Circular forms of developmentally excised DNA in *Euplotes crassus* have a heteroduplex junction

Lawrence A. Klobutcher,¹ Leah R. Turner,² and Janice LaPlante

Department of Biochemistry, University of Connecticut Health Center, Farmington, Connecticut 06030 USA

Extensive DNA elimination via a DNA breakage and rejoining process occurs during macronuclear development in the hypotrich ciliate *Euplotes crassus*. The excision process involves the removal of short, unique segments of DNA (internal eliminated sequences; IESs) and at least two highly repetitive families of transposon-like elements (Tec elements). Previous studies have demonstrated that circular forms of both IESs and Tec elements are generated following their developmental excision and that some flanking DNA sequences are retained at the circle junctions. In this study we have further analyzed the circle junctions of IESs. Analysis of polymerase chain reaction (PCR) products derived from IES circle junctions indicates that at least two sequence arrangements can be present. The circle junctions contain both of the direct repeats that define the ends of the IES separated by either 2 bp flanking the right end of the IES and 8 bp from the left-flanking region, or 8 bp from the right and 2 bp from the left. Using a method that we have termed "strand-biased PCR," we obtained evidence that the junctions of free circular IESs have a 6-base heteroduplex at their center, such that one strand of the DNA is derived from the left-flanking region of the IES and the other from the right. Models of IES excision are presented that incorporate these results and those of previous studies on the excision process.

[**Key Words:** Ciliated protozoa; macronuclear development; DNA rearrangement; DNA excision; circular DNA; heteroduplex DNA]

Received September 9, 1992; revised version accepted October 29, 1992.

Site-specific recombination events that result in the elimination of DNA segments and the concomitant rejoining of flanking DNA occur in a variety of organisms. Such processes are often linked to the formation of functional genes during the course of development. For example, DNA excision during sporulation in *Bacillus subtilis* is required to form a functional sporulation-specific σ -factor (Margolis et al. 1991). Similarly, site-specific recombination events are required to produce functional antibody and T-cell receptor genes in the immune cells of vertebrates (for review, see Lieber 1991). DNA deletion events mediated by site-specific recombination are generally limited to one or a small number of genes, or groups of genes, in most organisms. However, an extremely large number of such events occur during the development of a macronucleus in the hypotrich ciliated protozoa such that virtually all genes are involved.

The binucleate nature of ciliate cells forms the basis for the extensive DNA processing that occurs in these organisms (for review, see Klobutcher and Jahn 1991; Prescott 1992). Each cell contains both a macronucleus

and a micronucleus. The micronucleus has a complex genome that is organized into conventional eukaryotic chromosomes, but it is transcriptionally inert during asexual growth of the organism. The macronucleus provides for the transcriptional needs of the cell during vegetative growth. The DNA of the macronucleus is in the form of linear, gene-sized DNA molecules with an average size of ~2 kb. Each macronuclear DNA molecule appears to encode a single genetic function. Both nuclei replicate their genomes and divide during the vegetative growth of the organism. However, following mating (conjugation), the old macronucleus is destroyed and a new one is generated from a mitotic copy of the micronucleus. The development of a new macronucleus is a multistep process that occurs over a period of ~100 hr. During macronuclear development, the micronuclear chromosomes are fragmented to yield the macronuclear DNA molecules, a large fraction of the micronuclear genome is eliminated, and DNA amplification occurs.

Some of the DNA sequence elimination that occurs during macronuclear development is mediated by site-specific recombination such that interstitial segments of DNA are excised and flanking sequences are rejoined. These types of DNA breakage and joining events have been observed in all ciliate species examined to date (for review, see Yao 1989; Klobutcher and Jahn 1991; Pres-

¹Corresponding author.

²Present address: Department of Microbiology, SC42, University of Washington, Seattle, Washington 98195 USA.

cott 1992). In *Euplotes crassus*, the subject of the current study, two classes of excised DNA sequences have been defined. The first class consists of internal eliminated sequences (IESs), which are unique segments of DNA that range from 31 to 374 bp in size in *E. crassus* (Baird et al. 1989; L.A. Klobutcher, unpubl.). Greater than 10,000 IESs exist in the micronuclear genome and are excised during development. The second class consists of transposon-like elements that have been termed Tec elements (Jahn et al. 1988, 1989; Baird et al. 1989; Krikau and Jahn 1991). Tec elements are ~5.3 kb in size and have inverted repeats of ~700 bp at their ends. Two types of Tec elements have been identified, Tec1 and Tec2, and each is present in ~30,000 copies in the micronuclear genome.

There are a number of features shared between IESs and Tec elements. Both IESs and Tec elements are bordered by a direct repeat of the dinucleotide 5'-TA-3',¹ which may be equivalent to a target site duplication (Baird et al. 1989; Jahn et al. 1989; Krikau and Jahn 1991; Tausta et al. 1991). Both types of elements undergo precise excision during macronuclear development so as to leave one copy of the 5'-TA-3' direct repeat in the resulting macronuclear DNA molecule (Tausta et al. 1991). Following excision, both elements assume a free circular form with an unusual circle junction (Jahn et al. 1989; Tausta and Klobutcher 1989; Krikau and Jahn 1991). The circle junctions contain two copies of the terminal direct repeat separated by 10 bp. For the IESs, it is clear that this 10 bp is derived from the sequences that flank the IES in the chromosome and also becomes part of the macronuclear DNA molecule. This has led to the suggestion that the excision process involves staggered cuts in the DNA that are filled in so that the two products of excision can share some of the same nucleotides (Tausta and Klobutcher 1989; Klobutcher and Jahn 1991). The single major difference between the two classes of excised DNA elements is their time of developmental excision. Both classes of elements are excised during the polytene chromosome phase of macronuclear development, but the Tec elements are excised ~14 hr before the IESs (Jahn et al. 1989; Tausta and Klobutcher 1990; Krikau and Jahn 1991). Based on the similarities between IESs and Tec elements, it has been suggested that they are both excised by the same enzymatic machinery or at least machineries that share some of the same components (Hunter et al. 1989; Jahn et al. 1989; Tausta and Klobutcher 1989).

