Eco*RV for GATATC sequences, at pH 7.5 in the absence of metal ions. As mentioned in the introduction, Halford's group (*6*) have championed the idea that this enzyme shows no selectivity for its target site under these conditions. This observation, repeated by several investigators (*7*−*11*), is based on K_D evaluation using, most commonly, gel-shift analysis. This nonequilibrium method can suffer from artifacts, especially when binding is weak and, as a consequence, protein−DNA dissociation rapid. In particular, Jen-Jacobson's group (*23*) has criticized the use of diagnostic gels run at pH 8.3 (the standard methodology) to measure K_D values at pH 7.5 (see introduction) and has estimated a preference for GATATC sites of about 120-fold. Although no experiments were carried out at pH 7.5 and the figure of 120 is based on extrapolation of a graph of selectivity versus pH, the last measured point was at pH 7.4 (where selectivity was 155-fold). Therefore, the extrapolation is both reasonable and should be accurate. Since this publication Halford's group (*24*) has reevaluated the selectivity of R.*Eco*RV using strictly equilibrium fluorescence approaches and again observed no preference for GATATC sequences at pH 7.5, when divalent

metals were absent. Most recently, Perona and co-workers (*12*) measured a 5-fold preference for GATATC sites; although gel retardation was used, the pH of the diagnostic gel was selected to minimize protein−DNA dissociation during measurement.

Here we have used the equilibrium fluorescence anisotropy method to study R.*Eco*RV binding to specific and nonspecific sequences. We would like to give emphasis to the competition titrations, examples of which are shown in Figures 2, 3, and 4 and the results of which are summarized in Table 2. For three different sets of oligonucleotides we always observed a small preference for GATATC sites: 5.7-fold with the 16-mers, 1.5-fold with the 22-mers, and 6.5-fold with the 19-mers. Thus, four different groups, using a variety of methods, have observed selectivity values for R.*Eco*RV (under similar conditions: pH 7.5 and no divalent metal ion) that vary between 0 and 120. In all four cases, selectivity is defined as the ratio, K_D(specific)/K_D(nonspecific), and so the value quoted will depend on the determination of two equilibrium constants, for specific and nonspecific sequences. It has long been recognized that different target sites on the same plasmid are cut at varying rates by R.*Eco*RV, and this has been explained by flanking sequence effects, flexible flanks favoring hydrolysis (*48*, *49*). Jen-Jacobson (*23*) has reported an 18-fold variation in K_D values for GATATC-containing sequences, attributed to the influence of flanking sequences, and this parameter appears to have a large variability, strongly dependent on the bases surrounding the GATATC site. However, to produce a difference in selectivity, any variation in K_D(specific) must not be compensated by an equivalent change in K_D(nonspecific). Large differences in K_D(nonspecific) have been reported, and it is not yet clear if this reflects a true range or arises from experimental error. Nevertheless, if K_D(specific) showed a greater variability than K_D(nonspecific), as flanking sequence varied, then R.*Eco*RV would not show a unique preference for its target site but rather a range of values. Despite these complicating factors the selectivity values measured in this paper (between 1.5 and 6.5 for three oligonucleotides) are in the same region as the value of 5 seen for the 16-mer GGGAAAGATATCTTGG by the Perona group (*12*) and not too distant from the lack of specificity observed for the 12-mer GACGATATCGTC and 16-mer TCGACGATATCGTCA by Erskine and Halford (*24*) (although we do not agree with Halford that there is *absolutely* no selectivity for cognate sequences at pH 7.5 in the absence of metal ions). Clearly though, none of these values are as high as the 120-fold selectivity predicted by the Jen-Jacobson group (*23*). To allow us to make a direct comparison with Jen-Jacobson's work (*23*), we have measured the selectivity for the 16-mers at pH 7.2 (Figure 4, Table 2). These experiments were conducted by competitive titration using, in each case, the same fluorescent probe. We see little change in the K_D for the nonspecific oligonucleotide (145 nM at pH 7.5, 151 nM at pH 7.2) but a clear decrease in the K_D for the specific sequence (27.5 nM at pH 7.5, 13 nM at pH 7.2). These results agree in some respects with Jen-Jacobson's, which show little change in the K_D for nonspecific oligonucleotides as pH drops but tighter binding of specific sequences. The approximate 2-fold decrease in K_D we observe, on moving from pH 7.5 to pH 7.2 is also in line with measurements from this group. However, the selectivity of 11.6 for GTCCGGATATCACCTA over GTCCGCAATTGACCTA that we observe at pH 7.2 is about 15-fold less than the factor of 170 measured at the same pH for the specific and noncognate 22-mers, GGTAATACGGATATCCACAGAA and GGTAATACGGTTATCCACAGAA (*23*). The two sets of oligonucleotides contain the R.*Eco*RV target site in different sequence contexts, and we can only assume that the difference arises from flanking sequence effects discussed above.

