A distant 10-bp sequence specifies the boundaries of a programmed DNA deletion in *Tetrahymena*

Ronald Godiska, Charles James, and Meng-Chao Yao
Division of Basic Sciences, Fred Hutchinson Cancer Research Center, Seattle, Washington 98104 USA

Programmed DNA deletion occurs at thousands of specific sites in most ciliates studied. To understand the mechanism of this prominent DNA rearranging process, we analyzed one of the deletion elements (the M-element) in *Tetrahymena* by making specific mutations in cloned DNAs and testing their effects on rearrangement in vivo. We found that a 10-bp polypurine sequence \(5'-AAAAAGGGGG1\) plays a crucial role. This sequence is located at a short distance \(-45 \text{ bp}\) outside of the element on either side. Removal of it abolishes the deletion process. Moving it short distances away causes the deletion boundary to move with it. Insertion of this sequence into a site within the element induces new boundaries to form near the insertion site. Sequence analysis reveals that all new boundaries created are 41–54 bp away from the sequence. Thus, this sequence is necessary and sufficient to determine the boundaries of DNA deletion, and it does so from a short distance outside of the element. These results offer a possible explanation for the control of DNA deletion in ciliates and suggest that a new type of mechanism for site-specific DNA rearrangements is involved.

[Key Words: DNA rearrangements; ciliates; polypurine sequence; developmental regulation]

Received August 9, 1993; revised version accepted September 13 1993.

Programmed DNA rearrangements occur in a wide variety of organisms during development (for review, see Borst and Greaves 1987). Some of the most intriguing examples are found in ciliated protozoa. In these species the genome structure is drastically reorganized during the formation of the macronucleus (the equivalent of a somatic nucleus). These changes include dimerization, amplification, deletion, elimination, inversion (for review, see Yao 1986, 1989), and shuffling (Greslin et al. 1989) of specific DNA segments, as well as breakage of chromosomes and addition of new telomeres (for review, see Yao 1989). These processes appear to be well regulated, however, their molecular mechanisms remain largely unknown.

One of the most widely occurring DNA rearranging events in ciliates is internal DNA deletion. In *Tetrahymena thermophila* it occurs at \(>5000\) sites in the genome (Yao et al. 1984). These deletion elements range in size from several hundred base pairs to \(>10\) kb, and they consist of both single-copy and moderately repetitive sequences. They are removed from the chromosomes at a specific stage during macronuclear development and are degraded immediately (Austerberry et al. 1984). The deletion process can be very precise, leading to the formation of identical junction sequences in most occurrences for some of the elements studied (Austerberry and Yao 1987; Austerberry et al. 1989). Similar processes have been found in other ciliates, including *Paramaecium* (Forney and Blackburn 1988), *Oxytricha*, and *Euplotes* (for review, see Klobutcher and Prescott 1986; Jahn 1991).

Superficially, DNA deletion in ciliates resembles certain well characterized DNA rearrangement events, such as \(\lambda\) prophage excision, in that each element is precisely removed from the chromosome at a specific stage of development and the remaining flanking sequences are rejoined. However, the underlying mechanisms may be rather different. Sequences of two deletion elements in *Tetrahymena* characterized so far share very few common features (Austerberry and Yao 1987, 1988). They also contain no prominent structures often found with mobile genetic elements, such as long open reading frames, long terminal repeats, or inverted terminal repeats. The lack of obvious sequence features is common among deletion elements in ciliates (for review, see Yao 1989), with the exception of two special elements found in *Euplotes* and *Oxytricha*. These two contain long inverted terminal repeats and most likely are transposons (Herrick et al. 1985; Jahn et al. 1988, 1989).

One of the most critical questions regarding the regulation of DNA deletion is the determination of its site specificity. This is especially interesting considering that there are thousands of different deletion elements in each genome that are removed at about the same time.

1Present address: Icos Corporation, Bothell, Washington 98021 USA.
2Corresponding author.
To further understand this phenomenon we have carried out detailed studies on one of these elements, the M-element (referred previously to as M-region deletion sequence) in *Tetrahymena*. This element was first identified as one of three deletion elements present in a germ-line (micronuclear) DNA clone, cTt455 (Yao et al. 1984). It is deleted from the chromosome in two alternative forms: a 0.6- and a 0.9-kb form that share the same right end (M3) and differ in their left ends (M1 and M2) (Austerberry et al. 1984; Austerberry and Yao 1988) (Fig. 1). Creation of these two forms occurs at roughly equal frequencies during development. The element is ~24% in G + C (similar to the rest of the genome) and contains no long open reading frames. The 0.6-kb form is bounded by a pair of 8-bp direct repeats, one copy of which is removed during deletion. The 0.9-kb form contains 5 bp of the 8-bp sequence at its left end (Austerberry and Yao 1988, Austerberry et al. 1989). The presence of short terminal direct repeats is perhaps the most noticeable common feature of deletion elements in ciliates. The sequences and lengths of these repeats vary, and their importance is not yet clear (for review, see Klobutcher and Prescott 1986; Yao 1989).

