MOLECULAR AND CELLULAR BIOLOGY, Aug. 1991, p. 3823–3834
0270-7306/91/083823-12$02.00/0

Vol. 11, No. 8

Chromosome Structure: DNA Nucleotide Sequence Elements of a Subset of the Minichromosomes of the Protozoan *Trypanosoma brucei*

MICHAEL WEIDEN,[1] YVONNE N. OSHEIM,[2] ANN L. BEYER,[2] AND LEX H. T. VAN DER PLOEG[3]*

Department of Medicine[1] and Department of Genetics and Development,[3] College of Physicians and Surgeons, Columbia University, 701 West 168 Street, New York, New York 10032, and Department of Microbiology, University of Virginia, Charlottesville, Virginia 22908[2]

Received 18 December 1990/Accepted 30 April 1991

The genome of the protozoan *Trypanosoma brucei* contains a set of about 100 minichromosomes of about 50 to 150 kb in size. The small size of these chromosomes, their involvement in antigenic variation, and their mitotic stability make them ideal candidates for a structural analysis of protozoan chromosomes and their telomeres. We show that a subset of the minichromosomes is composed predominantly of simple-sequence DNA, with over 90% of the length of the minichromosome consisting of a tandem array of 177-bp repeats, indicating that these molecules have limited protein-coding capacity. Proceeding from the tip of the telomere to a chromosome internal position, a subset of the minichromosomes contained the GGGTTA telomere repeat, a 29-bp telomere-derived repeat, a region containing 74-bp G+C-rich direct repeats separated by approximately 155 bp of A+T-rich DNA that has a bent character, and 50 to 150 kb of the 177-bp repeat. Several of the minichromosome-derived telomeres did not encode protein-coding genes, indicating that the repertoire of telomeric variant cell surface glycoprotein genes is restricted to some telomeres only. The telomere organization in trypanosomes shares striking similarities to the organization of telomeres and subtelomeres in humans, yeasts, and plasmodia. An electron microscopic analysis of the minichromosomes showed that they are linear molecules without abnormal structures in the main body of the chromosome. The structure of replicating molecules indicated that minichromosomes probably have a single bidirectional origin of replication located in the body of the chromosome. We propose a model for the structure of the trypanosome minichromosomes.

Trypanosoma brucei is a protozoan parasite that lives in the bloodstream of its mammalian host. Trypanosomes can escape destruction by the humoral immune response because they are capable of periodically switching to the expression of an entirely new cell surface protein coat made of a single type of protein, the variant cell surface glycoprotein (VSG). This process, called antigenic variation, relies on a large repertoire of over 1,000 different VSG genes, only one of which is expressed in a trypanosome at any one time. The expressed VSG gene is invariably located at the very end of a chromosome, also referred to as the telomere, and telomeres containing expressed VSG genes have been identified on a few specific chromosomes only. These telomeres are referred to as VSG gene expression sites. Switching to the expression of a new VSG gene can be accomplished by different types of DNA recombinational events which can translocate a VSG gene to an expression site. Alternatively, differential transcriptional control of the expression sites can result in an antigenic switch (for reviews, see references 5, 13, 32, 33, and 41).

The genome of *T. brucei* contains roughly 20 large chromosomes, ranging in size from 200 kb to 5.7 Mb, and about 100 50- to 150-kb minichromosomes (MC) (21, 44, 45). Seven pairs of homologous chromosomes have been identified among the 20 larger chromosomes, while the genome has been proposed to be aneuploid for the MC (21). The chromosomes all share terminally located telomere repeats with the sequence $(GGGTTA)_n$ (4, 43), recently also identified in humans (9a, 28; for a review see reference 3). The MC comprise about 10% of the nuclear DNA but represent over 80% of the chromosomes of *T. brucei*. The MC are mitotically stable (44); however, their number may vary during genetic exchange (49).

The MC are believed to harbor a large number of telomeric VSG genes. The only protozoa identified thus far that contain MC are trypanosomes that undergo antigenic variation; hence, their only function may be to provide a set of telomerically located VSG genes (42). However, MC do not appear to be able to express VSG genes, and the MC-derived VSG genes need to be translocated to an expression site for transcription (15, 41, 42). The highly recombinogenic nature of trypanosome telomeres (1, 43) may lead to a more rapid accumulation of changes in the telomeric VSG genes compared with chromosome internally located VSG genes. A rapidly evolving group of VSG genes will have epidemiological implications and intuitively appears to be advantageous to the parasite. The small size of these molecules provides an opportunity to define the structure of a naturally occurring chromosome at the nucleotide level, thereby giving a detailed overview of the sequences required for chromosome functioning in trypanosomes. Our goal was to improve our understanding of the genetic elements that are necessary and sufficient for eukaryotic chromosome function.

MATERIALS AND METHODS

DNA isolation, enzyme digestion, and Southern blotting. *T. brucei* clone 118d (for Fig. 6B and C, 118 clone 1 was used) was grown in rats and purified from blood elements as described by Fairlamb et al. (20). Nuclear DNA was isolated from trypanosomes as described by Van der Ploeg et al. (46)

* Corresponding author.

and digested by adding a fourfold excess of restriction endonuclease for complete digestion; 10% of the recommended amount of enzyme was added for partial restriction enzyme digestions. For the partial restriction enzyme digestions, aliquots were taken at different times and were immediately phenol-chloroform-isoamyl alcohol extracted. Bal 31 digestions were similarly performed in a time series. The actual reduction in size of the fragments was calculated from rehybridization of the same Southern blots containing the Bal 31-treated DNA with a 177-bp repeat and dividing the observed reduction in size of small bands by 2 to give base pairs digested per end. Counts per minute on the filter were measured directly on a Betascope 603 (Betagen, Waltham, Mass.) Plasmids were isolated as described by Birnboim and Doly (2). Agarose gels were treated with 0.25% HCl prior to blotting. DNA was transfered to Hybond-N and hybridized with random-primed ^{32}P-labeled DNA. Plasmid DNA and size markers (1-kb ladder or phage lambda DNA digested with *Hin*dIII) were end labeled with [^{32}P]ATP and polynucleotide kinase. Final stringencies of posthybridization washing were 0.1× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate), for A2, 0.3× SSC for the 177-bp repeat, and 1.0× SSC for A4.

Sucrose gradient isolation of MC. A modification of the method published by Sloof et al. (39) was used. A 35-ml linear 5 to 20% sucrose gradient in 100 mM EDTA–25 mM Tris-HCl (pH 7.5)–1% sodium lauryl sarcosinate (SLS)–50 μg of proteinase K per ml was overlayered with 2 ml of a solution containing 200 mM EDTA, 1% SLS, and 1 mg of proteinase K per ml. On top of this solution we layered 4×10^8 trypanosomes in 1 ml of phosphate-buffered saline. The trypanosomes were allowed to lyse by incubating the gradient at 25°C for 3 h, after which the DNA was size separated by spinning at 10,000 rpm in an SW28 rotor for 16 h at 25°C; 1-ml fractions were removed from the top of the gradient with a sterile 1-ml syringe (omitting the needle). The fractions were assayed directly for MC content on pulsed-field gel electrophoresis (PFG) gels. Pure MC were usually obtained in fractions 6 to 10 of these gradients. The DNA was precipitated with 2 volumes of ethanol, resuspended in 0.5% of the initial volume in 10 mM Tris-HCl (pH 7.5), treated with RNase A (5 μg/ml) at 37°C for 15 min, and subjected to a second ethanol precipitation. Recovery of MC DNA was usually 2 to 4 μg per gradient.

Identification of bent DNA. Conditions for separation of bent DNA were as described by Marini et al. (17, 27, 50). The bending predictions were done on the DNA Star bend program (DNA Star, Madison, Wis.).

PFG. PFG gels were run as described by Van der Ploeg et al. (44, 45) with specific conditions as stated in the figure legends. The simple-sequence DNA shown in Fig. 6 was isolated by two-step purification. MC were cut from low-melting-point agarose PFG gels, and the blocks were washed in Tris-EDTA for 24 h. After washing of the blocks, the DNA was digested in the gel with a 100-fold excess of *Hin*fI for 12 h. Following digestion, the DNA was electrophoresed in a 0.7% regular low-melting-point gel, after which the high-molecular-weight DNA (over 30 kb) was isolated by melting the agarose at 65°C. The DNA was phenol extracted and then concentrated by ethanol precipitation. This DNA either was digested with *Hin*pI and end labeled with [^{32}P]dCTP and Klenow polymerase or was digested with *Alu*I, Southern blotted, and hybridized with random primer-labeled, purified MC.

DNA nucleotide sequence analysis. DNA nucleotide sequence analysis was performed by the dideoxy-chain termination method (36), using either Sequenase at 25°C with [^{35}S]dATP or *Taq* polymerase at 65°C with dye-labeled primers. Subclones of PT4 were inserted into the replicative forms of bacteriophage M13 or the Bluescript vector (Stratagene); both single-stranded and double-stranded sequencing were performed.

