Mechanisms of Mutagenesis by Exocyclic DNA Adducts. Transfection of M13 Viral DNA Bearing a Site-Specific Adduct Shows That Ethenocytosine Is a Highly Efficient RecA-Independent Mutagenic Noninstructional Lesion†

Vaseem A. Palejwala, Devendranath Simha, and M. Zafri Humayun*

Department of Microbiology and Molecular Genetics, University of Medicine and Dentistry of New Jersey-New Jersey Medical School, 185 South Orange Avenue MSB F607, Newark, New Jersey 07103-2714

Received April 17, 1991; Revised Manuscript Received June 26, 1991

ABSTRACT: It is widely accepted that mutagenic DNA lesions fall into two categories: *mispairing* lesions hydrogen bond with an incorrect incoming base, generally do not stop replication, and possess high mutagenic efficiency without any requirement for induced functions; *noninstructional* lesions lack accessible template information, act as strong blocks to DNA replication (and are therefore toxic), and their mutagenic effects are SOS-dependent. Our recent results show that ethenocytosine (εC), a noninstructional exocyclic DNA lesion induced by vinyl chloride, may have unusual mutagenic properties. To obtain more definitive experimental evidence for the observed effects, we have introduced a single εC residue at a specific site of coliphage M13AB28 replicative form DNA by a "single-stranded linker-ligation" technique. The resulting DNA was purified and transfected into appropriate *rec*A$^+$ or *rec*A$^-$ *Escherichia coli* host cells. The effect of εC on survival was determined from transfection efficiency. Both the frequency and specificity of mutations induced by εC were determined by direct sequence analysis of randomly picked progeny phage plaques. The results indicated that εC has little effect on the survival of M13 DNA. Approximately 30% of the progeny phage obtained by transfecting εC DNA had a base substitution mutation precisely at the lesion site. No such mutations were observed in progeny plaques obtained by transfecting the control DNA construct. All εC-induced mutations were either C-to-T transitions or C-to-A transversions. Neither survival nor mutagenic efficiency was significantly affected in *rec*A$^-$ host cells. The findings reported here and in the preceding paper suggest that ethenocytosine may represent a novel type of RecA-independent, highly mutagenic noninstructional DNA lesion. These and other results argue that a requirement for SOS functions is neither a defining attribute nor an exclusive attribute of noninstructional lesions. Our data also support the possibility that mutation recovery (i.e., lesion bypass) rather than misincorporation may be the major role of SOS functions in mutagenesis. Finally, the high mutagenic efficiency observed makes εC a reasonable candidate for mediating the genotoxic properties of vinyl chloride, a suspected human carcinogen and a major industrial chemical that is produced in large quantities around the world.

A number of bifunctional alkylating agents, including recognized mutagens and carcinogens, are now known to induce exocyclic DNA lesions. Ethenocytosine (εC)[1] and ethenoadenine (εA) are among the best studied examples of exocyclic DNA adducts because of their fluorescent properties. The original interest in εC and εA lesions was largely centered on their use as enzyme substrate analogues or as chemical probes of polynucleotide conformation [reviewed by Leonard (1984)]. Interest in the possible biological role of etheno adducts followed the realization that they are formed during exposure to carcinogens such as vinyl chloride (Green & Hathaway, 1978; Eberle et al., 1989), ethyl carbamate (urethane), and vinyl carbamate (Leithauser et al., 1990). Vinyl chloride is believed to be activated to 2-chloroethylene oxide (CEO), which spontaneously and rapidly rearranges to 2-chloroacetaldehyde (CAA). Both CAA and CEO are known to react with DNA to yield etheno adducts. However, in quantitative terms, by far the most abundant DNA adduct induced by vinyl chloride (via CEO) is 7-(oxoethyl)guanine, which, however, is not thought to be mutagenic on the basis of a lack of in vitro miscoding ability (Barbin et al., 1985b). In contrast, etheno adducts are quantitatively minor in vivo products. As briefly reviewed in the preceding paper (Simha et al., 1991), a number of laboratories have investigated the miscoding properties of synthetic polynucleotides containing etheno lesions.

Both CAA (McCann et al., 1975) and CEO (Barbin et al., 1985a) induce reversions in bacteria, but the adducts responsible for the mutations could not be deduced from these studies. Jacobsen et al. (1989) reported experimental evidence strongly suggesting that εC is mutagenic. The specificity of base substitutions elicited by ethenocytosine (εC) indicated that the lesions were being bypassed as noninstructional lesions, an observation consistent with the fact that two of the three Watson-Crick hydrogen-bonding positions are blocked in εC and that no satisfactory alternative base-pairing schemes are known. Paradoxically however, εC differed from noninstructional lesions in three regards: (a) these lesions were less toxic than expected for noninstructional lesions such as abasic sites or bulky DNA adducts; (b) the estimated mutagenic efficiency for εC was unusually high for an apparent nonin-

† This work was supported by USPHS Grant CA47234.

* Corresponding author.

[1] Abbreviations: εC, 3,N^4-ethenocytosine; εA, 1,N^6-ethenoadenine; CEO, 2-chloroethylene oxide; CAA, 2-chloroacetaldehyde; A, adenine; C, cytosine; G, guanine; T, thymine; RF DNA, replicative form DNA; ds DNA double-stranded DNA; ss DNA, single-stranded DNA; bp, base pair(s); nt, nucleotides; form II′ DNA, circular gapped duplex DNA constructed as described, with the gel electrophoretic mobility of M13AB28 form II (relaxed circular) DNA; form II′$_{εC}$, form II′ DNA containing a single εC residue at a defined position; form II′$_C$, a control version in which cytosine replaces the εC residue in form II′$_{εC}$.

0006-2960/91/0430-8736$02.50/0

structional lesion; and (c) mutagenesis was apparently independent of SOS induction (Jacobsen et al., 1989; Jacobsen & Humayun, 1990). Since these unexpected properties may yield insights into the mutagenic process, we have sought to obtain definitive information on the mutagenic properties of ϵC through a site-specific adduction approach. In the preceding paper, we have described experiments to assess the in vitro template characteristics of ϵC. Here, we report on experiments evaluating the survival and mutagenic effects of a single site-specific ϵC residue carried on a bacteriophage DNA molecule.

