Interaction of the *Eco*RV Restriction Endonuclease with the Deoxyadenosine and Thymidine Bases in Its Recognition Hexamer d(GATATC)†

Patrick C. Newman,‡ David M. Williams,§,∥ Richard Cosstick,§ Frank Seela,⊥ and Bernard A. Connolly*,‡

Department of Biochemistry (SERC Molecular Recognition Centre), University of Southampton, Southampton SO9 3TU, U.K., Department of Chemistry, University of Liverpool, Liverpool L69 3BX, U.K., and Laboratorium fur Organische and Bioorganische Chemie, Universitat Osnabruck, D-4500 Osnabruck, West Germany

Received April 11, 1990; Revised Manuscript Received July 20, 1990

ABSTRACT: A set of dA and T analogues suitable for the study of protein DNA interactions have been incorporated into the central d(ATAT) sequence within d(GACGATATCGTC). The individual analogues have one potential protein contact (either a hydrogen-bonding group or a CH_3 group capable of a van der Waals interaction) deleted. In general, the modified bases do not perturb the overall structure of the dodecamer, enabling results obtained to be simply interpreted in terms of loss of protein DNA contacts. We have used the modified oligodeoxynucleotide set to study the recognition of DNA by the *Eco*RV restriction endonuclease [recognition sequence d(GATATC)]. The k_{cat} and K_m values for the set have been determined, and a comparison with results seen with the parent oligodeoxynucleotide (containing no modified bases) has been carried out. Three classes of results are seen. First, some analogues lead to no change in kinetic parameters, meaning no enzyme contact at the altered site. Second (this is seen for most of the modified oligodeoxynucleotides), a drop in the k_{cat}/K_m ratio relative to the parent is observed. This comes mainly from a decrease in k_{cat}, implying that the endonuclease uses the interaction under study to lower the transition-state barrier rather than to bind the substrate. Analyses of these results show that the drop in k_{cat}/K_m is what would be expected for the simple loss of a hydrogen bond or a CH_3 contact between the enzyme and the oligodeoxynucleotide. This implies a contact of these types at these sites. Third, some analogue-containing oligodeoxynucleotides are not substrates; i.e., the k_{cat}/K_m drop is much greater than would be expected for loss of a single hydrogen bond or CH_3 contact. These results are interpreted in terms of a cooperative mechanism whereby the loss of one interaction causes a rearrangement at the enzyme active site leading to a consequent loss of further protein substrate contacts. However, in these cases gross structural changes in the oligodeoxynucleotide conformation cannot be excluded. It is found that the endonuclease makes very many interactions to the d(ATAT) sequence within its d(GATATC) recognition site, and these occur in both the major and minor grooves. The results obtained have been used to explain how the enzyme achieves the high degree of cleavage specificity for d(GATATC) as compared to all other sequences.

All biological processes depend on reversible and weak noncovalent interactions such as hydrogen bonds, salt bridges, the hydrophobic effect, and van der Waals dispersive forces. These forces are especially important in the binding of substrates by enzymes where they act to provide specificity and catalysis. X-ray crystallography is the best technique for determining the presence of noncovalent interactions between enzymes and substrates. Unfortunately, this method gives little indication of the strength of a particular interaction and the role it plays during each step of the catalytic cycle. To get an indication of the strengths of these interactions, it is necessary to perform binding or kinetic studies. The most popular approach uses substrate analogues in which the functional group responsible for the noncovalent interaction under study has been removed, usually by replacement with a hydrogen atom. Comparison of the binding and kinetic properties of the analogue with the parent gives an indication of the contribution that the group makes to binding and catalysis. More recently, site-directed mutagenesis has made it possible to selectively replace individual amino acids in proteins, leading to the production of "enzyme analogues" in which the roles of protein amino acids during catalysis can be probed. A kinetic analysis of enzymes altered by mutagenesis has been developed by Fersht (1985, 1987a,b, 1988). Although used in these cases for altered enzymes, it is also applicable to substrate analogues.

When enzymes are being considered (i.e., when catalysis rather than just binding is taking place), the contribution that a particular group on a substrate makes to the binding energy can be estimated from the equation

$$\Delta G_{app} = RT \ln [(k_{cat}/K_m)\text{analogue}/(k_{cat}/K_m)\text{parent}] \quad (1)$$

k_{cat}/K_m is the specifity constant in the Michaelis–Menten equation for either the parent (i.e., the unmodified natural substrate) or the analogue in which the group under study has been deleted. Usually the deleted group is actually replaced by a hydrogen atom (e.g., replacement of a hydrogen-bonding NH_2 group with H or conversion of a CH_3 group capable of van der Waals interactions to H). As the k_{cat}/K_m ratio de-

† B.A.C. was supported by grants from the U.K. SERC and particularly by the setting up of a molecular recognition initiative centre by this body. Additional funds were contributed by the EEC (Science Plan). B.A.C. is a Lister Institute research fellow. R.C. was supported by the U.K. SERC. P.C.N. and D.M.W. were SERC-supported research students.

* Address correspondence to this author.
‡ University of Southampton.
§ University of Liverpool.
∥ Present address: Max-Planck-Institut fur experimentelle Medizin, D-3400 Gottingen, West Germany.
⊥ Universitat Osnabruck.

0006-2960/90/0429-9902$02.50/0

termines the specificity that an enzyme shows for a particular substrate, the ΔG_{app} obtained from eq 1 is strictly a measure of the discrimination that the enzyme shows between the two substrates. In general, however, the measured value of ΔG_{app} does not equal ΔG_{bind}, which is the actual binding energy that the group under study contributes to the enzyme substrate complex. This arises because ΔG_{bind} relates the interaction of the group under study with the enzyme as compared to the interaction of the free enzyme and the free group with solvent. ΔG_{app} compares the interaction of the parent and the analogue in which the group has been deleted with the enzyme. In addition, problems can arise from injudicious choice of analogue. An ideal analogue should have the group under study replaced with no effects to the rest of the molecule (a conservative or nondisruptive change). In particular, no extra bulk should be introduced and the group should simply be removed and not replaced by another type (e.g., a potential hydrogen-bond donor should not be replaced by an acceptor). Even when the group under study is replaced by the small hydrogen atom, care is needed to ensure that other parts of the molecule are not perturbed. Despite these difficulties, the use of analogues to probe the roles of individual groups and evaluation of ΔG_{app} values using eq 1 is a useful biochemical method capable of giving information on how important a particular group on a substrate is in enzyme catalysis.

In the preceding paper we prepared a set of dA and T analogues suitable for studying protein nucleic acid interactions. Most of these analogues had potential hydrogen-bonding functions or CH_3 groups replaced by hydrogen. These were incorporated into the central d(GATATC) site of the self-complementary dodecamer d(GACGATATCGTC). d(GATATC) is the recognition site for the *Eco*RV restriction endonuclease and modification methylase (Kholmina et al., 1980; Schildkraut et al., 1984; D'Arcy et al., 1985; Nwosu et al., 1988), and the oligodeoxynucleotide set prepared was used to probe the recognition of dA and T bases by these enzymes. These initial studies involved determining rates at the single oligodeoxynucleotide concentration of 20 μM for the preliminary identification of important enzyme substrate contacts. This paper extends the studies with the endonuclease in that K_m and k_{cat} values have been determined and used to evaluate ΔG_{app} values by using eq 1. We have compared the ΔG_{app} values obtained with those seen for other proteins. Furthermore, the results have led to ideas as to how the endonuclease recognizes its d(GATATC) cognate site and discriminates against noncognate sites. It is known that cognate sites are cut at least 10^6 times faster than the best noncognate sequences (Taylor & Halford, 1989).