EM. For electron microscopy (EM), the DNA molecules were visualized with the drop diffusion method of Lang and Mitani (25). Forty-microliter drops of the following mixture were deposited on a clean plastic petri dish and allowed to diffuse for 20 to 60 min: 20 μl of DNA at 1 μg/ml; 15 μl of 0.01 M Tris–0.001 M EDTA (pH 8.5); 160 μl of 0.25 M ammonium acetate; and 5 μl of cytochrome *c* (1 mg/ml; Sigma type V, filtered) diluted 1:20 in distilled H_2O. Carbon-coated grids (300 mesh) were touched to the surface of each drop, washed in distilled H_2O, and stained for 1 min in uranyl acetate solution made by diluting a stock of 0.05 M uranyl acetate in 0.05 N HCl, 1:100 in 90% ethanol. Grids were then rinsed in 95% ethanol, air dried, and rotary shadow cast with platinum. The replicating molecule shown in Fig. 8 was prepared as described above, with the addition of NaCl to give a final concentration of 6 mM. Grids were viewed in a JEOL 100 CS transmission electron microscope at 80 kV.

RESULTS

Sequence analysis of the telomere domain. In previous studies, we isolated two telomere-derived clones, PT1 and PT4 (43). The physical maps of both clones are presented in Fig. 1A. One of these telomere clones (PT4) contained a tandem array of the GGGTTA telomere repeat (44). To identify the genetic elements that comprised a telomeric and subtelomeric domain, we determined the DNA nucleotide sequence of clone PT4 (Fig. 1B) and the ends of clone PT1 (43; data not shown). The nucleotide sequence of clone PT4 contains, from a telomeric position toward a chromosome internal location, 55 copies of the telomere GGGTTA repeat followed by six copies of a 29-bp GGGTTA-derived repeat (Fig. 1). The same 29-bp repeat was also present at the end of clone PT1. The significance of this 29-bp repeat is unclear, but 29-bp repeats of a different nucleotide sequence have also been found in human subtelomere repeat domains (16), and interestingly, the minimum size of a sequence which can support recombinational extension of telomeres in yeast cells is 28 bp (34).

Internal to the 29-bp subtelomere repeat are several stretches, approximately 155 bp long, which are composed of 74% A+T with many short runs of poly(dT). These A+T-rich repeats share little sequence similarity with each other. A 74-bp G+C-rich (65% G+C) domain is found in between the 155-bp A+T-rich repeats. There are three examples of this G+C-rich repeat in clone PT4 which are 80% similar to each other (numbered 1 to 3 in Fig. 1C). The 74-bp domain has a core of 17 identical nucleotides (Fig. 1C, boxed sequence). This 17-bp element is directly repeated and contains an *Ava*II site which we have used to verify the presence of this 17-bp domain on other telomeres (see below). This 17-bp sequence also contains 11 bp of the GGGTTA telomere repeat positioned in inverted orientation relative to the GGGTTA repeat at the end of the chromosome. Eleven base pairs of the telomere repeat is equivalent to one turn of the A-form double helix and allows a full GGGTTA repeat to be formed independent of the starting nucleotide. The organization of this subtelomere is reminiscent of the Y and X elements of *Saccharomyces cerevisiae*

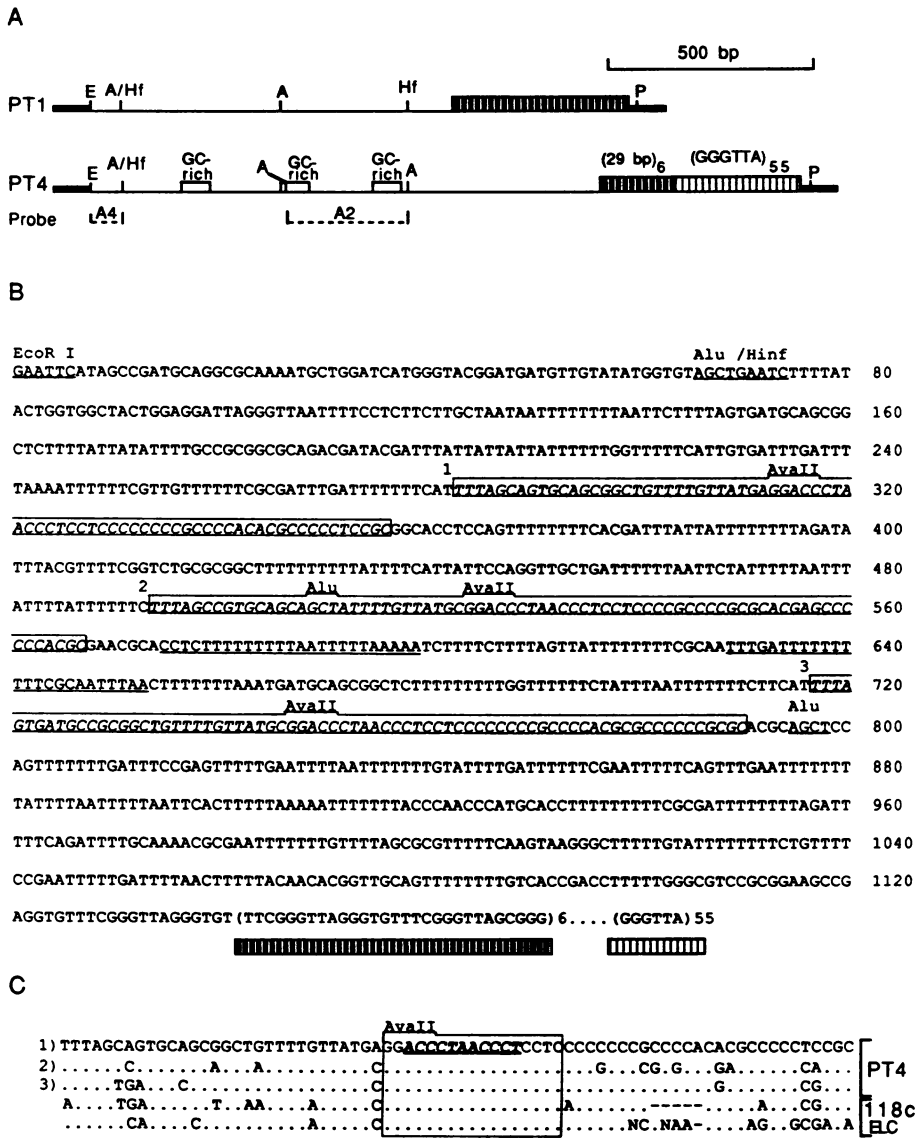

FIG. 1. (A) Restriction maps of telomere clones PT1 and PT4. The clones are oriented with the telomere sequence to the right. The black hatched rectangle represents the subtelomere 29-bp repeat, the open hatched rectangle represents the telomere repeat, and the three open boxes represent the G+C-rich subtelomere repeats. The thick line represents plasmid sequence. Below the restriction map of PT4 are the names of subclones of PT4 used as probes. Abbreviations: A, *Alu*I; E, *Eco*RI; Hf, *Hin*fI. (B) Sequence of telomere clone PT4, starting at the chromosome internal *Eco*RI site and extending 1,141 bp until a 29-bp telomere-derived repeat. The DNA nucleotide sequence of the junction region between the 29-bp repeat and the telomere repeat has not been determined. The sequences enclosed in boxes represent three 74-bp directly repeated elements which are 80% similar to each other and have a 65 to 69% G+C content. The boxes are numbered (1 to 3 in each panel) to facilitate comparison. The spacer regions have a G+C content of 21 to 27% and contain many short stretches of poly(dT). The sequences underlined between boxes 2 and 3 represent the predicted bent regions analyzed in Fig. 3. (C) Sequences of the 74-bp G+C-rich repeats (numbered 1 to 3) aligned with each other and with sequences from a previously isolated subtelomere repeat domain from a 1.5-Mb chromosome which contains the expression-linked copy of VSG 118c (15). Dots below the first repeat indicate nucleotide sequence identity; dashes indicate gaps in the sequence, introduced for optimal alignment. The box highlights the 17-bp element which is fully conserved, and the 11-bp example of telomere repeat sequence is underlined.

(9, 23, 48) and the subtelomeric repeats of *Plasmodium berghei* (18), with telomere repeat sequences interspersed among other conserved subtelomere repeat elements.

Evidence that the subtelomeric repeats with the characteristic A+T- and G+C-rich blocks observed in PT4 are representative of subtelomeres of most chromosomes in *T. brucei* is twofold. (i) All chromosome-sized DNA bands separated by PFG hybridize with a probe containing the subtelomeric A+T- and G+C-rich blocks (probe A2 of PT4; see Fig. 1 for probe location), suggesting that most chromosomes contain the repeat (Fig. 2, A2). Figure 2 shows PFG gels run at different conditions to optimally separate chromosomes in bands 1 to 12 (Fig. 2B) and to show the distribution of the larger chromosomes (Fig. 2A) in bands 13 to 19. The differences in the stoichiometry of the ethidium bromide (EtBr) staining versus the hybridization intensities,

