The terminal organization of macronuclear DNA in *Oxytricha fallax*

Ann F. Pluta*, Brian P. Kaine⁺ and Brian B. Spear[§]

Department of Biochemistry, Molecular and Cell Biology, Northwestern University, Evanston, IL 60201, USA

Received 20 August 1982; Revised and Accepted 9 November 1982

ABSTRACT

The macronucleus of the hypotrichous ciliate *Oxytricha fallax* is transcriptionally active and contains linear achromosomal DNA molecules that function as single-gene units. The terminal organization of macronuclear DNA was analyzed by chemical sequencing and S1 mapping. The terminal sequence of total macronuclear DNA was determined for molecules labeled at the 5' or 3' ends. Results indicate that the 5' sequence C₄A₄C₄A₄C₄ and the 3' sequence G₄T₄G₄T₄G₄T₄G₄ occur at both ends of all DNA molecules in the macronucleus. The discrepancy in the length of the common terminal sequence between the 5' and 3' ends was clarified by a limited S1 digestion experiment, which indicated the existence of a 16 nucleotide long single-stranded tail at the 3' ends.

INTRODUCTION

The genome of hypotrichous ciliates consists of two structurally and functionally distinct nuclei. The transcriptionally inactive micronucleus contains long multigenic DNA molecules typical of eukaryotic chromosomes. The transcriptionally active macronucleus is derived from the micronucleus by a complex process involving the polytenization of the micronuclear genome, and subsequent fragmentation and selective degradation of the polytene chromosomes followed by a 1000-fold duplication of the conserved sequences (1-4). The resulting mature macronucleus contains achromosomal DNA molecules ranging in size from 22,000 to 400 base pairs, that exist as linear DNA molecules and function as single-gene units (6-8).

The termini of macronuclear DNA molecules are of considerable interest. Upon thermal denaturation and reannealing, macronuclear DNA molecules form single-stranded circles, indicating the presence of terminal inverted repetitive sequences (7,9). Further characterization of these sequences indicated that a common terminal repeat is shared by the ends of all DNA molecules in the macronucleus (10). Electron microscopic and chemical nucleotide sequence analyses of macronuclear DNA from several hypotrichs has verified

the existence of a common terminal sequence at the ends of these molecules (9,11).

In this report we confirm these results with the demonstration of a shared terminal nucleotide sequence in total macronuclear DNA from Oxytricha fallax. Evidence is also presented for the presence of this sequence in a specific macronuclear DNA molecule, that coding for actin. Furthermore, we extended previous findings by demonstrating that 3' ends of total macronuclear DNA exist as free single-stranded protrusions.

MATERIALS AND METHODS

Culturing of *O. fallax* and DNA isolation

Oxytricha fallax were grown and macronuclear DNA was prepared as described by Rae and Spear (5). DNA from the recombinant plasmid pOfACT (1.6) was prepared as described by Kaine and Spear (8).

Radioactive end-labeling and DNA sequencing

All DNAs used for sequencing were separated from residual RNA and mononucleotides before end-labeling by filtration through a 5 ml Sepharose 2B column (Pharmacia) equilibrated in 10mM Tris-HCl, 1mM EDTA pH 7.5. DNA was radioactively labeled at the 5' end with $\gamma^{32}\text{P}$ -ATP and T4 polynucleotide kinase after dephosphorylation as described by Maxam and Gilbert (12) for protruding 5' ends, and at the 3' end with terminal transferase and $\alpha^{32}\text{P}$ -cordycepin 5' triphosphate according to the protocol provided with the New England Nuclear 3' end-labeling kit (cat# NEK 009). Chemical sequencing of end-labeled DNA and polyacrylamide gel electrophoresis were performed as described by Maxam and Gilbert (12), except that all base-specific modification reactions were carried out at 20°C.