Materials and Methods

Most of the materials and methods such as oligodeoxynucleotide and enzyme preparation, oligodeoxynucleotide purification and characterization, HPLC systems and the sources of many of the chemicals are given in the preceding paper. [γ-^{32}P]ATP (3000 Ci mmol^{-1}) was purchased from Amersham International, Aylesbury U.K. Polynucleotide kinase (from phage T4 infected *Escherichia coli*) was the product of Boehringer-Mannheim, Lewes, East Sussex, U.K. All the reagents used for gel electrophoresis were of Electran grade and were purchased from BDH Chemicals, Poole, Dorset, U.K.

Gel Electrophoresis. Polyacrylamide gels were prepared in volumes of 60 mL consisting of Tris–borate buffer, pH 8 (0.9% w/v Tris base, 0.46% w/v boric acid), 2 mM EDTA, 7 M urea, and 16.7% polyacrylamide [prepared from a deionized 30% polyacrylamide stock solution containing 29% acrylamide and 1% bis(acrylamide)]. Polymerization was initiated by the addition of 0.3 mL of a freshly prepared 10% (w/v) ammonium persulfate solution and 0.04 mL of *N,N,N',N'*-tetramethylethylenediamine. This solution was poured into a standard gel electrophoresis apparatus giving a gel of dimensions 20 cm × 20 cm × 1 mm. A buffer consisting of Tris–borate, pH 8 (1.08% w/v Tris base and 0.55% w/v boric acid), containing 2 mM EDTA was used to run these gels. Volumes of 5–10 μL of the quenched samples of the *Eco*RV endonuclease reaction mixtures were loaded and the gels run at 30-mA constant current until the bromophenol blue dye marker had migrated about two-thirds of the distance of the gel. The positions of these dodecamers and hexamers were determined by autoradiography. The radioactive gel slices corresponding to the 12-mers and 6-mers were excised from the gel and the counts in them determined by liquid scintillation counting.

Oligodeoxynucleotide Phosphorylation. Oligodeoxynucleotides were end labeled by 5'-phosphorylation using modifications of well-established procedures (Maniatis et al., 1982). In all cases two phosphorylations were carried out, one with nonradioactive ATP and the other with [γ-^{32}P]ATP. Nonradioactive phosphorylations were performed in volumes of 100 μL containing 50 mM Tris-HCl, pH 9.5, 10 mM $MgCl_2$, 5 mM dithiothreitol, 5% (v/v) glycerol, 20 mM ATP, 0.1 mM (10 nmol total) double-stranded dodecamer, and 10 units (as defined by the manufacturer) of polynucleotide kinase. These solutions were incubated at 30 °C for 48 h. Radioactive phosphorylations were performed in two stages. Initially, 100 μL of the above buffer again containing 0.2 mM (20 nmol total) double-stranded dodecamer but with 0.05 mM (5 nmol total) nonradioactive ATP and 100 μCi (about 0.03 nmol total) [γ-^{32}P]ATP was incubated with 10 units of polynucleotide kinase at 30 °C for 48 h. After this time, 1 μmol of nonradioactive ATP (giving an ATP concentration of 10 mM) and a further 10 units of kinase were added, and the mixture was incubated for 24 h at 30 °C. Phosphorylated oligodeoxynucleotides were purified by reverse-phase HPLC chromatography using octadecylsilyl (C18) columns (see previous paper for more details). These columns were developed with a linear gradient formed from 0.1 M triethylammonium bicarbonate, pH 8, containing 5% acetonitrile (buffer A) and 0.1 M triethylammonium bicarbonate, pH 8, containing 65% acetonitrile (buffer B). A gradient of 0–25% buffer B over 20 min at a flow rate of 1 mL min^{-1} and at room temperature was used. Excess ATP eluted at about 3 min and the required phosphorylated oligodeoxynucleotide at about 10 min. Traces (usually less than 2%) of the nonphosphorylated starting materials eluted at about 12 min. The peak corresponding to the desired phosphorylated oligodeoxynucleotide was collected and solvents removed by evaporation in a Savant SpeedVac. Excess triethylammonium bicarbonate was removed by coevaporation using 3 × 100 μL of methanol. The products were dissolved in 100 μL of H_2O and stored frozen at −20 °C. They were used as soon as possible and within 2 weeks as substrates for the *Eco*RV endonuclease.

Characterization of Phosphorylated Oligodeoxynucleotides. The purity of all the phosphorylated oligodeoxynucleotides was checked by reverse-phase HPLC using the systems given in the preceding paper for purity determination of the unphosphorylated parents. In addition, small samples of the ^{32}P-labeled oligodeoxynucleotides had their purity evaluated by polyacrylamide gel electrophoresis followed by autoradiography as detailed above. The concentrations of the stock solutions of the phosphorylated oligodeoxynucleotides were

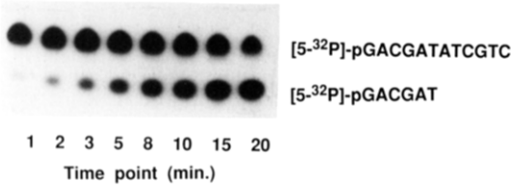

FIGURE 1: Autoradiogram of the time course of hydrolysis of [^{32}P]dp(GACGATATCGTC) to [^{32}P]dp(GACGAT) and dp(ATCGTC) by the *Eco*RV endonuclease.

determined by using absorbance at 254 nm. We have assumed that the extinction coefficients of the phosphorylated 12-mers are the same as those of their unphosphorylated dodecamer parents (see previous paper for these extinction coefficients). These concentration determinations showed that the yields of the phosphorylation reactions were about 95%. The specific activities of the products were about 5 μCi nmol^{-1}.

EcoRV Endonuclease Assays. The effect of endonuclease concentration on hydrolysis rates was determined by using [γ-^{32}P]dp(GACGATATCGTC). A volume of 50 μL containing 50 mM Hepes–KOH, pH 7.5, 100 mM NaCl, 10 mM $MgCl_2$, and 20 μM oligodeoxynucleotide (specific activity about 2 μCi nmol^{-1}) was used. *Eco*RV endonuclease was added to give final concentrations of 0.05, 0.1, 0.5, 0.8, and 1.2 μM. At appropriate times 5-μL aliquots were withdrawn and the reactions quenched by adding 10 μL of a formamide/water (9:1 v/v) mixture containing 0.02% bromophenol blue and heating at 100 °C for 5 min. The number of counts in the substrate dodecamer and product hexamer were determined by liquid scintillation counting after electrophoresis and autoradiography as described above. These figures were used to produce initial rates. The variation in enzyme-catalyzed rates as a function of Mg^{2+} concentration was investigated by using the parent phosphorylated dodecamer and the two phosphorylated dodecamers containing dU instead of T within the recognition hexamer. The above buffer system was used but the Mg^{2+} concentration was varied, and the NaCl concentration was also altered to maintain an ionic strength equal to the 10 mM $MgCl_2$ plus 100 mM NaCl buffer used previously. Five buffers were used that contained 2 mM $MgCl_2$/124 mM NaCl, 5 mM $MgCl_2$/115 mM NaCl, 10 mM $MgCl_2$/100 mM NaCl, 20 mM $MgCl_2$/70 mM NaCl, and 40 mM $MgCl_2$/10 mM NaCl. Once again, 20 μM concentrations of the three oligodeoxynucleotides (specific activity 1 μCi nmol^{-1}) were used. The concentration of *Eco*RV dimer was 0.1 μM for the parent phosphorylated dodecamer, 2 μM for dp(GACGAUATCGTC), and 5 μM for dp(GACGATAUCGTC). Aliquots were withdrawn at various times and rates determined as before. The dependence of *Eco*RV-catalyzed hydrolysis as a function of phosphorylated oligodeoxynucleotide concentration was investigated to allow K_m and k_{cat} determination. Assays were performed in 100-μL volumes of 50 mM Hepes–KOH, pH 7.5, containing 55 mM NaCl and 25 mM $MgCl_2$. The oligodeoxynucleotide concentrations varied between 0.25 and 10 μM (double stranded) as all K_m values were within this range. The specific activities of the substrate 12-mers used varied and were produced by adjusting the relative amounts of nonradioactive and radioactive 12-mers to give the final desired concentration. The lower substrate concentrations had more radioactive dodecamer relative to nonradioactive and so a greater specific activity as compared to the higher substrate concentrations. With this approach the decrease in counts due to the lower substrate concentrations is compensated by the increase in specific activity, and this makes counting and so rate determination more accurate. A constant amount of *Eco*RV endonuclease was used for each of the phosphorylated dodecamers (although this amount varied among different dodecamers). This amount was selected to be the minimum that gave measurable rates over reasonable time courses. Time courses varied from between 0 and 10 min for good substrates to between 0 and 360 min for poor substrates. The rates were determined as above. K_m and k_{cat} values were determined from graphs of substrate concentration divided by velocity against velocity.

