Interaction of the Restriction Endonuclease *Eco*RV with the Deoxyguanosine and Deoxycytidine Bases in Its Recognition Sequence†

Timothy R. Waters‡,§ and Bernard A. Connolly*,‖

Department of Biochemistry, The University of Southampton, Southampton SO9 3TU, U.K., and Department of Biochemistry and Genetics, The University of Newcastle upon Tyne, Newcastle upon Tyne NE2 4HH, U.K.

Received July 23, 1993; Revised Manuscript Received November 3, 1993®

ABSTRACT: The interaction of the *Eco*RV restriction endonuclease with the dG and dC bases in its recognition sequence (GATATC) has been studied using base analogues. These modified dG and dC bases each have a single potential protein contact removed. The analogues have been incorporated into the self-complementary dodecamer d(pGACGATATCGTC) at the appropriate positions (underlined). Many of the analogues caused no change in the T_m of the duplex or else lowered the T_m by a small amount such that a duplex was still formed at temperatures suitable for enzyme assay. However, the dG analogue 2-aminopurine-1-β-D-2′-deoxyriboside destabilized the duplex to such an extent that the 12′-mer could not be used for enzyme assays. To overcome this, a longer self-complementary 18′-mer was used with this modified base. The circular dichroism spectra of the modified base containing 12′-mers (and the 18′-mer in the case of 2-aminopurine) were very similar to the parent sequences lacking modified bases. This demonstrates the formation of B-DNA structures in all cases and similar overall conformations. The K_m and k_{cat} values for the various modified oligomers have been determined, and these data have been used to assess the roles that functional groups on the dG and dC bases play in the recognition and hydrolysis of GATATC sequences by the endonuclease. The results obtained here have been compared to the crystal structures of the *Eco*RV complexed with a GATATC sequence, and this has allowed a critical evaluation of the base analogue approach. Both methods indicate that the 6-keto oxygen and 7-ring nitrogen of dG exposed in the major groove are vital for DNA recognition and hydrolysis.

The *Eco*RV restriction endonuclease recognizes GATATC sequences in double-stranded DNA and cleaves between the central T and dA residues [for recent reviews, see Bennett and Halford (1989), Halford et al. (1993), and Vipond and Halford (1993)]. High-resolution crystal structures have been solved for both the free protein and complexes with oligodeoxynucleotides either lacking (enzyme–nonspecific DNA complex) or containing (enzyme–specific DNA complex) the GATATC sequence (Winkler, 1992; Winkler et al., 1993). The protein is dimeric with a deep cleft between the two subunits which forms the DNA-binding site. The enzyme–specific DNA complex shows that the DNA binds with its minor groove facing toward the bottom of the cleft and the major groove oriented toward the top. The protein makes most of its contacts to the DNA via two polypeptide loops. The first of these, centered on Asn 70, approaches the minor groove and contacts phosphate groups. The second loop is centered on Asn 185 and penetrates into the major groove of the DNA. Twelve direct hydrogen bonds (six from each subunit) are made between this loop and the GATATC sequence as shown in Table 1. The bound DNA is extremely distorted when compared to typical structural parameters seen in B-DNA. In contrast, the enzyme–nonspecific DNA complex shows very little distortion of the bound DNA, and the 12 direct hydrogen bonds between the 185 loop and the GATATC sequences are entirely absent.

Table 1: Direct Contacts Observed by Crystallography between the *Eco*RV Restriction Endonuclease and Its GATATC Cognate Sequence[a]

base	loci on base	protein contact	interaction	base analogue
G	6-O	Asn 185 (backbone)	H-bond	$d^{2am}P$
	7-N	Gly 184 (backbone)	H-bond	$d^{7C}G$
A	6-NH_2	Asn 185 (side chain)	H-bond	dP
	7-N	Asn 185 (side chain)	H-bond	$d^{7C}A$
T[b]				
A[b]				
T	4-O	Thr 186 (side chain)	H-bond	^{4H}T
	5-CH_3	Asn 185/Thr 186 (side chain)	van der Waals	dU
C	4-NH_2	Gly 182 (backbone)	H-bond	$d^{4H}C$

[a] The modified bases that can be used to probe these interactions are also listed. [b] No direct contacts between oligonucleotide and protein but extreme DNA distortion.

These structures are important in explaining the high specificity of the *Eco*RV endonuclease. Nevertheless, taken in isolation, they can be misleading. The most obvious interpretation is that the 12 hydrogen bonds seen only with cognate sequences ensures that they are much more tightly bound than all other sequences and that this leads to the observed discrimination. However, in the absence of the essential cofactor Mg^{2+}, the endonuclease binds to all DNA sequences with identical K_d values (Taylor et al., 1991). Thus the 12 extra hydrogen bonds seen only with cognate sequences contribute nothing to substrate binding. Despite this total lack of binding discrimination, GATATC sequences are cut at least 10^6 times more efficiently than all others (Taylor & Halford, 1989, 1992; J. Alves and A. Pingoud, personal communication). It has also been observed that the affinity of the endonuclease for Mg^{2+} varies depending on whether or not cognate sequences are being cut. For the hydrolysis of GATATC sites the K_d for Mg^{2+} is $\ll 1$ mM, whereas for

† This work was supported by the Molecular Recognition Initiative of the SERC. T.R.W. was an SERC supported Ph.D. student. B.A.C. is a Fellow of the Lister Institute of Preventive Medicine.

* Address correspondence to this author.
‡ The University of Southampton.
§ Current address: Department of Biochemistry and Molecular Biology, University College London, Gower Street, London WC1E 6BT, U.K.
‖ The University of Newcastle upon Tyne.

® Abstract published in *Advance ACS Abstracts,* January 15, 1994.

0006-2960/94/0433-1812$04.50/0

noncognate sites the K_d is $\gg$10 mM (Taylor & Halford, 1989). The combination of crystallographic and kinetic data has led to the following proposal for the mechanism of the endonuclease (Vipond & Halford, 1993; Halford et al., 1993). With GATATC sequences the binding energy available from the 12 additional hydrogen bonds is not used to strengthen substrate binding but rather to distort the DNA and produce the unusual (and what in free solution would be energetically unfavorable) enzyme-bound conformation. This distortion moves the scissile phosphate group toward the active site, part of which is made up from Asp 74 and Asp 90, and simultaneously assembles the Mg^{2+}-binding site. The DNA distortion is required for these three negatively charged groups to adopt the necessary orientation for Mg^{2+} binding. Thus only following DNA distortion, which is seen with GATATC sequences alone, does Mg^{2+} bind and hydrolysis take place.

