Proc. Natl. Acad. Sci. USA
Vol. 88, pp. 7719–7723, September 1991
Biochemistry

I-*Tev*I, the endonuclease encoded by the mobile *td* intron, recognizes binding and cleavage domains on its DNA target

(intron homing site/endonuclease–DNA interaction)

DEBORAH BELL-PEDERSEN*†‡, SUSAN M. QUIRK*, MARY BRYK*§, AND MARLENE BELFORT*¶‖

*Molecular Genetics Program, Wadsworth Center for Laboratories and Research, New York State Department of Health, Empire State Plaza, Albany, NY 12201-0509; †Biology Department, State University of New York at Albany, Albany, NY 12222; §Department of Microbiology and Immunology, Albany Medical College, Albany, NY 12208; and ¶School of Public Health, State University of New York at Albany, Empire State Plaza, Albany, NY 12201-0509

Communicated by Melvin I. Simon, June 7, 1991

ABSTRACT **Mobility of the phage T4 *td* intron depends on activity of an intron-encoded endonuclease (I-*Tev*I), which cleaves a homologous intronless (ΔIn) target gene. The double-strand break initiates a recombination event that leads to intron transfer. We found previously that I-*Tev*I cleaves *td*ΔIn target DNA 23–26 nucleotides upstream of the intron insertion site. DNase I-footprinting experiments and gel-shift assays indicate that I-*Tev*I makes primary contacts around the intron insertion site. A synthetic DNA duplex spanning the insertion site but lacking the cleavage site was shown to bind I-*Tev*I specifically, and when cloned, to direct cleavage into vector sequences. The behavior of the cloned duplex and that of deletion and insertion mutants support a primary role for sequences surrounding the insertion site in directing I-*Tev*I binding, conferring cleavage ability, and determining cleavage polarity. On the other hand, sequences around the cleavage site were shown to influence cleavage efficiency and cut-site selection. The role of cleavage-site sequences in determining cleavage distance argues against a strict "ruler" mechanism for cleavage by I-*Tev*I. The complex nature of the homing site recognized by this unusual type of endonuclease is considered in the context of intron spread.**

Mobility of the *td* and *sun*Y introns of phage T4 is mediated by endonucleases I-*Tev*I and I-*Tev*II, respectively (1–3). These enzymes, encoded by the introns themselves, recognize and cleave intronless alleles, generating double-strand breaks in the vicinity of the intron insertion sites. Cleavage initiates recombination between the cut strands and homologous exon sequences of the intron-containing allele, which acts as template for repair synthesis. This process, termed "intron homing" (4, 5) was initially described for the group I omega intron residing in the large rRNA gene of *Saccharomyces cerevisiae* mitochondria (for review, see refs. 4 and 6). The term "homing site" in reference to these mobile introns defines a composite of the intron insertion site, the endonuclease-binding and cleavage sites, as well as sequences involved in catalysis at the RNA level (5).

The endonucleases that promote intron mobility are quite different in pro- and eukaryotes (3, 7). Most notably, the eukaryotic endonucleases cleave within one residue of the intron insertion site (8–11), whereas the phage proteins cleave at a distance (3, 12). The *td* endonuclease I-*Tev*I cuts target DNA 23–26 nucleotides (nt) upstream of the intron-insertion site, generating 2- or 3-nt 3′-OH extensions.

Given the distance between the *td* intron insertion site and the cleavage site, it is of interest to determine how I-*Tev*I interacts with the homing site. In this work we show that sequences spanning the intron insertion site but lacking the cleavage site direct binding and cleavage by I-*Tev*I. However, cleavage-site sequences are also involved in recognition by I-*Tev*I and influence both cleavage efficiency and cut-site selection.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

MATERIALS AND METHODS

PCR Primers. The 5′ and 3′ *td*-specific primers are NBS4 (5′-AAAGTCGACTGCAGCATTACCGCCTTGTCAT-3′) and NBS5 (5′-AAATCTAGAGGTACCCCTGGAATAAGATTACAC-3′), respectively. The pBSM13⁻-specific (Stratagene) primers are NBS14 (5′-AAAATCGATAACGACGGCCACTGAATTGT-3′), NBS15 (5′-AAAATCGATCAAGCTCGGAATTAACCCTC-3′), and NBS16 (5′-AAAATCGATATTTCCCCGAAAAGTGCCAC-3′), complementary to nt 839–860, 952–973, and 3113–3132, respectively.

Phages, Plasmids, and PCR Fragments. Phage T4*td*ΔIn contains a precise deletion of the *td* intron (2); T2L is a naturally occurring intronless phage (13).