Materials and Methods

Phage and Bacterial Strains. Phage M13AB28 is similar to M13mp2 in that it can plate on Sup^0 *Escherichia coli* host cells to yield dark blue plaques on appropriate indicator plates, but it carries the polylinker sequence of M13mp8 (Sambamurti et al., 1988). The *E. coli* strains KH2 Sup^0 Δ*lac–pro trp*E9777 F′LacI^qZΔM15Pro^+ and KH2R Sup^0 Δ*lac–pro trp*E9777 F′LacI^qZΔM15Pro^+ Δ(*srl*R–*rec*A)306::Tn10(Tet^R) were gifts from K. Sambamurti constructed as follows. For constructing KH2, the *E. coli* strain NR3835 Sup^0 Δ*lac–pro trp*E9777 F′$Pro^+Lac^+Ara^+$ (from Dr. B. Glickman) was cured of its F′ factor by acridine orange, and the resulting F^- strain was mated with *E. coli* LL308 (from Dr. E. LeClerc) to introduce the F′LacI^qZΔM15Pro^+ factor (Sambamurti et al., 1988). The KH2R strain was constructed by P1 transduction of the Δ(*srl*R–*rec*A)306::Tn10(Tet^R) locus from *E. coli* K158 (gift of Dr. E. Witkin) into *E. coli* KH2. The Δ(*srl*R–*rec*A)306 deletion is a total deletion of the RecA gene (Csonka & Clark, 1979). Both strains were maintained on appropriately supplemented M9 minimal agar plates. Before transfection, the KH2R (*recA*$^-$) strain was grown on minimal agar plates with appropriate nutritional supplements and tetracycline (12.5 μg/mL), and its UV sensitivity was confirmed.

Construction of M13AB28 Gapped Duplex (Form II′) DNA Bearing ϵC at a Unique Position. Phage M13AB28 RF DNA was prepared as described (Sambamurti et al., 1988). Approximately 500 μg of RF DNA was digested with 2500 units of the restriction endonuclease *Eco*RI (New England Biolabs) for 3 h in buffer (50 mM NaCl, 100 mM Tris-HCl, pH 7.5, 5 mM $MgCl_2$, and 0.05% Triton X-100; supplied as a 10× concentrate by the vendor). A fraction of the digest was examined by agarose gel electrophoresis to verify complete linearization of RF DNA. The linearized DNA (200 μg) was incubated with 100 units of calf intestine alkaline phosphatase (Boehringer-Mannheim) in 750 μL of buffer (20 mM Tris-HCl, pH 8, 1 mM $MgCl_2$, and 1 mM $ZnCl_2$) for 60 min at 37 °C. An additional 100 units of the phosphatase was added, and the incubation continued for another 60 min, at the end of which the process was repeated one more time (final enzyme level, 300 units; total incubation time, 180 min). The efficiency of dephosphorylation was examined by subjecting an aliquot of the dephosphorylated DNA to self-ligation and analysis of the products by electrophoresis on 1% agarose gels. Absence of form II (relaxed circular DNA) or multimeric DNA was taken to indicate that complete dephosphorylation had occurred. DNA not subjected to dephosphorylation was almost entirely converted to circular and multimeric forms under the same ligation conditions.

Approximately 100 μg of the dephosphorylated DNA and 3.5 μg (approximately 10-fold molar excess) of the appropriate 25-mer were annealed in 300 μL of ligase buffer (50 mM Tris-HCl, pH 7.8, 10 mM $MgCl_2$, and 20 mM dithiothreitol) through the following sequential incubations in a thermal cycler: 2 min each at 50, 40, 30, and 20 °C; 10 min at 16 °C; and 5 min at 4 °C. Ligation was initiated by adding 30 μL of 10 mM ATP followed by 4000 vendor units of T4 DNA ligase (New England Biolabs; 400 units/μL; 60 vendor units equal 1 Weiss unit), and the contents were incubated at 4 °C for 16 h. A fraction of this DNA was examined to confirm successful ligation of the 25-mer as follows. A total of 2 μg of the DNA was digested with 10 units of restriction endonuclease *Alu*I or *Pvu*II (both from New England Biolabs), and the products were 5′-end-labeled with ^{32}P by standard procedures before fractionation by electrophoresis. This analysis revealed that, as expected, the end fragments produced by *Alu*I or *Pvu*II had increased in length by 25 nt as a result of the ligation.

The ligated DNA (about 80 μg) was incubated with 400 units of restriction endonuclease *Pst*I (New England Biolabs; 20 units/μL) at 37 °C for 3 h in 500 μL of buffer (50 mM Tris-HCl, pH 7.9, 10 mM $MgCl_2$, 100 mM NaCl, and 1 mM dithiothreitol; supplied as a 10× concentrate by the vendor). The large *Eco*RI–*Pst*I linear DNA fragment was separated from the small fragment (18 bp, with 4-nt overhangs at both ends in the plus strand) as well as from unligated excess 25-mer by electrophoresis on a preparative 1% agarose gel. The large fragment (referred to as stage II linear DNA) was recovered from the gel by electroelution, the ethidium bromide used for staining was removed by extraction with butanol, and the DNA was purified by ethanol precipitation (yield, approximately 35 μg).

Approximately 70 μg of the stage II linear DNA prepared as above and 0.3 μg (representing a 5-fold molar excess) of the "helper hexamer" (see Results) were annealed in 12 mL of the above ligase buffer by sequential incubations in a thermal cycler as described above. Ligation was initiated at 4 °C by the addition of 1.4 mL of 10 mM ATP and 0.672 mL of T4 DNA ligase (400 000 vendor units/mL; DNA concentration, 5 μg/mL; final ligase concentration, 19 200 vendor units/mL). The ligation was continued at 4 °C for 16 h, and fractions were analyzed to confirm the formation of form II′ DNA. The reaction was terminated by phenol extraction, and the products were fractionated by electrophoresis on preparative 0.8% agarose gels. The gapped duplex DNA, which had the mobility of form II (relaxed circular) DNA, was electroeluted, butanol extracted, and ethanol precipitated as above.

