the two adjacent base pairs of DNA and consequently its relative binding affinity and unwinding angle.

References

Bauer, W. R. (1978) *Annu. Rev. Biophys. Bioeng. 7*, 278.

Bloomfield, V., Crothers, D. M., & Tinoco, I. (1974) *Physical Chemistry of Nucleic Acids*, pp 406–420, Harper and Row, New York.

Cohen, G., & Eisenberg, H. (1969) *Biopolymers 8*, 45.

Davidson, M. W., Griggs, B. G., Lopp, I. G., Boykin, D. W., & Wilson, W. D. (1978) *Biochemistry 17*, 4220.

Deranleau, D. (1969) *J. Am. Chem. Soc. 91*, 4044.

Early, T. A., Kearns, D. R., Hillen, W., & Wells, R. D. (1981) *Biochemistry 20*, 3756, 3764.

Englander, S. W., Kallenbach, N. R., Heeger, A. J., Krumhansl, J. A., & Litwin, S. (1980) *Proc. Natl. Acad. Sci. U.S.A. 77*, 7222.

Gabbay, E. J. (1977) *Bioorg. Chem. 3*, 33.

Gabbay, E. J., Grier, D., Fingerle, R. E., Reimer, R., Levy, R., Pearce, S., & Wilson, W. D. (1976) *Biochemistry 15*, 2062.

Henry, D. W. (1976) in *Cancer Chemotherapy* (Sartorelli, A. C., Ed.) pp 15–57, American Chemical Society, Washington, DC.

Howe-Grant, M., & Lippard, S. J. (1979) *Biochemistry 18*, 5762.

Jones, R. L., & Wilson, W. D. (1980) *J. Am. Chem. Soc. 102*, 7776.

Jones, R. L., Lanier, A. C., Keel, R. A., & Wilson, W. D. (1980) *Nucleic Acids Res. 8*, 1613.

Mandel, C., Kallenbach, N. R., & Englander, S. W. (1979) *J. Mol. Biol. 135*, 391.

Manning, G. S. (1978) *Q. Rev. Biophys. 11*, 179.

McGhee, J. D., & von Hipple, P. H. (1974) *J. Mol. Biol. 86*, 469.

Müller, W., & Crothers, D. M. (1968) *J. Mol. Biol. 35*, 251.

Müller, W., & Crothers, D. M. (1975) *Eur. J. Biochem. 54*, 267.

Patel, D. J., & Canuel, L. L. (1978) *Eur. J. Biochem. 90*, 247.

Pigram, W. J., Fuller, W., & Hamilton, L. D. (1972) *Nature (London), New Biol. 235*, 17.

Pulleyblank, D. E., & Morgan, A. R. (1975) *J. Mol. Biol. 91*, 1.

Quigley, G. J., Wang, A. H.-J., Ughetto, G., van der Marel, G., van Boom, J., & Rich, A. (1980) *Proc. Natl. Acad. Sci. U.S.A. 77*, 7204.

Record, M. T., Jr., Anderson, C. F., & Lohman, T. M. (1978) *Q. Rev. Biophys. 11*, 103.

Revet, B., Schmir, M., & Vinograd, J. (1971) *Nature (London), New Biol. 229*, 10.

Sobell, H. M. (1973) *Prog. Nucleic Acid Res. Mol. Biol. 13*, 153.

Sobell, H. M., Reddy, B. S., Bhandray, K. K., Jain, S. C., Sakre, T. D., & Seshadri, T. D. (1977) *Cold Spring Harbor Symp. Quant. Biol. 42*, 87.

Wang, J. C. (1974) *J. Mol. Biol. 89*, 783.

Waring, M. J. (1972) in *The Molecular Basis of Antibiotic Action* (Gale, E. F., Cundliffe, E., Reynolds, P. F., Richmond, M. H., & Waring, M. J., Eds.) pp 173–277, Wiley, London.

Wilson, W. D., & Lopp, I. G. (1979) *Biopolymers 18*, 3025.

Wilson, W. D., & Jones, R. L. (1981) *Adv. Pharmacol. Chemother. 18*, 177.

Wilson, W. D., & Jones, R. L. (1982) in *Intercalation Chemistry* (Whittingham, S., & Jacobson, A., Eds.) Chapter 14, Academic Press, New York (in press).

Molecular Characterization of Small Polydisperse Circular Deoxyribonucleic Acid from an African Green Monkey Cell Line†

Arthur H. Bertelsen,‡ M. Zafri Humayun,* Stefan G. Karfopoulos, and Mark G. Rush

ABSTRACT: Several size classes of small polydisperse circular (spc) DNA from the African green monkey cell line BSC1 have been cloned into the bacterial plasmid pBR322. Analysis of the cloned spc DNA fragments as well as total spc DNA reveals that (a) most or all cloned spc DNAs share homologies with chromosomal sequences, (b) both unique and repetitive chromosomal sequences are represented in spc DNA, (c) the repetitive sequences in spc DNA include two known major repeat families (the α and the Alu) as well as a third, as yet unidentified, set of interspersed repetitive sequences, and (d) the α-like sequences are present in an oligomeric series of circular DNA molecules within the spc DNA population. The organizational features of repetitive DNA sequence-carrying circles suggest a mechanism for their generation.

It has been known for about a decade that many eukaryotic cells in culture (as well as some cells isolated from animal tissues) contain small polydisperse circular (spc) DNAs (Radloff et al., 1967; Smith & Vinograd, 1972; Stanfield & Helinski, 1976; Delap et al., 1978; Delap & Rush, 1978). The best characterized of the spc DNAs are those found in three lines of cultured cells: (1) HeLa from human, (2) Schneider line 2 from *Drosophila melanogaster*, and (3) BSC1 from African green monkey kidney. Table I summarizes the known physical characteristics of spc DNA isolated from these three sources. The common features of the spc DNA from these cell lines are the following: (1) average sizes in the range of 0.8–2.5 kilobase pairs (kbp), (2) copy numbers ranging from a few to perhaps a few thousand per cell, (3) the presence of

† From the Department of Microbiology, University of Medicine and Dentistry of New Jersey—New Jersey Medical School, Newark, New Jersey 07103 (M.Z.H. and S.G.K.), and the Department of Biochemistry, New York University Medical Center, New York, New York 10016 (A.H.B. and M.G.R.). *Received October 20, 1981.* This work was supported by a National Institutes of Health Biomedical Research Support grant (5SO7RR05393) to M.Z.H. and an American Cancer Society grant (NP 313A) to M.G.R. A.H.B. was a predoctoral trainee supported by a U.S. Public Health Service training grant (5T32GM 07238-04).

‡ Present address: Unigene Laboratories, Nutley, NJ 07110.