The complementary *td* homing-site oligodeoxynucleotides containing *Sal* I ends (see Fig. 2) were cloned into the *Sal* I site of pBSM13⁻ to yield pBSHS31⁺ (sense orientation relative to pT7) and pBSHS31⁻ (opposite orientation) (see Fig. 3). Plasmids pBS*td*ΔIn and pTZ*td*ΔIn contain the 1.8-kilobase (kb) *Eco*RI *td*ΔIn fragment in pBSM13⁻ (Stratagene) and pTZ18U (United States Biochemical), respectively, in transcriptional alignment with pT7 (2). Plasmids with deletions at the cleavage site, pBS*td*F1 and pBS*td*F5, were constructed by digesting pBS*td*ΔIn with I-*Tev*I, followed by deletion, ligation, transformation into *Escherichia coli* RRIΔ*thyA*::Km^R and selection for the Td⁻ phenotype (14) (see Fig. 4*B*). Construct pBS*td*HS+10 has a 10-base-pair (bp) insertion (3) and pBS*td*HSΔ21 has a 21-bp deletion at the intron insertion site (1) (Fig. 4*B*). To generate insertions and deletions between the cleavage and intron insertion sites, an *Xba* I site was created in pTZ*td*ΔIn lacking the *Xba* I site within the vector; this was achieved by inserting 2 bp (CT) at the site indicated in Fig. 5 to generate pTZ*td*IC+2. pTZ*td*IC+2 was digested with *Xba* I and mung bean nuclease to generate 3-nt and 4-nt deletions. A construct with a single cytosine insertion was also generated (see Fig. 5).

PCR amplification was used to generate 160- to 190-bp fragments of homing-site variants cloned for use in footprinting and mobility-shift assays. The *td*-specific primers NBS4 and NBS5 create fragments with 96 bp 5′ and 73 bp 3′ to the intron-insertion site, with *Sal* I and *Pst* I sites at the 5′ end and *Kpn* I and *Xba* I sites at the 3′ end. PCR fragments were generated with these primers from T4*td*ΔIn, T2L, pBS*td*F1, pBS*td*HS+10, and pBS*td*HSΔ21. A PCR fragment from pBS*td*F5 was obtained with pBSM13-specific primer NBS16,

Abbreviations: nt, nucleotide(s); *td*ΔIn, intronless copy of the *td* gene.

‡Present address: Department of Biochemistry, Dartmouth Medical School, Hanover, NH 03755.
‖To whom reprint requests should be addressed.

which contains a 5′ *Cla* I site, and NBS5. PCR fragments were synthesized from pBSHS31$^+$ and pBSHS31$^-$ with pBSM13-specific primers NBS14 and NBS15. These PCR fragments were cloned via corresponding restriction enzyme sites into pBSKS$^+$ (Stratagene) or pBSM13$^-$ and used in gel-shift assays (see Fig. 4*B*).

I-*Tev*I Preparations. I-*Tev*I was synthesized *in vitro* in rabbit reticulocyte lysates as described (2) or *in vivo* in *E. coli* BL21(DE3) after transcriptional amplification of the endonuclease coding sequence under T7 promoter control in pET-3b (ref. 15; to be described elsewhere). Active I-*Tev*I was isolated from inclusion bodies by solubilization with 6 M guanidine hydrochloride, followed by dialysis against 50 mM potassium phosphate (pH 8.0), 20 mM mercaptoethanol, and 0.5 M NaCl (16). One unit of activity is defined as the amount of enzyme required to cleave 250 ng of linearized target DNA (5.0 kb) to 50% completion in 1 min at 37°C. Cleavage-site mapping and activity assays were performed with the more active *in vitro* preparations (≈10 units/μl; 100 μg of total protein per μl), whereas binding studies were performed with more highly purified *in vivo* preparations (0.2–0.5 unit/μl; 0.5 μg/μl total protein). Cleavages mapped with both preparations were identical, although some differences in binding characteristics were noted (unpublished work).

DNase I Footprinting. PCR-amplified DNA (*td*ΔIn) digested with *Sal* I and *Xba* I was cloned into pBSKS$^+$ to yield pBSKS*td*ΔIn.2. *Sal* I–*Kpn* I- or *Xba* I–*Pst* I-digested pBSKS*td*ΔIn.2 was uniquely labeled on one strand with [α-^{32}P]dTTP and reverse transcriptase, and the fragments were isolated from native 8% polyacrylamide gels. Footprinting was done essentially as described (17). ^{32}P-labeled DNA fragment (0.1 pmol) was incubated with various amounts of I-*Tev*I (*in vivo* preparation) in 40 μl of binding buffer (50 mM Tris·HCl, pH 7.4/60 mM KCl/bovine serum albumin at 2 mg/ml/1 mM $MgCl_2$/1 mM 2-mercaptoethanol) for 5 min on ice. DNase I was added (16 μg/ml; BRL), samples were incubated for 2 min on ice, and reactions were terminated with 45 μl of DNase I stop solution (20 mM EDTA, sonicated salmon sperm DNA at 67 μg/ml). Samples were phenol-extracted, ethanol-precipitated, resuspended in 6 μl of loading buffer [95% (vol/vol) formamide/20 mM EDTA/0.05% xylene cyanol/0.05% bromphenol blue] and separated on an 8% acrylamide/8 M urea gel alongside Maxam–Gilbert sequencing ladders of the same fragments.