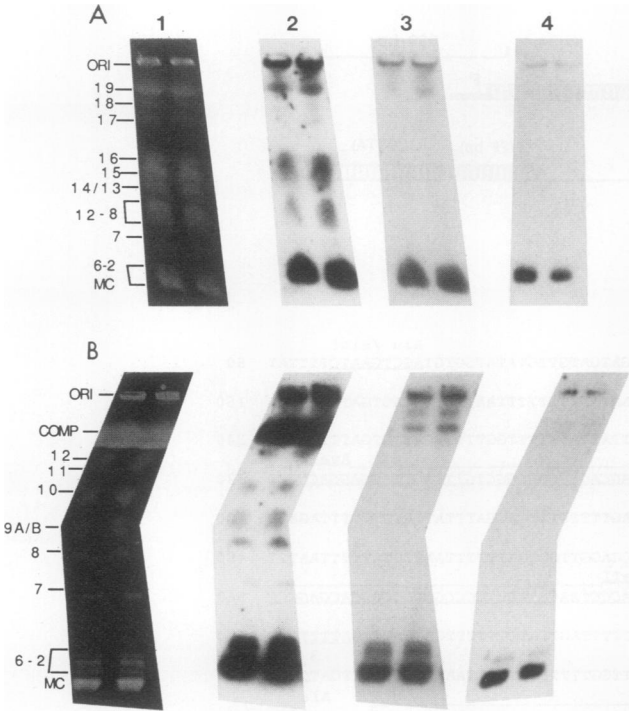

FIG. 2. PFG karyotypes of chromosome-sized DNA. (A) The gel was subjected to electrophoresis at 100 V with a pulse frequency of 3,600 s for 10 days to separate the larger chromosomes. (B) Electrophoresis was performed at 150 V, with a pulse frequency of 900 s for 5 days, to separate the intermediate-sized chromosomes. The chromosome bands are numbered at the left as described previously (21). Columns 2 through 4 represent Southern blots of the EtBr-stained PFG lanes. The DNA in column 2 was hybridized with the PT4-derived subclone A2. Column 3 shows the hybridization with a highly repetitive 177-bp *Alu*I repeat (38, 39) element. Column 4 represents different lanes of the same gel and shows hybridization with the MC-specific probe A4.

with the MC giving the strongest signal, reflect the differences in chromosome number per band. In addition, attrition of larger chromosomes (bands 13 to 19) during runs to separate these molecules leads to diminished hybridization intensity of these bands relative to the MC in Fig. 2A. Further support for the conserved nature of the subtelomere repeats comes from a sequence comparison of these repeats with the subtelomeric domain found adjacent to the transcriptionally active 118c VSG gene in the expression site (the expression-linked copy) on a 1.5-Mb chromosome (15). The G+C-rich blocks of this independently isolated telomere clone are 75% similar to those of PT4 (Fig. 1C), and they share the same 17-bp, absolutely conserved core. (ii) The notion that these repeats are subtelomeric is supported by previously published work which showed that most of the fragments that hybridize with probe A2 are sensitive to treatment of the DNA with the exonuclease Bal 31 (data not shown; 15, 43).

In contrast to the generalized distribution on many chromosomes of the repeats that occur on fragment A2, a small 66-bp *Eco*RI-*Alu*I restriction fragment, derived from clone PT4 (labeled A4 in Fig. 1), hybridized only with trypanosome MC and one additional, 300-kb molecule (Fig. 2, column 2; 4B). The same element was also present on clone PT1 (Fig. 1 and data not shown). Both telomeres represented by clones PT1 and PT4 are therefore likely to be MC derived. A second element which is predominantly localized in the MC and in chromosomes in bands 2 to 6 is a 177-bp *Alu*I repeat element (Fig. 2; 38, 39, 44). There is 177-bp repeat hybridization seen at and above the compression zone of the PFG gels (Fig. 2) which may represent minute quantities of 177-bp repeats in band 19.

The subtelomere repeats contain bent DNA. The poly(dT) motif that occurs in the A+T-rich spacer region of clones PT4 and the 118c expression-linked copy are similar to sequences that have been shown to induce sequence-specific bends in DNA (Fig. 1; 27). The subtelomeres of another protozoan, *P. berghei*, contain bent DNA (19). We were therefore interested in determining whether bent DNA was a characteristic of subtelomeric regions of *T. brucei*.

One of the physical properties of bent DNA is aberrantly slow migration through polyacrylamide gels, leading to a larger apparent molecular weight than is deduced either from sequence analysis or from migration of the same DNA in agarose gels. We first analyzed *Ava*II restriction enzyme digests of clone PT4 to test for the presence of bent DNA because this enzyme cuts in the highly conserved G+C-rich domain of the subtelomere, releasing the A+T-rich subtelomere spacers (Fig. 1). One way to demonstrate this aberrant migration in a single experiment is to run a two-dimensional (2D) gel in which the first dimension is 20% polyacrylamide and the second dimension is 2% agarose. DNA that has normal mobility will form a diagonal, while DNA that has aberrantly slow mobility in polyacrylamide will fall below the diagonal. Of the three *Ava*II restriction fragments of clone PT4 (labeled A to C in Fig. 3A), fragment B (calculated size, 223 bp), which is located between repeats 2 and 3 in Fig. 1, migrated at 223 bp in the agarose dimension and at 315 bp in polyacrylamide (Fig. 3A). Restriction fragment A showed only slightly aberrant migration, while restriction fragment C almost comigrated with the diagonal formed by the size standards. To test whether the aberrant migration of the *Ava*II restriction enzyme fragments was a feature both of the cloned DNA from PT4 and also of *Ava*II subtelomere restriction fragments in total genomic DNA, we repeated the 2D gel analysis with *Ava*II restriction enzyme-digested genomic DNA. This DNA was hybridized with the A2 fragment of clone PT4 to visualize the bent DNA *Ava*II repeats of many different chromosomes. Again the 220-bp band recognized by probe A2 in the agarose dimension runs at a higher apparent molecular weight in the polyacrylamide dimension. The same off-diagonal shift, suggestive of DNA bending, occurs in most if not all of the subtelomere domains of *T. brucei* (Fig. 3B, band B; the marker bands were visualized by scanning the filter with a beta scanner which generates a computer drawing of the signal distribution). The aberrant migration of fragment B, tested at different polyacrylamide concentrations, is shown in Fig. 3D. The characteristic increase in aberrant migration of bent DNA at higher polyacrylamide concentrations again shows that fragment B contains bent DNA.

It is possible that some structural feature other than bending is responsible for the aberrant migration observed in the polyacrylamide dimension. To test this, we analyzed the sequence of the most aberrantly migrating *Ava*II fragment (fragment B in Fig. 3A) with the bent subroutine of DNA Star, a program designed to predict bending in DNA (Fig. 3C). The predicted secondary structure of the DNA helix shows two strongly bent regions, one centering around 585 bp (Fig. 3C; underlined sequence in Fig. 1) and the other centering around 640 bp (Fig. 3C; underlined in Fig. 1; note

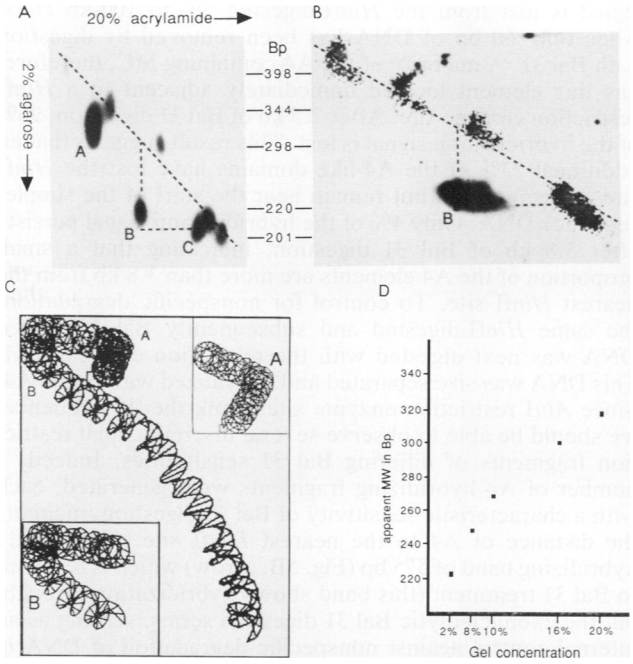

FIG. 3. (A) 2D gel of ^{32}P-end-labeled DNA comparing the mobility of fragments from an *Ava*II restriction enzyme digestion of clone PT4 with the mobility of 1-kb size markers. The first dimension was run at 4°C in 20% polyacrylamide; the second dimension was run in 2% agarose. The sizes and migration in agarose of the size markers are shown between panels A and B. The size markers form a diagonal which is highlighted by the dotted line. Bands marked with letters represent restriction fragments of clone PT4: A, 1 to 312 bp; B, 524 to 747 bp; and C, 313 to 524 bp. Band B migrates with an apparent size of 315 bp in 20% polyacrylamide (as demonstrated by the vertical dotted line) and migrates at 220 bp in agarose (as demonstrated by the horizontal dotted line). There is poor resolution between band C and the 220-bp size marker. (B) Southern blot of a 2D gel run under similar conditions. We loaded unlabeled genomic DNA digested with *Ava*II and introduced a ^{32}P-end-labeled 1-kb size marker into the same sample. The blot was hybridized with the PT4-derived subclone A2. The 1-kb marker diagonal was visualized by exposure of the filter on a beta scanner before hybridization, and alignment of the superimposed images was confirmed by overexposure of the filter that had been hybridized with probe A2. Visualization of the off-diagonal migration of the hybridizing band (labeled B) is facilitated by the vertical and horizontal dotted lines. There is poor resolution between the 220- and 201-bp marker bands in this particular 2D gel. (C) Computer-generated model of the secondary structure of the abnormally migrating *Ava*II fragment (524 to 747 bp in Fig. 1). The inserts highlight the region between 524 and 624 bp (insert A) and 574 and 674 bp (insert B). (D) Apparent sizes of the bent *Ava*II fragment in different concentrations of polyacrylamide (8 to 20%). The first point shows migration in 2% agarose, which agrees with the size predicted from the nucleotide sequence.

