STUDIES ON SEQUENCE RECOGNITION BY TYPE II RESTRICTION AND MODIFICATION ENZYMES

Author: **Paul Modrich**
Department of Biochemistry
Duke University Medical Center
Durham, North Carolina

Referee: Richard J. Roberts
Cold Spring Harbor Laboratory
Cold Spring Harbor, New York

I. INTRODUCTION

DNA restriction endonucleases and modification methylases are strain-specific enzymes responsible for the host-specific barriers to interstrain transfer of DNA that have been identified in numerous prokaryotic cell types. Foreign DNA entering a bacterial cell is subject to rapid endonucleolytic hydrolysis if it is devoid of the modification characteristic of the particular bacterial strain. The strain specific modification enzyme catalyzes methyl transfer from S-adenosyl-L-methionine (AdoMet) to a specific DNA sequence which is characteristic of the particular host specificity system. Thus, cellular DNA is rendered resistant to attack by the endogenous restriction enzyme by virtue of being methylated at DNA sites that are also recognized by the endonuclease.

Subsequent to the isolation in 1968 of restriction endonucleases responsible for host specificities of *Escherichia coli* strains K and B,[1,2] over 200 restriction activities have been identified upon in vitro analysis of gram negative and gram positive organisms.[3,4] Elucidation of enzyme structural features as well as cofactor requirements for restriction cleavage indicated that these enzyme systems could be classified into three types.[5,6] The class designated Type I is typified by the *Eco* B and *Eco* K enzymes which are responsible for the host specificities of *E. coli* strains B and K, respectively.† In these enzyme systems both restriction and modification activities appear to reside in a large (M_r 500,000 to 600,000) multifunctional complex comprised of three distinct polypeptide chains, with the activity of the complex being directed by the state of modification of the host specificity site (reviews 6,8). In vitro restriction by Type I enzymes is dependent on the presence of ATP, AdoMet and a divalent cation, and although interaction with the particular host specificity sequence is necessary, DNA cleavage occurs thousands of nucleotides away. In contrast, methyl transfer by the Type I enzyme complex is to the recognition site and requires only AdoMet, with ATP and Mg^{++} being only stimulatory.

The more recently characterized Type III restriction and modification enzymes[6] appear to be somewhat simpler than their Type I counterparts. As in the case of Type I enzymes, Type III restriction and modification activities may both reside in a large multifunctional complex which is composed of two distinct polypeptide chains. Like the more complex systems, DNA cleavage by Type III enzymes requires ATP and Mg^{++}, but

† Restriction and modification systems are designated according to the nomenclature suggested by Smith and Nathans.[7] Each restriction-modification system is designated by a three letter prefix indicating genus and species of the organism from which the enzymes have been obtained. Enzyme systems of differing specificity obtained from the same organism are differentiated by roman numerals following the prefix. For a recent listing of known restriction and modification enzymes see Roberts.[4]

in contrast to Type I restriction, AdoMet is not required, and DNA cleavage occurs near Type III host specificity sites. However, like Type I enzymes, methyl transfer by Type III complexes requires an AdoMet cofactor, with ATP and Mg^{++} being stimulatory.

The simplest class of restriction and modification enzymes, and the subject of this review, has been designated Type II. Although numerous Type II systems of distinct sequence specificities[3,4] (Table 1) have been identified, all bear features in common with the *Hin*d II system initially identified by Smith and his colleagues.[9-12] Thus, all are comprised of proteins which are relatively small relative to their Type I or Type III counterparts, and cofactor requirements are simple. In vitro restriction requires only Mg^{++} and unmodified DNA, while methyl transfer is dependent only on the presence of AdoMet. Moreover, DNA cleavage by Type II endonucleases bears a simple relationship to the recognition sequence, and for those systems examined in detail, endonuclease and methyltransferase activities have been found to reside in distinct (and separable) polypeptide chains.

The relative structural and catalytic simplicity of Type II enzymes, as well as the opportunity they afford for comparison of distinct polypeptides which recognize a common sequence, has rendered them attractive for study of mechanisms of sequence-specific protein-DNA interaction. Although it is the intention of this article to review literature bearing on DNA site recognition by Type II enzymes, it must be stated at the outset that the molecular mechanisms responsible for highly accurate and efficient DNA sequence recognition are not yet understood in the case of any protein. Nevertheless, recent progress in the study of several types of sequence-specific systems, only one class of which will be considered in this review, suggests that cautious optimism might be appropriate with respect to future advances in this area.

Two problems, which to some extent are related, are apparent upon consideration of specific DNA sequence recognition by proteins. First, one wishes to understand factors which account for the extreme specificity characteristic of this type of interaction. A minimal understanding of this facet of the problem requires knowledge of thermodynamic parameters characteristic of interaction of the protein with specific and nonspecific DNA sequences, as well as information on the nature of protein-DNA contacts governing interaction of a protein with these two classes of sites. The second problem is due to the fact that recognition sites of biological significance are embedded in a large background of nonspecific DNA. Despite the occurrence of recognition sites in such an environment, sequence-specific proteins typically locate their sites with apparent rates which approach or apparently exceed that expected for diffusion-limited association of uncharged macromolecules ($k \simeq 10^8\ M^{-1}\ s^{-1}$) suggesting the involvement of electrostatic or limited dimension diffusion mechanisms in specific site location.[66-69] An understanding of the mechanisms of site location requires elucidation of pathways utilized in location of specific sequences as well as parameters governing kinetic selection of the preferred path. In view of these problems, the following discussion will consider available information on enzyme and substrate structures, reaction mechanisms, and factors affecting specificity which may bear on these points.

II. STRUCTURES OF TYPE II RESTRICTION AND MODIFICATION ENZYMES

Some 200 site-specific endonucleases have been classified as Type II restriction enzymes.[3,4] Although the utility of these endonucleases as reagents has rendered this practice useful, this classification is misleading in the sense that only a few of these enzymes have been implicated in restriction in vivo. Furthermore, enzymes within this class may differ significantly with respect to modes of DNA-protein interaction. Indeed,

Table 1
PROPERTIES OF SELECTED TYPE II DNA RESTRICTION AND MODIFICATION SYSTEMS

Microorganism	Specificity	Enzyme	M_r[a]	Recognition sequence[b]	Ref.
B. amyloliquefaciens H	*Bam* HI	R	$(22,000)_2$ $(22,000)_4$	5′-G↓G-A-T-C*-C 3′-C-C*-T-A-G↑G	13—16
		M	N.D.		
B. caldolyticus	*Bcl* I	R	$(25,000)_2$	5′-T↓G-A-T-C-A 3′-A-C-T-A-G↑T	17
B. globigii	*Bgl* I	R	$(31,000)_1$	5′-G-C-C-N-N-N-N↓N-G-G-C 3′-C-G-G-N↑N-N-N-N-C-C-G	18—21
B. globigii	*Bgl* II	R	$(27,000)_2$	5′-A↓G-A-T-C-T 3′-T-C-T-A-G↑A	22, 23
B. sphaericus	*Bsp* I	R	$(35,000)_1$	5′-G-G-C-C 3′-C-C-G-G	24, 25
		M	~50,000		
B. stearothermophilus	*Bst* I	R	$(26,000)_1$ $(26,000)_2$	5′-G↓G-A-T-C-C 3′-C-C-T-A-G↑G	26—28
		M	~400,000		
B. subtilis	*Bsu* I	R	$(68,000)_1$ $(68,000)_2$	5′-G-G↓C*-C 3′-C-C*↑G-G	29—32
		M	?		
E. coli	*dam*[c]	M	$(31,000)_1$	5′-G-A*-T-C 3′-C-T-A-G*	33, 34
D. pneumoniae	*Dpn* I[d]	R	~20,000	5′-G-A*↓T-C 3′-C-T↑A-G*	35—37
D. pneumoniae	*Dpn* II[d]	R	~70,000	5′-G-A-T-C 3′-C-T-A-G	35—37
E. coli RY13	*Eco* RI[e]	R	$(31,000)_2$ ↑↓ $(31,000)_4$	5′-G↓A*-A-T-T-C 3′-C-T-T-A-A*↑G	38—46
		M	$(38,000)_1$		
E. coli R245	*Eco* RII	R	$(40,000)_2$	5′↓C-C*-A-G-G 3′-G-G-T-C*-C↑	47—48
		M	N.D.		
H. aegyptius	*Hae* III	R	N.D.	5′-G-G↓C*-C 3′-C-C*↑G-G	30, 49, 50
		M	N.D.		
H. haemolyticus	*Hha* I	R	N.D.	5′-G-C*-G↓C 3′-C↑G-C-G*	51, 52
		M	N.D.		

Table 1 (continued)
PROPERTIES OF SELECTED TYPE II DNA RESTRICTION AND MODIFICATION SYSTEMS

Microorganism	Specificity	Enzyme	M_r[a]	Recognition sequence[b]	Ref.
H. haemolyticus	*Hha* II	R	$(24,000)_2$	5′-G-A-N-T-C 3′-C-T-N-A-G	53, 54
		M	N.D.		
H. influenzae Rd	*Hind* II	R	~70,000	5′-G-T-Py↓Pu*-A-C 3′-C-A*-Pu↑Py-T-G	9—12
		M	~55,000		
H. influenzae Rd	*Hind* III	R	N.D.	5′A*↓A-G-C-T-T 3′-T-T-C-G-A↑A*	11, 12, 55
		M	N.D.		
H. parahaemolyticus	*Hph* I	R	N.D.	5′-G-G-T-G-A-N-N-N-N-N-N-N-N↓ 3′-C-C-A-C-T-N-N-N-N-N-N-N↑	56
H. parainfluenzae	*Hpa* I	R	$(29,000)_2$	5′-G-T-T↓A*-A-C 3′-C-A*-A↑T-T-G	57—60
		M	$(38,000)_1$		
H. parainfluenzae	*Hpa* II[f]	R	$(41,000)_2$	5′-C↓C*-G-G 3′-G-G-C*↑C	50, 57—61
		M	(39,000) (42,000)		
Moraxella nonliquefaciens	*Mno* I	R	?	5′-C↓C-G-G 3′-G-G-C↑C	62
Moraxella sp.	*Msp* I	R	?	5′-C↓C-G-G 3′-G-G-C↑C	63, 64
Staphylococcus aureus 3A	*Sau* 3A	R	?	5′↓G-A-T-C 3′-C-T-A-G↑	65

[a] In those cases where the enzyme has been isolated in pure form, stable aggregation states of the monomer are indicated. Approximate native M_r values, determined by gel filtration or sucrose gradient sedimentation, are shown for those enzymes which have been only partially purified. N.D.—not determined.

[b] Arrows indicate points of cleavage by the restriction endonuclease while bases methylated by the modification enzyme are indicated by asterisks. In all cases cleavage generates 5′-phosphoryl and 3′-hydroxyl termini, and modification yields N-6 methyladenine or 5-methylcytosine. In those cases where the particular base modified has been identified, both strands of the recognition site are shown as being modified. However, methylation of both DNA strands has been established only for *Eco* RI,[39] *Eco* RII,[48] *dam*,[33] *Hpa* I,[560] and *Hpa* II,[61] methyltransferases. In several cases sites of cleavage or methylation within the recognition site have not yet been determined.

[c] As mentioned in the text, although the *E. coli dam* methylase resembles other Type II modification enzymes, it does not appear to function in restriction-modification.

[d] *Dpn* I and *Dpn* II both recognize the sequence d(G-A-T-C). However, *Dpn* I is novel in that it cleaves within the sequence *only* if the adenines on both DNA strands are modified. In contrast, *Dpn* II only attacks sites devoid of the adenine modification.[35-37]

[e] Although there has been a claim that isolated *Eco* RI endonuclease is a monomer,[46] evidence in support of this view was not provided. Moreover, this finding has not been substantiated by several other laboratories.[40,41]

[f] Two DNA methylases of *Hpa* II specificity, which differ in molecular weight, have been identified. Since preliminary peptide mapping suggests that the two species are related, the smaller may have been derived from the larger.[61]

such differences are already apparent among the few Type II enzymes to be considered here, those which have been isolated in pure form or examined with respect to features of mechanism (Table 1). For this reason, those differences which may be of significance in the context of the sequence recognition question will be emphasized in this and following sections.

As summarized in Table 1, each Type II endonuclease that has been isolated in homogeneous form exhibits a single molecular weight species upon denaturation and reduction. This suggests that each is comprised of a single type of polypeptide chain, and as discussed below, this has been confirmed in the case of *Eco* RI endonuclease. These enzymes fall into several classes based on oligomeric states of the native proteins. The majority of Type II endonucleases exist in solution as oligomers of their constituent subunit, and in the case of several enzymes, more than one aggregation state has been observed. There are, however, exceptions to this generality, the most notable being *Bgl* I and *Bsp* I endonucleases. In these cases, only the monomer has been identified.[18,25] Since it is reasonable to assume that modes of DNA-protein interaction to some extent reflect biologically active aggregation state(s) of the protein in question, it has been suggested that these two structural classes of Type II endonucleases may differ with respect to mechanism.[18] However, attempts to assign biological activity to particular oligomeric states have been reported only in the case of the *Eco* RI enzyme, which exists in solution as an equilibrium mixture of dimers and tetramers.[41,70] In this system, the tetramer to dimer transition is governed by an equilibrium dissociation constant of approximately 10^{-7} *M*, and it has been found that the dimer remains the stable form at catalytic concentrations (10^{-10} *M*). This finding together with kinetic behavior of this endonuclease (below) indicates that the dimer possesses biological activity and interacts directly with the *Eco* RI sequence.

Structural analysis of Type II modification enzymes is still at an early stage. Although a number of such activities have been identified[4] (Table 1), only four have been isolated in pure form: *Eco* RI,[42] *Hpa* I,[60] *Hpa* II,[61] and dam.[34] While the latter enzyme, which is responsible for the majority of N-6-methyladenine in *E. coli* DNA,[71] does not appear to be involved in restriction-modification, it will be considered here because of its structural and mechanistic similarities to the former three enzymes. Each of these methyltransferases is comprised of a single polypeptide chain (M_r of 30,000 to 40,000), and in all four cases the stable form in solution is the monomer. Moreover, and as will be discussed below, kinetic studies indicate that the functional form of all of these enzymes is in fact the monomer. Thus, in the *Eco* RI system, and perhaps in *Hpa* I and *Hpa* II systems as well, the restriction endonuclease interacts with the recognition sequence as a dimer while the modification enzyme is functional as a monomer.† This observation led to the suggestion that, at least in the case of the *Eco* RI system, the endonuclease and methylase may differ significantly with respect to protein-nucleic acid interactions involved in specific site recognition.[42] This possibility might appear surprising in view of the complementary nature of the *Eco* RI enzymes as components of a restriction-modification system. Indeed, it has been established that continued presence of an active *Eco* RI methylase is necessary to maintain cell viability in the presence of the endonuclease.[72] This functional complementarity of two proteins which recognize the same DNA sequence led Boyer et al.[73] to propose that the two *Eco* RI enzymes might have a common evolutionary origin. In this view the two proteins would share substantial

† Although *Hpa* I and *Hpa* II methylases have been shown to function as the monomer[60,61] and only dimeric forms of the corresponding endonucleases have been observed, the possibility of dissociation of restriction enzyme dimers at dilute concentrations has not been assessed.

homology within DNA binding site regions as well as common features of polynucleotide interactions.

The recent determination of primary structures of the *Eco* RI enzymes has permitted direct assessment of structural similarities between the two proteins at the amino acid sequence level. The genes for these enzymes have been localized to a 2200 base pair segment of DNA, the sequence of which has been determined in two laboratories.[43-45] These determinations, together with partial amino and carboxyl terminal sequence analysis of the isolated enzymes,[74] have permitted assignment of structural genes within this DNA segment, as well as evaluation of the extent to which each polypeptide is processed subsequent to biosynthesis. As summarized in Figure 1, the *Eco* RI endonuclease gene encodes a 277 residue polypeptide (M_r = 31,063), and processing appears limited to removal of the NH_2-terminal formylMet. The methylase gene specifies a 326 residue protein (M_r = 38,048), but in this case terminal formylMet-Ala is removed.

Comparison of endonuclease and methylase amino acid and gene sequences for homologies has led to identification of a single limited and imperfect DNA sequence homology which is unlikely on the basis of chance.[44] However, the general lack of homology between the two polypeptides, and their corresponding genes, has led to the conclusion that the two proteins are of different evolutionary origin.[44,45] While the absence of amino acid sequence homology between *Eco* RI enzymes does not exclude existence of common features at higher levels of structure, the circular dichroism spectra of the two proteins indicate extensive differences at the level of secondary structure,[44] and secondary structures predicted for the two proteins by the probabalistic method of Chou and Fasman[75] are markedly different.[44,45] Thus, the dissimilar structures of *Eco* RI restriction and modification enzymes at primary, secondary, and quaternary levels are consistent with the view that the two proteins may interact with DNA in different ways. Additional evidence in support of this idea will be considered below.