Gel Mobility-Shift Assays. Complementary end-labeled oligodeoxynucleotides (≈1 pmol, 10,000 cpm) were hybridized in 80 μl of 70 mM Tris·HCl, pH 8.0/100 mM NaCl/10 mM dithiothreitol at 65°C for 30 min, followed by 30 min at 25°C. Bovine serum albumin and glycerol were then added to 0.3 mg/ml and 5%, respectively. Binding reactions (10 μl) containing hybridized oligodeoxynucleotides (4 μl) and I-*Tev*I (*in vivo* preparation), with unlabeled competitor duplexes (2 pmol) where indicated (see Fig. 2), were incubated at 25°C for 30 min and loaded onto a 5% acrylamide gel in Tris borate buffer (acrylamide/bisacrylamide, 20:1) (18, 19). Electrophoresis was at 25 W for 15 min at 4°C. For PCR fragments containing the cleavage site (see Fig. 4*B*), I-*Tev*I–DNA complexes were made by incubating I-*Tev*I (1.5 μg of protein, *in vivo* preparation) with 5 × 10^4 cpm of end-labeled DNA fragment that had been excised from the vector. Incubation was in 25 mM Tris·HCl (pH 8.0)/20 mM EDTA, bovine serum albumin at 2.5 mg/ml, 5% glycerol, poly(dI-dC) at 140 μg/ml, 50 mM NaCl, and 5 mM $MgCl_2$ for 30 min at 25°C. Complexes were separated as above.

Endonuclease Cleavage Assays. *Sca* I-digested, end-labeled target DNAs (10^5 cpm) were incubated with up to 100 units (10 μl) of I-*Tev*I prepared *in vitro*. Incubation was in cleavage buffer (50 mM Tris·HCl, pH 8.0/10 mM $MgCl_2$/100 mM NaCl) at 37°C in a volume of 70 μl. Aliquots of 10 μl were removed at intervals, added to 10 μl of SLB solution [50 μM EDTA/5% (wt/vol) SDS/25% (vol/vol) glycerol/0.025% bromphenol blue] and electrophoresed on a 1% agarose gel, which was dried and quantitated with a Betagen β-detector.

RESULTS

I-*Tev*I Binding Sequences. DNase I-footprinting experiments performed on the *td* homing site indicate that at low enzyme concentration I-*Tev*I protected the region spanning the intron insertion site on both strands (Fig. 1 *A* and *B*, lanes 4 and 5). However, with increased enzyme concentrations the upstream boundary of the footprint extended toward the cleavage site on both strands (Fig. 1 *A* and *B*, lanes 6 and 7). These results suggest the possibility that sequences surrounding the insertion site represent the primary binding site for I-*Tev*I, as considered further by the following experiments.

A Synthetic Duplex Binds and Directs Cleavage by I-*Tev*I. To determine whether the natural cleavage site is dispensable to I-*Tev*I binding and cleavage, complementary oligodeoxynucleotides containing 31 bp (HS31) spanning the intron inser-