The exact role of the 12 hydrogen bonds given in Table 1 can be further probed using base analogues. This method has been widely used to study a variety of DNA-binding proteins (Yansura et al., 1979; Goedell et al., 1978; Fisher & Caruthers, 1979) and several enzymes that act upon DNA especially restriction endonucleases (Dwyer-Hallquist et al., 1982; Ono et al., 1984; Yolov et al., 1985; Seela & Driller, 1986; Jiriciny et al., 1986; Brennan et al., 1986a,b; McLaughlin et al., 1987; Seela & Kehne, 1987; Ono & Ueda, 1987; Hayakawa et al., 1988; Lesser et al., 1990, 1993; Aiken et al., 1991; Zebala et al., 1992b). This technique consists of preparing a base analogue in which a group or atom that has the potential to interact with a DNA-binding protein (Seeman et al., 1976) is deleted. The analogue is incorporated into an oligodeoxynucleotide of appropriate sequence, and the binding or kinetic parameters are determined with the protein of interest. This method can reveal protein–DNA contacts and also be used to determine how much they contribute to specificity. However, the method has its pitfalls, the most serious one being that the base analogue may cause an overall change in the conformation of the oligodeoxynucleotide or interfere with any DNA distortion that is seen on protein–DNA complex formation. A good review discussing the method has recently been published (Aiken & Gumport, 1991).

A few years ago our group reported on the interaction of the *Eco*RV restriction endonuclease with oligodeoxynucleotides containing dA and T analogues (Connolly & Newman, 1989; Newman et al., 1990a,b; Cosstick et al., 1990). In these studies the self-complementary 12′-mer d(pGACGATATCGTC) was used as a substrate and dA and T analogues were incorporated at the central ATAT sequence that comprises four out of the six bases in the recognition sequence. Steady-state kinetic methods were used to determine K_m and k_{cat} values. Similar studies have been carried out by other groups (Fliess et al., 1986, 1988; Mazzarelli et al., 1989). All these experiments were performed before the crystal structures had been published, and in the absence of this information most of the results obtained were simply interpreted in terms of loss of contacts between the enzyme and its substrate. In this publication the effect of dG and dC base analogues on catalysis by the endonuclease is reported, completing the study of the GATATC site. The major change since the previous studies is the availability of crystal structures, and so we are able to discuss critically both how the endonuclease interacts the dG and C bases in its recognition sequence and also to review the previous results obtained with the dA and T analogues.

MATERIALS AND METHODS

The following methods have been extensively described in previous articles from this group (Connolly & Newman, 1989; Newman et al., 1990a,b; Connolly, 1991, 1992). (1) The synthesis, purification and characterization of oligodeoxynucleotides. (2) The accurate determination of the hyperchromicities, extinction coefficients, and concentrations of the oligodeoxynucleotides by digestion to their constituent deoxynucleosides. (3) The measurement of T_m values and circular dichroism spectra of the oligodeoxynucleotides. (4) The 5′-phosphorylation of the oligodeoxynucleotides. (5) High-pressure liquid chromatography (HPLC) methods. (6) The purification, characterization, and concentration determination of the *Eco*RV restriction endonuclease. The preparation of appropriately protected phosphoramidite derivatives of 2-aminopurine-(1-β-D)-2′-deoxyriboside ($d^{2am}P$), 6-thiodeoxyguanosine ($d^{6S}G$), and 2-pyrimidinone-(1-β-D)-2′-deoxyriboside ($d^{4H}C$) have been described (Connolly, 1991, 1992; Waters & Connolly, 1992a; Taylor et al., 1993). An appropriately protected derivative of 7-deazadeoxyguanosine ($d^{7C}G$), 5′-dimethoxytrityl-*N*-isobutyryl-7-deazadeoxyguanosine, was a kind gift from Dr. Philip Gottlieb (University of Delaware). This was converted to its 3′-phosphoramidite in the usual manner (Connolly, 1991, 1992). The preparation of 3-deazadeoxyguanosine ($d^{3C}G$) and its conversion to a protected phosphoramidite was as reported (Revanker et al., 1984; Seela & Lampe, 1991). Appropriately protected phosphoramidite derivatives of deoxyinosine (dI) and 5-methyldeoxycytidine ($d^{5Me}C$) were purchased from Cruachem Ltd. (Glasgow, Scotland).

Determination of Whether or Not Oligodeoxynucleotides Are Substrates for the EcoRV Endonuclease. 4 nmol (20 μM) of double-stranded oligodeoxynucleotide (not phosphorylated at the 5′-position) was incubated in 200 μL of 50 mM Hepes, pH 7.5, 10 mM $MgCl_2$, 100 mM NaCl, and 10 units (manufacturers units, special grade for molecular biology, obtained from Promega, Southampton, U.K.) of alkaline phosphatase in the presence of 340 pmol (1.7 μM) of endonuclease dimer. The reaction was monitored by reverse-phase HPLC (Newman et al., 1990a). The alkaline phosphatase removes the 5′-phosphate from one of the products, and this eases separation of the two product oligodeoxynucleotides. Controls showed that alkaline phosphatase alone did not give rise to cleavage of the substrate nor did it affect endonuclease catalyzed rates. If, after 1 h, no hydrolysis was observed, a further 680 pmol of endonuclease was added. If needed, this was followed by 680 pmol more after a further hour, and the mixture was then left overnight. When hydrolysis was seen, it was allowed to proceed to completion, and the product peaks were purified by HPLC and separately analyzed by base composition analysis (Newman et al., 1990a).

K_m and k_{cat} Determination. The K_m and k_{cat} values for the 5′-phosphorylated oligodeoxynucleotides were determined using a continuous spectrophotometric assay (based on cleavage of double-stranded substrates to single-stranded products) that has been previously described (Waters & Connolly, 1992b). Reactions were performed either in 1-mL volumes (1-cm pathlength) for oligodeoxynucleotides concentrations >1.2 μM or in 2-mL volumes (2-cm pathlength) for concentrations <1.2 μM. A buffer consisting of 50 mM Hepes, pH 7.5, 25 mM $MgCl_2$, and 55 mM NaCl at 25 °C was used. The oligodeoxynucleotide concentration varied between 0.2 and 8.3 μM and the enzyme concentration between 2 and 300 nM. In most cases the oligodeoxynucleotide concentration exceeded that of the enzyme by a factor of 10,