Determination of the 5' terminal nucleotide of macronuclear DNA

Macronuclear DNA labeled at the 5' ends with ^{32}P (10-50,000 cpm) was digested with 10 units DNase I at 10mM Tris-HCl pH 7.6, 10mM MgCl_2 at 37°C for 2 hours with unlabeled pBR322 DNA as carrier. The reaction mixture was then adjusted to 50mM Tris-HCl pH 9.0 and treated with 3×10^{-3} units snake venom phosphodiesterase (Sigma, Grade III) in 0.1M Tris pH 9.0, 0.1M NaCl, 14mM MgCl_2 , and 0.1M bovine serum albumin at 37°C for 2 hours. The reaction mixture was then spotted on a cellulose TLC plate (Kodak, 13254) and developed to 15 cm from the origin with 80% saturated NH_4SO_4 , 2% isopropanol, 0.18M sodium acetate. Nucleotides were identified by autoradiography and short wave UV illumination, and quantified by scintillation counting.

Conditions for S1 limited digestion of 3' end-labeled macronuclear DNA

3' end-labeled total macronuclear DNA was adjusted to 0.5 mM sodium acetate pH 4.5 and 0.5mM ZnSO_4 at 0°C. $\sim 2 \times 10^{-3}$ units of S1 nuclease (Sigma) in 0.01M sodium acetate pH 4.5 was added, and the reaction was warmed to room temperature. After 6 minutes, the samples were heated to 100°C to stop the reaction and mixed with sequencing gel dye for polyacrylamide gel electrophoresis.

RESULTS

Chemical sequencing of total macronuclear DNA purified by standard laboratory procedures through CsCl indicates the presence of a common terminal nucleotide sequence at both ends of essentially every DNA molecule in the macronuclei of *Oxytricha fallax*. Cleavage products of DNA radioactively labeled at the 5' or 3' ends were electrophoresed on polyacrylamide sequencing gels and visualized by autoradiography. Figure 1 shows the ladder of bands corresponding to the 20 nucleotide long sequence $\text{C}_4\text{A}_4\text{C}_4\text{A}_4\text{C}_4$ at the 5' ends of total macronuclear DNA (Fig. 1A) and the 36 nucleotide long sequence $\text{G}_4\text{T}_4\text{G}_4\text{T}_4\text{G}_4\text{T}_4\text{G}_4\text{T}_4\text{G}_4$ at the 3' ends of total macronuclear DNA (Fig. 1B). Incorrect spacing of the extra bands seen between the first and second nucleotides at the 5' end (Fig. 1A) suggest that these are not part of the terminal sequence, but rather represent a hydrazine artifact produced by incomplete β -elimination during chemical cleavage (12-14). The common terminal sequence terminates abruptly at a transition site 20 base pairs from the 5' ends (arrows, Figure 1), where bands appear non-specifically in all four nucleotide lanes of the sequencing gel.

Because the 5' terminal nucleotide of the common terminal sequence was consistently obscured on our sequencing gels, macronuclear DNA labeled at the 5' ends was digested extensively with DNase I and snake venom phosphodiesterase to generate free 5' terminal nucleotides. Upon thin layer chromatography of the digestion products, 70.8% of the radioactivity was associated with dCMP, 18.2% with TMP, 7.4% with dGMP and 3.7% with dAMP, indicating that the 5' terminal nucleotide of total macronuclear DNA is primarily dCTP.

Although the nucleotide sequences determined for the 5' and 3' ends of total macronuclear DNA are entirely complementary, they differ in overall length, as shown previously for *O. nova* (11). In Figure 1A, the 5' sequence becomes non-homogeneous at nucleotide position 21, whereas in Figure 1B, the 3' sequence becomes non-homogeneous at nucleotide position 37. This discrepancy in the number of nucleotides in the terminal sequence between 5' and 3'

Figure 1. Terminal nucleotide sequence of total macronuclear DNA. A. 5' terminal sequence of *Oxytricha fallax* macronuclear DNA; 20% polyacrylamide sequencing gel. Cleavage products electrophoresed on a sequencing gel with a ³²P_i marker gave a ladder of bands corresponding to the sequence C₄A₄C₄A₄C₄ at the 5' ends. Arrows indicate the abrupt transition between the common terminal sequence and the heterogeneous bulk macronuclear DNA. B. 3' terminal sequence of *Oxytricha fallax* macronuclear DNA; 20% polyacrylamide sequencing gel. Cleavage products electrophoresed on a sequencing gel gave a ladder of bands corresponding to the sequence G₄T₄G₄T₄G₄T₄G₄ at the 3' ends. Arrows indicate the abrupt transition between the common terminal sequence and the heterogeneous bulk macronuclear DNA.

ends implies that 16 bases in the terminal regions of macronuclear DNA molecules are not base-paired.

