The EMBO Journal vol.14 no.22 pp.5724–5735, 1995

Intron-encoded endonuclease I-*Tev*I binds as a monomer to effect sequential cleavage via conformational changes in the *td* homing site

John E.Mueller, Dorie Smith, Mary Bryk and Marlene Belfort[1]

Molecular Genetics Program, Wadsworth Center and School of Public Health, New York State Department of Health, PO Box 22002, Albany, NY 12201-2002, USA

[1]Corresponding author

I-*Tev*I, the intron-encoded endonuclease from the thymidylate synthase (*td*) gene of bacteriophage T4, binds its DNA substrate across the minor groove in a sequence-tolerant fashion. We demonstrate here that the 28 kDa I-*Tev*I binds the extensive 37 bp *td* homing site as a monomer and significantly distorts its substrate. *In situ* cleavage assays and phasing analyses indicate that upon nicking the bottom strand of the *td* homing site, I-*Tev*I induces a directed bend of 38° towards the major groove near the cleavage site. Formation of the bent I-*Tev*I–DNA complex is proposed to promote top-strand cleavage of the homing site. Furthermore, reductions in the degree of distortion and in the efficiency of binding base-substitution variants of the *td* homing site indicate that sequences flanking the cleavage site contribute to the I-*Tev*I-induced conformational change. These results, combined with genetic, physical and computer-modeling studies, form the basis of a model, wherein I-*Tev*I acts as a hinged monomer to induce a distortion that widens the minor groove, facilitating access to the top-strand cleavage site. The model is compatible with both unmodified DNA and glucosylated hydroxymethylcytosine-containing DNA, as exists in the T-even phages.

Keywords: DNA bending/double-strand intron endonuclease/minor groove interaction/nicked substrate/phage T4

Introduction

The thymidylate synthase (*td*) intron of the bacteriophage T4 is a mobile element by virtue of endonucleolytic activity encoded within the intron. The endonuclease I-*Tev*I recognizes and cleaves a cognate intronless allele; subsequent repair of the double-strand break results in intron inheritance (reviewed in Belfort, 1990; Mueller *et al.*, 1993).

I-*Tev*I interacts with a 35–37 bp region of the intronless *td* allele (Bryk *et al.*, 1995), referred to as the *td* homing site, and cleaves its substrate 23 and 25 nucleotides upstream of the intron insertion site (Bell-Pedersen *et al.*, 1990; Chu *et al.*, 1990; Figure 1A). Mutational analyses indicate that I-*Tev*I is extremely tolerant of base changes within the *td* homing site. Although no single base pair within the region of interaction is absolutely required for I-*Tev*I activity, the endonuclease manifests sequence preferences in distinct regions of the homing site. Two domains of the homing site, DI and DII, which are located downstream of the I-*Tev*I cleavage site and intron insertion site, respectively (Figure 2A), have been identified based on the relative intolerance of I-*Tev*I to mutations in these regions. The endonuclease is also somewhat sensitive to mutations in a third domain (DIII), which spans the intron insertion site (Bryk *et al.*, 1993; Figure 2A). In addition, I-*Tev*I exhibits sequence preferences at the remote cleavage site. The endonuclease is able to reach and pull back to a displaced cleavage site, suggesting that the DNA binding and catalytic portions of I-*Tev*I are tethered by a flexible hinge (Bryk *et al.*, 1995).

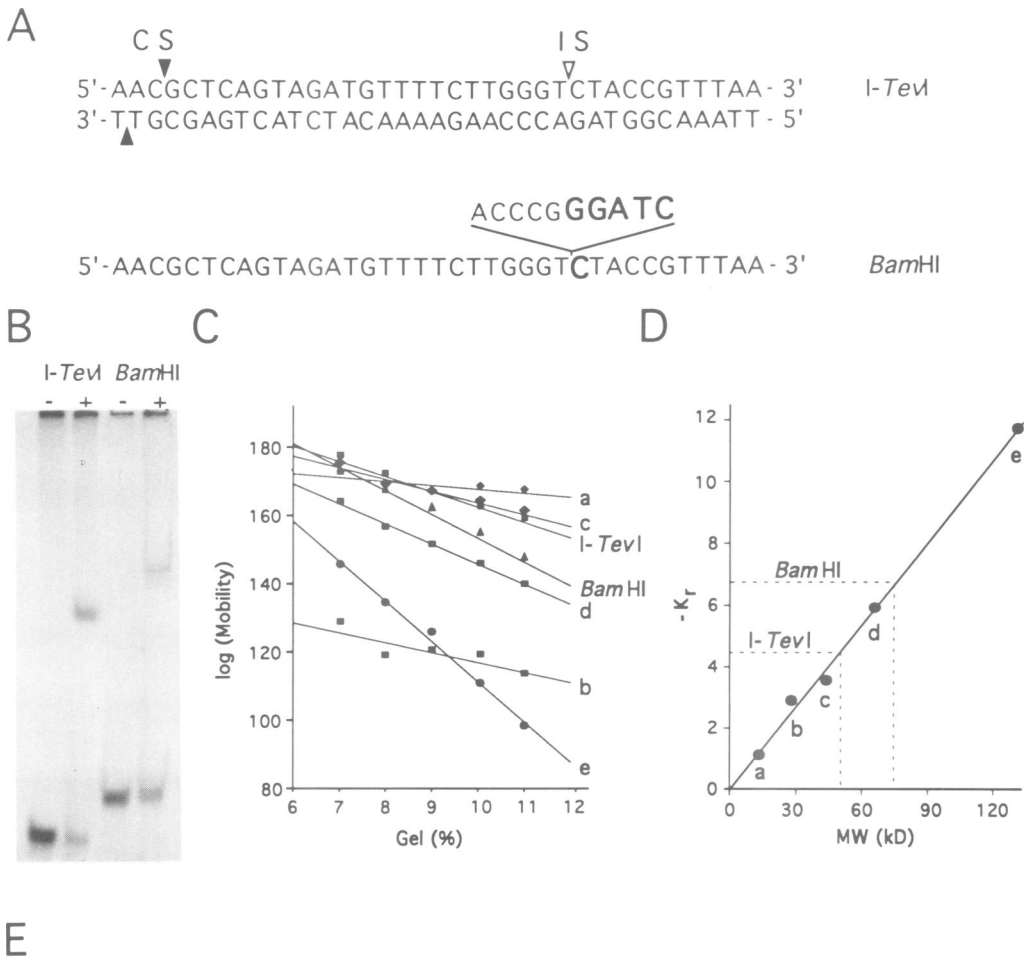
Chemical protection and interference studies indicate that I-*Tev*I contacts the *td* homing site predominantly across the minor groove, in close proximity to the sugar–phosphate backbone. The discontinuous regions of contact superimpose the genetically defined homing-site domains DI and DIII, while DII is partially contacted (Bryk *et al.*, 1993). Characteristics, that include the I-*Tev*I interaction across the minor groove where sequence discrimination is limited, the sequence tolerance of I-*Tev*I and the presence of genetically implicated noncontacted sequences within the *td* homing site, suggest that DNA features other than base sequence play a role in I-*Tev*I function (Bryk *et al.*, 1993).

Given the minor-groove nature of the interaction and the distance over which the enzyme acts, it was of interest to investigate the stoichiometry of the I-*Tev*I–DNA complex and to determine if the endonuclease alters the structure of the *td* homing site upon binding. Thus we have demonstrated that the 28 kDa I-*Tev*I binds its lengthy substrate as a monomer forming two distinct, catalytically active complexes, one of which exhibits a distortion within the DI domain near the I-*Tev*I cleavage site. The distortion includes a directed bend associated with nicking of the bottom strand at the cleavage site and with sequences flanking this site. A model of the I-*Tev*I–homing site interaction is presented, whereby formation of the nicked, catalytic complex facilitates the double-strand cleavage necessary to effect intron homing.

Results

I-TevI binds the td homing site as a monomer

A gel retardation assay based on a Ferguson (1964) analysis was used to determine the stoichiometry of the I-*Tev*I–DNA interaction, with a 37 bp fragment that represents the minimal homing site for endonuclease function (Bryk *et al.*, 1995). The molecular weight of the I-*Tev*I–homing site complex was determined by comparing its electrophoretic mobility with that of known protein standards on a series of nondenaturing gels of increasing



E

Protein	$-K_r$	MW Complex	Theoretical MW DNA	Theoretical MW Monomer	MW Protein	n
I-*Tev*I	4.56 (0.07)	51.06 (4.44)	24.01	28.22	27.05	0.96 (0.16)
*Bam*HI	6.67 (0.39)	76.56 (3.06)	30.50	24.57	46.06	1.87 (0.13)

Fig. 1. Stoichiometry of the I-*Tev*I–homing site interaction. (**A**) Sequence of DNA fragments. Fragments used to determine the stoichiometry of the I-*Tev*I–DNA and *Bam*HI–DNA (top strand) interactions are illustrated. *Bam*HI site, introduced by a 10 bp insertion at the intron insertion site (IS), is enlarged and bold. Open and filled arrowheads correspond to the insertion site and I-*Tev*I cleavage site (CS), respectively. (**B**) Representative mobility shift gel. DNA substrates were incubated in the absence (–) or presence (+) of I-*Tev*I or *Bam*HI and separated through a 10% polyacrylamide gel, to reveal the presence of the endonuclease–DNA complexes in the '+' lanes. Protein standards were separated alongside the complexes on gels ranging from 7 to 11% polyacrylamide (results not shown). (**C**) Representative Ferguson analysis. A logarithmic function of mobility for each of the protein standards and endonuclease–DNA complexes was plotted against the polyacrylamide concentration and fitted to a linear regression. Protein standards are indicated: a, α-lactalbumin (molecular weight 14.2 kDa); b, carbonic anhydrase (molecular weight 29.0 kDa); c, chicken egg albumin (molecular weight 45.0 kDa); d and e, bovine serum albumin monomer (molecular weight 66.0 kDa) and dimer (molecular weight 132.0 kDa), respectively. Endonuclease–DNA complexes, as indicated. (**D**) Representative plot of $-K_r$ versus molecular weight. $-K_r$ values for protein standards were derived by a Ferguson analysis and plotted as a function of molecular weight. Protein standards a–e are as indicated in (C). Interpolations of $-K_r$ values derived for the I-*Tev*I–DNA complex and the *Bam*HI–DNA complex (dashed lines) indicate molecular weights of ~51.06 and ~76.56 kDa, respectively. (**E**) Summary of data. Average molecular weights (kDa) of protein–DNA complexes were derived from the plots in (D). *n*, calculated stoichiometry of the reaction. Standard deviations for three independent trials are indicated in parenthesis.