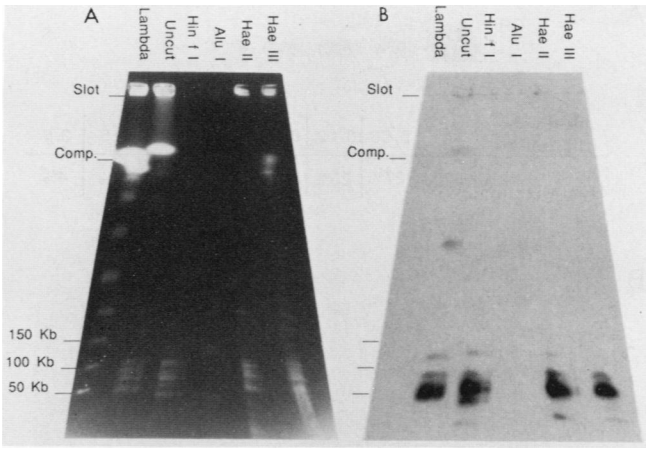

FIG. 4. PFG karyotype of trypanosome DNA electrophoresed for 36 h at 300 V with a 25-s pulse frequency, separating the MC. (A) EtBr-stained gel; (B) Southern blot hybridized with probe A4. The lanes contain, from left to right, lambda ladder size standard; uncut genomic DNA; and genomic DNA digested with *Hin*fI, *Alu*I, *Hae*II, and *Hae*III.

that the terminal location of the bent on the *Ava*II restriction fragment can lead to less drastic aberrant migrational behavior in polyacrylamide electrophoresis). The combination of aberrant migration in polyacrylamide and the outcome of computer modeling suggest that the subtelomeres in *T. brucei* are bent.

Physical mapping of the MC. The telomere and subtelomere sequences enabled us to walk toward the central portion of a MC to identify the complete array of sequence elements on the MC. Since fragment A4 was virtually MC specific, we could use this probe, derived from the most internal portion of clone PT4, to generate a restriction map of the interior of the MC. However, as is the case with subtelomeric sequences from several other organisms (18, 23), probe A4 identifies a subset of chromosomes only, hybridizing to many but not all of the roughly 100 MC (Fig. 2 and 4). The majority of MC recognized by probe A4 were relatively small and measured between 50 and 100 kb, while the larger MC (roughly between 100 and 150 kb) are underrepresented (compare the EtBr-stained gel and its hybridization pattern in Fig. 4 and 7).

The *Alu*I restriction enzyme site used to generate subclone A4 is 4 bp upstream from a *Hin*fI restriction enzyme site present in both telomeres represented by clones PT1 and PT4. Physical mapping indicated that for most of the MC, the next *Alu*I restriction enzyme site was located 200 bp upstream, since digestion of total DNA with *Alu*I and hybridization with probe A4 distributes a majority (>90%) of the hybridization signal to a 200-bp *Alu*I fragment (Fig. 5B, leftmost lane). However, the size of the MC was hardly affected by digestion of total genomic DNA with a series of restriction enzymes other than *Alu*I, even though the EtBr stain of the PFG gels showed that the nuclear DNA was cut to completion. Figure 4 shows a PFG gel in which genomic DNA digested with different restriction enzymes (*Hin*fI, *Alu*I, *Hae*II, and *Hae*III) is hybridized with fragment A4. It is obvious from these gels that probe A4 detects MC DNA, the majority of which does not contain *Hin*fI, *Hae*II, or *Hae*III restriction enzyme sites (Fig. 4). Most of the A4 hybridizing molecules ranged in size from 50 to 100 kb in the uncut lane, while the size of the *Hin*fI-digested molecules is roughly 45 to 100 kb. This indicates that a majority of the MC containing A4 are composed of an array of simple-sequence DNA which extends over 90% of the length of the molecule.

If the telomere represented by clone PT4 is representative of many MC, then the A4 element should be located at the very tip of most of the *Hin*fI-digested 50- to 100-kb MC (see Fig. 5C for a schematic representation of this molecule). The hybridization of 50- to 100-kb *Hin*fI-digested DNA with probe A4 should thus be highly susceptible to digestion with the exonuclease Bal 31 (A4 is relatively resistant to exonuclease treatment in total undigested genomic DNA as a result

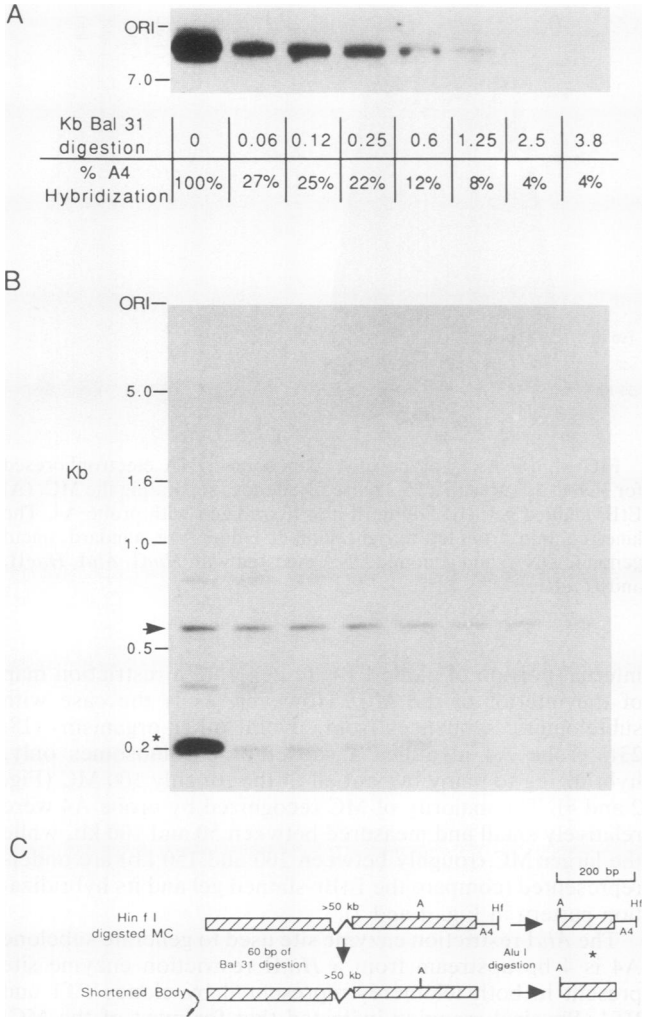

FIG. 5. (A) Southern blot of genomic DNA digested first with *Hin*fI and then with the exonuclease Bal 31. The DNA was size separated in a 1% agarose gel and hybridized with probe A4. The migration of size standards (in kilobases) is noted on the left; ORI denotes the position of the well. The top row in the table below each lane shows the amount of DNA in kilobases removed by Bal 31 from each end of the DNA molecule. The second row gives the relative hybridization intensity in the lane as a percentage of the counts per minute present on the filter in the first lane. (B) Blot in which the *Hin*fI-digested, Bal 31-treated DNA shown in panel A was digested with the restriction enzyme *Alu*I, after which the DNA was size separated and transferred to nitrocellulose and hybridized with probe A4. The table gives the number of base pairs digested by Bal 31 at each time point. The arrow on the left marks a band which persists over the range of Bal 31 digestion; the asterisk shows the position of the predominant class of Bal 31-sensitive, A4-containing molecules. (C) Schematic representation of the predominant molecules observed in the Bal 31 experiments. The asterisks indicates the first *Alu*I-*Hin*fI fragment from the MC, hybridizing in panel B, lanes 1 and 2.

of its subtelomeric location; 43; data not shown). To test whether A4 is at the junction between the subtelomere and the simple-sequence body of an MC, we digested total genomic DNA with the restriction enzyme *Hin*fI, subjected this DNA to a short incubation with the enzyme Bal 31, and hybridized it with probe A4 (Fig. 5A; controls are presented in Fig. 5B; see below). Indeed 73% of the A4 hybridization signal is lost from the *Hin*fI-digested 50- to 100-kb DNA when only 60 bp of DNA has been removed by digestion with Bal 31. A majority of the A4-containing MC, therefore, has this element located immediately adjacent to a *Hin*fI restriction enzyme site. After 2.5 kb of Bal 31 digestion, 96% of the hybridization signal is lost. This result suggests that an additional 23% of the A4-like domains have lost the *Hin*fI site adjacent to A4 but remain near the start of the simple-sequence DNA. Only 4% of the hybridization signal persists after 3.8 kb of Bal 31 digestion, indicating that a small proportion of the A4 elements are more than 3.8 kb from the nearest *Hin*fI site. To control for nonspecific degradation, the same *Hin*fI-digested and subsequently Bal 31-treated DNA was next digested with the restriction enzyme *Alu*I. This DNA was size separated and hybridized with probe A4. Since *Alu*I restriction enzyme sites flank the A4 sequence, we should be able to observe several discrete, small restriction fragments of differing Bal 31 sensitivities. Indeed, a number of A4 hybridizing fragments were generated, each with a characteristic sensitivity of Bal 31 digestion, mapping the distance of A4 to the nearest *Hin*fI site: (i) a faintly hybridizing band of 575 bp (Fig. 5B, arrow) which is resistant to Bal 31 treatment (this band shows hybridization throughout the exonucleolytic Bal 31 digestion series, serving as an internal control against nonspecific degradation of DNA in Fig. 5A and B); (ii) a 200-bp fragment (Fig. 5B and C, asterisks) representing the dominant Bal 31-sensitive PT4-like A4 species, which loses almost all hybridization after only 60 bp of Bal 31 digestion; and (iii) a minor 200-bp fragment which is, as judged from its Bal 31 sensitivity, located slightly more internally. We conclude that in the majority of MC, the A4 sequence is located immediately adjacent to a *Hin*fI restriction enzyme site, which marks the start of an array of simple-sequence DNA that lacks restriction enzyme sites for *Hin*fI, *Hae*II, and *Hae*III.