Finally, we have carried out K_D measurements in the presence of Ca^{2+}, an ion anticipated to promote the binding of specific but not nonspecific sequences (*9*, *12*). Unfortunately, we were unable to obtain an accurate measure of the K_D for GATATC-containing sequences when Ca^{2+} was present. This arises because of very tight binding which, using the direct approach (Figure 3A), gives a stoichiometric titration allowing the active site concentration of the protein to be assessed. With competitive titrations (Figure 3B) all of the nonspecific hex-labeled probe is displaced when the GATATC-containing competitor concentration just exceeds that of the enzyme. Under both conditions accurate K_D evaluation is not possible. The situation could be resolved, in principle, by using lower amounts of hex-labeled oligonucleotide. Currently though, we find 1 nM concentrations the practical limit, and this is still too high for an anticipated K_D of between 0.003 and 0.03 nM for GATATC sequences when Ca^{2+} is present (10^3-10^4-fold higher than the measured value minus Ca^{2+}). However, we have measured the K_D for the nonspecific 16-mer in the presence of Ca^{2+} by both direct (not shown) and competitive titration (Figure 3B). We observe that binding of this oligonucleotide improves 3.6-fold in response to this cation. Some caution needs to be attached to this figure as different hex-labeled probes were used for plus and minus Ca^{2+} (Table 2). Nevertheless, this figure agrees with a 2.5-fold improvement in nonspecific binding, in response to Ca^{2+}, seen by the Perona group (*12*). Even though the exact discrimination between specific and nonspecific sequences, when Ca^{2+} is present, cannot be measured by our approach, Figure 3B clearly shows that the GATATC sequence is strongly preferred.

Despite disagreements concerning exact K_D values for the interaction of R.*Eco*RV with specific and nonspecific sequences and the consequent degree of binding discrimination, a coherent mechanistic understanding is emerging. At pH 7.5, when divalent metal ions are absent, the preference for specific over nonspecific sequences is "small". Even the largest binding selectivity reported, about 10^2, will contribute only minimally to the overall cutting discrimination of about 10^6-10^7 (*23*, *48*, *50*). Selectivity is greatly improved by lowering the pH or adding Ca^{2+}, and this arises as GATATC-containing sequences become much more tightly bound (nonspecific sequences are barely affected) (*9*, *12*, *23*). A very similar picture is seen with the *Mun*I restriction endonuclease, and it was proposed that relief of charge repulsion at the protein−DNA interface gives rise to these effects (*20*, *21*). The charge repulsion mechanism was also suggested to be applicable to R.*Eco*RV (*12*). For both proteins negatively charged protein carboxylate side chains approach the DNA phosphate backbone on interaction with target sequences but not noncognate sites. With the crystallographically characterized R.*Eco*RV (*41*−*44*, *51*−*53*), two of the key carboxylates are D74 and D90, which occur in a