polyacrylamide concentration (Figure 1). In this analysis, as gel concentration increases, the decreased mobility of each species is influenced solely by the size and shape of the molecule, because ionic conditions within the gel series are constant (Orchard and May, 1993). Therefore, the molecular weight of an unknown species can be determined by comparison of its electrophoretic mobility with that of known protein standards. Relative mobilities of the protein standards and the I-*Tev*I–homing site complex were determined and plotted as a function of gel concentration (Figure 1B and C). The slope of each line represents the retardation coefficient (K_r) for that species and is inversely related to its molecular weight (Figure 1D). Thus, by comparing the $-K_r$ value of the I-*Tev*I–DNA complex with those of the protein standards, we determined the molecular weight of the complex to be

51.06 kDa (with a standard deviation of ±4.44, over three trials). Taking into account the molecular weight of the DNA substrate (theoretical molecular weight 24.01 kDa), the molecular weight of the protein component of the complex was determined to be 27.05 kDa, indicating that I-*Tev*I (theoretical molecular weight 28.22 kDa) binds the *td* homing site as a monomer (Figure 1D and E).

To confirm the accuracy of the analysis, we determined the stoichiometry of restriction endonuclease *Bam*HI complexed to its target site. In this experiment, the DNA substrate was of similar sequence to the I-*Tev*I minimal homing site, but with a 10 bp insertion that introduces a *Bam*HI recognition site into the fragment (Figure 1A). Using Ferguson analyses and $-K_r$ comparisons (Figure 1B–E), the molecular weight of the *Bam*HI component of the protein–DNA complex was estimated to be 46.06 kDa (±3.06, over three trials), indicating that the restriction endonuclease, with a monomer molecular weight of 24.57 kDa, binds as a dimer, as demonstrated previously (Strzelecka *et al.*, 1995).

To verify independently that I-*Tev*I binds its DNA substrate as a monomer, mixed mobility shift experiments (Hope and Struhl, 1987) were performed with wild-type I-*Tev*I (28.22 kDa) and a larger derivative of the endonuclease (70.61 kDa) with I-*Tev*I fused to maltose binding protein (G.H.Silva, unpublished data). Complexes formed with a mixture of the two I-*Tev*I species exhibited electrophoretic mobilities characteristic of each protein individually. Similar results were obtained when wild-type I-*Tev*I and a truncated form of the endonuclease (13.21 kDa) that retains DNA binding function (V.Derbyshire, unpublished data) were used in mixing mobility shift experiments. Because no intermediate-sized complexes were detected (data not shown), the data corroborate the Ferguson analysis, indicating that I-*Tev*I binds as a monomer.

Circular permutation studies: a distortion near the cleavage site

Given that I-*Tev*I binds its 37 bp target sequence as a monomer, and considering that many minor-groove binding proteins bend their substrates (see Discussion), a conformational analysis of the DNA in the complex was undertaken. The anomalous electrophoretic mobility of curved or bent DNA through nondenaturing polyacrylamide gels is related to the end-to-end distance of the distorted molecule (Lerman and Frisch, 1982; Lumpkin and Zimm, 1982). Thus, DNA fragments containing a circularly permutated protein binding site can be used in gel retardation assays to study the effects of protein binding on DNA substrate conformation (Wu and Crothers, 1984). In these studies, as the protein binding site is moved from the end to the middle of the DNA fragment, if the helical axis of the substrate is distorted then the end-to-end distance of the fragment decreases, as does its electrophoretic mobility. It is worth noting that the observed differences in mobility do not result necessarily from DNA bending in a directed orientation, but may reflect isotropic bending of the helical axis or other structural alterations in the DNA or protein (Kerppola and Curran, 1993; Avitahl and Calame, 1994; Kahn *et al.*, 1994).

A 50 bp insert containing the *td* homing site (Figure 2A) was cloned into the pBend2 vector (Kim *et al.*, 1989). As shown in Figure 2B, restriction enzyme cleavage at sites that flank the insert results in the circularly permutated placement of the *td* homing site within the fragment. When these bending probes (Figure 2B, a–k) were incubated with I-*Tev*I in the absence of exogenous magnesium and electrophoresed through polyacrylamide gels under nondenaturing conditions, three *td*-related bands were evident: the most rapidly migrating band corresponds to unbound DNA (UB), whereas the two upper bands correspond to I-*Tev*I-complexed DNA (U_F and U_S). The anomalous migration of I-*Tev*I–DNA complexes indicates that I-*Tev*I induces a significant distortion in its DNA substrate within the slower-migrating complex (U_S; Figure

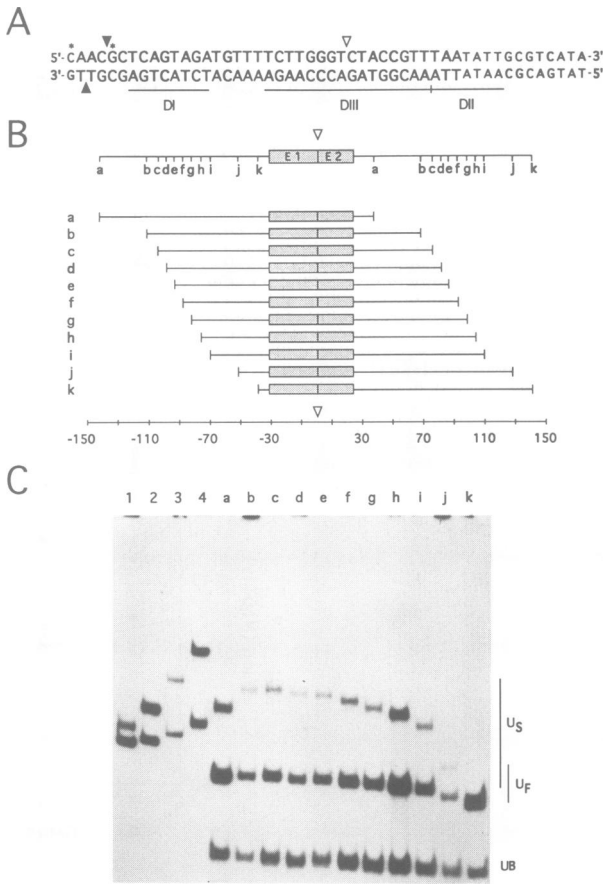


Fig. 2. Circular permutation analysis. (**A**) Sequence of the *td* homing site. Larger letters represent the extent of the *td* homing site sequences required for I-*Tev*I cleavage (Bryk *et al.*, 1995). Hypomutable domains DI, DII and DIII are underlined. Open and filled arrowheads are as explained in the legend to Figure 1. (*) Base pairs implicated in homing site flexibility. (**B**) Circular permutation probes for *td* homing site. Shaded boxes (E1, exon 1; E2, exon 2) represent a 50 bp *td* fragment, illustrated in (A), cloned into the pBend2 vector. Probes a–k were generated by cleavage with the restriction enzymes *Bam*HI, *Nru*I, *Stu*I, *Sma*I, *Pvu*II, *Eco*RV, *Dra*I, *Xho*I, *Spe*I, *Nhe*I and *Mlu*I, respectively. The numerical scale denotes the base pair positions of the *td* homing site in the bending probes, such that base pairs upstream of the insertion site are numbered from –1 and base pairs downstream of the insertion site are numbered from +1. (**C**) Gel mobility shift assay. Lanes 1–4 contain A_6-tract standards carrying two, three, four or five A_6 tracts, respectively (Thompson and Landy, 1988). Lanes a–k contain bending probes a–k incubated with I-*Tev*I and electrophoresed through an 8% polyacrylamide gel. U_S indicates the slower-migrating upper complex. U_F indicates the faster-migrating upper complex. UB indicates unbound probe.

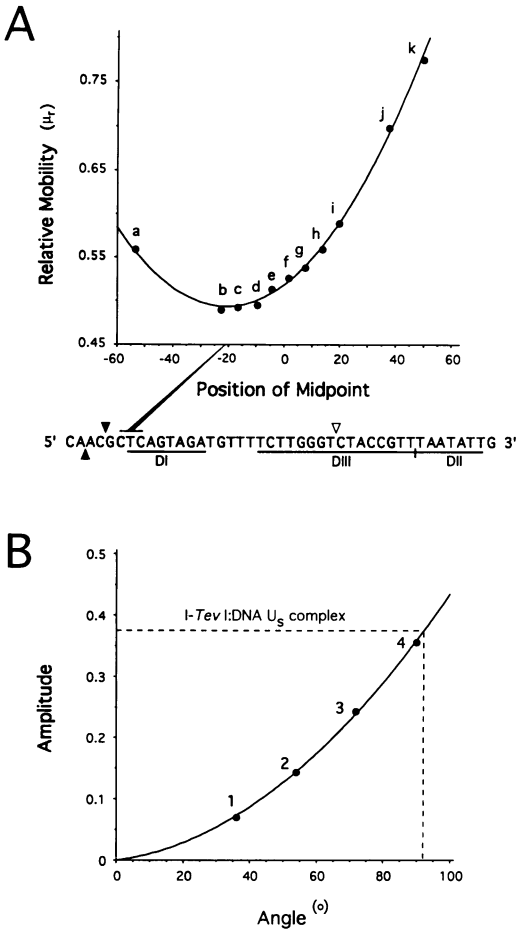


Fig. 3. Localization and magnitude of the distortion. (**A**) Mapping of the I-*Tev*I-induced distortion. The relative mobility (μ_r) of each of the bending probes (a–k) was plotted as a function of the midpoint position of the probe according to the scale in Figure 2B. The minimum of the curve reflects the position of the center of the distortion within the homing site. The average curve minimum for nine independent experiments was mapped to –20.8 (black diagonal line) ± 0.8 bp (horizontal line). (**B**) Magnitude of the I-*Tev*I-induced distortion. The amplitude [1 – (μ_{MID}/μ_{END})] for each of the A_6-tract standards (1–4) was plotted as a function of the estimated induced angle (see text). The graph corresponds to data from the experiment shown in Figure 2C, lanes 1–4. The dashed line represents the interpolation of the amplitude for the U_S complex from a best-fit circular permutation function (see Materials and methods, Equation 1).