Another notable feature associated with the M-element is a 10-bp purine sequence, 5'-AAAAAGGGGG-3', one copy of which is located ~45 bp away from each of the two left ends (M1 and M2). An inverted copy of a similar sequence (5'-AATAAGGGGG-3') is present at a similar distance from the right end of the element (M3) (Austerberry and Yao 1988). These sequences (referred to as AsGs for simplicity) have not been found near any other deletion element studied so far.

To determine the cis-acting sequences that regulate DNA deletion, we have developed a transformation system through which a DNA construct can be introduced into the developing macronucleus of *Tetrahymena*, and its ability to go through rearrangement can be analyzed in the transformed cell (Tondravi and Yao 1986; Yao and Yao 1989). A cloned copy of the M-element is able to undergo rearrangement efficiently and accurately when tested by this method (Godiska and Yao 1990). This rearrangement is dependent on certain sequences that normally flank the element. M-elements with 65 bp or more of the left flanking sequence (sequence outside of M1, Fig. 1) can carry out the 0.9-kb deletion, whereas those with 33 bp or less cannot. The left-most copy of the A5G5 sequence is present between these two points. Substitution of this copy with A5C5G2 abolishes the 0.9-kb deletion. Thus, the A5G5 sequence must play an important role in DNA deletion. The 0.6-kb deletion occurs normally in all of these cases (Godiska and Yao 1990).

In this study we carried out further analysis to determine the role of the A5G5 sequence. We found that A5G5 is the only sequence signal that is required to specify M-element boundaries. It does so in a distance- and orientation-specific manner. Moving an existing copy of this sequence away from the element, or inserting additional copies at a new site within the element, causes deletions to occur with new termini. Remarkably, these new termini are always at specific distances (42–54 bp) from the 3' end of the A5G5 sequence. Other sequences at the termini, including the direct repeats, appear to have no essential role. This finding provides a clear example for the determination of DNA deletion site specificity in ciliates. It also reveals a new type of sequence signal that regulates DNA rearrangement sites at a distance.

### Results

**A5G5 specifies deletion boundaries at a distance**

We have shown previously that following its introduction into developing cells, the M1 or M2 terminus of an unaltered M-element in the rDNA vector is spliced to M3, accurately reproducing the normal process to eliminate 0.9 or 0.6 kb of the internal region (Godiska and Yao 1990). Further analysis revealed that the sequence A5G5m, which is located in both flanking regions of the M-element, is essential for deletion to occur. The sequences outside of the A5G5m tracts are not needed for this process. It was not known whether the sequences internal to these tracts, including the direct repeats at the deletion junctions, were important. Neither was it known how A5G5 worked to regulate DNA deletion. These issues were investigated further in the following experiments.

We constructed a DNA clone containing an M-element that lacked a stretch of 29 bp extending from within the direct repeat at M3 toward the flanking purine tract (Fig. 2) (for details, see Materials and methods). Synthetic oligonucleotides of varying sizes were

---

*Figure 1.* Diagram of the M-element and its deletion products. The micronuclear structure of M (unspliced) is presented at the top. The thick line represents the deletion element (the region to be deleted from the chromosome). Triangles depict the position and orientation of A5G5 tracts. Small arrowheads represent direct repeats at splice junctions. StuI indicates the position of the unique StuI site in the element. M1 and M2 are the two alternate left ends; M3 is the right end of the deletion element. The two alternative products of the deletion in the macronucleus are presented at the bottom: splicing between M2 and M3 to delete 0.6 kb, and splicing between M1 and M3 to delete 0.9 kb. Except for the symbols and the sequences that they represent, the diagram is drawn to the scale indicated.
changes made at M3, as aberrant deletions were also observed using unaltered M-elements [Godiska and Yao 1990]. In some cases, smearing bands were observed. This is probably attributable to aberrant rearrangements that remove some vector sequences downstream, thus linking this sequence to the heterogeneous telomeric sequences of this vector.