Presence of a 177-bp repeat array immediately adjacent to the telomere. In a large population of MC, the A4 sequence specifically marks the start of the simple-sequence body of an MC. This allowed us to map the simple-sequence interior of the MC by performing an indirect end-labeling analysis (40) of partially digested MC DNA. By using A4 as a specific end label, a restriction map of all hybridizing fragments can be drawn by simply measuring the size of each fragment in the partial digestion series (Smith-Birnstiel indirect end-label mapping; 40). The size difference between each fragment in the series will thus indicate the distance between the restriction enzyme sites on the fragment. The simple-sequence body of these MC is likely to have *Alu*I restriction enzyme sites, since digestion with this enzyme abolishes A4 hybridization from the high-molecular-weight DNA and leaves very little DNA in the size range between 50 and 150 kb (Fig. 4A). We first digested total genomic DNA with *Hin*fI and subsequently performed a partial restriction enzyme digestion on this DNA with enzymes that can cut the 177-bp repeat. This DNA was size separated and hybridized with the A4 probe, which marks the beginning of the simple-sequence DNA. When either *Cfo*I or *Alu*I was used to partially digest the simple-sequence DNA, a ladder of fragments with a 177-bp periodicity occurred, with the smallest band being 220 bp for *Cfo*I (Fig. 6A) and 200 bp for *Alu*I (data not shown). A discrete banding pattern, with a periodicity of 177 bp, was observed in the region between 0.2 to 4 kb, after which a smear of hybridizing DNA was detected that continued uninterrupted to the compression zone of these gels (over 23 kb in size).

These data indicate that a large region of the MC is

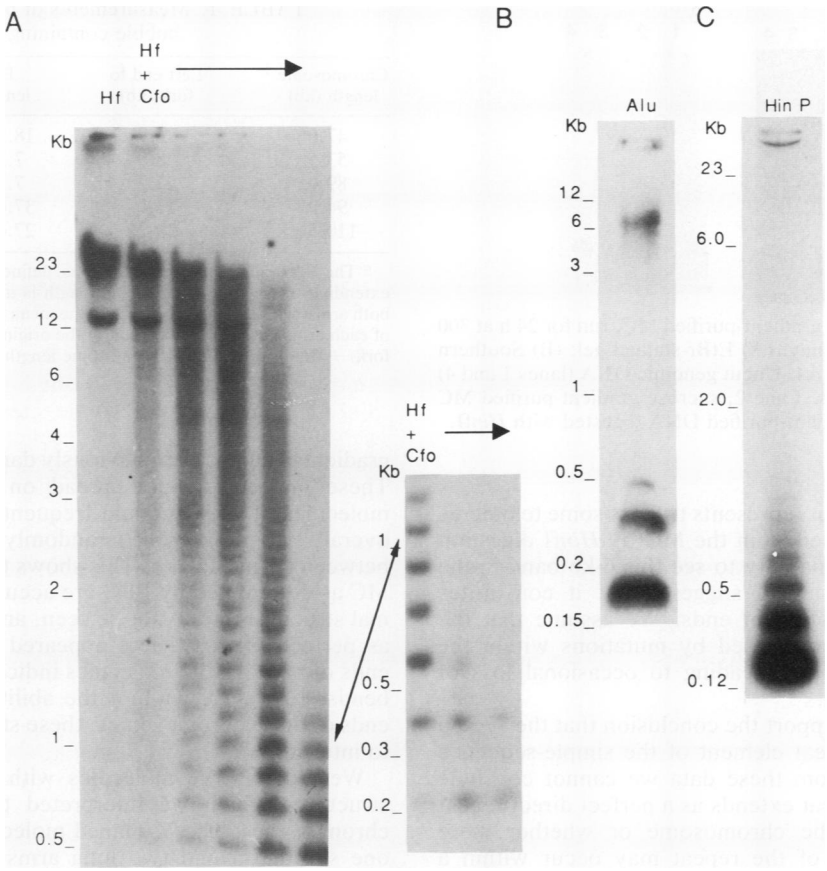

FIG. 6. (A) Southern blot of genomic DNA digested to completion with *Hin*fI (Hf), partially digested with the restriction enzyme *Cfo*I, and then hybridized with probe A4. The migration of size standards is shown to the left of each panel. The first gel is 0.7% agarose, showing the ladder of *Cfo*I hybridizing bands that range from 0.5 to 23 kb. The second panel is the low-molecular-weight portion of a 1% agarose gel. The partial digestions create a 177-bp ladder. (B) Southern blot of purified *Hin*fI-digested, 50- to 150-kb simple-sequence DNA from MC, subsequently digested with the restriction enzyme *Alu*I and size separated on a 1% agarose gel. Hybridization was with ^{32}P-labeled, purified simple-sequence MC DNA (labeling by random priming). Posthybridization washes were in 0.1× SSC at 65°C. (C) Autoradiogram of a dried 0.8% agarose gel containing ^{32}P-end-labeled *Hin*pI-digested simple-sequence MC DNA (purified *Hin*fI-digested simple-sequence DNA from MC; *Hin*pI is an isoschizomer of *Cfo*I).

composed of a 177-bp repeat element. The most likely explanation of these results is that the simple-sequence body of the MC is made from a highly repetitive element with *Alu*I and *Cfo*I restriction enzyme sites, lacking *Hin*fI and *Hae*II and *Hae*III restriction enzyme sites. An abundant 177-bp repeat with *Alu*I and *Cfo*I restriction enzyme sites that was predominantly located in MC had been described in *T. brucei* (44; see ref. 38 for DNA nucleotide sequence of the 177-bp repeat; see Fig. 2 for 177-bp repeat location in MC). The simple sequence of the MC therefore most likely consists of this 177-bp repeat.

The presence of a heterogeneous sequence element within the simple-sequence array would appear as a gap in this ladder if it were at a fixed distance from the A4 sequence. The absence of gaps in the ladder suggests that the simple-sequence body is exclusively composed of 177-bp repeats. However, this analysis does not exclude randomly placed heterogeneous simple-sequence elements in this population of MC.

The most abundant repeat of purified MC measures 177 bp. The Smith-Birnstiel physical mapping analysis described above was limited in its ability to identify all of the genetic elements of the body of an MC. We were interested in identifying any non-177-bp elements in the body of these chromosomes, since they could be centromeres or origins of replication. We therefore used an independent approach to assess the composition of the simple sequence of the MC. We purified MC by PFG, *Hin*fI digested these molecules, and again isolated the large (50- to 100-kb) DNA from PFG gels. We then digested this DNA with restriction enzymes that will cut in the 177-bp repeat. The composition of this collection of fragments was assayed by Southern blotting, in which the filter is hybridized with ^{32}P-labeled purified 50- to 100-kb *Hin*fI simple-sequence MC DNA (Fig. 6B); alternatively, the digested DNA was end labeled and size separated (Fig. 6C). Both strategies will represent the stoichiometry of repetitive elements which exists in the body of the 50- to 100-kb *Hin*fI fragments. In the Southern blot, the hybridization signal of the bands will be proportional to the mass of an element in the sample; in the ^{32}P-end-labeled DNA, the signal intensity of a band will be roughly proportional to the molarity of any fragment in the sample. Both experiments demonstrated predominant bands at 177 bp only, in addition to bands that represent multimers of 177 bp. In the Southern blot, a single additional repeat class can be detected at about 6 kb. Since this DNA hybridizes with trypanosome telomere

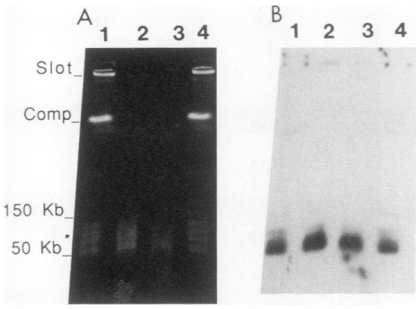

FIG. 7. PFG of sucrose gradient-purified MC, run for 24 h at 300 V with a 20-s pulse frequency. (A) EtBr-stained gel; (B) Southern blot hybridized with probe A4. Uncut genomic DNA (lanes 1 and 4) was used as size standards. Lane 2, sucrose gradient-purified MC DNA; lane 3, sucrose gradient-purified DNA digested with *Hin*fI.

repeats, we believe that it represents trypanosome telomeres that could not be released from the MC by *Hin*fI digestion (data not shown). The inability to see this 6-kb band in the ^{32}P-end-labeled DNA sample suggests that it constitutes only a small molar fraction of ends. We assume that the multimers of 177 bp are created by mutations within the 177-bp repeat element (38), leading to occasional loss of restriction enzyme sites.

These experiments support the conclusion that the 177-bp repeat is the major repeat element of the simple-sequence DNA of these MC. From these data we cannot conclude whether the 177-bp repeat extends as a perfect direct repeat to the other end of the chromosome or whether more complex arrangements of the repeat may occur within a single chromosome. Proof that short stretches of other heterogeneous elements are absent from the simple-sequence body of an MC awaits the cloning of an entire MC in yeast artificial chromosomes (29), followed by comparison of the structure of the cloned molecule with data obtained from the structural analysis of MC described here.