2C, lanes a–k). Bending probes with the I-*Tev*I binding site located near the center form complexes that migrate substantially more slowly (Figure 2C, probes b and c, lanes b and c, respectively) than those whose binding sites are located near an end (Figure 2C, probe k, lane k). The single bound complex shown in Figure 1B corresponds to the situation in lane k, where the cleavage site is close to the end of the fragment (note that the single upper band in lane k corresponds to comigrating U_F and U_S complexes).

Representative data from a circular permutation analysis are displayed graphically in Figure 3A. The mobility of each of the U_S I-*Tev*I complexes relative to the unbound DNA was plotted as a function of the distance between the insertion site and the fragment midpoint for each of the bending probes. When the data from nine independent experiments were fitted by a second-order polynomial, the center of the I-*Tev*I-induced distortion mapped to position –20.8 ± 0.8 bp on the *td* homing site. Similar results (–20.2 ± 0.7 bp) were obtained when the data were fitted to a trigonometrically derived, circular permutation function (see Materials and methods, Equation 1; Kerppola and Curran, 1993). Thus, the I-*Tev*I-induced distortion was mapped to the DI domain, proximal to the I-*Tev*I cleavage site.

In the circular permutation analyses, slightly anomalous electrophoretic mobilities of the UB fragments and U_F complexes indicate that a subtle helical nonlinearity is also associated with these species. A distortion of the UB fragment, which becomes more apparent when separated with a uniform size marker in each lane (data not shown), was inferred previously from hydroxyl radical footprinting of the naked homing site (Bryk *et al.*, 1993). However, the subtlety of these flexures precludes measurement of the distortion.

The magnitude of the distortion

The magnitude of the I-*Tev*I-induced distortion was assessed by comparison with a set of bending standards (Thompson and Landy, 1988) that was separated on the mobility shift gels (Figure 2C, lanes 1–4) alongside the circularly permuted, I-*Tev*I–homing site complexes (lanes a–k). Each pair of bending standards consists of DNA fragments containing two to five A_6 tracts located near the end or middle of each fragment. When the A_6 tracts are located in the middle of the fragment, electrophoretic mobility is reduced. Consistent with previous studies, we have used a value of 18° as the bend angle generated by each A_6 tract (Koo *et al.*, 1990). Thus, the difference in electrophoretic mobilities between fragments can be correlated to a corresponding bend angle. A representative curve generated from the differential mobilities of the A_6-tract standards in Figure 2C is shown in Figure 3B. By interpolating the amplitude of the circular permutation function (Equation 1) for the I-*Tev*I–*td* homing site U_S complex, we estimated an I-*Tev*I-induced distortion magnitude of ~94 ± 1° from four independent trials. This result is in agreement with estimates of 90 ± 1° generated when the data were fitted to a cosine-related function describing the circular permutation analysis (see Materials and methods, Equation 2; Thompson and Landy, 1988; Kerppola and Curran, 1993). The magnitude of the I-*Tev*I-induced distortion reflects the electrophoretic consequences of a directed bend in the DNA substrate, other structural anomalies within the U_S complex, or both.

The orientation and magnitude of a directed I-TevI-induced bend

To determine if and to what extent the I-*Tev*I-induced distortion (Figures 2 and 3) resulted from a directed bend, we performed a phasing analysis on the I-*Tev*I–homing site interaction (Zinkel and Crothers, 1987; Salvo and Grindley, 1988; Kerppola and Curran, 1991a,b). Phasing probes containing the *td* homing site were generated by incrementally increasing, over a single helical turn, the distance between the center of the I-*Tev*I-induced distortion (–21 bp) and the center of a known intrinsic bend, consisting of three phased A_5 tracts (Figure 4A). If I-*Tev*I binding causes a directed bend, then the end-to-end distance of the probe and the electrophoretic mobility of the complex will be at a minimum when the I-*Tev*I-

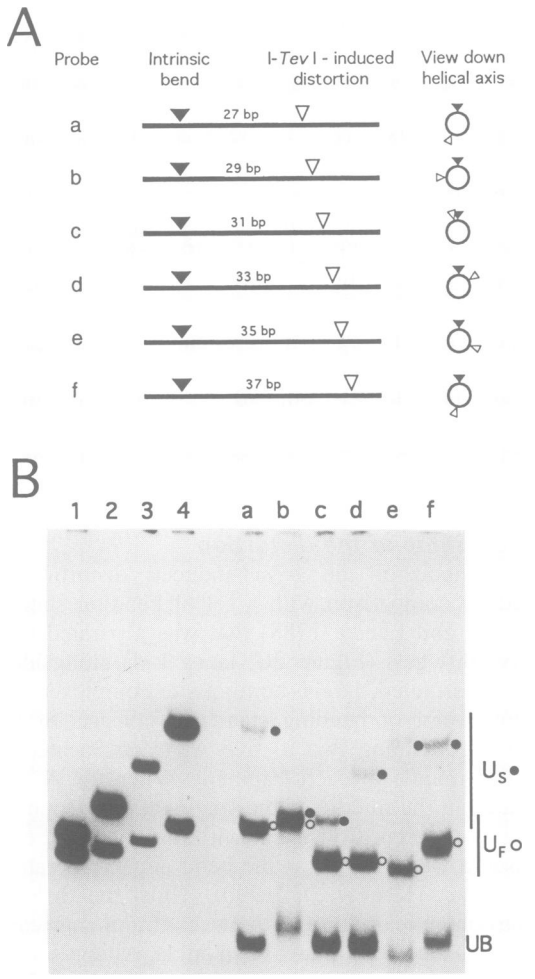


Fig. 4. Phasing analysis. (**A**) Schematic of phasing probes. Above each phasing probe a–f is the center-to-center distance (bp) between the I-*Tev*I-induced distortion (▽) and the center of an intrinsic A_5-tract bend (▼). The circles provide a view down the helical axis of the substrate, indicating the relative placement of the two bends. (**B**) Gel mobility shift analysis of phasing probes. Lanes 1–4 contain A_6-tract standards carrying two, three, four or five A_6 tracts, respectively. Lanes a–f contain phasing probes a–f incubated with I-*Tev*I and electrophoresed through an 8% polyacrylamide gel. UB, U_F and U_S are as in the legend to Figure 2. (●) U_S; (○) U_F. For phasing probe b (lane b), the U_F and U_S complexes comigrate, as determined by *in situ* cleavage analysis (data not shown).

induced bend and the intrinsic bend act cooperatively. Similarly, the end-to-end distance of the fragment and the electrophoretic mobility of the complex will be at a maximum when the two bends counteract. Considering that A_5 tracts bend DNA towards the minor groove (Zinkel and Crothers, 1987), in-phase cooperative bending and decreased electrophoretic mobility would indicate an I-*Tev*I-induced compression of the minor groove, while in-phase counteractive bending and increased electrophoretic mobility would indicate compression of the major groove.

Phasing probes, with center-to-center distances of 27, 29, 31, 33, 35 and 37 bp between the I-*Tev*I-induced bend and the intrinsic A_5-tract bend, were incubated with I-*Tev*I and separated using polyacrylamide gels (Figure 4B). For the fragment in which the center-to-center distance between the two bends was in phase (31 bp, assuming 10.5 bp per helical turn), the mobility of the U_S complex was fastest (Figure 4B, lane c). Conversely, the slowest mobilities were observed when the center-to-center distances were out of phase (Figure 4B, lanes a, e and f, representing separations of 27, 35 and 37 bp, respectively). These results indicate that I-*Tev*I bends the *td* homing site in a directed fashion towards the major groove. Interestingly, as seen in the circular permutation analysis, the variation in the mobilities of the unbound probes and the U_F complexes during the phasing analysis underscores the nonlinearity associated with the *td* homing site.

By fitting the data to a trigonometric phasing function (see Materials and methods, Equations 3 and 4; Kerppola and Curran, 1993), the magnitude of the I-*Tev*I-induced directed bend was estimated at 38 ± 2° from three independent trials. Differences between the observed distortion and the directed bending of the DNA target upon I-*Tev*I binding will be discussed.