Results and Discussion

In the preceding paper we measured the rates of *Eco*RV endonuclease catalyzed hydrolysis of oligodeoxynucleotides using an HPLC-based procedure. However, assays based on labeling of dodecamers with ^{32}P at their 5′-hydroxyl position and separation of the radioactive substrates and products by gel electrophoresis are faster, more sensitive, and more accurate than those based on HPLC and so are more suited to K_m and k_{cat} determination. As an example Figure 1 shows the time course for the hydrolysis of [^{32}P]dp(GACGATATCGTC) (for the exact conditions see figure legend). Assays of this kind have been previously used to study endonuclease-catalyzed hydrolysis of oligodeoxynucleotides (Dwyer-Hallquist et al., 1982; Ono et al., 1984; Kita et al., 1985; Yolov et al., 1985; Jiricny et al., 1986; Brennan et al., 1986; McLaughlin et al., 1987; Ono & Ueda, 1987; Hayakawa et al., 1988; Mazzarelli et al., 1989). We have used standard enzymological methods based on polynucleotide kinase and ATP to prepare phosphorylated 12-mers (Maniatis et al., 1982). No difficulties were experienced in the preparation of nonradioactive phosphorylated 12-mers as the use of large quantities of kinase, a large excess of ATP, and long incubation time gave essentially 100% yields of the desired products. In preparing ^{32}P-labeled derivatives, some care must be taken to ensure that the phosphorylation proceeds in high yields (approaching 100%) and that the product is of the requisite high specific activity. The two-stage approach given under Materials and Methods where 20 nmol of oligodeoxynucleotide is initially incubated with 5 nmol of ATP containing 100 μCi of ^{32}P and then subsequently incubated with a large excess of unlabeled ATP was optimal in giving both 100% phosphorylation yields and a high specific activity. Increasing the 5 nmol of ATP in the initial incubation (while keeping the 100 μCi of ^{32}P present constant) reduced the specific activity of the final product. However, decreasing this 5 nmol at a constant 100 μCi of ^{32}P also decreased the specific activity of the product. This is probably because the low levels of ATP present in the initial incubation lead to low levels of enzyme-catalyzed phosphorylation and ^{32}P incorporation at this stage. After purification of the phosphorylated oligodeoxynucleotides by reverse-phase HPLC, we obtained products that were extremely pure both by HPLC (>98% of the UV-absorbing material and >98% of the ^{32}P eluted in a single sharp peak for all the phosphorylated oligodeoxynucleotides, data not shown) and by gel electrophoresis (>98% of the ^{32}P-labeled material ran in a single band of the gel corresponding to a dodecamer, data not shown).

Table I shows that *Eco*RV endonuclease concentrations of between 0.05 and 0.5 μM were required for K_m and k_{cat} determination. Using [γ-^{32}P]dp(GACGATATCGTC), we have demonstrated that the cleavage rate increases in a linear fashion for *Eco*RV concentrations betwen 0.05 and 1.2 μM (data not shown). This linear increase avoids any complications due to concentration-dependent changes in specific activity of the enzyme. In a recently published paper, Halford

Table I: k_{cat}, K_m, and k_{cat}/K_m Values Found for the Parent Oligodeoxynucleotide and Those Containing Modified Bases[a]

	[*Eco*RV] (μM)	k_{cat} (min^{-1})	K_m (μM)	k_{cat}/K_m [S^{-1} M^{-1} (%)]	ΔG_{app} (kJ mol^{-1})
d(GACGATATCGTC)	0.05	6.9	3.8	30.2×10^3 (100)	–
GPTATC	–	too slow to measure			–
G[^{7C}A]TATC	–	0.0	–	– (0)	–
G[^{3C}A]TATC	0.2	0.057	4.0	238 (0.8)	−6.0
GAUATC	0.5	0.013	3.8	56 (0.2)	−7.8
GA[^{4S}T]ATC	0.1	0.016	0.5	533 (1.8)	−5.0
GA[^{4H}T]ATC	–	0.0	–	– (0)	–
GA[^{2S}T]ATC	0.1	1.87	1.0	32.9×10^3 (108)	+0.1
GATPTC	0.25	0.067	8.0	139 (0.5)	−6.7
GA[^{7C}A]TC	0.05	1.9	0.8	39.6×10^3 (131)	+0.3
GAT[^{3C}A]TC	0.05	0.1	1.4	1.2×10^3 (4)	−4.8
GATAUC	0.5	0.006	6.5	15.4 (0.05)	−9.4
GATA[^{4S}T]C	–	0.0	–	– (0)	–
GATA[^{4H}T]C	–	0.0	–	– (0)	–
GATA[^{2S}T]C	0.1	0.1	4.0	417 (1.4)	−5.3

[a] The [*Eco*RV] used in these determinations is also given. The ΔG_{app} values have been calculated from eq 1. Initial values were divided by 2 to give a figure for loss of one interaction (see text).

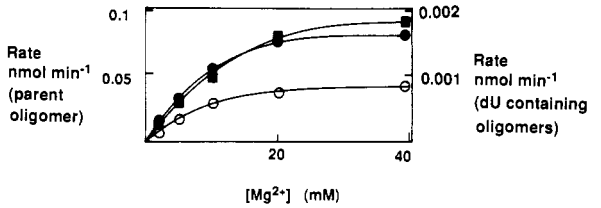

FIGURE 2: Effect of [Mg^{2+}] on the hydrolysis of dp(GACGATATCGTC) (●), dp(GACGAUATCGTC) (■), and dp(GACGATAUCGTC) (○) by the *Eco*RV endonuclease. The concentration of enzyme used was 0.1 μM for the parent oligodeoxynucleotide, 2 μM for dp(GACGAUATCGTC), and 5 μM for dp(GACGATAUCGTC). It should be noted that in this graph the rates have not been normalized to units of [enzyme] to allow plotting of the three lines on a single graph.