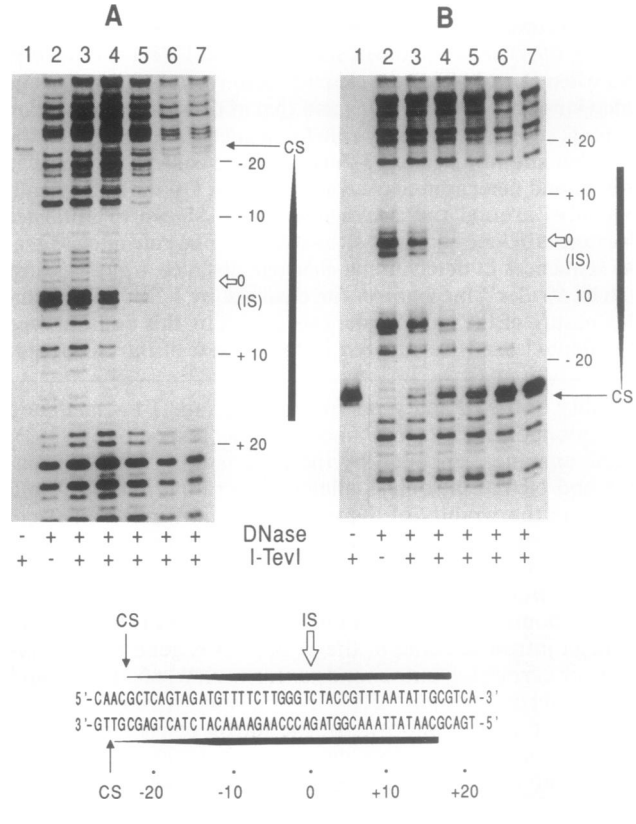


FIG. 1. DNase I footprinting of the I-*Tev*I binding site. A cloned PCR fragment encompassing the intron insertion site (IS) and I-*Tev*I cleavage site (CS) was labeled with ^{32}P on the top (*A*) or bottom (*B*) strands. The *Pst* I–*Xba* I fragment in *A* was labeled at the *Xba* I end, and the *Sal* I–*Kpn* I fragment in *B* was labeled at the *Sal* I end. Fragments were incubated with (+) or without (−) DNase I in the presence (+) or absence (−) of I-*Tev*I (0.12, 0.25, 0.5, 1.2, or 2.5 μg in lanes 3–7, respectively, and 1.2 μg in lanes 1). Identity of bands was inferred from Maxam–Gilbert sequencing ladders run in parallel (data not shown). The protected area is demarcated by solid bars in *A* and *B* and on the interpretive diagram. Bars taper in the region extending toward the CS, where protection occurred only at elevated I-*Tev*I concentrations. The relative degree of protection was assessed by using comparative scans from a densitometer and phosphorimager. Primary cleavages map to positions -23 and -25 on the top and bottom strands, respectively. The alternative cleavage at -26 on the bottom strand (ref. 3) was not seen in these experiments or those reported in Fig. 5.

tion site (15 bp 5′ and 16 bp 3′) but lacking the cleavage site were synthesized (Fig. 2). To assess binding by I-*Tev*I, end-labeled HS31 duplex was incubated with I-*Tev*I (Fig. 2, lanes 5–8) or with extracts lacking the endonuclease (lanes 1–4) and subjected to gel-retardation analysis. I-*Tev*I-specific complexes were observed in the absence of competitor DNA (lane 5) and in the presence of nonspecific competitor duplexes (lanes 7 and 8), but the complexes were displaced with an excess of unlabeled HS31 duplex (lane 6).

Once binding of I-*Tev*I to HS31 was established, the duplex was cloned into pBSM13$^-$ in both orientations to test whether it could direct cleavage into vector sequences. Both recombinants pBSHS31$^+$ and pBSHS31$^-$ (3.2 kb), which had been linearized with *Sca* I, were cleaved by I-*Tev*I (Figs. 3 and 4*A*, class c) into two smaller fragments (≈1.8 and 1.4–1.5 kb). Mapping of the cleavage sites on both strands indicated that cleavage occurred within vector sequences upstream of the intron insertion site on opposite sides of the cloning site: at 23–25 nt and 22–24 nt from the insertion site in pBSHS31$^+$ and pBSHS31$^-$, respectively (Fig. 3). Thus cleavage was at the normal distance for pBSHS31$^+$ and displaced by 1 nt on both strands for pBSHS31$^-$. These results indicate that, whereas sequences external to HS31 can influence cleavage distance, the insertion-site sequences contained within HS31 have a primary role in determining I-*Tev*I binding as well as cleavage ability and directionality.

Cleavage and Binding of Mutant Homing Sites. To determine the relative effects of sequence alterations surrounding the insertion and cleavage sites, cleavage was assayed with homing-site variants. The variants have been grouped into classes a, b, c, and d, where a represents the most efficiently cleaved homing sites, and d represents those sites that remain uncleaved (Fig. 4*A*). For both wild-type *td* homing sites, of T4*td*ΔIn and T2, 50% cleavage occurred in <1 min with 10 units of enzyme (class a). Despite naturally occurring sequence polymorphisms, the cleavage rates for these substrates were indistinguishable (Fig. 4*A* and data not shown).

Cleavage efficiencies of the cleavage-site variants were reduced to different extents (classes b and c). For pBS*td*F1, with a 2-bp deletion at the cleavage site, and pBS*td*F5, with