motif, PDX_{10-30}(E/D)XK, common to many restriction endoncleases (*54*, *55*). With R.*Mun*I, crystal structures (*56*) have shown that the corresponding amino acids are D83 and E98, also present in the conserved motif. Crystallographic studies with both R.*Eco*RV and R.*Mun*I have demonstrated that the function of these conserved carboxylates is to bind the cofactor Mg^{2+}, which is essential for catalysis. Charge−charge repulsion, between the carboxylates and the DNA, leads to weak binding of specific sequences with both R.*Eco*RV and R.*Mun*I. Elimination of the negative charge on the carboxylates, either by protonation (i.e., lowering the pH) or by complexation with Ca^{2+}, attenuates the charge repulsion and so increases the binding of specific sequences. Mutagenesis of these carboxylates to uncharged amino acids for both R.*Mun*I (*20*) and R.*Eco*RV (*57*) improves the binding of specific DNA sequences, lending support to the charge repulsion model. It is tempting to speculate (*21*) that the essential cofactor Mg^{2+} acts analogously to H^+ and Ca^{2+} and also promotes the binding of specific sequences, leading to efficient hydrolysis. It is impossible to directly measure K_D values when Mg^{2+} is present, due to rapid turnover. Using inactive mutants, large enhancements in specificity were observed on Mg^{2+} addition (*58*, *59*). However, as these mutations actually involve changes to the carboxylates that coordinate the metal ion, enhancement is unlikely to arise by the same mechanism as shown by wild-type R.*Eco*RV (*12*). An alternative approach using completely hydrolysis resistant oligonucleotides (containing base, sugar, and phosphate analogues) showed only minor improvement in binding following Mg^{2+} addition (*10*, *23*). It was suggested that wholly inactive analogues may bind in an aberrant conformation that eliminates or weakens metal binding sites (*12*), as recently shown crystallographically for an oligonucleotide containing a noncleavable phosphorothiolate (*52*). Using active, but slowly cleaved oligonucleotides containing modified bases, Mg^{2+} has actually been observed to improve and promote binding (*12*), suggesting that it may also relieve charge repulsion and so behave in a manner similar to that of H^+ or Ca^{2+}.

Although the charge repulsion model may apply to all restriction endonucleases that show poor selectivity for their target sites at pH 7.5 in the absence of divalent metal ions, it is clear that a large number of restriction endonucleases (*2*−*5*) show a pronounced specificity under these conditions. The PDX_{10-30}(E/D)XK motif, containing the carboxylates responsible for charge repulsion, is present in restriction endonucleases that fall into both the R.*Eco*RV and R.*Eco*RI category. Furthermore, crystal structures of enzymes other than R.*Eco*RV and R.*Mun*I, R.*Eco*RI (*60*, *61*), R.*Bam*HI (*62*, *63*), R.*Pvu*II (*64*, *65*), R.*Fok*I (*66*), and R.*Bgl*I (*67*), have shown that the relative positions of the two carboxylates and the scissile phosphate are almost identical in every case and that these acidic amino acids clearly serve as Mg^{2+} ligands (*54*, *55*, *65*, *68*). Even when a PDX_{10-30}(E/D)XK motif is absent, e.g., for R.*Cfr*10I, structural and mutagenic studies (*69*, *70*) have revealed spatial conservation of active site carboxylates. Why, if the presence and positions of the acidic amino acids responsible for charge repulsion are conserved, do restriction endonucleases fall into two classes with different binding selectivities for their target sites at pH 7.5? The simplest explanation is that the degree of charge repulsion varies, being high for the R.*Eco*RV class of nucleases and lower for those that behave like R.*Eco*RI. The electrostatic interaction energy (E) (equivalent to the degree of charge repulsion) between two charges is given by $E = e_c e_p/Dr$ (where e_c and e_p represent the charge on the carboxylate and phosphate, respectively, D is the distance between them, and r is the dielectric constant of the medium. As the distance between the carboxylates and the scissile phosphate, revealed by X-ray crystallography, does not vary significantly between the two endonuclease categories, it is unlikely that the separation term "*D*" contributes to any variation in "*E*". However, two interrelated factors could alter charge repulsion. As suggested for R.*Mun*I (*21*), the pK_a values of the carboxylate groups, and hence the degree to which they are negatively charged at pH 7.5, may vary for different enzymes (this equates to a change in e_c). Thus R.*Eco*RI-type enzymes bind their target sequences tightly at pH 7.5 because the carboxylates are more protonated than the corresponding residues in the R.*Eco*RV/R.*Mun*I group. Although this requires anomalously high pK_a values, for both enzyme classes, this is not unprecedented, and indeed it has been suggested that one of the active site carboxylates in R.*Mun*I has a $pK_a > 7.0$ (*21*). However, the pK_a values of the groups responsible for metal ion binding in R.*Eco*RV have a $pK_a < 6$ (*71*). Alternatively, the R.*Eco*RI class may generate a medium with a higher dielectric constant than enzymes in the R.*Eco*RV group, thereby reducing electrostatic repulsion. Both the carboxylate pK_a values and the dielectric constant depend on the entire enzyme structure, i.e., relative dispositions of charged, polar, and hydrophobic regions to the key carboxylates and the DNA phosphate backbone, and this may well vary between the restriction endonuclease categories. On this interpretation it would be difficult to divide restriction endonucleases into two fundamentally distinct mechanistic groups; rather they differ in a single particular, the degree of charge repulsion at different conditions of pH and divalent metal ion. Nevertheless, even if the distinction is not a fundamental one, it remains difficult to explain why it should occur at all.