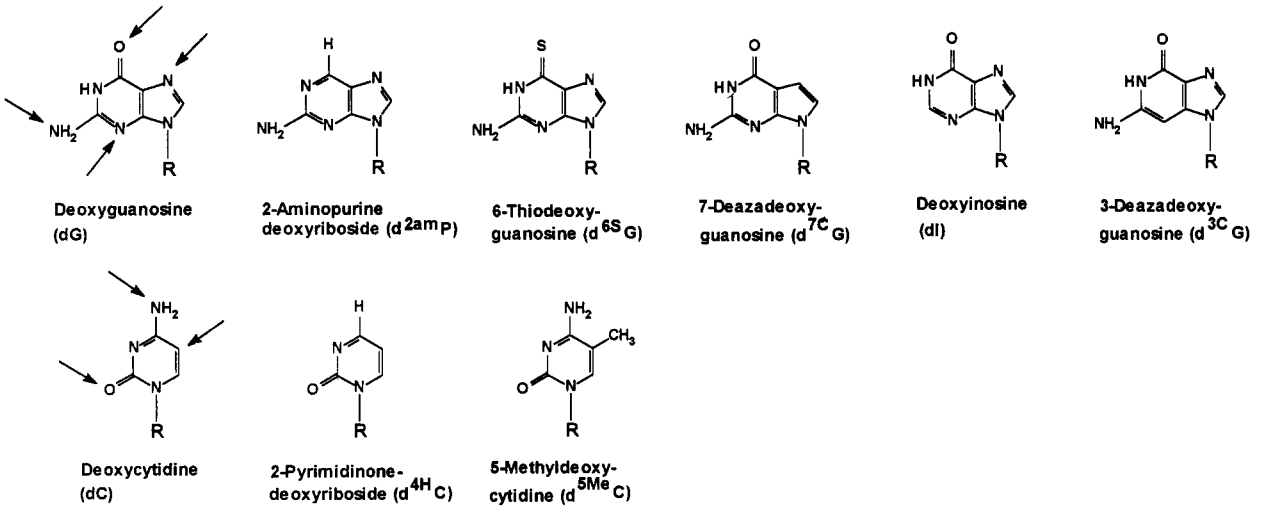

FIGURE 1: dG and dC analogues used in this study. Each analogue is derived from the parent base and has one of the potential protein contacting sites (indicated by arrows for the parents) altered. R = 1-β-D-(2′-deoxyribosyl).

Table 2: Properties of the Oligodeoxynucleotides Used in This Study[a]

oligodeoxynucleotide	base composition	T_m (°C)	hyperchromicity	extinction coefficient 254 nm (M^{-1} cm^{-1})	substrate
d(GACGATATCGTC)	$A_{3.0}C_{3.0}G_{3.2}T_{2.9}$	53	1.48	1.66×10^5	yes
d(GAC[6SG]ATATCGTC)	$A_{2.7}C_{3.1}G_{2.0}T_{2.9}[^{6S}G]_{1.2}$	43	1.39	1.70×10^5	yes
d(GAC[2amP]ATATCGTC)	$A_{2.8}C_{3.3}G_{2.1}T_{3.0}[^{2am}P]_{0.9}$		1.19	1.88×10^5	no
d(GAC[7CG]ATATCGTC)	$A_{3.0}C_{3.0}G_{1.8}T_{3.1}[^{7C}G]_{0.9}$	52	1.41	1.73×10^5	no
d(GAC[3CG]ATATCGTC)	$A_{2.9}C_{3.0}G_{2.0}T_{3.1}[^{3C}G]_{1.0}$	47	1.42	1.64×10^5	yes
d(GACIATATCGTC)	$A_{2.9}C_{3.1}G_{2.0}T_{3.0}I_{1.0}$	46	1.45	1.64×10^5	yes
d(GACGATA[4HC]GTC)	$A_{2.9}C_{2.2}G_{3.2}T_{2.9}[^{4H}C]_{0.7}$	33	1.34	1.75×10^5	yes
d(GACGATA[5MeC]GTC)	$A_{2.8}C_{2.2}G_{3.0}T_{3.0}[^{5Me}C]_{0.9}$	54	1.42	1.71×10^5	yes
d(TGACGATATCGTC)	$A_{2.9}C_{3.2}G_{3.0}T_{4.0}$	not determined	1.41	1.83×10^5	yes
d(GTCGACGATATCGTCGAC)	$A_{4.2}C_{4.8}G_{5.0}T_{4.0}$	56 and 72	1.55	2.36×10^5	yes
d(GTCGAC[2amP]ATATCGTCGAC)	$A_{3.9}C_{4.9}G_{4.2}T_{4.1}[^{2am}P]_{1.0}$	44 and 70	1.40	2.46×10^5	no

[a] No double- to single-stranded transition was observed for d(GAC[2amP]ATATCGTC). The extinction coefficients are given per mole of double-stranded oligodeoxynucleotide.

but with some of the poorly cleaved oligomers this was not possible. Stock solutions of the endonuclease were diluted using 50 mM Hepes, pH 7.5, 25 mM, $MgCl_2$, 55 mM NaCl, and 1 mg mL^{-1} bovine serum albumin immediately prior to the experiment and kept on ice. Omission of the albumin from these dilute solutions of the endonuclease resulted in its slow inactivation. Controls established that the albumin did not hydrolyze oligodeoxynucleotides. Values of K_m and k_{cat} were obtained from plots of substrate concentration/velocity versus substrate concentration (Fersht, 1985).

RESULTS AND DISCUSSION

The dG and dC analogues we have used to probe the interaction of the *Eco*RV restriction endonuclease with its recognition sequence are illustrated in Figure 1. The locations on these bases that have the potential to interact with proteins (Seeman et al., 1976) are indicated by arrows. These groups are mostly hydrogen bond donors of acceptors and so have the ability to form hydrogen bonds with DNA-binding proteins. Contacts between the groups indicated by arrows in Figure 1 (and also analogous functions in the dA-T base pair) and a variety of repressor proteins (Wolberger et al., 1988; Aggarwal et al., 1988; Jordan & Pabo, 1988; Brennan et al., 1990; Schulz et al., 1991) and restriction endonucleases (Rosenberg, 1991; Winkler, 1992; Winkler et al., 1993) have been observed by X-ray crystallography. Interactions of this sort have been termed direct readout (Matthews, 1988; Brennan & Matthews, 1989) and form an important component of specificity in protein–DNA interactions. It is possible to probe direct readout using base analogues in which potential protein contacting groups are deleted or altered. For dG the exocyclic 6-keto oxygen and 2-amino groups have been replaced with a hydrogen atom giving 2-aminopurine (1-β-D)-2′-deoxyriboside ($d^{2am}P$) and deoxyinosine (dI), respectively. Similarly, the 3 and 7 ring nitrogen atoms of dG have been substituted with a CH group to give 3- and 7-deazadeoxyguanosine ($d^{3C}G$ and $d^{7C}G$). In a slight variation to simple deletion of these groups, we have also prepared 6-thiodeoxyguanosine ($d^{6S}G$). Here a keto oxygen is replaced with a thioketo sulfur. With dC the 4-amino group has been replaced by hydrogen giving 2-pyrimidinone-(1-β-D)-2′-deoxyriboside ($d^{4H}C$). Additionally the 5′ region of dC has been studied by the introduction of a CH_3 group to form 5-methyldeoxycytidine ($d^{5Me}C$). With thymidine a methyl group occupies this position, and it is often important in protein–DNA interactions (Ivarie, 1987).