and his co-workers demonstrated that part of the discrimination between cognate and noncognate sequences shown by *Eco*RV endonuclease relied on changes in enzyme affinity for the essential cofactor Mg^{2+} (Taylor & Halford, 1989). These experiments were carried out with two circular pBR322 type plasmids that contained either a single *Eco*RV cognate recognition hexamer d(GATATC) or a nearly cognate sequence d(GTTATC). By use of these two plasmids, it was demonstrated that the K_m for Mg^{2+} at the cognate d(GATATC) site was <1 mM, while that found for the d(GTTATC) sequence was >10 mM. As Mg^{2+} is an absolute requirement for catalysis by the *Eco*RV endonuclease, it is clear that part of the reason for the slower cleavage of nearly cognate against cognate sequences lies in a decreased affinity for Mg^{2+}. We have investigated the effects of Mg^{2+} concentration on the rate of hydrolysis of the parent oligodeoxynucleotide and the two modified oligodeoxynucleotides in which the T bases in the d(GATATC) recognition sequence have been replaced by dU. These two modified oligodeoxynucleotides are among the most slowly cleaved, and if the affinity for Mg^{2+} is decreased with poor substrates, one would expect it to be most pronounced with these two oligomers. The results are summarized in Figure 2, which shows the initial rates of the hydrolysis of the three oligodeoxynucleotides as a function of Mg^{2+} concentration. Although the two dU-containing oligomers are cut much more slowly than the parent, all three oligodeoxynucleotides show a similar rate dependence as the Mg^{2+} concentration varies. We have not performed similar experiments with our other modified base containing oligodeoxynucleotides but have assumed that they too have a similar Mg^{2+} rate dependence. In the subsequent experiments for K_m and k_{cat} determination we have used a Mg^{2+} concentration of 25 mM, a level at which hydrolysis is Mg^{2+} independent. Thus, any effects in K_m and k_{cat} variations observed can be completely ascribed to the base substitutions.

We have determined the K_m and k_{cat} values for the *Eco*RV endonuclease catalyzed hydrolysis of the dodecamer set by measuring the initial velocities as a function of oligodeoxynucleotide concentration. To analyze accurately these steady-state measurements, two standard criteria have to be met. First, the substrate concentration must be greater than that of the enzyme, usually by at least a factor of 10. Second, the substrate concentration used should be varied around its K_m value. Table I shows the K_m values for the entire oligodeoxynucleotide set lie between 0.5 and 8 μM, with 2–4 μM being typical. Unfortunately, with some of the poor substrate oligodeoxynucleotides, using these substrate levels and the consequent 10 times reduced enzyme levels gave rates that were sometimes too slow to be measured accurately over a reasonable time course. The actual enzyme and substrate levels used for each experiment are shown in Table I. Each K_m and k_{cat} determination involved five oligodeoxynucleotide concentrations, and in all cases the top four were at least 10-fold higher than the enzyme concentration utilized. With some of the slower oligodeoxynucleotides (as can be seen from Table I) the lowest substrate level used was only 4 or 5 times higher than the endonuclease levels. Under these conditions we have been able to obtain K_m and k_{cat} values for all the substrate oligodeoxynucleotides except one. Unfortunately, dp(GACGPTATCGTC) was cleaved so slowly by the endonuclease that it proved impossible to determine K_m and k_{cat} values under sensible conditions. This is not due to this oligomer having a much higher K_m value than the others shown in Table I as rate determinations at 10 and 25 μM produced about equal velocities, implying a K_m value below 10 μM and presumably within the range of values given in Table I. Figure 3 gives some representative examples on which the data shown in Table I are based. Figure 3a shows the amounts of product formed as a function of time for five different concentrations of the parent dodecamer dp(GACGATATCGTC). Figure 3b shows a replot of these data as substrate concentration/velocity versus substrate concentration for the determination of K_m and k_{cat} values. Parts c and d of Figure 4 present the same data for dp(GACGATPTCGTC), and finally parts e and f of Figure 4 show the results seen with the dp(GACGATAUCGTC) (the poorest substrate for which K_m and k_{cat}

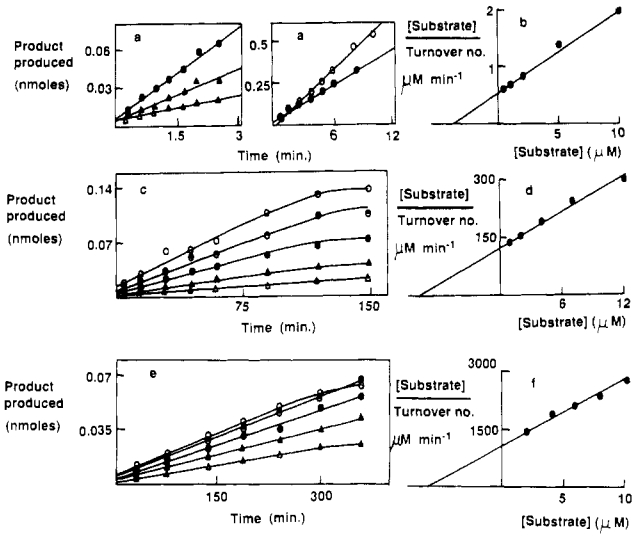

FIGURE 3: K_m and k_{cat} determination for selected oligodeoxynucleotides. (a) shows the amount of product formed with time for the parent 12-mer dp(GACGATATCGTC) at the following substrate levels: 0.5 (△), 1 (▲), 2 (●), 5 (⊖), and 10 μM (○). The concentration of endonuclease used was 0.05 μM. (b) is a replot of these data ([substrate]/turnover number versus [substrate]) to allow K_m and k_{cat} determination. (c) shows product formation against time for dp(GACGATPTCGTC) at oligonucleotide levels of 1 (△), 2 (▲), 4 (●), 7 (⊖-⊖-⊖), and 12 μM (○). The *Eco*RV concentration used was 0.25 μM. (d) is a replot of (c) for kinetic constant determination. (e) gives dp(GACGATAUCGTC) (the slowest substrate for which kinetic constants could be determined) at substrate levels of 2 (△) 4 (▲), 6 (⊖), 8 (⊖), and 10 μM (○). An enzyme concentration of 0.5 μM was used. (f) is a replot of the data in (e) for kinetic constant evaluation.

determinations were possible). As can be seen from Table I, the K_m value for the parent dodecamer is 3.8 μM and the K_m values for the oligodeoxynucleotides containing modified bases all lie between 0.5 and 8 μM. Thus, the K_m values seen are relatively invariant. The k_{cat} values show a much larger variation. A value of 6.9 min^{-1} is seen for the parent, and this is lower for all the analogue oligomers. This decrease is often very large, for example, about 1000-fold in the case of the slowest substrate dp(GACGATAUCGTC). Table I also gives the k_{cat}/K_m ratios which reflect the specifity of an enzyme for alternate substrates (Fersht, 1985). A value of 30.2×10^3 s^{-1} M^{-1} is seen for the parent oligodeoxynucleotide, and two oligomers, namely dp(GACGA[^{2S}T]ATCGTC) and dp(GACGAT[^{7C}A]TCGTC), have similar values, implying equally efficient recognition by the endonuclease and therefore no enzyme substrate contacts at these loci. With all the other oligodeoxynucleotides a large drop in this constant is observed, and as mentioned above most of this decrease is due to lowered k_{cat} values. Also shown in Table I are ΔG_{app} values in kilojoules per mole obtained by using eq 1 in the introduction. It should be noted that the true substrate for the endonuclease is a self-complementary duplex and that the endonuclease acts as a dimer. Therefore, each enzyme substrate complex contains a symmetrical arrangement in which one DNA strand in the double helix interacts with one protein subunit and the other strand the second subunit. The analogue oligodeoxynucleotides therefore are different to the parent in two positions during catalysis. The k_{cat}/K_m values given in Table I therefore represent two changes in the analogue as compared to the parent oligomer. However, we have divided the initial ΔG_{app} value obtained from a ratio of k_{cat}/K_m values (calculated by using eq 1) by 2. Thus, the ΔG_{app} values given in this table refer to the loss of one and not two enzyme oligodeoxynucleotide contacts. This is to allow a simpler comparison with ΔG_{app} values in the literature. The significance of these k_{cat}/K_m ratios and the ΔG_{app} values are discussed under Conclusion.