We determined the sequences of five rejoined junctions from these transformed Tetrahymena and found that the right boundaries of these deletions were scattered within a 13-bp region near the remnant of the terminal repeat at M3 [Fig. 4]. These sites shared sequence identity 3 bp or less with the native direct repeat. It is significant that the splice sites were not at the remnant terminal repeat but to the left of it, thus maintaining the characteristic distance between the junction and the A5G6 sequence. This result suggests that the distance from A5G6, and not the specific sequence of the splice site, is important. To explore this idea further, we made and tested a construct in which an additional 12 bp was inserted into the same location [S12−, see Fig. 2]. Deletions also occurred efficiently in this construct (Fig. 3b). Three junctions produced from it were analyzed. Sig-

![Figure 2](image_url)

**Figure 2.** DNA constructs with altered distances between the polypurine tract and the direct repeat of M3. A stretch of 29 bp between the polypurine tract and the direct repeat of M3, including part of the direct repeat, was deleted and replaced by oligonucleotides of various lengths. Clones containing inserts of the following lengths were analyzed: 20, 32, 89, and 103 bp. The inserts are indicated by thick lines in these diagrams and their sequences are given in Fig. 4. (Right) The names of the clones. The symbols are as described in Fig. 1.

then inserted to replace the missing bases. These clones also contained 250 bp of the left flanking sequence (outside of M1) and 0.9 kb of the right flanking sequences (outside of M3) [Godiska and Yao 1990]. The constructs were inserted into the micronuclear rDNA vector pD5H8 and introduced back into Tetrahymena by injection. As described previously, this rDNA is processed and amplified in the developing macronucleus, which confers paromomycin resistance to the injected cells ['Yao and Yao 1989]. These transformed cells were identified and their DNAs analyzed by Southern hybridization using an oligonucleotide containing a sequence just outside of M1 as a probe. Hybridizations using a 2.1-kb probe that contained the entire inserted sequence gave similar results [data not shown].

In the first construct tested, a 20-bp linker was inserted into the modified M3 junction, effectively moving the A5G6 motif 9 bp closer to the center of the element. Cells transformed with this molecule efficiently spliced either M1 or M2 to points near the altered M3 [Fig. 3a]. Of the 11 transformed lines analyzed, 9 carried out deletions between M1 and M3 [producing the 1.3-kb band] and 6 carried out deletions between M1 and M3 [producing the 1.6-kb band]. The levels of deletion are similar to those observed earlier using an unaltered M-element in the same system. Thus, the 29-bp sequence removed from the M3 junction, which includes part of the terminal repeat, is apparently not required for the process. Together with earlier results, this result indicates that A5G6 is the major, and likely the only, sequence within the flanking regions that is essential for DNA deletion to occur. Some aberrant deletions were also observed, but they are probably not related to the

![Figure 3](image_url)

**Figure 3.** Rearrangement of M-elements with altered M3 junctions. Whole-cell DNA [2−5 µg] from Tetrahymena transformants was cut with BamHI, fractionated by electrophoresis in 1% agarose gels, blotted to membrane filters, and hybridized to an oligonucleotide probe, which includes the sequence of the M1 polypurine tract [oligonucleotide M9AG]. Each lane contains DNA from a single transformed line. The results from injection of construct S12− are given in a, those from construct NSX12-8 and NSX12-10 in c. The numbers at right indicate the estimated sizes [in kb] of fragments migrating at those positions. The band at 2.2 kb is derived from unprocessed molecules, that at 1.6 kb from M2−M3 splicing, and that at 1.3 kb from M1−M3 splicing.
Figure 4. Nucleotide sequences at deletion junctions of altered M-elements. The positions of the junctions produced from deletion of the native, as well as the altered, M-elements are compared on the right. The entire sequence from the polyuridine to the splice site of the un-rearranged [micronuclear or injected] DNA is displayed for these regions. The synthetic oligonucleotides inserted into these DNAs are indicated by thin underlines. The sites at which a deletion begins and ends (the deletion boundaries) are indicated by arrows below the sequences. In many cases, these boundaries are ambiguous due to the presence of terminal direct repeats, which are indicated by thick underlines. The CsTs (or CsT2AT2} motifs that flank these boundaries are aligned to each other and enclosed in boxes. The direct repeats at all boundaries (M1, M2, and M3} and their orientations are marked by arrows above them. These repeats, as well as their remnants at M3, are also enclosed in boxes. (Left) The location of the corresponding boundary of each deletion whose other boundary is shown at right. The lowercase letters beneath the arrows identify the two corresponding boundaries from a given deletion event. For example, the two boundaries indicated by a in the left and the right half of a sequence identify the two ends of one deletion event.

Significantly, the right deletion boundaries were shifted to a region located to the right of (external to) the remnant of the direct repeat at M3, again maintaining the characteristic distance from A5G5 [Fig. 4]. Moreover, increasing the length of the insertion to 89 or 103 bp did not inhibit splicing [Fig. 3c). Two deletion boundaries produced from the 103-bp insertion clone were analyzed. Again, they retained the characteristic positional relationships to A5G5 [Fig. 4]. These findings clearly demonstrate that the A5G5 tract is responsible for specifying deletion boundaries. It does so at a short and slightly variable distance. The direct repeat and the sequences between it and A5G5 do not appear to play any significant role.