Transfection, Determination of Survival, and Mutagenic Effects. DNA was transfected by the procedures of Chung et al. (1989) as follows. The appropriate *E. coli* strain was grown overnight in LB broth (10 g of Bacto tryptone, 5 g of Bacto yeast extract, and 10 g of sodium chloride per liter), the culture was diluted 1:100 in fresh LB broth, and incubation continued with vigorous aeration at 37 °C until an OD_{600} value of 0.3–0.4 was reached (approximately 8×10^7 cells/mL). The cells were pelleted by centrifugation at 4 °C in a Sorvall SS34 rotor at 2500 rpm for 10 min. The cells were washed by resuspension in fresh sterile LB broth (original volume) and pelleting again as above. The washed cells were gently resuspended in one-tenth of the original volume of ice-cold LB broth containing 10% (w/v) poly(ethylene glycol) 3350 (J. T. Baker), 5% (v/v) dimethyl sulfoxide (Sigma), and 20 mM $MgCl_2$, pH 6.5 (transformation and storage solution). For transfection, 1-mL aliquots of the cell suspension were transferred to prechilled 15-mL polypropylene tubes (Falcon No. 2059) and mixed with 50 ng of DNA contained in a volume of 5–10 μL. After incubation of the contents at 4 °C for 30 min, the transfection mix either was directly plated as described below or was diluted before plating by gently mixing in an additional 4 mL of cold transformation and storage

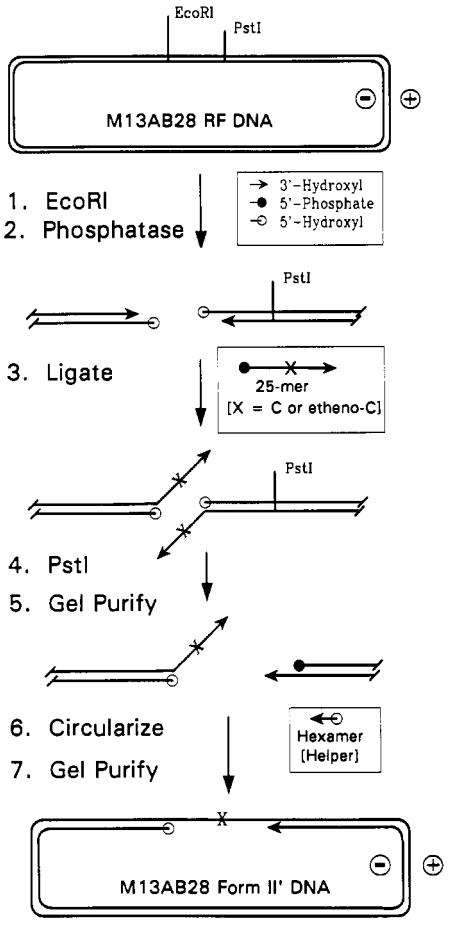

FIGURE 1: Strategy used for introducing a single ϵC residue at a specific site in M13AB28 circular gapped duplex (form II′) DNA. The viral (+) and complementary (−) DNA strands are identified. For convenience, the linear DNA obtained after step 5 (linear DNA containing a 25-mer ligated to the *Eco*RI end) is called stage II linear DNA in the text. While this article was in review, an analogous procedure for constructing a gapped duplex DNA containing an 8-oxodeoxyguanosine adduct was reported by Moriya et al. (1991).

solution. Aliquots (0.01–0.4 mL) of the transfected cells were mixed with 0.2 mL of fresh culture grown to saturation in LB medium, 0.06 mL of inducer-dye solution [16.7 mM isopropyl β-D-1-thiogalactopyranoside, 3.3% (w/v) of Xgal, 83.3% (v/v) dimethylformamide], and 3 mL of top agar (8 g of Bacto minimal agar and 8 g of sodium chloride per liter) and plated on M9 minimal agar. The plates were incubated for 16 h at 37 °C for observation of phage plaques. For DNA sequence analysis, individual plaques were subjected to a second cycle of plaque purification. Methods for DNA isolation and sequencing have been described (Refolo et al., 1987; Sambamurti et al., 1988). A 15 nt long primer (P6368: 5′-GTGCGGGCCTCTTCG), which primes complementary strand synthesis 123 nt downstream of the lesion site, was used to determine the DNA sequence.

Results

Construction and Purification of a Circular Duplex M13 DNA Molecule Containing an ϵC Residue within a 17 nt Long Single-Stranded Region. The strategy for introduction of a single ϵC residue into M13AB28 RF DNA is outlined in Figure 1, while the relevant DNA sequences are shown in Figure 2. The single-stranded 25-mer containing a centrally located ϵC residue (Figure 2, top) was constructed and characterized as described in the preceding paper (Simha et al., 1991). The form II′ DNA shown at the bottom of Figure 1 was constructed from M13AB28 RF DNA and the 25-mer

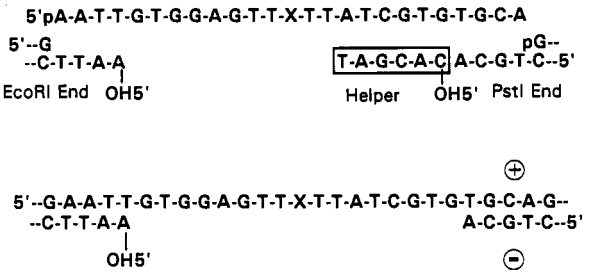

FIGURE 2: DNA sequence of the M13AB28 RF DNA termini, the 25-mer, and the (boxed) helper hexamer. Note that the 5′- and 3′-ends of the 25-mer are ligated at different stages (see Figure 1) rather than in one step and that the hexamer itself cannot be ligated because it has a 5′-hydroxyl group. The viral (+) and complementary (−) strands are identified.

in the following four stages. (i) M13 RF DNA was cut with *Eco*RI and the 5′-phosphates were removed with use of a phosphatase. (ii) 5′-phosphorylated 25-mer was ligated with the *Eco*RI-cut, dephosphorylated linear DNA. Under these conditions, the 5′-phosphate end of the 25-mer is expected to be ligated to the 3′-hydroxyl ends of the linearized RF DNA, as verified by subsequent restriction endonuclease digestion and gel analysis (Materials and Methods, data not shown). (iii) The ligated linear DNA was cut with *Pst*I, and the resulting large DNA fragment was purified by agarose gel electrophoresis. (iv) This DNA was self-ligated at low concentrations in the presence of a helper hexamer (see below) to obtain circular gapped DNA, which was purified by electroelution.