0006-2960/82/0421-2076$01.25/0

Table I: Properties of spc DNA from Three Sources

organism	copy no.	av mol mass (megadaltons) [contour length (μm)]	buoyant density (g/cm³)	reassociation kinetics	amplification: stationary phase	amplification: cycloheximide treatment	ref
human (HeLa)	50–200[a]	1.0 (0.32)	1.692[d]	20–30% renatures with kinetic complexity of 3×10^6 bp	yes[e]	yes[e]	Smith & Vinograd (1972)
Drosophila (Schneider line 2)	3–40[a]	2.2 (1.1)[c]	1.703[d]	85% renatures rapidly with kinetic complexity of 2×10^4 bp, 15% renatures slowly	no	yes[f]	Stanfield & Helinski (1976)
monkey (BSC1)	1000–2000[b]	0.75 (0.24)	1.700[d]	kinetic complexity of 1.5×10^5 bp	yes[g]	NT[h]	Delap et al. (1978)

[a] Copies per exponentially growing cell. [b] Copies per stationary phase cell. [c] The average molecular weight decreases to 0.9×10^6 upon amplification. [d] These values are close to or identical with those for high molecular weight DNA. [e] 10–30-fold increase in copy number. [f] 30-fold increase in copy number. [g] 10-fold increase in copy number. [h] NT = not tested.

more than one complexity class, (4) "amplification" in response to continued culture at stationary phase or after treatment with cycloheximide or puromycin, and (5) buoyant densities not appreciably different from that of total chromosomal DNA.

The origin, cellular location, and function of this DNA from a majority of sources are unknown. Some possible origins of this DNA are the following: (1) preparative artifacts (formation during preparation; prokaryotic contamination), (2) defective viruses, (3) as yet unrecognized organelles or symbionts, (4) byproducts of dead and dying cells, or (5) a directed phenomenon of biological significance such as intermediates or byproducts of genetic rearrangements including deletion, amplification, or transposition. There is evidence to suggest that spc DNAs from the cultured cell lines mentioned above are not preparative artifacts, are not due to prokaryotic contamination, and are not entirely derived from defective viruses or from unknown organelles (Smith & Vinograd, 1972; Stanfield & Helinski, 1976; Delap et al., 1978). Moreover, since varying fractions of these spc DNAs are homologous to chromosomal DNA sequences, the last two of the five possibilities mentioned above appear to be the most plausible explanation for the occurrence of spc DNA (Stanfield & Lengyel, 1979; Delap et al., 1978).

Despite considerable speculation on the nature and function of spc DNAs, very little molecular information is available regarding them. This is not surprising since (a) spc DNA represents a small fraction of total cellular DNA and (b) spc DNAs are heterogeneous in size and sequence. In this paper, we describe the cloning and characterization of some DNA sequences represented in spc DNA isolated from African green monkey cell line (BSC1). The results support a chromosomal origin for most or all spc DNAs in BSC1 cells and might contain clues to the formation and functional significance of this spc DNA.

Experimental Procedures

Cells. BSC1 cells were grown to confluence in modified Eagle's medium supplemented with 10% calf serum at 37 °C in roller bottles. Cells were harvested by trypsinization for 10 min at 37 °C, followed by pelleting at 4 °C for 10 min at 1500*g* (gravity). The pelleted cells were washed twice by resuspension in cold TD buffer [50 mM tris(hydroxymethyl)aminomethane (Tris), 10 mM KCl, 1.5 mM sodium phosphate, and 150 mM NaCl, pH 7.4) and finally pelleted at 4 °C for 10 min at 2500*g*. The pelleted cells either were used immediately or were stored at −20 °C for later use.

Preparation of BSC1 Form I DNA. The cell pellet obtained as described above was resuspended in 15 volumes of 0.15 M ethylenediaminetetraacetic acid (EDTA) (pH 8.0) and lysed by addition of 10% sodium dodecyl sulfate ($NaDodSO_4$) to a final concentration of 0.5%. The lysate was stirred vigorously at room temperature for 15 min and then alkali denatured by the addition of 1.25 N NaOH to a final pH of 12.2. The pH was maintained at 12.2 for 15 min, and the solution was rapidly brought to pH 7.8 by the addition of 6 N HCl. The mixture was extracted with 0.5 volume of phenol (previously equilibrated with 0.15 M EDTA, pH 8.0), and the nucleic acids were precipitated by adding 3 volumes of cold ethanol and incubation overnight at −20 °C. The precipitate was harvested by centrifugation at 5000*g* for 20 min at 4 °C, washed sequentially with 70% ethanol, 95% ethanol, and ether, air-dried, and then resuspended in a minimal volume of 10 mM EDTA (pH 8.0). RNase A was added to a final concentration of 200 μg/mL and the solution incubated at 70 °C for 10 min and then at 37 °C for 30 min. The material was then subjected to chromatography on a Sephadex G-100 column [750 mL; equilibrated with 0.3 M NaCl and 0.03 M trisodium citrate (2 × SSC)]. The material in the void volume was passed through a nitrocellulose column (25 mL; equilibrated with 25 volumes of 2 × SSC) and the effluent concentrated, dialyzed against TE buffer (10 mM Tris and 1 mM EDTA, pH 8.0), and then subjected to sequential (two or three) equilibrium centrifugations in cesium chloride–ethidium bromide gradients. Form I material was dialyzed against TE buffer after removing the ethidium bromide by three 1-propanol extractions. Extensive electron microscopy (Davis et al., 1971) revealed that 99% of the DNA molecules were circular.

Isolation and Fractionation of spc DNA. BSC1 form I DNA was fractionated by electrophoresis in 4% polyacrylamide gels. This step separates spc DNA from mitochondrial DNA and resolves the spc component into at least four broad size classes of about 1.5, 1.2, 0.8, and 0.3 kbp identifiable by ethidium bromide staining, which were recovered from the gel (Delap et al., 1978). General procedures used for gel electrophoresis and recovery of DNA have been described (Humayun et al., 1977). The isolated DNA fractions were finally dialyzed against TE buffer and stored frozen at −20 °C.

Cloning of spc DNA. In each reaction, 50 ng of three separated size classes (1.5, 1.2, and 0.8 kbp) of spc DNA was cleaved with an excess of restriction enzyme (1 unit) under conditions specified by the supplier, and the reaction was terminated by one phenol extraction followed by three ether extractions. *Taq*I and *Hpa*II were used to cleave each size class of spc DNA in two separate reactions. The restricted DNA was ligated to 50 ng of pBR322 plasmid previously

digested with *Cla*I (which cleaves the plasmid once) under the following conditions: restricted spc DNA (1 μg/mL) and *Cla*I-cut pBR322 (1 μg/mL) were ligated in 50 μL of ligase buffer (recommended by supplier) with 8 units/mL T4 DNA ligase for 12 h at 15 °C, and the reaction was terminated by heating at 60 °C for 5 min. The ligated DNA preparation either was used directly for transformation of competent cells or was cleaved with *Cla*I prior to transformation. A frozen (−70 °C) competent cell preparation (Morrison, 1979) from *Escherichia coli* C600 SF8 was used for transformation by standard procedures [e.g., see Humayun & Chambers (1978)]. Apparent recombinant DNA clones were isolated by selecting for amp^r tet^s phenotype.