The sequencing results also show that the left boundaries of these deletions, with one exception, all fell within or very near the direct repeats at M1 or M2 [Fig. 4]. Thus, changing the sequence at the right boundary mainly alters the junctions created at that boundary. It has only very limited effect on junctions created at the left boundaries. Because the left boundaries occurred at or near normal sites, we believe that the rearrangements we have observed occurred by a normal process. It is noteworthy that none of the left deletion boundaries were at exactly the same point. This microheterogeneity may have been generated because of the lack of the native 5- or 8-bp direct repeat in the right boundaries. However, it should also be pointed out that the presence of direct repeats at each of the native sites makes it impossible to determine the exact junction points within the repeats, as long as both junctions occur at the same point within each repeat. Thus, the function of the direct repeats may be to "fine-tune" the deletion process, and/or
Figure 5. Insertion of a polypurine tract into the interior of the M-element. A synthetic oligonucleotide containing an A5G5 motif and flanking restriction sites (for sequence, see Materials and methods) was ligated in either orientation into the StuI site of the M-element. The clone containing the insert in the CsTs orientation actually contains three tandem copies of this oligonucleotide. A clone containing the insert in the AsG5 orientation was modified further by placing a 0.4 kb stuffer fragment adjacent to the polypurine insert.

to obscure small variations in splicing among a population of molecules, thus enhancing the homogeneity of the deletion junctions.

Inserted A5G5 induces deletion

Based on the studies described above, we suspected that A5G5 might be the only signal required to determine the site of the deletion junction. To test this idea directly, a synthetic oligonucleotide containing this sequence was inserted within the eliminated portion of an M-element clone that included 250 bp to the left of M1 and 70 bp to the right of M3. M-element clones with this amount of flanking sequence were known to undergo deletion, albeit at somewhat reduced efficiencies (Godiska and Yao 1990). If the inserted oligonucleotide can promote deletion at its new location, shorter M-element deletions should occur.

The inserted oligonucleotide contained A5G5 between restriction sites (see Materials and methods). It was inserted in either orientation into the unique StuI site between M2 and M3 (Fig. 5). Oligonucleotides inserted in the 5'-A5G5 orientation (as read from M1 toward M3) were similar in orientation to the native copies at M1 and M2 and thus might direct splicing to the M3 junction if polarity of the sequence is important to this process. Those in the opposite, or CsTs, orientation resemble the polypurine tract at M3; consequently, they might be spliced to M1 or M2. In the particular clone tested for the CsTs orientation, three tandem copies of the oligonucleotide were inserted.

One of the probes used for the Southern hybridization analysis of the transformed cell DNA is the synthetic oligonucleotide itself. If A5G5 functions to induce splicing to a downstream (3' side) site, it should not be removed by the induced deletion. Molecules that have undergone splicing to the induced sites will therefore hybridize to the probe. However, those molecules that have undergone normal splicing of M1 or M2 to the junction at M3 will not be detected, because they will have deleted the region containing the inserted oligonucleotide.

The analysis of the insertion in the CsTs orientation demonstrates that the synthetic sequence does appear to have induced a splice site at the predicted location (Fig. 6b). The bands at 1.7 and 1.4 kb represent deletions of 0.4 and 0.7 kb, respectively. These correspond to splicing from M2 or M1 to the induced site, M-a, which is very close to the inserted oligonucleotide. A large amount of unprocessed material is also present (the 2.1-kb band), reflecting the fact that not all of the injected molecules were processed. This finding is typically observed in our experiments, and it is presumably related both to injecting an excess of substrates and to the shortened flanking sequence present in these clones (Godiska and Yao 1990). Reprobing this gel blot with a 2.1-kb fragment containing the M-element and flanking sequences reveals the products derived from splicing of the M1 and M2 sites to M3, in addition to the bands noted previously (Fig. 6c). Similar results were obtained when probing these samples with an oligonucleotide complementary to...
to a sequence outside of M1 [data not shown]. Thus, individual molecules within the injected population underwent any one of the four combinations of splicing from M1 or M2 to the sites at M-a or M3.

To ascertain that the inserted signal does not induce deletion between Ma and M3, a similar gel blot was hybridized with an oligonucleotide probe complementary to a sequence between M2 and M3 (oligomer K861M). As expected, the probe hybridized mainly to the un-rearranged materials. No band corresponding to the expected deletion products (1.8 kb) was detected [data not shown]. Thus, AsGs does not appear to induce deletion in the other orientation.