A nick in the bottom strand is associated with the bend

To examine further the structure of the distorted I-*Tev*I–*td* homing site intermediate, we generated, by PCR, 304 bp bending probes that place the bend locus near the end and at the middle of the respective fragments (Figure 5A, probes m and n). Representative data of I-*Tev*I binding to fragments with the bend 10% from the end (probe m) and in the middle (probe n) of the fragment are shown in Figure 5B. As with the pBend2 probes (Figure 2), the position of the bend in the PCR-generated fragments had little influence on the electrophoretic mobility of the U_F complex. However, positioning greatly affected mobility of the U_S complex, reflecting a distortion. Taking advantage of our ability to isolate catalytically active I-*Tev*I–DNA complexes formed in the absence of divalent cations from polyacrylamide gels (Bryk *et al.*, 1995), we performed *in situ* cleavage assays to evaluate the cleavage status of the complexes generated with the two bending probes. Thus, unbound DNA and complexes formed with the two probes m and n (Figure 5B, UB, U_F and U_S) were treated within the gel with buffer containing either EDTA or Mg^{2+}, gel purified and re-electrophoresed through denaturing polyacrylamide gels (Figure 6A). The data indicate that in the faster-migrating, minimally distorted complexes (U_F), the DNA was largely intact, and that cleavage occurred after incubation with Mg^{2+}-containing buffer (Figure 6A). However, in the slower-migrating, distorted complexes (U_S), the DNA was nicked on the bottom strand at the cleavage site prior to the addition of exogenous Mg^{2+}. After incubation with the divalent cation, cleavage of the top strand took place, producing a double-strand break (more clearly evident in Figure 6A, END). In experiments where only the bottom strand was labeled, re-electrophoresis through denaturing gels indicated that >95% of the DNA in the U_S complex was nicked on the bottom strand (data not shown). *In situ* cleavage assays have been performed on a number of different *td* homing site fragments. In all cases, irrespective of the position of the bend, the slower-migrating bent complex (U_S) was associated with a nick in the bottom strand, whereas the faster-migrating unbent complex (U_F) was not (Bryk *et al.*, 1995; Figure 6A; data not shown).

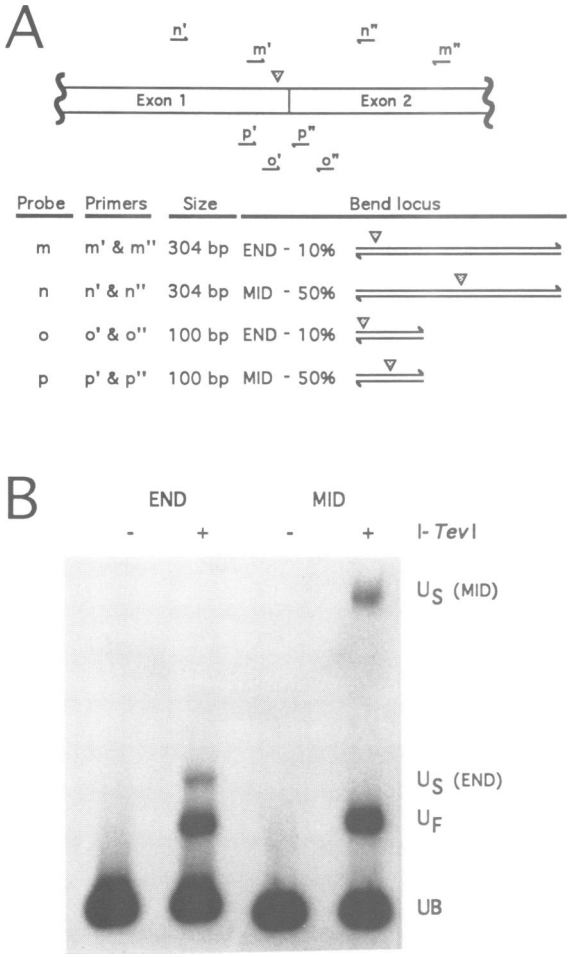


Fig. 5. Gel mobility shift analysis of PCR-generated I-*Tev*I bending probes. (**A**) I-*Tev*I bending probes. Oligonucleotide primers complementary to exon 1 (m′–p′) and exon 2 (m″–p″) of the *td*ΔIn gene were used to generate, via PCR, two 304 bp fragments (probes m and n) and two 100 bp fragments (probes o and p) which contain the *td* homing site. The I-*Tev*I-induced bend locus (stippled triangle) was placed either near the upstream end (10%) (probes m and o) or in the middle (50%) (probes n and p). (**B**) Representative mobility shift gel. Labeled bending probes m (END) and n (MID) were incubated in the absence (–) or presence (+) of I-*Tev*I and electrophoresed through 8% polyacrylamide gels. UB, U_F and U_S are as in Figure 2.

The role of the nick in bend formation

To examine the effect of a single nick on the structure of the *td* homing site, we performed a two-step, non-denaturing electrophoretic analysis (Figures 5B and 6B). In this experiment, we isolated the nicked *td* homing site from the slower-migrating I-*Tev*I complex (Figure 5B, U_S) and re-electrophoresed the naked DNA through native polyacrylamide gels. This was performed with DNA fragments that carry the I-*Tev*I cleavage site at the middle and near the end of the molecule, thus producing a permuted, nicked *td* homing site. Upon re-electrophoresis, the migration of the nicked homing sites was similar to that of the unnicked, intact counterparts (Figure 6B, compare U_S with UB). This suggests that while the bend within the active I-*Tev*I–DNA complex is associated with a break in the sugar–phosphate backbone, the helical distortion and altered mobility result directly from I-*Tev*I binding.

Additionally, DNA was isolated from the nicked U_S complex and used as the substrate for I-*Tev*I binding. After separation on a nondenaturing gel, the U_S complex, in which the homing site is distorted, was the predominant form (Figure 6C). These experiments suggest that the nick favors the formation of the bent complex.

The role of cleavage site sequences on distortion of the td homing site DNA

Mutational analyses of the *td* homing site indicate that sequences flanking the I-*Tev*I cleavage site affect the mobility of the nicked, catalytically active complexes through nondenaturing gels (Bryk *et al.*, 1995). In a study in which the 6 bp encompassing the cleavage site (5′-CAACGC-3′) were changed individually to each other possible base pair, it became apparent that base substitutions at positions G:C–23 and C:G–27 led to anomalous mobilities (Figure 2A, asterisks, and Table I). To examine a possible role for these sequences in conformational changes of the homing site, we performed gel mobility shift assays on bending probes (Figure 5A, probes o and p) carrying these variant homing sites either near the end or at the middle of each fragment. Of the six variants at positions G:C–23 and C:G–27 tested, the G–23C and all three C:G–27 (C–27A, C–27G and C–27T) mutants formed U_S complexes, albeit at greatly reduced efficiencies (<25% of wild type; Bryk *et al.*, 1995; Table I). Furthermore, the relative mobilities of these variant U_S complexes differed by 3–7% from that of the wild type (Table I, μ_R ranging from 1.03 to 1.07), representing a decrease in magnitude of the I-*Tev*I-induced distortion of between 2.8 and 6.6° relative to wild type. This was in contrast to the control mutant at position A:T–25, A–25C, which migrates like the wild-type *td* homing site (Bryk *et al.*, 1995; Table I). Thus, mutations at two positions, G:C–23 and C:G–27, which flank the sites of cleavage on the top and bottom strands reduce both the formation and the degree of distortion of the U_S complex, indicating that these base pairs act either directly or indirectly to influence the conformation of the homing site in the presence of I-*Tev*I.

Discussion

We have demonstrated that I-*Tev*I endonuclease binds the *td* homing site as a monomer, changing the conformation of the site as it forms the nicked, catalytically active U_S complex. Remarkably, the 28 kDa endonuclease is able to span over three turns of the helix to effect catalysis, consistent with genetic and physical data suggesting that I-*Tev*I interacts with its substrate via a flexible hinge (Bryk *et al.*, 1995). Interestingly, the unbound *td* homing site exhibits a small degree of helical nonlinearity (Figure 4B). This intrinsic structural feature might serve as a recognition element for the endonuclease, which binds in a sequence-tolerant fashion; alternatively, the structural anomaly might predispose the helix to bending upon contact with I-*Tev*I. The pronounced distortion in the U_S catalytic complex is associated with a directed bend and a sequence-dependent I-*Tev*I-induced nick at the bottom-strand cleavage site. The distortion near the cleavage site maps to the DI domain (Figure 3A), which was shown previously by mutational and physical analyses to be important for the formation of the catalytically active I-*Tev*I complexes (Bryk *et al.*, 1993, 1995). The data are