In the current study we have further analyzed the junctions of free circular IESs to learn more about the mechanism of IES excision. Using a polymerase chain reaction [PCR (Saiki et al. 1985)] procedure to amplify the circle junctions of IESs, we have found that at least two types of circle junctions exist. The two classes of circle junction sequences are identical except for their central 6 bp,

which is derived from either sequences flanking the left or right end of the IES. We have also utilized a specialized form of PCR, which we have dubbed "strand-biased PCR", to provide evidence that the circle junctions are heteroduplex in nature and that the excision process involves staggered cuts in the DNA that generate 5'-overhanging ends. On the basis of these results, models for the IES excision process are presented.

Results and Discussion

Analysis of the 374-bp V1 IES circle junctions

In this study we focused on two IESs that exist within a region of the micronuclear genome represented by clone LEMICV (Fig. 1A). The 374-bp V1 IES resides in the precursor of a 3.85-kbp macronuclear DNA molecule of unknown function designated V1. The 144-bp V3 IES exists in the precursor of the 1.8-kbp V3 macronuclear DNA molecule. Macronuclear V3 appears to encode a protein with homology to protein kinases (L.A. Klobutcher, M. Jacobs, and C. Hale, unpubl.).

In a previous study (Tausta and Klobutcher 1989) we had used PCR to specifically amplify the circular forms of these two IESs that are generated during macronuclear development. The PCR strategy employed pairs of oligonucleotide primers complementary to the IES but directing DNA synthesis divergently, so that only a circular form of an IES is exponentially amplified. Direct sequencing of the PCR products produced by this procedure revealed an unexpected circle junction. Circle junctions were found to possess two copies of the IES terminal direct repeat 5'-TA-3', separated by 10 bp. The 10 bp corresponds to sequences that flank the IES in the micronucleus and is also retained in the macronuclear DNA molecules from which the IESs are excised. It has been proposed that the 10 bp shared by the two products of excision (the free IES and the macronuclear DNA molecule) arises as the result of a staggered cut in the DNA during IES excision, with subsequent fill-in of the single-stranded ends before joining (Tausta and Klobutcher 1989; Klobutcher and Jahn 1991). We initially interpreted the IES circle junction sequences as indicating that 8 bp of the central 10 bp was derived from sequences flanking one side of the IES, while the remaining 2 bp was derived from the opposite end of the IES (Tausta and Klobutcher 1989). However, the apparent quality of the sequencing in the circle junction region was poor; that is, secondary bands were observed in the sequencing ladder (Fig. 2A). The secondary bands were initially believed to be sequencing artifacts, perhaps owing to secondary structure within the region. We have found, however, that the secondary bands are highly reproducible. The positions in the sequencing ladders representing the central 6 bases of the junction appear to display both a major

¹5'-TA-3' can be viewed as either a direct or inverted repeat. We choose to view it as a direct repeat because IESs in other species of hypotrich ciliates have direct repeats that are nonpalindromic. In addition, Tec elements and some IESs in *E. crassus* can be viewed as having direct repeats of up to 4 bp. In this report we have chosen to view the direct repeat as 2 bp in length because it allows the data on different elements to be generalized.

Figure 1. The 374-bp V1 and 144-bp V3 IESs. (A) Restriction maps of a portion of micronuclear recombinant clone LEMICV and all or part of the three macronuclear DNA molecules derived from this region (V1–V3). Micronuclear map: (Open boxes) Regions that will give rise to macronuclear DNA molecules; (hatched box) eliminated spacer DNA; (solid boxes) IESs, with their sizes in bp below. (H) *HindIII*; (E) *EcoRI*; (Ss) *SstI*; (X) *XbaI*; (B) *BglII*; (Pv) *PvuII*; (Xm) *XmnI*. (B) Maps of 374-bp V1 and 144-bp V3 IESs. The terminal 5′-TA-3′ direct repeats are indicated; arrows denote oligonucleotides used as primers for PCR and sequencing. Oligonucleotides VICG–Eco and V1CC–Bam (not shown) are homologous to the same regions of the 374-bp V1 IES as VICG and V1CC, respectively. (H) *HindIII*; (Ha) *HaeIII*; (X) *XbaI*. All references to the right and left ends of the IESs in the text are based on the orientations shown.

and minor band at each position. For the 374-bp V1 IES, the sequence of the major bands corresponds to a circle junction with 8 bp derived from rightward-flanking sequences and 2 bp derived from the left side of the IES (Fig. 2A,C). The minor sequence corresponds to a circle junction with 2 bp derived from the right end of the IES and 8 bp from the left end (Fig. 2A,C). To determine the relative amounts of the two sequences, densitometric scanning was used to determine the relative intensities of the two bands at the first position in the junction region that displayed a dual sequence. [This corresponds to an A in the major sequence and a G in the minor sequence (Fig. 2A). Other positions in the circle could not be sufficiently resolved to provide reliable data.] The results indicate that the minor sequence corresponds to ~23% of the total.

This result suggests that two distinct classes of DNA molecules differing in the central 6 bp of the circle junction are generated during PCR of the circular forms of the 374-bp V1 IES. To directly confirm this, PCR products derived from the V1 IES were cloned into M13 or pUC vectors and the sequences of the circle junctions of nine independent clones were determined. The sequences of two such clones are shown in Figure 2B. Clone pJL2 was found to contain a circle junction composed of a 5′-TA-3′ direct repeat followed by 8 bp of right-flanking sequence, 2 bp of left-flanking sequence,

and a second 5′-TA-3′ direct repeat (Fig. 2B). This corresponds to the major sequence seen in the direct sequencing of the bulk PCR products. In all, seven of the nine clones displayed this sequence at the circle junction (Fig. 2C). The second clone shown, pJL6, has a circle junction corresponding to the minor sequence; that is, the 10-bp core of the circle junction consists of 2 bp from the right flank of the IES followed by 8 bp from the left flank (Fig. 2B). Two of the nine clones contained the minor sequence at the circle junction (Fig. 2C).

The two different classes of circle junction sequences can be viewed as reciprocal products in regard to retention of leftward- and rightward-flanking sequences. For simplicity, we shall refer to these types of arrangements as 8 + 2 circle junctions. The circle junctions may also be viewed as being identical, with the exception of the central 6 bp, which can be derived from either left- or right-flanking sequences.