Insertion of the same synthetic sequence in the opposite [AsGs] orientation simulates the sequence signals at M1 and M2 junctions; it would therefore be expected to induce splicing to the M3 site, deleting 0.2 kb. However, initial attempts using this arrangement showed no detectable induction of such a site [data not shown]. Because previous work suggested that there might be a lower limit on the size of a fragment that can be deleted (R. Godiska and M.-C. Yao, unpubl.), a 0.4-kb “stuffer” fragment was inserted between the synthetic polyuridine tract and M3. This stuffer DNA was from a neighboring region of the genome and was not known to be involved in DNA rearrangement. In this case, induction of the expected splice site M-b was observed, reducing the size of the injected molecule by 0.6 kb [Fig. 7b]. Hybridization of a 2.1-kb probe containing the M-element and flanking sequences [Fig. 7c], or an oligonucleotide probe to the sequence flanking M1 [data not shown], detected the products created by splicing from M1 and M2, in addition to those from M-b. Several aberrantly sized molecules have also been detected. To determine whether any of them are produced from deletions between Mb and M1 or M2, an oligonucleotide complementary to an M-element sequence between M2 and M3 was used to probe a similar blot containing the same transformed line DNA. No bands corresponding to these deletions (2.1 kb for Mb–M2 deletion and 1.8 kb for Mb–M1 deletion) were detected [Fig. 7d]. Thus, although some degeneracy was introduced into the process, the creation of a deletion by the inserted oligonucleotide emphasizes the ability of AsGs to direct splicing to a specific site in an orientation-dependent manner.

Junction sequences of the induced deletions
To further understand the nature of the induced deletions, the sequences of the rejoined junctions in several of the transformed lines were determined [Fig. 4]. In all cases involving the inserted AsGs tract, one end of the deleted region resides within the direct repeat (M2 or M3), whereas the other occurs at a short distance from the inserted sequence. In transformants containing the insert at the AsGs orientation, the distance is 43–47 bp from the end of the AsGs sequence. This spacing agrees very well with the distance between AsGs and the splice site in each of the native sites and the artificial sites described above [Fig. 4]. For transformants with the in-

---

**Figure 7.** Induction of a splice site by insertion of an AsGs tract. [a] Schematic drawing of the injected DNA molecule (clone 731AG-C1) and its induced processed forms, showing positions of the inserted oligonucleotide, stuffer fragment, induced splice site M-b, and positions of the oligonucleotide sequences used as probes (indicated by asterisks). The other symbols are the same as those described in Fig. 1. [b] Whole-cell DNA from individual transformants (lanes 1–19) was digested with EcoRI, fractionated by electrophoresis, blotted, and hybridized with an oligonucleotide probe that contained the inserted sequence. EcoRI cuts at vector sequences, which are outside of the DNA shown in the diagram, to generate a fragment contains the inserted sequence plus 0.74 kb of the insert. The estimated sizes of DNA fragments in the gel are indicated at right (in kb). The band at 2.4 kb represents unprocessed molecules, the band at 1.8 kb is derived from splicing of M-b to M3, deleting the stuffer (0.4 kb) plus the StuI–M3 region (0.2 kb). [c] Reprobing the same blot with a 2.1-kb fragment containing M-element and flanking sequences reveals splicing of M1 to M3 (1.4-kb band), M2 to M3 (1.4-kb band), and M-b to M3 (1.8-kb band). The other bands have not been analyzed. [d] Probing a similar gel blot with an oligonucleotide (M1022) containing a sequence between Mb and M3 shows that most of the deletion products do not retain this sequence. In this particular gel, the hybridizing DNA in lane 8 is under-represented because of a failure in restriction digestion.
insert in the C₅T₅ orientation, the situation was slightly more complicated, as three tandem copies of the polypurine tract had been inserted. Although the effect of the additional copies is not clear, the location of the splice site 48 bp from the end of the most distal copy corroborates the previous results well. Taken together, these data confirm that the sequence A₅G₅ induces deletion to occur at specific distances downstream from it. The spacing of 42–53 bp between the polypurine sequence and the splice site appears to be an intrinsic property of this deletion mechanism. The data further indicate that this sequence signal functions even when placed outside of its original locations at the termini of the element. Thus, A₅G₅ itself appears to carry sufficient sequence information for determining the boundaries of the M-element deletion.

The above results also demonstrate that splicing does not require a perfect copy of the direct repeat at each of the native sites. Deletions induced by inserts in the A₅G₅ orientation have 2, 3, or 4 bp of the repeat at their termini. This is probably more than a coincidence and suggests that short repeats might have a nonessential minor role. In the C₅T₅ orientation, the single induced junction analyzed contains no terminal direct repeats [Fig. 4].