The analogues shown in Figure 1 were incorporated into the self-complementary dodecamer d(GACGATATCGTC) at the dG and dC positions within the GATATC *Eco*RV recognition site. After appropriate HPLC purification all the oligodeoxynucleotides (listed in Table 2) were ≥97% pure by reverse phase HPLC (date not shown). Base composition analysis gave the expected ratios of both the four standard bases and the modified base (Table 2). The hyperchromicity values and the extinction coefficients are also listed in Table 2. The *Eco*RV restriction endonuclease recognizes double-

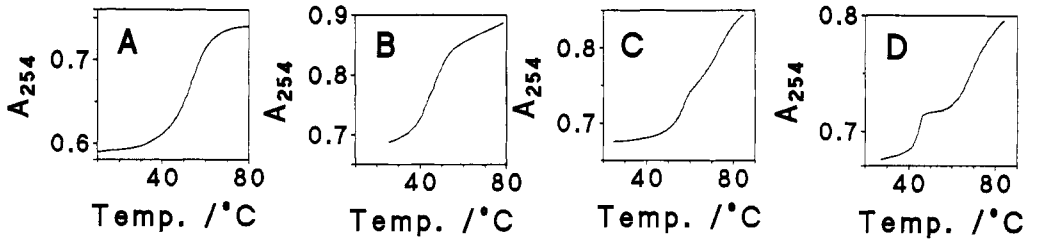

FIGURE 2: Melting curves (recorded in 50 mM Hepes, pH 7.5, 10 mM $MgCl_2$, and 100 mM NaCl) of the oligodeoxynucleotides (4–6 μM): (A) d(GACGATATCGTC); (B) d(GAC[^{3}C]GATATCGTC); (C) d(GTCGACGATATCGTCGAC); (D) d(GTCGAC[^{2am}P]ATATCGTCGAC). In cases where it was difficult to determine start and end points, the T_m values were obtained from differential plots. The T_m values seen for all the oligodeoxynucleotides (obtained from similar plots, not shown) used are summarized in Table 2.

stranded DNA, and so it is important to show that the dodecamers used are predominantly double stranded at the assay temperature (25 °C). Figure 2 shows, as examples, the melting curves obtained for d(GACGATATCGTC) and d(GAC[^{3}C]GATATCGTC), and all the T_m values are summarized in Table 2. The parent 12′-mer has a T_m of 53 °C (Connolly & Newman, 1989; Newman et al., 1990a). Oligonucleotides containing d^{7}CG and d^{5Me}C (which do not disrupt Watson–Crick hydrogen bonds) had virtually identical T_ms. The analogue d^{3}CG does not directly interfere with Watson–Crick base pairs but gives an oligodeoxynucleotide with a 6 °C lowered T_m perhaps caused by alterations to base stacking. The analogues dI and d^{4}HC both result in the loss of a Watson–Crick hydrogen bond to their partner bases on the complementary strand, and d^{6}SG also interferes with Watson–Crick base pairing. Oligodeoxynucleotides containing these bases all showed lowered T_m values. Reduction in melting temperatures for oligodeoxynucleotides containing dI and d^{4}HC have been observed previously (Ono & Ueda, 1987; Gildea & McLaughlin, 1989). The base d^{2am}P deletes two Watson–Crick hydrogen bonds, and the 12-mer containing this base did not undergo a melting transition on heating and so probably exists in the single-stranded form at 25 °C. Similar destabilization with d^{2am}P has been previously seen (McLaughlin et al., 1987). To overcome this difficulty, two 18-mers, d(GTCGACGATATCGTCGAC) and d(GTCGAC[^{2am}P]-ATATCGTCGAC), have been prepared. As shown in Figure 2 these showed biphasic melting curves. This was very pronounced for the d^{2am}P containing oligonucleotide but also visible for the control. It was possible to obtain two T_m values from these curves, which are given in Table 2. This biphasic behavior may result from the initial melting of either the ATAT region (in the control) or the [d^{2am}P]ATATC region (for the analogue substituted oligomer) followed by the melting of the remainder of the duplex. These regions would be expected to melt first as they contain relatively few Watson–Crick hydrogen bonds. The lower transitions of 44 and 56 °C found suggest, however, that both species will be predominantly double stranded at the 25 °C used for the assay. The circular dichroism spectra of the parent 12-mer and those containing d^{6}SG, d^{4}HC, and dI are shown in Figure 3. All the spectra are reminiscent of B-DNA and have roughly equally sized positive and negative peaks centered at 280 and 250 nm, respectively, and a crossover point at around 265 nm (Ivanov et al., 1973; Fairall et al., 1989; Gray et al., 1992). With d^{4}HC and d^{6}SG which have UV maxima well removed from the normal bases (d^{4}HC, 305 nm; d^{6}SG, 345 nm), additional circular dichroism transitions are visible at around or above 300 nm. With the dI containing 12-mer an extra positive peak is observed at 250–260 nm. This has previously been seen with other dI containing species (Ono & Ueda, 1987) and may arise as the UV absorbance maximum of dI occurs at 248 nm. This is somewhat lower than the four normal

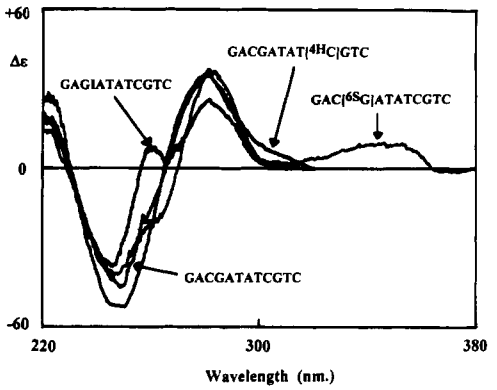

FIGURE 3: Circular dichroism spectra (recorded in 25 mM KH_2PO_4, pH 7.2, 10 mM $MgCl_2$, and 100 mM NaCl) of some of the oligodeoxynucleotides (at concentrations of 5 μM) used in this study. All of the other oligodeoxynucleotides had spectra very similar to that of the parent 12-mer (not shown) (5 μM).

bases and may cause it to show up as a separate peak in the circular dichroism spectrum. The circular dichroism spectra of all the other oligodeoxynucleotides were very similar (not shown) to the parent 12-mer. It must be emphasized that both the T_m and circular dichroism methods are low-resolution methods. Although these approaches give good evidence that the modified base containing oligodeoxynucleotides (providing an 18-mer is used for d^{2am}P) exist as double-stranded B-DNA structures, they do not reveal any subtle changes to the DNA conformation that may be caused by the presence of the analogue.

When the oligodeoxynucleotides given in Table 2 were incubated with large amounts of the *Eco*RV endonuclease, most of them were found to be substrates (data not shown). The only oligonucleotides that could not be hydrolyzed were the 12-mers containing d^{2am}P and d^{7}CG and the 18-mer containing d^{2am}P. The lack of cleavage of d(GAC[^{2am}P]-ATATCGTC) is probably due to it being single stranded, but this criticism cannot be applied to the 18-mer that contains this base. A previous study of the *Eco*RV endonuclease using a self-complementary 10-mer containing d^{2am}P also gave no reaction (Mazzarelli et al., 1989). However this 10-mer, like our 12-mer, appeared to be single stranded. There have been no previous studies with the endonuclease using d^{7}CG. On the basis of the sensitivity of the HPLC assay, we estimate that oligonucleotides containing the above two base analogues have maximal turnover numbers of 2×10^{-5} min^{-1}. This is more than 1 000 000-fold less than that of the control oligonucleotide and corresponds to the rate reductions of about 10^6 observed using plasmid substrates containing noncognate sites (Taylor & Halford, 1989). All the other oligonucleotides given in Table 2 were substrates and could be completely hydrolyzed to two products (data not shown). In all cases base composition analysis of the products showed cutting at