Analysis of 144-bp V3 IES circle junctions

To determine whether the 8 + 2 sequence arrangement at the circle junction was common to other IESs, we performed a similar analysis on the circular forms of the 144-bp V3 IES. In this case, 21 independent clones were generated from PCR products of the circular forms of the V3 IES and sequenced (Fig. 3). Of the 21 clones, 11 displayed 8 + 2 circle junctions analogous to those observed for the 374-bp V1 IES. Of these, seven had circle junctions with 2 bp derived from right-flanking sequences and 8 bp from left-flanking sequences (2R + 8L junctions) while four displayed the reciprocal arrangement (8R + 2L junctions).

The remaining 10 clones possessed 18 bp between the direct repeats of the circle junction. Eight of the clones contained 8 bp of right-flanking sequence followed by 10 bp of left-flanking sequence, whereas the remaining two clones had 2 bp of right-flanking sequence plus 16 bp of left-flanking sequence. It is noteworthy that each of these larger circle junctions possesses a “secondary” 5′-TA-3′ direct repeat (Fig. 3) between the direct repeats that we normally view as defining the ends of the IES (“primary” direct repeats). Ten base pairs separate this secondary direct repeat from the primary direct repeat derived from the right end of the IES. Moreover, relative to this pair of direct repeats, the intervening sequences fall into either the 8R + 2L or 2R + 8L classes defined above. This suggests that the secondary repeat is being used alternatively at some step of the excision process, such as the initial strand scission event of IES removal or the subsequent processing of the IES ends that results in circle formation. The implications of these larger circle junctions for models of the IES excision process are discussed below.

Overall, the above results indicate that the 8 + 2 sequence arrangement is generally present at the junctions of free circular IESs. It might be argued that these results are an artifact of the PCR procedure employed. Artifacts are indeed possible, particularly because both macronuclear and micronuclear DNA forms are present as con-

Figure 2. Sequence analyses of the circle junctions of the 374-bp V1 IES. (A) Sequence of the circle junction regions of the total PCR products derived from the 374-bp V1 IES. PCR was carried out using oligonucleotides V1CC and V1CG, and the PCR products were sequenced using V1CG as the primer. The major sequence of the circle junction region is shown at *left* in uppercase letters with the direct repeats boxed. The central 6 bases of the circle junction region displays light, or minor, bands. The sequence of the minor bands is indicated in lowercase letters. (B) Junction sequences of two cloned PCR products (clones pJL2 and pJL6) derived from the free circular 374-bp V1 IES. The sequence of the junction region is shown beside the gels, with the direct repeats boxed and lines denoting the 6 bp within the junctions that differ between the two clones. Clone pJL2 contains a circle junction with the major sequence seen above (A); clone pJL6 displays the minor sequence. (C) Sequences at the ends of the 374-bp V1 IES are shown (*top*). The sequence of the IES, including both 5'-TA-3' direct repeats (boxed), is shown in lowercase letters; flanking, macronuclear-destined sequences are shown in uppercase. The rightward-flanking sequence is highlighted by a stippled background. The nucleotides that are joined to form the two classes of circle junctions are bracketed, with the arrows denoting the direction in which the sequences are displayed on the above sequencing gels. The 6 bases at either end of the IES that differ in the two classes of circle junctions are underlined. (*Bottom*) Circle junction sequences of the nine clones analyzed. The central 10 bp of the circle junction can be seen as consisting of 8 bp from the right of the IES plus 2 bp from the left (8R + 2L) or 2 bp from the right and 8 bp from the left (2R + 8L).

taminants in the substrate DNAs derived from developing macronuclei, and this may allow for the formation of "hybrid" PCR products (Arnheim and Erlich 1992). There are, however, two arguments against the results being artifactual. First, we have shown previously that PCR products of circle junctions are produced only from developing macronuclear DNA after the time of IES excision, and not from vegetative cellular DNAs (Tausta and Klobutcher 1989). Thus, the types of circle junctions that we have observed are at least dependent on some DNA form generated at the time of IES excision. Second, two different IESs were found to produce 8 + 2 circle

junctions. The two IESs are generally quite different in sequence at their extreme ends, as is the flanking DNA. It is thus difficult to explain the repeated observation of the same type of circle junctions based on hybrid PCR products. In the case of the larger circle junctions observed for the 144-bp V3 IES, it is more difficult to exclude the possibility of a PCR artifact, as they were observed for only one of the IESs analyzed.

Strand-biased PCR of IES circles

The above results indicate that at least two different

Figure 3. Sequence analysis of the 144-bp V3 IES circle junction clones. (Top) Sequences at the ends of the 144-bp V3 IES. The sequence of the IES, including both 5'-TA-3' direct repeats (boxed), is shown in lowercase letters; flanking, macronuclear-destined sequences are shown in uppercase. The rightward-flanking sequence is highlighted by a stippled background. The nucleotides that are joined to form circle junctions are bracketed, with the associated numbers denoting the four classes of circle junction sequences displayed below. The arrows on the brackets denote the direction of the sequences displayed below. The underlined sequences flanking the ends of the IES represent the bases that differ in the circle junctions of class 1 and class 2 clones. (Bottom) Circle junction sequences of 21 clones of PCR products produced from the V3 IES using oligonucleotides V3CA and V3CT. For the clones with >10 bp between the direct repeats of the circle junction, the "secondary" 5'-TA-3' direct repeat (see text) is doubly underlined. The 8 + 2 arrangement of bases relative to direct repeats is also indicated.

DNA sequences can reside at the junctions of IES circles: either the 8R + 2L or 2R + 8L arrangements. It is important to note that the nature of the circle junctions was investigated indirectly through the use of PCR. As a result, there are at least two situations that could explain the results. First, two populations of free circular double-stranded IESs could be generated following excision that differ only in the central 6 bp of the circle junction. PCR would have amplified both populations of circular IESs to give rise to the two classes of clones that we observed. Alternatively, there may be only one type of circular IES generated, but the circular form contains a 6-base heteroduplex at the center of its junction, with the 6 bases on one strand being derived from right-flanking sequences and the 6 bases of the other strand from left-flanking sequences. Because of the replicative nature of the process, PCR would then have copied each strand of the free circles to generate the two classes of clones observed. There is a precedent for the heteroduplex hypothesis, as work on the bacterial conjugative transposon Tn916 indicates that it forms a circle with a heteroduplex junction following its excision (Caparon and Scott 1989). If this is the case for IESs, one would ideally expect to observe two circle junction sequences of equal intensity in the bulk sequencing of PCR products derived from the V1 IES, which was not the case (Fig. 2A).