Discussion

In this study we present direct evidence that the polypurine sequence 5’-AAAAAGGGGG-3’, or close variants of it, determines the boundaries of a programmed DNA deletion in Tetrahymena. This type of developmentally regulated DNA deletion occurs at thousands of specific sites in the genomes of Tetrahymena and other ciliates [for review, see Yao 1989; Jahn 1991]. Although the structure of several such deletion elements has been characterized, the molecular mechanism of this process is not known. The M-element of Tetrahymena is perhaps the best-characterized example of these elements. By direct functional analysis we have shown previously that the regions adjoining this element, including the A₅G₅ motif, are important for deletion to occur [Godiska and Yao 1990]. In this study we show that the A₅G₅ sequence plays one of the most crucial roles in the process. By inserting copies of this motif into a new location and by moving one of the existing copies various distances away from the element, we show that the boundaries of deletion are determined solely by this motif in an orientation- and distance-specific manner. This result offers a simple explanation for how the specificity of this DNA deletion is determined. It also serves as a clear example for understanding the deletion of other elements in ciliates. It is likely that other deletion elements are regulated by a similar fashion: Their site specificities are determined by a pair of flanking sequences in a distance- and orientation-specific manner. Because A₅G₅ is not present near other deletion elements characterized so far, it cannot be the only controlling signal. Because there are thousands of deletion elements, we propose that there are more than a few such flanking regulatory sequences, and each controls a family of elements. This could explain why A₅G₅ is not present in the other deletion elements characterized so far. Our recent analysis of the neighboring R-element renders further support to this view. The R-element shares no sequence identity with the M-element [Austerberry and Yao 1987, 1988], yet its deletion requires its immediately flanking sequences but not sequences at its junctions [A. La Terza, A. Wilson, and M.-C. Yao, unpubl.]. Thus, the R-element may be a member of a different family, which uses a similar mechanism but a different signal to regulate its deletion.

It should be pointed out that the studies presented here do not address the issue of whether any sequence within the M-element is also required for deletion to occur. Although A₅G₅ is a sufficient signal to determine the locations of deletion junctions, other sequences may still be required for deletion to occur efficiently. A comparison of the lower efficiencies in the formation of the induced junctions with the normal junctions in Figure 6 and 7 supports this possibility. In any case, if other sequences are required, they will not have an essential role in determining the locations of the deletion boundaries.

The sequence motif A₅G₅ determines the deletion boundaries in a novel fashion. The action site [the boundary] of the deletion is always a short distance away from its recognition sequence A₅G₅, which is located outside of the element itself. This distance is 44–51 bp between the end of the native A₅G₅ and the deletion boundary [the midpoint of the terminal direct repeats] at M1, M2, or M3, and it is 41–53 bp for the 15 artificially induced junctions analyzed in this study. The positions of the junctions are apparently not affected significantly by neighboring sequences, including those at the junction. The 13-bp range represents a degree of microheterogeneity that not only exists among junctions generated in different constructs but also among different junctions from the same construct. Thus, in this system the action site of deletion is located at a short and somewhat variable distance from its recognition signal. This type of control is rarely seen in DNA rearrangement systems, suggesting that a novel mechanism may be involved. The machinery that carries out the process probably contains at least two interacting components, one recognizing the sequence signal and the other executing the enzymatic reactions. They could be two domains of a single molecule or two interacting molecules. Although there are examples of recombination systems in which the recognition sequence and the action sites are separated, such as the transposition of Tn7 in Escherichia coli [for review, see Craig 1989], they usually have a shorter and invariable distance between the two DNA sites.

The presence of short terminal direct repeats is the most significant feature common to deletion elements in Tetrahymena and other ciliates [for review, see Klobutcher and Prescott 1986; Yao 1989; Jahn 1991]. Whereas their prevalence suggests a possible role, our studies here show clearly that they are not essential for deletion to occur. Removal of a part of this sequence at the M3 junc-
tion does not prevent deletion from occurring. Deletion boundaries also occur readily at many new sites that do not contain the repeat. However, closer examinations of our results suggest a possible role for these sequences in influencing the exact boundaries of a deletion. For instance, the normal M-element contains a pair of 8-bp repeats at the M2 and M3 junctions. During deletion, one copy of these repeats is invariably and precisely removed (Austerberry et al. 1989). When the copy at M3 is partially destroyed, deletion junctions no longer occur at exactly the same locations. They are scattered within the short range specified by AsGs. Upon closer examination, one found that in 12 of 15 cases these new junctions occur at sites with direct repeats of 1–4 bp in length. On the basis of these results, we suggest that short direct repeats, if placed at a proper distance from AsGs, are preferred sites for deletion to occur. One can imagine a mechanism in which AsGs determines a cut site some 45 bp away. The two exposed ends are then brought together and aligned with the help of the direct repeats, when present, before they are rejoined.