Conclusion

(a) Comparison of Our Results with Published Data Using Oligodeoxynucleotides Containing Modified Bases To Study the EcoRV Endonuclease. In a recent study with the *Eco*RV endonuclease a K_m of 4 μM and a k_{cat} of 7 min^{-1} were seen for the unmodified oligodeoxynucleotide d(AAAGATATCTT) (Fliess et al., 1986, 1988). This is in excellent agreement with the 6.9 min^{-1} and 3.8 μM we have determined for d(GACGATATCGTC) despite differences in both oligodeoxynucleotide sequence and reaction conditions (11 °C and 10 mM $MgCl_2$ for the former oligomer, 25 °C and 25 mM $MgCl_2$ here). These investigators also performed experiments using d^{7C}A replacement for the two dA bases within the d(GATATC) recognition site of d(AAAGATATCTT), and again the results are in excellent agreement with ours. Thus, substitution of the first dA by d^{7C}A produces an oligodeoxynucleotide that is not a substrate, whereas changing the second dA to d^{7C}A has very little effect on the k_{cat}/K_m value as compared to the parent. Finally, dU for T substitutions in d(AAAGATATCTT) were reported to completely abolish substrate activity. In our hands extremely poor substrates are obtained with the T → dU change, and we have been able to determine K_m and k_{cat} values. A second study with the *Eco*RV endonuclease based on the self-complementary decamer d(CTGATATCAG) has also been undertaken (Mazzarelli et al., 1989). A K_m of 0.46 μM and a k_{cat} of 0.7 min^{-1} were observed. These differ from the values we observed with d(GACGATATCGTC). A possible rationalization is a differential interaction of the endonuclease with the bases flanking the d(GATATC) recognition seqeunce. These were d(CT)-d(AG) for Mazarelli and d(AC)-d(GT) in our case. One might envisage the endonuclease making different interactions with these flanking bases, leading to alterations in K_m and k_{cat} values. Although these interactions are not vital for specificity, they could affect the overall strength of the enzyme substrate interaction. This idea is supported by endonuclease-catalyzed cleavage of plasmids containing several restriction sites where different reaction rates at each site are often seen, e.g., for *Eco*RI. These sites vary only in flanking sequences which are assumed to be responsible for these effects (Halford, 1983). Unfortunately, the low rate of 0.7 min^{-1} seen by Mazzarelli makes the accurate measurement of slow hydrolysis of oligodeoxynucleotides containing modified bases very difficult. Thus, d(CTGATATCAG) derivatives that contain dP for dA or dU for T give "traces" of cleavage. These are in agreement with our results which show that these changes greatly slow the rates. However, we have been able to determine k_{cat} and K_m values for three of the four modified oligodeoxynucleotides that contain dP or dU, in contrast to Mazzarelli. In our case only d(GACGPTATCGTC) was hydrolyzed too slowly for K_m and k_{cat} determination and so corresponds to the trace activity seen by Mazzarelli.

(b) Comparison of EcoRV Hydrolysis of Oligodeoxynucleotides and Plasmids. The k_{cat} and K_m values of 6.9 min^{-1} and 3.8 μM we obtained for dp(GACGATATCGTC) give a k_{cat}/K_m ratio of 30.2×10^{-3} M^{-1} s^{-1}. Recently these parameters have been determined by using plasmid pAT153 which contains a single GATATC *Eco*RV site, and values of k_{cat} of 0.9 min^{-1} and K_m of <0.5 nM were found, giving a k_{cat}/K_m ratio of $>3 \times 10^7$ M^{-1} s^{-1}. With plasmid DNA the rate-limiting step in *Eco*RV-catalyzed hydrolysis has been proposed to be product release (Halford & Goodall, 1988), as deter-

mined from single-turnover and transient kinetics. As we have only undertaken steady-state kinetics we do not as yet know the slowest step for endonuclease-catalyzed hydrolysis of small oligodeoxynucleotides. However, the most likely reason for the differences in k_{cat} values is that the 0.9 min^{-1} seen with the plasmid reflects rate-limiting product release. The increase to 6.9 min^{-1} observed with the dodecamer probably means that product release is no longer the slowest step. Although it is tempting to speculate that the chemical hydrolysis step is rate limiting with oligodeoxynucleotides, we have no proof for this at present. The micromolar K_m value seen with the oligomer as compared to the nanomolar values obtained with plasmids is typical for restriction endonucleases. The value of 3.8 μM we have found with dp(GACGATATCGTC) is about 4 orders of magnitude higher than that of <0.5 nM seen for the d(GATATC) site on pAT153. Similar results have been seen with *Eco*RI, where d(GAATTC) cognate sites in many plasmids were cut with K_m values in the 2–35 nM range (Greene et al., 1975; Modrich & Zabel, 1976; Halford & Johnson, 1981), whereas the values seen for the (GAATTC) sequence in short oligodeoxynucleotides (8–12 base pairs in length) varied between 0.1 and 7 μM (Greene et al., 1975; Brennan et al., 1986; McLaughlin et al., 1987). Several reasons could account for these changes. First, the short oligodeoxynucleotides may not entirely fill the binding site with nonspecific protein contacts being made to, for example, phosphate groups in the DNA flanking the recognition sequence. This has been observed for *Eco*RI by both protection from reaction with ethylnitrosourea (Lu et al., 1981) and X-ray crystallography (McClarin et al., 1986). However, the most likely reason for the K_m variation is the presence of facilitated diffusion of the enzyme when plasmids are used. This can occur by the enzyme binding nonspecifically to noncognate sites and then "sliding" along the DNA until the cognate site is found (von Hippel & Berg, 1989). No such sliding can occur with short oligodeoxynucleotide substrates, and hence their higher K_m values reflect the absence of the nonspecific binding crucial to this process.

As well as differences in K_m and k_{cat} values for the *Eco*RV endonuclease with plasmids and oligodeoxynucleotides, we have also observed different behavior toward Mg^{2+}. This metal ion is an essential cofactor for hydrolysis, probably being bound to the scissile phosphate group. With the d(GATATC) site on pAT153 the K_m for Mg^{2+} was below 0.5 mM (Halford & Goodall, 1988; Taylor & Halford, 1988). With a noncognate site on pAT153 [in this case d(GTTATC)] a much higher level of Mg^{2+} was required (K_m > 10 mM) for maximum rates. This noncognate site is cut much more poorly than the cognate (k_{cat}/K_m for cognate site >3 × 10^7 M^{-1} s^{-1}; k_{cat}/K_m for noncognate site 36 M^{-1} s^{-1}). With dp(GACGATATCGTC) and the two oligodeoxynucleotides that contained dU in place of T, we observed the same dependence on Mg^{2+} levels for the hydrolysis, and in all cases high Mg^{2+} levels (>10 mM) were required for maximum rates. These dU-containing oligodeoxynucleotides are much poorer substrates than the parent and can therefore be considered analogous to the noncognate site on pAT153, where a base substitution has taken place. It is clear that in plasmids changing from a cognate to a noncognate sequence raises the K_m for Mg^{2+} but that this does not happen in oligodeoxynucleotides. At present, the reasons underlying these differences in behavior with Mg^{2+} for oligodeoxynucleotides and plasmids remain obscure. These changes make the kinetics seen with plasmids and oligodeoxynucleotides difficult to compare directly. A full understanding of these effects is, however, clearly required for a thorough explanation of how the endonuclease works.