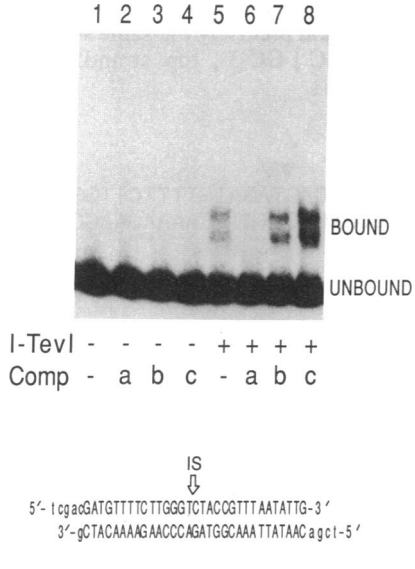


FIG. 2. Binding of I-*Tev*I to synthetic homing-site sequences. End-labeled *td* HS31 duplex (*Bottom*) was incubated with I-*Tev*I (lanes 5–8) or with control extract (lanes 1–4). Unlabeled competitor (Comp) DNAs are HS31 duplex (a); polylinker duplex (b); *nrd*B splice-junction duplex (c). For lanes 5–8, 0.3 μg of total protein (*in vivo* I-*Tev*I preparation) was added. The nature of the two bound species is not known. Lowercase letters at the ends of the duplex represent *Sal* I linkers.

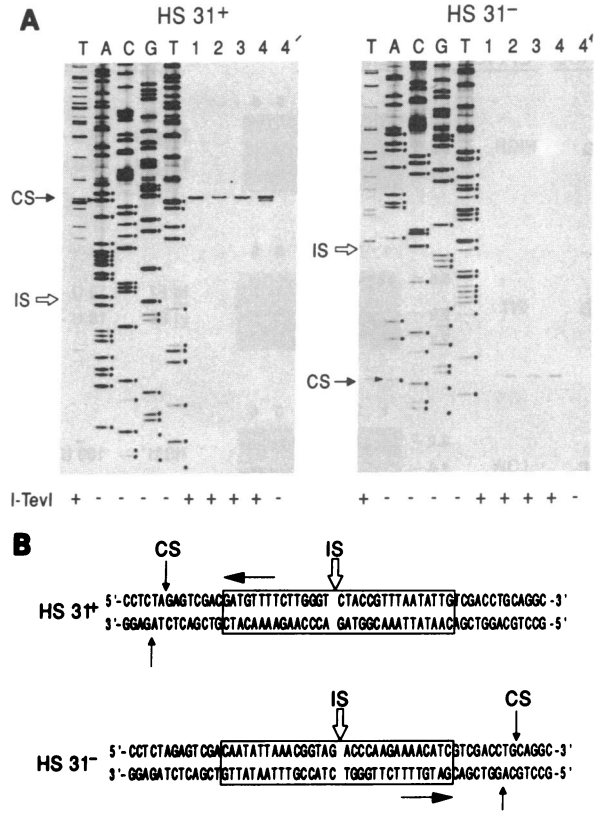


FIG. 3. Cleavage-site mapping of pBSHS31$^+$ and pBSHS31$^-$. (*A*) Mapping was as described (3, 10), with double-stranded plasmid DNA. DNA was primed downstream of the cloned duplex and extended with Sequenase (United States Biochemical) in the presence of dNTPs (including ^{35}S-labeled dATP) to determine cut sites on the bottom strands. Cleavage reactions were done on the resulting double-stranded DNAs for 2.5, 5, 10, and 20 min (lanes 1–4, respectively). Lanes 4′ show substrate incubated with unprogrammed extract. Sequencing ladders were generated with dideoxynucleotides. The precise cleavage site (CS) was determined by adding cleaved substrate (20-min incubation) to the first T lane. (*B*) Interpretive diagrams. The cut sites on the bottom and top strands, the latter determined with upstream primers, are indicated. Horizontal arrows indicate duplex orientation. Cleavage sites were identical with substrates labeled with [^{32}P]dATP.

a 1,646-bp deletion between the cleavage site and a 5′ site within pBSM13$^-$, 50% cleavage occurred at 1 min and 2 min, respectively, with 10 units of I-*Tev*I (class b). There was a much more dramatic reduction in cleavage efficiency for pBSHS31$^+$ and pBSHS31$^-$ containing hybrid homing sites, with 50% cleavage requiring 100 units of I-*Tev*I and 5 min or 20 min, respectively. This result reflects a >20-fold decrease in cleavage efficiency from the wild type (class c).

In contrast to the cleavage-site variants, deletion and insertion mutants at the intron insertion site were resistant to cleavage (class d). Both pBS*td*HSΔ21, with a 21-nt deletion (1), and pBS*td*HS+10, with a 10-nt insertion (3), remained uncleaved with 100 units of enzyme. Thus, alterations at the cleavage site reduced activity to various degrees, whereas interruption of the insertion site completely abolished cleavage.