Since each *Eco* RI enzyme must discriminate two G·C and four A·T pairs, the possible occurrence of amino acid sequence homologies within each enzyme has also been examined.[44,45] The endonuclease contains a single homology which is statistically unlikely, the overlapping tandem repeat Ile-Met-Phe-Glu-Ile-Met-Phe-Asp-Ile. Similarly, the methylase contains a 4-fold repeat of Leu-Ile-Lys. This tripeptide repeat occurs at defined intervals with 21 residues separating both the first two and last two occurrences, and additional homologous residues are evident in regions adjacent to Leu-Ile-Lys sequences.[44] Although organization of these improbable repeats within *Eco* RI polypeptides suggests that they might be of importance in endonuclease or methylase structure and function, this possibility will remain a matter of speculation until the significance of such regions is assessed by chemical or genetic alteration.

It is noteworthy that availability of primary sequences of *Eco* RI enzymes should facilitate crystallographic analyses of these proteins. Although crystallization of *Eco* RI methylase has not been reported, crystals of the endonuclease suitable for high resolution structure analysis have been obtained. Rosenberg et al.[76] have obtained high quality crystals of the endonuclease and have reported results of preliminary X-ray diffraction analysis. The enzyme crystallizes in space group C2 with unit cell parameters of a = 209 Å, b = 129 Å, c = 50 Å and β = 98.4°. The crystals appear to contain four endonuclease monomers (M_r = 31,000) per asymmetric unit which is consistent with the proclivity of the enzyme to associate into tetramer at high protein concentrations.

An apparently identical crystalline form has been also obtained by Young, et al.[77] utilizing different crystallization conditions. In addition, this group has also successfully crystallized a complex between the endonuclease and the *Eco* RI recognition sequence hexanucleotide d(pGpApApTpTpC). The crystalline complex diffracts to a nominal resolution of 2.8 Å and is of space group $P42_12$ with $a = b$ = 183.2 Å, c = 49.7 Å, and $\alpha = \beta = \gamma = 90°$. The unit cell contains four endonuclease monomers plus two duplex

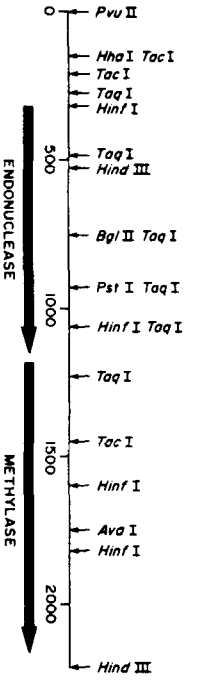

FIGURE 1. Organization of structural gene sequences for *Eco* RI restriction and modification enzymes.[44,45]

DNA fragments in an asymmetric unit. As discussed below, this stoichiometry of one endonuclease dimer per duplex recognition sequence is consistent with the mode of interaction of the enzyme with DNA in solution. It is also interesting to note that unit cells of both the enzyme and the complex crystals have a *c* axial length of 50 Å. Young et al.[77] have pointed out that this together with the space group symmetry suggests that the endonuclease tetramer and the complex tetramer have 50 Å as one of their dimensions.

It is apparent that the rapid acquisition of structural information on *Eco* RI enzymes will facilitate understanding of DNA-protein interactions in this system. In particular, elucidation of crystal structures of the endonuclease and the endonuclease-DNA complex should provide substantial insight into molecular contacts involved in DNA sequence recognition in this system, as well as protein conformational transitions that might occur during this process.

Lastly, it should be pointed out that although structural analyses of Type II restriction and modification enzymes has proceeded most rapidly in the case of the *Eco* RI system, detailed structural information on other systems is rapidly accumulating. Genes specifying the *Hha* II restriction and modification enzymes[78] and the *E. coli dam* methylase[34] have also been cloned, and DNA sequence determinations are in progress.[79,80] Consequently, the primary structures of these proteins will soon be available. It seems likely that similar information on other systems will be forthcoming as well. As is evident from inspection of only the partial list presented in Table 1, restriction-modification systems originating in different organisms may recognize sequences that are identical or which are substantially homologous. For example, more than 20 endonuclease have been identified which recognize GATC, or a sequence in which GATC is the internal tetranucleotide.[4] Thus, comparison of polypeptide sequences of such enzymes should yield insight into evolution of related systems and perhaps into features of polypeptide structure important in specific interactions.

III. RECOGNITION SEQUENCES OF TYPE II RESTRICTION AND MODIFICATION ENZYMES

The recognition sequences of numerous Type II systems have been determined, and these have been reviewed.[3,4] The specificities of selected systems are presented in Table 1. Type II recognition sites vary from four to six base pairs in length, corresponding to expected frequencies of occurrence in a random sequence of one per 256 base pairs to one per 4096 base pairs, respectively. With few exceptions, these sequences exhibit twofold rotational symmetry, and sites of DNA strand scission or methylation are symmetrically disposed relative to the twofold axis (Table 1). Although the significance of the frequent occurrence of rotational symmetry in DNA sequences recognized by proteins is not known, several possible functions have been entertained:

1. Perhaps the simplest possibility has been proposed by Kelly and Smith[10] who suggested that symmetry within the recognition sequence may correspond to symmetry within the protein. Thus, a protein comprised of one type of polypeptide chain could interact equivalently with both strands of the DNA helix.
2. A related hypothesis has been proposed by Dickson et al.[81] who noted that rotationally symmetric sequences could facilitate site recognition if the protein is capable of bidirectional diffusion along the DNA helix. As mentioned above, such unidimensional diffusion mechanisms have been suggested to play a role in the kinetic path by which specific DNA sequences are located by proteins.[66-69] As detailed below, such a mechanism appears to be operable in the case of *Eco* RI endonuclease.
3. It is possible that the conformation of DNA is to some extent sequence dependent (for reviews see References 82 and 83). A special case of this hypothesis is the suggestion that twofold symmetrical sequences may form special structures which would be easily recognizable along the DNA helix. Gierer[84] and Sobell[85] have noted that twofold symmetrical sequences are at least in principle capable of intrastrand base pairing to form cruciform structures protruding from the helix axis (Figure 2). Although such structures would in general be thermodynamically unstable due to the associated requirement for local helix disruption, their formation could be energetically driven by superhelical free energy[86,87] and once formed, they could be stabilized by interaction with a protein.

An alternate model for special structure formation by twofold symmetrical sequences, and one which does not require disruption of base pairs, has recently been suggested.[88,89] This model is based on the proposal by McGavin[90] of a hydrogen bonding scheme by which like Watson-Crick base pairs might interact and which could permit homologous duplex regions to assume a four-stranded configuration. Model building studies have led Strasiak and Klopotowski[89] to suggest that even short (four to six base pair) symmetrical recognition sequences may adopt a four-stranded "cage" structure by virtue of this secondary bonding scheme (Figure 2). A particularly interesting feature of the cage model is that methylation of the C^5 of cytosine or the N^6 of adenine (Table 1) would, in the case of many Type II recognition sequences, hinder tetrad formation as a consequence of steric effects or by precluding formation of the secondary hydrogen bonds.[89] This model would thus imply recognition of a cage structure by the endonuclease, with protection against restriction by methylation being due at least in part to disruption of the four-stranded structure. Unfortunately, available information does not permit evaluation of this model.

Although one or more of the above hypotheses may account for the frequent occurrence of rotational symmetry in DNA recognition sequences, it is clear that such

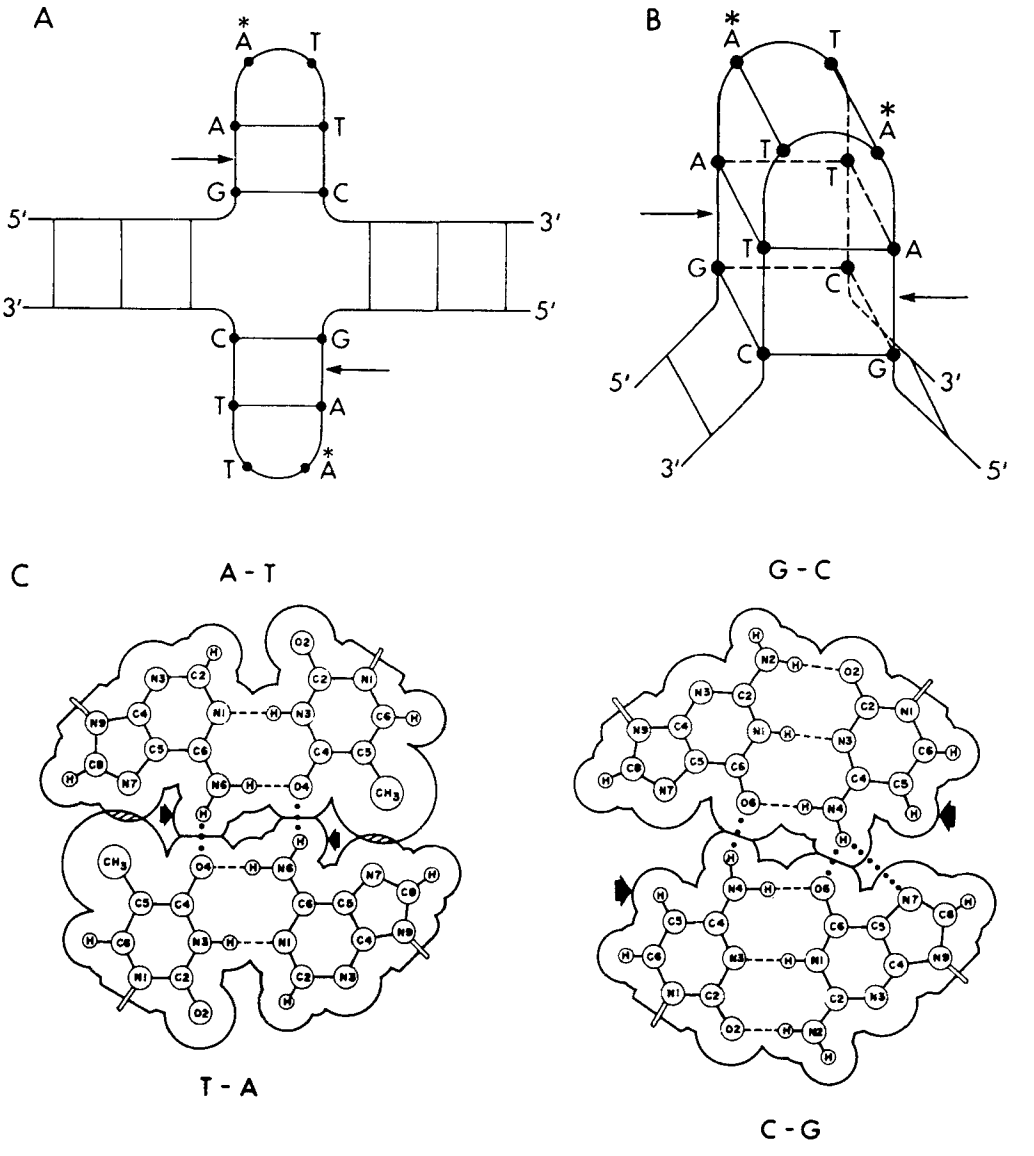

FIGURE 2. Proposed special structures for symmetrical recognition sequences. (A) Hairpin structure for *Eco* RI site.[84,85] Arrows and asterisks indicate sites of cleavage and modification, respectively. (B) Cage structure.[88,89] This structure is based on the hydrogen bonding scheme proposed by McGavin[90] shown in panels (C) and (D) by which like base pairs might interact. Large arrows in panels (C) and (D) indicate base determinants which are subject to modification by Type II methyltransferases.

symmetry is not a general requirement for site-specific DNA cleavage as evidenced by the class of enzymes typified by *Hph I*[56] (Table 1). This endonuclease, and two similar enzymes, *Hga* I[91,92] and *Mbo* II,[93,94] recognize five base pair sequences devoid of symmetry and cleave the DNA helix five to ten nucleotides to one side of their recognition sequences.

Models invoking *major* conformational transitions in DNA recognition sequences as being important in sequence recognition by proteins, such as those cited above for twofold symmetrical sequences, have received substantial attention in the literature. At present, however, evidence implicating major rearrangements in helix structure as being essential in such processes is generally lacking. In particular, the involvement of

cruciform structures appears to be excluded in the several systems that have been examined in detail. Although relatively long inverted repeat sequences may adopt cruciform structures at least transiently when present in negatively supertwisted DNA,[86,87] involvement of such a structure has been excluded during recognition of the relatively long *lac* operator sequence.[95] In the case of the four to six base pair sequences recognized by Type II restriction and modification enzymes, the adoption of a cruciform conformation is expected to be highly unlikely on steric grounds, and occurrence of such structures has been ruled out during recognition by *Eco* RI endonuclease,[40], *Eco* RI methylase,[96] and *Hpa* II endonuclease.[62]

The preceding comments are not meant to exclude local fluctuations in DNA conformation as being important in sequence-specific interactions. Indeed, sequence dependent variation in helix stability and structure is well documented,[82,83] and the importance of such effects in protein-DNA interactions would not be surprising. On the other hand, it has been noted that adequate discrimination of the four base pairs can be achieved by simple interactions between protein side chains and DNA determinants exposed only in the major and minor grooves of the B- form helix.[97,98]

At least a partial resolution of the DNA conformation question should be forthcoming with solution of crystal structures of protein-DNA complexes. As mentioned above, an *Eco* RI endonuclease-DNA complex has been crystallized.[77] Moreover, Dickerson and his colleagues[99] have recently reported a partially refined crystal structure analysis of the self-complementary dodecamer d(CpGpCpGpApApTpTpCpGpCpG) which contains a centrally located *Eco* RI sequence. Although conclusions concerning deoxypentose pucker cannot be drawn at the current stage of refinement, the oligonucleotide duplex is essentially a Watson-Crick B helix with an average rise of 3.4 Å per residue and 10.1 base pairs per turn.[99] Furthermore, no sequence-dependent variation in B helix geometry have been observed. While such variations may become apparent as refinement continues, they presumably would constitute rather small effects. Perhaps the most interesting departure of this structure from classical B helix geometry is a slight "propeller twist" of each base pair which serves to increase stacking interactions between neighboring base pairs. Since electric dichioism measurements indicate a propeller twist of base pairs for DNA in solution[100,101] it is argued that the solid state structure is close to that which exists in solution.[99] It must be pointed out, however, that the helical repeat observed in the crystal is somewhat less than the 10.4 ± 0.1 base pairs per turn determined by Wang[102] for a natural DNA segment in solution, although the significance of this difference is not yet clear. It thus seems evident that the next several years will see substantial progress in the structural analysis of oligodeoxyribonucleotides and complexes between such molecules and sequence-specific proteins. Advances in this area will certainly serve to clarify DNA conformational transitions which are associated with DNA sequence recognition by proteins and which persist in crystalline complexes.

In the context of conformational variation of DNA structure it is appropriate to consider possible effects of site-specific DNA methylation on helix structure. As mentioned above, such an effect is implicit in the "cage" structure or tetrad model for Type II recognition sites as proposed by Strusiak and Klopotowski[88], although it must be emphasized that no evidence is presently available for the involvement of a tetrad conformation in Type II recognition. Nevertheless, the potential may exist for conformational transition in the helix as a consequence of adenine or cytosine methylation. Engel and von Hippel[103] have demonstrated that base pairs between thymine and N-6-methyladenine are destabilized by 0.4 to 0.9 kcal/mol relative to A-T pairs, presumably as a consequence of rotation about the C^6-N bond. It is also possible that transfer of a hydrophobic methyl group to the N^6 of adenine could substantially alter major groove water structure in the vicinity of the modified base. Recent experiments, however, indicate that while a detectable DNA conformational transition is associated

with adenine methylation, its magnitude is extremely small, corresponding to an unwinding of the Watson-Crick helix by somewhat less than 1° per methyl group at 23° in the absence of Mg^{+2} at an ionic strength of approximately 0.1.[104] Although these particular experiments do not necessarily bear on the cage structure hypothesis, they do indicate that adenine methylation results in a minimal local alteration of helix parameters.

Similar measurements on the magnitude of helix perturbation as a consequence of 5-methylation of cytosine have not yet been performed. However, it has been recently found that poly d(G-m^5C) undergoes the transition from right-handed B form to left-handed Z form at much lower cation concentrations than does poly d(G-C).[135] For example, in 50 m*M* NaCl, 5 m*M* Tris-HCl (pH8.0) the midpoint of the B to Z transition of poly d(G-C) occurs at 700 m*M* Mg^{+2} while in the case of poly d(G-m^5C) the transition occurs at 0.6 m*M*. Although the density of 5-methylcytosine in the copolymer is very high, this profound effect certainly occurs in the physiological range of Mg^{+2} concentration.

IV. DNA CLEAVAGE BY TYPE II ENDONUCLEASES

Of the numerous Type II endonucleases known,[4] the *Eco* RI enzyme has been examined in greatest detail with respect to mechanism of DNA cleavage. Therefore the following discussion will focus on this endonuclease, with information on other systems being cited as available. In vitro restriction by Type II endonucleases, which requires only Mg^{++} ion and unmodified DNA, involves hydrolysis of two phosphodiester bonds (Table 1). Furthermore, the reaction mediated by the *Eco* RI enzyme is kinetically simple. Reaction velocities are first order with respect to endonuclease concentration and kinetics are Michaelis-Menten.[40,41,105] Since the stable form of the endonuclease at catalytic concentrations is the dimer,[41,70] this simple kinetic behavior implies that an active enzyme-substrate complex results from interaction of an endonuclease dimer with the *Eco* RI sequence. At 37°, K_m's for the single *Eco* RI sites of plasmids ColE1 and pBR322 are 8 n*M* and 4 nM, respectively, and the turnover number in both cases is about four double-strand scissions per min per dimer.[41,106,145] At 15°, the K_m for the *Eco* RI site of SV40 DNA has been found to be 30 n*M* with a turnover number of 1.5 min^{-1}.[40] Comparable values can be calculated from the data of Berkner and Folk[105] who examined the kinetics of cleavage of bacteriophage λ DNA: at 37° the K_m for *Eco* RI sites is 10 n*M* with a turnover number of 1.3 min^{-1}. These surprisingly low values for the turnover number reflect inherently slow catalysis since, as detailed below, at least 30 to 50% of isolated endonuclease molecules are active in site-specific binding and DNA cleavage.