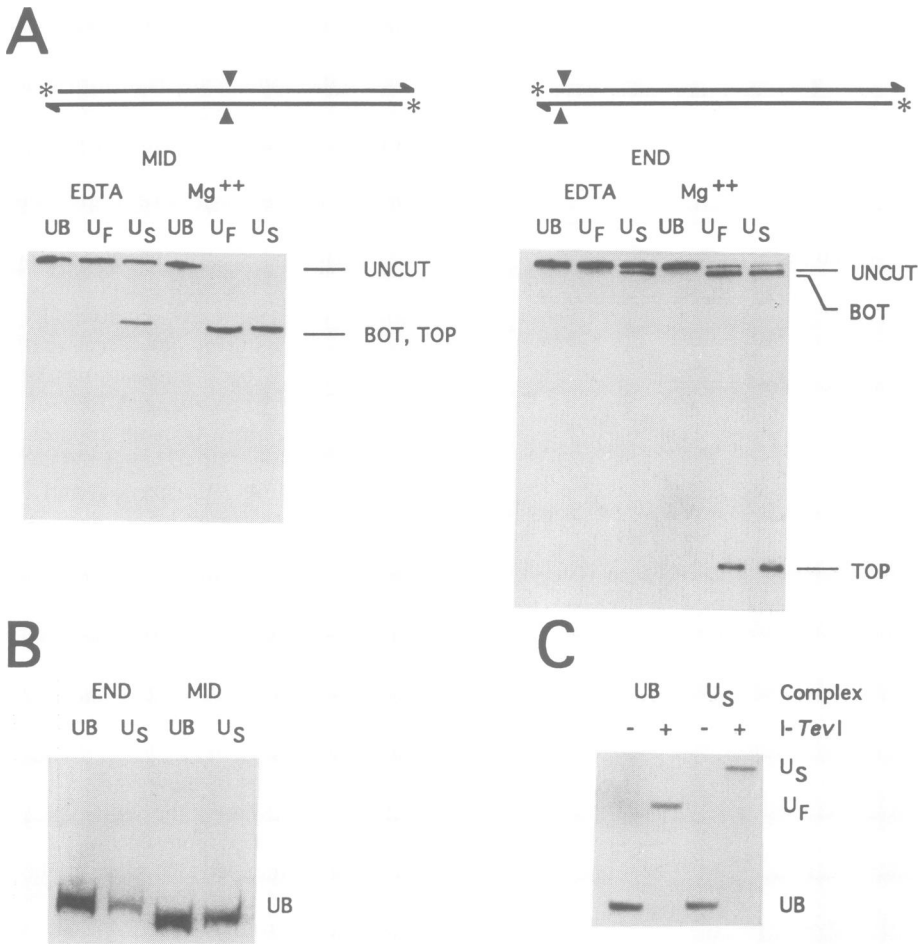


Fig. 6. Re-electrophoresis of I-*Tev*I–bending probe complexes. (**A**) Re-electrophoresis through denaturing polyacrylamide gels. Two individual mobility shift assays were performed as in Figure 5B with bend probes m (END) and n (MID). Gels were treated with either EDTA buffer or Mg^{2+} buffer. DNAs were eluted and re-electrophoresed through 8% polyacrylamide gels containing 8.3 M urea. A schematic representation of the homing sites is illustrated above each gel. Filled arrowheads, UB, U_F and U_S are as detailed in the legend to Figure 2. UNCUT, uncut substrate. BOT and TOP, labeled product resulting from I-*Tev*I cleavage of the bottom and top strands, respectively. Predictably for MID, cleavage products of the two strands comigrate. (**B**) Re-electrophoresis through native polyacrylamide gels. A mobility assay, as in Figure 5B, was performed. The gel was treated with EDTA, and purified DNAs from the UB and U_S bands were re-electrophoresed through 8% nondenaturing polyacrylamide gels. Labels are as in Figure 5B. (**C**) Gel mobility shift analysis of intact and nicked I-*Tev*I substrates. A 100 bp *td* homing site fragment which contains the I-*Tev*I-induced distortion in the middle of the fragment (Figure 5A, probe p) was incubated with I-*Tev*I and electrophoresed through a native 12% polyacrylamide gel. DNA from UB (intact) and U_S (nicked) complexes was purified from the gel, re-incubated in the absence (–) or presence (+) of I-*Tev*I and re-electrophoresed through a native 12% polyacrylamide gel, as indicated.

Table I. The effect of variant *td* homing sites on I-*Tev*I-induced flexure

Homing site[a]	Sequence[b]	Mobility[c]	U_S complex[d]	μ_R[e]	Change in distortion (°)[f]
Wild type	5′-CAACGC-3′	normal	++	1.00	0.0
G–23A	5′-CAAC**A**C-3′	aberrant	–	NA	NA
G–23C	5′-CAAC**C**C-3′	aberrant	+	1.07	–6.6
G–23T	5′-CAAC**T**C-3′	aberrant	–	NA	NA
C–27A	5′-**A**AACGC-3′	aberrant	+	1.07	–6.6
C–27G	5′-**G**AACGC-3′	aberrant	+	1.07	–6.6
C–27T	5′-**T**AACGC-3′	aberrant	+	1.03	–2.8
A–25C	5′-CA**C**CGC-3′	normal	++	1.00	0.0

[a]Homing sites analyzed were the wild-type and cleavage site variants listed (Bryk *et al.*, 1995).
[b]Base substitutions are bold in the top strand of the variants (nucleotides –27 to –22, 5′ to 3′).
[c]Mobility of the U_S and U_F complexes (Bryk *et al.*, 1995).
[d]U_S complex present in an amount similar to wild type (++), present in an amount substantially lower (<25%) than wild type (+), or not detected (–).
[e]Two 100 bp *td* homing site fragments which contain the I-*Tev*I-induced bend near the end (Figure 5A, probe o) or in the middle (Figure 5A, probe p) of the fragment were incubated with I-*Tev*I and electrophoresed through 12% polyacrylamide gels. μ_R represents the relative mobility of the slower-migrating nicked complex (U_S) for those fragments carrying the bend locus at the middle relative to fragments carrying the bend locus near the end, normalized to that of wild type. NA, not applicable, applies to those variants for which no U_S was detected.
[f]The magnitude of the change in distortion values is calculated based on a wild-type value of 94°.

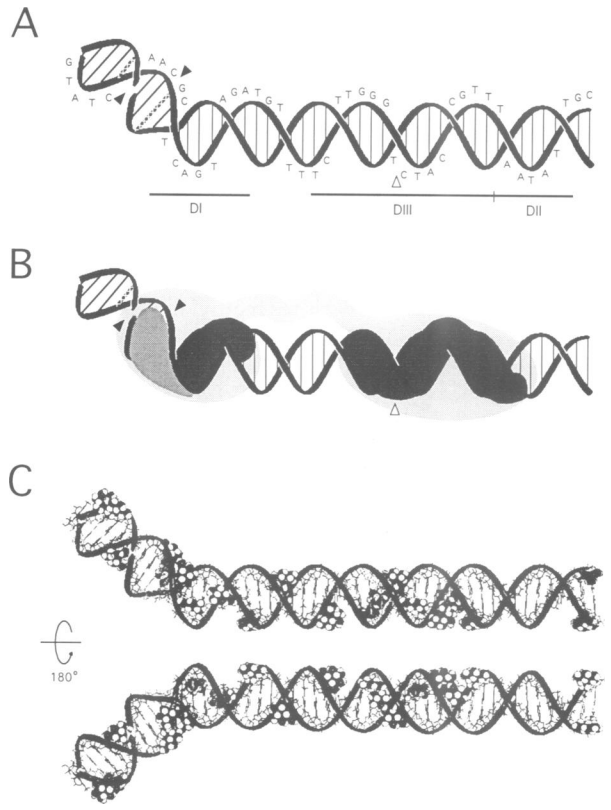


Fig. 7. Model of the *td* homing site. (**A**) Directed bending of the *td* homing site. The *td* homing site in the catalytically active nicked U_S complex was modeled from genetic and physical studies of the interaction with I-*Tev*I and the behavior of other minor-groove binding proteins that bend their DNA substrates (see Materials and methods). The model depicts a directed bend in the helical axis of the DNA of ~38°. Stippled base pairs represent G:C–23 and C:G–27, which have been implicated in the distortion of the homing site. Other labels are as in Figures 1 and 2. (**B**) Model of I-*Tev*I–*td* homing site interactions. Light shading represents I-*Tev*I; dark shading indicates regions of I-*Tev*I–DNA contact across the minor groove (Bryk *et al.*, 1993); intermediate shading represents regions of inferred contact based on functional assays (Bryk *et al.*, 1995) and depicts the protein accessing the top-strand cleavage site through the widened minor groove. The model represents a hinged monomer with a flexible association of I-*Tev*I subdomains (Bryk *et al.*, 1995). (**C**) Molecular modeling of the cytosine-modified *td* homing site. DNA is shown in a ball-and-stick configuration, with cytosines modified at the C5 positions and glucosylated hydroxymethyl groups as space-filled molecules. The lower duplex is a 180° rotation of the upper, exposing the opposite face of the helix.

consistent with a role for DI in positioning the I-*Tev*I-induced bend, thus facilitating access to the top-strand cleavage site via the minor groove (see below and Figure 7).

The role of the nick in distorting the homing site DNA

A number of studies have addressed how discontinuities in the sugar–phosphate backbone affect DNA flexibility. NMR and crystallographic analyses show that the structure of a nicked DNA duplex is similar to that of canonical B-DNA with minor perturbations flanking the break (Pieters *et al.*, 1989; Aymami *et al.*, 1990). Similarly, gel mobility analyses suggest no difference in helical rigidity between intact and nicked duplexes (Mills *et al.*, 1994). However, other studies have demonstrated that a nick in a 139 bp fragment of pBR322 DNA induces a bend of ~60°. The magnitude of the bend is enhanced further with protein binding (Le Cam *et al.*, 1994). Our results suggest that although the nick in the bottom strand of the *td* homing site has little effect on DNA conformation in the absence of protein (Figure 6B), it does facilitate the formation of the distorted complex by I-*Tev*I (Figure 6C).

As indicated by phasing analyses, the nick in the bottom strand of the U_S complex increases the flexibility of the bound substrate and allows I-*Tev*I to direct a 38° bend in the DNA towards the major groove (Figure 4). Interestingly, the anomalous electrophoretic mobilities observed during circular permutation studies reveal a distortion which correlates to a bend angle of ~94° (Figures 2 and 3). The difference between these two values indicates that structural anomalies, such as changes in helical conformation or local DNA melting, are present within the U_S complex, in addition to the directed bend.

Although I-*Tev*I forms two distinct cleavage-competent complexes (U_F and U_S) with its DNA substrate, only U_S is associated with the nick and the bend. The observed activity of the intact U_F complex in the presence of Mg^{2+} during *in situ* cleavage assays (Figure 6A) may result from conformational changes that occur within the gel matrix. Alternatively, U_F and U_S might represent two different modes of cleavage by the endonuclease. Indeed, in the case of restriction endonuclease *Eco*RV, the preference for an either sequential or concerted cleavage mechanism has been linked to divalent cation availability (Halford and Goodall, 1988; Taylor and Halford, 1989). Similarly, Mg^{2+} concentrations may influence the mechanism by which I-*Tev*I catalysis occurs.

A second phage T4 homing endonuclease I-*Tev*II, encoded by the *sun*Y intron, interacts with its DNA substrate across the minor groove, distorting the DNA helix upon binding (Loizos *et al.*, 1995). Interestingly, I-*Tev*II also forms two catalytically active complexes with its DNA substrate. However, unlike I-*Tev*I, both intact and nicked complexes are associated with a measurable distortion. It is unclear whether the observed distortion of the intact I-*Tev*II complex simply reflects the extent of a structural anomaly that is more pronounced than in the intact I-*Tev*I complex, or whether it signifies fundamental mechanistic differences between I-*Tev*I and I-*Tev*II catalysis.