ACKNOWLEDGMENT

The authors are grateful to Pauline Heslop for skilled technical assistance. B.A.C. thanks S. E. Halford (Bristol), J. J. Perona (Santa Barbara), and L. Jen-Jacobson (Pittsburgh) for comments on the manuscript and unpublished observations and the suggestion from L. Jen-Jacobson that variations in the binding selectivity of R.*Eco*RV may arise from differential effects of flanking sequences on the binding of specific and nonspecific sequences.

REFERENCES

1. Zebala, J. A., Choi, J., Trainor, G. L., and Barany, F. (1992) *J. Biol. Chem. 267*, 8106−8116.
2. Halford, S. E., and Johnson, N. P. (1980) *Biochem. J. 191*, 593−604.
3. Terry, B. J., Jack, W. E., and Modrich, P. (1983) *J. Biol. Chem. 258*, 9820−9825.
4. Lesser, D. R., Kurpiewski, M. R., and Jen-Jacobson, L. (1990) *Science 250*, 776−786.
5. Lesser, D. R., Kurpiewski, M. R., Waters, T., Connolly, B. A., and Jen-Jacobson, L. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 7548−7552.
6. Taylor, J. D., Badcoe, I. G., Clarke, A. R., and Halford, S. E. (1991) *Biochemistry 30*, 8743−8753.

