Strong sequence conservation of a 38 bp region near the center of the extrachromosomal rDNA palindrome in different Tetrahymena species

J. Engberg

Department of Biochemistry B, The Panum Institute, University of Copenhagen, Blegdamsvej 3, DK-2200 Copenhagen N, Denmark

Received 18 May 1983; Revised and Accepted 30 June 1983

ABSTRACT

The restriction-endonuclease map and the nucleotide sequence of the central region in the extrachromosomal rDNA palindrome of two micronuclear and one a-micronucleate species of Tetrahymena has been determined. The sequence data show that the different species investigated have a 24 or 26 nucleotide sequence region at the very center of the rDNA molecule which is non-palindromic. Comparison of the present sequence data with the published data of another micronucleate species reveal that a segment of 38 base pairs just outside the non-palindromic center is highly conserved in all the different species, while the rest of the central region show little sequence homology. The relevance of this conserved region to the amplification process of the rDNA molecule is discussed.

INTRODUCTION

Two types of nuclei exist in the unicellular micronucleate species of the ciliated protozoan, Tetrahymena; a polyploid (somatic) macronucleus which is responsible for the transcriptional activity during vegetative growth and a diploid (germinal) micronucleus which maintains the genetic continuity of the organism (reviewed in ref. 19). During the sexual cycle (conjugation) the micronucleus undergoes several processes including meiosis and mitosis to give rise to both the micronucleus and the macronucleus of the offspring. In contrast, certain Tetrahymena species exist which do not contain a micronucleus. These a-micronucleate species can only reproduce vegetatively. In the macronucleus of both micronucleate and a-micronucleate species of Tetrahymena the genes coding for ribosomal RNA (rRNA) exist in multiple copies as extrachromosomal, linear molecules of about 20 kb with two transcription units for the rRNA precursor arranged as palindromes (1-3). In the micronucleus, however, the rDNA is integrated in the chromosome and exists as a single copy of approximately half the size of the palindromic rDNA molecule (4). Therefore, in the case of the micronucleate species the macronuclear rDNA palindromes are generated from the integrated rDNA copy by
an excision and amplification process during macronuclear development. By comparison of the nucleotide sequence of the central and terminal regions of the rDNA palindrome with those of the junction regions of the integrated rDNA it might be possible to obtain information about the molecular mechanisms involved in the generation of extrachromosomal rDNA palindromes. We report here sequence data of the central region of the rDNA palindrome from a number of species. Comparison of our sequence data with that previously published (12,13) suggests that the extrachromosomal rDNA from all species of Tetrahymena contain a non-palindromic region at the center of the molecule flanked by a short sequence highly conserved in all species. These sequence comparisons should be important in helping to understand the mechanism by which extrachromosomal rDNA is generated.

MATERIALS AND METHODS

*T. pyriformis* (GL-C), *T. pigmentosa* (UM 1286) and *T. hyperangularis* (10 IEN) were originally provided by Prof. E. Zeuthen of the Carlsberg Biological Institute, Copenhagen and Dr. E. Simon of the University of Illinois at Urbana-Champaign, respectively. The strains were cultivated axenically in a complex proteose peptone, yeast extract medium at 28°C as previously described (8). Macronuclear rDNA was isolated by the previously described selective, hot-phenol extraction procedure (2) except that 5 volumes of the phenol-cresol mixture was used instead of 10 volumes. Cells were grown in batches of 3 l and about 30 l of each strain were used in the present study which yielded about 200 μg of pure rDNA from each strain. Restriction enzymes were purchased from Boehringer-Mannheim, New England Biolabs and Bethesda Research Laboratories and used according to the suppliers' specifications. Electrophoretic fractionation of DNA fragments on agarose or polyacrylamide gels and subsequent visualization with ultraviolet light was done according to standard procedure as described previously (1). For DNA sequence analysis, DNA restriction fragments with recessed 3'-ends were labeled with 5'-[α-32P] deoxyynucleosidetriphosphates (NEN) using the large fragment of *E. coli* DNA pol I (kindly provided by Prof. H. Klenow), protruding 3'-ends and fragments with flush ends were labeled with 5'-[α-32P] cordycepintriphosphate (NEN) using terminal deoxynucleotidyltransferase (NEN) while protruding 5'-ends were labeled with 5'-[γ-32P] ATP (NEN) using polynucleotide kinase (Boehringer-Mannheim). Labeled fragments were isolated from polyacrylamide gels following autoradiography by cutting out the appropriate bands, crushing the gel pieces with a spatula and eluting the DNA in 2 volumes of
25 mM Tris-HCl, pH 8.5, 250 mM NaCl and 25 mM EDTA overnight at 37°C followed by phenol extractions and ethanol precipitations. The complementary strands of the central fragments were separated by electrophoresis on 20% urea/polyacrylamide gels at room temperature. DNA sequence analysis was done by the method of Maxam and Gilbert (9).

