Characterization of cloned human alphoid satellite with an unusual monomeric construction: evidence for enrichment in HeLa small polydisperse circular DNA

Richard S.Jones and S.Steven Potter

Department of Molecular Biology and Biochemistry, Wesleyan University, Middletown, CT 06457, USA

Received 11 July 1984; Revised and Accepted 4 January 1985

ABSTRACT

A recombinant DNA plasmid library was constructed from HeLa cell extrachromosomal circular DNA and the sequence organization of one family of clones, which contain sequences enriched in HeLa small polydisperse circular (spc) DNA, was studied by restriction mapping and base sequence analysis. Restriction mapping revealed each clone to be composed solely of imperfect tandem repeats of ca. 170 bp. The entire DNA sequence of one clone was determined and found to be alphoid satellite with a variant monomeric construction.

INTRODUCTION

Small polydisperse circular (spc) DNA is a class of extrachromosomal circular DNA which is quite heterogeneous in size and sequence composition and seems to be ubiquitous in eucaryotic cells. The cell sources from which spcDNA has been identified include: human and boar spermatazoa (1); lymphoid tissues from humans (2,3), mice (4,5), and chickens (5,6); human warts (7); primary cell lines of mouse retina and liver (4); *Drosophila* embryos (8) and tissue culture (Schneider's) cells (8,9); HeLa cells (10); African green monkey (AGM) kidney BSC-1 cells (11,12,13); Chinese hamster ovary (CHO) cells (14); and numerous cell lines from mice (4,15). Sizes of these circular DNAs range from 150bp to over 20kb, with each cell source having its own characteristic size distribution. In the cell culture systems in which it has been determined, the total sequence complexity of a population of spcDNA is considerably greater than the average molecular mass, unlike eucaryotic cytoplasmic organelle associated DNAs and procaryotic plasmids, which consist of populations of DNA molecules of essentially identical sequence.

The most thoroughly characterized spcDNA is that from three established cell lines: *Drosophila* Schneider's line 2 cells (8,9), AGM BSC-1 cells (11,12,13), and HeLa cells (10). The small circular DNAs from Schneider's cells have a mean length of 3.3kb and range in size from approximately 270bp

to over 20kb. Exponentially growing or stationary Schneider's cells contain an average of 3 to 40 small circular DNA molecules per cell. Reassociation kinetics studies reveal two complexity classes: 82% of the spcDNA has a complexity of 1.8×10^4 and 18% has a significantly higher complexity (8). Stationary monkey BSC-1 cells contain an average of 1000 to 2000 spcDNA molecules per cell with a mean length of approximately 1.1kb and a total kinetic complexity of 1.5×10^5 (11). On average, exponentially growing HeLa cells contain 50 to 200 small circular DNA molecules ranging from 150bp to over 6kb in length with a mean size of about 1.5kb. HeLa cell spcDNA contains multiple complexity classes. 30-40% has a very low complexity, 20-30% has a moderate complexity, and 20-30% has a very high complexity (10). An enigmatic characteristic of spcDNA is its amplification in numbers per cell in response to treatment with specific drugs (e.g., cycloheximide or puromycin) or continued culture at stationary phase (9,10,11). With the exception of Schneider's cells, spcDNA is generally localized to the cytoplasm of fractionated cells. However, recent studies suggest that this may be due to leakage of the small DNA molecules from the nuclei during preparation (14,15).

Several lines of evidence suggest a chromosomal origin for most, if not all, spcDNA. In each of the well-characterized cell lines, the buoyant density of spcDNA is virtually indistinguishable from that of total chromosomal DNA. Reassociation kinetics experiments have shown the vast majority of spcDNA from Schneider's cells to be homologous to moderately repetitive genomic sequences (approximately 120 copies per haploid genome) (8). With one exception (14), all recombinant DNA clones derived from spcDNA have proven to be homologous to genomic sequences. The sequences so far identified in spcDNA clones encompass all genomic complexity classes: the highly repetitive alphoid satellite and Alu families from mammals (13); the copia family of transposable elements from *Drosophila* (16,17), the mammalian KpnI repeat family (Jones and Potter, in preparation), and as yet unidentified interspersed repetitive sequences from mammalian cells (12,14); as well as unique chromosomal sequences (12, Jones and Potter, in preparation). It is highly likely that spcDNA includes intermediates and/or by-products of interesting genetic events such as transposition, deletion and extrachromosomal amplification.

In this study, we have characterized a family of clones derived from HeLa cell spcDNA which is composed entirely of divergent alphoid satellite monomer repeats. We demonstrate that the sequences carried in these clones are enriched in spcDNA relative to genomic chromosomal DNA. By studying clones

whose sequences are relatively abundant in HeLa cell spcDNA, we hope to learn what subclass of genomic sequences are preferentially circularized, and from this to gain insight into the mechanism(s) of circularization.