Table 3: Kinetic Parameters of the Oligodeoxynucleotides with the *Eco*RV Endonuclease

oligodeoxynucleotide	k_{cat} (min^{-1})	K_m (μM)	k_{cat}/K_m ($s^{-1}\ M^{-1}$)	k_{cat}/K_m (%)
d(pGACGATATCGTC)	33	2.9	1.9×10^5	100
d(pGAC[^{6S}G]ATATCGTC)[a]	poor substrate; turns over at 0.05% of control			
d(pGAC[^{2am}P]ATATCGTC)	0			0
d(pGAC[^{7C}G]ATATCGTC)	0			0
d(pGAC[^{3C}G]ATATCGTC)	0.28	3.5	1.3×10^3	0.7
d(pGACIATATCGTC)	0.63	1.8	5.8×10^3	3.1
d(pGACGATAT[^{4H}C]GTC)	38	2.0	3.2×10^5	168
d(pGACGATAT[^{5Me}C]GTC)	21	1.0	3.5×10^5	184
d(GTCGACGATATCGTCGAC)	132[b]			
d(GTCGAC[^{2am}P]ATATCGTCGAC)	0[b]			

[a] The oligonucleotide that contained d[^{6S}G] turned over too slowly to allow K_m and k_{cat} evaluation. [b] Turnover number determined at a single concentration of 5 μM.

the correct position, i.e., between the central T and dA bases in the GATATC sequence (data not shown).

Previously (Newman et al., 1990a), we have determined the K_m and k_{cat} values for the turnover of oligonucleotides containing dA and T analogues using ^{32}P labeling and separation of substrates and products by gel electrophoresis. More recently we developed a continuous spectrophotometric assay based on the increase in absorbance at 260 nm following the hydrolysis of double-stranded substrates to single-stranded products (Waters & Connolly, 1992). Using this assay, d(pGACGATATCGTC) gave a K_m of 2.9 μM (Table 3), in good agreement with the value of 3.8 μM seen previously (Newman et al., 1990b). One difference between the two studies is in the turnover number of the endonuclease. Previously values of 6.9 min^{-1} were seen, but currently we obtain values of 33 min^{-1}, a 5-fold increase. This difference has been traced to the final concentration step of the endonuclease purification. Earlier the protein was concentrated by $(NH_4)SO_4$ precipitation, but at present membrane filtration is being used. The latter method consistently gives more active enzyme. However, in both studies the hydrolysis of analogue containing oligonucleotides was compared to that of the parent using identically prepared enzyme. Thus the differences in turnover numbers between the two studies does not affect the conclusions reached in either. The kinetic parameters for the oligodeoxynucleotides are summarized in Table 3. All substrates showed Michaelis–Menten kinetics over the concentration range tested, and linear plots of [substrate]/velocity versus [substrate] were obtained (data not shown).

The oligodeoxynucleotide of d(GAC[^{6S}G]ATATCGTC) was a substrate, but it was hydrolyzed extremely slowly. We estimate (on the basis of comparisons with the control at a single concentration of 20 μM) that it is turned over at about 0.05% of the rate of the parent. This slow rate made K_m and k_{cat} determination impossible. This result reinforces that seen with the 18-mer containing d^{2am}P and confirms that the 6-keto oxygen atom of the dG is extremely important for catalysis. The findings seen with d^{6S}G, d^{2am}P, and d^{7C}G, all of which cause alterations to the major groove determinants of dG, clearly demonstrate the importance of this region for the correct recognition and hydrolysis of DNA by the enzyme. The contacts made by the *Eco*RV endonuclease to this dG, as determined by X-ray crystallography, are illustrated in Figure 3 (Winkler, 1992; Winkler et al., 1993). Hydrogen bonds are formed between the 6-keto oxygen and the 7-ring nitrogen of dG and the main chain nitrogen atoms of Asn 185 and Gly 184, respectively. Removing these 7-ring nitrogen or 6-keto oxygen atoms will obviously lead to the loss of these interactions, and this is almost certainly the cause of the poor substrate properties. Deletion of enzyme–substrate hydrogen bonds that do not involve charged groups or introduce unfavorable steric interactions (as occurs when d^{2am}P and d^{7C}G are used) has been found to result in an energy penalty of 2–7.5 kJ mol^{-1} in very many cases (Fersht, 1987a,b, 1988; Wells & Fersht, 1986). With the *Eco*RI endonuclease the removal of enzyme–DNA contacts with base analogues typically costs 5–7 kJ mol^{-1} (Lesser et al., 1993). The symmetrical modified base containing oligonucleotides differ in two positions (one on each strand) as compared to the parent, and so one might expect energetic penalties of 4–15 kJ mol^{-1}. The energetic consequences of base analogues can be evaluated using the expression $\Delta\Delta G_{app} = -RT\ln[(k_{cat}/K_m)_{analogue}/(k_{cat}/K_m)_{parent}]$. This predicts k_{cat}/K_m values for the modified base containing oligonucleotides of between 20% and 0.25% of those seen with the control. Clearly, far lower values are obtained with d^{2am}P and d^{7C}G, and so these bases give much worse substrates than would be expected on the basis of simple hydrogen bond loss. One possible explanation is a cooperative mechanism where the loss of a single protein–DNA interaction is accompanied by a rearrangement of the DNA-binding apparatus and a consequent loss of further interactions. This has previously been proposed for the *Eco*RV endonuclease (Winkler, 1992; Winkler et al., 1993) on the basis of the crystal structures. In particular, Winkler has argued that "the participation of main chain atoms, the small size of the recognition loop residues (Gly-Ser-Gly-Asn-Thr) and the presence of interactions within and between symmetry related recognition loops (of the two sub-units of the endonuclease) generates a highly co-operative set of interactions". A cooperative mechanism would be expected to increase the enzymes specificity.

The bases dI and d^{3C}G both alter the minor groove of dG. Oligonucleotides that contained these bases were substrates for the endonuclease. However the k_{cat} values were reduced to about 1% of the values observed with the parent. The K_m values were unaffected. Previously (Newman et al., 1990b) in a study with dA and T analogues it was observed that poor substrates always had lowered k_{cat} values and unchanged K_ms, and this occurs with the analogues used in this study. This agrees with the idea that the *Eco*RV endonuclease does not discriminate between DNA sequences at the substrate-binding stage but rather at the catalytic step. The k_{cat}/K_m for the dI containing 12-mer is 3% of that seen with the parent. Previous studies using dI have given a value of 11% (Fliess et al., 1988) or very little change (Mazzarelli et al., 1989). No previous studies have used d^{3C}G, which has a k_{cat}/K_m about 1% of that seen with the parent. These results suggest that the minor groove region of dG is important for DNA recognition by the endonuclease, but not as critical as the major groove. The