Sequences flanking the cleavage site influence formation of the bent complex

The relative inefficiency with which I-*Tev*I forms the U_S complex upon binding base substitution variants at homing site positions –23 and –27 may result from altered protein–DNA contacts, although none have been mapped directly to these base pairs (Bryk *et al.*, 1993, 1995; Table I). Alternatively, changes at these positions may inhibit the distortion directly by increasing the rigidity of the DNA helix in the region of the cleavage site. Thus, by analogy to the sequence-dependent bending of the *lac* promoter by the *Escherichia coli* catabolite activator protein (Gartenberg and Crothers, 1988), the flexibility of the *td* homing site may be influenced by the nucleotide sequence.

Model of the I-TevI–homing site interaction

Functional studies on modified substrates suggest that I-*Tev*I interactions at both insertion and cleavage sites

occur via the minor groove (Bryk *et al.*, 1993, 1995). A number of studies investigating the interaction of minor-groove binding proteins and their substrates point to protein-induced DNA bending (Robertson and Nash, 1988; Suck *et al.*, 1988; Thompson and Landy, 1988; White *et al.*, 1989; Giese *et al.*, 1992; Kim *et al.*, 1993a,b; King and Weiss, 1993; Loizos *et al.*, 1995; Werner *et al.*, 1995). Structural (Suck *et al.*, 1988; Giese *et al.*, 1992; Kim *et al.*, 1993a,b; Feng *et al.*, 1994; Werner *et al.*, 1995) and modeling (White *et al.*, 1989) studies have indicated that proteins such as DNase I, LEF-1, TBP, Hin recombinase, IHF, HU and SRY distort their DNA substrates and widen the minor groove at the bend locus. Similarly, our results allow us to assign directionality to the I-*Tev*I-induced bend. Analogy to the above-mentioned minor-groove binding proteins coupled with our experimental data (Figures 1–6; Bryk *et al.*, 1993, 1995) have led us to propose a working model in which I-*Tev*I binding across the minor groove of the *td* homing site (Figure 7B, dark shading) results in a distortion of the helical path approaching the cleavage site. The depicted bend involves major-groove compression accompanied by widening of the minor groove, in a way that may increase accessibility to the endonuclease to effect cleavage of the top strand (Figure 7A and B, intermediate shading).

The conserved amino acid motif GIY–YIG has been identified in I-*Tev*I and fungal, mitochondrial intron open reading frames, as well as in the bacteriophage T4 site-specific endonuclease SegA (reviewed in Mueller *et al.*, 1993). This motif, which is proposed to be involved in endonucleolytic activity, is predicted by secondary structure analyses to form part of a β-ribbon at the N-terminus of I-*Tev*I (Bryk *et al.*, 1993). Interestingly, TBP, HU and Hin recombinase adopt a β-structure which is implicated in both DNA interactions across the minor groove and DNA bending. In a similar manner, our model proposes that I-*Tev*I interacts across the minor groove and distorts the *td* homing site upon nicking the bottom strand. The contacts of I-*Tev*I across the minor groove are depicted in Figure 7B in two regions, flanking the intron insertion site (DIII–DII) and approaching the cleavage site (DI), in accordance with previous genetic and physical studies (Bryk *et al.*, 1993, 1995).

The model depicts I-*Tev*I binding its DNA substrate as a monomer with a flexible hinge, reminiscent of the monomeric restriction endonuclease *Fok*I which cleaves its target nine and 13 nucleotides away from its recognition site (Li *et al.*, 1992; Li and Chandrasegaran, 1993). I-*Tev*I cleavage site displacement studies (Bryk *et al.*, 1995), together with the stoichiometric data presented here (Figure 1), serve as the basis for the proposed flexible tether between functional domains of the endonuclease. Furthermore, biochemical analyses indicate that I-*Tev*I comprises separate DNA binding and catalytic domains (V.Derbyshire, personal communication). Considering that polypeptide linkers between functional domains can lack secondary structure (Flick *et al.*, 1994; Klemm *et al.*, 1994), a tether between the primary DNA binding and endonucleolytic domains of I-*Tev*I would provide this 28 kDa endonuclease with the necessary flexibility to make contacts at the insertion site and to extend to DI and cleavage site sequences to effect catalysis (Figure 7B).

Naturally occurring modifications in the major groove of DNA are consistent with protein-induced distortions. Transcription factor TFI from the bacteriophage SP01, which is related to the minor-groove binding proteins HU and IHF (White *et al.*, 1989), binds and bends 5-hydroxymethyluracil-containing DNA (Sayre and Geiduschek, 1990; Schneider *et al.*, 1991). Likewise, the restriction enzyme *Sma*I bends both unmethylated and C5-methylated substrates (Withers and Dunbar, 1993). A *dam* methylation site introduced near the cleavage site of the *td* homing site had no deleterious effects on I-*Tev*I binding or cleavage when either methylated or unmethylated (Bryk *et al.*, 1995). Thus, the I-*Tev*I-induced distortion is not compromised by the presence of a methyl group within the major groove (at the –22 bp position) near the bend locus. Notably, bacteriophage T4 DNA contains glucosylated 5-hydroxymethylcytosine that imparts upon the DNA a wider, shallower major groove than canonical B-DNA (Saenger, 1984). Considering that the substrates used in our experimental system do not contain hydroxymethylcytosine, it remains unclear how the structural differences of phage DNA might affect I-*Tev*I binding and bending of the *td* homing site. However, competition studies indicate that I-*Tev*I has similar binding affinities for modified and unmodified substrates (data not shown). Furthermore, from a preliminary molecular modeling study, glucosylated hydroxymethyl groups at the C5 position of cytosines are consistent with the I-*Tev*I-induced bend (Figure 7C). Minimization data indicate that major-groove compression of the modified DNA is permitted without atomic overlap. The compatibility of major-groove compression with the modification of cytosine residues within the *td* homing site implies that the endonuclease functions on both natural and unmodified substrates in a manner similar to that of other minor-groove binding proteins.

Helical distortions, as observed with the I-*Tev*I–*td* homing site interaction, are energetically unfavorable (Liu-Johnson *et al.*, 1986; Zinkel and Crothers, 1991). Although the presence of the nick and the flexibility of homing site sequences contribute to the formation of the catalytically active U_S complex, free energy would still be required and remain stored within the bent complex. The potential role of this energy in the ensuing steps of the endonucleolytic reaction, as, for example, in the cleavage of the top strand or the release of bound I-*Tev*I from cleaved substrate to facilitate the homing process, remains a subject of speculation.

Materials and methods

DNA fragments and bending standards

The DNA fragments used to determine the stoichiometries were generated by PCR using end-labeled primers 5′-AACGCTCAGTAGATG-3′ and 5′-TTAAACGGTAG-3′. Because the cleavage site is located at the upstream end of the fragment, DNA distortion is of no consequence to the mobility of the fragment (Figures 1–3) for the purposes of the stoichiometric analysis. The I-*Tev*I target site was synthesized from template pBS*td*ΔIn (Bell-Pedersen *et al.*, 1989), and the *Bam*HI target site was synthesized from template pSU*td*ΔInOP21 (Clyman and Belfort, 1992).

Bending probes containing wild-type or variant *td* homing sites were generated by PCR in the presence of [α-^{32}P]dCTP unless otherwise stated. Primers complementary to sequences flanking the pBend2 cloning site, a′ (5′-CGGTGCCTGACTGCGTTAGC-3′) and a″ (5′-GGCGTATCACGAGGCCC-3′), were used to amplify the 50 bp *td* homing site.

PCR products were digested with the appropriate restriction enzymes to generate bending probes a–k (Figure 2B), which were then purified using 8% polyacrylamide gels. The 304 bp bending probes (Figure 5A, probes m and n) were generated using primers m′ (5′-GGCTATTTGGATTTGCAGTTG-3′) and m″ (5′-CTTTAGGCGTAAGTTTAAGAAC-3′), and primers n′ (5′-TATTGATCGTATTAAAAAACTGCC-3′) and n″ (5′-ACATTTTTCTACGTGATTC-3′), respectively. The 100 bp bending probes (Figure 5A, probes o and p) were generated using primers o′ (5′-TGCGTAATGGCTATTTGGATTTG*CAGCTG*TA-3′) and o″ (5′-GAATAAGATTACACATCTTAGCTAC-3′), and primers p′ (5′-CTATCAGTTTAATGTGCGTAATGGC-3′) and p″ (5′-AGCATATGACGCAATATTAAACGG-3′), respectively. Primers o′ and o″ generated a 126 bp fragment and introduced a *Pvu*II site (in italic in primer o′, variant nucleotides in bold) upstream of the I-*Tev*I cleavage site. Restriction with *Pvu*II generated a 100 bp fragment. This design allowed the synthesis of variant *td* homing sites (Table I) with a single set of primers. The 304 and 100 bp wild-type probes were generated from the template pBS*td*ΔIn (Bell-Pedersen *et al.*, 1989). Variant probes were generated from plasmids containing the *td* homing site with single base substitutions (Bryk *et al.*, 1995).