MATERIALS AND METHODS

Cell culture

Seed cultures of S3-HeLa cells were kindly provided by D. Giard. HeLa cells were grown in suspension in MEM (Gibco) supplemented with 10% calf serum (Gibco) and 50 ug/ml Gentamycin (Schering).

Isolation of HeLa Extrachromosomal Circular DNA

Cycloheximide was added to 5 l of cell culture (3×10^9 cells) at a concentration of 50 ug/ml approximately 12 hours prior to harvesting. The cells were pelleted (GSA rotor, 2k, 5 min, 4°C) and washed twice with cold isotonic buffer (0.14 M NaCl, 5 mM KCl, 7 mM Na_2HPO_4 , 25 mM Tris-HCl [pH 7.4]), and then resuspended in 10x volume (of pellet) of cold hypotonic buffer (10 mM NaCl, 10 mM Tris-HCl [pH 7.4], 1 mM EDTA). After sitting on ice 10 min, the cell suspension was dounced to disrupt the cells. Sucrose was added to a final concentration of 0.2 M. Nuclei and unbroken cells were removed by centrifugation (SS34 rotor, 2.2 K, 5 min, 4°C). The crude cytoplasmic supernatant was removed and recentrifuged twice. 1/10 volume of 10% SDS was added to the supernatant, which was then incubated 20 min at 37°C . 1/15 volume of 3 M Na Acetate was added to the lysate, which was then extracted once each with neutralized phenol and chloroform. Nucleic acid was then precipitated with EtOH and resuspended in 2.5 ml of 1 mM Tris-HCl (pH 7.4), 1 mM EDTA on ice. The nucleic preparation was then subjected to a modification of Zasloff et al's (18) acid phenol extraction procedure as previously described for the purification of plasmid DNA (19). Following acid phenol extraction, the DNA was banded in a CsCl/EtBr isopycnic gradient. A single band was observed and collected by side puncture, the ethidium bromide was removed by isopropanol extraction, and the CsCl removed by dialysis against TE (10 mM Tris-HCl [pH 7.4], 1 mM EDTA). The nucleic acid was precipitated with EtOH and the pellet resuspended in 30 ul of TE.

HeLa Circle Library Construction

One third of the extrachromosomal circular DNA from a single preparation (2-5 ug DNA) was digested to completion with HindIII. The pBR322 DNA, after linearization with HindIII, was treated with calf intestinal phosphatase to prevent recircularization of plasmid molecules without inserts (20). After ligation of plasmid and HeLa cell DNA, and transformation of *E. coli* HB101

Nucleic Acids Research

(21) an aliquot was plated out onto ampicillin containing (50 ug/ml) agar plates to determine the initial size of the library.

Isolation of HeLa Chromosomal DNA

High molecular weight DNA was extracted from the crude nuclear pellet as described for the isolation of cytoplasmic supercoiled DNA, with the exception that these cells had not been treated with cycloheximide. The pellet was re-dounced and re-pelleted (SS34 rotor, 2.2 K, 5 min, 4°C). 10% SDS was added to 1% in addition to proteinase K (approximately 200 ug/ml) and incubated 15 min at 37°C. Due to great viscosity, the lysate was diluted 1:5 with Final Buffer (0.15 M Na Acetate, 0.05 M Tris-HCl [pH 7.4], 0.025 M EDTA, 0.25 M sucrose) and extracted with neutralized phenol twice and once with chloroform, then EtOH precipitated and resuspended in TE.

Purification of Plasmid and Bacteriophage DNA

Plasmid DNAs were purified by the rapid boil method of Holmes and Quigley (22), with the modifications previously described (23). This DNA was excellent for restriction mapping and sequencing, but it contained an inhibitor of the Bal31 exonuclease which could be removed by CsCl/EtBr isopycnic centrifugation.

Lambda bacteriophage DNAs were prepared from mini (4.0 ml) lysates as described by Maniatis et al (20).

Restriction Mapping

Restriction sites were located by a combination of end labelling and partial restriction enzyme digestion as described by Smith and Birnstiel (24).

DNA Sequencing

DNA sequencing was performed according to the partial chemical cleavage method of Maxam and Gilbert (25), using the G, A+G, T+C, A+C, and C only reactions. DNA was labelled at the 5' ends by polynucleotide kinase and at the 3' ends by use of the Klenow fragment of DNA polymerase I.

Introduction of Restriction Sites

In order to facilitate the DNA sequencing, restriction sites were artificially introduced into clone pc21 using Bal31 exonuclease, as previously described (26). pc21 was opened at the Cla I site of pBR322 and a new Cla I site was introduced approximately 250 bp in from the left end of the insert following Bal31 digestion.