Small amounts of plasmid DNA were isolated from 10–50-mL cultures by using one of several rapid isolation procedures (Colman et al., 1978; Kahn et al., 1979). For large preparations (1 L or more), scaled-up versions of the same procedures or that described by Clewell (1972) were used. Preliminary restriction maps were constructed for spc DNA recombinant plasmids by the usual procedures (Humayun et al., 1977).

"Gradient Blotting". Total form I DNA (~2 μg) prepared from BSC1 cells together with 2 μg of low multiplicity passaged SV40 DNA (internal marker) was centrifuged to equilibrim in 7 mL of a cesium chloride–ethidium bromide density gradient for 36 h at 38 000 rpm at 20 °C in a Beckman 50 Ti rotor. Gradients were fractionated dropwise from the bottom of the tube into about 30 equal fractions.

(A) Gradient Dot Blotting. Aliquots (10 μL) from each fraction were removed, diluted with 10 μL of 2 × SSC, and treated sequentially with 2 μL of 1.25 N HCl (5 min), 10 μL of 1 N NaOH (10 min), and enough 1.25 N HCl (about 6 μL) to neutralize the solution. A 10-μL aliquot of each denatured DNA sample was spotted onto a nitrocellulose filter strip, which was then baked for 2–4 h at 80 °C under vaccum. The filter was then soaked for 2–4 h at 65 °C with Denhardt's 2 × SSC (Denhardt, 1966) containing 40 μg/mL sonicated calf thymus DNA. The filter was blotted dry and hybridized with DNA labeled by nick translation with ^{32}P to a specific activity of (1–4) × 10^8 cpm/μg (for procedures, see below).

(B) Gradient Southern Blotting. A 3–5-μL sample of each gradient fraction collected as above was diluted with 10 volumes of electrophoresis buffer [1 × TEA (40 mM Tris, 1 mM EDTA, and 5 mM sodium acetate, pH adjusted to 7.65 with acetic acid)] and the DNA fractionated on a 1% agarose gel. The fractionated DNA was transferred to a nitrocellulose membrane strip and hybridized with ^{32}P-labeled probe DNA by the procedure of Southern (1975).

Preparation of High Molecular Weight (HMW) DNA from BSC1 Cells. Cells were grown to confluence in petri dishes in modified Eagle's medium supplemented with 10% calf serum. Cells were washed once with modified Eagle's medium and lysed by adding 3–4 mL/dish of 0.2% $NaDodSO_4$ in 10 mM Tris-HCl, pH 7.9, 10 mM NaCl, 10 mM EDTA, and 50 μg/mL proteinase K. The lysate was scraped into 150-mL glass centrifuge bottles and incubated overnight at 37 °C. The lysate was shaken with 1 volume of phenol–chloroform (1:1), and the clear aqueous layer which separated after a 10-min centrifugation at 5000 rpm was removed. The DNA was precipitated by adding 2 volumes of 95% ethanol, followed by pelleting at 4000*g* for 10 min. The precipitate was washed by resuspension in 70% ethanol, pelleting, and a 95% ethanol rinse of the pellet. The final pellet was dried under vacuum and resuspended in a minimal volume of 1 mM Tris-HCl and 0.5 mM EDTA, pH 7.9. RNase A was added to a concentration of 5 μg/mL and the suspension incubated overnight at 45 °C before dialysis against 1 mM Tris-HCl and 0.5 M EDTA, pH 7.9, at 4 °C. The usual yield was 100–200 μg/dish, and the average size of this DNA was 30 kbp.

Preparation of α Repetitive DNA from BSC1 Cells. BSC1 high molecular weight DNA (100 μg) was digested with 500 units of *Hin*dIII overnight at 37 °C in a buffer recommended by the supplier and was fractionated on a 5% polyacrylamide gel. Ethidium bromide staining of the gel revealed a series of evenly spaced bands characteristic of α family sequences (Rosenberg et al., 1978). The DNA in bands corresponding to the monomeric, dimeric, and trimeric α sequences was recovered by the usual procedures (Humayun et al., 1977).

Hybridization and Other Procedures. Labeled DNA probes were prepared by nick translation using the ^{32}P-labeled deoxynucleotide triphosphates (Amersham), DNase I (Worthington), and *E. coli* DNA pol I (Boehringer-Mannheim) essentially by the procedures of Maniatis et al. (1975). Southern blots were made by slight modification (Krowlewski et al., 1981) of the procedure developed by Southern (1975). Colony hybridizations were performed by the procedure of Grunstein & Wallis (1979) with some modifications.

DNA was sequenced by the procedure of Maxam & Gilbert (1977). DNA–DNA reassociation kinetic experiments were performed as previously described (Delap et al., 1976). Serially passaged and low multiplicity passaged covalently closed SV40 DNAs were prepared from cesium chloride–ethidium bromide gradient-resolved Hirt supernatants (Hirt, 1967) of infected BSC1 cells. pSN2 (a 1.2-kbp *Staphylococcus aureus* plasmid) DNA was prepared essentially by the procedure of Kahn et al. (1979).

Results

Cloning of African Green Monkey spc DNA. The heterogeneity of BSC1 spc DNA makes construction of restriction maps difficult. Therefore, two tetranucleotide-recognizing restriction enzymes, *Hpa*II (CCGG) and *Taq*I (TCGA), were arbitrarily chosen to cleave the spc DNA for initial cloning. Since a tetranucleotide is likely to be represented frequently in a heterogeneous population, any resultant cloned fragment should be more representative of the population than a clone bank derived from an enzyme recognizing a hexanucleotide sequence. DNA from spc DNA size classes of 1.5, 1.2, and 0.8 kbp was used in these cloning experiments. (Ethidium bromide staining of form I BSC1 spc DNA resolved on agarose or polyacrylamide gels reproducibly reveals four major abundance classes of 0.3, 0.8, 1.2, and 1.5 kbp.) As will be shown in this paper, many more distinct size classes of spc DNA can be identified by more sensitive techniques. spc DNA fragments obtained by digestion with the above two enzymes were cloned into the *Cla*I site of pBR322 by conventional procedures. spc DNA clones were initially identified by the amp^r tet^s phenotype expected as a result of insertion of heterologous sequences at the *Cla*I (ATCGAT) site of the vector plasmid, pBR322. It should be noted that *Hpa*II, *Taq*I, and *Cla*I produce the same sticky ends. spc DNA clones were further characterized by restriction analysis of isolated recombinant plasmid DNAs. Table II summarizes the cloning data and reveals that the cloned DNA fragments range in size from 0.02 to about 1.7 kbp. A large majority of the cloned segments were shorter than unit length (1.5, 1.2, or 0.8 kbp), indicating that those molecules that were cut had more than one site for the enzyme. Alternatively, it is conceivable that the shorter segments could have been the result of deletions after the initial cloning event (Brutlag et al., 1977). The information collected during insert sizing experiments involving