The most probable structures to accommodate these 16 unpaired bases are a protruding single-stranded 3' tail or an internal single-stranded loop. To distinguish between these two possibilities, total macronuclear DNA labeled at the 3' ends was digested with the single strand-specific enzyme, S1 nuclease. Reaction conditions were adjusted to limit enzymatic cutting to approximately one cut per molecule and to reduce molecular breathing at the ends of the molecules. When the products of this limited S1 digest were separated on a polyacrylamide sequencing gel, they formed a ladder of bands corresponding to the nucleotides susceptible to digestion. Figure 2 shows the resulting ladder of bands which correspond to the first 16 nucleotides at each 3' end of macronuclear DNA molecules consistent with the existence of protruding 3' single-stranded tails, rather than an internal loop.

DISCUSSION

Based on results from chemical sequence analyses, we have demonstrated that the ends of macronuclear DNA from Oxytricha fallax share a common terminal inverted repeat. The essentially homogeneous sequence for the terminal repeat consists of an alternating string of 4C's and 4A's at the 5' ends, and the complementary 4G's and 4T's at the 3' ends. S1 nuclease mapping of the 3' ends of total macronuclear DNA has established that the 16 extra nucleotides detected at the 3' ends exist as single-stranded tails. Our findings, summarized in Figure 3, describe the generalized structure of the ends of extensively purified macronuclear DNA. Although this interpretation is valid for the situation in vitro, we recognize that it is not necessarily representative of the structure of these molecules in vivo, as will be discussed further.

Chemical sequence analysis of the cloned 1.6 kb macronuclear DNA molecule coding for the protein actin produced similar, though non-identical, results concerning terminal sequence organization (8). Among the differences observed was the absence of four terminal G residues at one end of the insert. Although this difference could represent an artifact produced during the cloning process, the possibility exists that the length of the common terminal sequence is heterogeneous among macronuclear DNA molecules. It appears, in fact, that the terminal self-complementary sequence in the actin molecule actually extends inward an additional 2 base pairs from the common repeat shown for total macronuclear DNA. Furthermore, the sequence 3' AACTT

Figure 2. Limited S1 digestion of 3' end-labeled macronuclear DNA; 20% polyacrylamide sequencing gel. Digestion products from two independent experiments were electrophoresed on sequencing gels and produced a ladder of bands corresponding to the nucleotides susceptible to the single-strand specific enzyme. When compared to an adjacent lane of 3' G cleavage products, S1 products indicate that a single-stranded tail of 16 nucleotides exists at 3' ends of macronuclear DNA molecules.

Figure 3. Terminal organization of *Oxytricha fallax* macronuclear DNA molecules.

5', which appears in the terminal repeat of *Euplotes aediculatus*, a related ciliate, at a position 18 nucleotides from the 3' ends, was found at the downstream end of the *O. fallax* actin sequence, at a position 22 nucleotides from the 3' end.

Our results, along with the sequencing data of Klobutcher *et al.* (11), allow us to eliminate several proposals concerning the nature of the ends of macronuclear DNA in *O. fallax*. S1 mapping of labeled 3' ends confirms that no internal single-stranded loops exist at the ends of macronuclear DNA molecules. Although we cannot eliminate the possibility that internal single-stranded tails exist, as have been reported for cauliflower mosaic virus (12), this seems unlikely as no such sequences were detected within the cloned macronuclear DNA coding for actin. Furthermore, we wish to point out that although the generalized structure depicted in Figure 3 features two single-stranded tails for each macronuclear DNA molecule, a fraction of the macronuclear DNA population could contain molecules having one or no 3' tail, and would be indistinguishable from bipolar molecules in our S1 assay.

The S1 sensitivity demonstrated for the 3' ends of macronuclear DNA clearly indicates the single-stranded nature of the 3' ends of purified molecules *in vitro*. The possibility exists, however, that a small terminal base-paired RNA or DNA might fill in the terminal gap *in vivo*, but because of its size is selectively lost during RNase digestion and subsequent purification of total macronuclear DNA. Recently, Dawson and Herrick (13) have identified a class of single-stranded DNA fragments 22 to 11 bases long that segregate from bulk *O. fallax* macronuclear DNA on denaturing gels, and have suggested that such fragments may be associated with macronuclear tails *in vivo*. This would not be inconsistent with our results.