Filter Hybridization

Nitrocellulose filters were pre-soaked, hybridized, and washed post-hybridization in 1x Denhardt's (0.02% bovine serum albumin, 0.02% polyvinylpyrrolidone, 0.02% ficoll), 4x SSCP (0.6 M NaCl, 0.06 M Na citrate, 0.08 M Na phosphate, pH 7.0) with shaking at 65°C. The presoak was for 4 to 8

hours and the hybridization for 12 to 15 hours. Denatured salmon sperm DNA (200 ug/ml) was included in the hybridization. Following hybridization, filters were washed with 2 to 3 l of 1x Denhardt's, 4x SSCP at 65°C with at least 6 changes of solution and then with 2x SSC (0.3 M NaCl, 0.03 M Na citrate) for 30 minutes at room temperature. Where specified, Blur8 plasmid DNA (27) and/or HeLa chromosomal DNA were nicked with mild DNAase I treatment, denatured, and added to the hybridization at a concentration of approximately 6 ug/ml.

To remove hybridized probe from filters, which had been kept moist during autoradiographic exposure, they were washed in 0.5 M NaOH, then 0.5 M Tris-HCl (pH 7.4), and finally with 2x SSC. Filters were washed in each solution for 2 minutes with 6 to 8 changes. Following the wash with 2x SSC, filters were ready to pre-soak for another hybridization. Approximately 90% of hybridized probe can be removed with this procedure.

MATERIALS

Enzymes were purchased from Bethesda Research Labs and New England Biolabs and isotopes were from New England Nuclear. Acrylamide was from Biorad and agarose and low gel temperature agarose were from Sigma.

RESULTS

Isolation and Cloning of Extrachromosomal Circular DNA from HeLa Cells

Extrachromosomal circular DNA was purified from approximately 10⁹ cycloheximide treated HeLa cells by removing the nuclei and then subjecting the remaining cytoplasmic supernatant to acid phenol extraction. The covalently closed circular DNA was further purified by isopycnic centrifugation in a CsCl/EtBr gradient. A single band of DNA was observed and collected.

Restriction analysis of the DNA preparation showed that the predominant DNA species was mitochondrial DNA (mtDNA), as expected. Electron microscopy revealed the presence of numerous small circular DNA molecules, as well as a few linear molecules.

In order to study individual species of spcDNA, the extrachromosomal circular DNA (including mtDNA) was cloned into the plasmid pBR322. One third of the DNA from a single preparation was digested to completion with HindIII, ligated into the HindIII site of pBR322, and then used to transform E. coli HB101 cells. Although many of the spcDNA species (average size is about 1.5 kb [10]) will not be cleaved by HindIII, and therefore will not be cloned, the

advantage of this strategy is that most of the molecules that are cleaved are only cut once, thus allowing the purification of complete circles. The initial size of the library was determined to be approximately 5×10^4 independent transformants, with approximately 10% containing no insert in the pBR322 (still resistant to tetracycline). Of the clones with inserts, approximately 90% contained mtDNA and 10% carried non-mtDNA (and therefore presumably spcDNA) inserts. This was determined by doing "mini"-plasmid preparations of more than 50 randomly selected tetracycline sensitive clones, digesting the plasmid of each with HindIII, and electrophoresing to size the inserts.

Isolation of Human Genomic Clones Containing DNA Sequences Enriched in HeLa Cell spcDNA

We were interested in identifying chromosomal sequences that are preferentially circularized. We anticipated that the mechanism of circle formation might be revealed by analyzing the organization of these preferentially circularized chromosomal sequences, and by then comparing corresponding circle and chromosomal DNAs. We employed a method of duplicate screening to identify those sequences highly enriched in the spcDNA. The degree to which a cloned DNA sequence hybridizes to a complex probe reflects the abundance of that sequence in the probe. By comparing the hybridization of a cloned sequence to two different complex probes, such as HeLa chromosomal DNA vs HeLa extrachromosomal circular DNA probes, it was possible to identify sequences enriched in spcDNA relative to total genomic DNA. Therefore, we first screened a human genomic library (kindly provided by T. Maniatis [28]) sequentially with HeLa extrachromosomal circular DNA probe (a sample of the DNA preparation from which the circle DNA library was constructed) and then (after removing the hybridized probe with alkaline wash treatment) with HeLa chromosomal DNA probe. Cold competitor DNAs were used to decrease hybridization of sequences which are highly repetitive in the human genome (e.g., Alu elements are present in about 5×10^5 copies per haploid genome and dispersed so that greater than 90% of genomic clones contain at least one copy) so that differential hybridization of less repetitive sequences might be detected. Cold HeLa chromosomal DNA and Blur8 plasmid DNA (a cloned member of the Alu family of human repetitive elements [27]) were denatured and added to both hybridizations (6 ug/ml of each). Twelve clones were selected which appeared to show relatively greater hybridization to the circle probe, based on comparison with other clones on the same filters, than to the chromosomal DNA probe. These twelve chromosomal DNA clones were therefore postulated to