(*c*) *Interpretation of the Results Seen with Oligodeoxynucleotides Containing Modified Bases.* From Table I it can be seen that the modified base containing oligodeoxynucleotides can be divided into three classes. The first class consists of two dodecamers [recognition sequences d(GA-[^{2S}T]ATC and d(GAT[^{7C}A]TC] that have a k_{cat}/K_m ratio similar to that of the parent. These results imply that the endonuclease makes no contact with the 2-keto oxygen of the first T and the 7-ring nitrogen of the second dA in the d(GATATC) sequence. As mentioned in the previous paper, from CD spectroscopy, it is not clear whether ^{2S}T-containing oligodeoxynucleotides have the same overall structure as the parent. The similar values of the constants seen here for d(GA[^{2S}T]ATC) and d(GATATC) would seem to argue against a large conformational difference, which would be expected to change the k_{cat}/K_m values. The second class of oligodeoxynucleotides, comprising most of the members listed in Table I, are substrates for the endonuclease, but have lowered k_{cat}/K_m values as compared to the parent. As can be seen from Table I, the decrease in k_{cat}/K_m values is largely due to a drop in k_{cat} values (which vary over 3 orders of magnitude) rather than any change of K_m (which is relatively constant with only a 10-fold variation). The term k_{cat} describes the process by which the ES complex is converted to EP, providing all the binding steps are fast and only one ES complex is present. The much faster k_{cat} value seen with the cognate dodecamer as opposed to most of the oligodeoxynucleotides containing modified bases implies that many of the individual binding interactions are used to reduce the transition-state energy barrier (and thereby speed up the ES to EP conversion) rather than to increase substrate binding (which would affect K_m). We have performed only steady-state kinetics, and if product release is rate limiting, it can contribute to k_{cat} values. This would seem to be unlikely here as all the analogues are turned over more slowly that the parent. These modified substrates must turn over to modified products, and one would not expect that these would be released more slowly than the cognate products, leading to the observed drop in k_{cat}. Often when product release is rate limiting, analogues are handled faster than the true substrate as the analogue product formed is rapidly released. Nevertheless, contributions to k_{cat} due to product release steps cannot be entirely eliminated.

Many of the base analogues we have used delete a potential hydrogen-bonding partner with the endonuclease. Many studies have been carried out to evaluate the strengths of enzyme substrate hydrogen bonds. These have involved determination of the k_{cat}/K_m ratios for parent and analogue and subsequent evaluation of the ΔG_{app} values using eq 1. Usually the deletion of a hydrogen bond that does not leave a charged partner costs between 2 and 7.5 kJ mol^{-1} (Fersht, 1987a,b, 1988). Abolition of a hydrogen bond leaving a charged partner resulted in a loss of binding energy of between 12 and 25 kJ mol^{-1}. With d(G[^{3C}A]TATC), d(GATPTC), and d(GAT-[^{3C}A]TC) a hydrogen-bonding group on the oligodeoxynucleotide is deleted, and the ΔG_{app} values found consequent on these changes lie between −4.8 and −6.7 kJ mol^{-1}. The simplest explanation is that the endonuclease forms a hydrogen bond to each of these positions. The ΔG_{app} values also suggest that the groups on the protein responsible for these bonds are uncharged. Evidence for this comes from the crystal structure of the endonuclease (F. Winkler and S. Halford, personal communication), which shows that most of the amino acids near the DNA binding site are uncharged. With d(GPTATC) we have been unable to obtain k_{cat} and K_m values, as it is a very poor substrate. However, the 6-NH_2 group of the first

dA is clearly important for the endonuclease, and we would postulate a direct enzyme substrate hydrogen bond at this site also. Poor substrates are also obtained when the T residues in the d(GATATC) sequence are replaced with dU, a conversion of a CH_3 group to H. This gives ΔG_{app} values of −7.8 kJ mol^{-1} in the first position and −9.4 kJ mol^{-1} in the second. The energy contributed by a substrate CH_3 group has been analyzed for several proteins and varies between 6.5 kJ mol^{-1} for chymotrypsin (Dorovska et al., 1972) and 14 kJ mol^{-1} for various aminoacyl-tRNA synthetases (Loftfield & Eigner, 1966; Owens & Bell, 1970; Fersht & Dingwall, 1979; Fersht et al., 1980). The values of −14 kJ mol^{-1} kJ mol^{-1} seen with tRNA synthetases indicate an extremely well formed CH_3 binding hole that makes a very tight interaction with the methyl group. The values we see of −7.8 and −9.0 kJ mol^{-1} show a poorer fit between the endonuclease and the two CH_3 groups on T as compared to synthetases. Nevertheless, it is obvious that the enzyme interacts with both CH_3 groups and that these contacts are important in determining specificity. The importance of the CH_3 group of thymidine in a variety of protein nucleic acid complexes has been discussed (Ivarie, 1987).

We also obtained poor substrates when some keto oxygen atoms were converted to sulfur. This situation is more difficult to analyze than simple deletion of a functional group by replacement with hydrogens. First, the C=S bond distance is 0.04 nm longer than the C=O, and sulfur (0.185 nm) has a larger van der Waals radius than oxygen (0.14 nm) (Saenger, 1984), leading to the introduction of some degree of steric hindrance. Second, the thioketo group can still form hydrogen bonds, although to a lesser extent than the parent keto function (Saenger, 1984). With d(GATA[^{2S}T]C) a much lower k_{cat}/K_m as compared to the parent was seen as may be expected if the enzyme contacts the oligodeoxynucleotide in this position. However, as detailed in the previous paper this oligodeoxynucleotide has an irregular CD spectrum that is possibly indicative of an overall conformational change. Therefore, we feel it would be premature to postulate an enzyme contact to the 2-oxygen atom of the second T at this stage. With d(GA[^{4S}T]ATC) a large lowering of k_{cat}/K_m is measured. This is indicative of the importance of this 4-oxygen atom. This drop in k_{cat}/K_m could be due to either an enzyme substrate hydrogen bond or the inability of the endonuclease to tolerate any bulk at this site.

The third category of base-modified oligodeoxynucleotides gives no measurable cleavage by the endonuclease. Into this group fall d(G[^{7C}A]TATC), d(GA[^{4H}T]ATC), d(GATA-[^{4H}T]C, and d(GATA[^{4S}T]C). With the exception of the last of this group, all these changes involve the deletion of a hydrogen-bonding function. If the consequence of this was simply to eliminate an enzyme substrate hydrogen bond, one would expect a ΔG_{app} of between 2 and 7.5 kJ mol^{-1}, giving poor but nevertheless measurable substrates (as is found for several members of the oligodeoxynucleotides in Table I). Two explanations are possible. First, the above oligodeoxynucleotides have altered conformations as compared to the parent [discussed further under (e)]. Second, a cooperative mechanism of discrimination proposed for the *Eco*RI endonuclease (McClarin et al., 1986; Rosenberg et al., 1987) may be in operation. Here the loss of one contact between the enzyme and the oligodeoxynucleotide leads to the loss of several other favorable interactions. A possible mechanism involves displacement of the amino acid that forms the contact and a consequent alteration of the structure of the enzyme binding site. This structural change could result in either the further loss of enzyme substrate interactions or the production of an active site incapable of catalyzing hydrolysis.