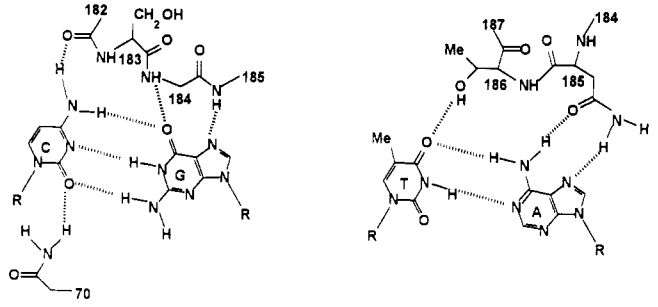

FIGURE 4: Contacts formed between the *Eco*RV endonuclease and the dG-dC or the outer dA-T base pairs in the GATATC recognition sequence as determined by X-ray crystallography. The amino acids are numbered and consist of Gly 182–Ser 183–Gly 184–Asn 185–Thr 186–Thr 187 and Asn 70. The 185 loop makes the hydrogen bonds illustrated. The 5-methyl group of the T base is also in a hydrophobic pocket. The Asn 70 shown hydrogen bonding to the O-2 of dC comes from the other subunit of the dimeric methylase as compared to residues 182–186. This side chain may additionally/alternatively hydrogen bond to the N-3 of the dA illustrated. No direct contacts are observed between the inner T-dA base pairs and the protein.

crystal structure indicates that the minor groove of the DNA is oriented toward the endonuclease and that most of the interface between the minor groove and the protein is accessible to solvent water molecules. No direct hydrogen bonds are seen between the 2-NH_2 or the 3-N function and the protein, but water-mediated hydrogen bonds are a possibility. The slow rates observed may reflect the positioning of the minor groove toward the protein rather than the solvent. In this orientation slight changes to the DNA conformation due to the modified base would be expected to disrupt the protein–DNA interface and produce changes in hydrolysis rates. Alternative explanations include disruption of water-mediated hydrogen bonds between the protein and the minor groove determinants of dG or, given the sensitivity of the major groove area of this dG, small alterations to the overall conformation that result in the movement of these major groove functions.

The bases $d^{5Me}C$ and $d^{4H}C$ probe the major groove of dC. These analogues gave excellent substrates with kinetic properties very similar to the parent. Previously, it has been observed that $d^{5Me}C$ or the isosteric 5-bromodeoxycytidine gave good substrates for the endonuclease (Fliess et al., 1988). However, the role of the 4-amino group of dC has not been probed using analogues. These results suggest that the major groove region of dC is unimportant for recognition and catalysis by the enzyme. However the X-ray structure (Figure 4) shows a hydrogen bond between the 4-amino group of dC and the main chain carbonyl oxygen of Gly 182. As discussed above, one would expect deletion of this hydrogen bond to give a substrate that turned over at between 0.25% and 20% of the rate seen with the parent, and at present it is not clear why this does not happen. This is the only occurrence with the *Eco*RV endonuclease where the crystal structure indicates an enzyme–DNA contact but using the appropriate base analogue produces a good substrate. A possible explanation is a compensation mechanism where loss of the direct contact is balanced by the gain of some other favorable interaction. With the *Eco*RI endonuclease (recognition sequence GAATTC) using purine (dP) in place of the first dA produces a better substrate (Lesser et al., 1993) despite the 6-NH_2 group of this dA being involved in hydrogen bonding with the endonuclease (Rosenberg, 1991). In this case the DNA is distorted on binding, and it was proposed that the loss of the enzyme–substrate hydrogen bond was balanced by facilitation of the DNA deformation. Almost identical effects were observed with the trp repressor when dP was used, and similar arguments were put forward in explanation (Mazzarelli et al., 1992). A similar mechanism could be operating with *Eco*RV. Substantial DNA distortion is seen on binding, and this may be made easier by the removal of 4-NH_2 from dC and the consequent loss of a Watson–Crick hydrogen bond to its partner dG. The 5-position of dC (usually occupied with hydrogen) tolerates the presence of a methyl group, and the structural data indicate that extra bulk can be tolerated in this region. Thus the 100% turnover rates of oligodeoxynucleotides containing $d^{5Me}C$ are easily rationalized.

CONCLUSION

This study probes the interaction of the *Eco*RV restriction endonuclease with the dG and dC bases in its recognition sequence, GATATC, using base analogues and completes our previous studies with dA and T analogues. (Newman et al., 1990a,b; Cosstick et al., 1990). Although the use of base analogues can be very powerful, the method does have its limitations. With most of the analogues a potential protein-contacting function has been deleted in order to probe the effect of this particular part of the base on protein–DNA interactions. The perfect situation occurs when the base analogue only results in a change to the site under study and does not cause further conformational changes to the remainder of the DNA molecule or alter base stacking interactions. It is extremely difficult to fully ensure that this actually takes place as we (Newman et al., 1990a,b) and others (Brennan et al., 1986a,b; Aiken & Gumport, 1991) have pointed out. Small changes to these parameters may occur, and this could lead to modified oligonucleotides being poor substrates for reasons other than loss of direct enzyme substrate contacts.

Nevertheless, modified bases have often been used, in the absence of other structural information, to try to identify protein–DNA contacts. With the availability of crystal structures for *Eco*RV the reliability of this approach can be assessed. The dG and dC bases in the GATATC recognition sequence have been discussed earlier. Figure 4 shows the contacts made by the endonuclease to the outer dA-T base pairs. In the major groove a bidentate hydrogen bond is made between the amide side chain of Asn 185 and the 6-NH_2 and 7-N groups of dA. The hydroxyl side chain of Thr 186 hydrogen bonds to the 4-O of T, and the 5-methyl group of T is located in a loose hydrophobic pocket. The N3 of this dA may be hydrogen bonded to the side chain of Asn 70. Good agreement is seen between the base analogue results that we reported previously (Newman et al., 1990a,b) and the crystallographic data for this base pair. Alterations to any of the major groove determinants gave extremely poor or nonsubstrates as did changing the N3 of dA. We explained this in terms of loss of enzyme–substrate contacts that were subsequently seen in the crystal structures. However, discrepancies were seen between the modified base approach and the crystallographic data with the central T and dA bases. The 5-methyl group of this T is in a hydrophobic pocket, and its removal gave a poor substrate presumably caused by the unfavorable positioning of an H_2O molecule in the hydrophobic region. There are no direct hydrogen bonds between these two bases and the endonuclease. Conversion of the 7-ring nitrogen of dA to a CH group gave a good substrate as would be expected from the lack of contact. However, deletion of the 6-NH_2 group and 4-keto oxygens of dA and T gave extremely poor substrates. Before the availability of the crystal structure, we interpreted the results as indicating hydrogen bonds between these locations on the DNA and the endo-

nuclease, but this conclusion appears to be incorrect. Thus with the *Eco*RV endonuclease many of enzyme–DNA contacts were picked up using base analogues. In some cases a modified base gave a low rate, and this was erroneously assigned to a contact. In one instance ($d^{4H}C$) the normal substrate properties would have most simply been used to, incorrectly, indicate no enzyme–substrate contact had a crystal structure not been available. The *Eco*RI endonuclease has also been evaluated using modified bases (Brennan et al., 1986a; Fliess et al., 1986, 1988; Seela & Kehne, 1987; McLaughlin et al., 1987; Lesser et al., 1990, 1993) and had its structure determined (Rosenberg, 1991). With this protein the agreement between the two methods in the major groove is good. This may be due to the *Eco*RI endonuclease contacting all the possible major groove sites in its GAATTC recognition sequence. However, some alterations to the minor groove gave poor substrates even though there are no direct contacts and this region faces the solvent. All in all, the results seen with these two enzymes confirm the view expressed in a review of modified bases (Aiken & Gumport, 1991) that the method can give an indication of interactions between a protein and its DNA substrate but in the absence of other information cannot be taken as conclusive.