Our results present a simple explanation for how the site specificity of M-element deletion is achieved, which might serve as a general model for DNA deletion in ciliates. The mechanism includes several distinctive features. First, the critical recognition signal is located outside of the elements to be deleted, which is distinct from most mobile genetic elements, in which the recognition signals are located inside the element. Second, the sequence signal AsGs does not resemble any sequence signal found previously. It is possible that this polypurine sequence has a special secondary structure that is important for its function. Third, the recognition signal of this element is separated from its action site by a short and variable distance, which implies unusual features in its mechanism. One must therefore contemplate the possibility that DNA deletion in ciliates involves a new type of process. Elucidating its mechanism could lead to a better understanding of DNA rearrangements in general.

Materials and methods

Cells and cultures

*T. thermophila* strains used in this study were CU427 and CU428 of inbreeding line B obtained from Peter Bruns of Cornell University (Ithaca, NY). Cells were maintained and grown in axenic media as described by Gorovsky et al. (1975).

Microinjection

Microinjection of DNA into developing macronuclear anlagen was carried out using the method first described by Tondravi and Yao (1986) and later modified by Godiska and Yao (1990). Between 150 and 500 cells were typically injected with each DNA sample to generate the transformants produced (with transformation frequencies of 1–6%). Individual injected cells were cloned and grown for 2–3 days before drug selection. Paromomycin-resistant clones were grown to saturation in 5–10 ml of media for an additional 2–3 days before being harvested for DNA analysis.

DNA analysis

*Tetrahymena* DNA was isolated by phenol extraction and polyethylene glycol precipitation (Austerberry and Yao 1987). Restriction enzyme digestion, agarose gel electrophoresis, Southern blotting, and hybridization were carried out following standard protocols as described earlier (Austerberry et al. 1984). DNA and oligonucleotides were labeled with 32P by nick translation or T4 polynucleotide kinase, respectively. DNA sequences of the rearranged junctions were determined from one strand on double-stranded plasmids by the chain termination method of Sanger, using the Sequenase v.2 kit (U.S. Biochemical Corp.).

Plasmid construction

The M-element (referred to previously as M-region deletion sequence) was originally cloned by Yao et al. (1984). A 2.1-kb segment containing the element was subcloned in either orientation into a vector bearing adjacent *KpnI* and *BstEI* sites, facilitating unidirectional deletion with λ exonuclease III (Henikoff 1984). Two sets of unidirectional deletions were thereby created, extending from either end of the subclone to various points within the M-element (Godiska and Yao 1990). Sequencing the resulting deletion products revealed one clone that lacked the sequences downstream of the direct repeat at junction M3 and another that lacked most of the region upstream of the C7T7 tract at M3. These two fragments were ligated in the presence of a synthetic oligonucleotide (5'-GGCATCAG-GAAATCCCATC) that introduced a unique *EcoRI* site between them. To increase the length of this insert, a *SalI* linker (5'-GGTCGACC) was inserted into the *EcoRI* site. The length was increased further by inserting multiple copies of another synthetic oligonucleotide (5'-GGCGGGGGGGGGGCTAGGT-TACCGGTACCGCGGCC) into this *SalI* site.

To construct M-elements containing extra copies of AsGs, a synthetic double-stranded oligonucleotide was inserted into the *Stul* site within the M-element in a clone that contained 250 bp upstream of M1 and 70 bp downstream of M3 (Godiska and Yao 1990). The sequence of the oligonucleotide was 5'-GATAT-CAAAAAAGGGGGATGGGGATCCAAGCTTA, which contains the AsGs sequence flanked by an *EcoRV* site upstream and a *BamHI* and *HindIII* site downstream. The clones were sequenced to determine the orientation and the copy number of the insert.

The sequence of the oligonucleotide probe specific for the AsGs tract near M1 is 5'-TAAAAAAGGGGGTAAAATAA-TAAGG-3' (oligonucleotide M9AG). The two oligonucleotide probes that contain sequences specific for the internal region of the M-element are M1022, 5'-CATATGATCAGGTCCAGCAG-3' and K861M, 5'-TCTTATAAGAGACCCAACTGTTATTAG-3'. The sequence in M1022 is located 10 bp from the M3 junction. K861M contains the sequence located 280 bp from the M2 junction, as well as an added *KpnI* site at its 5' end. The vector pD5H8 was used in this study for transformation of *Tetrahymena*. It was modified from T94701, which was created originally by inserting a 14-kb fragment of the micronuclear DNA encoding the rRNAs of *Tetrahymena* (Yao and Yao 1989) into a pUC-based plasmid. The rDNA allele used (C3 type) is preferentially maintained over the host rDNA type (B type) (Larson et al. 1986) and contains a mutation conferring resistance to paromomycin (Bruns et al. 1985). pD5H8 was created by recombining two modified versions of 94701: the 5' half of one in which all but 199 bp of the 5' flanking sequences were deleted by exonuclease III digestion (T947-400) (Yao et al. 1990), and the 3' half of the other, which contains a polylinker.
at its 3' spacer region (947H8) [Sweeney and Yao 1989], into which fragments of interest can be inserted.