Fig. 1. Ethidium stained gel and autoradiographs showing restriction fragments homologous to sequences abundant in HeLa chromosomal DNA and HeLa spcDNA. Mini (4.0 ml) lysates were prepared from human genomic lambda phage clones which appeared to contain sequences preferentially represented in HeLa spcDNA. Phage DNA was prepared from each mini lysate, triple restriction digested with Eco RI, BglII and PvuII, electrophoresed in (a) duplicate 1.0% agarose gels, Southern transferred and hybridized to nick translated (b) HeLa chromosomal DNA or (c) HeLa extrachromosomal circular DNA probes. Both hybridizations contained 6 ug/ml HeLa chromosomal DNA as cold competitor. (*) Fragments in clone H11 (lane f) showing preferential hybridization to the extrachromosomal circular DNA probe. (←) Restriction fragment from H11 subsequently used as probe. (A) Restriction fragments which hybridized most intensely to Blur8 probe (fig. 2a).

carry sequences enriched in our preparation of spcDNA and therefore preferentially circularized.

In order to confirm that these genomic clones contained sequences enriched in spcDNA, and to identify restriction fragments responsible for the differential hybridizations, DNA was prepared from "mini" (4.0ml) lysates of the 12 positive clones. An aliquot of each DNA was triple digested with BglII, EcoRI, and PvuII, and equal amounts of each clone's digest were run in duplicate in 1.0% agarose gels and Southern blotted. One of the EtBr stained gels is shown in figure 1a. The filters were then hybridized separately to either the HeLa chromosomal DNA probe or the HeLa circular DNA probe (figures 1b and 1c respectively). HeLa chromosomal DNA was added to these hybridizations as the sole competitor (6 ug/ml), to dampen the hybridization of the more repeated sequences. In other words the effect of the competitor is to slightly reduce the differences in signal produced by sequences of different repetition. It will not, however, allow the signals produced by less repeated sequences to surpass those of more highly repeated sequences and cannot, therefore, account for the dramatic differences seen in figure 1. Six

Fig. 2. Autoradiographs showing restriction fragments of human genomic clones in fig. 1 that are homologous to (a) Blur8 and (b) H11. Hybridized probes were removed from the Southern transfers shown in fig. 1 by alkaline wash treatment, and the filters rehybridized to nick translated (a) Blur8 or (b) a restriction fragment from H11 (↔).

of the 12 clones (lanes b to g, figures 1b and 1c) very clearly showed relatively more intense hybridization to the circle DNA probe than to the chromosomal DNA probe. Note that in figure 1b the bands in lane j, which are homologous to Alu repeats (identified by hybridization to Blur8 probe in figure 2a and designated by 'A's in figures 1b and 1c) are much more intense than the bands in lane f (designated with asterisks). However, in figure 1c the top two bands in lane f are significantly more intense than the bottom Alu band in lane j, indicating enrichment of their sequences in the circle DNA probe relative to the chromosomal DNA probe.

Clone H11 (figure 1, lane f), with chromosomal DNA sequences that are enriched in spcDNA, was selected for further study. First, in order to determine if any of the other chromosomal clones carrying sequences enriched in spcDNA were homologous to H11, the following experiment was done. A restriction fragment (marked by an arrow in figure 1c) was isolated from an agarose gel, nick translated, and hybridized to the alkaline-washed filter in figure 1b. As can be seen in figure 2b, at least two other clones (lanes e and h) contain sequences homologous to H11.

Isolation and Characterization of HeLa spcDNA clones homologous to Human Genomic Clone H11

In order to isolate homologous spcDNA clones, this same H11 probe was used to screen the HeLa circle DNA library. Five homologous clones were isolated. Conventional restriction mapping proved impossible since

Fig. 3. Restriction maps of clones isolated from a HeLa circular DNA library carrying sequences homologous to human genomic clone H11. A plasmid (pBR322) library was constructed from HeLa extrachromosomal circular DNA and screened with H11 probe ($\leftarrow$ in figures 1 and 2). Each clone was mapped using the end labelling, partial restriction digestion method (22). Insert sizes for pc6, pc23 and pc21 are 3.4, 2.1 and 0.85 kb respectively.

restriction enzymes either did not hit the inserts at all or produced a small number of low molecular weight bands, which consisted of multiple co-migrating restriction fragments. These results suggested an internally repetitious sequence organization. The end-labelling, partial restriction digestion method of Smith and Birnstiel (24) was used to construct restriction maps of the three clones illustrated in figure 3 (two other clones were isolated which appeared to be identical to pc6 from preliminary restriction enzymes digestions). Each clone is constructed solely of imperfect tandem direct repeats of ca. 170bp. Comparison with the consensus human EcoRI-alphoid satellite dimer repeat sequence of 340bp (29,30) suggested that these clones might be divergent alphoid satellite repeats. Alphoid satellite DNA is a complex satellite found in primates with a monomer repeat of about 170bp (31-37, for review see ref. 38). In humans, the majority of alphoid satellite is arranged as a dimer of 340bp that is made up of two subunits that are about 27% divergent from each other, with only about 1% divergence between dimers (30). To further investigate the possible relatedness of our circle DNA clones to alphoid satellite, we sequenced clone pc21 (852bp in length) using the partial chemical cleavage method of Maxam and Gilbert (25). 98% of the clone was sequenced on both strands according to the sequencing strategy