What does the information obtained with base analogues when considered in conjunction with crystallographic and other data tell us about DNA recognition by the *Eco*RV endonuclease? It must be realized that all of our studies have used steady-state kinetic methods and as such they measure the slowest step in the catalytic cycle. At present the rate-determining step with short oligodeoxynucleotides and the *Eco*RV endonuclease is not known. With large plasmids the rate-limiting step is product release, and k_{cat} values of 0.9 min^{-1} are observed (Halford & Goodall, 1988). A much faster k_{cat} of 33 min^{-1} is seen with d(pGACGATATCGTC). This increase in rate may be due to either faster product release with the oligonucleotide as compared to the plasmid or to a change in the rate-determining step. The use of steady state data limits the interpretation that can be made to the data obtained in this publication, and this should be borne in mind. All the modified oligonucleotides that are poor substrates have a lowered k_{cat} with very little change in K_m, and this is consistent with the endonuclease binding to all sequences with equal affinity (Taylor & Halford, 1989; Taylor et al., 1991) and achieving discrimination after substrate binding, presumably at the transition state. An identical pattern has been observed with *Taq*I endonuclease using modified bases, and this enzyme also binds to all DNA sequences with similar affinities (Zebala et al., 1992a,b). This behavior contrasts with that of *Eco*RI (Brennan et al., 1986a; Fliess et al., 1986; Seela & Kehne, 1987; McLaughlin et al., 1987; Lesser et al., 1990, 1993) and *Rsr*I (Aiken et al., 1991) where analogues cause reductions in both K_m and k_{cat} and specific complexes are formed between the enzymes and their cognate sequences. It has been proposed that restriction endonucleases can be divided into two categories (Zebala et al., 1992b), those which achieve discrimination at both the binding and catalytic stages (*Eco*RI, *Rsr*I, and *Bam*HI) and those generating selectivity solely at the cleavage step (*Eco*RV and *Taq*I).

With the *Eco*RV endonuclease many of the modified bases give k_{cat}/K_m ratios that are very much poorer than would be expected for the simple deletion of an enzyme–substrate contact. This could be due to a cooperative mechanism as suggested by the crystal structure. The very tight interface between the protein and the DNA involving both protein–DNA and protein–protein interactions suggests that the loss of a single contact would result in a major structural reorganization and the subsequent loss of several more interactions. All the contacts (except the 4-amino group of dC) seen by crystallography are absolutely essential for catalysis, and deletion of any one of them gives a very poor or nonsubstrate. This suggests that all the interactions much be formed in order to drive the distortion of the DNA, assemble the active site, and give hydrolysis. Perhaps the most fascinating aspect of the *Eco*RV endonuclease is its exquisite recognition of the central T and dA bases without any direct contacts. Not only do base analogues (which are structurally very similar to the parents) cause large rate reductions but so also do substitutions by the normal bases (J. Alves and A. Pingoud, personal communication). Our view is that this is not due to the free conformation of the oligonucleotide containing a central TpdA step being vastly different from other sequences, composed of either modified or normal bases. Rather, only the TpdA sequence is able to adopt and support the distorted bound conformation. This is an example of indirect readout in which the bound conformation of this TpdA step becomes part of the recognition process (Otwinowski et al., 1988; Matthews, 1988; Brennan & Matthews, 1989; Luisi & Sigler, 1990). With the *Eco*RV endonuclease, DNA distortion is centered on the middle T and dA bases which are unstacked and form poor Watson–Crick base pairs. This distortion brings the 4-keto oxygen and 6-amino groups of the two adjacent T and dA bases into close proximity, and this may be favored by the partial negative charge on the oxygen and the partial positive charge on the amino groups. This could explain why alterations to these functions (using either base analogues or the natural bases) give poor substrates. An analogous indirect recognition of T-dA base pairs has been observed with the phage 434 cro protein (Wolberger et al., 1988) and the met J repressor (Phillips, 1991).

Currently we are studying the *Eco*RV endonuclease with modified bases using longer nonsymmetrical oligonucleotides and single-turnover kinetics. When combined with a crystal structure, these approaches can lead to a very thorough and quantitative understanding of sequence discrimination. This has been recently demonstrated for *Eco*RI (Lesser et al., 1990, 1993; Jen-Jacobson et al., 1991). The importance of these experiments arises because of the very different ways *Eco*RV and *Eco*RI recognize and cut their cognate sequences.

ACKNOWLEDGMENT

Thanks are due to Fritz Winkler, Alfred Pingoud, and Steven Halford for preprints and comments. The gift of 7-deazadeoxyguanosine derivatives from Philip Gottlieb is gratefully acknowledged. A special thanks goes to Margaret Davidson for typing the manuscript.

REFERENCES

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

Aiken, C. R., & Gumport, R. I. (1991) *Methods Enzymol. 208*, 433–457.

Aiken, C. R., McLaughlin, L. W., & Gumport, R. I. (1991) *J. Biol. Chem. 266*, 19070–19078.

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

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

Brennan, C. A., Van Cleve, M. D., & Gumport, R. I. (1986b) *J. Biol. Chem. 261*, 7278–7286.

Brennan, R. G., & Matthews, B. W. (1989) *J. Biol. Chem. 264*, 1903–1906.