Table II: Sizes of Cloned DNA Fragments Derived from Three Size Classes of BSC1 spc DNA

spc DNA size class[a] (kbp)	restriction enzyme used for cloning	total no. of recombinants	no. of recombinants with insert sizes of		
			10–100 bp	100–500 bp	500 bp and longer[b]
1.5	*Hpa*II	25	14	10	1 (1030)
	*Taq*I	21	8	7	6 (1170)
1.2	*Hpa*II	18	6	10	2 (1750)
	*Taq*I	2	0	2 (346)	0
0.8	*Hpa*II	22[c]	11	5	1 (710)
	*Taq*I	59[c]	21	35	2 (1070)

[a] Sizes are averages for each size class (Delap et al., 1978). [b] The figures in parentheses are the sizes in base pairs of the largest cloned fragments. [c] The sizes of spc DNA inserts in some of the recombinants were not determined.

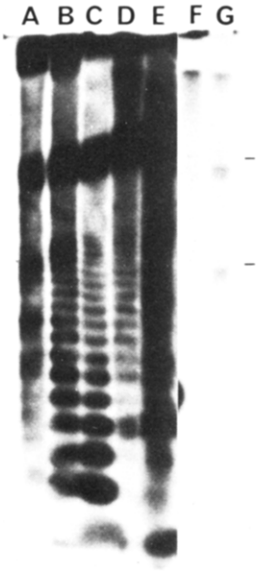


FIGURE 1: Autoradiograph of a Southern blot demonstrating homology of spc clone pTC001 DNA with BSC1 and SV40 DNAs. Restriction-digested BSC1 HMW (lanes A–E) and form I SV40 DNA derived from low and high multiplicity passages of the virus (lanes F and G, respectively) were fractionated by electrophoresis through a 1% agarose gel, blotted onto nitrocellulose, and fixed. The filter was hybridized at 65–67 °C in 2 × SSC with 1×10^7 cpm of a ^{32}P-labeled DNA fragment recovered from the spc DNA containing clone pTC001 by *Taq*I restriction digestion. Lanes A–E contain 10 μg of BSC1 HMW DNA digested with *Bam*HI (A), *Eco*RI (B), *Hin*dIII (C), *Hpa*II (D), and *Taq*I (E). Lanes F and G contain 0.5 μg of low multiplicity of infection-passaged SV40 DNA and serially passaged SV40 DNA, respectively. The hybridization pattern for the HMW DNA lanes is characteristic of that obtained for a highly repeated, tandemly arranged sequence. Note that the absence of a discrete ladder banding pattern in lanes A and E is expected for enzymes which do not cut a tandemly repeated sequence with a high degree of regularity. Hybridization of probe to serially passaged but not to low multiplicity passaged SV40 DNA indicates that the positive hybridization is due to host cell sequences contained in the defective virus population of the serially passaged virus. The bars at the right margin mark the position of form I and form II wild-type SV40 DNA. Note that defective SV40 genomes are smaller than wild-type genomes. Accordingly, the positive hybridization seen in lane G is with DNA species smaller than the corresponding wild-type component. The dark band extending diagonally across the upper portion of lanes A–E, the high molecular weight bands in lanes F and G, and weak hybridizations in lane F are all artifacts.

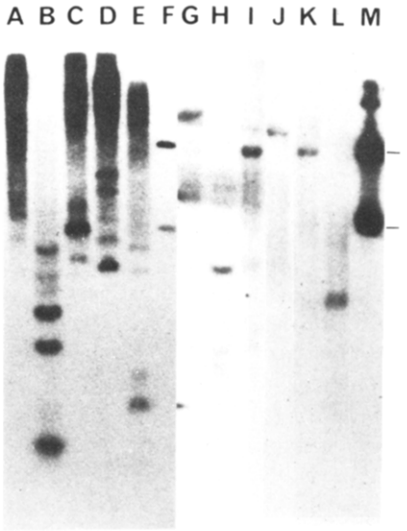


FIGURE 2: Composite autoradiograph of Southern blots demonstrating homology of three cloned BSC1 spc DNA sequences with chromosomal DNA. High molecular weight (HMW) DNA from BSC1 cells was digested with an excess of restriction endonuclease, electrophoresed through 1% agarose gels, and transferred to nitrocellulose strips. Each of three identical strips was then hybridized with 1×10^7 cpm of a different ^{32}P-labeled DNA probe [$(0.5–5) \times 10^8$ cpm/μg] in 2 × SSC at 65–67 °C. Probes were prepared from DNA fragments isolated from spc DNA clones. In this composite, lanes A–E contain 10 μg of HMW DNA digested with *Eco*RI, *Hae*III, *Hin*dIII, *Hpa*II, and *Taq*I, respectively. Lanes F and M contain 1 ng of pBR322 DNA (forms I and II). Lanes A–F were probed with a DNA fragment from the clone pHD005 which was isolated from a *Hpa*II digest of the recombinant. This fragment contains the entire inserted DNA fragment as well as flanking pBR322 DNA sequences originally found between positions (3901-161) of the vector. Lanes G–I contain 10 μg of HMW DNA digested with *Hin*dIII, *Hae*III, and *Eco*RI, respectively, and probed with the insert fragment from clone pTB009. This fragment was isolated by *Taq*I restriction of pTB009 and only contains insert sequences. Lanes J–L contain 10 μg of HMW DNA digested with *Eco*RI, *Hin*dIII, and *Taq*I, respectively. Lanes J–M were probed with the insert fragment derived by *Hpa*II restriction of DNA from clone pHB113. This fragment contains the entire segment of inserted DNA as well as flanking sequences derived from pBR322 DNA (see pHD005 above). The apparent difference in hybridization between lanes F and M results from the almost 50-fold greater exposure time required to detect the putative single-copy sequence present in the pTB009 and pHB113 insert segments. Control hybridization of BSC1 HMW DNA with ^{32}P-labeled pBR322 DNA failed to detect pBR322 homologous sequences in monkey DNA (data not shown), and control hybridization of blotted BSC1 mitochondrial DNA with each of the probes failed to detect homology.

a number of other tetranucleotide sequence recognizing enzymes (data not shown) did not reveal any consistent pattern in the distribution of such sites. The absence of a discernible pattern in the distribution of restriction sites as well as the heterogeneity of the insert sizes themselves suggests that the DNA sequences present in spc DNAs are not entirely made up of simple repeats.