RESULTS

The restriction maps of rDNA isolated from *T. pyriformis*, *T. pigmentosa* and *T. hyperangularis* have been determined for a number of restriction enzymes some of which is depicted in Fig. 1A. The maps were deduced from agarose gel analysis of the digestion products generated by different combinations of double digestions as described by us (2) and by others (7,6,11).

Eco RI digestions of *T. pigmentosa* and *T. hyperangularis* gave rise to a conveniently small sized central fragment which could be end-labeled and used for subsequent restriction mapping analysis and sequencing. With respect to *T. pyriformis* rDNA, restriction sites close to the center of the molecule were investigated by the snap-back test: The isolated central Hind III fragment was cut with different restriction enzymes followed by labeling of the newly generated ends with 5'-[γ-32P]ATP and polynucleotide kinase. Aliquots of the fragment mixture were heat-denatured and quickly cooled followed by polyacrylamide gel analysis in parallel with a non-denatured sample. The fragment which can snap back on itself is converted to half of its original size and is thereby identified as the fragment in the mixture containing the center of the palindrome (1). These analyses revealed the sites indicated in Fig. 1B. The sites distal to the central Kpn I site have been verified by E. Niles and T. Higashinakagawa (7,10). The central Tag I sites in *T. pigmentosa* and *T. hyperangularis* were found in analogous experiments starting with the isolated central EcoRI fragments. In the case of *T. hyperangularis* the about 300 bp large central fragment generated by EcoRI digestion proved to be a mixture of two very similar sized molecular species one of which could snap back following the heat denaturation treatment while the other could not.

A heat denaturation step was therefore included in the preparative isolation of the central EcoRI fragment to insure its purity. The map position of the contaminating EcoRI fragment from *T. hyperangularis* on the rDNA molecule has not been determined accurately but DNA-DNA hybridization experiments using this DNA as probe against Southern blots of Bgl II, Bam HI and Hind III digested rDNA showed homology with the distal regions of the rDNA molecule (results not shown). Sequence analysis of the central rDNA fragments from the
Figure 1. Physical maps of the rDNA palindromes of T. pyriformis, T. pigmentosa and T. hyperangularis. (A) A collection of restriction enzyme cleavage maps are shown. Restriction enzymes used in (A) and (B): BamHI (B), BglII (Bg), EcoRI (R), HaeIII (Hae), HindIII (H), HphI (Hph), KpnI (K), Sau3A (S), TaqI (T) and XhoI (X). The vertical axis indicate the center of the palindromes. The thick horizontal arrows indicate the transcriptional region of pre-rRNA (2,7,10). (B) Selected restriction sites close to the center of the rDNA palindromes are shown. The arrows indicate the strategy used in determination of the nucleotide sequence. Each set of DNA sequencing reactions was repeated several times.

different strains was initiated assuming that the molecules were perfect palindromes so that the sequencing reactions could be performed on the central DNA fragments labeled at either end (3'- or 5'-end). This approach yielded unequivocal sequence determinations until a certain position was reached close to the center of the molecule. From thereon overlapping bands were observed in the gel sequence ladder corresponding to the different base specific cleavage reactions which indicated that we were not dealing with
Figure 2. Comparison of the nucleotide sequence in the central rDNA region of T. pyriformis, T. pigmentosa, T. hyperangularis and T. thermophila. The nucleotide sequence of T. thermophila was taken from Kiss and Pearlman (12) and the sequence of another T. pigmentosa strain was determined by Kan and Gall (13). The sequences represent the slow moving bands on strand-separating gels. Arrows indicate the center of the palindromes, while the symbols V, V and A mark the non-palindromic regions of T. pigmentosa/T. hyperangularis, T. pyriformis and T. thermophila, respectively. Space in T. thermophila is inserted in order to demonstrate maximum homology. Nucleotide residues are numbered in the direction from 5' to 3'.
perfect palindromes. This was further indicated by the observation that the complementary strands of the central fragments could be separated on a strand separating gel. Thus, the remaining DNA sequence analysis was performed on the separated strands of the central HaeIII fragment (T. pyriformis) and the central Taq I fragments (T. pigmentosa and T. hyperangularis) labeled at their 3'-ends. Both strands showed that 24 nucleotides (T. pyriformis) or 26 nucleotides (T. pigmentosa and T. hyperangularis) surrounding the center were not palindromic. The sequence strategy used is indicated in Fig. 1B and the actual sequences are shown in Fig. 2 along with the T. thermophila sequence which was taken from ref. 12. The sequences indicated represent the sequences of the slow moving bands on the strand separating gels.