The early studies of Smith and Wilcox[9] demonstrated that DNA cleavage by *Hin*d II endonuclease proceeded without extensive accumulation of an intermediated incised in only one strand of the recognition sequence. Although this work did suggest the involvement of such an intermediate in the *Hin*d II reaction, it was evident that the two single-strand events required for double-strand cleavage were coupled. More recently, the direct involvement of an intermediate species containing a single-strand scission in the recognition sequence has been demonstrated in the reaction catalyzed by *Eco* RI endonuclease.[41,107–109] In particular, results of steady state and transient kinetic experiments are consistent with the mechanism for double strand cleavage shown in Figure 3.[108]

The path of the *Eco* RI endonuclease reaction is dependent on both the nature of the DNA substrate as well as reaction temperature, with effects of these parameters being manifest at the level of the E·II intermediate. Under steady state conditions at 37°, double strand cleavage of ColE1, G4 RFI, and pBR322 substrates (covalently closed

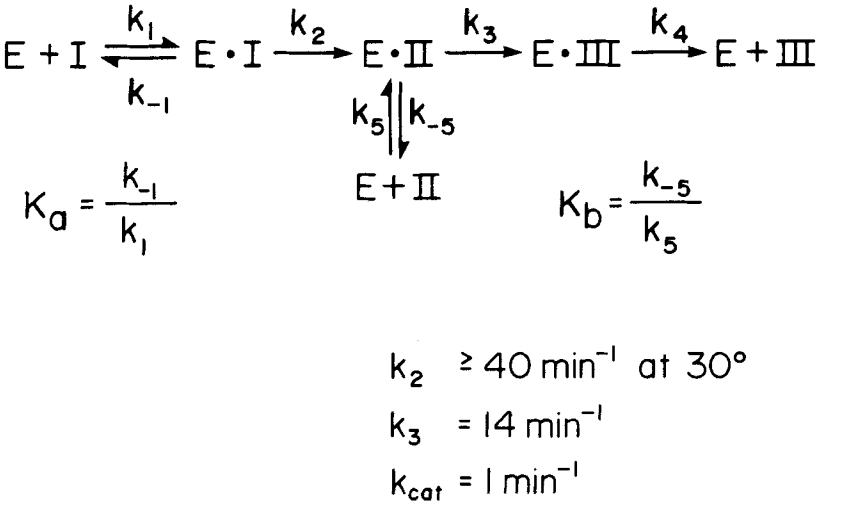

FIGURE 3. Proposed mechanism of DNA cleavage by *Eco* RI endonuclease, I, II, and III represent, respectively, an intact *Eco* RI sequence; an intermediate species which has been subject to cleavage in only one DNA strand; and a site which has suffered double strand hydrolysis. Rate constants shown were determined at 30°.[41] Other details are provided in the text. Adapted from Rubin and Modrich.[108]

duplex circular DNA molecules each of which contains a single *Eco* RI site) proceeds without detectable dissociation of the E·II intermediate.[41,106,108,110] Under these conditions the form II intermediate has been shown to remain enzyme-bound, and the two single-strand events required for double-strand cleavage are coupled by virtue of this fact.[41] These findings indicate that the *Eco* RI endonuclease can interact with both DNA strands during a single binding event under these conditions. Given that the *Eco* RI sequence is twofold symmetrical and that the active form of the endonuclease is a dimer of a single polypeptide chain, these results have been interpreted in terms of symmetry of the complex between endonuclease and the *Eco* RI sequences.[8] More recent evidence substantiating this view is cited below (Section VIII.)

In contrast to the mechanism of cleavage of DNA substrates cited above, cleavage within *Eco* RI sites of SV40 or pMB9 DNA proceeds with significant dissociation of the form II intermediate from the enzyme surface.[107-109] However, there are differences in the literature concerning the extent to which dissociation occurs. Ruben et al.[107] observed a rapid conversion of 60% of form I SV40 molecules to form II prior to generation of form III linear molecules, which appeared only after a lag of several minutes. These findings led to the conclusion that first and second single-strand scissions are mediated by different enzyme molecules. Similar conclusions have been subsequently drawn by Halford et al.[109] for pMB9 DNA. In analogous experiments Rubin and Modrich[108] also observed generation of free form II intermediate during cleavage of SV40 DNA. However, intermediate accumulation in these experiments was much more limited, and form III product appeared without a lag. These observations together with analysis of comparative kinetics of hydrolysis of form I SV40 and the free form II intermediate led to the conclusion that double-strand cleavage of this viral DNA proceeds with dissociation of the E·II intermediate in only 25% of catalytic events. While the reason for this discrepancy is unclear, it is noteworthy that endonuclease preparations utilized in the several studies were isolated by three different procedures. Only in the latter study was a

well characterized enzyme preparation[41,44,70,74] utilized and single strand breaks within form II molecules localized to the *Eco* RI sequence by unambiguous criteria.†

Although introduction of the two single-strand breaks within the recognition site by *Eco* RI endonuclease are typically coupled at 37°, coupling is markedly reduced at lower reaction temperatures.[41] At 0°, extensive dissociation of the E·II intermediate occurs, and form II molecules represent the major reaction product. This property can be employed to introduce site-specific single strand breaks into duplex DNA.

Kinetic analysis of a single turnover of *Eco* RI endonuclease with ColE1 DNA has permitted quantitation of several first order rate constants relating intermediate species alluded to above (Figure 3). At 30° k_2 is in excess of 40 min^{-1} while k_3 has a value of 14 min^{-1}, corresponding to a half life of 3 s.[41] Since both rate constants are much greater than the turnover number of 0.7 min^{-1} determined under identical conditions, it has been concluded that the rate limiting step in catalysis is k_4, or release of endonuclease from its DNA product. Additional support for this conclusion has been provided by comparison of the steady state rate of hydrolysis of form I DNA with that of the form II intermediate species. Under conditions of near substrate saturation, both forms of the *Eco* RI sequence are attacked at equivalent rates indicating that k_3 or k_4 is rate limiting. Thus, available evidence is consistent with the mechanism of Figure 3, with k_4 being rate limiting at physiological temperature.

While occurrence of an unmodified *Eco* RI recognition sequence appears sufficient for double-strand scission, it is evident that efficiency of cleavage is dependent on additional factors of DNA structure. In 1975, Thomas and Davis[111] initially observed differences in efficiency of double strand cleavage of the five *Eco* RI sites of bacteriophage λ DNA. Similar results have been obtained with adenovirus DNA by Forsblum et al.[112] who attributed differences in double strand cleavage efficiency to K_m effects. In both studies, conditions known to stabilize DNA secondary structure were found to enhance differences in efficiency of double strand cleavage among multiple substrate sites. In this respect it is noteworthy that cleavage of the *duplex* form of d(pTpGpApApTpTpCpA) by the endonuclease at 15° is governed by the same V_{max} as cleavage of SV40 DNA.[40] However, the K_m for the oligonucleotide duplex is 200-fold greater than that for the viral molecule. All of these observations clearly indicate that interaction of the endonuclease with its recognition sequence is modulated by sequences external to the *Eco* RI site. The substrate dependence of the *Eco* RI cleavage mechanism considered above has also been interpreted in terms of outside sequence effects[108] although it is not clear that this phenomenon is related to differential efficiency of double-strand cleavage. Since *Eco* RI endonuclease interacts with about 10 base pairs (below), such phenomena may simply reflect nucleotide sequences flanking the *Eco* RI sites in question. Indeed, it has been noted that the sequence-dependence of the mechanism of cleavage by *Eco* RI correlates with the extent to which purine-pyrimidine symmetry persists beyond the *Eco* RI site.[108] However, factors other than nearest neighbor effects could certainly be involved.

Unfortunately, features of mechanism analogous to those considered above have not been pursued in detail for other Type II endonucleases. However, evidence for involvement of a cleavage intermediate has been obtained in several systems. The two single-strand scissions required for double strand cleavage by *Bam* HI and *Sal* I endonucleases have been reported to be coupled at 37°.[109] As in the case of *Eco* RI, the cleavage intermediate has been found to dissociate from *Bam* HI endonuclease at low temperature.[16]

† It will become apparent to the reader that different preparations of *Eco* RI endonuclease have been reported to differ not only at the level of subtle variation in mechanism, but in terms of absolute specificity as well. It is likely that these differences reflect alterations in enzyme structure which presumably occur during isolation. Availability of DNA sequences for *Eco* RI polypeptides[44,45] has demonstrated that *Eco* RI enzymes isolated by at least one method have not been subject to proteolysis during fractionation.[74]

In contrast, cleavage intermediates have been found to dissociate from *Hpa* II[107] and *Hind* III[109] enzymes at physiological temperature.

Thus, although limited, current information suggests that an intermediate containing a scission in only one DNA strand is a common feature of DNA cleavage by Type II endonucleases. It is evident that in such a mechanism interaction between endonuclease and intermediate is necessarily asymmetric. However, it is also possible that asymmetry may occur at the level of introduction of the first single-strand break. Indeed, strand preferences have been observed during cleavage of short DNA duplexes by *Hpa* II[62] and *Msp* I[64] endonucleases. In both studies, cleavage of a nine base pair DNA duplex was monitored:

5′ pGpApApCpCpGpGpApGpA
3′ TpTpGpGpCpCpTpCpTp

In the case of *Hpa* II, the rate of cleavage within the lower pyrimidine-rich strand was found to be about three times that for cleavage of the upper strand.[62] *Msp* I exhibited the opposite preference, cleaving the upper purine-rich strand about twice as rapidly as the other.[64] In contrast, both strands were attacked at an equal rate by *Mno* I endonuclease.[62] Since these three endonucleases all recognize the internal d(pCpCpGpG) sequence, the strand preferences of *Hpa* II and *Msp* I presumably reflect differential effects of nucleotides surrounding this sequence. These findings therefore suggest that these two endonucleases interact with nucleotides beyond the required tetranucleotide, but do so in different ways. They would also indicate that while proteins may interact equivalently with the d(pCpCpGpG) sequence within the two DNA strands, a certain degree of asymmetry may exist in protein-DNA complexes by virtue of interaction with outside nucleotides. This interpretation is similar to that cited above with respect to substrate dependence of the mechanism of *Eco* RI endonuclease.

Since biological DNA molecules typically contain multiple recognition sites for a restriction endonuclease, it is possible that such enzymes might act in a processive fashion. To date, this possibility has been explored only in the case of the *Eco* RI enzyme. Cleavage of plasmid pVH153 DNA, which contains two *Eco* RI sites, has been found to proceed with extensive accumulation of a free intermediate which has suffered only one double strand scission.[8,70] While this type of experiment is only qualitative, it is evident that the endonuclease exhibits little, if any, processive behavior over the distances separating *Eco* RI sites in this plasmid (the shorter distance being 3500 base pairs, 7600 the longer). Since these distances are on the order of the expected average distance between *Eco* RI sites in DNA, it would appear that processive behavior is not an important feature of the mechanism of this endonuclease. The mode of cleavage of λ DNA by the enzyme has led to similar conclusions.[113]

As discussed above, unmodified recognition sites are subject to Type II restriction while sites methylated on both DNA strands are resistant (Table 1). However, the interaction of an endonuclease with recognition sequences methylated on only one DNA strand is also of interest for several reasons. First, such sites are of biological interest since they will result from semiconservative DNA synthesis. Secondly, efficacy of hemi-methylated sites as substrates for an endonuclease provides insight into the mode of enzyme-DNA interaction in a given system. Type II endonucleases examined in this respect fall into two classes. Restriction endonucleases *Eco* RI,[42] *Dpn* I,[114] and *Dpn* II[114] have been shown to be inactive on hemi-methylated recognition sites. Hemi-methylated duplexes have also been shown to be resistant to double-strand cleavage in the case of the *Eco* RII system,[48] but the possibility of single-strand cleavage has not been examined. In contrast, it has been found that recognition sites methylated on only one DNA strand are subject to single-strand scission within the unmodified strand in the *Bsp* I system.[25]

Similarly, a duplex d(pGpApTpC) site containing 5-methylcytosine in only one strand is subject to single-strand cleavage in the unmethylated strand by *Sau* 3A.[115] In the case of this latter enzyme it is not known whether the cytosine methylation represents the natural modification. The basis for these distinct responses to hemi-methylated sites is not yet known. However, it can be noted that *Eco* RI and *Bsp* I, which have been obtained in pure form, differ with respect to stable oligomeric states in solution (Table 1).

V. METHYL TRANSFER BY TYPE II MODIFICATION ENZYMES

Only four Type II methylases, *Eco* RI,[42] *dam*,[34] *Hpa* I[60] and *Hpa* II,[61] have been isolated in pure form. Each of these enzymes is kinetically simple. In each case the rate of methyl transfer is first order with respect to enzyme concentration and Michaelis-Menten behavior is obeyed with respect to both AdoMet and DNA substrates. Since the stable form of each enzyme is the monomer (Table 1), these findings have led to the conclusion that all three methyltransferases function as the monomer.[34,42,60,61] Like Type II endonucleases, Type II modification enzymes are also slow enzymes with turnover numbers of 3 (*Eco* RI), 19 *(dam),* and 1 *(Hpa I)* methyl transfers per minute per monomer at 37°.[34,42,156]

In the case of *Eco* RI and *dam* enzymes, it has also been demonstrated that the mechanism of methyl transfer is consistent with the monomer being the functional form of these proteins. To determine whether one or two methyl groups are transferred to an unmodified site per DNA binding event, Rubin and Modrich[42] examined conversion of *Eco* RI sites to an endonuclease-resistant form as a function of average extent of methylation. In similar experiments, Herman[34] monitored conversion of d(pGpApTpC) sequences to a form sensitive to *Dpn* I endonuclease as a function of extent of *dam* methylation. These studies indicated that both modification enzymes transfer a single methyl group to a recognition site per binding event, and in the case of the *Eco* RI enzyme, also demonstrated that kinetic parameters for methylation of a hemi-modified site are not more favorable than those for methyl transfer to the unmodified sequence.[42] In the case of the *dam* methylase, a slight preference for methyl transfer to hemi-methylated sites was observed.[34] These findings are in contrast to the reaction mediated by Type I modification enzymes in which a marked preference for half-modified sites has been demonstrated.[116]

These four Type II methylases are novel in the sense that they represent the first examples of proteins which interact with twofold symmetrical DNA sequences in the monomeric state. In the absence of internal symmetry within the protein monomers, these findings would imply that complexes between such enzymes and their recognition sequences are necessarily asymmetric. For the *Eco* RI system this work, together with comparable studies with the endonuclease (above), has served to demonstrate that the two enzymes interact with their common recognition sequence in different ways.[42]

VI. SPECIFICITY OF TYPE II RESTRICTION AND MODIFICATION ENZYMES

Analysis of a variety of sequence-specific proteins has demonstrated that such proteins typically bind to DNA molecules lacking recognition sites, albeit with reduced affinity.[97,117] In principle, the spectrum of such interactions may range from those which are truly nonspecific to those in which a significant portion of the interaction energy is dictated by the DNA sequence. The latter class is expected to include interactions with a hierarchy of DNA sites that are to some extent homologous with the recognition sequence of the protein in question, (see, for example, Reference 117). In the case of

several Type II enzymes it has been possible to monitor a presumably related phenomenon, namely catalysis characterized by reduced sequence specificity in vitro.

Because this phenomenon bears on the fidelity of Type II enzymes, it seems appropriate to comment on this question first. The frequency with which Type II endonucleases cleave endogenous cellular DNA in vivo would appear to be very low. For example, it has been demonstrated that in vivo *Eco* RI restriction of bacteriophage λ DNA destroys biological activity of the viral molecule with a probability of 0.8–0.9 per *Eco* RI site present.[118,119] If one assumes that double strand cleavage within the *E. coli* chromosome (4×10^6 base pairs) would confer lethality with a similar probability, it is evident that in vivo restriction cleavage errors would be quite rare.† It must be noted, however, that the error rate of in vivo restriction may not solely reflect fidelity of sequence recognition by the restriction enzyme. It can be argued that the restriction endonuclease and chromosome are differentially compartmentalized within the cell, or alternatively, that secondary sites subject to low efficiency restriction are also subject to modification and thus rendered resistant to the endogenous endonuclease. Indeed, some evidence for in vivo modification of secondary sites by *Eco* RI methylase has been obtained.[120]

The analysis of in vitro fidelity of Type II enzymes is still at an early stage and has not yet been assessed in a systematic way in any system. It must be stated at the outset, however, that caution must be exercised in attributing restriction or modification events at secondary sites to the same enzyme species that mediate activity at canonical sequences, particularly in those cases involving poorly characterized enzyme preparations. Since it is well known that proteins can be subject to chemical alteration during isolation, the spectre of chemically distinct enzymatic species in a purified preparation must be carefully considered in fidelity experiments. This is not a trivial point since, as will be detailed below, evidence for such effects already exists in the case of DNA cleavage by the *Eco* RI endonuclease.[121,122]

Although double strand cleavage by Type II endonucleases exhibits a strict requirement for established recognition sequences under optimal reaction conditions, two modes of DNA cleavage characterized by reduced specificity have been observed in vitro. Several laboratories have found that under optimal reaction conditions certain DNA sites can be subject to single but not double-strand cleavage by Type II endonucleases.[121–124] Sequences "nicked" under these conditions are typically largely homologous to the canonical recognition site, differing by one base pair. This phenomenon has been analyzed in greatest detail in the case of the *Eco* RI enzyme. Bishop[121] initially reported that the sequence $\frac{d(G\downarrow A\text{-}A\text{-}T\text{-}T\text{-}A)}{d(C\text{-}T\text{-}T\text{-}A\text{-}A\text{-}T)}$ is subject to single strand cleavage by *Eco* RI, and suggested that occurrence of the sequence G-A-A-T-T is sufficient for cleavage within a DNA strand. However, subsequent studies[122] demonstrated that this reaction is mediated only by certain preparations of the endonuclease. Hence, the nicking activity, designated *Eco* RI^n, is distinct from the enzymatic species which catalyzes double strand cleavage within the canonical site.