Early experiments indicated that formation of form II′ molecules at a level detectable by ethidium bromide staining of agarose electrophoretic gels required the presence of the helper hexamer, which had a sequence complementary to bases 16–21 of the 25-mer (Figure 2, boxed oligonucleotide). In Figure 3, panel a shows that significant levels of form II′ DNA are only formed in the presence of the hexamer (compare lane 2 with 3 or lane 4 with 5). Note that the hexamer itself has a 5′-hydroxyl group and is therefore nonligatable and is expected to be removed during the subsequent gel electrophoretic purification of the ligation products. [5′-phosphorylation of the hexamer does not improve ligation efficiency further (data not shown).] Although we did not characterize the helper effect further, we assume that efficient ligation of a nick by T4 DNA ligase requires that the nick be flanked by a minimum length of ds DNA.

Preparative-scale self-ligation of the stage II linear DNA was carried out in the presence of the hexamer as described under Materials and Methods, and aliquots were analyzed to confirm formation of form II′ DNA (lane 2 in Figure 3b,c). ϵC-containing form II′ DNA (form $II'_{\epsilon C}$) and the control DNA (form II'_C) containing normal cytosine in place of ϵC were purified by preparative agarose gel electrophoresis followed by electroelution. Figure 3d shows that gel-purified form $II'_{\epsilon C}$ (lane 3) and form II'_C DNA are substantially pure. Although difficult to see in Figure 3d, we found that trace amounts of the linear dimer (D) as well as linear monomer (L) persisted through the gel purification step. Figure 4 shows densitometric tracings (from a polaroid photographic negative) of two of the lanes in the gel shown in Figure 3d (top trace, markers; bottom trace, purified form $II'_{\epsilon C}$). Densitometric analysis confirmed that the form $II'_{\epsilon C}$ DNA was >95% pure. The final yield of the purified gapped duplex DNA was about 8% of the input stage II linear DNA for the ϵC-containing DNA and about 14% for the corresponding control DNA. For the purpose of comparing transfection efficiencies of circular and linear DNA

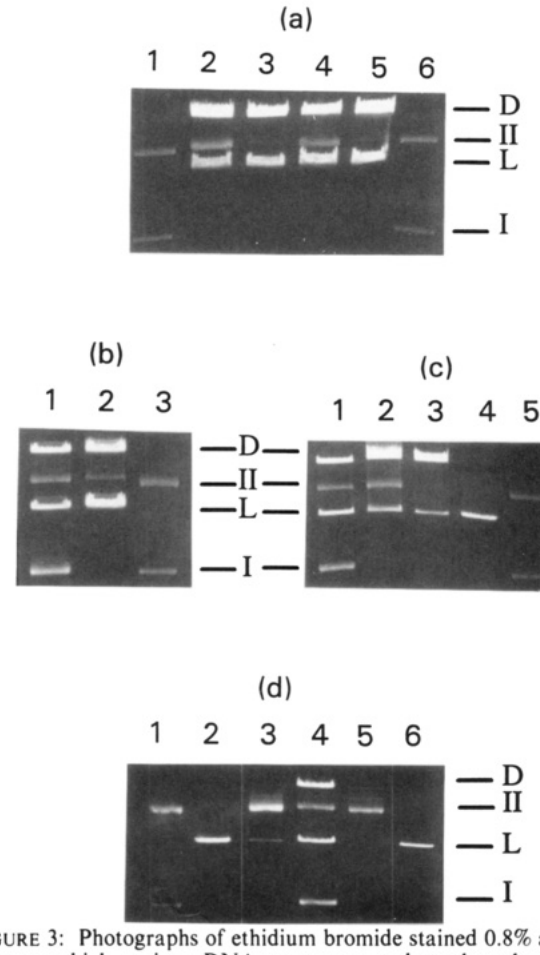

FIGURE 3: Photographs of ethidium bromide stained 0.8% agarose gels on which various DNA constructs and markers have been fractionated. D, linear dimeric DNA; II, form II (and form II′) DNA; L, linear monomeric DNA; I, form I marker. Panel a: Lanes 1 and 6 are identical marker lanes showing the positions of form I (covalently closed supercoiled) DNA and form II (relaxed circular DNA). Lane 2 represents self-ligation of stage II linear DNA (see legend for Figure 1) at a concentration of 10 μg/mL in the presence of hexamer as described under Materials and Methods. Lane 3 is the same as lane 2 except for the omission of the hexamer. Lanes 4 and 5 are the same as lanes 2 and 3, respectively, except that the DNA concentration at ligation was 5 μg/mL. Panels b and c: Lane 2 is the analysis of an aliquot of preparative-scale ligation as described to show the formation of form II′$_C$ (panel b) and of form II′$_{\epsilon C}$ (panel c) DNA. All other lanes are markers. Panel d: Lanes 1 and 4 are markers. Lane 3 is the gel-purified form II′$_{\epsilon C}$ DNA, and lane 2 is the corresponding gel-purified stage II linear DNA (i.e., the DNA remaining uncircularized after ligation); lane 5 is the gel-purified form II′$_C$ DNA; lane 6 is the corresponding gel-purified linear DNA.

forms, linear monomer DNA that escaped ligation was also recovered as above (lanes 2 and 6, Figure 3d).

Phage M13AB28, which is similar to M13mp2, yields dark blue plaques on appropriate media containing Xgal (Sambamurti et al., 1988). The strategy used for building form II′ DNA as described above replaces 26 nt of M13AB28 DNA with 25 nt, in effect creating a −1 frameshift in the LacZ(α) gene segment. Therefore, progeny phage derived from transfecting form II′ DNA are expected to have a colorless phenotype.