Cloned spc DNA Sequences Show Homology to Chromosomal Sequences. In order to determine whether there were any sequence homologies between spc and chromosomal DNAs, we selected and examined a number of spc DNA clones for homology to chromosomal sequences. High molecular weight (chromosomal) DNA was isolated from BSC1 cells and digested overnight with various restriction enzymes. The digests were fractionated on agarose gels, and the DNA was transferred to nitrocellulose strips and hybridized to ^{32}P-labeled

cloned spc DNA fragments. The results of experiments using four different radiolabeled probes of cloned spc DNA are shown in Figures 1 and 2. Three distinct patterns of hybridization were observed. Figure 1 shows the patterns obtained when the insert fragment of clone pTC001 was used as a probe. These patterns (Figure 1, lanes A–E) are typical of a highly repeated, tandemly arranged chromosomal sequence, since an evenly spaced population of fragments hybridized to the probe. The size of the monomeric repeat unit, based on mobility in the gel, is about 170 bp. Figure 1 also includes the results of an experiment in which the insert fragment of clone PTC001 was used to probe DNA isolated from both high and low multiplicity of infection passaged simian virus 40 (SV40). (The rationale for this experiment is presented in a following section.) Figure 2 shows the patterns obtained when the insert fragments of clones pTB009 (lanes G–I) and pHB113 (lanes J–L) were used as probes. These patterns are typical of low copy number or unique chromosomal sequences since only a few fragments hybridized to the probes. Figure 2 also shows the patterns obtained when the insert fragment of clone pHD005 was used as a probe. These patterns (Figure 2, lanes A–E) are typical of interspersed chromosomal repeated sequences, since many smaller clearly resolved, and larger unresolved, fragments hybridized to the probe (Young, 1979).

Identification of spc DNA Clones Homologous to Highly Repeated Chromosomal DNA Sequences. In order to determine the representation of highly repetitive chromosomal sequences in the spc DNA clone bank, we carried out colony hybridization of the 138 clones with ^{32}P-labeled total monkey high molecular weight DNA. Such experiments were expected to reveal cloned sequences (in colonies) homologous to highly and moderately repetitive chromosomal sequences. It was anticipated that the extent of hybridization would be proportional to the extent of repetition in the chromosome such that there would be a gradient of autoradiographic intensity, with the strongest hybridizations indicating highly repeated chromosomal sequences and the weakest indicating low copy number chromosomal sequences. The results of such colony hybridizations revealed that 12 out of 138 spc DNA containing colonies exhibited strong hybridization while the remaining exhibited various degrees of weaker hybridization. It should be noted that clones containing very small segments of repetitive sequences and those containing highly diverged repetitive sequences could have escaped detection under the conditions of these experiments. Among the 12 strong positives was the spc DNA clone (pTC001) which had previously been identified as being homologous to a highly repeated tandemly arranged chromosomal sequence. To determine if the other 11 strongly positive clones were homologous to pTC001, we repeated the colony hybridization by using the insert fragment from pTC001 (about 350 bp) as the labeled probe. Only one clone, pTD022, hybridized with this labeled DNA.

Identification of Two Major Monkey Repeat Family Sequences in the spc DNA Clone Bank. In order to determine if any of the 12 BSC1 spc DNA clones previously identified as being homologous to repetitive chromosomal sequences belonged to the primate Alu family (Jelinek et al., 1980; Schmid & Jelinek, 1981), we carried out colony hybridization with concensus human Alu DNA sequences as a radioactive probe. The results showed that six out of the 12 indeed shared homology with Alu sequences. As previously noted, the spc DNA sequence contained in clone pTC001 was homologous to a tandemly repeated chromosomal DNA sequence. For determination of whether pTC001 and the related clone

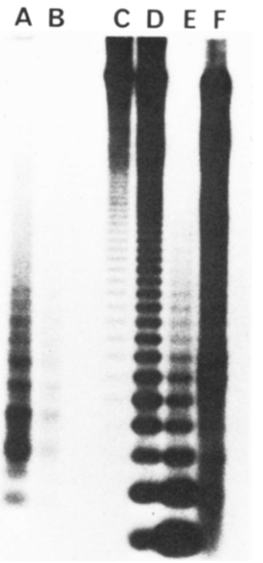


FIGURE 3: Autoradiograph of a Southern blot demonstrating the localization of BSC1 α family sequences in spc DNA. Total BSC1 form I DNA (lanes A and B) and restriction endonuclease digested BSC1 high molecular weight DNA (lanes C–F) were fractionated by electrophoresis through a 1% agarose gel, blotted onto nitrocellulose, and fixed. The filter was then hybridized with ^{32}P-labeled α family DNA (monomer) in 2 × SSC at 65–67 °C. Lanes A and B contain 0.6 and 0.1 μg of BSC1 form I DNA, respectively. Lanes C–F each contain 0.2 μg of BSC1 HMW DNA digested with *Hpa*II (C), *Eco*RI (D), *Hin*dIII (E), and *Taq*1 (F). The patterns observed in lanes C–F are comparable with a cloned BSC1 spc DNA sequence (pTC001) derived from the α repetitive family. Lanes A and B show size classes of α-homologous spc DNA. As many as 15–20 bands are observed on the original films. These lanes also show the absence of α-homologous sequences in mitochondrial DNA which runs near the top of the lanes and represents more than 50% (by mass) of the total form I BSC1 DNA. The smallest sized spc DNA band corresponds in mobility to a chromosomal α dimer (for example, lane E; second band from the bottom), allowing for the slightly greater mobility of form I circular DNA as compared to linear DNA.

pTD022 (neither of which show homology to Alu) shared sequence homology with the monkey α family, monkey chromosomal DNA was digested with *Hin*dIII, fractionated on agarose gels, transferred to nitrocellulose membranes, and hybridized with radiolabeled spc DNA inserts isolated from plasmids pTC001 or pTD022. The results (complete data not shown; for example, see lane C of Figure 1) showed that hybridization patterns with either probe coincided with the characteristic ethidium bromide staining pattern (Rosenberg et al., 1978) of a *Hin*dIII digest of chromosomal DNA, indicating homology of these two clones with the monkey α family sequence. In a related experiment, colony hybridization using ^{32}P-labeled chromosomal "consensus" α monomeric DNA (see Experimental Procedures) revealed that only pTC001 and pTD022 exhibited hybridization (data not shown).