A second type of variation in the specificity of Type II enzymes has been observed under conditions of altered reaction environment. These effects, which generally involve "relaxation" of specificity, have been documented for number of endonucleases, including *Eco* RI,[125–129] *Bam HI*,[129,130] *Bst* 1,[28] *Bsu* 1,[31] *Hha* I,[129] and *Hind* III.[126] This relaxation of specificity is typically observed under conditions of high endonuclease concentration and has been promoted by variation of pH and ionic strength,[125,127,129] replacement of Mg^{+2} by Mn^{+2},[126–129] and supplementation of reactions with glycerol, dimethylsulfoxide, or other water-miscible organic solvents.[127,129,130] The best studied endonuclease in this

† Although it should be possible to determine effects on cell viability of the presence of a restriction-modification system, such experiments have not been reported.

respect has been that of *Eco* RI specificity, and only work with this enzyme will be considered here.

Polisky et al.[125] initially demonstrated a reduction in specificity of highly purified *Eco* RI endonuclease under conditions of elevated pH and low ionic strength. Analysis of DNA termini generated by this activity (designated *Eco* RI*) suggested that cleavage was occurring within the tetranucleotide sequences d(X↓A-A-T-T-Y).[125] However, even under *Eco* RI* conditions DNA cleavage was observed to occur much more rapidly at the canonical *Eco* RI sequence, with cleavage at alternate sites exhibiting a gradient of preference for nucleotide X: dG (0.59) > dA (0.26) > dT (0.14) ⩾ dC (0.006).

In related studies Hsu and Berg[126] and Tikchonenko et al.[127] demonstrated a similar reduction in specificity in the presence of Mn^{+2} ion or organic solvents (glycerol, dimethylsulfoxide, dimethylacetamide, and dimethylformamide). It would appear that a similar set of DNA sites is subject to cleavage under the several conditions of reduced specificity,[127,128] and synergistic effects have been observed.[127] Although it is difficult to generate limit *Eco* RI* digests under conditions of low ionic strength and elevated pH, or at moderate salt concentrations in the presence of Mn^{+2}, near limit digests have been observed under the former conditions in the presence of Mn^{+2}.[127]

More recently von Hippel and colleagues[128] have exploited availability of the total DNA sequence of ϕX174,[131] which is devoid of *Eco* RI sites, to assess sequences cleaved under *Eco* RI* conditions. In these experiments, *Eco* RI* sites within the duplex form of the DNA were localized by restriction mapping, and corresponding regions of the viral DNA sequence inspected for sequences homologous to the canonical *Eco* RI site. This analysis suggested that the most labile *Eco* RI* sites within this DNA are d(G-G-A-T-T-T), d(A-A-A-T-T-T), d(G-A-A-T-T-T), and d(G-A-A-T-T-A). However, observed restriction patterns of ϕX174 DNA were not consistent with *Eco* RI* cleavage at all d(A-A-A-T-T-T) sequences,[128] and d(G-G-A-T-T-T) sites, which were identified as the most labile sequences in the ϕX174 molecule, appear refractory to cleavage in the case of SV40 DNA.[126,128] Although these discrepancies may reflect the nature of flanking sequences, it is evident that elements of DNA sequence governing reduced specificity cleavage by *Eco* RI endonuclease have been only partially defined.

Relaxation of specificity of Type II endonucleases has generally been attributed to effects of altered reaction environment on the nature of complexes between substrate and a chemically authentic form of the endonuclease. This explanation is consistent with the presence of secondary cleavage activity in highly purified, and in several cases homogeneous endonuclease preparations,[125,126,128-130] cofractionation of *Eco* RI and *Eco* RI* activities[127,132] and similar effects of Mn^{+2} ion, glycerol, and other solvents on several endonucleases of differing specificity. In this case the hierarchy of susceptibilities of secondary sites would be expected to bear on the mechanisms of sequence recognition by Type II enzymes, and in the case of the *Eco* RI endonuclease, a model for recognition based on specificity of *Eco* RI* cleavage has been proposed.[128]

In contrast, Tikchonenko et al.[127] have suggested that *Eco* RI and *Eco* RI* activities may reflect distinct enzymatic species. They have reported that *Eco* RI and *Eco* RI* activities of purified preparations are differentially labile during prolonged storage; that *Eco* RI endonuclease isolated in the presence of nonionic detergents contains more *Eco* RI* activity than enzyme prepared in their absence; and that *Eco* RI* activity is abolished by parachloromercuribenzoate while *Eco* RI activity is unaffected by this reagent. Although these observations do not prove that *Eco* RI and *Eco* RI* cleavage reactions are mediated by distinct enzymatic species, it is apparent that the origin and nature of *Eco* RI* activity have not yet been completely resolved.

Reduction in specificity of methyl transfer by *Eco* RI modification enzyme has also been observed at high enzyme concentrations or in the presence of glycerol.[133,134] Under these conditions methyl transfer occurs exclusively to adenine residues.[133]

Since $d(A\text{-}T)_n \cdot d(A\text{-}T)_n$ accepts methyl groups slowly under such conditions while $d(A)_n \cdot d(T)_n$ does not, it has been suggested that the minimal substrate requirement for the methylase is an $\begin{smallmatrix}A\text{-}T\\A\text{-}T\end{smallmatrix}$ duplex.[134] Furthermore, the fact that natural DNA methylated under reduced specificity conditions is largely resistant to *Eco* RI* cleavage indicates that an overlapping set of sites are recognized by reduced specificity forms of the two enzymes.[133,134] Indeed, it has been proposed that reduced specificity modification may function physiologically to protect noncanonical sites against *Eco* RI* cleavage in vivo.[133] That in vivo modification of secondary sites may occur is suggested by the finding that DNA isolated from a strain bearing the *Eco* RI restriction-modification system is a poor acceptor for noncanonical methylation in vitro, as compared to DNA isolated from *E. coli* strains which do not carry this restriction-modification system.[134] However, a comparative analysis of the sensitivity of such DNA preparations to in vitro *Eco* RI* cleavage has not been reported.

The specificity of Type II enzymes with respect to polynucleotide conformation has also received substantial attention. Most of this work has dealt with activity of such enzymes on single-stranded as opposed to duplex DNA. It has been found in several laboratories that single-stranded circular DNAs isolated from virions of f1, M13, and ϕX174 are subject to cleavage at reduced rates by some, but not all, Type II endonucleases.[136138] Moreover, the pattern of DNA fragments obtained from cleavage of the single-stranded molecules is the same as that obtained from restriction of the duplex forms. The mechanism of this single-strand cleavage reaction has been assessed only in the cases of *Hae* III and *Hha* I,[139] and *Msp* I[64] endonucleases, with results of these two analyses leading to slightly different interpretations.

Several lines of evidence led Blakesley et al.[139] to conclude that cleavage of ϕX174 single-strands by *Hae* III or *Hha* I does not involve cleavage of recognitions sequences in the single-stranded state, but rather occurs within regions of duplex helical structure which may be transiently stable or otherwise. Cleavage of single-strands by these two endonucleases was exquisitely sensitive to inhibition by low levels of actinomycin D and netropsin, both of which bind only to duplex DNA. Secondly, cleavage of single-stranded and duplex forms of ϕX174 were differentially sensitive to the temperature. For example, *Hae* III hydrolysis of single-strands exhibited a much lower temperature optimum (47°) than did the reaction with duplex circles (72°). Moreover, near the respective temperature optima certain sites in the single-stranded substrate were rendered resistant to hydrolysis while all sites within the double helical substrate remained sensitive. These findings strongly suggest that *Hae* III and *Hha* I sites within ϕX174 single strands are located within regions of secondary structure which are differentially stable.[139]

In more recent experiments Yoo and Agarwal[64] analyzed cleavage of the complementary oligonucleotides d(pGpApApCpCpGpGpApGpA) and d(pTpCpTpCpC-pGpGpTpT) by *Msp* I endonuclease. While the nine base pair duplex formed between the oligonucleotides was readily cleaved by the endonuclease, it was demonstrated that the individual oligonucleotides were also subject to cleavage at reduced rates by the enzyme. Since the self-complementarity of each oligonucleotide is limited to the d(pCpCpGpG) recognition sequence, the latter reaction presumably cannot reflect formation of stable duplex structures by individual strands.[64] Furthermore, examination of DNA sequences surrounding d(pCpCpGpG) recognition sites within the ϕX174 molecule[131] indicated that a six base pair duplex is the longest region of secondary structure involving an *Msp* I site that can be formed by intramolecular base pairing within the viral single-strand.[64] Hence, it is highly unlikely that *Msp* I cleavage of ϕX174 single-strands reflects preexisting regions of stable duplex structure.

While these findings may indicate cleavage within single-stranded regions by *Msp* I,

rates of cleavage of the individual oligonucleotide strands mentioned above were found to decrease markedly with increasing temperature in the range of 4° to 37°, suggesting a more complex mechanism.[64] Based on these observations and the properties of other Type II endonucleases, Yoo and Agarwal[64] proposed the following model for *Msp* I action: a form of the endonuclease characterized by two-fold symmetry (postulated to be the dimer) is capable of interaction with d(pCpCpGpG) sequences in either the single-stranded or duplex conformation, with complexes involving duplex sites possessing elements of twofold symmetry. When the substrate is a single strand, it is proposed that two identical DNA · protein complexes interact to yield an intermediate similar to that resulting from interaction of the enzyme with preformed duplex substrate. It is argued that in either case, enzyme-induced transient complexes so formed are rapidly converted to an equimolar complex of recognition sequence and enzyme subunit, with interactions among subunits and between subunits and recognition sequence occurring in a concerted manner to determine a rate of DNA cleavage. Thus, the distinction between this model and that of Blakesley et al.[139] cited above is that the latter involves endonuclease recognition only within regions of preformed secondary structure. It should be emphasized that the model of Yoo and Agarwal[64] is based on the assumption that duplex *Msp* I sites could not have pre-existed in their experiments. Although the formation of stable regions of secondary structure was excluded in this work, the authors note that transient formation of duplex recognition sites, which could be subsequently stabilized by interaction with enzyme, was not ruled out.

In the case of those endonucleases which are inactive on natural DNA in the "single-stranded" configuration, it could be argued that failure to observe cleavage reflects inhibition by the high levels of nonspecific single-strand sequences present. However, analysis of oligonucleotide cleavage by *Eco* RI and *Hpa* II endonucleases indicate that at least these enzymes cleave only the duplex form of their recognition sites. In the case of *Eco* RI, Greene et al.[40] demonstrated cleavage of the self-complementary octanucleotide d(pTpGpApApTpTpCpA), but the dependence of rate of cleavage on oligonucleotide concentration was greater than first order. Since the duplex and single-stranded forms of the octanucleotide were in reversible equilibrium at reaction temperature, it was argued that the anomalous rate dependence reflected a requirement for duplex structure. Analysis of the kinetic data in terms of the equilibrium constant governing octanucleotide association was consistent with this view. Similarly, Baumstark et al.[62] have demonstrated that while the nine base pair duplex formed between d(pGpApApCpCpGpGpApGpA) and d(pTpCpTpCpCpGpGpTpT) is readily cleaved by *Hpa* II, the individual single strands are not. It is interesting to note that, as discussed above, these oligonucleotides are those which were subject to cleavage by *Msp* I in the "single-stranded" form. Thus, *Hpa* II and *Msp* I differ in this respect despite their common recognition sequence.

The requirement for a duplex recognition sequence in the Type II modification reaction has not been examined in detail. *Hae* III, *Hpa* II, and *dam* modification enzymes have been found to methylate denatured salmon sperm or T7 DNAs, albeit at rates somewhat less than that for duplex substrates.[34,52] However, the possibility that methylation of single-stranded DNA by these enzymes occurs within regions of secondary structure has not been explored. In contrast, *Eco* RI and *Hpa* I modification enzymes do not methylate denatured DNA.[60,96] Perhaps the most intriguing observations in this respect have been made with the *Hha* I methylase.[52] This enzyme has been found to methylate a random copolymer of $d(G, C)_n$, with the extent of methylation being consistent with the expected frequency occurrence of the tetranucleotide recognition sequence of this enzyme within the random polymer. While modification of this polymer could reflect secondary structure, the enzyme also catalyzes methyl transfer to random

d(N-acetylG, C)$_n$ and d(X, C)$_n$. Since these latter polymers were devoid of Watson-Crick secondary structure as judged by the absence of thermal melting transitions, these observations suggest that the *Hha* I enzyme can recognize and methylate a recognition site in the single-stranded conformation. However, both rate and extent of methyl transfer to denatured *M. luteus* DNA were found to be only 30% of those observed with the duplex form of this substrate.[52] The reduced extent of methyl transfer by the *Hha* I enzyme to single-stranded natural DNA suggests that at least some sites are not subject to modification in this form.

The potential for cleavage of DNA · RNA hybrid duplexes by Type II endonucleases has also been examined in one study.[140] In these experiments RNA · DNA hybrids prepared by reverse transcription were subjected to hydrolysis by a number of Type II enzymes, with cleavage of the isotopically labeled DNA product being monitored. Several endonucleases, including *Eco* RI, *Hae* III, *Hha I*, *Hin*d II, and *Msp* I, were found to cleave a substantial fraction of the labeled DNA, with the size of products being consistant with location of recognition sequences within the RNA template utilized for reverse transcription. Unfortunately, neither substrates nor products of endonuclease action were characterized with respect to DNA · RNA hybrid structure. It is possible that DNA duplexes represented a portion of the reverse transcription product, or alternatively that hydrolysis of the RNA component of RNA · DNA duplexes (by RNase contamination of restriction enzyme preparations) resulted in DNA single strands which could have been cleaved by several of the endonucleases tested.

VII. TYPE II ENZYME-DNA INTERACTION: THERMODYMIC AND KINETIC PARAMETERS

The analysis of structures and catalytic properties of Type II endonucleases and methyltransferases outlined above has provided a crude picture of the modes of interaction of such proteins with their recognition sites. However, they have not indicated whether sequence-specificity is manifested at the level of DNA binding or catalysis. Furthermore, alternate methods are required to delineate contacts between protein and polynucleotide which are important in establishing specificity. This section will consider thermodynamic and kinetic parameters which govern interaction of Type II enzymes with DNA in the absence of catalysis. That which follows will summarize information available on DNA contacts which may be important in specific DNA-protein complexes.

The binding of Type II enzymes to DNA has been monitored by several methods. In the absence of a divalent cation, conditions under which DNA cleavage does not occur, *Eco* RI endonuclease forms site specific complexes with DNA which are retained on nitrocellulose membranes.[8,41,43,141,142] Such complexes are site-specific since their formation is markedly reduced if DNAs bear the *Eco* RI modification[8,43,142] or if the *Eco* RI recognition sequence has been selectively deleted from the molecule.[43,141] The only other Type II endonuclease which has been examined using this assay is the *Bgl* I enzyme.[18] While this enzyme forms nitrocellulose-retainable complexes with SV40 DNA in the absence of Mg^{+2}, these complexes have not been shown to be site-specific.

In the case of the *Eco* RI enzyme, specific binding as monitored by this method is hyperbolic and exhibits requisate criteria for an equilibrium process.[8,41,43,141,142] Furthermore, the simple nature of specific binding observed during titration of DNA with increasing endonuclease indicates that oligomerization of enzyme is not required for specific complex formation. Since the stable form of the *Eco* RI enzyme is the dimer in the concentration range employed in these experiments (Section II), these findings, like the kinetic studies cited above, demonstrate that the dimer interacts directly with the *Eco* RI sequence.