However, it is quite possible that the two primers used in the PCR procedure did not work with equal efficiency, so that an overabundance of PCR products was generated from one of the two strands.

To discriminate between the two models, we developed a procedure that we refer to as strand-biased PCR (Fig. 4A). The method involves two successive rounds of amplification. The first round of amplification is carried out under PCR conditions, but using only one oligonucleotide as a primer. This linear amplification step is expected to generate an overabundance of the complement of one strand of the target sequence. In the case of a circular IES with a heteroduplex circle junction, this first round of amplification would create an overabundance of one of the two sequences that can be present at the circle junction. The second round of amplification is conventional PCR using two oligonucleotide primers chosen to amplify the region of interest. This second exponential round of amplification should maintain any bias generated during the first round of amplification. Direct sequencing of the PCR products is then used to reveal whether any bias has been generated. Bias toward one of the two circle junction sequences will occur if a heteroduplex circle junction is present. (Note that biasing would also result if there were two populations of IES circles that differed in their junction sequences and one population had a gap or nick on one strand in the junction region while the other population had a gap or nick on the opposite strand.) If two populations of duplex IES circles are present, no strand bias would be observed. In addition, if a heteroduplex is present at the IES circle junction, the strand-biased PCR procedure will provide information on the nature of the excision process. As we discuss in more detail below, both a heteroduplex circle junction and bases shared between the two products of IES excision can be explained most easily if IES excision is initiated by staggered cuts in the DNA. The direction of bias (i.e., which of the two sequences becomes more prominent) when one of the primer oligonucleotides is used in the first round of amplification can thus be used to infer whether an intermediate with 3'- or 5'-overhanging ends is generated during the excision process.

Strand-biased PCR was used to analyze the free circular forms of the 374-bp V1 IES (Fig. 4B). When V1CG was used for 20 cycles of round I amplification, the products generated after round II PCR using primers V1CG and V1CC displayed almost exclusively the major circle junction sequence (8R + 2L; Fig. 4B). Densitometric scanning of the gel indicated that the minor sequence was now reduced to <4% of the total circle junction sequences. When V1CC was used for 15 cycles of round I amplification, a bias in the sequence was again observed, but in this instance the bias was less complete (Fig. 4B). The minor junction sequence (2R + 8L) increased in intensity, so that it was present in an amount similar to that of the major circle junction sequence (~41% of the total). Similar results were obtained using oligonucleotide V13 (Fig. 1B) in round I amplification (data not shown). In each case, the direction of bias is consistent with 5'-overhanging ends being produced at

Figure 4. Strand-biased PCR. (A) The strategy for strand-biased PCR. (Top) A DNA molecule with a heteroduplex region (BBBB/XXXX). In the first round of amplification, only one primer (stippled arrowhead) is used so that only one strand of the DNA substrate is replicated. In the second round of amplification using both primers (stippled and solid arrowheads), both of the DNA strands are amplified but on the basis of their relative concentrations in the reaction. The result is a population of PCR-amplified DNA molecules containing predominantly the sequence and the complement of one of the original two strands of the heteroduplex. Lowercase b and x indicate bases complementary to B and X, respectively. (B) Sequences of the strand-biased PCR products generated from the 374-bp V1 IES. Oligonucleotides V1CG and V1CC, V1CG alone, and V1CC were used in round I of strand-biased PCR. Each

of these reactions was then amplified further in round II PCR with both oligonucleotides, and the PCR products were sequenced directly using oligonucleotide V1CG as the sequencing primer. Bars, central 6 bases of the circle junctions.

some point during the excision process (this is visualized most easily by referring to the model of excision shown in Fig. 5A, below).

There are at least two possible explanations for our observation of only a partial bias when the V1CC oligonucleotide was used in round I amplification. We believe the most likely explanation for this result is that the strand-biased PCR procedure does not work as ideally as diagramed in Figure 4A. It is very unlikely that every circular IES substrate molecule is copied in each cycle of round I amplification. Thus, with 15 cycles of round I synthesis, one is not achieving the maximum 15 : 1 ratio of newly synthesized DNA strands to substrate strands. We have tried to enhance the bias by carrying out additional cycles of round I amplification but have found that this greatly enhanced the amount and number of non-specific PCR products following round II PCR while at the same time reducing the amount of the specific PCR product. A second explanation for our result is that heteroduplex, circular IESs are formed following excision, but some are subjected to DNA repair within the cell such that one strand of the heteroduplex is preferentially

repaired. In this case, only a subpopulation of the circular IESs would have heteroduplexes that could contribute to a bias.

In any event, both the ability to obtain a bias toward one of the sequences in the above experiment and the fact that the bias is in different directions for the two oligonucleotide primers argue strongly for the existence of a heteroduplex at some IES circle junctions. The data are also consistent with two populations of IES circles that both differ in their junction sequences and contain nicks or gaps on opposing DNA strands at the junction, but we do not favor this interpretation because of data reported in the accompanying paper by Jaraczewski and Jahn (this issue). These workers have taken advantage of the high copy number of the Tec transposon-like elements, and the resultant large number of free circular forms generated during development, to perform direct enzymatic analyses on circle junctions. They have found that the Tec circles are sensitive to a number of single-strand nucleases, and that the point of nuclease sensitivity maps to the circle junction region. We have independently confirmed that the Tec1 circle junctions are sen-

sitive to S1 nuclease (L.A. Klobutcher, unpubl.). In addition, they provide evidence that both strands of the Tec circles are covalently closed. As noted in the introductory section, there are a number of similarities between IESs and Tec elements in *E. crassus* and they are probably both excised by the same, or similar, mechanisms. The results of these investigators thus both provide further evidence for an unusual structure at the circle junctions of developmentally excised DNA in *E. crassus*, and are most consistent with the hypothesis of a heteroduplex circle junction. The repetitive nature of the Tec elements has made it impossible to determine the derivation of the 10 bp between the direct repeats at their circle junctions (Jahn et al. 1989; Jaraczewski and Jahn, this issue); however, on the basis of our results, we suggest that they are also derived from flanking sequences in the same manner as the IESs.