7. Thielking, V., Selent, U., Köhler, E., Landgraf, A., Wolfes, H., Alves, J., and Pingoud, A. (1992) *Biochemistry 31*, 3727−3732.
8. Vermote, C. L. M., Vipond, I. B., and Halford, S. E. (1992) *Biochemistry 31*, 6089−6097.
9. Vipond, I. B., and Halford, S. E. (1995) *Biochemistry 34*, 1113−1119.
10. Szczelkun, M. D., and Connolly, B. A. (1995) *Biochemistry 34*, 10724−10733.
11. Alves, J., Selent, U., and Wolfes, H. (1995) *Biochemistry 34*, 11191−11197.
12. Martin, A. M., Horton, N. C., Lusetti, S., Reich, N. O., and Perona, J. J. (1999) *Biochemistry 38*, 8430−8439.
13. Xu, S., and Schildkraut, I. (1991) *J. Biol. Chem. 266*, 4425−4429.
14. Aiken, C. R., Fisher, E. W., and Gumport, R. I. (1991) *J. Biol. Chem. 266*, 19063−19069.
15. Withers, B. E., and Dunbar, J. C. (1995) *J. Biol. Chem. 270*, 6496−6504.
16. Zebala, J. A., Choi, J., and Barany, F. (1992) *J. Biol. Chem. 267*, 8097−8105.
17. Siksnys, V., and Pleckaityte, M. (1993) *Eur. J. Biochem. 217*, 411−419.
18. Skirgaila, R., and Siksnys, V. (1988) *Biol. Chem. 379*, 595−598.
19. Kong, H., Roemer, S. E., Waite-Rees, P. A., Benner, J. S., Wilson, G. G., and Nwanko, D. O. (1994) *J. Biol. Chem. 269*, 683−690.
20. Lagunavicius, A., and Siksnys, V. (1997) *Biochemistry 36*, 11086−11092.
21. Lagunavicius, A., Grazulis, S., Balciunaite, E., Vainius, D., and Siksnys, V. (1997) *Biochemistry 36*, 11093−11099.
22. Gormley, N. A., Bath, A. J., and Halford, S. E. (2000) *J. Biol. Chem. 275*, 6928−6936.
23. Engler, L. E., Welch, K. K., and Jen-Jacobson, L. (1997) *J. Mol. Biol. 269*, 82−101.
24. Erskine, S. G., and Halford, S. E. (1998) *J. Mol. Biol. 275*, 759−772.
25. Bougueleret, L., Techini, M. L., Botterman, J., and Zabeau, M (1985) *Nucleic Acids Res. 13*, 3823−2839.
26. Grasby, J. A., and Connolly, B. A. (1992) *Biochemistry 31*, 7855−7861.
27. D'Arcy, A., Brown, R. S., Zabeau, M., Van Resandt, R. W., and Winkler, F. K. (1985) *J. Biol. Chem. 260*, 1987−1990.
28. Newman, P. C., Nwosu, V. U., Williams, D. M., Cosstick, R., Seela, F., and Connolly, B. A. (1990) *Biochemistry 29*, 9902−9910.
29. Connolly, B. A. (1991) in *Oligonucleotide Analogues—a Practical Approach* (Eckstein, F., Ed.) pp 155−183, IRL press, Oxford, U.K.
30. Connolly, B. A. (1992) *Methods Enzymol. 211*, 36−53.
31. GraFit (version 3.09a) (1992) Erithacus Software, Staines, U.K.
32. Müller, B., Restle, T., Reinstein, J., and Goody, R. S. (1991) *Biochemistry 30*, 3709−3715.
33. Maury, F., Divita, G., Morvan, F., Imbach, J.-L., and Goody, R. S. (1993) *Biochim. Biophys. Acta 1216*, 1−8.
34. Scientist (version 2) (1995) MicroMath Scientific Software, Salt Lake City, UT.
35. Krebs, R., Immendörfer, U., Thrall, S., Wöhrl, B. M., and Goody, R. S. (1997) *Biochemistry 36*, 10292−10300.
36. Wenner, J. R., and Bloomfield, V. A. (1999) *Anal. Biochem. 268*, 201−212.
37. Jameson, D. M., and Sawyer, W. H. (1995) *Methods Enzymol. 246*, 283−300.
38. Gutfreund, H. (1995) in *Kinetics for the Life Sciences; Receptors, Transmitters and Catalysts*, pp 201−295, Cambridge University Press, U.K.
39. Eftink, M. R. (1997) *Methods Enzymol. 278*, 221−257.
40. Powell, L. M., Connolly, B. A., and Dryden, D. T. F. (1998) *J. Mol. Biol. 283*, 947−961.