At 37° and ionic strengths of 0.07 to 0.15 *M*, the apparent equilibrium dissociation constant for specific complexes between *Eco* RI endonuclease and plasmid pBR322 (one *Eco* RI site) is on the order of 10^{-11} to 10^{-10} *M*.[8,43,142] Moreover, binding curves obtained under conditions of invariant endonuclease concentration have demonstrated that a large fraction of isolated endonuclease (as the dimer) is active in site-specific binding.[43] Since such complexes undergo rapid cleavage upon addition of Mg^{+2},[41] this implies that a large fraction of purified *Eco* RI endonuclease is catalytically active. At 22° and comparable ionic strengths, similar studies with derivatives of bacteriophage λ DNA containing only one *Eco* RI site yielded apparent dissociation constants on the order of 10^{-9} *M*.[141] It should be noted that these binding constants generally have been estimated assuming a simple two state model in which enzyme not involved in specific complex formation is presumed to be free in solution. However, the endonuclease also binds to DNA sites other than the *Eco* RI sequence (below). Since the effects of this type of interaction are neglected in the two state model, these dissociation constants should be regarded as only approximate.

Analysis of salt dependence of specific binding as monitored by the nitrocellulose method has indicated that the apparent affinity of the endonuclease for DNA containing *Eco* RI sites increases with increasing ionic strength up to about 0.15 *M* and then decreases at higher salt concentration.[141,142] The increase in affinity with increasing ionic strength has led to the suggestion that electrostatic interactions are actually unfavorable in endonuclease-DNA binding, perhaps reflecting the presence of negative charges within the DNA binding site of the enzyme.[142] However, several other interpretations of these findings are also plausible. For example, as mentioned above, dissociation constants have been estimated assuming a simple two state model which neglects nonspecific interactions. It is possible that the increase in affinity with increased ionic strength is only apparent in that it may not reflect tighter binding to an *Eco* RI site, but rather reduction in affinity for nonspecific sites within the DNA that compete with specific complex formation.

The interaction of *Eco* RI endonuclease with nonspecific DNA has been examined by several techniques. Woodhead and Malcolm[143] have demonstrated that nonspecific DNA sequences can protect *Eco* RI endonucleases against inactivation by methyl acetimidate. Utilizing this protection assay, equilibrium dissociation constants were estimated for several DNA species. At 20° in the absence of Mg^{+2} and an ionic strength of 0.2 *M*, this method yielded equilibrium dissociation constants of 1.2 μ*M* (φX174 DNA), 0.7 μ*M* (SV40 DNA), 15 μ*M* (λ DNA), 290 μ*M* ($d(G\text{-}C)_n$), and 300 μ*M* ($d(A\text{-}T)_n$), with concentrations expressed in terms of base pair equivalents. Although both SV40 and λ DNA contain *Eco* RI sequences, protection afforded by these molecules was largely due to interaction with other DNA sequences since endonuclease was present in large molar excess over recognition sites.[143]

Interaction of *Eco* RI endonuclease with sequences other than the *Eco* RI site has also been monitored by the nitrocellulose membrane assay. Equilibrium competition experiments, in which specific complex formation with pBR322 was competed by a derivative of this plasmid lacking the *Eco* RI sequence at an ionic strength of 0.07 *M*, has also indicated a nonspecific dissociation constant in the μ*M* range[96] which is similar to the values mentioned above obtained by the protection method under somewhat different conditions.

Competitive inhibition of DNA cleavage has also demonstrated that *Eco* RI endonuclease binds nonspecifically to a variety of polynucleotides in the presence of Mg^{+2},[144] although enzyme affinity for nonspecific sites may be reduced slightly by this cation.[141] Other than natural DNA containing *Eco* RI sites, highest affinity inhibitors identified in these studies were several alternating DNA copolymers and tRNA. Hence, as in case of *lac* repressor,[82,177] nonspecific interaction with polynucleotide is a characteristic of

Eco RI endonuclease. The possible significance of such interactions in the mechanism by which the enzyme locates its recognition sequence is considered below.

The foregoing discussion implies that a substantial component of the specificity of *Eco* RI endonuclease can be attributed to its affinity for the *Eco* RI sequence relative to that for other DNA sites. As mentioned above, the affinity of the enzyme for *Eco* RI sites is on the order of 10^4 to 10^6 times greater than that for nonspecific sequences. In addition, methylation of the *Eco* RI site of pBR322 appears to reduce the affinity of the endonuclease for the DNA to a level similar to that for the derivative plasmid from which the *Eco* RI site has been deleted,[43,142] Thus, not only is a substantial component of the specificity of the enzyme attributable to its high affinity for the *Eco* RI sequence, modification of the recognition site markedly reduces this affinity.

In contrast to the results outlined above, which imply that *Eco* RI endonuclease displays increased affinity for its recognition site relative to other DNA sequences, analysis of oligonucleotide binding by the endonuclease have led Goppelt et al.[145] to a different set of conclusions. This group examined binding of d(GpGpApApTpTpCpC), d(GpApApTpTpC), d(TpTpApCpApT) and d(TpApApApTpG) by monitoring reduction in ellipticity of the CD spectrum of the enzyme upon complex formation. Each oligomer was bound by these criteria, and in all cases the estimated equilibrium dissociation constant was about 10^{-7} *M* (10^{-6} *M* in nucleotide) at 5° in the absence of Mg^{+2}. On the other hand, half-maximum inhibition of cleavage of the *Eco* RI octamer, d(pGpGpApApTpTpCpC), by d(GpApApTpTpC) required a 50-fold excess of the hexamer at 20° in the presence of Mg^{+2}. Based on these findings, Goppelt et al.[145] concluded that specific binding is Mg^{+2}-dependent. There are, however, several difficulties in interpretation of this data. First, use of the CD assay necessitated use of high (μM) concentrations of enzyme and oligonucleotides, concentrations comparable to dissociation constants determined by other methods for nonspecific interactions (above). Moreover, compelling evidence that binding of the *Eco* RI octamer in the absence of Mg^{+2} represented an equilibrium, rather than a stoichiometric reaction, was not provided. Secondly, the *Eco* RI octamer employed in cleavage experiments was phosphorylated at the 5′-terminus while that employed in the CD experiments evidently was not.[145] As discussed below (Section VIII), this phosphate appears to be utilized as a contact by the endonuclease. Indeed, work with *Hpa* I endonuclease has demonstrated that a phosphorylated *Hpa* I octamer is a much better substrate than the unphosphorylated oligomer in this system. Hence, the comparison by these authors[145] of octamer and hexamer affinities by CD versus kinetic methods may be invalid. In view of these reservations concerning evidence to the contrary, it seems safe to conclude that *Eco* RI endonuclease displays specificity for binding to its recognition sequence both in the presence and absence of a divalent cation. Additional evidence supporting this view is presented below (Section VIII).

The nature of nonspecific interactions between *Eco* RI endonuclease and DNA are not yet understood. Woodhead and Malcolm[143] have suggested that nonspecific interactions between endonuclease and DNA primarily reflect binding to *Eco* RI* sequences. As yet, no evidence to support this contention is available. Indeed, it has been shown that nonspecific interactions are not affected by extensive methylation of the DNA by *Eco* RI modification enzyme under conditions of reduced specificity methyl transfer,[43] which is known to protect at least some secondary sites against *Eco* RI* cleavage.[133,134] It could be argued, however, that in contrast to effects modification on binding of endonuclease to the canonical *Eco* RI sequence, methylation of Eco RI* sites may not reduce their affinity for the endonuclease.

In addition to their high thermodynamic stability, specific complexes between *Eco* RI endonuclease and DNA formed in the absence of Mg^{+2} are also kinetically very stable. Specific complexes between the enzyme and plasmid pBR322 DNA dissociate with a first

order rate constant of 7×10^{-4} sec^{-1} (at 37° and ionic strength of 0.1) corresponding to a half-life of 16 min.[43,146] Surprisingly, the kinetic stability of such complexes has been found to be markedly dependent on the chain length of the DNA in which the *Eco* RI site is embedded. Jack[146] has analyzed dissociation kinetics of specific complexes formed with eight linear DNA fragments ranging in size from 34 to 4362 base pairs, all of which were derived from pBR322 and contained the *Eco* RI site of this plasmid in a centrally located position. These experiments demonstrated an increase in the first order rate constant governing dissociation from a value of 8×10^{-5} sec^{-1} ($t_{1/2}$ = 140 min) for the 34 base pair molecule to 7×10^{-4} sec^{-1} for the 4362 base pair full length linear duplex. Therefore, DNA sequences outside the *Eco* RI site kinetically destabilize specific complexes.

In contrast to the marked dependence of dissociation rate on DNA chain length, competition experiments demonstrated that the intrinsic equilibrium constant governing interaction of the endonuclease with the *Eco* RI sequence is independent of DNA chain length. The equilibrium ratio of specific complexes formed with long and short fragments was 1:1 when the two DNAs were present at equimolar concentrations in the same reaction.[146] Furthermore, as predicted by microscopic reversibility the rate of specific complex formation was also enhanced by DNA chain length. Thus, in reactions containing equimolar long and short molecules, specific complex formation by the endonuclease favored the longer DNA chain during initial stages of the reaction, with an equilibrium ratio of 1:1 achieved upon further incubation.[146] The magnitude of the initial preference for long DNA molecules (ratio of long:short DNA molecules bound) in such reactions reflects the relative rate constants governing association of enzyme with the two sizes of DNA molecule. Analyses of seven different pairwise combinations of DNA fragments demonstrated that the initial preference for binding to the longer DNA was always identical to the ratio of first order rate constants governing dissociation of enzyme from the fragments in question.[146] In other words, the ratio of association rate constants for two DNA fragments was always found to be equivalent to the ratio of their dissociation rate constants, a finding consistent with the observed lack of dependence of the equilibrium constant on DNA chain length.

These findings imply that DNA sequences outside the recognition site are involved in the major kinetic path by which *Eco* RI endonuclease locates and leaves its recognition sequence, although it must be emphasized that to date such effects have been demonstrated only in the absence of Mg^{+2} ion. The simplest interpretation of association and dissociation rate enhancement by outside DNA sequences is that interaction of the endonuclease with nonspecific sites is involved in the major path by which the endonuclease locates and leaves it recognition sequence. Indeed, the results cited above are perfectly consistent with the theoretical treatments of Richter and Eigen[68] and Schranner and Richter[69] for a unidimensional directed diffusion mechanism in which a sequence specific protein initially binds to nonspecific sites and then locates its recognition sequence by random diffusion along the DNA helix. In particular, results obtained with the *Eco* RI enzyme suggest that the average distance scanned by the random walk mechanism is about 1000 base pairs per DNA binding event under the experimental conditions employed. It is important to note that evidence for involvement of such a mechanism in specific site location by a protein is not limited to *Eco* RI endonuclease. von Hippel and colleagues[147] have examined the DNA chain length dependence of kinetics in the case of *lac* repressor and have drawn similar conclusions for this system.

Despite the ability of *Eco* RI endonuclease to form site-specific complexes in the absence of Mg^{+2}, attempts to detect specific binding by *Eco* RI methylase in the absence of AdoMet have not yet been successful.[8,43,143] While titration of DNA with methylase can lead to retention of a large fraction of input DNA on nitrocellulose membranes, a

large molar excess of methylase is required. Similarly, titration of a fixed concentration of methylase with DNA leads to retention of DNA-protein complexes on filters, but only at a level of 10 to 20% that expected based on input enzyme. Lastly, formation of filter-retained complexes in either case occurs with DNA containing unmodified or modified *Eco* RI sites, or with DNA devoid of the *Eco* RI sequence.[96,143] These findings could reflect formation of nonspecific DNA-methylase complexes which are retained on nitrocellulose with low efficiency, or perhaps less likely, the presence of a low level DNA binding contaminant in purified preparations of the enzyme. This failure to detect specific methylase-DNA complexes may be a consequence of assay procedures employed. For example, it is possible that specific complexes are not retained by nitrocellulose membranes. Alternatively, as in the case of Type I enzymes,[6,8] site specific binding of the methylase may require the presence of AdoMet. This would imply an ordered addition of substrates during the methyl transfer reaction. Such a mechanism would be consistent with the ability of the methylase to transfer only one methyl group per DNA binding event (Section V). However, these possibilities remain unresolved at the present time.

VIII. TYPE II ENZYME-DNA INTERACTION: DNA DETERMINANTS IMPORTANT IN SPECIFIC RECOGNITION

Two approaches have been utilized in attempts to define DNA contacts important in specific interactions between Type II enzymes and their recognition sequence. One approach has been to examine effects of base analog substitution within the recognition sequence on specific interactions. A number of analogs are available which form good Watson-Crick base pairs and which differ in simple ways from the common bases. Interaction of Type II enzymes with analog-substituted sites can be examined under catalytic, as well as noncatalytic, conditions in order to define DNA substituent effects on parameters governing specific recognition and catalysis.

The alternate approach, which was developed by Gilbert and colleagues,[148] relies on partial chemical alkylation of DNA fragments and is related to the chemical method for DNA sequence analysis.[123] In one variation, the alkylation protection method, an end-labeled DNA fragment is subjected to partial dimethylsulfate alkylation in the presence and absence of a DNA binding protein. Subsequent depurination and strand cleavage[123] permit identification of those purines whose reactivities toward the reagent are altered by the bound protein. Since this reagent methylates the N^7 of dG (major groove) and the N^3 of dA (minor groove),[149] this procedure is used to infer presence of polypeptide in the vicinity of affected nitrogens.

In the second variation, alkylation interference,[150] an end-labeled DNA fragment is partially alkylated prior to interaction with the protein. After addition of the protein, the DNA is separated (by filtration through a nitrocellulose membrane) into a fraction which is capable of binding to the protein and one which is not. Subsequent analysis permits identification of those nucleotides which, in the alkylated form, interfere with DNA-protein interaction. In addition to the use of dimethylsulfate, alkylation interference studies have also employed ethylnitrosourea, as an ethylating agent for backbone phosphates.[150,151] The alkylation interference technique, which can be seen to be analogous to base analog substitution, thus allows identification of 3-nitrogens of dA, 7-nitrogens of dG, and phosphates which may be utilized as contacts by the protein.

Although covalent alteration of a nucleotide determinant utilized as a contact during sequence discrimination would be expected to interfere with recognition by a protein, the possibility of indirect effects must also be considered. Available information concerning dependence of DNA secondary structure on nucleotide sequence is incomplete, but in at least one model system, a major effect on helix conformation as a consequence of

Table 2
POTENTIAL DNA CONTACTS WITHIN THE *Eco* RI SEQUENCE

Base	Position	Groove	Methylase contact	Endonuclease contact	Ref.
G	N_7	Major	?	Yes	155
	2-NH_2	Minor	Yes	No	154
A	N^3	Minor	?	No	155
A	6-NH_2	Major	Yes	Yes	39
	N^3	Minor	?	Yes	155
T, T [b]	5-CH_3	Major	Yes	No	105
C	5-H	Major	No	No	105, 152

[a] Positive assignment as a contact is subject to the reservations cited in the text.
[b] Assignment of the 5-CH_3 of dT residues as a methylase contact is based on analysis of PBS2 DNA which contains dU in place of dT.[153] Since both dT residues are presumably substituted within *Eco* RI sites of this viral molecule, it is not clear whether the 5-CH_3 of one or both nucleotides is required.

extensive 5-methylcytosine substitution has been documented.[135] The possibility of conformational effects has generally not been excluded in the analog and alkylation experiments discussed below. Similarly, indirect steric effects are also possible in alkylation interference studies or in experiments employing analogs with bulky substituent groups not present on the four common bases. For example, modification of a dC or dA residue by Type II methylase (Table I) would interfere with interaction between the corresponding restriction enzyme and the base determinant methylated (6-NH_2 of dA or C^5 of dC). However, analysis of molecular models indicates that the presence of such methyl groups may also interfere sterically with protein bonding to other determinants of the base pair modified, with the potential for steric interference extending to interactions with adjacent base pairs as well.

A. *Eco* RI Restriction-Modification Enzymes

Kaplan and Nierlich[152] and Berkner and Folk[105] have examined *Eco* RI cleavage and methylation of several bacteriophage DNAs which naturally contain analogs of dT and dC, with substituent variation at the 5-position of the pyrimidine ring. Since this position is exposed in the major groove of B-DNA, the simplest interpretation of substituent effects at this position is to define major groove contacts between protein and polynucleotide. Bacteriophage T4 DNA, which contains glucosylhydroxymethylcytosine in place of cytosine, is resistant to cleavage by *Eco* RI endonuclease.[105,152] Since nonglucosylated T4 DNA is hydrolyzed by the enzyme,[152] this resistance evidently reflects steric hinderance by the presence of glucosyl residues in the major groove. Furthermore, the fact that nonglucosylated DNA is a substrate for the enzyme implies that the presence of a hydroxymethyl substituent at the 5-position of dC does not block cleavage. However, since cleavage of the nonglucosylated substrate was only analyzed qualitatively, the possibility of some interference with recognition or catalysis by this substituent cannot be excluded. Results obtained with this and other analog-containing substrates are summarized in Table 2.