Fig. 4. Sequencing strategy for clone pc21. Each arrow represents a separate end-labelling experiment. Open circle tails indicate 5' end labelling by polynucleotide kinase and closed circle tails indicate 3' end labelling using the Klenow fragment. One restriction site (*) was introduced by the method of Bal 31 *in vitro* mutagenesis as described in Materials and Methods.

in figure 4. Clone pc21 is composed of five tandem imperfect repeats of 171bp, 171bp, 171bp, 169bp, and 170bp (arranged in register in figure 3b), which are approximately 20 to 25% divergent from each other in all pairwise combinations and approximately 25% divergent from the consensus human alphoid satellite dimer subunits. This is the first discovered DNA sequence of divergent monomeric human alphoid satellite. A similar monomeric alphoid satellite construction was subsequently found in a HeLa genomic clone (39) isolated due to its homology to pc6.

	10	20	30	40	50	60	70	80	90				
I	AATTC	TCAGTAACTC	CTTGTG	TGTGTG	TATTC	AACTC	CACAGAGT	TGAACG	ATCCTTT	TACACAGAGCAGAGCT	TGAAACACTC	TTTTGT	
II	GATTC	CAGAACTC	CTTGTG	TGATGTG	TGCGGT	TCAACT	CACAGAGT	TAACT	CTTTCTTT	CATAGACAGCT	TAGGAAACACTC	TGTTGT	
1	CTTCT	GAGAACTC	TTTGTG	TGATGTG	TGCTCT	CATCT	CACAGAGT	TGAACCT	TTCTTTG	ATTG	TGAACCTTTG	GAAACAGTA	TTTTGA
2	CATTT	GAGAAA	CTTGTG	ATGCACA	CTTCA	ATT	TCACAGAGT	TGAAT	CTTTCTTT	TGAATGAGCAGCT	TAGAAACAGTC	TTTTGT	
3	CATTC	TGAGAACTC	CTTGTG	TATATGC	CAATC	ATCTC	CACAGAGT	TGAACCT	TTCTTTT	GATTG	CAGCAGTTT	TGAAACAGTC	TTTTGT
4	ATTTCT	GAGAACTC	TTTGTG	TGATGTG	CTTTCA	CTC	CACAGT	TGAACCT	TTCTTTT	GATTG	CAGCAGTTT	TGAAACAGTC	TTTTGT
5	ATTTCT	GAAAACTC	TTTGTG	TGATGTG	GAATTC	ACGTC	CACAGAA	TCAACCT	TTCTTT	TGAGT	TGACGAGTT	TGAAACAGTC	TTTTGT

	100	110	120	130	140	150	160	170				
I	GGAATT	TGCAAGTGC	GATTTTCAG	CCGC	TTTGAGGT	CAATGCT	AGAAATAT	CTTCTAT	GAAACTAG	CACAAT	TAGACAGAAT	
II	AAAGCT	CTCAAGTGC	ATATTCAG	CCTC	TTTGAGGC	CTTGTG	TGAAACGG	GATT	TCTTCAT	ATTATG	CTAGACAGAAG	
1	AGAATC	TCAAAAGCG	GATATTTGGA	GAGC	TTTGAGGC	CTACAGT	TAAAAACGG	AAATATAT	GCACAT	AAAAAC	GAAGACAGAAG	171
2	ACAATC	TGCAAAAGCG	GATATTTGCT	AGCC	CAATTGAGGC	CTTATGCT	TGAAAA	CGAAATAT	CTGCACAT	AAAACT	GTACACAGAAG	342
3	AGAATC	TGCAA	GCATATTTGCT	AGCC	CTTTATG	CCCTGCT	GAAATAT	GAAATAT	CTTCACAT	AAAACT	GTACACAGAAG	513
4	AGAATC	TGCAAAAGCG	GATATTTGGA	GCTA	TTTGAGGC	CTTGTG	TGAAAAAG	AAAGTAT	CTTCACAT	AAAACT	GTACACAGAAG	682
5	AGAAGCT	GCAAAAGG	AAATTTCT	TAGCG	TTTGAGGC	CTATAGT	AAAAAGAAA	TATCTT	CACAT	AAAACT	GTACACAGAAG	852

Fig. 5. Base sequence of clone pc21 and comparison with the human consensus alphoid satellite dimer repeat. The entire base sequence of pc21, beginning with the left HindIII site, is presented so that each monomer repeat is aligned with the consensus repeats. The 171bp and 169bp subunits of the consensus human EcoRI delineated dimer repeat are designated I and II, respectively (sequence of the consensus dimer repeat was determined by Wu and Manuelidis [30]). The five divergent monomeric repeats of pc21 are designated 1, 2, 3, 4 and 5 respectively. The numbers above the sequence refer to the position in the first (I) human consensus repeat and the numbers to the right indicate the distance from the left HindIII site in pc21.