Effect of ϵC on the Survival of M13AB28 Form II′ DNA. Table I summarizes the results of transfection of linear and gapped duplex DNA bearing a site-specific ϵC lesion (or cytosine in the control) into unirradiated *E. coli* KH2 (*rec*A$^+$) or KH2R (*rec*A$^-$) cells. The data in Table I show that in *E. coli* KH2 (wild-type) cells, the transfection efficiency of form II′$_C$ DNA averages to 125% (60–157%) of that of the marker

Table I: Effect of a Single Site-Specific ϵC Residue on Survival of M13 Circular Gapped Duplex DNA

strain	expt[a] no.	type[b] of DNA	pfu[c]/50 ng of DNA	percent of marker
E. coli KH2 (*rec*A$^+$)	I	linear	93	0.12
		form II′$_{\epsilon C}$	85 727	110
		marker	77 938	100
	II	linear	120	0.09
		form II′$_C$	83 324	60
		marker	138 880	100
	III	linear	57	0.08
		form II′$_{\epsilon C}$	73 300	99.28
		marker	73 830	100
	IV	linear	118	0.31
		form II′$_C$	59 500	157.97
		marker	37 665	100
	Va	linear	31	0.06
		form II′$_{\epsilon C}$	83 600	150.09
		marker	55 700	100
	Vb	linear	95	0.17
		form II′$_C$	87 700	157.45
E. coli KH2R (*rec*A$^-$)	VI	linear	ND[d]	
		form II′$_{\epsilon C}$	10 953	68.46
		marker	16 000	100
	VII	linear	76	0.55
		form II′$_C$	21 825	158.15
		marker	13 800	100
	VIII	linear	47	0.13
		form II′$_{\epsilon C}$	24 633	67.49
		marker	36 500	100
	IX	linear	78	0.07
		form II′$_C$	50 047	47.31
		marker	105 786	100
	Xa	linear	26	0.16
		form II′$_{\epsilon C}$	25 450	161.08
		marker	15 800	100
	Xb	linear	72	0.46
		form II′$_C$	38 850	245.89

[a] Each experiment is completely independent, involving a different competent cell preparation. In experiments Va and Vb, the same *E. coli* KH2 competent cells were used for transfecting all DNAs. Similarly, in experiments Xa and Xb, the same *E. coli* KH2R competent cells were used. [b] Linear DNA is the uncircularized monomeric DNA [stage II linear DNA; see legend for Figure 1) recovered from the same preparative gels from which the circular gapped duplex DNA (form II′) was isolated (see lane 6, Figure 3d)]. Marker DNA is a preparation of M13AB28 RF DNA that was transfected into each competent cell preparation for the purposes of comparing transfection efficiencies of the various DNAs across different transfection experiments. [c] pfu, plaque forming units per 50 ng of transfected DNA. Transfection mixes were plated either directly (linear DNA) or after appropriate dilution (form II′ and marker DNAs) on Xgal indicator plates so as to obtain 25–400 plaques per plate as described. Values shown for form II′ DNAs are averages for 10–19 plates, and those for linear and marker DNAs are averages for 2–5 plates. As expected, all plaques from marker DNA transfection were blue, and almost all from form II′ transfection were colorless. About 9% (1–24%) of the plaques obtained by transfecting linear DNA were blue. These blue plaques are most likely from low levels of wild-type M13AB28 RF DNA surviving the various procedures. About 0.7% (0–1.5%) of the plaques induced by the form II′ DNAs were also similarly blue. [d] ND, not determined.

(M13AB28 RF) DNA, while that of form II′$_{\epsilon C}$ DNA averages to 120% (99–150%) of the transfection efficiency of the marker. These figures imply that the ϵC residue does not confer significant lethality in *rec*A$^+$ cells. In *E. coli* KH2R (*rec*A$^-$) cells, the transfection efficiency of form II′$_C$ DNA averages to 150% (47–246%) of the marker, while that of form II′$_{\epsilon C}$ DNA averages to 99% (68–161%). These data imply that the absence of a functional RecA gene product does not significantly affect the survival of DNA containing an ϵC residue. However, it can be argued, by comparing the average efficiency of form II′$_{\epsilon C}$ DNA to that of form II′$_C$ DNA in *rec*A$^-$ cells, that a reduction in survival of about 33% may be at-

Table II: DNA Sequence Analysis of Progeny Phage Obtained by Transfection of Form II′$_{\epsilon C}$ and Form II′$_C$ DNAs

E. coli host strain	DNA	no. sequenced	no. of phage in which base X is replaced by[a]				large deletion[b]	mutation frequency[c] (%)
			C	A	T	G		
KH2 (*recA*$^+$)	form II′$_{\epsilon C}$	53	36	10	7	0	0	32
	form II′$_C$	39	38	0	0	0	1	–
KH2R (*recA*$^-$)	form II′$_{\epsilon C}$	56	40	8	7	0	1	27
	form II′$_C$	30	30	0	0	0	0	–

[a] Base X is ϵC in form II′$_{\epsilon C}$ and C in form II′$_C$ (see Figure 2 for sequence). Note that X-to-C change is equivalent to no mutation. Other than the two large deletions, the only observed mutations mapped opposite base X. No other sequence changes were observed. [b] The large deletions are phages that have lost the sequence site for the primer as deduced by a total absence of DNA synthesis. Although we have not investigated these deletions further, these may be similar to the previously observed 126-bp deletion and larger deletions previously observed as background events in this experimental system (Sambamurti et al., 1988). These deletions presumably arise from inter- or intramolecular recombination events. [c] We have ignored the large deletions because they are believed to be background events, but the numbers do not change significantly even if they are included.