In related experiments Berkner and Folk[105] demonstrated that virus PBS2 DNA is hydrolyzed at normal rates by the *Eco* RI enzyme. Since this DNA contains uracil in place of thymine, this finding implies that the 5-methyl of thymine is not required for recognition of hydrolysis by the enzyme. In contrast, DNA from bacteriophage ϕe,

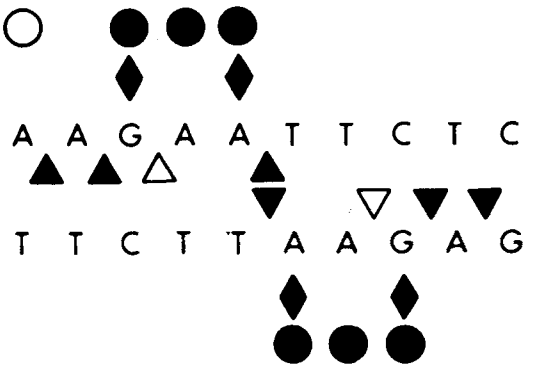

FIGURE 4. Results of alkylation protection and alkylation interference analysis on specific complexes between *Eco RI* endonuclease and the *Eco* RI sequence of pBR322 DNA.[155] O, ●—purines protected against alkylation by *Eco* RI endonuclease; ◆—purines, the methylation of which inter-fere with specific complex formation; △, ▲—phosphates, the ethylation of which interfere with specific binding. Closed symbols indicate large effects while open symbols indicate weaker but nevertheless significant effects. Shown is only that region about the *Eco* RI site over which alkylation effects were observed.

which contains hydroxymethyluracil in place of thymine, was found to be an extremely poor substrate for *Eco* RI endonuclease.[105] This effect appears to reflect a block at the level of DNA cleavage rather than recognition since the K_m for hydrolysis of ϕe DNA is similar to that for λ DNA, with the defect being evident in the V_{max}. In this respect it is noteworthy that bromouracil-containing λ DNA has also been found to be subject to cleavage at reduced rates,[153] although the kinetic basis for this effect has not been determined. Thus, while the 5-methyl of thymine is not essential for recognition or catalysis by the endonuclease, its replacement by bromo or hydroxymethyl groups substantially reduces rates of DNA cleavage. Reduced activities with the latter substituents could be the consequence of a conformational transition in the substrate, or alternatively may be due to electronic or steric effects reflecting their increased polarity or size.

In order to minimize the possibility of conformational perturbation of the DNA substrate in analog studies, Modrich and Rubin[154] selectively replaced the dG residues within the *Eco* RI site of Col El DNA with dI without altering other dG residues in the molecule. This substitution, which results in loss of the 2-NH_2 of dG exposed in the minor groove of B-form DNA, was without effect on rates of cleavage by *Eco* RI endonuclease. Hence, this DNA functional group is not required for recognition or catalysis by the enzyme.

Although the analog studies summarized above have not identified DNA determinants important in sequence recognition by *Eco* RI endonuclease, application of the alkylation protection and alkylation interference methods discussed above has suggested DNA contacts important in this process.[155] Results of these studies are summarized in Figure 4 and Table 2. It should be reiterated that application of these techniques is dependent on formation of stable and isolable complexes between protein and DNA. Hence, such experiments with *Eco* RI endonuclease have been performed in the absence of Mg^{+2}. As shown in Figure 4, the N^7 position (major groove) of the dG residue and the N^3 position (minor groove) of the two dA residues within the *Eco* RI sequence were protected by the

endonuclease against dimethylsulfate alkylation. In the absence of substantial helix perturbation upon endonuclease binding, these results suggest the presence of polypeptide in both major and minor grooves in the vicinity of these purine nitrogens. Moreover, interaction in the minor groove outside the *Eco* RI site was suggested by weak protection of the N^3 of a dA residue two base pairs 5′ to the recognition sequence.

Complementary alkylation interference analyses demonstrated that methylation of the N^7 of *Eco* RI site dG residue or the N^3 of dA adjacent to the dyad axis resulted in a marked reduction in specific complex formation, suggesting that these purine nitrogens may be utilized as contacts by the endonuclease.[155] Although the 5′-penultimate dA in the *Eco* RI sequence was protected against chemical alkylation in the presence of the enzyme, methylation of this purine did not interfere with specific complex formation. Therefore, the N^3 of this dA residue does not bond to the enzyme. In this context it is appropriate to note that methylation of the *Eco* RI sequence by the modification enzyme (Table 1) represents a natural example of methylation interference. Again, the simplest explanation for the resulting block to endonuclease recognition is that the 6-NH_2 (major groove) of the central dA is an essential contact for the enzyme. Although indirect steric effects as consequence of the natural modification cannot be excluded, the possibility of an associated DNA conformational transition in this case is highly unlikely (Section III).

Ethylation interference studies have suggested that four phosphates in each DNA strand are important in specific DNA endonuclease complexes[155] (Figure 4). It is interesting that two of these four residues are 5′-external to the *Eco* RI sequence. If it is assumed that these phosphates are involved in ionic bonding to the protein, this finding implies that a substantial fraction of the electrostatic component of binding energy is due to interaction with backbone phosphates outside the recognition site. The third phosphate, the alkylation of which markedly suppressed specific binding, lies on the dyad axis of *Eco* RI sequence. The last of the four phosphates implicated in this study is the one involved in the phosphodiester bond which is hydrolyzed by the enzyme. Surprisingly, ethylation of this phosphate did not inhibit specific binding as severely as did alkylation of the other three. While the significance of this observation is not yet clear, it is likely that this phosphate is coordinated to Mg^{+2} in the cleavage complex. Consequently, the nature of interaction between polypeptide and this phosphate may well differ from the mode of bonding to other phosphate residues.

A representation of the spacial distribution of endonuclease "contacts" along the helix can be obtained by projecting the B DNA surface onto a cylinder and then unwrapping the cylindrical projection[150] (Figure 5). Inspection of this projection, or space filling models, illustrates several interesting features concerning the placement of DNA determinants implicated in specific recognition. For example, potential endonuclease contacts are accessible primarily from one side of the helix. Also particularly evident are potential contacts within the major groove between the points of cleavage of the two DNA strands. This representation also demonstrates the proximity (4 Å) of 6-NH_2 substituents of dA residues modified by *Eco* RI methylase.

Several lines of evidence considered above led to the suggestion that the *Eco* RI endonuclease dimer interacts with its recognition sequence to yield a complex possessing elements of two-fold symmetry. The alkylation experiments summarized in Figs. 4 and 5 confirm the presence of symmetry at the fine structure level and establish equivalence of interaction with the two strands of the six base pair *Eco* RI sequence. However, this work has also shown that the DNA binding site of the enzyme encompasses about 10 base pairs. As mentioned above, it is possible that interaction with nucleotides external to the recognition sequence may impose a certain degree of asymmetry on specific endonuclease-DNA complexes.

Since specific complexes between *Eco* RI methylase and DNA have not been

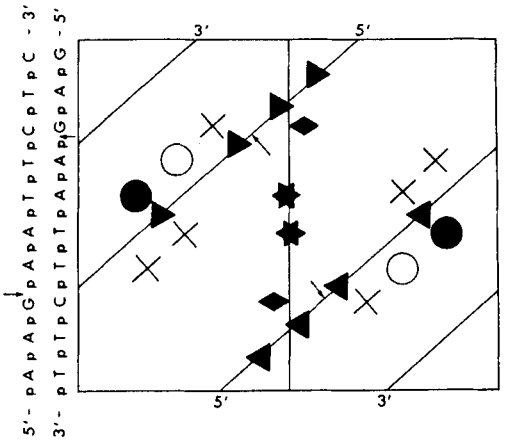

FIGURE 5. Potential contacts between *Eco* RI enzymes and their recognition sequence. The representation shown is based on a cylindrical projection of the DNA surface which has been opened along a line parallel to the helix axis.[150] Major and minor grooves are oriented diagonally. The helix repeat was assumed to be 10.5 base pairs per turn.[102] Points of endonuclease cleavage are indicated by arrows. *Endonuclease contacts* —▲— phosphates identified by virtue of ethylation interference; ●—N^3 of dA residues within the *Eco* RI site identified on the basis of methylation interference. ○—N^3 of dA residues which are protected against dimethylsulfate alkylation by the endonuclease, but which do not interfere with specific binding when methylated. ◆—N^7 of dG residues, the alkylation of which interferes with binding. ★—6-NH_2 of dA residues which are methylated by *Eco* RI modification enzyme. *Methylase* — Contacts identified by base analog substitution are indicated by X's. These are not required by the endonuclease. Those in the minor groove correspond to the 2-NH_2 of dG, those in the major groove of the 5-CH_3 of dT. Although the 5-CH_3 of both dT residues in each strand are indicated, only one may be required (text). The 6-NH_2 of dA residues indicated by ★ are sites of modification by the methylase and, hence, are clearly required by the enzyme.

demonstrated under noncatalytic conditions, attempts to elucidate DNA contacts utilized by this enzyme have involved analysis of analog effects on kinetics of methyl transfer. The results of these studies, which are summarized in Table 2 and Figure 5, are in marked contrast to those obtained with the endonuclease. Thus, bacteriophage T4 DNA, which is resistant to cleavage by the endonuclease, was found to be an excellent substrate for the modification enzyme.[105] This surprising finding has demonstrated that the presence of the bulky glucosylhydroxymethyl moiety in the major groove does not interfere with interaction between this enzyme and its recognition site. The methylase also differs from the endonuclease on dU-containing PBS2 DNA, which in the case of the modification enzyme is an extremely poor substrate.[105] However, it has not been determined whether the uracil for thymine replacement in this DNA affects site-specific binding or subsequent catalysis. Nevertheless, this result suggests that the 5-CH_3 of one or both dT residues within the *Eco* RI site has an important role in the methylase reaction. The substitution of dI within the *Eco* RI site of Col El DNA also had markedly

different effects on the two enzymes.[154] In contrast to results with the endonuclease, the dI-containing site was an extremely poor substrate for the modification enzyme, with the block being at the level of sequence recognition. Since analog substitution in these experiments was extremely limited and since steric effects can be excluded, it seems likely that 2-NH_2 of dG exposed in the minor groove is utilized as a contact by the methylase. The distribution of these potential methylase contacts on the B form structure is presented in Figure 5.

Examination of Table 2 and Figure 5 illustrates the striking differences between the two *Eco* RI enzymes with respect to DNA substituent effects on recognition and/or catalysis. The single exception may be a common requirement for 6-NH_2 groups of adenines adjacent to the twofold axis. This lends further support to the conclusion cited above that the two proteins interact with their common recognition sequence in different ways, and in particular, suggests that the two proteins employ different mechanisms to discriminate a given base pair.

B. *Hpa* I Restriction and Modification Enzymes

The only other Type II system that has been examined in some detail with respect to DNA substituent effects is that of *Hpa* I specificity (Table 1). Utilizing homogeneous preparations of *Hpa* I enzymes, Dwyer-Hallquist and Agarwal[156] have determined steady state kinetic parameters governing cleavage and methyl transfer to d(GpGpTpTpApApCpC) and derivatives of this self-complementary octanucleotide containing base analog substitutions.

The octanucleotide d(GpGpTpBUpApApCpC) was readily cleaved by *Hpa* I endonuclease, but the K_m and k_{cat} for this substrate were both reduced five-fold relative to the corresponding parameters for d(GpGpTpTpApApCpC). The parallel decrease in these parameters suggests that although the mechanism of cleavage of the BUdR-containing oligonucleotide may be similar to that of the control, some nonproductive interaction between the enzyme and the substrate analog may occur.[156] In contrast, d(GpGpTpUpApApCpC) and d(GpIpTpTpApApCpC) were not substrates for the endonuclease. While failure to cleave the dI-containing oligonucleotide may be at least partially due to reduced stability of the duplex form, the failure to utilize the dU-containing substrate analog implies a requirement for the 5-CH_3 of the dT residue adjacent to site of cleavage for recognition and/or catalysis by this endonuclease.[156]

Berkner and Folk[157] have also obtained evidence suggesting that the 5-CH_3 of thymine is important in the *Hpa* I endonuclease reaction. This work demonstrated that initial rates of double strand cleavage of PBS2 DNA, which contains dU in place of dT, were ten-fold lower than for λ DNA under comparable conditions. However, kinetics of double-strand cleavage of the dU containing substrate was anamolous, being biphasic with respect to time and enzyme concentration. Thus, at moderate enzyme concentrations, products arising as a consequence of double-strand cleavage were only observed after a lag of several hours, but their subsequent rate of appearance exceeded that of the λ control. A similar phenomenon was observed with hydroxymethyluracil-containing ϕe DNA. These kinetics were attributed to introduction of single strand scissions during early stages of reaction followed by their conversion to double strand breaks during the second phase.[157] Unfortunately this possibility was not directly addressed, rendering it difficult to compare the PBS2 and ϕe results with those obtained in the oligonucleotide experiments mentioned above. It is, nevertheless, not easy to reconcile results of these two approaches, although it may be pertinent to note that experiments with analog-containing viral DNAs employed preparations of *Hpa* I which were only partially purified, and sequences subject to cleavage were not established.

While analysis of phosphate contacts of *Hpa* I endonuclease is not yet complete, it has

Table 3
CLEAVAGE OF BASE-SUBSTITUTED DNAs BY RESTRICTION ENDONUCLEASES[a]

Endonuclease	Recognition sequence	Relative rate of DNA cleavage[b]			
		ϕe (HmdU)	PBS2 (dU)	T4 (GlcHmdC)	λ, P22(BUdR)
*Bam*H I	5′-G↓G-A-T-C-C	Poor	Good	Poor	Fair
Hha I	5′-G-C-G↓C	Good	Fair[c]	Inert	—
Hpa II	5′-C↓C-G-G	Good	Poor[c]	Inert	Good
Hind II	5′-G-T-Py↓Pu-A-C	Fair	Poor	Poor	Good
Hind III	5′-A↓A-G-C-T-T	Fair	Poor	Poor	Fair

[a] Adapted from Berkner and Folk.[157] Results of these workers with *Eco* RI[105] and *Hpa* I[157] enzymes are discussed in the text.

[b] Rates are relative to hydrolysis of λ DNA at comparable enzyme and DNA concentrates. Relative rates are Good (50 to 100%), Fair (14 to 33%), and Poor (≤ 10%). The last column summarizes results with BUdR-substituted λ or P22 DNAs.

[c] Reduced rates at least partially reflect a low density of cleavage sites.

been found that d(pGpGpTpTpApApCpC) is hydrolyzed much more rapidly by the enzyme than d(GpGpTpTpApApCpC).[156] This implicates phosphate residues 5′-external to the recognition sequence in enzyme-DNA interaction, a finding identical to that cited above for the *Eco* RI endonuclease.

In contrast to the behavior of *Hpa* I endonuclease with d(GpGpTpUpApApCpC) and d(GpGpTpBUpApApCpC), homogeneous *Hpa* I modification enzyme methylated both analog-containing oligonucleotides with steady state kinetic parameters which were almost identical to those for the control substrate d(GpGpTpTpApApCpC).[156] However, like the endonuclease, the methyltransferase failed to utilize d(GpIpTpTpApApCpC), although, as in the case of the restriction enzyme, this may at least partially reflect the reduced T_m of the dI-containing oligonucleotide.[156] The differential response of *Hpa* I endonuclease and methylase to dU and BUdR-containing oligonucleotides is reminiscent of analogous results with the *Eco* RI system (above). It would thus appear that in both systems, the restriction endonuclease and the modification methylase interact with their recognition sequence in different ways.

C. Other Type II Endonucleases

Only limited information is available on DNA determinants important in recognition or catalysis by other Type II restriction enzymes. Mann and Smith[50] have utilized a synthetic duplex oligonucleotide containing overlapping *Hpa* II and *Hae* III sites to examine the effects of abnormal methylation of the 5-position of dC on subsequent cleavage by the restriction endonuclease. For example, a portion of the oligonucleotide contained the sequence:

5′-CpCpGpGpCpC
Hpa II
Hae III

Hence, it was possible to assess effects of the *Hpa* II modification (Table 1) on *Hae* III restriction, and vice versa. Although only qualitative, these experiments demonstrated that *Hae* III endonuclease will cleave 5′-GpGpCpC as opposed to the normally modified

site 5′-GpGpĊpC, which is resistant. Similarly *Hpa* II was found to cleave 5′-ĊpCpGpG but not 5′-CpĊpGpG. These experiments are of interest for two reasons. First, they exclude endonuclease contacts with the 5-position of the dC residue which is not subject to modification in each system. Secondly, they render unlikely the possibility that resistance to restriction is a result of perturbation of DNA secondary structure which might occur (Section III) as a consequence of modification of this base.

Limited information on a number of systems has also been provided by Marchionni and Roufa[153] and Berkner and Folk[157] who have analyzed patterns of cleavage of viral DNAs which naturally contain base analogs or which have been prepared in such a way as to contain a BUdR for dT substitution. The results of these experiments, which are summarized in Table III, are somewhat difficult to interpret for several reasons. Being primarily survey in nature, these studies did not attempt to attribute rate variations with substrate to changes in kinetic parameters. Secondly, with the exception of BUdR-substituted DNAs, the number of recognition sites within an analog-containing viral molecule is unknown in the absence of a positive result. Hence, reduced rates of cleavage relative to the λ standard employed[157] may reflect a density of recognition sites less than that expected on statistical grounds.† For example, reduced rates of cleavage of PBS2 DNA (dU in place of dT) by *Hha* I and *Hpa* II, which do not contain dT in their recognition sequences, were attributed to such an effect.[157] However, the possibility that rate reductions were due to the presence of dU in outside sequences was not excluded. Indeed, BUdR substitution has been observed to reduce rates of cleavage by several endonucleases which recognize sequences lacking dT,[153,157] and effect which is clearly due to the presence of analog in external sequences. Because of these ambiguities, these studies will not be considered further here beyond noting that the differential response of the several enzymes to analog substitution (Table III) is consistent with the view that discrimination of a given base pair can occur by different mechanisms.