Fig. 6. Southern transfer of undigested HeLa extrachromosomal circular DNA showing size distribution of alphoid satellite containing circles. Undigested HeLa extrachromosomal circular DNA was electrophoresed in a 1.0% agarose gel, Southern transferred to nitrocellulose, and hybridized with nick translated pc6 insert probe. The sizes on the left were derived from linear size markers and therefore do not correspond directly to the sizes of the hybridizing molecules, which were probably supercoiled or nicked circles.

Size Distribution of alphoid satellite DNA in HeLa spcDNA

Bertelsen *et al* (12) have previously shown that alphoid satellite DNA occurs as an oligomeric series of discretely sized closed circular DNAs in AGM BSC-1 cells, which appear to be multimers of the 172bp satellite repeat unit. Alphoid satellite constitutes about 20% of the AGM genome (compared to about 2% of the human genome) and exists as tandem arrays of 172 bp monomers, with only about 3% divergence of unit repeats from the consensus sequence (40,41).

To determine the size distribution of alphoid satellite in the spcDNA of HeLa cells (in which the spcDNA had been amplified with cycloheximide treatment), we hybridized an alphoid satellite probe (pc6) to a Southern blot of undigested HeLa extrachromosomal circular DNA (figure 6). In contrast to Bertelsen *et al*'s (12) results, we primarily see a smear of hybridization indicating very heterogeneously sized alphoid satellite circles. Although somewhat obscured by this smear, several distinct bands can be seen in the upper portion of figure 6, suggesting that at least some discrete size classes do exist.

DISCUSSION

We have shown that certain chromosomal sequences are highly enriched in the spcDNA of cycloheximide-treated HeLa cells. Several spcDNA clones of one

such sequence family were analyzed by restriction mapping and the entire DNA sequence of one clone was determined. Each clone was found to be composed entirely of divergent monomeric alphoid satellite repeats.

Alphoid satellite is a complex satellite found in a wide variety of primates (see ref. 38 for review). Unlike the prototypical simple sequence satellite DNAs, alphoid satellite is composed of longer repeat units of about 170bp with no significant internal redundancies. The base composition of alphoid satellite is very similar to that of total chromosomal DNA. It is classified as satellite DNA due to its tandem repeat construction and centromeric localization (32,42). In humans, the majority of alphoid satellite is arranged as a 340bp dimer with 171 and 169bp subunits that are 27% divergent from each other, with only about 1% divergence between dimers (29,30). In contrast, the tandem array of alphoid satellite repeats in clone pc21 show no dimeric construction. Each monomer repeat is 20-25% divergent from the other four repeats as well as from each of the two subunits of the consensus dimer (30).

If unequal crossing over, as proposed by Smith (43) and experimentally supported by a number of investigators (44,45), is an important mechanism in the maintenance of homogeneity in tandem arrays of satellite repeats, then those repeats near the ends of satellite would be less efficiently rectified and tend to become more divergent. It is therefore possible that the divergent monomeric satellite described here was derived from terminal regions of satellite. It should be noted that we cannot determine the relative amounts of satellite with conserved dimeric vs divergent monomeric repeats in spcDNA, since alphoid satellite with a conserved dimeric construction was inadvertently selected against in the cloning process due to lack of a HindIII recognition site. Location of the HindIII sites in clone pc21, at the junction between two repeats, supports, but does not prove the hypothesis that pc21 originated as a circular DNA molecule of divergent alphoid-satellite sequence that was linearized at a single HindIII site.

The existence of extrachromosomal circular satellite DNA might explain how identical satellite sequences can exist on nonhomologous chromosomes and at non-centromeric locations. Once circularized, satellite sequences could be randomly inserted elsewhere in the genome, where it could then be eliminated or amplified by unequal crossing over and acted upon by natural selection. Also, extrachromosomal satellite sequences could participate in gene conversion events.

Cycloheximide induced amplification of spcDNA does not involve de novo

synthesis of DNA, but rather the conversion of pre-existing DNA, presumably chromosomal, from linear to a closed circular form (10). Based on these results, Smith and Vinograd (10) proposed homologous recombination between directly oriented repeats as a likely circularization mechanism. The products of such intrachromosomal recombination events would be a deleted chromosome, now containing only one copy of the repeat, and a circular DNA molecule containing one repeat and the sequence flanked by the repeats.