Models of IES excision

We have attempted to develop a model of IES excision on the basis of the data currently available. Any model of excision must account for the current and previously observed features of the process, which include the following: (1) IES elements undergo precise excision such that one copy of the 5'-TA-3' direct repeat is retained in

the resulting macronuclear DNA molecule (Tausta et al. 1991); (2) the excised element contains a circle junction with two copies of the direct repeat separated by 10 bp of flanking sequence; and (3) the central 6 bases of the circle junction can be derived from either left- or right-flanking DNA. A model that conforms to these data and postulates a circular IES with a heteroduplex region is shown in Figure 5A. IES excision begins with a staggered cut at the ends of the excised element. One strand is cut at the inside edge of the 5'-TA-3' direct repeat, whereas the other strand is cut in sequences flanking the element, to generate ends with 10-bp 5' overhangs. As noted above, 5' overhangs are indicated by the strand-biased PCR data. The ends that are destined to form the macronuclear DNA molecule associate via pairing of the 5'-TA-3' residues at their termini. The gaps in this structure are then filled in, and ligation generates a macronuclear DNA molecule that retains one direct repeat. The DNA ends that generate the free circular forms are processed somewhat differently. Initially, the two element ends are aligned such that the terminal 6 bases of the 5' overhangs overlap. The gaps in this structure are then filled in by a DNA polymerase activity, and ligation produces the free circular form. The resulting junction of the free circle contains two 5'-TA-3' direct repeats separated by 10 bp of flanking sequence and a central 6 base

Figure 5. Models of IES excision. (A) (Top) The sequences surrounding the 374-bp V1 IES are shown, with the IES direct repeats boxed and the nucleotides involved in forming the putative heteroduplex region indicated by lines. The sequences flanking the right end of the IES are highlighted with a stippled background. The positions of the cuts that generate staggered ends are indicated by arrowheads. (Below) The two ends destined to form the macronuclear DNA molecule associate via their terminal 5'-TA-3' residues and are then filled in and ligated, resulting in a macronuclear DNA molecule with one copy of the direct repeat. The ends forming the circular IES associate with a 6-base overlap and the gaps are filled in and ligated. The result is a free circular IES with two copies of the direct repeat and a 6-base heteroduplex region. (B) A modified model to accommodate the larger circle junctions observed for the 144-bp V3 IES. The sequences flanking the V3 IES are shown, with conventions as in A. (Below) Two alternative ways of processing the ends to form a circular IES are shown based on overlapping the single-stranded tails relative to two different pairs of 5'-TA-3' direct repeats. The secondary 5'-TA-3' direct repeat is doubly underlined. The ends destined to form the macronuclear DNA molecule are not shown.

heteroduplex. Although we propose this model specifically for IES excision, we believe it applies to Tec element excision as well, because the Tec excision products can be viewed as identical to those produced by IES excision.

This model is attractive not only because it explains the current data but also because it generally does not require any unknown enzymatic activities. The initial DNA cutting event to generate 5' overhangs is a common feature of many restriction endonucleases, and it is also involved in the transposition of some transposable elements (Mizuuchi 1984; Benjamin and Kleckner 1989; Bainton et al. 1991). Subsequent DNA synthesis steps all occur in the conventional 5' → 3' direction. The one unusual feature of the model involves the initial association of the element ends. Association of the overhanging ends cannot be based on conventional base-pairing, as the two ends differ in sequence. We postulate that a component of the excision machinery is responsible for bringing the two ends together in the suggested fashion so that the next steps of circle formation can proceed.

The above model does not account for the larger circle junctions that were sometimes observed with the 144-bp V3 IES (Fig. 3). In these cases, one can find a secondary direct repeat within the circle junction such that the intervening sequences conform to the 8 + 2 arrangement of nucleotides (see above). The function of the 5'-TA-3' direct repeat is unknown, but its conservation at all IES and Tec element ends argues that it plays some essential role. The sequence itself is too short to serve as the sole recognition sequence for the excision machinery, but it could serve to orient the assembled machinery in regard to either the strand scission event or the subsequent steps of circle formation. If the 5'-TA-3' direct repeat is responsible for defining the point of strand scission, the larger 144-bp V3 IES circle junctions could be explained by minor alterations of the above model. For example, the excision machinery might, in some cases, cleave the DNA relative to this secondary 5'-TA-3' repeat in flanking DNA. If the ends were processed as described in the above model, this would result in an IES with a larger circle junction (relative to the conventional direct repeats) and a correspondingly smaller macronuclear DNA molecule. We do not, however, favor this model because the predicted smaller V3 macronuclear DNA molecules are not observed (Tausta et al. 1991). It remains possible that the macronuclear DNA ends formed by the alternative DNA cleavage process fail to rejoin and, hence, are not retained in the mature macronucleus. An alternative variation in the model that would accommodate the larger V3 IES circle junctions, as well as the standard 8 + 2 circle junctions, involves a fragmentation event that leaves longer single-stranded ends (Fig. 5B). The single-stranded ends would need to be at least 18 bases in length to include all the extra bases present in the larger V3 IES circle junctions. The ends destined to form the macronuclear DNA molecule would then be processed essentially as described previously (see Fig. 5A), but the ends that are joined to form the circular IES could be aligned relative to either the primary or other secondary repeats such that 6 bases

overlap. This alignment would again result from interaction of the excision machinery with the DNA ends. Following trimming of the 5' ends, fill-in and ligation could proceed to generate the different classes of circle junctions.

The proposed models share some similarities with those proposed recently for the excision of the bacterial conjugative transposons Tn916 and Tn1545 (Caparon and Scott 1989; Poyart-Salmeron et al. 1990), which also appear to excise to generate circular forms with heteroduplex circle junctions of 5 or 6 bases. These models differ, however, in that the heteroduplex regions are thought to be formed by first excising the integrated element such that 5- or 6-base overhanging ends composed of flanking DNA sequences are generated. The single-stranded ends are then viewed as aligning with complete overlap so that ligation may directly generate the closed circular forms. Moreover, the DNA ends at the "empty site" are treated in the same manner. Our models differ both in that they must incorporate longer single-stranded tails that do not completely overlap during circle formation to account for the types of junctions observed and in that the ends destined to form the macronuclear DNA molecule are processed in a different manner.