Phasing substrates were generated by cloning the duplexes formed with three different pairs of oligonucleotides [(i) 5′-TCGACCAACGCTCAGTAGATGTTTTCTTGGGTCTACCGTTTAATATTGCGTCATAT-3′ and 5′-CTAGATATGACGCAATATTAAACGGTAGACCCAAGAAAACATCTACTGAGCGTTGG-3′; (ii) 5′-TCGACGCGTCAACGCTCAGTAGATGTTTTCTTGGGTCTACCGTTTAATATTCATAT-3′ and 5′-CTAGATATGAATATTAAACGGTAGACCCAAGAAAACATCTACTGAGCGTTGACGCG-3′; and (iii) 5′-TCGACGCGTCATACAACGCTCAGTAGATGTTTTCTTGGGTCTACCGTTTAATATTT-3′ and 5′-CTAGAAATATTAAACGGTAGACCCAAGAAAACATCTACTGAGCGTTGTATGACGCG-3′] into the *Xba*I–*Sal*I site of pTK401-26 and pTK401-28 (Kerppola and Curran, 1991b). Six plasmids were produced, each carrying a 50 bp insert that contains the *td* homing site, with spacing between the center of the I-*Tev*I-induced distortion and the center of a bend intrinsic to the vector increasing incrementally (Figure 4A). Primers a′ and a″ were used to amplify these plasmids in the presence of [α-^{32}P]dCTP. PCR products were then digested with *Bam*HI to generate the appropriate phasing probes.

Plasmids carrying the A_6-tract bending standards (Thompson and Landy, 1988) were cleaved with *Nhe*I or *Bam*HI to position the bend locus near the middle or end, respectively, of each fragment. Bending standards were labeled with T4 DNA polymerase in the presence of [α-^{32}P]dCTP. All DNA fragments were purified from polyacrylamide gels prior to mobility shift assays.

Stoichiometric analysis

Protein–DNA (^{32}P-labeled) complexes were separated on a series of 7, 8, 9, 10 and 11% polyacrylamide gels alongside 10 μg of nondenatured protein molecular weight standards (Sigma). Gels were stained with Coomassie blue, destained, dried and exposed to X-ray film. For each species, 100[log(100μ)] was determined and plotted against gel concentration, where μ is equal to the mobility of the species relative to that of the bromophenol blue tracking dye. The negative slope or retardation coefficient ($-K_r$) was then plotted as a function of molecular weight for each protein standard (Figure 1D). The molecular weight of the DNA, assuming that a single base pair has a molecular weight of 649 Da, was subtracted from the estimated molecular weight of the protein–DNA complexes. This difference was then divided by the molecular weight of a protein monomer to determine the number of protein monomers bound to the DNA (n). Similar results were obtained when 0.25× and 1.00× TBE (90 mM Tris-borate, 2 mM EDTA, pH 8.0) were used.

Gel retardation analyses

I-*Tev*I was synthesized *in vitro* using wheat germ extracts, as described previously (Bell-Pedersen *et al.*, 1991). Mobility shift analyses were performed according to Bryk *et al.* (1993). Unbound probe and I-*Tev*I–DNA complexes were separated on 8 or 12% (29:1) polyacrylamide gels.

Circular permutation and phasing quantitation

Quantitative analyses for circular permutation and phasing studies have been described previously (Kerppola and Curran, 1993). For circular permutation analyses, data were fitted to a cosine function:

$$\mu = \mu_{max}[(A_{CP}/2)(\cos\{[(D-C_D)/P_{CP}]\ 2\pi\}-1) + 1], \quad (1)$$

where μ is the mobility of the I-*Tev*I–DNA complex relative to the unbound substrate, μ_{max} is the theoretical maximum mobility of the complex, A_{CP} is the amplitude of the cosine function, D is the distance, in bp, between the insertion site and the nearest end of the probe, with distances left of the insertion site given a negative (–) value, C_D is the center of distortion and P_{CP} is the period of the circular permutation function.

The magnitude of the distortion was estimated from the cosine function:

$$A_{CP} = 1-\cos(k\alpha_D/2), \quad (2)$$

where α_D is the magnitude of the distortion and k is the coefficient to adjust for electrophoretic conditions. Under our conditions, the value of k, derived from the electrophoretic mobilities of the A_6-tract bending standards, ranged from 1.11 to 1.18.

For phasing analyses, data were fitted to a phasing function:

$$\mu = \mu_{AVE}[(A_{PH}/2)\cos\{[(S-S_t)/P_{PH}]2\pi\} + 1], \quad (3)$$

where μ is the mobility of the I-*Tev*I–DNA complex relative to the unbound substrate, μ_{AVE} is the theoretical average mobility of the phasing complexes, A_{PH} is the amplitude of the phasing function, S is the center-to-center distance between the I-*Tev*I-induced distortion and the intrinsic A_5-tract bend, St is the center-to-center distance in which the I-*Tev*I-induced bend and the intrinsic A_5-tract bend counteract to maximize the end-to-end distance of the probe and P_{PH} is the period of the phasing function.

The magnitude of the directed bend angle was determined from the trigonometric function:

$$\tan(k\alpha_B/2) = (A_{PH}/2)/\tan(k\alpha_C/2), \quad (4)$$

where α_B is the magnitude of the I-*Tev*I-induced bend and α_C is the magnitude of the intrinsic bend, consisting of three phased A_5 tracts (54°).

In situ cleavage assays

Following gel retardation analyses, gels were painted with either Mg^{2+} buffer (50 mM Tris–HCl, pH 8.0, 10 mM $MgCl_2$, 100 mM NaCl) or EDTA buffer (50 mM Tris–HCl, pH 8.0, 30 mM EDTA) and allowed to incubate at room temperature for 20 min. Excess buffer was removed with absorbent paper, and gels were re-painted with EDTA buffer. After 5 min of incubation and the removal of excess buffer, gels were exposed to X-ray film. Bands containing unbound DNA and I-*Tev*I–DNA complexes were excised from the gel; DNAs were eluted, resuspended in 90% formamide, 20 mM EDTA, 0.05% bromophenol blue and 0.05% xylene cyanol and separated on 8% (19:1) polyacrylamide gels containing 8.3 M urea in 1× TBE buffer.

Two-step electrophoretic analysis

Following gel retardation analyses, gels were soaked in EDTA buffer for 20 min and exposed to X-ray film. DNAs from the appropriate bands were eluted and re-electrophoresed through polyacrylamide gels in 1× TBE buffer at 4°C. Intact and nicked substrates were subjected to gel retardation analyses as described above.

Molecular modeling

Molecular modeling images and data were generated on a Silicon Graphics (SGI) workstation using Insight II software (Biosym Technologies Inc.) and modeled from data describing the interaction of I-*Tev*I with its DNA substrate (Bryk *et al.*, 1993, 1995; this work) as well as studies involving other minor-groove binding proteins (Suck *et al.*, 1988; White *et al.*, 1989; Kim *et al.*, 1993a,b; Feng *et al.*, 1994; Suzuki and Yagi, 1995; Werner *et al.*, 1995). Because chemical footprinting and interference analyses (Bryk *et al.*, 1993) and minimal homing site studies (Bryk *et al.*, 1995) indicate that sequences upstream of the I-*Tev*I cleavage site are not essential for I-*Tev*I function, the model assumes the simplest interpretation of the data, i.e. a single bend locus. The unmodified and glucosylated hydroxymethylated *td* homing sites were generated using the Biopolymer and Builder programs, with base pairs T:A–21, C:G–22 and G:C–23 built into the *td* homing site with a twist of –24° and base pair T:A–21 with a positive roll of 40°. These changes allowed us to generate a directed bend of 38° to widen the minor groove of the DNA approaching the I-*Tev*I-induced nick and top-strand cleavage site, increasing the O4′ to O4′ distance from 6.9 to 10.5 Å, and to maintain the helical properties of canonical B-DNA throughout the remainder of the homing site. These changes were modeled after minor-groove binding proteins that bend their DNA substrates, where structural studies indicate that directed bends result from the introduction of a positive base pair roll that is associated with an asymmetric unwinding of the DNA in the region flanking the bend locus (Kim *et al.*, 1993a,b; Suzuki and Yagi, 1995; Werner *et al.*, 1995). Minimization calculations were performed using the Discovery program, as recommended by the manufacturer.

Acknowledgements

We are grateful to Dr Vicky Derbyshire and Mr George H.Silva for I-*Tev*I derivatives and sharing unpublished data, and to Dr Nick Loizos for sharing unpublished data. We thank Dr Sankar Adhya for the pBend2 circular permutation vectors, Drs John Thompson and Art Landy for the A_6-tract bending standards and Dr Tom Kerppola for the phasing vectors and helpful discussions regarding circular permutation and phasing analyses. We also thank members of the Belfort laboratory for critical comments about the manuscript, and Maryellen Carl for preparing the manuscript. This work was funded by National Institutes of Health grants GM39422 and GM44844 to M.B. and GM15454 to J.E.M.

References

Avitahl,N. and Calame,K. (1994) The C/EBP family of proteins distorts DNA upon binding but does not introduce a large directed bend. *J. Biol. Chem.*, **269**, 23553–23562.

Aymami,J., Coll,M., van der Marel,G.A., van Boom,J.H., Wang,A.H.-J. and Rich,A. (1990) Molecular structure of nicked DNA: a substrate for DNA repair enzymes. *Proc. Natl Acad. Sci. USA*, **87**, 2526–2530.

Belfort,M. (1990) Phage T4 introns: self-splicing and mobility. *Annu. Rev. Genet.*, **24**, 363–385.

Bell-Pedersen,D., Quirk,S.M., Aubrey,M. and Belfort,M. (1989) A site-specific endonuclease and co-conversion of flanking exons associated with the mobile *td* intron of phage T4. *Gene*, **82**, 119–126.

Bell-Pedersen,D., Quirk,S., Clyman,J. and Belfort,M. (1990) Intron mobility in phage T4 is dependent upon a distinctive class of endonucleases and independent of DNA sequences encoding the intron core: mechanistic and evolutionary implications. *Nucleic Acids Res.*, **18**, 3763–3770.

Bell-Pedersen,D., Quirk,S.M., Bryk,M. and Belfort,M. (1991) I-*Tev*I, the endonuclease encoded by the mobile *td* intron, recognizes binding and cleavage domains on its DNA target. *Proc. Natl Acad. Sci. USA*, **88**, 7719–7723.