In support of such a mechanism, others have discovered deletion hot spots in E. coli (46) and transposable element deletions in eucaryotes (47) to be flanked by direct repeats. In addition, chromatin diminution in Ascaris suum involves mainly the loss of satellite DNA sequences (48), and human diploid fibroblasts have been reported to lose approximately 24% of their alphoid satellite during their finite replicative life span in vitro (49). Although other circularization mechanisms such as replicon misfiring (50) or reverse transcription (51) may also be active in spcDNA formation, virtually every cloned spcDNA molecule, which has been sufficiently characterized, could possibly have arisen via recombination. This includes sequence analyses of clones containing members of the Alu repeat family (13) and alphoid satellite sequence (this report and ref. 12) from primate cell lines, and an extensive examination of cloned circular copies of the transposable element copia from Drosophila cell lines (16,17). The circular copia molecules so far studied are capable of extrachromosomal replication under positive selective pressure (52), but do not appear to be functional transposition intermediates (17). If homologous intrachromosomal recombination is primarily responsible for spcDNA formation, sequences with an abundance of tandem repeats, such as satellite, should be preferentially circularized.

Nevertheless, the discovery that alphoid satellite sequences were enriched in HeLa cell spcDNA was somewhat surprising in light of previously published work in which Bertelsen et al (12) reported alphoid satellite to be underrepresented in the spcDNA of monkey RSC-1 cells. The apparent discrepancies in our results could reflect inherent differences in the cell lines or differences in treatment of cells prior to harvesting. RSC-1 cells contain a pauci-disperse population of spcDNA with several prominent size classes, whereas HeLa cell spcDNA has a very heterogeneous size distribution. These size classes in RSC-1 cells do not seem to be composed simply of individual species of circles. Still, abundance classes could represent subsets of sequences which have been clonally amplified in RSC-1 cells. Such amplification would reduce the relative representation of other sequences,

which have a heterogeneous size distribution, such as is the case for alphoid satellite (fig. 6, and ref. 12). Alternatively, differences in treatment of cells prior to harvesting might alter the sequence composition of the spcDNA population. Bertelsen et al (12) isolated spcDNA from stationary BSC-1 cells, whereas we treated the HeLa cells with cycloheximide. Cycloheximide is a potent inhibitor of translation and also blocks DNA synthesis in late S-phase (53,54). Discontinuous DNA replication has been shown to be ubiquitous for at least the lagging chromosome arm in eucaryotes (for review see ref. 55). "Okazaki fragments" of 40-290 nucleotides are initially synthesized which then join to form intermediate (replicon) sized DNA chains, which subsequently elongate as replicons meet. It is this second elongation step which cycloheximide seems to block (54). Suspended in this pre-replicon-fusion state, single stranded DNA ends would be free to initiate intrachromosomal recombination. DNA sequences such as satellite, which contain an abundance of directly oriented repeats, would then be particularly prone to cycloheximide induced amplification. Alternatively, cycloheximide-amplified spcDNA may be a by-product of cell death and chromosome degradation. That cycloheximide amplification might produce changes in the spcDNA population is supported by studies in which the spcDNA of Drosophila culture cells is reduced from a mean size of 3.3 kb to 1.3 kb upon amplification (9). Nevertheless, the buoyant density of this amplified spcDNA remains unaltered, so the change is probably not due to circularization of just satellite sequences. It may be that multiple mechanisms are responsible for spcDNA formation, but that only one or a few mechanisms are enhanced by cycloheximide treatment. That subset of sequences most prone to circularization by that particular mechanism(s) would then be preferentially circularized.

ACKNOWLEDGEMENTS

This work was supported by grant GM30816 from the N.I.H. and by grant CA 00852 from the N.C.I..

REFERENCES

1. Hotta, Y. and Bassel, A. (1965) *Proc. Nat. Acad. Sci. (USA)* 53, 356-362.
2. Calabretta, B., Robberson, D.L., Barrera-Saldana, H.A., Lambrou, T.P. and Saunders, G.F. (1982) *Nature* 296, 219-225.
3. Schmookler-Reis, R.J., Lumpkin Jr, C.K., McGill, J.R., Riabowol, K.T. and Goldstein, S. (1983) *Nature* 301, 394-398.
4. Yamagishi, H., Kunisada, T. and Tsuda, T. (1982) *Plasmid* 8, 299-306.
5. Tsuda, T., Yamagishi, H., Ohnishi, N., Yamada, Y., Izumi, H. and Mori, K.J. (1983) *Plasmid* 10, 235-241.