To test binding capability, gel-retardation analyses were done as in Fig. 2 with 160- to 190-bp PCR derivatives of the homing-site variants (summarized in Fig. 4*B*). Although small differences in binding affinity would not have been detectable, the wild-type and cleavage-site variants (classes a–c) are clearly proficient at binding I-*Tev*I. In contrast, the insertion-site mutants (class d) are incapable of binding the

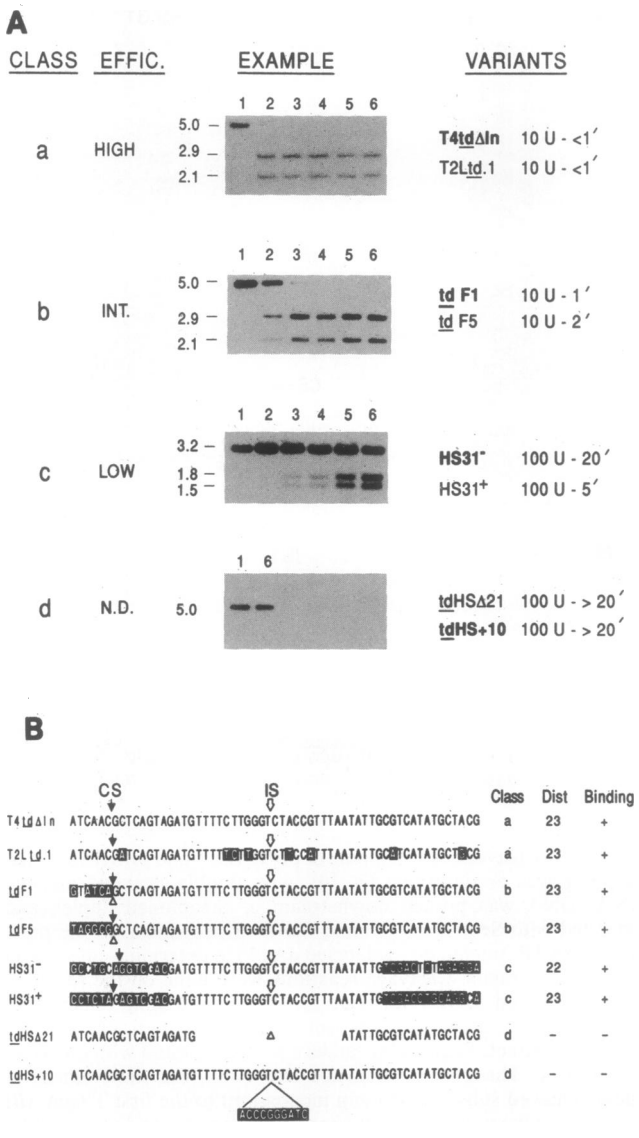


FIG. 4. Characteristics of substrate variants. (*A*) I-*Tev*I cleavage efficiency (EFFIC.) on wild-type and mutant substrates. Cleavage classes a–d, grouped according to cleavage efficiency (INT., intermediate; N.D., not detectable), are represented in the third column by those variants indicated in boldface type in the fourth column. All variants are derivatives of pBSM13⁻ (containing restriction fragments or HS31 duplex), except for T2L*td*.1, which is a pBSKS⁺ derivative (containing a PCR fragment). *Sca* I-linearized DNA was end-labeled and incubated with 10 units and 100 units of I-*Tev*I, and aliquots were removed at 0, 1, 2.5, 5, 10, and 20 min (lanes 1–6, respectively). Data shown for a and b were with 10 units of I-*Tev*I and for c and d with 100 units of I-*Tev*I; time required for 50% cleavage is indicated in the fourth column. (Variants in classes a and b were completely cleaved with 100 units of enzyme in <1 min, whereas mutants in class c were largely uncleaved at 20 min with 10 units of I-*Tev*I.) (*B*) Summary of substrate variants including binding by I-*Tev*I. White-on-black letters reflect deviations from the T4*td*ΔIn sequence; Δs represents sites of deletion. Cleavage sites (CS) are marked by black arrows. Distance of CS from the insertion site (IS) for the top strand appears at right, beside the cleavage class and ability to bind I-*Tev*I in gel-retardation experiments. These mobility-shift experiments were done as in Fig. 2 with 160- to 190-bp PCR fragments of the corresponding homing-site variants (+, efficient binding; −, no binding detectable).

enzyme, consistent with disruption of the critical recognition/binding domain.

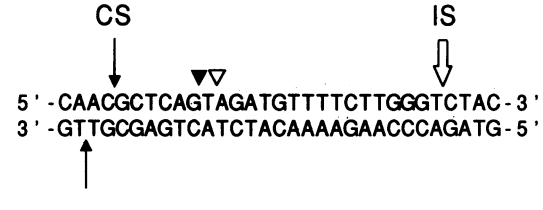
Distance Effects. To examine the role of cleavage-site sequences in cut-site selection, insertions and deletions were created between the cleavage and intron insertion sites. Insertions of 1 and 2 bp increased the cleavage distance on both strands by 1 and 2 nt, respectively, whereas 3- and 4-bp deletions reduced the distance accordingly (Fig. 5). In all cases cleavage occurred at the normal site (5′-CAAC↓GC-3′, top strand; 5′-GCGT↓TG-3′, bottom strand), indicating a possible role for these sequences in cut-site selection.