(*d*) *Discrimination between Cognate and Noncognate Sites by the EcoRV Endonuclease.* Using the oligodeoxynucleotides containing modified bases, we have been able to postulate many contacts between the endonuclease and the T and dA bases within the d(GATATC) recognition sequence. In the major groove these are the 6-NH_2 and 7N of the first dA, and the 6-NH_2 of the second dA and the 4-keto oxygen and 5-CH_3 groups of both T bases. In the minor groove, the two N-3 ring nitrogens of dA appear to contact the enzyme. The high specificity of cleavage shown by the endonuclease is mainly due to only the cognate sequence making the correct and maximal number of interactions to the enzyme. The lowering of k_{cat}/K_m on loss of an interaction is mainly due to a decrease in k_{cat}, and the enzyme primarily uses these binding interactions to lower the transition state of the reaction rather than to bind the substrate in the ground state. Only with the cognate sequence are all the enzyme protein contacts made, the k_{cat}/K_m maximized, and the transition-state barrier lowered enough to allow efficient catalysis. Discrimination could also be amplified by the cooperative mechanism outlined above. Noncognate sequences arising from natural base substitutions (in contrast to analogues) not only fail to make the full number of interactions but also introduce unfavorable clashes, for example, by placing a pair of hydrogen-bond donors or acceptors in close proximity. These unfavorable factors would be expected to reinforce any discrimination seen due to simple loss of an interaction. Therefore, the endonuclease should distinguish between its cognate sequence and natural variants even more strongly than against variants containing modified bases. It should be noted that other discrimination mechanisms—not studied here—are also used by the *Eco*RV endonuclease. In elegant studies using plasmids, Halford and his group (Taylor & Halford, 1989) have shown that the K_m for Mg^{2+} is much higher for noncognate as compared to cognate sequences. With noncognate sequences the dimeric endonuclease is therefore likely to only contain one Mg^{2+} per dimer and so only cleave one strand of duplex DNA. This nicked DNA can be repaired by cellular DNA ligase. In contrast, cognate sequences have endonuclease containing two Mg^{2+} per dimer, and therefore both DNA strands are simultaneously cut. This double-cut structure cannot be repaired. This means that incorrect cleavage at noncognate sequences can be repaired, adding to enzyme fidelity.

(*e*) *Validity of the Use of Oligodeoxynucleotides Containing Modified Bases To Determine Enzyme Substrate Contacts.* In the above discussion we have assumed that the presence of modified bases does not perturb the overall oligodeoxynucleotide conformation, and this has allowed us to postulate several enzyme oligodeoxynucleotide contacts. How valid is this assumption, and what are the possible consequences if it does not hold? Studies of these kinds are best carried out in conjunction with a crystal structure. Thus, an X-ray structure can identify contacts conclusively, and use of analogues can measure their strengths. This has so far been carried out only with the *Eco*RI endonuclease (Seela & Driller, 1986; Brennan et al., 1986; Seela & Kehne, 1987; McLaughlin et al., 1987). The crystal structure shows contacts to the 6-keto group and 7-ring nitrogen of dG and to the 6-NH_2 and 7-ring nitrogen of both dA's in the GAATTC recognition sequences (McClarin et al., 1986). These interactions were found by using appropriate base analogues, and when a contact is seen in the crystal structure, the use of analogues seems to reproduce the result. Unfortunately, removal of the 2-amino group from dG and the 5-CH_3 groups from the two T residues also resulted in a drop in k_{cat}/K_m, even though no contacts were

seen in the crystal structure. Thus, the base analogue approach may yield false positives. Possible explanations include the crystal structure being performed in the absence of Mg^{2+} and the active site not being assembled. On Mg^{2+} addition a conformational change could take place to produce the active site and more enzyme substrate contacts be formed. Alternatively, the base analogues could change the overall DNA structure, giving a different oligodeoxynucleotide conformation to the parent. This undoubtedly remains the weakness of this approach. Clearly, when this occurs, a drop in k_{cat}/K_m could be due as much to the change in conformation as to the specific loss of an enzyme substrate contact. In the previous paper we showed essentially normal CD and T_m curves for the entire oligodeoxynucleotide set, indicating that they have structures that are similar overall. However, these methods may not reflect small local distortions that could be critical for endonuclease activity. These small structural distortions could lead to interpretations of the results other than the ones we have given, i.e., on the basis of simple loss of enzyme oligodeoxynucleotide contacts. This would appear to be particularly a problem with ^{4H}T, which cannot form any Watson–Crick hydrogen bonds with dA. Small independent movements in the relative positions of these base pairs could result in a small local perturbation of structure relative to that present with a dA–T base pair. This small perturbation would slightly alter the ability of the enzyme to effectively contact the pyrimidine methyl or the purine 6-NH_2 or N7 groups and eliminate the sequence as a substrate even though no contact to the 4-keto group is made. To a much lesser extent this criticism applies to all the base analogues. A final possibility is that the modified bases may prevent any distortion of the oligodeoxynucleotide that is a consequence of binding the enzyme and necessary for catalysis. With the *Eco*RI endonuclease X-ray crystallography has shown rather large distortion of the DNA on binding (McClarin et al., 1986). In one instance (binding of an operator sequence to phage 434 repressor; Aggarwal et al., 1988) the DNA was distorted, and this resulted in extra (i.e., non-Watson–Crick) hydrogen bonds between bases. Thus, with analogues that lack hydrogen-bonding functions it is possible that these extra interbase hydrogen bonds will be unable to form and therefore oligodeoxynucleotide distortion on protein binding will be prevented, or at least hindered. It would be difficult to distinguish experimentally between loss of an enzyme oligodeoxynucleotide hydrogen bond and loss of an extra interbase hydrogen bond that is necessary to achieve substrate distortion on binding.

It is hoped that the crystal structure of the *Eco*RV endonuclease will be published soon (F. Winkler, personal communication). The high-resolution structure, when published, together with the analogue work described here should together go a long way to explaining the mechanism of *Eco*RV endonuclease action and how specificity is achieved.

Acknowledgments

We thank Dr. S. Halford (University of Bristol) and Dr. F. Winkler (Hoffman-La Roche Ltd., Basel) for discussions and making unpublished data available. We also thank Miss Clare Philo for typing the manuscript.

References

Aggarwal, A. K., Rodgers, D. W., Drottar, M., Ptashne, M., & Harrison, S. C. (1988) *Science 242*, 899–907.

Brennan, C. A., Van Cleve, M. D., & Gumport, R. I. (1986) *J. Biol. Chem. 261*, 7270–7278.

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

Dorovska, V. N., Varfolomeyev, S. D., Kazanskaya, N. F., Klyosov, A. A., & Martinek, K. (1972) *FEBS Lett. 23*, 122–124.

Dwyer-Hallquist, P., Kezdy, F. S., & Agarwal, K. L. (1982) *Biochemistry 21*, 4693–4700.

Fersht, A. R. (1985) in *Enzyme Structure and Mechanism*, 2nd ed., p 105, Freeman New York.

Fersht, A. R. (1987a) *Trends Biochem. Sci. 12*, 301–304.

Fersht, A. R. (1987b) *Biochemistry 26*, 8031–8037.

Fersht, A. R. (1988) *Biochemistry 27*, 1577–1588.

Fersht, A. R., & Dingwall, C. (1979) *Biochemistry 18*, 1238–1245.

Fersht, A. R., Shindler, J. S., & Tsui, W. C. (1980) *Biochemistry 19*, 5520–5524.

Fliess, A., Wolfes, H., Rosenthal, A., Schwellus, K., Blocker, H., Frank, R., & Pingoud, A. (1986) *Nucleic Acids Res. 16*, 3643–3474.

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

Greene, P. J., Poonian, M. S., Nussbaum, A. L., Tobias, L., Garfin, D. E., Boyer, H. W., & Goodman, H. M. (1975) *J. Mol. Biol. 99*, 237–261.

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

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

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

Hayakawa, T., Ono, A., & Ueda, T. (1988) *Nucleic Acids Res. 16*, 4761–4776.

Ivarie, R. (1987) *Nucleic Acids Res. 15*, 9975–9983.

Jiriciny, J., Wood, S. G., Martin, D., & Ubasawa, A. (1986) *Nucleic Acids Res. 14*, 6579–6590.

Kholmina, G. V., Rebentish, B. A., Skoblov, Y. S., Mironov, A. A., Yankovskii, Y., Kozlov, Y. I., Glatman, L. I., Moroz, A. E., & Debabov, V. G. (1980) *Dokl. Acad. Nauk. SSSR 253*, 495–497.