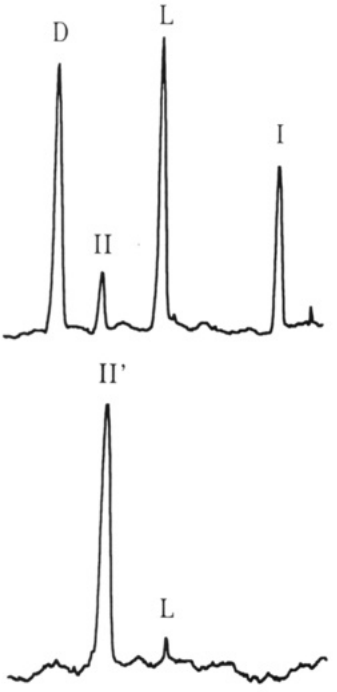

FIGURE 4: Densitometric analysis of a photographic negative of the gel shown in Figure 3d. The top trace shows markers (lane 4 of Figure 3d): D, linear dimeric DNA; II, form II (relaxed circular DNA); L, linear monomeric DNA. In the bottom trace, L represents co-purifying uncircularized stage II linear DNA and II′ represents form II′ DNA with a mobility similar to relaxed circular (form II) DNA shown in the top trace.

tributable to ϵC (see Discussion). Finally, the data in Table I show that the transfection efficiency of form II′ DNA is 2–3 orders of magnitude higher than that for the linear DNA, implying that the low levels of linear DNA copurifying with form II′ DNAs will have negligible impact on transfection efficiency. [Linear dimeric DNA had a similar low transfection efficiency (data not shown).]

DNA Sequence Analysis of Phage Progeny Obtained by Transfecting Form II′ DNA. Since wild type as well as mutant progeny of form II′ DNA (except for any genetic +1 frameshifts) produce colorless plaques, no phenotype screen is available for distinguishing mutants from wild-type phage. We have, therefore, determined both the mutation frequency and specificity by a direct sequencing approach. Figure 5 shows parts of the DNA sequences of the three types of progeny phage observed, while Table II summarizes the results of sequence analysis. The data in Table II show that approximately 30% of the progeny obtained by transfecting form II′$_{\epsilon C}$ DNA were mutant, with C-to-T transitions and C-to-A transversions accounting for all mutations. No such mutations are observed in the progeny obtained by transfecting form II′$_C$ DNA. Finally, neither mutation frequency nor specificity is affected in the *recA*$^-$ strain.

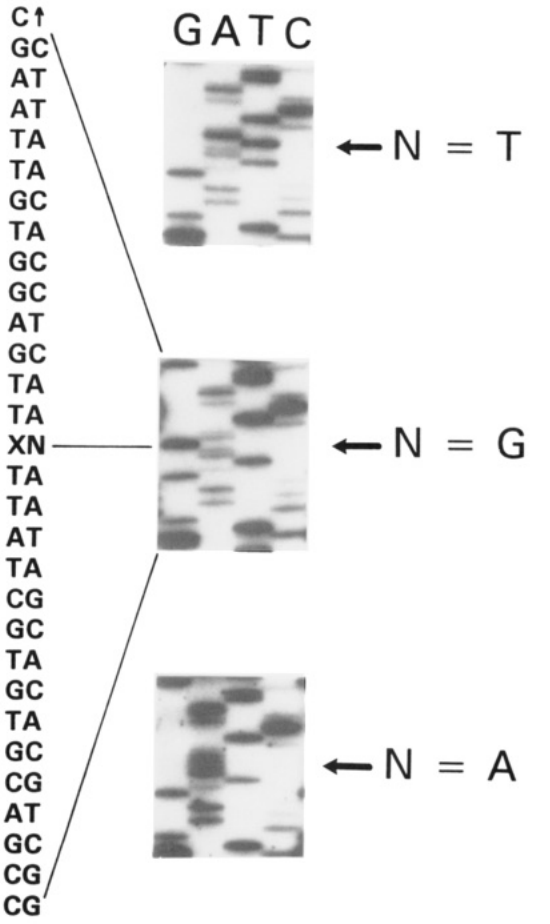

FIGURE 5: Autoradiographs of sequence analysis gels showing the three types of progeny phage molecules observed upon transfection of form II′$_{\epsilon C}$ DNA. A part of the DNA sequence of the template strand (X = ϵC), and the synthesized strand (N = G, A, or T) is shown at left. The three sequences differ only at the position corresponding to the template base X.

DISCUSSION

ϵC Does Not Significantly Affect Survival. Hitherto available experimental evidence supports the prevalent assumption that noninstructional lesions stop DNA replication, leading to SOS induction and, eventually, to "error-prone SOS repair" of the lesion [for reviews, see Radman (1975), Witkin (1976), Walker (1984), and Loeb and Preston (1986)]. In the absence of SOS induction, replication arrest leads to loss (death) of the genome. For example, survival of M13 single-stranded DNA is reduced to <1% in unirradiated cells by the introduction of a single T–T cyclobutane dimer (Banerjee et al., 1988) or a single abasic site (Lawrence et al., 1990).

The data here make it clear the εC DNA does not reduce survival to a detectable extent in unirradiated wild-type host cells. Depending on which control transfection is used for comparison, there is either no drop in survival in *recA*⁻ cells, or, at best, there is a 33% loss in survival. It is difficult to evaluate the significance of this 33% drop, because it is within the transfection-to-transfection variability observed in M13 system (Kunkel, 1984; Refolo et al., 1987) and differs significantly from the >99% loss of survival observed for abasic sites and UV dimers as reviewed above. These observations support the hypothesis that εC does not represent a strong block to replication (Jacobsen & Humayun, 1990).

εC Is a Highly Efficient, RecA-Independent, Mutagenic Lesion. Site-specific adduction experiments as well as random DNA adduction experiments generally reveal mutagenic efficiency ranging from under 1% to under 10% of surviving genomes [e.g., Banerjee et al. (1988), Burnouf et al. (1990), and Moriya et al. (1988); see review by Basu and Essigmann (1988)]. Furthermore, for most bulky lesions examined, there is a strong SOS dependence for efficient mutagenesis [e.g., Banerjee et al. (1988), Burnouf et al. (1990), and Lawrence et al. (1990)]. O^6-Methylguanine, the classic example of a mispairing mutagenic lesion, induces mutations (exclusively G-to-A transitions) at frequency of 0.4% when introduced as a site-specific lesion in ss DNA. Under conditions of depleted repair, O^6-methylguanine induces G-to-A transitions at an efficiency reaching or exceeding 20% (Bhanot & Ray, 1986; Basu & Essigmann, 1988).