Representation of the α Family Sequence in Total, Uncloned spc DNA. The experiments described above suggested the presence of α family sequences in spc DNA. In order to determine whether this family was present in any particular size classes of spc DNA, we performed the following experiment. Total BSC1 form I (covalently closed circular) DNA preparations were fractionated on agarose gels, transferred to nitrocellulose membranes, and hybridized to radioactively labeled consensus α family monomeric sequences. The results (Figure 3, lanes A and B) showed that α family sequences were apparently represented in all size classes of spc DNA. Lanes A and B of Figure 3 reveal a series of discrete bands

with sizes (approximately 0.3–3.0 kbp) expected for BSC1 spc DNAs (note that this banding pattern was obtained by using undigested form I spc DNA preparations). Despite the relatively sharp resolution, it is difficult to ascertain if the bands represent multimers of the smallest unit. Nevertheless, the spacing between the spc DNA bands hybridizing to an α family sequence probe suggests that BSC1 spc DNA contains an oligomeric series of α family DNA circles, with the smallest member being a dimer of the basic 172-bp sequence. Figure 3 also reveals that mitochondrial DNA, which constitutes a significant mass fraction (about 65%) of form I BSC1 DNA, is not homologous to α sequences.

From the results of these experiments, it appeared that the α family was represented in all size classes of monkey spc DNA. For assessment of total α family sequence content of form I BSC1 spc DNA preparations, reassociation kinetics analysis was employed. Total spc DNA resolved from mitochondrial DNA by electrophoresis through 4% polyacrylamide gels was radiolabeled and driven into hybrid with a 10^4-fold (10 μg/mL) excess of α monomer DNA (C_0t = 1.0). This analysis (data not shown) indicated that 2–4% of the total spc DNA was homologous to α family sequences. (This result has been confirmed by similar experiments using a cloned representative of the α family monomer DNA as driver.)

Since this result implies an underrepresentation of the α family in BSC1 spc DNA (relative to the chromosomal frequency of 10–20%), the total content of highly repetitive DNA in spc DNA was also assessed. For these experiments, a ^{32}P-labeled sample of total spc DNA was driven into hybrid in a 10^5-fold (100 μg/mL) excess of total BSC1 chromosomal DNA (C_0t = 10). The results of these experiments indicated that spc DNA contained a total of 6–12% of highly repetitive DNA sequences, again an apparent underrepresentation relative to their reported frequency in chromosomal DNA.

Confirmation That α Family Sequences Are Present in Small Circular DNA Species. The form I spc DNA preparations used in the experiments described here were purified by several consecutive centrifugations through cesium chloride–ethidium bromide density gradients (Experimental Procedures) and were free of chromosomal DNA contamination as determined by electron microscopy. We therefore consider it unlikely that the two α-DNA-containing clones (pTC001 and pTD022) were derived from contaminating chromosomal fragments. In addition, the experiments described in the preceding section (Figure 3) indicate the presence of α family sequences in DNA with the same mobility as spc DNA. Nevertheless, for confirmation of the presence of α family sequences in form I spc DNA, two related strategies were employed. BSC1 form I preparations (which consist of 65% mitochondrial and 35% spc DNA by mass) were centrifuged to equilibrium in cesium chloride–ethidium bromide density gradients together with low multiplicity passaged SV40 DNA as an internal marker. A total of 80% of the marker was present as form I (closed circular) and 20% as form II (open circular) DNA. The gradients were fractionated from the bottom, and the DNA in each fraction was analyzed by two methods. In "gradient dot blot" experiments, the DNA in each individual fraction was denatured, spotted onto duplicate strips of nitrocellulose, and hybridized to radioactively labeled SV40 or α monomer DNAs. Hybridization with labeled SV40 DNA (Figure 4) revealed two peaks of hybridization which correspond in position to form I and form II species of the SV40 DNA marker. Hybridization with labeled α family consensus (monomeric) DNA also revealed (Figure 4) two hybridization peaks coinciding with the form I and form II positions defined

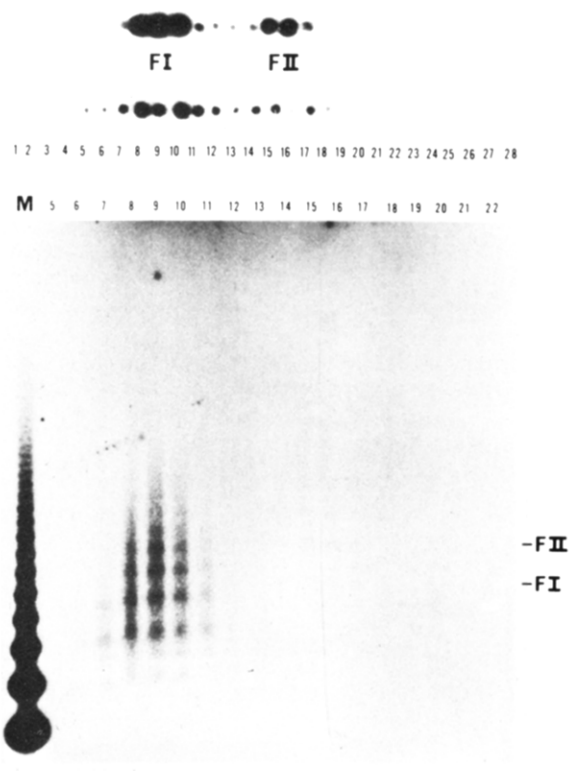
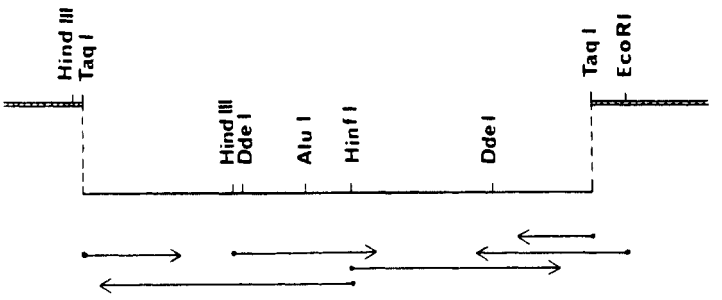