The types of circle junctions that we have observed, and the models for their formation, also differ from other transposon and site-specific recombination systems, where free, covalently closed circular forms have been observed. For example, circular forms have been observed for the Tc1 transposon of *Caenorhabditis elegans* (Ruan and Emmons 1984), the σ -factor rearrangement system in *B. subtilis* (Kunkel et al. 1990), and the nitrogen fixation gene rearrangement system of *Anabaena* (for review, see Haselkorn 1989). In each of these cases, the circle junction consists of one of the direct repeats that define the boundaries of the integrated element. These types of circles could simply be formed by a reciprocal recombination event between the direct repeats at the ends of the integrated element. Alternatively, excision may be initiated by staggered cuts at each end of each direct repeat, and the circles formed by association and ligation of the resulting complementary single-stranded tails. Circles with a somewhat different structure are formed during antibody gene rearrangement in vertebrates (for review, see Lieber 1991). In this case, the vast majority of circle junctions consist of a head-to-head linkage of the recombination signal sequences that border the excised DNA segment. Current models postulate a double-stranded cut at the outside ends of the recombination signal sequence followed by direct ligation of the ends to form these circles. Although this is again different from the type of excision process that we propose, the antibody gene rearrangement system does provide one example for the differential processing of the pairs of ends that our models entail. While the ends destined to form the free circles appear to be directly ligated, the ends that are joined to form the immunoglobulin gene-coding sequence are often modified by the addition and/or deletion of nucleotides before ligation.

The circular junctions of *E. crassus* IESs and Tec elements also differ from those of developmentally excised DNA segments in other ciliates. In *Oxytricha trifallax*, free circular forms of the TBE1-ot transposon-like element have recently been observed during macronuclear development (G. Herrick, pers. comm.). The TBE1-ot circle junctions appear to consist of only one of the 3-bp direct repeats that bound the integrated element. Similarly, in *Tetrahymena thermophila*, a developmentally excised 0.6-kb DNA segment within the M region (Austerberry et al. 1989) forms a circle with a single copy of its terminal 8-bp direct repeat at the junction (M.-C. Yao, pers. comm.). These results suggest that multiple excision systems exist in the ciliates. It is not yet clear whether these are species-specific differences or whether multiple excision systems can operate in the same organism. There have, however, been indirect indications that distinct classes of IESs exist in hypotrichs based on either different conserved sequences near their termini (Ribas-Aparicio et al. 1987) or their times of developmental excision (Tausta et al. 1991). The recent observation that some internal segments of DNA undergo elimination, in concert with the reordering of flanking DNA sequences in *Oxytricha nova* (for review, see Prescott 1992), can also be viewed as an additional indication of multiple excision systems operating in the same organism.

In summary, we have provided evidence that the excised circular forms of IESs in *E. crassus* have a circle junction that is partially heteroduplex in nature, and we provide models of the excision process. Some aspects of the models are similar to other site-specific recombination and transposon systems; but, in toto, the excision system in *E. crassus* appears to be unique. Although the models explain all current data on the excision process, at this point they are speculative and require confirmation. The detection and characterization of intermediates in the process, as well as the development of an in vitro rearrangement system, will be required to test the models.

Materials and methods

Cell lines, cell culture, and isolation of developing macronuclear DNA

Cell lines used for matings were from the CC series (Baird and Klobutcher 1988), the ST series (Tausta et al. 1991), and cell line 14 (Jahn et al. 1989), which was kindly provided by Dr. C. Jahn (Northwestern University Medical School). Cells were grown as described by Roth et al. (1985), except that vitamin B12 was omitted and, in some cases, Reef Crystals (Aquarium Systems, Mentor, OH) were substituted for Instant Ocean (Aquarium Systems). Cells were mated, and developing macronuclear DNA was isolated as described by Roth et al. (1985), with minor modifications (Tausta and Klobutcher 1990; Tausta et al. 1991).

Molecular biological techniques

Restriction enzymes and DNA-modifying enzymes were purchased from either Life Technologies, Inc. (Gaithersburg, MD), or New England Biolabs (Beverly, MA) and used under condi-

tions described by the manufacturer. DNAs were sized by electrophoresis through agarose or low-melting-point (LMP) agarose (Life Technologies, Inc.) gels prepared and run in TBE buffer (89 mM Tris base, 89 mM boric acid, 2 mM EDTA at pH 8.0). Plasmid DNA minipreparations used for DNA sequencing were prepared using the Magic Miniprep DNA purification system (Promega, Madison, WI). Other standard molecular biological methods were performed as described by Maniatis et al. (1982).

PCR

The following oligonucleotides for PCR (Sakei et al. 1985) were synthesized on a Cyclone dual column DNA synthesizer (Biosearch, Inc.): V1CC, 5'-AATACTCAAAGTTATGTACC-3'; V1CC-Bam, 5'-TAGGATCCAATACTCAAAGTTATGTACC-3'; V1CG, 5'-GAACAAATAATGGGTAAAAG-3'; V1CG-Eco, 5'-GCCAATTCGAACAAATAATGGGTAAAAG-3'; V3CA, 5'-TTTATGACTCAAAGTCTCTA-3'; V3CT, 5'-AATTAGAA-TATCTGAGGCTT-3'.

PCR was carried out using 0.2–1.0 µg of template DNA and reagents from a GeneAmp PCR kit (Perkin-Elmer Cetus). In some cases, a reaction buffer consisting of 50 mM KCl, 10 mM Tris-HCl (pH 8.8), 3.5 mM MgCl₂, and 0.1% Triton X-100 was substituted. Additional details of PCR reactions for various experiments are given below.