DISCUSSION

DNase I-protection studies and mobility-shift analyses indicated that the major determinants for I-*Tev*I binding reside in sequences surrounding the intron insertion site. Nucleotides at the cleavage site were only protected at increased enzyme concentrations (Fig. 1), suggesting that I-*Tev*I–DNA contacts at the cleavage site are secondary to or less stable than those at the insertion site. Consistent with a primary role for insertion-site sequences in I-*Tev*I binding, cleavage-site variants with an intact insertion site were binding-proficient (Fig. 4*B*, classes b and c), whereas the insertion-site mutants with an intact cleavage site were unable to bind I-*Tev*I (Fig. 4*B*, class d). These data are in accord with intron-plus variants being resistant to cleavage by the intron-encoded endonuclease (ref. 4; unpublished work).

The importance of insertion-site sequences in I-*Tev*I recognition was also reflected in the binding capability of the HS31 duplex, which lacks a cleavage site (Fig. 2), and in the ability of the cloned duplex to direct cleavage into vector sequences (Fig. 3). The HS31–plasmid fusions further revealed that insertion-site sequences dictate directionality of the cleavage. Although the natural cleavage site appears dispensable, the sequence context of the site can have a dramatic effect (>20-fold) on I-*Tev*I activity. Furthermore, displacement of the cleavage site by 1 nt on both strands in the pBSH31⁻ variant relative to the wild type (Figs. 3 and 4*B*) suggests that these 5′ sequences may also play a role in cleavage specificity.

Several lines of evidence indicate that cleavage-site sequences may play a role in determining the distance of the cut site from the insertion site. (*i*) For insertions of 1 and 2 bp and deletions of 3 and 4 bp between the intron insertion site and cleavage site, cleavage occurred at an altered distance from the insertion site but at precisely the same sequence as for the wild type (5′-CAAC↓GC-3′, top strand) (Fig. 5). (*ii*) In



Site	Change	Mutation	Distance Top	Distance Bottom
▽	+2 nt	+CT	25	27
▼	+1 nt	+C	24	26
	0 nt	None	23	25
▼	-3 nt	-TAG	20	22
▼	-4 nt	-TAGA	19	21

FIG. 5. Effect of insertions and deletions between intron insertion site (IS) and cleavage site (CS) on cleavage distance. Positions of changes are indicated by open and closed triangles above the sequence. The 2-nt CT insertion at the site indicated created an *Xba* I site (TCTAGA), used to generate the deletions. Cleavage-site distance is represented for each strand with the particular mutation at the indicated site (+, insertion; −, deletion). Cleavage sites were mapped as described in the legend to Fig. 3.

pACYCHS24, which contains a 24-bp fragment, including 12 bp to each side of the intron insertion site in the *Pvu* I site of pACYC184 (20), cleavage occurred 17 nt upstream of the insertion site (top strand), rather than the normal 23 nt, at a site identical to the normal cut site (5'-CAAC ↓ G-3') (21). (*iii*) Based on the observation that cleavage consistently appears to occur 5' to a guanine on the top strand (Fig. 4*B*), we changed this guanine to a thymine, while altering a thymine 2 nt downstream to a guanine. Both the cleavage efficiency and fidelity of this mutant were reduced; cutting occurred preferentially at the normal distance but also at spurious locations both upstream and downstream (S.M.Q., unpublished work). These results argue strongly against a model in which a purely distance-based "ruler" mechanism for cleavage is operative. Rather, they support a model in which cleavage-site sequences have a role in determining the location of I-*Tev*I cleavage.

Despite the demonstrated role of cleavage-site sequences in cut-site selection, the sequence polymorphisms tolerated around the cleavage site (Fig. 4*B*) suggest that I-*Tev*I has sequence preferences rather than strict sequence requirements at the cleavage site. Additionally, it is noteworthy that in mutants lacking the natural cleavage sequences cutting occurs, albeit inefficiently, at ≈23 nt from the insertion site (Fig. 4*B*; D.B.-P. and S.M.Q., unpublished work). This observation suggests that in the absence of preferred sequences I-*Tev*I may cleave with modest efficiency at a preferred distance.

Precisely which nucleotides are important for recognition and cleavage by I-*Tev*I remain to be determined. Likewise the relative contributions of DNA sequence versus helix conformation in I-*Tev*I recognition and cleavage are yet to be ascertained. However, because cleavage distances in the entire range from 17 to 25 nt have been observed (see above), I-*Tev*I must not require interaction with a specific face of the helix for cleavage. Furthermore, because both unmodified DNA and glucosylated hydroxymethylcytosine-containing DNA are cleaved efficiently and with high specificity (21), critical contacts in this unusual DNA–protein interaction must be independent of bulky side groups in the major groove.