D. Other Type II Methylases

Mann and Smith[52] have utilized the random copolymer $d(G, C)_n$ and analogs of this polynucleotide as substrates for partially purified *Hha* I, *Hpa* II, and *Hae* III modification enzymes. While random $d(G, C)_n$ was methylated by all three enzymes, the $d(I, C)_n$ analog supported methylation by *Hha* I and *Hpa* II enzymes only at a low rate. In contrast, the rate of methyl transfer to $d(I, C)_n$ by the *Hae* III enzyme was comparable to that observed with $d(G, C)_n$. Although steady state kinetic parameters were not evaluated, these findings suggest that like the *Eco* RI enzyme, *Hha* I and *Hpa* I methylases require the 2-NH_2 of dG while the *Hae* III enzyme evidently does not. The activity of these three methyltransferases were also examined with random $d(N\text{-acetylG}, C)_n$ and $d(X, C)_n$.[53] While rates of methyl transfer to these copolymers by the *Hha* I enzyme were comparable to those observed with $d(G, C)_n$, these analog polymers were poor substrates for the *Hpa* II and *Hae* III activities. Since $d(N\text{-acetylG}, C)_n$ and $d(X, C)_n$ copolymers employed did not display thermal melting transitions, suggesting the absence of stable secondary structure,[53] it is not clear whether the reduced activity of the latter enzymes on such substrates is due to perturbation of DNA secondary structure or to alteration of a DNA contact.

† This was not a problem with the *Eco* RI studies employing PBS2 considered earlier. Although the methylase was inert on PBS2 DNA, the number of *Eco* RI sites was readily monitored by virtue of sensitivity of the viral DNA to *Eco* RI cleavage.[105]

CONCLUSION

While substantial information is now available concerning modes of interaction of sequence-specific proteins with DNA, our understanding of the molecular mechanisms involved in this phenomenon is still at an early stage. Several problems are apparent from the enzymological point of view. A central theme that has emerged from the work reviewed here is that multiple mechanisms exist for discrimination of base pairs. Hence, elucidation of the set of protein-DNA bonding schemes that may be employed in sequence discrimination will require analysis of multiple systems.

The mechanisms which allow Type II enzymes to discriminate their canonical sequence from the relatively high background of sites which differ from the canonical sequence at only one position remains unresolved. While this problem must be considered in the case of any sequence-specific protein, it can to some extent be circumvented in the case of those proteins which recognize single copy sequences by overspecification of the recognition site.[117] Thus, while a 12 base pair recognition site would be expected to comprise a statistically unique sequence within the *E. coli* chromosome, the binding site for the *lac* repressor appears to encompass about 17 base pairs (reviewed in Reference 117).

It is evident from the work reviewed here, as well as from analysis of other systems, that although a large amount of information has been accumulated concerning potential DNA contacts utilized during sequence discrimination by proteins, little or no information is available on peptide determinants involved in their DNA binding sites. This aspect of the problem is certainly amenable to study in several systems. It could be argued that this information will be forthcoming with the crystallographic analysis of site-specific proteins and complexes between such proteins and recognition site oligonucleotides. It must be recalled, however, that such proteins also exhibit nonspecific modes of interaction with DNA. Hence, interpretation of co-crystal structures as representing bona fide site-specific complexes must rely on results of solution biochemistry for requisite criteria on which to base such conclusions.

ACKNOWLEDGMENTS

I am indebted to Kan Agarwal, Patricia Greene, Sung-Hou Kim, Richard Roberts, Hamilton Smith, and Peter von Hippel for useful discussions and communication of unpublished results, and to Vickers Burdett for comments on the manuscript. I also wish to acknowledge Soo-Chen Cheng, William Jack, A-Lien Lu, Andrea Newman, and Robert Rubin for their ideas and contributions to the studies performed in our laboratory. Our work has been supported by grants from the National Institutes of Health and the National Science Foundation.

REFERENCES

1. **Meselson, M. and Yuan, R.,** DNA restriction enzyme from *E. coli, Nature,* 217, 1110, 1968.
2. **Linn, S. and Arber, W.,** In vitro restriction of phage fd replicative form, *Proc. Natl. Acad. Sci. U.S.A.,* 59, 1300, 1968.
3. **Roberts, R. J.,** Restriction endonucleases, *CRC Crit. Rev. Biochem.,* 4, 123, 1976.
4. **Roberts, R. J.,** Restriction and modification enzymes and their recognition sequences, *Nucl. Acids Res.,* 8, r63, 1980.
5. **Boyer, H. W.,** DNA restriction and modification mechanisms in bacteria, *Annu. Rev. Microbiol.,* 25, 153, 1971.
6. **Yuan, R.,** The structure and mechanism of multifunctional restriction endonucleases, *Annu. Rev. Biochem.,* 50, 285, 1981.

7. **Smith, H. O. and Nathans, D.,** A suggested nomenclature for bacterial host modification and restriction systems and their enzymes, *J. Mol. Biol.*, 81, 419, 1973.
8. **Modrich, P.,** Structures and mechanisms of DNA restriction and modification enzymes, *Q. Rev. Biophys.*, 12, 315, 1979.
9. **Smith, H. O. and Wilcox, K. W.,** A restriction enzyme from *Haemophilus influenzae*. I. Purification and general properties, *J. Mol. Biol.*, 51, 379, 1970.
10. **Kelly, T. J. and Smith, H. O.,** A restriction enzyme from *Haemophilus influenzae*. II. Base sequence of the recognition site, *J. Mol. Biol.*, 51, 393, 1970.
11. **Roy, P. H. and Smith, H. O.,** DNA methylases of *Haemophilus influenzae* Rd. I. Purification and properties, *J. Mol. Biol.*, 81, 427, 1973.
12. **Roy, P. H. and Smith, H. O.,** DNA methylases of *Haemophilus influenzae* Rd. II. Partial recognition site base sequences, *J. Mol. Biol.*, 81, 445, 1973.
13. **Wilson, G. A. and Young, F. E.,** Isolation of a sequence-specific endonuclease (*Bam* I) from *Bacillus amyloliquefaciens, J. Mol. Biol.*, 97, 123, 1975.
14. **Roberts, R. J., Wilson, G. A., and Young, F. E.,** Recognition sequence of specific endonuclease *Bam* HI from *Bacillus amyloliquefaciens* H, *Nature*, 265, 82, 1977.
15. **Hattman, S., Keisler, T. and Gottehrer, A.,** Sequence specificity of DNA methylases from *Bacillus amyloliquefaciens* and *Bacillus brevis, J. Mol. Biol.*, 124, 701, 1978.
16. **Smith, L. A. and Chirikjian, J. G.,** Purification and characterization of the sequence-specific endonuclease *Bam* HI. *J. Biol. Chem.*, 254, 1003, 1979.
17. **Bingham, A. H. A., Atkinson, T., Sciaky, D., and Roberts, R. J.,** A specific endonuclease from *Bacillus caldolyticus, Nucl. Acids Res.*, 5, 3457, 1978.
18. **Lee, Y. H. and Chirikjian, J. G.,** Sequence-specific endonuclease *Bgl* I: modification of lysine and arginine residues of the homogeneous enzyme, *J. Biol. Chem.*, 254, 6838, 1979.
19. **van Heuverswyn, H. and Fiers, W.,** *Gene*, 9, 195, 1980.
20. **Bickle, T. A. and Ineichen, K.,** *Gene*, 9, 205, 1980.
21. **Lautenbergen, J. A., White, C. T., Haigwood, N. L., Edgell, M. H., and Hutchison, C. A. III,** *Gene*, 9, 213, 1980.
22. **Pirotta, V.,** Two restriction endonucleases from *Bacillus globiggi, Nucl. Acids Res.*, 3, 1747, 1976.
23. **Bickle, T. A., Pirotta, V., and Imber, R.,** Purification and properties of the *Bgl* I and *Bgl* II endonucleases, Part I, *Methods Enzymol.*, 65, 132, 1980.
24. **Kiss, A., Sain, B., Csordas-Toth, E., and Venetianer, P.,** A new sequence-specific endonuclease (Bsp) from *Bacillus sphaericus, Gene*, 1, 323, 1977.
25. **Koncz, C., Kiss, A., and Venetianer, P.,** Biochemical characterization of the restriction-modification system of *Bacillus sphaericus, Eur. J. Biochem.*, 89, 523, 1978.
26. **Catterall, J. F. and Welker, N. E.,** Isolation and properties of a thermostable restriction endonuclease (Endo R. *Bst* 1503), *J. Bact.*, 129, 1110, 1977.
27. **Levy, W. P. and Welker, N. E.,** Purification and properties of a modification methylase from *Bacillus stearothermophilus, Fed. Proc.*, 37, 1414, 1978.
28. **Clarke, C. M. and Hartley, B. S.,** Purification, properties, and specificity of the restriction endonuclease from *Bacillus stearothermophilus, Biochem. J.*, 177, 49, 1979.
29. **Bron, S., Murray, K., and Trautner, T. A.,** Restriction and modification in *B. subtilis*. Purification and general properties of a restriction endonuclease from strain R, *Mol. Gen. Genet.*, 143, 13, 1975.
30. **Bron, S. and Murray, K.,** Restriction and modification in *B. subtilis*. Nucleotide sequence recognized by restriction endonuclease R · *Bsu* R from strain R, *Mol. Gen. Genet.*, 143, 25, 1975.
31. **Heininger, K., Hörz, W. and Zachau, H. G.,** Specificity of cleavage by a restriction nuclease from *Bacillus subtilis, Gene*, 1, 291, 1977.
32. **Bron, S., and Hörz, W.,** Purification and properties of *Bsu* I endonuclease, Part I, *Methods Enzymol.*, 65:112, 1980.
33. **Geier, G. E. and Modrich, P.,** Recognition sequence of the *dam* methylase of *Escherichia coli* K12 and mode of cleavage of *Dpn* I endonuclease, *J. Biol. Chem.*, 254, 1408, 1979.
34. **Herman, G. E.,** Ph.D. thesis, Duke University, 1981.
35. **Lacks, S. and Greenberg, B.,** A deoxyribonuclease of *Diplococcus pneumoniae* specific for methylated DNA, *J. Biol. Chem.*, 250, 4060, 1975.
36. **Lacks, S. and Greenberg, B.** Complementary specificity of restriction endonucleases of *Diplococcus pneumoniae* with respect to DNA methylation, *J. Mol. Biol.*, 114, 153, 1977.
37. **Lacks, S. A.,** Purification and properties of the complementary endonuclease *Dpn* I and *Dpn* II, Part I, *Methods Enzymol.*, 65, 138, 1980.
38. **Hedgpeth, J., Goodman, H. M., and Boyer, H. W.,** DNA nucleotide sequence restricted by the RI endonuclease, *Proc. Natl. Acad. Sci. U.S.A.*, 69, 3448, 1972.

39. **Dugaiczyk, A., Hedgpeth, J., Boyer, H. W. and Goodman, H. M.,** Physical identity of the SV40 deoxyribonucleic acid sequence recognized by the *Eco* RI restriction endonuclease and modification methylase, *Biochemistry,* 13, 503, 1974.
40. **Greene, P. J., Poonian, M. S., Nussbaum, A. L., Tobias, L., Garfin, D. E., Boyer, H. W., and Goodman, H. M.,** Restriction and modification of a self-complementary octanucleotide containing the *Eco* RI site, *J. Mol. Biol.,* 99, 237, 1975.
41. **Modrich, P. and Zabel, D.,** *Eco* RI endonuclease: physical and catalytic properties of the homogeneous enzyme, *J. Biol. Chem.,* 251, 5866, 1976.
42. **Rubin, R. A. and Modrich, P.,** *Eco* RI methylase: physical and catalytic properties of the homogeneous enzyme, *J. Biol. Chem.,* 252, 7265, 1977.
43. **Jack, W. E., Rubin, R. A., Newman, A., and Modrich, P.,** Structures and mechanisms of *Eco* RI DNA restriction and modification enzymes, in *Gene Amplification and Analysis,* Vol. 1: *Restriction Endonucleases,* Chirikjian, J. G., Ed., Elsevier-North Holland, 1980, 165.
44. **Newman, A. K., Rubin, R. A., Kim, S. H., and Modrich, P.,** DNA sequences of structural genes for *Eco* RI DNA restriction and modification enzymes, *J. Biol. Chem.,* 256, 2131, 1981.
45. **Greene, P. J., Gupta, M., Boyer, H. W., Brown, W. E. and Rosenberg, J. M.,** Sequence analysis of the DNA encoding *Eco* RI endonuclease and methylase, *J. Biol. Chem.,* 256, 2143, 1981.
46. **Bingham, A. H. A. and Atkinson, A.,** Restriction endonucleases and modification methylases in bacteria, *Biochem. Soc. Trans.,* 6, 315, 1978.
47. **Bigger, C. H., Murray, K. and Murray, N.,** Recognition sequence of a restriction enzyme, *Nature New Biol.,* 244, 7, 1973.
48. **Roulland-Dussoix, D., Yoshimori, R., Greene, P., Betlach, M., Goodman, H. M., and Boyer, H.,** R factor-controlled restriction and modification of deoxyribonucleic acid, in *American Society for Microbiology Conference on Bacterial Plasmids, Microbiology,* ASM Press, 1974, 187.
49. **Middleton, J. H., Edgell, M. H. and Hutchison, C. A. III,** Specific fragments of ϕX174 deoxyribonucleic acid produced by a restriction enzyme from *Haemophilus aegyptius,* endonuclease Z, *J. Virol.,* 10, 42, 1972.
50. **Mann, M. B. and Smith, H. O.,** Specificity of *Hpa* II and *Hae* III DNA methylases, *Nucl. Acids Res.,* 4, 4211, 1977.
51. **Roberts, R. J., Myers, P. A., Morrison, A., and Murray, K.,** A specific endonuclease from *Haemophilus haemolyticus, J. Mol. Biol.,* 103, 199, 1976.
52. **Mann, M. B. and Smith, H. O.,** Specificity of DNA methylases from *Haemophilus* sp., in *Transmethylation,* Usdin, E., Borchardt, R., and Kreveling, C., Eds., Elsevier-North Holland, 1979, 483.
53. **Mann, M. B., Rao, R. N. and Smith, H. O.,** Cloning of restriction and modification genes in *E. coli:* the *Hha* II system from *Haemophilus haemolyticus, Gene,* 3, 97, 1978.
54. **Mann, M. B. and Smith, H. O.,** The *Hha* II restriction-modification system, *Fed. Proc.,* 38, 293, 1979.
55. **Old, R., Murray, K., and Roizes, J.,** Recognition sequence of restriction endonuclease III from *Haemophilus influenzae, J. Mol. Biol.,* 92, 331, 1975.
56. **Kleid, D., Humayun, Z., Jeffrey, A. and Ptashne, M.,** Novel properties of a restriction endonuclease isolated from *Haemophilus parahaemolyticus, Proc. Natl. Acad. Sci. U.S.A.,* 73, 293, 1976.
57. **Sharp, P. A., Sugden, B. and Sambrook, J.,** Detection of two restriction endonuclease activities in *Haemophilus parainfluenzae* using analytical agarose-ethidium bromide electrophoresis, *Biochemistry,* 12, 3055, 1973.
58. **Garfin, D. E. and Goodman, H. M.,** Nucleotide sequences at the cleavage sites of two restriction endonucleases from *Haemophilus parainfluenzae, Biochim. Biophys. Res. Commun.,* 59, 108, 1974.
59. **Hines, J. L. and Agarwal, K. L.,** Purification of *Hpa* I and *Hpa* II, *Fed. Proc.,* 38, 294, 1979.
60. **Agarwal, K. L.,** personal communication.
61. **Yoo, O. J. and Agarwal, K. L.,** Isolation and characterization of two proteins possessing *Hpa* II methylase activity, *J. Biol. Chem.,* 255, 6445, 1980.
62. **Baumstark, B. R., Roberts, R. J. and RajBhandary, U. L.,** Use of short synthetic DNA duplexes as substrates for the restriction endonucleases *Hpa* II and *Mno* I, *J. Biol. Chem.,* 254, 8943, 1979.
63. **Van Montagu, M., Sciaky, D., Myers, P. A., and Roberts, R. J.,** unpublished experiments.
64. **Yoo, O. J. and Agarwal, K. L.,** Cleavage of single strand oligonucleotides and bacteriophage ϕX174 DNA by *Msp* I endonuclease, *J. Biol. Chem.,* 255, 10559, 1980.
65. **Sussenbach, J. S., Monfoort, C. H., Schiphof, R., and Stobberingh, E. E.,** A restriction endonuclease from *Staphylococcus aureus, Nucleic Acids Res.,* 3, 3193, 1976.
66. **Riggs, A. D., Bourgeois, S., and Cohen, M.,** The *lac* repressor-operator interaction. III. Kinetic studies, *J. Mol. Biol.,* 53, 401, 1970.
67. **Lin, S. and Riggs, A. D.,** *lac* repressor binding to non-operator DNA: detailed studies and comparison of equilibrium and rate competition methods, *J. Mol. Biol.,* 72, 671, 1972.
68. **Richter, P. H. and Eigen, M.,** Diffusion controlled reaction rates in spheroidal geometry. Application to repressor-operator association and membrane bound enzymes, *Biophys. Chem.,* 2, 255, 1974.