DISCUSSION

The sequence of the central rDNA region of the T. pigmentosa strain used in this study (UM 1286) appears to be identical to that of another T. pigmentosa strain (ALP 15) which was sequenced by Kan and Gall (13). Furthermore, as indicated in Fig. 2, the T. pigmentosa sequence is very homologous to that of T. hyperangularis: only 3 differences occur in 141 nucleotides (at position 118, 129 and 174 in the numbering of Fig. 2). However, when the nucleotide sequence of T. pigmentosa is compared to that of T. thermophila or T. pyriformis very little homology is observed except in a region close to the center of the molecule just outside of the non-palindromic region. In this region a stretch of 38 nucleotides of almost perfect homology is observed when sequences from all the species (micronucleate as well as a-micronucleate) are compared. On this basis we suggest that this structural feature may be common to all Tetrahymena species. It is unlikely that the conserved sequence is involved in replication function since the origin of replication has been localized outside the conserved sequences in different Tetrahymena species. In T. thermophila the origin was shown to be about 600 bp and in T. pyriformis less than 200 bp from the center of the rDNA molecule (5,6). Using a computer program to search for sequence homologies, we have compared the sequence of the central 368 bp Sau 3A fragment of the T. pyriformis rDNA which may include the origin of replication (5,6) with the sequence of T. thermophila rDNA 500 to 1000 bp from the center of the molecule (R.E. Pearlman, personal communication). This region in T. thermophila includes the in vivo origin of replication (6) and also functions as an autonomously replicating sequence (ars) in yeast (14). No obvious homologies other than short homooligomeric stretches were observed. Similarly, comparisons of
DNA sequences known to contain replication origins such as the 2 μm yeast circle and the ars1 element (15,16) to those of the Sau3A fragment from T. pyriformis rDNA did not reveal any striking homologies. It appears that the regions confining the sequences involved in initiation of replication have to be defined more precisely before sequence comparisons become meaningful. Inspection of the three upper sequences of Fig. 2 at the non-palindromic/palindromic junction reveals that the region of homology extends further towards the center than indicated by the boxed region which was drawn to show maximum homology between all of the Tetrahymena sequences. Thus, it appears that the size of the conserved region may vary slightly but that it is located right next to the non-palindromic region in all cases. This situation resembles very much the overall genome organization found in palindromic rho- mitochondrial DNAs of yeast where recent experiments have shown that the palindromes were not symmetrical right up to the junction points but that the junctions consist of an asymmetrical sequence of variable length with inverted repeats at both ends (17). On the basis of these structural observations a model was proposed to explain the generation of the palindromes in which unequal excision of complementary strands was involved. A related model was proposed originally by Yao and Gall (18) to explain the generation of the extrachromosomal rDNA molecules of Tetrahymena from the single, non-repetitive rDNA copy of the germinal chromosomes. The sequence data from the present work allows the precise prediction that two copies of the conserved region or parts of it should be present as inverted repeats at the chromosomal junction if the model is correct. Work is in progress in our laboratory to test this prediction. The present sequence data show that the presence of a non-palindromic junction region as well as a conserved region adjacent to it are common features of both micronuclear and α-micronuclear strains of Tetrahymena. Most probably, these similarities reflect that the extrachromosomal rDNA molecules arose in a similar manner in the two cases. Since the α-micronuclear strains do not undergo the conjugal amplification of their rDNA the occurrence of extrachromosomal rDNA palindromes in T. pyriformis could be explained by assuming that this strain represents a former micronuclear strain which once lost its micronucleus. This could happen as a result of an impaired conjugal process or as a spontaneous event as has been reported for laboratory strains during vegetative cultivation (19). Considering that the sequences outside the conserved region (Fig. 2) as well as the overall restriction enzyme map of T. pyriformis rDNA have diverged considerably relative to all the known micronucleate Tetrahymena species (2,7),
it becomes unlikely that T. pyriformis derived recently from one of the known micronucleate species by means of losing a micronucleus. More likely, T. pyriformis represents a separate Tetrahymena species whose micronuclear counterpart has not yet been found in nature.

ACKNOWLEDGEMENTS

I am indebted to Dr. R.E. Pearlman, York University, Toronto, for communicating sequence data prior to publication; to Dr. F.G. Hansen, The Technical University of Denmark, for teaching me the Maxam-Gilbert technique; to Dr. M. Johnsen, Institute of Microbiology, for running the computer analysis; to Dr. B. Bonvén, Institute of Molecular Biology, University of Aarhus, Denmark, for providing samples of T. pyriformis rDNA during the initial part of this work; to Mr. F. Frenzel, for technical assistance and to Ms. E. Hofyer for typing the manuscript.

This work was supported by the Danish Research Council for Natural Sciences.

REFERENCES