41. Winkler, F. K., Banner, D. W., Oefner, C., Tsernoglu, D., Brown, R. S., Heathman, S. P., Bryan, R. K., Martin, P. D., Petratos, K., and Wilson, K. S. (1993) *EMBO J. 12*, 1781−1795.
42. Kostrewa, D., and Winkler, F. K. (1995) *Biochemistry 34*, 683−696.
43. Perona, J. J., and Martin, A. M. (1997) *J. Mol. Biol. 273*, 207−225.
44. Horton, N. C., and Perona, J. J. (1998) *J. Biol. Chem. 273*, 21721−21729.
45. Thorogood, H., Grasby, J. A., and Connolly, B. A. (1996) *J. Biol. Chem. 271*, 8855−8862.
46. Heyduk, T., and Lee, J. C. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 1744−1748.
47. Hill, J. J., and Royer, C. A. (1997) *Methods Enzymol. 278*, 390−416.
48. Taylor, J. D., and Halford, S. E. (1992) *Biochemistry 31*, 90−97.
49. Vermote, C. L. M., and Halford, S. E. (1992) *Biochemistry 31*, 6082−6089.
50. Taylor, J. D., and Halford, S. E. (1989) *Biochemistry 28*, 6198−6207.
51. Martin, A. M., Sam, M. D., Reich, N. O., and Perona, J. J. (1999) *Nat. Struct. Biol. 6*, 269−277.
52. Horton, N. C., Connolly, B. A., and Perona, J. J. (2000) *J. Am. Chem. Soc. 122*, 3314−3324.
53. Horton, N. C., Newberry, K. J., and Perona, J. J. (1998) *Proc. Natl. Acad. Sci U.S.A. 95*, 13489−13494.
54. Aggarwal, A. K. (1995) *Curr. Opin. Struct. Biol. 5*, 11−19.
55. Pingoud, A., and Jeltsch, A. (1997) *Eur. J. Biochem. 246*, 1−22.
56. Deibert, M., Grazulis, S., Janulaitis, A., Siksnys, V., and Huber, R. (1999) *EMBO J. 18*, 5805−5816.
57. Selent, U., Reuter, T., Koehler, E., Liedtke, M., Thielking, V., Alves, J., Oelgeschlaeger, T., Wolfes, H., Peters, F., and Pingoud, A. (1992) *Biochemistry 31*, 4808−4815.
58. Thielking, V., Selent, U., Köhler, E., Landgraf, A., Wolfes, H., Alves, J., and Pingoud, A. (1992) *Biochemistry 31*, 3727−3732.
59. Jeltsch, A., Maschke, H., Selent, U., Wenz, C., Köhler, E., Connolly, B. A., Thorogood, H., and Pingoud, A. (1995) *Biochemistry 34*, 6239−6246.
60. Rosenberg, J. M. (1991) *Curr. Opin. Struct. Biol. 1*, 104−113.
61. Kim, Y., Grable, J. C., Choi, P. J., Greene, P., and Rosenberg, J. M. (1990) *Science 249*, 1307−1309.
62. Newman, M., Strzelecka, T., Dorner, L. F., Schildkraut, I., and Aggarwal, A (1994) *Nature 368*, 660−664.
63. Newman, M., Strzelecka, T., Dorner, L. F., Schildkraut, I., and Aggarwal, A. (1995) *Science 269*, 656−663.
64. Athanasiadis, A., Vlassi, M., Kotsiafki, D., Tucker, P. A., Wilson, K. S., and Kokkinidis, M. (1994) *Nat. Struct. Biol. 1*, 469−475.
65. Cheng, X., Balendiran, K., Schildkraut, I., and Anderson, J. E. (1994) *EMBO J. 13*, 3927−3935.
66. Wah, D. A., Hirsch, J. A., Dorner, L. F., Schildkraut, I., and Aggarwal, A. K. (1997) *Nature 388*, 97−100.
67. Newman, M., Lunnen, F., Wilson, G., Greci, J., Schildkraut, I., and Phillips, S. E. V. (1998) *EMBO J. 17*, 5466−5476.
68. Venclovas, C., Timinskas, A., and Siksnys, V. (1994) *Proteins: Struct., Funct., Genet. 20*, 279−282.
69. Bozic, D., Grazulis, S., Siksnys, V., Huber, R. (1996) *J. Mol. Biol. 255*, 176−186.
70. Skirgaila, R., Grazulis, S., Bozic, D., Huber, R., and Siksnys, V. (1998) *J. Mol. Biol. 279*, 473−481.
71. Stanford, N. P., Halford, S. E., and Baldwin, G. S. (1999) *J. Mol. Biol. 288*, 105−116.

BI001956P