6. Delap, R.J. and Rush, M.G. (1978) *Proc. Nat. Acad. Sci. (USA)* 75, 5855-5859.
7. Delap, R.J., Zouzias, D. and Rush, M.G. (1976) *Biochim. Biophys. Acta* 418, 257-265.
8. Stanfield, S.W. and Lengyel, J.A. (1979) *Proc. Nat. Acad. Sci. (USA)* 76, 6142-6146.
9. Stanfield, S.W. and Helinski, D.R. (1976) *Cell* 9, 333-345.
10. Smith, C.A. and Vinograd, J. (1972) *J. Mol. Biol.* 69, 163-178.
11. Delap, R.J., Rush, M.G., Zouzias, D. and Khan, S. (1978) *Plasmid* 1, 508-521.
12. Bertelsen, A.H., Humayun, M.Z., Karfopoulos, S.G. and Rush, M.G. (1982) *Biochemistry* 21, 2076-2085.
13. Krolewski, J.J., Bertelsen, A.H., Humayun, M.Z. and Rush, M.G. (1982) *J. Mol. Biol.* 154, 399-415.
14. Stanfield, S.W. and Helinski, D.R. (1984) *Mol. Cell. Biol.* 4, 173-180.
15. Kunisada, T., Yamagishi, H. and Sekiguchi, T. (1983) *Plasmid* 10, 242-250.
16. Flavell, A.J. and Ish-Horowicz, D. (1981) *Nature* 292, 591-595.
17. Flavell, A.J. and Ish-Horowicz, D. (1983) *Cell* 34, 415-419.
18. Zasloff, M., Grindler, G.D. and Felsenfeld, G. (1978) *Nucl. Acids Res.* 5, 1139-1152.
19. Truett, M.A., Jones, R.S. and Potter, S.S. (1981) *Cell* 24, 753-763.
20. Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) *Molecular Cloning*, p. 133, Cold Spring Harbor Laboratory.
21. Potter, S.S., Brorein, W.J., Dunsmeir, P. and Rubin G.M. (1979) *Cell* 17, 415-427.
22. Holmes, D.S. and Quigley, M. (1981) *Analytical Biochem.* 114, 193-197.
23. Potter, S.S. (1982) *Mol. Gen. Genet.* 188, 107-110.
24. Smith, H.O. and Birnstiel, M.L. (1976) *Nucl. Acids Res.* 3, 2387-2398.
25. Maxam, A.M. and Gilbert, W. (1980) *Meth. Enzym.* 65, 499-560.
26. Potter, S.S. (1982) *Nature* 297, 201-204.
27. Rubin, C.M., Houck, C.M., Deininger, P.L., Friedman, T. and Schmid, C.W. (1980) *Nature* 284, 373-374.
28. Lawn, R.M., Fritsch, E.F., Parker, R.C., Blake, G. and Maniatis, T. (1978) *Cell* 15, 1157-1174.
29. Manuelidis, L. and Wu, J.C. (1978) *Nature* 296, 92-94.
30. Wu, J.C. and Manuelidis, L. (1980) *J. Mol. Biol.* 142, 363-386.
31. Maio, J.J. (1971) *J. Mol. Biol.* 56, 579-595.
32. Kurnit, D.M. and Maio, J.J. (1974) *Chromosoma* 45, 387-400.
33. Donehower, L. and Gillespie, D. (1979) *J. Mol. Biol.* 134, 835-842.
34. Gillespie, D. (1977) *Science* 196, 889-891.
35. Singer, D.S. and Donehower, L. (1979) *J. Mol. Biol.* 134, 835-842.
36. Musich, P.R., Brown, F.L. and Maio, J.J. (1980) *Chromosoma* 80, 331-348.
37. Maio, J.J., Brown, F.L. and Musich, P.R. (1981) *Chromosoma* 83, 103-125.
38. Singer, M.F. (1982) *Int. Rev. Cytol.* 76, 67-112.
39. Potter, S.S. and Jones, R.S. (1983) *Nucl. Acids Res.* 11, 3137-3153.
40. Rosenberg, Singer, M.F. and Rosenberg, M. (1978) *Science* 200, 394-402.
41. Thayer, R.E., Singer, M.F. and McCutchan, T. (1981) *Nucl. Acids Res.* 9, 169-181.
42. Manuelidis, L. (1978) *Chromosoma* 66, 23-32.
43. Smith, G.P. (1976) *Science* 191, 528-535.
44. Petes, T.D. (1980) *Cell* 19, 765-774.
45. Szostak, J.W. and Wu, R. (1980) *Nature* 284, 426-450.
46. Albertini, A.M., Hofer, M., Calos, M.P. and Miller, J.H. (1982) *Cell* 29, 319-328.
47. Roeder, S.G. and Fink, G.R. (1983) In "Mobile Genetic Elements",

- Shapiro, J.A., Ed., pp. 299-327. Academic Press, Inc., New York, New York.
48. Roth, G.E. and Moritz, K.B. (1981) *Chromosoma* 83, 169-190.
 49. Schmookler Reis, R.J. and Goldstein, S. (1980) *Cell* 21, 739-749.
 50. Varshavsky, A. (1981) *Proc. Natl. Acad. Sci.* 78, 3673-3677.
 51. Gilboa, E., Mitra, S.W., Goff, S. and Baltimore, D. (1979) *Cell* 18, 93-100.
 52. Sinclair, J.H., Sang, J.H., Burke, J.E. and Ish-Horowicz, D. (1983) *Nature* 306, 198-200.
 53. Hyodo, M., Koyama, H. and Ono, T. (19) *Exptl. Cell. Res.* 67, 461-463.
 54. Schwartzman, J.B. and Van't Hof, J. (1982) *Nucl. Acids Res.* 10, 6191-6205.
 55. DePamphilis, M.L. and Wassarman, P.M. (1980) *Ann. Rev. Biochem.* 49, 627-666.