Cleavage-at-a-distance and the observation that cleavage can occur in a range of sequences are features to be considered in the context of intron spread. According to the double-strand-break repair model, all sequences between a distant cleavage site and the intron must be coconverted during intron transfer, as has indeed been demonstrated for the *td* intron (2, 7). Provided coconverted exons do not disrupt gene function, they ensure acquisition of exon sequences needed to maintain exon–intron RNA pairings (P1 or P10, at the 5' and 3' junctions, respectively) required for accurate splicing of the relocated intron. In addition to coconversion, the eukaryotic introns may preserve exon–intron contacts by virtue of their endonucleases having long and specific recognition sequences (up to 18 bp) (4, 6), which encompass exon nucleotides involved in forming P1 and P10. Thus target-site specificity would help guarantee that P1 and P10 remain intact to guide faithful splicing of a translocated intron. In contrast, I-*Tev*I has a more relaxed target specificity, recognizing highly polymorphic homing sites and cleaving at secondary sites (Fig. 4*B*; 2, 3, 7, 12, and 21). This relaxed specificity combined with a range of sequence options for distant cleavage may effectively allow the intron to sample double-strand breaks in different sequence contexts. Those that allow intron–exon pairings through coconversion, while simultaneously maintaining gene function, would ultimately play host to the mobile intron.

We thank John Mueller for help with I-*Tev*I preparations and useful discussions, Joe Salvo for advice on DNA footprinting, Maryellen Carl and Carolyn Wieland for expert manuscript preparation, Doris Dixon and Maureen Belisle for technical assistance, and Tim Coetzee for comments on the manuscript. This work was funded by grants from the National Institutes of Health (GM44844 and GM39422) and the National Science Foundation (DMB8502961). S.M.Q. is supported by a National Institutes of Health Postdoctoral Fellowship (GM13582).

1. Quirk, S. M., Bell-Pedersen, D. & Belfort, M. (1989) *Cell* **56,** 455–465.
2. Bell-Pedersen, D., Quirk, S., Aubrey, M. & Belfort, M. (1989) *Gene* **82,** 119–126.
3. Bell-Pedersen, D., Quirk, S., Clyman, J. & Belfort, M. (1990) *Nucleic Acids Res.* **18,** 3763–3770.
4. Dujon, B. (1989) *Gene* **82,** 91–114.
5. Dujon, B., Belfort, M., Butow, R. A., Jacq, C., Lemieux, C., Perlman, P. S. & Vogt, V. M. (1989) *Gene* **82,** 115–118.
6. Perlman, P. S. & Butow, R. A. (1989) *Science* **246,** 1106–1109.
7. Belfort, M. (1990) *Annu. Rev. Genet.* **24,** 363–385.
8. Colleaux, L., D'Auriol, L., Galibert, F. & Dujon, B. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 6022–6026.
9. Delahodde, A., Goguel, V., Becam, A. M., Creusot, F., Perea, J., Banroques, J. & Jacq, C. (1989) *Cell* **56,** 431–441.
10. Wenzlau, J. M., Saldanha, R. J., Butow, R. A. & Perlman, P. S. (1989) *Cell* **56,** 421–430.
11. Muscarella, D. E., Ellison, E. L., Ruoff, B. M. & Vogt, V. M. (1990) *Mol. Cell. Biol.* **10,** 3386–3396.
12. Chu, F. K., Maley, G., Pedersen-Lane, J., Wang, A.-M. & Maley, F. (1990) *Proc. Natl. Acad. Sci. USA* **87,** 3574–3578.
13. Quirk, S. M., Bell-Pedersen, D., Tomaschewski, J., Ruger, W. & Belfort, M. (1989) *Nucleic Acids Res.* **17,** 301–315.
14. Bell-Pedersen, D., Salvo, J. L. G. & Belfort, M. (1991) *J. Bacteriol.* **173,** 1193–1200.
15. Studier, F. W., Rosenberg, A. H., Dunn, J. J. & Dubendorff, J. W. (1990) *Methods Enzymol.* **185,** 60–89.
16. West, D. K., Changchien, L.-M., Maley, G. F. & Maley, F. (1989) *J. Biol. Chem.* **264,** 10343–10346.
17. Ross, W. & Landy, A. (1979) *Cell* **18,** 297–307.
18. Revzin, A. (1989) *Biotechniques* **7,** 346–355.
19. Singh, H., Sen, R., Baltimore, D. & Sharp, P. A. (1986) *Nature (London)* **319,** 154–158.
20. Chang, A. C. Y. & Cohen, S. N. (1978) *J. Bacteriol.* **134,** 1141–1156.
21. Bell-Pedersen, D. (1991) Ph.D. thesis (State University of New York at Albany).