EM analysis of MC. We considered it useful to visualize the molecules by EM as part of our structural analysis. Hairpins, loops, knots, or cruciforms that could be part of these naturally occurring chromosomes might thus be identified. However, the PFG-isolated MC used in the analyses described above were of insufficient quality to allow their visualization by EM. This was due mainly to DNA damage by exposure to UV but was also due to contaminating agarose (data not shown). However, the relatively small (50 to 150 kb) size of these molecules makes them stable in solution and therefore allowed their isolation on sucrose sedimentation gradients after in situ lysis of *T. brucei* in 1% SLS supplemented with 1 mg of proteinase K per ml (39).

Sucrose gradients were therefore used to isolate microgram quantities of MC that had been stripped of their associated proteins. The sedimentation characteristics of the MC were sufficiently different from those of larger chromosomes to allow fractionation of virtually pure populations of MC. An EtBr-stained PFG gel of these purified MC showed that sucrose-purified MC were indistinguishable from MC that had been prepared in agarose blocks (Fig. 7; compare lanes 1 and 4 with lanes 2). Digestion of these MC with *Hin*fI showed that only some of the larger MC (100 to 150 kb), which do not hybridize with A4, are shortened by *Hin*fI digestion, while the smaller MC (50 to 100 kb) cannot be cut by *Hin*fI. A Southern blot of this gel hybridized with the A4 element confirmed that the DNA isolated from the sucrose

TABLE 1. Measurements of putative replication bubble-containing MC[a]

Chromosome length (kb)	Left end to fork (kb)	Bubble length (kb)	Fork to right end (kb)
47.6	6.9	18.2 ± 0.5	22.5
53.5	8.9	7.2 ± 0.1	37.5
89.0	27.2	7.7 ± 0.1	54.0
94.9	13.0	37.6 ± 0.7	44.4
110.5	39.2	27.4 ± 0.4	43.9

[a] The left end of the molecule was defined as having the shortest arm extending to the fork. The bubble length is a mean of the measurements of both arms, with the variation around the mean giving the actual measurements of each arm. The relative position of the origin was calculated by [(left end to fork) + (bubble length/2)]/chromosome length.

gradients had not been obviously damaged by the procedure. These molecules were spread on EM grids. Full-length molecules of 150 kb could frequently be detected, and the overall size range of 50 randomly chosen MC measured between 50 and 150 kb. This shows that the size estimates of MC as determined by PFG are accurate. In addition, abnormal structures could not be seen, and the molecule behaved as perfect linears. There appeared to be structures at the ends of many of the molecules indicative of small knobs and bends. However, without the ability to mark the extreme ends of these chromosomes, these structures remain difficult to interpret.

We did observe molecules with bubbles and Y-shaped structures which we interpreted to represent replicating chromosomes. The Y-shaped molecules predominantly had one short arm and two long arms, but it was sometimes difficult to trace each arm unambiguously to its end. Mapping of the origin of replication could therefore not be done on these molecules. However, since 25 Y-shaped molecules observed contained only one replication fork, the data support the notion that only a single origin of replication is present on each MC. The five molecules containing a putative replication bubble were more informative, allowing low-resolution mapping of the origin of replication in this class of molecules. First, all molecules analyzed contained only a single putative replication bubble. Second, a summary of the lengths in kilobases of the components of these molecules, which allows the tentative mapping of an origin of replication, is given in Table 1. A good example of a molecule with a putative replication bubble is shown in Fig. 8. This 47.6-kb (3.12 kb/μm) molecule has two forks, with the length of DNA in each arm of the bubble being 17.7 and 18.7 kb (denoted as 18.2 ± 0.5 in Table 1). The variability of arm lengths within a bubble was never more than 3% in any molecule analyzed. Therefore, we interpret this structure as representing a replicating chromosome, presumably with a single bidirectional origin of replication located about 34% of the way into the chromosome (assuming a single initiation event and equivalent migration of both replication forks). When this analysis was done on all five bubble-containing chromosomes, the mean position of the putative origin was 35% ± 5% of the way into the molecule, with a 95% confidence interval for the origin of replication being located between 20 and 50% of the distance from the end of the MC. This analysis thus suggests that the MC contains a single origin of replication which lies in the simple-sequence body of the molecule. We will need a larger sample size to map the origin of replication more accurately in these molecules. We will also need to analyze the complete nucleotide sequence

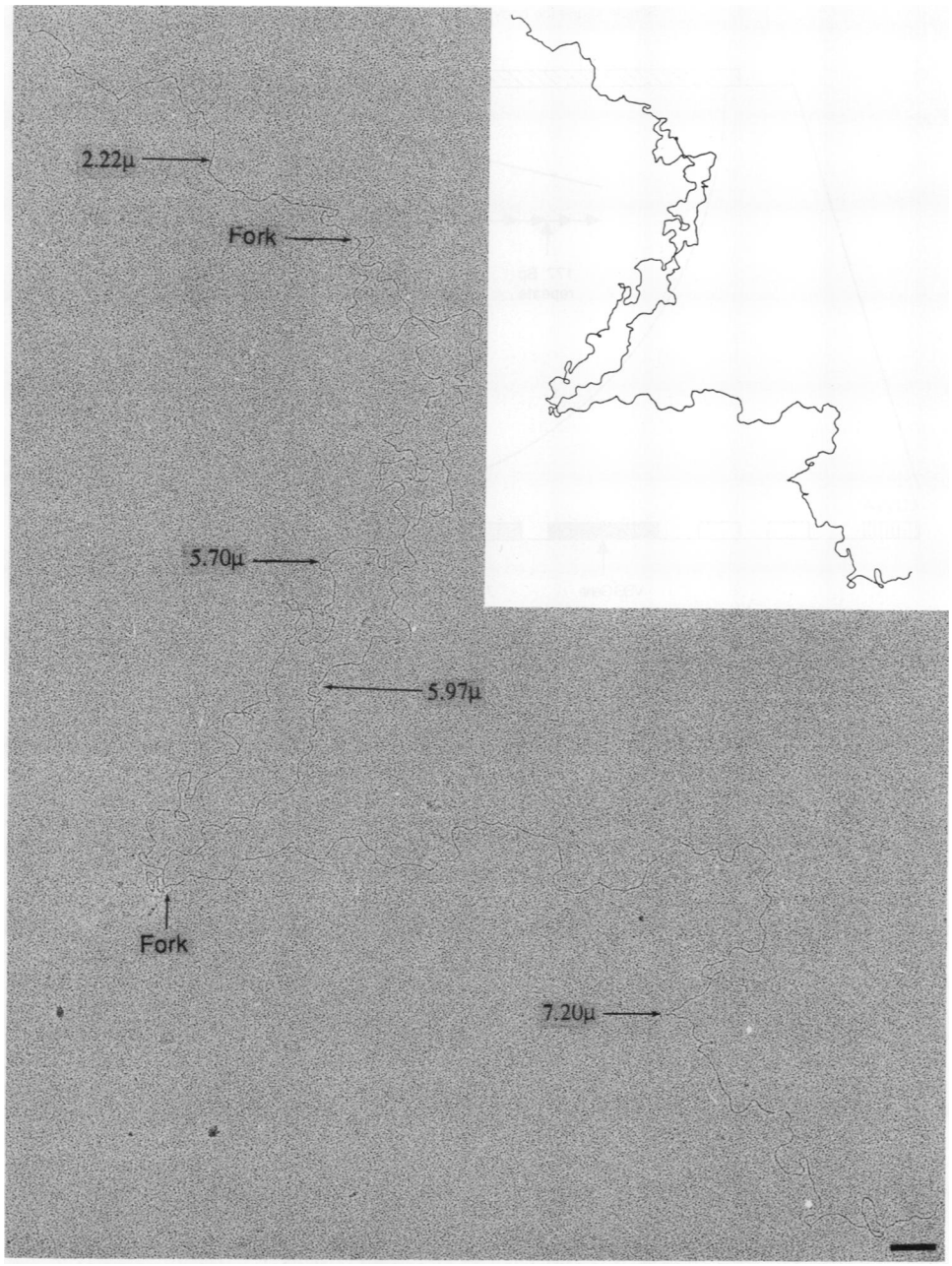

FIG. 8. (A) Electron micrograph of a 47.6-kb molecule with a single bubble; (B) line drawing at one-half scale. The lengths of each segment in micrometers are given, with a conversion factor of 3.12 kb/μm. The bar represents 0.2 μm.

of a single MC, cloned in yeast cells, to identify any nonhomogeneous element that may serve as an origin of replication buried in the 177-bp-repeat simple-sequence body of the MC.

The presence of replicating molecules may also account for the DNA that shows 177-bp-repeat hybridization that we have observed at the wells, the compression zone, and between the well and the compression zone of these PFG gels (Fig. 2, column 3, bands between well and the compression zone). These hybridizing molecules may represent abnormally migrating MC due to the presence of Y structures, or bubbles as described here. Support for this conclusion also comes from the observation that fractions from the top of the sucrose gradients, which contain MC-sized DNA, have 177-bp-repeat hybridizing molecules which comigrated with the 5-Mb chromosomes in PFG gels, and fractions which have a greater proportion of abnormally migrating 177-bp hybridizing molecules in PFG gels also have a greater percentage of Y structures and bubbles in EM analysis (data not shown).