Kita, K., Hiraoka, N., Kimizaka, F., Obayashi, A., Kojima, H., Takahashi, H., & Saito, F. (1985) *Nucleic Acids Res. 13*, 7015–7024.

Loftfield, R. B., & Eigner, E. A. (1966) *Biochim. Biophys. Acta 130*, 426–448.

Lu, A. L., Jack, W. E., & Modrich, P. (1981) *J. Biol. Chem. 256*, 13200–13206.

Maniatis, T., Fritsch, E. F., & Sambrook, J. (1982) in *Molecular Cloning—a Laboratory Manual*, pp 122–123, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Mazzarelli, J., Scholtissek, S., & McLaughlin, L. W. (1989) *Biochemistry 28*, 4616–4622.

McClarin, J. A., Frederick, C. A., Wang, D., Greene, P., Boyer, H. W., Grable, J., & Rosenberg, J. M. (1986) *Science 234*, 1526–1541.

McLaughlin, L. W., Benseler, F., Graeser, E., Piel, N., & Scholtissek, S. (1987) *Biochemistry 26*, 7238–7245.

Modrich, P., & Zabel, D. (1976) *J. Biol. Chem. 251*, 5866–5874.

Nwosu, V. U., Connolly, B. A., Halford, S. E., & Garnett, J. (1984) *Nucleic Acids Res. 16*, 3705–3720.

Ono, A., & Ueda T. (1987) *Nucleic Acids Res. 12*, 3059–3072.

Ono, A., Sato, M., Ohtani, Y., & Ueda, T. (1984) *Nucleic Acids Res. 12*, 8939–8949.

Owens, S. L., & Bell, F. E. (1970) *J. Biol. Chem. 245*, 5515–5523.

Rosenberg, J. M., McClarin, J. A., Frederick, C. A., Grable, J., Boyer, H. W., & Greene, P. J. (1987) *Gene Amplif. Anal. 5*, 119–145.
Saenger, W. (1984) in *Principles of Nucleic Acid Structure*, p 185, Springer-Verlag, Berlin.
Schildkraut, I., Banner, C. D., Rhodes, C. S., & Parekh, S. (1984) *Gene 27*, 327–329.
Seela, F., & Driller, H. (1986) *Nucleic Acids Res. 14*, 2319–232.
Seela, F., & Kehne, A. (1987) *Biochemistry 26*, 2232–2238.
Taylor, J. D., & Halford, S. E. (1989) *Biochemistry 28*, 6198–6207.
von Hippel, P. H., & Berg, O. G. (1989) *J. Biol. Chem. 264*, 675–678.
Yolov, A. A., Vinogradova, M. N., Gromova, E. S., Rosenthal, A., Cech, D., Veiko, V. P., Metelev, V. G., Kosykh, V. G., Buryanov, Y. I., Bayev, A. A., & Shaborova, Z. A. (1985) *Nucleic Acids Res. 13*, 8983–8997.

Cross-Linking of the 25- and 20-Kilodalton Fragments of Skeletal Myosin Subfragment 1 by a Bifunctional ATP Analogue†

Shinsaku Maruta,‡ Morris Burke,§ and Mitsuo Ikebe*,‡

Department of Physiology and Biophysics, School of Medicine, and Department of Biology, Case Institute of Technology, Case Western Reserve University, Cleveland, Ohio 44106

Received January 4, 1990; Revised Manuscript Received July 24, 1990

ABSTRACT: The bifunctional photoreactive ATP analogue azidonitrobenzoyl-8-azido-ATP (ANB-8-N_3-ATP) was synthesized. This ATP analogue carriers photoreactive azido groups at the eighth position of the adenine ring and at the 3′ position of ribose. Photolysis of this analogue in the presence of skeletal muscle α-chymotryptic subfragment 1 (S-1) resulted in a new 120-kDa band, while photolysis in the presence of the tryptic S-1 produced a new 45-kDa band. The 45-kDa peptide was shown to be combined with the 25-kDa N-terminal and 20-kDa C-terminal fragments since it was labeled with a monoclonal antibody specific for the N-terminal 25-kDa segment of the S-1 heavy chain, and it was also found to retain the fluorescence of (iodoacetamido)fluorescein attached specifically to the SH-1 thiol of the C-terminal 20-kDa segment. These results indicate that the 25- and 20-kDa peptides are in close contact with the ATPase active site.

It is well established that muscle contraction is performed by the relative sliding of thick and thin filaments (Huxley, 1969) in which the energy is supplied by ATP hydrolysis catalyzed by myosin. Although this scenario is well established, the molecular mechanism of energy transduction is still obscure. To shed light on this problem, it is essential to characterize the structure of the myosin active site. The assignment of the functionally important domains has been developed by using limited proteolysis. Initially, it was found that the head region of the myosin molecule [subfragment 1 (S-1)][1] (Margossian & Lowey, 1973) contains both the ATPase and actin binding sites. The 95-kDa heavy chain of S-1 is further cleaved by trypsin into three substructures connected by two protease-sensitive hinges (Balint et al., 1978; Mornet et al., 1979; Yamamoto & Sekine, 1979). Beginning with the N-terminus of the heavy chain, the approximate sizes of these three fragments are 25, 50, and 20 kDa. The N-terminal 25-kDa peptide contains the reactive lysine residue (Mornet et al., 1981a; Hozumi & Muhlrad 1981), while an activity-critical carboxyl group (Korner et al., 1983) and part of the actin binding site (Mornet et al., 1981b; Sutoh, 1983) have been located on the 50-kDa peptide. The two reactive thiols of myosin, termed SH-1 and SH-2, are located on the 20-kDa peptide (Balint et al., 1978), and it has been known that the modification of these thiols significantly alters the ATPase activity of myosin (Sekine et al., 1962). For example, the cross-linking of these thiols by bifunctional cross-linking reagents in the presence of Mg^{2+}-ADP abolishes the ATPase activity (Burke & Reisler, 1977) which is due to the stable trapping of Mg^{2+}-ADP on the enzyme (Wells & Yount, 1979).

To localize the ATP binding site on the myosin molecule, several photoaffinity labeling ATP analogues have been used. Szilagyi et al. (1979) showed that arylazido-β-alanyl-ATP specifically labeled the 25-kDa tryptic peptide. The 25-kDa peptide was also labeled by NANDP, a photoaffinity analogue of ADP (Nakamaye et al., 1985; Okamoto & Yount, 1985) and Mant-2-N_3-ADP (Maruta et al., 1989). On the other hand, Mahmood et al. (1984) showed that Bz_2-ATP containing a photoprobe at the ribose ring labeled the 50-kDa peptide. Recently, Maruta et al. (1989) reported that Mant-8-N_3-ATP labeled the 20-kDa peptide of myosin and suggested the close proximity of the 20-kDa peptide to the ATP binding site. In this paper, the bifunctional photoreactive ATP analogue

† This work was supported by NIH Grant AR 38888. M.I. is an Established Investigator of the American Heart Association and is a Syntex Scholar.
* Address correspondence to this author.
‡ Department of Physiology and Biophysics.
§ Department of Biology.

[1] Abbreviations: S-1, subfragment 1; ANB, azidonitrobenzoyl; NANDP, *N*-(4-azido-2-nitrophenyl)-2-aminoethyl diphosphate; Bz_2, 4-benzoylbenzoyl; IAF, 5-(iodoacetamido)fluorescein; Mant, methylanthraniloyl; SDS, sodium dodecyl sulfate; PAGE, polyacrylamide gel electrophoresis; TLC, thin-layer chromatography; IR, infrared; SH-1, reactive thiol 1; SH-2, reactive thiol 2; PBS, phosphate-buffered saline; DMF, dimethylformamide; MOPS, 4-morpholinepropanesulfonic acid.

0006-2960/90/0429-9910$02.50/0