Bryk,M., Quirk,S.M., Mueller,J.E., Loizos,N., Lawrence,C. and Belfort,M. (1993) The *td* intron endonuclease makes extensive sequence-tolerant contacts across the minor groove of its DNA target. *EMBO J.*, **12**, 2141–2149.

Bryk,M., Belisle,M., Mueller,J.E. and Belfort,M. (1995) Selection of a remote cleavage site by I-*Tev*I, the *td* intron-encoded endonuclease. *J. Mol. Biol.*, **247**, 197–210.

Chu,F.K., Maley,G., Pedersen-Lane,J., Wang,A. and Maley,F. (1990) Characterization of the restriction site of a prokaryotic intron-encoded endonuclease. *Proc. Natl Acad. Sci. USA*, **87**, 3574–3578.

Clyman,J. and Belfort,M. (1992) *Trans* and *cis* requirements for intron mobility in a prokaryotic system. *Genes Dev.*, **6**, 1269–1279.

Feng,J., Johnson,R.C. and Dickerson,R.E. (1994) Hin recombinase bound to DNA: the origin of specificity in major and minor groove interactions. *Science*, **263**, 348–355.

Ferguson,K.A. (1964) Starch gel electrophoresis – application to the classification of pituitary proteins and polypeptides. *Metabolism*, **13**, 985–1002.

Flick,K.E., Gonzalez,L.,Jr, Harrison,C.J. and Nelson,H.C.M. (1994) Yeast heat shock transcription factor contains a flexible linker between the DNA-binding and trimerization domains. *J. Biol. Chem.*, **269**, 12475–12481.

Gartenberg,M.C. and Crothers,D.M. (1988) DNA sequence determinants of CAP-induced bending and protein binding affinity. *Nature*, **333**, 824–829.

Giese,K., Cox,J. and Grosschedl,R. (1992) The HMG domain of lymphoid enhancer factor 1 bends DNA and facilitates assembly of functional nucleoprotein structures. *Cell*, **69**, 185–195.

Halford,S.E. and Goodall,A.J. (1988) Modes of DNA cleavage by the *Eco*RV restriction endonuclease. *Biochemistry*, **27**, 1771–1777.

Hope,I.A. and Struhl,K. (1987) GCN4, a eukaryotic transcriptional activator protein, binds as a dimer to target DNA. *EMBO J.*, **6**, 2781–2784.

Kahn,J.D., Yun,E. and Crothers,D.M. (1994) Detection of localized DNA flexibility. *Nature*, **368**, 163–166.

Kerppola,T.K. and Curran,T. (1991a) DNA bending by Fos and Jun: the flexible hinge model. *Science*, **254**, 1210–1214.

Kerppola,T.K. and Curran,T. (1991b) Fos–Jun heterodimers and Jun homodimers bend DNA in opposite orientations: implications for transcription factor cooperativity. *Cell*, **66**, 317–326.

Kerppola,T.K. and Curran,T. (1993) DNA bending by Fos and Jun: structural and functional implications. In Eckstein,F. and Lilley,D.M.J. (eds), *Nucleic Acid and Molecular Biology 7*. Springer-Verlag, New York, pp. 70–105.

Kim,J., Zwieb,C., Wu,C. and Adhya,S. (1989) Bending of DNA by regulatory proteins: construction and use of a DNA bending vector. *Gene*, **85**, 15–23.

Kim,J.L., Nikolov,D.B. and Burley,S.K. (1993a) Co-crystal structure of TBP recognizing the minor groove of a TATA element. *Nature*, **365**, 520–527.

Kim,Y., Geiger,J.H., Hahn,S. and Sigler,P.B. (1993b) Crystal structure of a yeast TBP/TATA-box complex. *Nature*, **365**, 512–519.

King,C.-Y. and Weiss,M.A. (1993) The SRY high-mobility-group box recognizes DNA by partial intercalation in the minor groove: a topological mechanism of sequence specificity. *Proc. Natl Acad. Sci. USA*, **90**, 11990–11994.

Klemm,J.D., Rould,M.A., Aurora,R., Herr,W. and Pabo,C.O. (1994) Crystal structure of the Oct-1 POU domain bound to an octamer site: DNA recognition with tethered DNA-binding modules. *Cell*, **77**, 21–32.

Koo,H., Drak,J., Rice,J.A. and Crothers,D.M. (1990) Determination of the extent of DNA bending by an adenine–thymine tract. *Biochemistry*, **29**, 4227–4234.

Le Cam,E., Fack,F., Menissier-de Murcia,J.M., Cognet,J.A.H., Barbin,A., Sarantoglou,V., Revet,B., Delain,E. and de Murcia,G. (1994) Conformational analysis of a 139 base-pair DNA fragment containing a single-stranded break and its interaction with human poly(ADP-ribose) polymerase. *J. Mol. Biol.*, **235**, 1062–1071.

Lerman,L.S. and Frisch,H.L. (1982) Why does the electrophoretic mobility of DNA in gels vary with the length of the molecule? *Biopolymers*, **21**, 995–997.

Li,L. and Chandrasegaran,S. (1993) Alteration of the cleavage distance of *Fok*I restriction endonuclease by insertion mutagenesis. *Proc. Natl Acad. Sci. USA*, **90**, 2764–2768.

Li,L., Wu,L.P. and Chandrasegaran,S. (1992) Functional domains in *Fok*I restriction endonuclease. *Proc. Natl Acad. Sci. USA*, **89**, 4275–4279.

Liu-Johnson,H.-N., Gartenberg,M.R. and Crothers,D.M. (1986) The DNA binding domain and bending angle of the *E. coli* CAP protein. *Cell*, **47**, 995–1005.

Loizos,N., Silva,G.H. and Belfort,M. (1995) The intron-encoded endonuclease I-*Tev*II binds across the minor groove and induces two distinct conformational changes in its DNA structure. *J. Mol. Biol.*, in press.

Lumpkin,O.J. and Zimm,B.H. (1982) Mobility of DNA in gel electrophoresis. *Biopolymers*, **21**, 2315–2316.

Mills,J.B., Cooper,J.P. and Hagerman,P.J. (1994) Electrophoretic evidence that single-stranded regions of one or more nucleotides dramatically increase the flexibility of DNA. *Biochemistry*, **33**, 1797–1803.

Mueller,J.E., Bryk,M., Loizos,N. and Belfort,M. (1993) Homing endonucleases. In Linn,S.M., Lloyd,R.S. and Roberts,R.J. (eds), *Nucleases*. 2nd edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 111–143.

Orchard,K. and May,G.E. (1993) An EMSA-based method for determining the molecular weight of a protein–DNA complex. *Nucleic Acids Res.*, **21**, 3335–3336.

Pieters,J.M.L., Mans,R.M.W., van den Elst,H., van der Marel,G.A., van Boom,J.H. and Altona,C. (1989) Conformational and thermodynamic consequences of the introduction of a nick in duplexed DNA fragments: an NMR study augmented by biochemical experiments. *Nucleic Acids Res.*, **17**, 4551–4565.

Robertson,C.A. and Nash,H.A. (1988) Bending of the bacteriophage lambda attachment site by *Escherichia coli* integration host factor. *J. Biol. Chem.*, **263**, 3554–3557.

Saenger,W. (1984) *Principles of Nucleic Acid Structure*. Springer-Verlag, New York.

Salvo,J.J. and Grindley,N.D.F. (1988) The gd resolvase bends the res site into a recombinogenic complex. *EMBO J.*, **7**, 3609–3616.

Sayre,M.H. and Geiduschek,E.P. (1990) Effects of mutations at amino acid 61 in the arm of TF1 on its DNA-binding properties. *J. Mol. Biol.*, **216**, 819–833.

Schneider,G.J., Sayre,M.H. and Geiduschek,E.P. (1991) DNA-bending properties of TF1. *J. Mol. Biol.*, **221**, 777–794.

Strzelecka,T.E., Hayes,J.J., Clore,G.M. and Gronenborn,A.M. (1995) DNA binding specificity of the Mu Ner protein. *Biochemistry*, **34**, 2946–2955.

Suck,D., Lahm,A. and Oefner,C. (1988) Structure refined to 2 Å of a

nicked DNA octanucleotide complex with DNase I. *Nature*, **332**, 464–468.

Suzuki,M. and Yagi,N. (1995) Stereochemical basis of DNA bending by transcription factors. *Nucleic Acids Res.*, **23**, 2083–2091.

Taylor,J.D. and Halford,S.E. (1989) Discrimination between DNA sequences by the *Eco*RV restriction endonuclease. *Biochemistry*, **28**, 6198–6207.

Thompson,J.F. and Landy,A. (1988) Empirical estimation of protein-induced DNA bending angles: application to lambda site-specific recombination complexes. *Nucleic Acids Res.*, **16**, 9687–9705.

Werner,M.H., Huth,J.R., Gronenborn,A.M. and Clore,G.M. (1995) Molecular basis of human 46X,Y sex reversal revealed from the three-dimensional solution structure of the human SRY–DNA complex. *Cell*, **81**, 705–714.

White,S.W., Appelt,K., Wilson,K.S. and Tanaka,I. (1989) A protein structural motif that bends DNA. *Proteins: Struct. Funct. Genet.*, **5**, 281–288.

Withers,B.E. and Dunbar,J.C. (1993) The endonuclease isoschizomers, *Sma*I and *Xma*I, bend DNA in opposite orientations. *Nucleic Acids Res.*, **21**, 2571–2577.

Wu,H.-M. and Crothers,D.M. (1984) The locus of sequence-directed and protein-induced DNA bending. *Nature*, **308**, 509–513.

Zinkel,S.S. and Crothers,D.M. (1987) DNA bend direction by phase sensitive detection. *Nature*, **328**, 178–181.

Zinkel,S.S. and Crothers,D.M. (1991) Catabolite activator protein-induced DNA bending in transcription initiation. *J. Mol. Biol.*, **219**, 201–215.

Received on September 27, 1994; revised on August 1, 1995