DISCUSSION

We have defined a majority of the sequence elements which constitute a protozoan chromosome, starting with a sequence analysis of the telomere and subtelomere domains

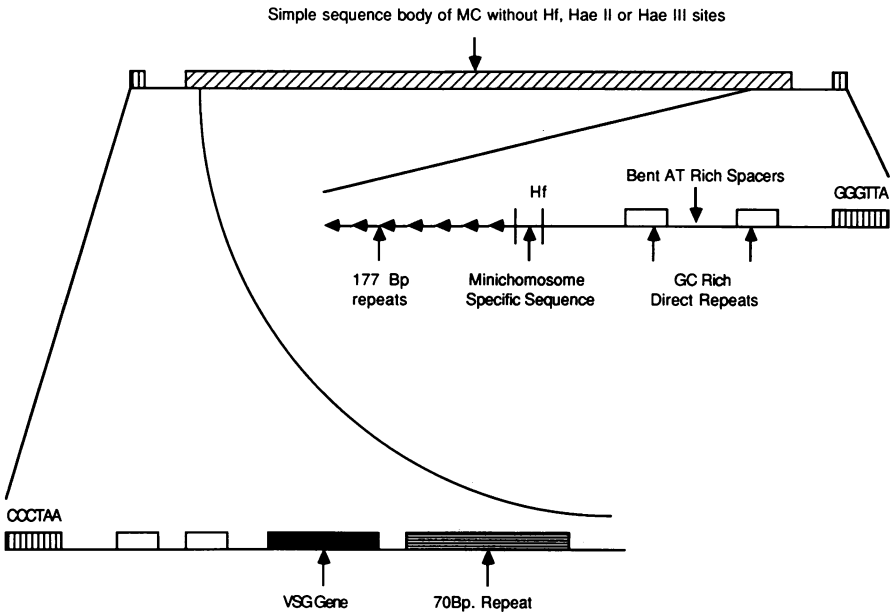

FIG. 9. Model of an MC combining the information obtained about these molecules. Data showing the occurrence of 70-bp repeats and VSG genes in MC have been described previously (44).

of cloned MC molecules and extending this into a physical map of an MC. A large array of simple-sequence DNA, composed predominantly of a 177-bp *Alu*I repeat, comprises over 90% of the length of a subset of MC. One of the many unusual aspects of this set of molecules is that the subtelomere domain can be compressed into a 1.4-kb region which is immediately adjacent to simple-sequence DNA. The exact composition of the other telomere of the MC is open to speculation. The MC, including those between 50 and 100 kb, can contain telomeric protein-coding VSG genes in a transcriptionally inactive form, along with a simple-sequence 70-bp repeat that flanks the VSG genes (8, 37, 42). We have collated this information into a preliminary model of a trypanosome MC (Fig. 9). One of the surprising conclusions of these experiments is that there is room for only one or a few (VSG) genes on these MC and that the repertoire of telomeric VSGs at MC may be limited because of the presence of empty telomeres. It is unclear whether only one telomere of each MC encodes a VSG gene or whether MC that are entirely devoid of VSG genes exist. Since VSG genes are present in the smaller MC as well (44), it is obvious that the smaller 50- to 100-kb MC do encode VSG genes, as reflected in the model presented in Fig. 9. The function of the MC may be to allow random assortment of many VSG genes, without the need for recombination, during genetic exchange. In addition, the MC can provide the cell with a pool of rapidly evolving telomeric VSG genes, an advantage in an area where trypanosomiasis is endemic. MC-derived VSG genes do not appear to be transcribed. Our unpublished data indicate a low level of transcription for some of the subtelomere repeats (35, 48a). Because of the repetitive nature of these probes, it is not possible to determine whether these transcripts are MC derived.

Despite the small size of the MC-derived telomere clones, we have shown that many, if not all, subtelomeres in *T. brucei* share features with the structures that we observed in clone PT4. Probes that contain the G+C-rich element found in PT4 hybridized to all chromosomes, and this element was also present in the sequence of a subtelomere derived from a 1.5-Mb chromosome (15). The significance of the subtelomeric A+T- and G+C-rich blocks is unclear. However, subtelomeres of organisms as distantly related as yeasts, trypanosomes, and plasmodia are similar in this respect: subtelomere domains in *P. berghei* (18) and the X and Y elements of *S. cerevisiae* have clusters of telomere repeats separated by spacer regions (9, 23, 48). However, in both of these examples only some subtelomeres contain these elements, and the internal representation of the telomere sequence is, in contrast to that of *T. brucei*, directly repeated relative to the repeats at the end of the chromosome.

The X and Y elements in yeast cells can function as autonomous replication sequences and have also been identified as mobile genetic elements with a circular episomal intermediate form (23). The *P. berghei* subtelomeric elements are also mobile and can be transferred to subtelomeres of other chromosomes by an as yet unknown mechanism (31). Recombination facilitated by the presence of these subtelomere repeats may be responsible for the instability observed in the telomere domains of *T. brucei* (1, 30, 43) and other protozoans (14).

The presence of bent DNA in the subtelomere repeats of many trypanosome telomeres as well as the subtelomeres of other eukaryotes (19) suggests a specific role for the bent DNA. Possible functions of DNA bending in the subtelomere could be as recognition sites for binding to the nuclear scaffold or as binding sites for DNA topoisomerases (20, 22, 26), perhaps serving a function in determining the specificity observed of the subnuclear localization of telomeric sequences in *T. brucei* (10).

We have presented evidence for the notion that the main body of a subset of the MC consists of simple-sequence DNA, directly flanked by the subtelomere repeat elements. At the moment we do not have evidence for other nonhomogeneous sequences within the body of the MC, and the 177-bp repeat sequence therefore becomes a reasonable candidate for a trypanosome centromere. If this is the case,

the centromere in the MC more closely resembles that of *Schizosaccharomyces pombe* (12) than that of *S. cerevisiae* (11). The larger chromosomes would have a different centromeric sequence since they lack the long arrays of the 177-bp repeats. An interesting observation from analysis of in situ hybridization of trypanosome nuclei in the G_1, S, or G_2 phase was that the 177-bp repeats in G_1- and S-phase nuclei were associated with the nuclear envelope, while these repeats aligned at the center of the nucleus in some G_2-phase nuclei (10). The localization of 177-bp repeats in G_2-phase nuclei preceded the condensation of the majority of the DNA at the metaphase plate, as determined by 4,6-diamidino-2-phenylindole staining. This observation is in line with the notion that MC behave independently of the other chromosomes during division.

The EM data suggest that the origin of replication lies in the simple-sequence body of the MC. One possible explanation for this observation is that DNA replication initiates in the 177-bp repeat sequence and uses distance from a matrix attachment region as the signal for initiation (47). Alternatively, there may be a specific element buried in the simple-sequence DNA which acts as an origin of replication (6, 7, 24).

Finally, the size range of the MC appears to be constrained between 50 and 150 kb. We may now be able to address the mechanisms leading to these chromosome size limitations. (i) The size of the MC may be determined, at the lower limit, by reduced mitotic stability of very small chromosomes, as has been shown in yeast cells (29); and (ii) if only a single, centrally located origin of replication is present, the upper limit of the size of the MC may be constrained by the time it takes to copy the entire chromosome, once per cell cycle.

ACKNOWLEDGMENTS

We thank all colleagues for critical reading of the manuscript. M.W. thanks Peter Caldwell for his encouragement.

M.W. was supported by NIH grant T32 HL 07018 and is a James Burns Amberson Fellow of the New York Lung Association. This work was supported by NIH grant A1 21784 to L.H.T.V.D.P., by a grant from the John D. and Catherine T. MacArthur Foundation to L.H.T.V.D.P., and by grants from the NSF (DCB 9004860) and American Cancer Society (NP493) to A.L.B. L.H.T.V.D.P. is a Burroughs Wellcome Scholar in Molecular Parasitology.

REFERENCES

1. **Bernards, A., P. A. M. Michels, C. R. Lincke, and P. Borst.** 1983. Growth of chromosome ends in multiplying trypanosomes. Nature (London) **303**:592–597.
2. **Birnboim, H. C., and J. Doly.** 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. **7**:1513–1523.
3. **Blackburn, E. H.** 1990. Telomeres and their synthesis. Science **249**:489–490.
4. **Blackburn, E. H., and P. B. Challoner.** 1984. Identification of a telomeric DNA sequence in Trypanosoma brucei. Cell **36**:447–457.
5. **Borst, P.** 1986. Discontinuous transcription and antigenic variation in trypanosomes. Annu. Rev. Biochem. **55**:701–732.
6. **Brewer, B. J., and W. L. Fangman.** 1987. The location of replication origins on ARS plasmids in *S. cerevisiae*. Cell **51**:463–471.
7. **Burhans, W. C., L. T. Vassilev, M. S. Caddle, N. H. Heintz, and M. L. DePamphilis.** 1990. Identification of an origin of bidirectional DNA replication in mammalian chromosomes. Cell **62**: 955–965.
8. **Campbell, D. A., M. P. Van Bree, and J. C. Boothroyd.** 1984. The 5′-limit of transposition and upstream barren region of a trypanosome VSG gene: tandem 76 base-pair repeats flanking $(TAA)_{90}$. Nucleic Acids Res. **12**:2759–2774.
9. **Chan, C. S. M., and B.-K. Tye.** 1983. Organization of DNA sequences and replication origins at yeast telomeres. Cell **33**: 563–573.

9a. **Cheng, J. F., C. L. Smith, and C. R. Cantor.** 1989. Isolation and characterization of a human telomere. Nucleic Acids Res. **17**:6109–6127.