FIGURE 4: Cesium chloride–ethidium bromide gradient blotting. (Upper panel) Total form I DNA (2 μg) from BSC1 cells and 2 μg of SV40 (80% form I and 20% form II) DNA were centrifuged to equilibrium in a 7.0-mL ethidium bromide–cesium chloride density gradient. The gradient was fractionated from the bottom into 28 equal fractions, and 10 μL of each fraction was removed, diluted with 10 μL of 2 × SSC, and treated sequentially with 2 μL of 1.25 N HCl, 10 μL of 1 N NaOH, and 6 μL of 1.25 N HCl (to neutral pH). Of each treated fraction, 9.5 μL was then spotted and fixed onto a nitrocellulose strip and probed with ^{32}P-labeled DNA in 2 × SSC at 65 °C. The upper strip was probed with ^{32}P-labeled SV40 DNA and the lower with ^{32}P-labeled α family DNA (monomer) prepared as described (Experimental Procedures). After being washed, each filter was subjected to autoradiography at −70 °C. The figure illustrates the resolution to form I (FI) and form II (FII) DNAs and demonstrates the presence of sequences homologous to the α family probe in the total form I DNA from BSC1 cells. (Lower panel) Of each gradient fraction, 3.5 μL was diluted to 35 μL with electrophoresis buffer (1 × TEA) and electrophoresed through a 1% agarose gel. The DNA was then transferred to a nitrocellulose filter, fixed, and hybridized to ^{32}P-labeled α family DNA as described above. Lane M contains 50 ng of *Hin*dIII-digested BSC1 HMW DNA included as linear size markers (bands spaced at integer multiples of 172 bp) and as a control for hybridization. FI and FII at the right of the lower panel mark the distance migrated for forms I and II DNA of pSN2, a 1.2-kbp enterotoxin plasmid of *S. aureus*. This panel shows the distribution of α family homologous sequences within the spc DNA of BSC1 cells. It further shows the absence of homology to low multiplicity passaged SV40 DNA present in the gradient and the absence of α family homology to BSC1 mitochondrial DNA which comprises more than 50% of the total mass of form I DNA present (see text).

by the SV40 marker. Identical results were obtained when the insert fragments of clones pTC001 and pTD022 were used as probes in this type of analysis (data not shown). In "gradient Southern blot" experiments, DNA from the (cesium chloride–ethidium bromide gradient) fractions used in dot blotting was fractionated on 1% agarose gels, transferred to nitrocellulose strips, and hybridized to the same labeled consensus α-DNA probe. The results of such experiments (Figure 4) showed that α family DNA was present in form I (fractions 7–12) and form II (14–18) peaks. The gel fractionation pattern was similar to that obtained in Southern blots where total form I DNA was directly fractionated on gels (Figure




```
          74          84          94          104         114         124         134         144
Consensus: TCGCTAAGGC TGTTCTTGTG GAATTGGCAA AGGGATATTT GGAAGCCCAT AGAGGGCTAT GGTGAAAAAG GAAATATCTT
Insert   : ---A----A- ---------- -----T-T-- --T------- ---------- --T------- ---------- ---G-X---
           1          10          20          30          40          50          60          70

          154         164         2           12          22          32          42          52
Consensus: CCGTTCAAAA CTGGAAAGAA GCTTTCTGAG AAACTGCTCT GTGTTCTGTT AATTCATCTC ACAGAGTTAC ATCTTTCCCT
Insert   : ---A-A---- T---------- ---------- ---------- ---------T- -----G--- ----G----- -----A----- -G--------C
           80          90          100         110         120         130         140         150

          62          72          82          92          102         112         122         132
Consensus: TCAAGAAGCC TTTCGCTAAG GCTGTTCTTG TGGAATTGGC AAAGGGATAT TTGGAAGCCC ATAGAGGGCT ATGGTGAAAA
Insert   : ---------- ---- --G--- A--C------ G------T-G ----T-T--- ---------- -C------ ---X----------
           160         170         180         190         200         210         220         229

          142         152         162         172         10          20          30          40
Consensus: AGGAAATATC TTCCGTTCAA AACTGGAAAG AAGCTTTCTG AGAAACTGCT CTGTGTTCTG TTAATTCATC TCACAGAGTT
Insert   : ---T------ ---A --GA-- -----A---- ----AC---- ---------G TG-------- G---------- ----------
           239         249         259         269         279         289         299         309

          50          60          70          80
Consensus: ACATCTTTCC CTTCAAGAAG CCTTTCGCTA AGGCT
Insert   : ---A-C--T- TC------A -------A
           319         329         339     346
```

FIGURE 5: Nucleotide sequence of the spc DNA fragment in clone pTC001. The upper panel shows a partial restriction map of the insert and indicates (arrows) the strategy employed to determine the nucleotide sequence. Hatched regions bounding the *Taq*I sites represent pBR322 DNA. Note that the insert is oriented with the *Hind*III site of pBR322 DNA on the left-hand side. The lower panel compares the African green monkey α family consensus sequence (Rosenberg et al., 1978) with the spc DNA insert. Equivalence of the sequences is indicated by a dash in the line labeled insert. Where the sequences differ, the base pair of the insert is given. An X in the line labeled insert denotes a deleted base pair.

3), and as expected, the bands in the form I region migrated more rapidly than those corresponding to form II. No hybridization was observed at positions corresponding to either mitochondrial or SV40 DNAs. These results confirm that α family sequences occur in small covalently closed circular DNAs and that they are represented by discrete, almost evenly spaced bands suggestive of an oligomeric series. It is assumed that form II material represents in vitro nicking of the original form I preparation.

The failure to detect low multiplicity passaged SV40 DNA markers in BSC1 spc DNA gradient Southern blot experiments probed with either α monomer, pTC001, or pTD022 DNAs was consistent with the results obtained from the experiment presented in lane F of Figure 1 where low multiplicity passaged SV40 DNA was probed with the pTC001 insert fragment. It should also be noted that, in agreement with published reports (Kelly & Nathans, 1977; McCutchan et al., 1979; Wakamuya et al., 1979), the results obtained from the experiment presented in lane G of Figure 1, where serially passaged SV40 DNA was probed with the pTC001 insert, demonstrated that some defective SV40 genomes, known to accumulate during serial passage of the virus, contained substituted α family sequences.

Analysis of Chromosomal Repetitive DNA Sequences Contained in Two spc DNA Clones. For further definition of the structure of spc DNAs, the nucleotide sequences of the DNA inserts in two clones were determined. The Alu-homologous DNA insert in pTD120 is 379 bp long and contains an entire Alu element together with about 25 bp of non-Alu DNA flanking each end of the Alu sequence. Characterization of the Alu-homologous spc DNA clones has been described in a separate communication (Krowlewski et al., 1982).

The α-homologous spc DNA clone pTC001 was derived by cloning a *Taq*I digest of the 1.2-kbp size class of BSC1 spc DNA. The nucleotide sequence of the entire 346-bp-long spc DNA insert was determined by the methods of Maxam & Gilbert (1977) and is shown in Figure 5. Figure 5 includes a comparison of the consensus α sequence with the cloned insert. An examination of these DNA sequences reveals the following: (1) the 346-bp-long insert, which is composed entirely of α family sequences, is a permuted dimer of the chromosomal α family sequence containing one complete α monomeric sequence flanked by two partial repeats in a head-to-tail arrangement; (2) the internal monomeric repeat differs from the consensus sequence by 20 base mismatches and one single base-pair deletion; (3) the two partial repeats diverge in sequence compared to both the consensus and the internal monomeric sequence.