69. **Schranner, R. and Richter, P. H.,** Rate enhancement by guided diffusion. Chain length dependence of repressor-operator association rates, *Biophys. Chem.*, 8, 135, 1978.
70. **Rubin, R. A. and Modrich, P.,** Purification and properties of *Eco* RI endonuclease, Part I, *Methods Enzymol.*, 65, 96, 1980.
71. **Marinus, M. G. and Morris, N. R.,** Isolation of Deoxyribonucleic acid methylase mutants of *Escherichia coli* K-12, *J. Bacteriol.*, 114, 1143, 1973.
72. **Humphreys, G. O., Willshaw, G. A., Smith, H. R., and Anderson, E. S.,** Mutagenesis of plasmid DNA with hydroxylamine: isolation of mutants of multi-copy plasmids, *Mol. Gen. Genet.*, 145, 101, 1976.
73. **Boyer, H. W., Greene, P. J., Meagher, R. B., Betlach, M. C., Russel, D., and Goodman, H. M.,** FEBS Symposium Vol. 34, Akademiai Kiado, Budapest, Hungary, 1975, 23.
74. **Rubin, R. A., Modrich, P., and Vanaman, T. C.,** Partial amino and carboxyl terminal sequence analyses of *Eco* RI DNA restriction and modification enzymes, *J. Biol. Chem.*, 256, 2140, 1981.
75. **Chou, P. Y. and Fasman, G. D.,** Prediction of the secondary structure of proteins from their amino acid sequence, *Adv. Enzymol.*, 47, 45, 1978.
76. **Rosenberg, J. M., Dickerson, R. E., Greene, P. J., and Boyer, H. W.,** Preliminary X-ray diffraction analysis of crystalline *Eco* RI endonuclease, *J. Mol. Biol.*, 122, 241, 1978.
77. **Young, T.-S., Modrich, P., Jay, E., and Kim, S.-H.,** Preliminary X-ray diffraction studies of *Eco* RI restriction endonuclease-DNA complex, *J. Mol. Biol.*, 145, 607, 1981.
78. **Hausler, B.,** Mapping and sequencing of cloned *Hha* II restriction and modification genes, *Fed. Proc.*, 39, 1784, 1980.
79. **Hausler, B. and Smith, H. O.,** unpublished experiments.
80. **Roberts, R. J.,** unpublished experiments.
81. **Dickson, R. C., Abelson, J., Barnes, W. M., and Reznikoff, W. S.,** Genetic regulation: the *Lac* control region, *Science*, 87, 27, 1975.
82. **Jovin, T. M.,** Recognition mechanisms of DNA-specific enzymes, *Annu. Rev. Biochem.*, 45, 889, 1976.
83. **Wells, R. D., Goodman, T. C., Hillen, W., Horn, G. T., Klein, R. D., Larson, J. E., Müller, U. R., Neuendorf, S. K., Panayotatos, N., and Stirdivant, S. M.,** DNA structure and gene regulation, *Prog. Nucl. Acid Res. Mol. Biol.*, 24, 168, 1980.
84. **Gierer, A.,** Model for DNA and protein interactions and the function of the operator, *Nature*, 212, 1480, 1966.
85. **Sobell, H. M.,** Molecular mechanism of genetic recombination, *Proc. Natl. Acad. Sci. U.S.A.*, 69, 2483, 1972.
86. **Lilley, D. M. J.,** The inverted repeat as a recognizable structural feature in supercoiled DNA molecules, *Proc. Natl. Acad. Sci. U.S.A.*, 77, 6468, 1980.
87. **Panayotatos, N. and Wells, R. D.,** Cruciform structures in supercoiled DNA, *Nature*, 289, 466, 1981.
88. **Lim, V.I. and Mazanov, A.L.,** Tertiary structure for palindromic regions of DNA, *FEBS Lett.*, 88, 118, 1978.
89. **Stasiak, A. and Klopotowski, T.,** Four-stranded DNA structure and DNA base methylation in the mechanism of action of restriction endonucleases, *J. Theor. Biol.*, 80, 65, 1979.
90. **McGavin, S.,** Models of specifically paired like (homologous) nucleic acid structures, *J. Mol. Biol.*, 55, 293, 1971.
91. **Brown, N. L. and Smith, M.,** Cleavage specificity of the restriction endonuclease isolated from *Haemophilus gallinarum (Hga* I), *Proc. Natl. Acad. Sci. U.S.A.*, 74, 3213, 1977.
92. **Sugisaki, H.,** Recognition sequence of a restriction endonuclease from *Haemophilus gallinarum, Gene*, 3, 17, 1978.
93. **Gelinas, R. E., Myers, P. A., and Roberts, R. J.,** Two sequence-specific endonucleases from *Moraxella bovis, J. Mol. Biol.*, 114, 169, 1977.
94. **Brown, N. L., Hutchison, C. A., and Smith, M.,** The specific non-symmetrical sequence recognized by restriction endonuclease *Mbo* II, *J. Mol. Biol.*, 140, 143, 1980.
95. **Wang, J. C., Barkley, M. D., and Bourgeois, S.,** Measurements of unwinding of lac operator by repressor, *Nature*, 251, 247, 1974.
96. **Rubin, R. A. and Modrich, P.,** unpublished experiments.
97. **von Hippel, P. H. and McGhee, J. D.,** DNA-protein interactions, *Annu. Rev. Biochem.*, 41, 231, 1972.
98. **Seeman, N. C., Rosenberg, J. M., and Rich, A.,** Sequence-specific recognition of double helical nucleic acids by proteins, *Proc. Natl. Acad. Sci. U.S.A.*, 73, 804, 1976.
99. **Wing, R., Drew, H., Takano, T., Broka, C., Tanaka, S., Itakura, K., and Dickerson, R. E.,** Crystal structure analysis of a complete turn of B-DNA, *Nature*, 287, 755, 1980.
100. **Levitt, M.,** How many base pairs per turn does DNA have in solution and in chromatin? Some theoretical calculations, *Proc. Natl. Acad. Sci. U.S.A.*, 75, 640, 1978.
101. **Hogan, M., Dattagupta, N., and Crothers, D. M.,** Transient electric dichroism of rod-like DNA molecules, *Proc. Natl. Acad. Sci. U.S.A.*, 75, 195, 1978.
102. **Wang, J. C.,** Helical repeat of DNA in solution, *Proc. Natl. Acad. Sci. U.S.A.*, 76, 200, 1979.

103. **Engel, J. D. and von Hippel, P. H.,** Effects of methylation on stability of nucleic acid conformations: studies at the polymer level, *J. Biol. Chem.*, 253, 927, 1978.
104. **Cheng, S. C. and Modrich, P.,** unpublished experiments.
105. **Berkner, K. L. and Folk, W. R.,** *Eco* RI cleavage and methylation of DNAs containing modified pyrimidines in the recognition sequence, *J. Biol. Chem.*, 252, 3185, 1977.
106. **Jack, W. E. and Modrich, P.** unpublished experiments.
107. **Ruben, G., Spielman, P., Tu, C. D., Jay, E., Siegel, B., and Wu, R.,** Relaxed circular SV40 DNA as cleavage intermediate of two restriction endonucleases, *Nucl. Acids Res.*, 4, 1803, 1977.
108. **Rubin, R. A. and Modrich, P.,** Substrate dependence of the mechanism of *Eco* RI endonuclease, *Nucl. Acids Res.*, 5, 2991, 1978.
109. **Halford, S. E., Johnson, N. P., and Grinstead, J.,** The reactions of *Eco* RI and other restriction endonucleases, *Biochem. J.*, 179, 353, 1979.
110. **Langowski, J., Pingoud, A., Goppelt, M., and Maas, G.,** Inhibition of *Eco* RI by polynucleotides. A characterization of non-specific binding of the enzyme to DNA, *Nucl. Acids Res.*, 8, 4727, 1980.
111. **Thomas, M. and Davis, R. W.,** Studies on the cleavage of bacteriophage lambda DNA with *Eco* RI restriction endonuclease, *J. Mol. Biol.*, 91, 315, 1975.
112. **Forsblum, S., Rigler, R., Ehrenberg, M., Pettersson, U., and Philipson, L.,** Kinetic studies on the cleavage of adenovirus DNA by restriction endonuclease *Eco* RI, *Nucl. Acids Res.*, 3, 3255, 1976.
113. **Halford, S. E., Johnson, N. P., and Grinsted, J.,** The *Eco* RI restriction endonuclease with bacteriophage λ DNA. Kinetic studies, *Biochem. J.*, 191, 581, 1980.
114. **Vovis, G. F. and Locks, S.,** Complementary action of restriction enzymes Endo R · *Dpn* I and Endo R · *Dpn* II on bacteriophage fl DNA, *J. Mol. Biol.*, 115, 525, 1977.
115. **Streeck, R. E.,** Single-strand and double-strand cleavage at half-modified and fully modified recognition sites for the restriction nucleases *Sau* 3A and *Taq* 1, *Gene*, in press.
116. **Vovis, G. F., Horiuchi, K., and Zinder, N. D.,** Kinetics of methylation by a restriction endonuclease from *E. coli* B, *Proc. Natl. Acad. Sci. U.S.A.*, 71, 3810, 1974.
117. **von Hippel, Peter H.,** On the molecular bases of the specificity of interaction of transcriptional proteins with genome DNA, *Biol. Recog. Devel.*, 1, 279, 1979.
118. **Murray, N. E. and Murray, K.,** Manipulation of restriction targets in phage λ to form receptor chromosomes for DNA fragments, *Nature*, 251, 476, 1974.
119. **Thomas, M., Cameron, J. R., and Davis, R. W.,** Viable molecular hybrids of bacteriophage lambda and eukaryotic DNA, *Proc. Natl. Acad. Sci. U.S.A.*, 71, 4579, 1974.
120. **Woodbury, C. P., Downey, R. L., and von Hippel, P. H.,** DNA site recognition and overmethylation by the *Eco* RI methylase, *J. Biol. Chem.*, 255, 11526, 1980.
121. **Bishop, J. O.,** A DNA sequence cleaved by restriction endonuclease R · *Eco* RI in only one strand, *J. Mol. Biol.*, 128, 545, 1979.
122. **Bishop, J. O. and Davies, J. A.,** Plasmid cloning vectors that can be nicked at a unique site, *Mol. Gen. Genet.*, 179, 573, 1980.
123. **Maxam, A. M. and Gilbert, W.,** Sequencing end-labeled DNA with base-specific chemical cleavages, Part I, *Methods Enzymol*, 65, 499, 1980.
124. **Sutcliffe, J. G.,** personal communication.
125. **Polisky, B., Greene, P., Garfin, D. E., McCarthy, B. J., Goodman, H. M., and Boyer, H. W.,** Specificity of substrate recognition by the *Eco* RI restriction endonuclease, *Proc. Natl. Acad. Sci. U.S.A.*, 72, 3310, 1975.
126. **Hsu, M. and Berg, P.,** Altering the specificity of restriction endonuclease: effect of replacing Mg^{2+} with Mn^{2+}, *Biochemistry*, 17, 131, 1978.
127. **Tikchonenko, T. I., Karamov, E. V., Zavizion, B. A., and Naroditsky, B. S.,** *Eco* RI* activity: enzyme modification or activation of accompanying endonuclease, *Gene*, 4, 195, 1978.
128. **Woodbury, C. P., Hagenbüchle, O., and von Hippel, P. H.,** DNA site recognition and reduced specificity of the *Eco* RI endonuclease, *J. Biol. Chem.*, 255, 11534, 1980.
129. **Malyguine, E., Vannier, P., and Yot, P.,** Alteration of the specificity of restriction endonucleases in the presence of organic solvents, *Gene*, 8, 163, 1980.
130. **George, J., Blakesley, R. W., and Chirikjian, J. G.,** Sequence-specific endonuclease *Bam* HI: effect of hydrophobic reagents on sequence recognition and catalysis, *J. Biol. Chem.*, 255, 6521, 1980.
131. **Sanger, F., Coulson, A. R., Friedmann, T., Air, G. M., Barrell, B. G., Brown, N. L., Fiddes, J. C., Hutchison III, C. A., Slocombe, P. M., and Smith, M.,** The nucleotide sequence of bacteriophage φX174, *J. Mol. Biol.*, 125, 255, 1978.
132. **Greene, P. J.,** unpublished experiments.
133. **Berkner, K. L. and Folk, W. R.,** Overmethylation of DNAs by the *Eco* RI methylase, *Nucleic Acids Res.*, 5, 435, 1978.
134. **Woodbury, C. P., Downey, R. L., and von Hippel, P. H.,** DNA site recognition and overmethylation by the *Eco* RI methylase, *J. Biol. Chem.*, 255, 11526, 1980.

135. **Behe, M. and Felsenfeld, G.,** Effects of methylation on a synthetic polynucleotide: The B-Z transition in poly(dG-m^{5}dC) · poly(dG-m^{5}dC), *Proc. Natl. Acad. Sci. U.S.A.,* 78, 1619, 1981.
136. **Horiuchi, K. and Zinder, N. D.,** Site-specific cleavage of single-stranded DNA by a *Haemophilus* restriction endonuclease, *Proc. Natl. Acad. Sci. U.S.A.,* 72, 2555, 1975.
137. **Blakesley, R. W. and Wells, R. D.,** "Single-stranded" DNA from ϕX174 and M13 is cleaved by certain restriction endonucleases, *Nature,* 257, 421, 1975.
138. **Godson, G. N. and Roberts, R. J.,** A catalogue of cleavages of ϕX174, S13, G4, and ST-1 DNA by 26 different restriction endonucleases, *Virology,* 73, 561, 1976.
139. **Blakesley, R. W., Dodgson, J. B., Nes, I. F., and Wells, R. D.,** Duplex regions in "Single-stranded" ϕX174 DNA are cleaved by a restriction endonuclease from *Haemophilus aegyptius, J. Biol. Chem.,* 252, 7300, 1977.
140. **Molloy, P. L. and Symons, R. H.,** Cleavage of DNA · RNA hybrids by Type II restriction enzymes, *Nucleic Acids Res.,* 8, 2939, 1980.
141. **Halford, S. E. and Johnson, N. P.,** The *Eco* RI restriction endonuclease with bacteriophage λ DNA. Equilibrium binding studies, *Biochem. J.,* 191, 593, 1980.
142. **Rosenberg, J. M., Boyer, H. W. and Greene, P.,** The structure and function of the *Eco* RI restriction endonuclease, in *Gene Amplification and Analysis, Vol. I: Restriction Endonucleases,* Chirikjian, J. G., Ed., Elsevier-North Holland, 1980, 131.
143. **Woodhead, J. L. and Malcolm, A. D. B.,** Non-specific binding of restriction endonuclease *Eco* RI to DNA, *Nucleic Acids Res.,* 8, 389, 1980.
144. **Langowski, J,., Pingoud, A., Goppelt, M., and Maass, G.,** Inhibition of *Eco* RI action by polynucleotides. A characterization of the non-specific binding of the enzyme to DNA, *Nucleic Acids Res.,* 8, 4727, 1980.
145. **Goppelt, M., Pingoud, A., Maass, G., Mayer, H., Koster, H. and Frank, R.,** The interaction of the *Eco* RI endonuclease with its substrate. A physico-chemical study employing natural and synthetic oligonucleotides and polynucleotides, *Eur. J. Biochem.,* 104, 101, 1980.
146. **Jack, W. E. and Modrich, P.,** in preparation.
147. **von Hippel, P. H.,** personal communication.
148. **Gilbert, W., Maxam, A. and Mirzabekov, A.,** Control of ribosome synthesis, in *Alfred Benzon Symposium XIII,* Kjeldgaard, N. O. and Maaloe, O., Eds., Copenhagen: Munksgaard, 139.
149. **Singer, B.,** The chemical effects of nucleic acid alkylation and their relation to mutagenesis and carcinogenesis, *Prog. Nucl. Acid Res. Mol. Biol.,* 15, 219, 1975.
150. **Siebenlist, U., Simpson, R. B., and Gilbert, W.,** *E. coli* RNA polymerase interacts homologously with two different promoters, *Cell,* 20, 269, 1980.
151. **Sun, L. and Singer, B.,** The specificity of different classes of ethylating agents toward various sites of HeLa cell DNA in vitro and in vivo, *Biochemistry,* 14, 1795, 1975.
152. **Kaplan, D. A. and Nierlich, D. P.,** Cleavage of nonglucosylated bacteriophage T4 deoxyribonucleic acid by restriction endonuclease *Eco* RI, *J. Biol. Chem.,* 250, 2395, 1975.
153. **Marchionni, M. A. and Roufa, D. J.,** Digestion of 5-bromodeoxyuridine-substituted λ-DNA by restriction endonucleases, *J. Biol. Chem.,* 253, 9075, 1978.
154. **Modrich, P. and Rubin, R. A.,** Role of the 2-amino group of deoxyguanosine in sequence recognition by *Eco* RI restriction and modification enzymes, *J. Biol. Chem.,* 252, 7273, 1977.
155. **Lu, A.-L., Jack, W. E., and Modrich, P.,** DNA determinants important in sequence recognition by *Eco* RI endonuclease, *J. Biol. Chem.,* 256, 13200, 1981.
156. **Dwyer-Hallquist, P. and Agarwal, K. L.,** unpublished experiments.
157. **Berkner, K. L. and Folk, W. R.,** The effects of substituted pyrimidines in DNAs on cleavage by sequence-specific endonucleases, *J. Biol. Chem.,* 254, 2551, 1979.