10. **Chung, H. M., D. B. Tse, S. Fields, R. N. Taub, C. Shea, and L. H. T. Van der Ploeg.** 1990. Architectural organization in the interphase nucleus of the protozoan *Trypanosoma brucei*: location of telomeres and mini-chromosomes. EMBO J. **9**:2611–2619.
11. **Clarke, L., and J. Carbon.** 1980. Isolation of a yeast centromere and construction of functional small circular chromosomes. Nature (London) **287**:504–509.
12. **Clarke, L., and M. Baum.** 1990. Functional analysis of a centromere from fission yeast: a role for centromere-specific repeated DNA sequences. Mol. Cell. Biol. **10**:1863–1872.
13. **Clayton, C. E.** 1988. The molecular biology of kinetoplastidae. Genet. Eng. News **7**:1–56.
14. **Corcoran, L. M., J. K. Thompson, D. Walliker, and D. J. Kemp.** 1988. Homologous recombination within subtelomeric repeat sequences generates chromosome size polymorphisms in *P. falciparum*. Cell **53**:807–813.
15. **De Lange, T., J. M. Kooter, P. A. M. Michels, and P. Borst.** 1983. Telomere conversion in trypanosomes. Nucleic Acids Res. **11**:8149–8165.
16. **De Lange, T., L. Shiue, R. M. Myers, D. R. Cox, S. L. Naylor, A. M. Killery, and H. E. Varmus.** 1990. Structure and variability of human chromosome ends. Mol. Cell. Biol. **10**:518–527.
17. **Dickmann, S., and D. A. Zarling.** 1987. Unique poly(dA): poly(dT) B′-conformation in cellular and synthetic DNAs. Nucleic Acids Res. **15**:6063–6074.
18. **Dore, E., T. Pace, M. Ponzi, L. Picci, and C. Frontali.** 1990. Organization of subtelomeric repeats in *Plasmodium berghei*. Mol. Cell. Biol. **10**:2423–2427.
19. **Dore, E., T. Pace, M. Ponzi, R. Scotti, and C. Frontali.** 1988. A site of intrinsic bending in a highly repeated element of *P. berghei* genome. Mol. Biochem. Parasitol. **27**:201–206.
20. **Fairlamb, A. H., P. O. Weisvogel, J. H. J. Hoeijmakers, and P. Borst.** 1978. Isolation and characterization of kinetoplast DNA from bloodstream form *Trypanosoma brucei*. J. Cell Biol. **76**:293–309.
21. **Gottesdiener, K., J. Garcia-Anoveros, M. G.-S. Lee, and L. H. T. Van der Ploeg.** 1990. Chromosome organization of the protozoan *Trypanosoma brucei*. Mol. Cell. Biol. **10**:6079–6083.
22. **Homberger, H. P.** 1989. Bent DNA is a structural feature of scaffold-attached regions in *Drosophila melanogaster* interphase nuclei. Chromosoma **98**:99–104.
23. **Horowitz, H., and J. E. Haber.** 1985. Identification of autonomously replicating circular subtelomeric Y′ elements in *Saccharomyces cerevisiae*. Mol. Cell. Biol. **5**:2369–2380.
24. **Huberman, J. A., L. D. Spotila, K. A. Nawotka, S. M. El-Assouli, and L. R. Davis.** 1987. The in vivo replication origin of the yeast 2μm plasmid. Cell **51**:473–481.
25. **Lang, D., and M. Mitani.** 1970. Simplified quantitative electron microscopy of biopolymers. Biopolymers **9**:373–379.
26. **Linial, M., and J. Shlomai.** 1988. A unique endonuclease from *Crithidia fasciculata* which recognizes a bend in the DNA helix. J. Biol. Chem. **263**:290–297.
27. **Marini, J. C., S. D. Levene, D. M. Crothers, and P. T. Englund.** 1982. Bent helical structure in kinetoplast DNA. Proc. Natl. Acad. Sci. USA **79**:7664–7668.
28. **Moyzis, R. K., J. M. Buckingham, L. S. Cram, M. Dani, L. L. Deaven, M. D. Jones, J. Meyne, R. L. Ratliff, and J. R. Wu.** 1988. A highly conserved repetitive DNA sequence, $(TTAGGG)_n$, present at the telomeres of human chromosomes. Proc. Natl. Acad. Sci. USA **85**:6622–6626.
29. **Murray, A. W., and J. S. Szostak.** 1983. Construction of an artificial chromosome in yeast. Nature (London) **305**:189–193.
30. **Myler, P. J., R. F. Aline, Jr., J. K. Scholler, and K. D. Stuart.** 1988. Changes in telomere length associated with antigenic

variation in *Trypanosoma brucei*. Mol. Biochem. Parisitol. **29:** 243–250.

31. **Pace, T., M. Ponzi, E. Dore, C. Janse, B. Mons, and C. Frontali.** 1990. Long insertions within telomeres contribute to chromosome-size polymorphisms in *P. berghei*. Mol. Cell. Biol. **10:** 6759–6764.
32. **Parsons, M., R. G. Nelson, and N. Agabian.** 1984. Antigenic variation in African trypanosomes: DNA rearrangements program immune evasion. Immunol. Today **5:**43–50.
33. **Pays, E.** 1989. Pseudogenes, chimeric genes and the timing of antigenic variation in trypanosomes. Trends Genet. **5:**389–391.
34. **Pluta, A. F., and V. A. Zakian.** 1989. Recombination occurs during telomere formation in yeast. Nature (London) **337:**429–433.
35. **Rudenko, G., and L. H. T. Van der Ploeg.** 1989. Transcription of telomere repeats in protozoa. EMBO J. **8:**2633–2838.
36. **Sanger, F., A. R. Coulsen, B. G. Barrell, A. J. H. Smith, and B. Roe.** 1980. Cloning in single stranded bacteriophage as an aid to rapid DNA-sequencing. J. Mol. Biol. **143:**161–178.
37. **Shah, S. J., J. R. Young, B. E. Kimmel, K. P. Iams, and R. O. Williams.** 1987. The 5′ flanking sequence of a *Trypanosoma brucei* variable surface glycoprotein. Mol. Biochem. Parisitol. **24:**163–174.
38. **Sloof, P., J. L. Bos, A. F. J. M. Konings, H. H. Menke, P. Borst, W. E. Gutteridge, and W. Leon.** 1983. Characterization of satellite DNA in *Trypanosoma brucei* and *Trypanosoma cruzi*. J. Mol. Biol. **167:**1–21.
39. **Sloof, P., H. H. Menke, M. P. M. Caspers, and P. Borst.** 1983. Size fractionation of *Trypanosoma brucei* DNA: localization of the 177-bp repeat satellite DNA and variant surface glycoprotein gene in a mini-chromosomal DNA fraction. Nucleic Acids Res. **11:**3889–3901.
40. **Smith, H. O., and M. L. Birnstiel.** 1976. A simple method for DNA restriction site mapping. Nucleic Acids Res. **3:**2387–2398.
41. **Van der Ploeg, L. H. T.** 1990. Antigenic variation of African trypanosomes: genetic recombination and transcriptional control of VSG genes, p. 51–97. *In* B. Hames and D. L. Glover (ed.), Frontiers in molecular biology: genome rearrangements and amplification. IRL Press, Oxford.
42. **Van der Ploeg, L. H. T., A. W. C. A. Cornelissen, J. D. Barry, and P. Borst.** 1984. Chromosomes of Kinetoplastida. EMBO J. **3:**3109–3115.
43. **Van der Ploeg, L. H. T., A. Y. C. Liu, and P. Borst.** 1984. Structure of the growing telomeres of trypanosomes. Cell **36:** 459–468.
44. **Van der Ploeg, L. H. T., D. C. Schwartz, C. R. Cantor, and P. Borst.** 1984. Antigenic variation in *Trypanosoma brucei* analyzed by electrophoretic separation of chromosome sized DNA molecules. Cell **37:**77–84.
45. **Van der Ploeg, L. H. T., C. L. Smith, R. I. Polvere, and K. Gottesdiener.** 1989. Improved separation of chromosome sized DNA from *Trypanosoma brucei* stock 427-60. Nucleic Acids Res. **17:**3217–3227.
46. **Van der Ploeg, L. H. T., D. Valerio, T. De Lange, A. Bernards, P. Borst, and F. G. Grosveld.** 1982. Analysis of cosmid clones of nuclear DNA from *Trypanosoma brucei* shows that the genes for variant surface glycoproteins are clustered in the genome. Nucleic Acids Res. **10:**5905–5923.
47. **Vaughn, J. P., P. A. Dijkwei, and J. L. Hamlin.** 1990. Replication initiates in a broad zone in the amplified CHO dihydrofolate reductase domain. Cell **61:**1075–1087.
48. **Walmsley, R. W., C. S. M. Chan, B.-K. Tye, and T. D. Petes.** 1984. Unusual DNA sequences associated with the ends of yeast chromosomes. Nature (London) **310:**157–160.

48a. **Weiden, M., and L. H. T. Van der Ploeg.** Unpublished data.

49. **Wells, J. M., T. D. Prospero, L. Jenni, and R. F. E. LePage.** 1987. DNA contents and molecular karyotypes of hybrid *Trypanosoma brucei*. Mol. Biochem. Parasitol. **249:**103–116.
50. **Wu, H., and D. M. Crothers.** 1984. The locus of sequence-directed and protein induced DNA bending. Nature (London) **308:**509–513.