The terminal sequence of *O. fallax* DNA is similar to the repeating terminal C_4A_2 sequence of the rDNA from the ciliate *Tetrahymena thermophila*. However, we have evidence to suggest that *O. fallax* and *T. thermophila* have different terminal organizations. Based on results from restriction analysis and *in vitro* DNA synthesis in *Tetrahymena*, Blackburn and Gall (15) proposed a terminal organization consisting of a hairin loop interrupted by single-nucleotide gaps. Several lines of evidence indicate that this structure does

not exist at the ends of Oxytricha macronuclear DNA. First, our sequencing data preclude the possibility of any intramolecular base-pairing of the single-stranded 3' end with itself, which would necessitate pairing between G's and T's, or in the actin gene, with an internal complementary sequence. Second, if a Tetrahymena-like hairpin "cap" arrangement were in fact the case for Oxytricha macronuclear DNA termini, limited S1 digestion of O. fallax DNA should yield a lack of S1 sensitivity and hence, no ladder of nucleotide bands upon polyacrylamide gel electrophoresis. Furthermore, a small palindromic polynucleotide fragment of homogeneous length that would correspond to the Tetrahymena hairpin cap would appear. Although no such fragment has been identified on any of our gels, it is possible that such a fragment could be lost during purification of the DNA if it were both weakly base-paired to the rest of the molecule and very small. Alternatively, if such a fragment did exist and it was very large (≥ 100 nucleotides), we could not detect it on a gel of S1 digestion products. Nevertheless, we can rule out this possibility because chemical sequence analysis of total macronuclear DNA produces a coherent homogeneous nucleotide sequence. If a non-covalently attached cap was present, its sequence would be superimposed on that of bulk macronuclear DNA, resulting in a non-homogeneous ladder of bands.

The terminal organization of Oxytricha fallax macronuclear DNA reported here is unique with regard to other linear eukaryotic and viral genomes. Only very closely related species of hypotrichs share a common terminal nucleotide sequence and presumably, a 3' single-stranded tail with O. fallax (9,11). Indeed, the hairpin cap arrangement reported for Tetrahymena is significantly dissimilar to our findings (15). An alternative arrangement of terminal DNA is represented by the covalently crosslinked termini in yeast chromosomes (16) and in the mitochondrial DNA of Paramecium aurelia (17,18). The presence of identical single-stranded tails in O. fallax prohibits terminal crosslinking of macronuclear DNA by normal base-pairing. More extensive characterizations of linear viral DNAs have revealed long highly-conserved terminal palindromes in the parvoviruses (19,20) and a covalent association between a terminal unpaired 5' deoxycytidine residue and a 55,000 dalton protein in the adenoviruses (21,22). The termini of O. fallax macronuclear DNA do not contain any identifiable palindromic sequences characteristic of the parvoviruses, but we have no other basis for comparison of the O. fallax terminal structure with that of the adenoviruses.

The uniqueness of the ends of Oxytricha fallax macronuclear DNA is of considerable interest. Yet perhaps even more intriguing is the homogeneity

of the macronuclear DNA population with regard to this terminal organization. Since it appears that most or all DNA molecules in the macronucleus have identical terminal structures, we can speculate on several possible functions. During macronuclear development, the chromosomes of the micronucleus undergo polytenization and subsequent fragmentation. After a selective degradation of this DNA, the conserved "gene-sized" pieces undergo 1000-fold replication to form the new macronuclear genome (1-4). Therefore, it is possible that the inverted terminal repeat provides a recognition sequence for macronuclear sequences to be "rescued" from the micronuclear genome. Recently, however, it has been reported that in Oxytricha nova, micronuclear sequences that are retained throughout macronuclear development lack the inverted terminal repeats of their macronuclear counterparts, suggesting that the terminal repeats are added after excision of macronuclear sequences from the micronuclear genome (23). Alternatively, the specific terminal organization of O. fallax macronuclear DNA may be intimately involved in the enigmatic replication of the termini of these molecules, or may represent a protein binding site. Finally, it has been reported that macronuclear terminal repeats in the hypotrichs Stylonichia mytilus and Oxytricha nova are involved in circle or concatemer formation in vitro to yield long, multistranded aggregates (24,25). Because the single-stranded 3' tails contain a sequence that cannot readily base-pair with itself or with an identical sequence in another molecule, we feel that such an interpretation is unlikely for Oxytricha fallax.