Cloning of 374-bp V1 IES PCR products

To clone circle junctions of the 374-bp V1 IES, PCR was carried out using oligonucleotides V1CG-Eco and V1CC-Bam as the primers and 50-hr developing macronuclear DNA isolated from a mating of cell lines CC102 and ST2 as the substrate. Thirty-five cycles of PCR were carried out on a model 50 Tempcycler (Coy Laboratory Products, Inc.), with a cycle consisting of 94°C for 1 min, 46°C for 1 min, and 72°C for 50 sec. The PCR reaction was then run on a LMP agarose gel, and the circle junction PCR products were purified (Kuhn et al. 1979). The purified PCR products were doubly digested with the restriction enzymes *Hae*III and *Hind*III, cloned into *Sma*I- plus *Hind*III-cut M13mp10 or pUC12 vectors (Vieira and Messing 1982), and transformed into *Escherichia coli* JM103.

Cloning of 144-bp V3 IES circle junction PCR products

PCR was carried out using oligonucleotides V3CA and V3CT as the primers and 46-hr developing macronuclear DNA isolated from a mating of cell line CC100 and cell line 14 as the substrate. Thirty cycles of PCR were carried out, with a cycle consisting of 94°C for 1 min, 43.5°C for 1 min, and 72°C for 50 sec. Following PCR, the products were made blunt-ended by adjusting the reaction mixture to 5 mM MgCl₂, adding 2 units of Klenow polymerase, and incubating at room temperature for 20 min. The PCR products were then subjected to LMP agarose gel electrophoresis and purified (Kuhn et al. 1979). The 5' ends of the purified PCR products were phosphorylated using polynucleotide kinase (Maniatis et al. 1982), and this material was ligated into *Sma*I-digested pUC12 vector.

Strand-biased PCR

Strand-biased PCR was carried out using 46.5-hr developing macronuclear DNA isolated from a mating of cell lines CC103 and CC91 as the substrate and the hot start procedure (Newton et al. 1989). For round I amplification, a cycle consisted of 94°C for 30 sec, 46°C for 1 min, and 72°C for 45–60 sec. Twenty-five microliters of round I reactions (total reaction volume was 100 µl)

was then used as the substrate for round II PCR using oligonucleotides V1CC and V1CG as the primers. Round II PCR was carried out for 27–28 cycles using the touchdown PCR procedure (Don et al. 1991). The touchdown procedure involved carrying out the initial cycles of PCR at an annealing temperature 6°C above the final annealing temperature and progressively lowering the temperature to the final annealing temperature during the first three to five cycles of PCR. The temperatures and times for the final cycles of round II PCR were as described for round I PCR. In some experiments, the PCR reaction mix contained 50 μ M tetramethylammonium chloride.

DNA sequencing and densitometric scanning

Dideoxy chain-termination sequencing of cloned PCR products was performed using a Sequenase 2.0 kit (U. S. Biochemical, Cleveland, OH) as described by the manufacturer. Following the sequencing reactions, a terminal deoxynucleotidyltransferase extension step (Fawcett and Bartlett 1990) was performed to reduce artifact banding. For the direct sequencing of PCR products derived from the 374-bp V1 IES, the PCR products were run on a LMP agarose gel and the specific product was purified from the gel as described by Qian and Wilkinson (1991). This material was then sequenced as described by Kusakawa et al. (1990), except that the sequencing reagents were from a Sequenase 2.0 kit. The V1CG oligonucleotide was used as the sequencing primer in all cases.

Autoradiographs of sequencing gels were scanned using a Bio-Rad (Richmond, CA) video densitometer equipped with 1-D Analyst software. To determine the percent of minor and major sequences at the circle junction region, the first position in the sequencing ladder where two sequences were present was scanned. This corresponds to an A residue in the major sequence and a G residue in the minor sequence (see Fig. 2A). Other positions within the circle junction could not be sufficiently resolved. The ratio of the signal intensity in the G lane as compared with the total signal intensity in the A and G lanes was then calculated to determine the percentage of the minor sequence. Note that this is an inexact method owing to variations in sequencing band intensity due to sequence context effects. Nevertheless, it provides a relative measure of the amounts of the major and minor sequences.

Acknowledgments

We thank Drs. Ann Cowan and Gordon Carmichael for their critical reading of the manuscript and Mr. Ken Hoppe for expert technical assistance. We also thank an anonymous reviewer of the manuscript for suggesting a useful modification to the excision model. This work was supported by U.S. Public Health Service grant GM33277 to L.A.K.