Acknowledgments

We thank Ching-Ho Yao for her excellent technical help in part which fragments of interest can be inserted.

American Cancer Society during part of this work.

National Institute of General Medical Sciences (GM26210).

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

Austerberry, C.F. and M.C. Yao. 1987. Nucleotide sequence structure and consistency of a developmentally regulated DNA deletion in Tetrahymena thermophila. Mol. Cell. Biol. 7(1): 435–443.

Austerberry, C.F. and M.C. Yao. 1988. Site-specific chromosome breakage and DNA deletion in Tetrahymena thermophila. Mol. Cell. Biol. 8(9): 3947–3950.

Austerberry, C.F., C.D. Allis, and M.C. Yao. 1984. Programmed DNA rearrangements in synchronously developing nuclei of Tetrahymena. Proc. Natl. Acad. Sci. 81(23): 7383–7387.

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(18): 7263–7272.

Borst, P. and P.R. Greaves. 1987. Programmed gene rearrangements altering gene expression. Science 235: 658–667.

Bruns, P.J., A.L. Katzen, L. Martin, and E.H. Blackburn. 1985. A drug resistant mutation maps in the ribosomal DNA of Tetrahymena. Proc. Natl. Acad. Sci. 82: 2844–2846.

Craig, N. 1989. Transposon Tn7. In Mobile DNA (ed. D. Berg and M. Howe), pp. 211–225. American Society for Microbiology Press, Washington, D.C.

Forney, L.A. and D.M. Prescott. 1986. The special case of the hypotrichs. In The molecular biology of ciliated protozoa, pp. 111–154. Academic Press, Orlando, FL.

Yao, M.-C., C.-H. Yao, and B. Monks. 1990. The controlling 3' spacer region (947H8) [Sweeney and Yao 1989], into which fragments of interest can be inserted.

DNA deletion in Tetrahymena

Jahn, C.L., M.F. Krikau, and S. Shyman. 1989. Developmentally coordinated en masse excision of a highly repetitive element in Euplotes crassus. Cell 59: 1009–1018.

Klobutcher, L.A. and D.M. Prescott. 1986. The special case of the hypotrichs. In The molecular biology of ciliated protozoa, pp. 111–154. Academic Press, Orlando, FL.

Larson, D.D., E.H. Blackburn, P.C. Yeager, and E. Orias. 1986. Control of rDNA replication in Tetrahymena involves a cis–acting upstream repeat of a promoter element. Cell 47: 229–240.

Sweeney, R. and M.C. Yao. 1989. Identifying functional regions of rRNA by insertion mutagenesis and complete gene replacement in Tetrahymena thermophila. EMBO J. 8(3): 933–938.

Tondravi, M.M. and M.C. Yao. 1986. Transformation of Tetrahymena thermophila by microinjection of ribosomal RNA genes. Proc. Natl. Acad. Sci. 83(12): 4369–4373.

Yao, M.-C. 1986. Amplification of ribosomal RNA genes. In The molecular biology of ciliated protozoa, pp. 179–201. Academic Press, Orlando, FL.

Yao, M.C. and C.H. Yao. 1989. Accurate processing and amplification of cloned germ line copies of ribosomal DNA injected into developing nuclei of Tetrahymena thermophila. Mol. Cell. Biol. 9(3): 1092–1099.

Yao, M.C., J. Choi, S. Yokoyama, C.F. Austerberry, and C.H. Yao. 1984. DNA elimination in Tetrahymena: A developmental process involving extensive breakage and rejoining of DNA at defined sites. Cell 36(2): 433–440.

Yao, M.-C., C.-H. Yao, and B. Monks. 1990. The controlling sequence for site-specific chromosome breakage in Tetrahymena. Cell 63: 763–772.

GENES & DEVELOPMENT 2365
A distant 10-bp sequence specifies the boundaries of a programmed DNA deletion in Tetrahymena.

R Godiska, C James and M C Yao

*Genes Dev.* 1993, 7:
Access the most recent version at doi:10.1101/gad.7.12a.2357

---

**References**  
This article cites 21 articles, 3 of which can be accessed free at:  
http://genesdev.cshlp.org/content/7/12a/2357.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](#).