We have previously estimated the efficiency of mutagenesis at εC lesions to be 40% on the basis of adduct quantitation and relative mutation frequency calculations (Jacobsen & Humayun, 1990). The present study confirms that the mutagenic efficiency is indeed a high 30% in wild-type as well as *recA*⁻ cells. The previous study had also indicated that relative mutation frequency was higher at εC flanked by purines (Pu-εC-Pu) than at εC flanked by pyrimidines (Py-εC-Py; Jacobsen & Humayun, 1990). Since the sequence context in the current experiments is T-εC-T, it is possible that the 30% figure might represent a conservative estimate.

Strong initial evidence that εC residues as well as their hydrated derivatives (εC·H_2O) acted as noninstructional lesions came from the observation that both lesion forms induced predominantly C-to-T transitions and C-to-A transversions (Jacobsen et al., 1989; Jacobsen & Humayun, 1990). These observations, coupled to the fact that two of the three Watson–Crick hydrogen-bonding positions are eliminated by the etheno bridging, made this a reasonable hypothesis. The data presented in the preceding paper (Simha et al., 1991) establish that εC indeed had the properties of a noninstructional lesion in vitro. Specifically, no base incorporation opposite εC is detectable under conditions in which normal cytosine readily permits incorporation of guanine. At elevated nucleotide precursor concentrations, base incorporation opposite εC follows patterns established for noninstructional lesions, namely, a preferential addition of adenine followed by thymine and other bases.

The current experiments confirm that εC predominantly induces C-to-T transitions and C-to-A transversions. However, the ratio of C-to-T transitions to C-to-A transversions determined from mutation spectra is 3:1, whereas the ratio obtained here is closer to 1:1. This difference may point to a sequence context effect, since in the mutation spectra, while C-to-T transitions constitute a consensus majority, C-to-A transversions predominate at specific sites [e.g., C6338 and C6379; Figure 4 of Jacobsen et al. (1989)]. [The exact sequence context used in the site-specific experiments (T-εC-T) does not exist in the target sequence previously used for obtaining the mutational spectra.] It is alternatively possible that the observed specificity difference is attributable to the relatively small number of mutants sequenced in the present study.

A Requirement for SOS Functions Is neither a Definitive nor an Exclusive Attribute of Noninstructional Mutagenic Lesions. Even though SOS dependence for survival and mutagenesis has come to be accepted as a defining characteristic of a noninstructional mutagenic DNA lesion, the properties of εC argue that a requirement for SOS functions cannot be a defining attribute of such lesions. Because mutational mechanisms have been intensively studied for only a small number of lesions, it is likely that εC will prove to be not so much a unique mutagenic lesion as the first clear example of a hitherto unrecognized class of mutagenic lesions. Accumulating experimental evidence indicates that SOS dependence may be an independent property of mutagenic lesions unrelated to whether the lesion is mispairing or noninstructional in its template characteristics. More specifically, there are SOS-dependent as well as SOS-independent mispairing lesions (Sambamurti et al., 1988) just as there are SOS-dependent and SOS-independent noninstructional lesions. Finally, the highly efficient mutagenesis opposite εC lesions in unirradiated cells as well as in the cells totally lacking the RecA protein strongly supports the idea that the primary function of the SOS phenomenon in mutagenesis might be in mutation recovery (lesion bypass) rather than in base incorporation opposite lesions. This proposal is similar to that of Bridges and Woodgate (1985) except that there is no indirect (i.e., via SOS) or direct requirement for the RecA protein for base incorporation opposite some noninstructional lesions.

While the available data are consistent with the above conclusions regarding the mutagenic properties of εC, two caveats remain to be experimentally addressed. The first caveat is that all experiments here used unirradiated cells. Therefore, it is possible that even though SOS functions may not be required for mutagenesis at εC, SOS functions can still modulate mutagenesis. The spectrum data provided by Jacobsen and Humayun (1990) do not show a significant modulating effect of SOS induction on mutational frequency or specificity attributable to εC residues. However, more definitive information on the effect of SOS induction on mutagenesis at εC residues is desirable. Finally, the possibility exists that the high survival and mutagenesis of form $II'_{εC}$ DNA in the absence of SOS functions is attributable to an experimental strategy that places the lesion in a 17 nt long single-stranded DNA region (gap; see Figure 1). Here, one can argue that translesion synthesis occurs during gap-filling DNA synthesis rather than during normal replication and that gap-filling synthesis differs from normal replication in ways that can account for the observed high mutagenic efficiency of εC as well as its low toxicity in the absence of SOS functions.

Mechanisms of Error Avoidance at εC Residues. In vitro evidence presented in the preceding paper indicates that εC does not permit significant guanine incorporation in vitro and that the efficiency of lesion bypass after guanine incorporation is low (Simha et al., 1991). This result is entirely consistent with the expectation that etheno bridging seriously compromises the Watson–Crick base-pairing ability of εC. Nevertheless, as noted by Jacobsen and Humayun (1990) and confirmed by the data in Table II, guanine accounts for a major (55–70%) fraction of the in vivo base incorporation opposite the lesion site. Two highly speculative hypotheses

can account for this discrepancy. According to the first hypothesis, the replication apparatus has evolved error-avoidance mechanisms capable of recognizing the parent base from which some types of noninstructional lesions arise. No evidence is presently available to support this possibility. Repair of ϵC is an alternative possibility. An excision pathway, even if available, is unlikely to account for the results here because excision before translesion synthesis will inactivate form II'$_{\epsilon C}$ DNA through linearization. Excision after translesion synthesis cannot change the pattern of base incorporation, and furthermore, translesion synthesis will result in a covalently closed minus strand capable of initiating the phage replication cycle. Effective repair must therefore be able to restore normal cytosine without cleaving phosphodiester bonds in a manner analogous to that of photolyases (Sancar & Sancar, 1988) or alkyltransferases (Lindahl et al., 1988). Either possibility can also account for the paradoxical properties of ϵA, which is neither lethal nor highly mutagenic despite the blockage of both Watson–Crick hydrogen-binding positions (Jacobsen et al., 1989). The possibility of direct repair can in principle be tested by saturation of the repair pathway before transfection of form II'$_{\epsilon C}$ DNA. Even though the high toxicity of chloroacetaldehyde makes this particular experiment difficult, available genetic and biochemical tools should enable one to experimentally address the possibility of direct repair of etheno lesions.