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

References

- Arnheim, N. and H. Erlich. 1992. Polymerase chain reaction strategy. *Annu. Rev. Biochem.* **61**: 131–156.
- Austerberry, C.F., R.O. Snyder, and M.-C. Yao. 1989. Sequence microheterogeneity is generated at junctions of programmed DNA deletions in *Tetrahymena thermophila*. *Nucleic Acids Res.* **17**: 7263–7272.
- Bainton, R., P. Gamas, and N.L. Craig. 1991. Tn7 transposition in vitro proceeds through an excised transposon intermediate generated by staggered breaks in DNA. *Cell* **65**: 805–816.
- Baird, S.E. and L.A. Klobutcher. 1988. Genetic characterization and use of a restriction fragment length variant in the hypotrichous ciliate *Euplotes crassus*. *J. Protozool.* **35**: 459–465.
- Baird, S.E., G.M. Fino, S.L. Tausta, and L.A. Klobutcher. 1989. Micronuclear genome organization in *Euplotes crassus*: A transposonlike element is removed during macronuclear development. *Mol. Cell. Biol.* **9**: 3793–3807.
- Benjamin, H.W. and N. Kleckner. 1989. Intramolecular transposition by Tn10. *Cell* **59**: 373–383.
- Caparon, M.G. and J.R. Scott. 1989. Excision and insertion of the conjugative transposon Tn916 involves a novel recombination mechanism. *Cell* **59**: 1027–1034.
- Don, R.H., P.T. Cox, B.J. Wainwright, K. Baker, and J.S. Mattick. 1991. "Touchdown" PCR to circumvent spurious priming during gene amplification. *Nucleic Acids Res.* **19**: 4008.
- Fawcett, T.W. and S.G. Bartlett. 1990. An effective method for eliminating "artifact banding" when sequencing double-stranded DNA templates. *BioTechniques* **9**: 46–48.
- Haselkorn, R. 1989. Excision of elements interrupting nitrogen fixation operons in Cyanobacteria. In *Mobile DNA* (ed. D.E. Berg and M.M. Howe), pp. 735–742. American Society for Microbiology, Washington, D.C.
- Hunter, D.J., K. Williams, S. Cartinhour, and G. Herrick. 1989. Precise excision of telomere-bearing transposons during *Oxytricha fallax* macronuclear development. *Genes & Dev.* **3**: 2101–2112.
- Jahn, C.L., L.A. Nilles, and M.F. Krikau. 1988. Organization of the *Euplotes crassus* micronuclear genome. *J. Protozool.* **35**: 590–601.
- Jahn, C.L., M.F. Krikau, and S. Shyman. 1989. Developmentally coordinated en masse excision of a highly repetitive element in *E. crassus*. *Cell* **59**: 1009–1018.
- Klobutcher, L.A. and C.L. Jahn. 1991. Developmentally controlled genomic rearrangements in ciliated protozoa. *Curr. Opin. Genet. Dev.* **1**: 397–403.
- Krikau, M.F. and C.L. Jahn. 1991. Tec2: A second transposon-like element demonstrating developmentally programmed excision in *Euplotes crassus*. *Mol. Cell. Biol.* **11**: 4751–4759.
- Kuhn, S., H.-J. Fritz, and P. Starlinger. 1979. Close vicinity of IS1 integration sites in the leader sequence of the *gal* operon of *E. coli*. *Mol. Gen. Genet.* **167**: 235–241.
- Kunkel, B., R. Losick, and P. Stragier. 1990. The *Bacillus subtilis* gene for the developmental transcription factor σ^K is generated by excision of a dispensable DNA element containing a sporulation recombinase gene. *Genes & Dev.* **4**: 525–535.
- Kusakawa, N., T. Uemori, K. Asada, and I. Kato. 1990. Rapid and reliable protocol for direct sequencing of material amplified by the polymerase chain reaction. *BioTechniques* **9**: 66–72.
- Lieber, M.R. 1991. Site-specific recombination in the immune system. *FASEB J.* **5**: 2934–2944.
- Maniatis, T., E.F. Fritsch, and J. Sambrook. 1982. *Molecular cloning: A laboratory manual*. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
- Margolis, P., A. Driks, and R. Losick. 1991. Differentiation and the establishment of cell type during sporulation in *Bacillus subtilis*. *Curr. Opin. Genet. Dev.* **1**: 330–335.
- Mizuuchi, K. 1984. Mechanism of transposition of bacteriophage Mu: Polarity of the strand transfer reaction at the initiation of transposition. *Cell* **39**: 395–404.
- Newton, C.R., A. Graham, L.E. Heptinstall, S.J. Powell, C. Summers, N. Kalsheker, J.C. Smith, and A.F. Markham. 1989. Analysis of any point mutation in DNA. The amplification

- refractory mutation system (ARMS). *Nucleic Acids Res.* **17**: 2503–2516.
- Poyart-Salmeron, C., P. Trieu-Cuot, C. Carlier, and P. Courvalin. 1990. The integration-excision system of the conjugative transposon Tn1545 is structurally and functionally related to those of lambdoid phages. *Mol. Microbiol.* **4**: 1513–1521.
- Prescott, D.M. 1992. Cutting, splicing, reordering, and elimination of DNA sequences in hypotrichous ciliates. *BioEssays* **14**: 317–324.
- Qian, L. and M. Wilkinson. 1991. DNA fragment purification: Removal of agarose 10 minutes after electrophoresis. *Bio-Techniques* **10**: 736–738.
- Ribas-Aparicio, R.M., J.J. Sparkowski, A.E. Proulx, J.D. Mitchell, and L.A. Klobutcher. 1987. Nucleic acid splicing events occur frequently during macronuclear development in the protozoan *Oxytricha nova* and involve the elimination of unique DNA. *Genes & Dev.* **1**: 323–336.
- Roth, M., M. Lin, and D.M. Prescott. 1985. Large scale synchronous mating and the study of macronuclear development in *Euplotes crassus*. *J. Cell Biol.* **101**: 79–84.
- Ruan, K. and S.W. Emmons. 1984. Extrachromosomal copies of transposon Tc1 in the nematode *Caenorhabditis elegans*. *Proc. Natl. Acad. Sci.* **81**: 4018–4022.
- Saiki, R.K., S. Scharf, F. Faloona, K.B. Mullis, G.T. Horn, H.A. Erlich, and N. Arnheim. 1985. Enzymatic amplification of β -globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science* **230**: 1350–1354.
- Tausta, S.L. and L.A. Klobutcher. 1989. Detection of circular forms of eliminated DNA during macronuclear development in *E. crassus*. *Cell* **59**: 1019–1026.
- . 1990. Internal eliminated sequences are removed prior to chromosome fragmentation during macronuclear development in *Euplotes crassus*. *Nucleic Acids Res.* **18**: 845–853.
- Tausta, S.L., L.R. Turner, L.K. Buckley, and L.A. Klobutcher. 1991. High fidelity developmental excision of Tc1 transposons and internal eliminated sequences in *Euplotes crassus*. *Nucleic Acids Res.* **19**: 3229–3236.
- Vieira, J. and J. Messing. 1982. The pUC plasmids, an M13mp7-derived system for insertion mutagenesis and sequencing with universal primers. *Gene* **19**: 259–268.
- Yao, M.-C. 1989. Site-specific chromosome breakage and DNA deletion in ciliates. In *Mobile DNA* (ed. D.E. Berg and M.M. Howe), pp. 715–734. American Society for Microbiology, Washington, D.C.

Circular forms of developmentally excised DNA in *Euplotes crassus* have a heteroduplex junction.

L A Klobutcher, L R Turner and J LaPlante

Genes Dev. 1993, **7**:

Access the most recent version at doi:[10.1101/gad.7.1.84](https://doi.org/10.1101/gad.7.1.84)

References

This article cites 32 articles, 8 of which can be accessed free at:
<http://genesdev.cshlp.org/content/7/1/84.full.html#ref-list-1>

License

Email Alerting Service

Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#).

Focused on your science.

USC
SCIENTIFIC
SIDE BY SCIENCE