Brennan, R. G., Roderick, S. L., Takeda, Y., & Matthews, B. M. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 8165–8169.
Connolly, B. A. (1991) in *Oligonucleotides and Analogues, A Practical Approach* (Eckstein, F., Ed.) pp 155–183, IRL Press, Oxford.
Connolly, B. A. (1992) *Methods Enzymol. 211*, 36–53.
Connolly, B. A., & Newman, P. C. (1989) *Nucleic Acids Res. 17*, 4957–4974.
Cosstick, R., Tuli, D. K., Williams, D. M., Connolly, B. A., & Newman, P. C. (1990) *Nucleic Acids Res. 18*, 4771–4778.
Dwyer-Hallquist, P., Kezdy, F. J., & Agarwal, K. L. (1982) *Biochemistry 21*, 4693–4700.
Fairall, L., Martin, S., & Rhodes, D. (1989) *EMBO J. 8*, 1809–1817.
Fersht, A. R. (1985) *Enzyme Structure and Mechanism*, 2nd ed., 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.
Fisher, E. F., & Caruthers, M. H. (1979) *Nucleic Acids Res. 7*, 401–416.
Fliess, A., Wolfes, H., Rosenthal, A., Schwellus, K., Blocker, H., Frank, R., & Pingoud, A. (1986) *Nucleic Acids Res. 14*, 3463–3474.
Fliess, A., Wolfes, H., Seela, F., & Pingoud, A. (1988) *Nucleic Acids Res. 16*, 11781–11793.
Gildea, B., & McLaughlin, L. W. (1989) *Nucleic Acids Res. 17*, 2261–2281.
Goeddel, D. V., Yansura, D. G., Winston, C., & Caruthers, M. H. (1978) *J. Mol. Biol. 123*, 661–687.
Gray, D. M., Ratliff, R. L., & Vaughan, M. R. (1992) *Methods Enzymol. 211A*, 389–400.
Halford, S. E., & Goodall, A. J. (1988) *Biochemistry 27*, 1771–1777.
Halford, S. E., Taylor, J. D., Vermote, C. L. M., & Vipond, I. B. (1993) in *Nucleic Acids and Molecular Biology* (Eckstein, F., & Lilley, D. M. J., Eds.) Vol. 7, pp 47–69, Springer-Verlag, Berlin.
Hayakawa, T., Ono, A., & Ueda, T. (1988) *Nucleic Acids Res. 16*, 4761–4776.
Ivanov, V. I., Michenkova, L. E., Schyolkina, A. K., & Poletayev, A. I. (1973) *Biopolymers 12*, 89–110.
Ivarie, R. (1987) *Nucleic Acids Res. 15*, 9975–9983.
Jen-Jacobson, L., Lesser, D. R., & Kurpiewski, M. R. (1991) in *Nucleic Acids and Molecular Biology* (Eckstein, F., & Lilley, D. M. J., Eds.) pp 141–170, Springer-Verlag, Berlin.
Jiriciny, J., Wood, S. G., Martin, D., & Ubasawa, A. (1986) *Nucleic Acids Res. 14*, 6579–6590.
Jordan, S. R., & Pabo, C. O. (1988) *Science 242*, 893–899.
Lesser, D. R., Kurpiewski, M. R., & Jen-Jacobson, L. (1990) *Science 250*, 776–786.
Lesser, D. R., Kurpiewski, M. R., Waters, T., Connolly, B. A., & Jen-Jacobson, L. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 7548–7552.
Luisi, B. F., & Sigler, P. B. (1990) *Biochim. Biophys. Acta 1048*, 113–126.
Matthews, B. W. (1988) *Nature 335*, 294–295.
Mazzarelli, J. M., Scholtissek, S., & McLaughlin, L. W. (1989) *Biochemistry 28*, 4616–4622.
Mazzarelli, J. M., Rajur, S. B., Iadarola, P. L., & McLaughlin, L. W. (1992) *Biochemistry 31*, 5925–5936.
McLaughlin, L. W., Benseler, F., Graeser, E., Piel, N., & Scholtissek, S. (1987) *Biochemistry 26*, 7238–7245.
Newman, P. C., Nwosu, V. U., Williams, D. M., Cosstick, R., Seela, F., & Connolly, B. A. (1990a) *Biochemistry 29*, 9891–9901.
Newman, P. C., Williams, D. M., Cosstick, R., Seela, F., & Connolly, B. A. (1990b) *Biochemistry 29*, 9902–9910.
Ono, A., & Ueda, T. (1987) *Nucleic Acids Res. 15*, 3059–3072.
Ono, A., Masayuki, S., Ohtani, Y., & Ueda, T. (1984) *Nucleic Acids Res. 12*, 8939–8949.
Otwinowski, Z., Schevitz, R. W., Zhang, R. G., Lawson, C. L., Joachimiak, A., Marmorstein, R. Q., Luisi, B. F., & Sigler, P. B. (1988) *Nature 335*, 321–339.
Phillips, S. E. V. (1991) *Curr. Opin. Struct. Biol. 1*, 89–98.
Revankar, G. R., Gupta, P. K., Adams, A. D., Dalley, N. K., McKernan, P. A., Cook, P. D., Cononico, P. G., & Robins, R. K. (1984) *J. Med. Chem. 27*, 1389–1390.
Rosenberg, J. M. (1991) *Curr. Opin. Struct. Biol. 1*, 104–113.
Schulz, S. C., Shields, G. C., & Steitz, T. A. (1991) *Science 253*, 1001–1007.
Seela, F., & Driller, H. (1986) *Nucleic Acids Res. 14*, 2319–2332.
Seela, F., & Kehne, A. (1987) *Biochemistry 26*, 2232–2238.
Seela, F., & Lampe, S. (1991) *Helv. Chim. Acta 74*, 1790–1800.
Seeman, N. C., Rosenberg, J. M., & Rich, A. (1976) *Proc. Natl.* Acad. *Sci. U.S.A. 73*, 804–808.
Taylor, J. D., & Halford, S. E. (1989) *Biochemistry 28*, 6198–6207.
Taylor, J. D., & Halford, S. E. (1992) *Biochemistry 31*, 90–97.
Taylor, J. D., Badcoe, I. G., Clarke, A. R., & Halford, S. E. (1991) *Biochemistry 30*, 8743–8753.
Taylor, C., Ford, K., Connolly, B. A., & Hornby, D. P. (1993) *Biochem. J. 291*, 493–504.
Vipond, I. B., & Halford, S. E. (1993) *Mol. Microbiol. 9*, 225–231.
Waters, T. R., & Connolly, B. A. (1992a) *Nucleosides Nucleotides 11*, 1561–1574.
Waters, T. R., & Connolly, B. A. (1992b) *Anal. Biochem. 204*, 985–998.
Wells, T. N. C., & Fersht, A. R. (1986) *Biochemistry 25*, 1881–1886.
Winkler, F. K. (1992) *Curr. Opin. Struct. Biol. 2*, 93–99.
Winkler, F. K., Banner, D. W., Oefner, C., Lahm, A., Tsernoglu, D., Brown, R. S., Heathman, S. P., Bryan, R. K., Martin, P. D., Petratos, K., & Wilson, K. S. (1993) *EMBO J. 12*, 1781–1795.
Wolberger, C., Dong, Y., Ptashne, M., & Harrison, S. C. (1988) *Nature 335*, 789–795.
Yansura, D. G., Goeddel, D. V., Cribbs, D. L., & Caruthers, M. H. (1979) *J. Mol. Biol. 133*, 117–135.
Yolov, A. A., Vinogradova, M. N., Gromova, E. S., Rosenthal, A., Cech, D., Veiko, V. P., Metelev, V. G., Kosykyh, V. G., Buryanov, Ya. I., Bayev, A. A., & Shaborova, Z. A. (1985) *Nucleic Acids Res. 13*, 8983–8998.
Zebala, J. A., Choi, J., & Barany, F. (1992a) *J. Biol. Chem. 267*, 8097–8105.
Zebala, J. A., Choi, J., Trainor, G. L., & Barany, F. (1992b) *J. Biol. Chem. 267*, 8106–8116.