Role of ϵC in the Genotoxicity of the Carcinogen Vinyl Chloride. Even though etheno adducts constitute a quantitatively minor fraction of the adducts induced by metabolites of vinyl chloride, the high efficiency of mutagenesis in *E. coli* makes ϵC a reasonable candidate for mediating the biological activity of vinyl chloride. Jacobsen and Humayun (1990) have previously argued that mutagenesis induced by 2-chloroethylene oxide (Barbin et al., 1985a), the major initial metabolite of vinyl chloride, may in fact be mediated by ϵC. In contrast to ϵC, ϵA is relatively less mutagenic in *E. coli*. On the other hand, although in vivo mutagenic effects have not yet been investigated, N^2,N^3-ethenoguanine readily miscodes for thymine in vitro (Singer et al., 1987), raising the possibility that it might induce mutations with a predicted exclusive specificity of G·C to A·T transitions. In view of the known mutagenic properties of guanine N7 adducts (Refolo et al., 1987; Sambamurti et al., 1988; Sahasrabudhe et al., 1989, 1990), it is possible that 7-(oxoethyl)guanine, the major vinyl chloride adduct, is mutagenic despite its lack of in vitro miscoding ability (Barbin et al., 1985b).

Since etheno adduction of cytosine involves atoms that are normally hydrogen bonded in duplex DNA, it is expected that mutagenesis induced by vinyl chloride metabolites through ϵC will be highly nonrandom. Target sequences are likely to be those that can provide access to unpaired cytosines, such as heavily transcribed sequences, H-DNA sequences, B–Z conformational junction sequences, and inverted repeat sequences capable of forming cruciform structures in vivo [see Wells et al. (1988) for a review). This possibility is strongly supported by the known preferential reaction of the related agent 2-bromoacetaldehyde with unusual DNA structures in vitro and in vivo (Kohwi-Shigematsu et al., 1983; Noirot et al., 1990). The frequent association of unusual DNA structures with regulatory sequence elements confers obvious biological significance to mutagenic mechanisms capable of targeting such sites.

Acknowledgments

We thank Drs. Glickman, LeClerc, and Witkin for strains and Dr. Sambamurti for constructing the strain set used here.

References

Banerjee, S. K., Christensen, R. B., Lawrence, C. W., & LeClerc, J. E. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 8141–8145.

Barbin, A., Besson, F., Perrard, M.-H., Bereziat, J.-C., Kaldor, J., Michel, G., & Bartsch, H. (1985a) *Mutat. Res. 152*, 147–156.

Barbin, A., Laib, R. J., & Bartsch, H. (1985b) *Cancer Res. 45*, 2440–2445.

Basu, A. K., & Essigmann, J. M. (1988) *Chem. Res. Toxicol. 1*, 1–18.

Bhanot, O. S., & Ray, A. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 7348–7352.

Bridges, B. A., & Woodgate, R. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 4193–4197.

Burnouf, D., Gauthier, C., Chottard, J. C., & Fuchs, R. P. P. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 6087–6091.

Chung, C. T., Niemela, S. L., & Miller, R. H. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 2172–2175.

Csonka, L. N., & Clark, A. J. (1979) *Genetics 93*, 321–343.

Eberle, G., Barbin, A., Laib, R. J., Ciroussel, F., Thomale, J., Bartsch, H., & Rajewsky, M. F. (1989) *Carcinogenesis 10*, 209–212.

Green, T., & Hathaway, D. E. (1978) *Chem.-Biol. Interact. 22*, 211–224.

Jacobsen, J. S., & Humayun, M. Z. (1990) *Biochemistry 29*, 496–504.

Jacobsen, J. S., Perkins, C. P., Callahan, J. T., Sambamurti, K., & Humayun, M. Z. (1989) *Genetics 121*, 213–222.

Kohwi-Shigematsu, T., Gelinas, R., & Weintraub, H. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 4389–4393.

Kunkel, T. A. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 1494–1498.

Lawrence, C. W., Borden, A., Banerjee, S. K., & LeClerc, J. E. (1990) *Nucleic Acids Res. 18*, 2153–2157.

Leithauser, M. T., Liem, A., Stewart, B. C., Miller, E. C., & Miller, J. A. (1990) *Carcinogenesis 11*, 463–473.

Leonard, N. J. (1984) *CRC Crit. Rev. Biochem. 15*, 125–199.

Lindahl, T., Sedgwick, B., Sekiguchi, M., & Nakabeppu, Y. (1988) *Annu. Rev. Biochem. 57*, 133–157.

Loeb, L. A., & Preston, B. D. (1986) *Annu. Rev. Genet. 20*, 201–230.

McCann, J., Simon, V., Streitweiser, D. & Ames, B. N. (1975) *Proc. Natl. Acad. Sci. U.S.A. 72*, 3190–3193.

Moriya, M., Takeshita, M., Johnson, F., Peden, K., Will, S., & Grollman, A. P. (1988) *Proc. Natl. Acad. Sci U.S.A. 85*, 1586–1589.

Moriya, M., Ou, C., Bodepudi, V., Johnson, F., Takeshita, M., & Grollman, A. P. (1991) *Mutat. Res. 254*, 281–288.

Noirot, P., Bargonetti, J., & Novick, R. P. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 8560–8564.

Radman, M. (1975) in *Molecular Mechanisms for Repair of DNA* (Hanawalt, P. C., & Setlow, R. B., Eds.) pp 355–367, Plenum, New York.

Refolo, L. M., Bennett, C. B., & Humayun, M. Z. (1987) *J. Mol. Biol. 193*, 609–636.

Sahasrabudhe, S., Sambamurti K., & Humayun, M. Z. (1989) *Mol. Gen. Genet. 217*, 20–25.

Sahasrabudhe, S., Luo, X., Humayun, M. Z. (1990) *Biochemistry 29*, 10899–10905.

Sambamurti, K., Callahan, J., Luo, X., Perkins, C. P., Jacobsen, J. S., & Humayun, M. Z. (1988) *Genetics 120*, 863–873.

Sancar, A., & Sancar, G. B. (1988) *Annu. Rev. Biochem. 57*, 29–67.

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

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

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

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

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

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

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

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

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

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