ACKNOWLEDGEMENTS

We wish to thank M. Dolan and J.D. Engel for their patient instruction in the fine art of sequencing. This work was supported by NIH grant GM23440 to B.B.S.

*Fred Hutchinson Cancer Research Center, Seattle, WA 98104, USA

+Department of Genetics and Development, University of Illinois, Urbana, IL 61801, USA

§Department of Molecular Biology 93-D, Abbott Laboratories North Chicago, IL 60064, USA (To whom reprint requests should be addressed)

REFERENCES

1. Ammermann, D. (1971) *Chromosoma* 33, 209.
2. Lauth, M.R., Spear, B.B., Heumann, J. and Prescott, D.M. (1976) *Cell* 7, 67.
3. Prescott, D.M. and Murti, K.G. (1973) Cold Spring Harbor Symp. Quant. Biol. 38, 609.
4. Spear, B.B. and Lauth, M.R. (1976) *Chromosoma* 54, 1.

5. Rae, P.M.M. and Spear, B.B. (1978) *Proc. Natl. Acad. Sci. USA* 75, 4992.
6. Spear, B.B. (1980) *Chromosoma* 77, 193.
7. Wesley, R.D. (1975) *Proc. Natl. Acad. Sci. USA* 72, 678.
8. Kaine, B.P. and Spear, B.B. (1980) *Proc. Natl. Acad. Sci. USA* 77, 5336.
9. Oka, Y., Shiota, S., Nakai, S., Nishida, Y. and Okubo, S. (1980) *Gene* 10, 301.
10. Lawn, R.M. (1977) *Proc. Natl. Acad. Sci. USA* 74, 4325.
11. Klobutcher, L.A., Swanton, M.T., Donini, P. and Prescott, D.M. (1981) *Proc. Natl. Acad. Sci. USA* 78, 3015.
12. Franck, A., Guilley, H., Jonard, G., Richards, K. and Hirth, L. (1980) *Cell* 21, 285.
13. Dawson, D. and Herrick, G. (1982) *Nucl. Acid Res.* 10, 2911.
14. Maxam, A.M. and Gilbert, W. (1980) *Methods Enzymol.* 65, 499.
15. Escarmis, C. and Salas, M. (1981) *Proc. Natl. Acad. Sci. USA* 78, 1446.
16. Ohmori, H., Tomizawa, J. and Maxam, A.M. (1978) *Nucleic Acids Res.* 5, 1479.
17. Blackburn, E.H. and Gall, J.G. (1978) *J. Mol. Biol.* 120, 33.
18. Forte, M.A. and Fangman, W.L. (1979) *Chromosoma* 72, 131.
19. Goddard, J.M. and Cummings, D.J. (1977) *J. Mol. Biol.* 109, 327.
20. Pritchard, A.R., Herron, L.M. and Cummings, D.J. (1980) *Gene* 11, 43.
21. Berns, K.I., Hauswirth, W.W., Fife, K.H. and Lusby, E. (1979) *Cold Spring Harbor Symp. Quant. Biol.* 43, 781.
22. Astell, C.R., Smith, M., Chow, M.B. and Ward, D.C. (1979) *Cell* 17, 691.
23. Rekosh, D.M.K., Russell, W.C., Bellett, A.J.D. and Robinson, A.J. (1977) *Cell* 11, 283.
24. Robinson, A.J., Younghusband, H.B. and Bellett, A.J.D. (1973) *Virology* 56, 54.
25. Boswell, R.E., Klobutcher, L.A. and Prescott, D.M. (1982) *Proc. Natl. Acad. Sci. USA* 79, 3255.
26. Lipps, H.J. (1980) *Proc. Natl. Acad. Sci. USA* 77, 4104.
27. Lipps, H.J. Grussem, W. and Prescott, D.M. (1982) *Proc. Natl. Acad. Sci. USA* 79, 2495.