Discussion

More than a decade after its initial discovery, eukaryotic spc DNA remains an enigmatic biological phenomenon. Nevertheless, an appreciation of various prokaryotic genetic phenomena involving circular DNA structures coupled with existing information regarding eukaryuotic spc DNA has led to a number of speculations concerning its origin and function, including the possibility that spc DNA represents intermediates or byproducts of genetic rearrangements. In any case, spc DNA appears to be a widely represented biological phenomenon, and thus a determination of the nucleotide sequences

present in this DNA appeared to be an appropriate significant step toward its characterization. For the intial studies, we cloned spc DNA from an African green monkey cultured cell line (BSC1).

Strategy Used for Cloning spc DNA. The initial cloning strategy illustrates the problems encountered in working with spc DNA. The heterogeneity of sequences contained in spc DNA and the small quantities of material preclude the construction of restriction maps. This leaves one with the option of adopting various shotgun cloning approaches, involving fragmentation of spc DNA either by a nonspecific nuclease (e.g., DNase I) or with specific restriction endonucleases. We chose two tetranucleotide-recognizing restriction enzymes (*Taq*I and *Hpa*II) for fragmenting spc DNA. Assuming a more or less random sequence representation, it was reasoned that some spc molecules of all size classes would be cut at least once by either enzyme. The major difficulty anticipated with this approach is that intact full-length circles may not be cloned. Nevertheless, an analysis of fragmented cloned spc DNA sequences has yielded useful information.

Examination of the spc DNA inserts in our clone bank did not reveal any patterns in the size of the inserts, which ranged from about 20 to 1700 bp. In those cases examined, the distribution of internal restriction endonuclease sites did not reveal any pattern. These results implied that spc DNA molecules contained no common "core" sequence, although microheterogeneity in such a sequence could also give the above results. Nevertheless, the sizes and restriction patterns of cloned sequences indicated that complex rather than simple sequences were represented in spc DNA.

Nature of Sequences in Cloned BSC1 spc DNA. All of the spc DNA clones examined were capable of hybridizing with chromosomal sequences, suggesting that most if not all spc DNA was derived from chromosomes. These results indicate that spc DNA (1) is not due to prokaryotic contamination, (2) is not exclusively made up of exogenous defective viruses, and (3) is unlikely to be derived from any unknown organelles or symbionts.

The hybridization experiments also revealed that both unique and repetitive DNA sequences were represented in spc DNA. Cloned fragments homologous to highly repeated chromosomal sequences were identified by colony hybridization using ^{32}P-labeled high molecular weight BSC1 DNA. Twelve out of 138 clones appeared to contain DNA homologous to highly repeated chromosomal sequences. Out of these, six were shown to be homologous to the human Alu, two (pTC001 and pTD022) to the major African green monkey α repeat family sequence, and four (including pHD005) to as yet unidentified repeat family sequences.

In higher mammals, as much as 40% of the haploid genome might consist of repeating DNA elements (Britten & Kohne, 1968; Lewin, 1980). The major monkey repeat family sequence, α, might represent 10–20% of the genomic DNA (Rosenberg et al., 1978). If the representation of the Alu sequence in the monkey is similar to that in the human, then the α and Alu sequences alone should account for 25% of total genomic DNA. The finding that only 12 out of 138 clones shared sequence homology with highly repetitive chromosomal sequences might mean that spc DNA has a sequence representation very different from that of the chromosome or that our clone bank is not representative of the sequences contained in spc DNA. In addition, it is possible that highly diverged but still identifiable versions of highly repeated sequences may have escaped detection because of the nonspecific probe used. In any event, from the reassociation data and the analysis of the clone bank, it is clear that spc DNA contains sequences, both repeated and unique, other than the Alu and α families.

While reassociation experiments and Southern blotting reveal the extent of representation and organization of repeated sequences in spc DNA, an examination of the actual nucleotide sequences present in a number of spc DNA molecules will be necessary to determine any specific features distinguishing spc DNA from the genomic consensus α sequences. Determination of the sequence of the spc DNA insert in pTC001 (Figure 5) showed this particular 346-bp segment to be a permuted dimer of the consensus α sequence published by Rosenberg et al. (1978). However, a comparison of the internal complete monomeric sequence with the consensus revealed a large divergence. Specifically, this sequence differs from the consensus at 21 positions, with 20 base mismatches and one single-base deletion. This represents a 12% divergence and is significantly greater than the 3% divergence reported for several cloned genomic α family members (Thayer et al., 1981). However, it would be premature to attach any significance to these data before the nucleotide sequences contained in a number of spc DNA molecules become available.

Representation of the α Family Sequences in BSC1 Form I spc DNA. The representation of the α sequence in total BSC1 spc DNA was investigated by a variety of techniques. Thus, gradient dot blotting and gradient Southern blotting were used to demonstrate that α family sequences were actually present in form I spc DNA molecules. The Southern blots of spc DNA fractionated on agarose gels and hybridized to labeled α family sequences revealed that they were represented in all major size classes of spc DNA and that the spacing of the α-homologous spc DNA bands was fairly regular. It is important to note that this spacing was observed on blotted form I spc DNA (i.e., no restriction endonuclease digestion was involved) and is highly suggestive of an oligomeric series of circular molecules at least partly made up of α family sequences. We are concurrently determining whether any of the spc DNA circles are composed entirely of α sequences. A Southern blot experiment using form I DNA cut with α and probed with a consensus α sequence reveals a ladder pattern similar to that given by chromosomal DNA and contains at least 12 bands (L. M. Refolo and M. Z. Humayun, unpublished experiments). This result is consistent with the existence, within the spc DNA population, of an oligomeric series of DNA molecules differing from each other by the number of α sequence repeats.

Significance of Sequences Present in BSC1 spc DNA. What, if any, is the significance of finding repetitive DNA sequences in spc DNA? It is known that one major highly repetitive sequence of higher mammals (the Alu family) has a conserved, 14-bp segment homologous to a sequence occurring at the origins of the replication of papova viruses (Jelinek et al., 1980). A subset of the yeast-repeating sequence *Ty1* is reported to have origins of replication [e.g., see Chan & Tye (1980)]. Many repetitive sequences seem to be mobile (Young, 1979; Potter et al., 1979). Saltatory replication of repetitive sequences has been proposed as one possible mechanism for both tandem duplication as well as dispersal of repeat sequences (Brutlag, 1980). Amplification of certain genes is assumed to involve circular DNA species (Hourcade et al., 1973). One conceivable consequence of all these predicted and observed eukaryotic DNA rearrangements could be the generation of circular DNA intermediates. Indeed, spc DNA might represent a collection of disparate phenomena which happen to share a structural feature, circularity. However, it is also possible that the presence of repetitive

sequences in spc DNA could be due to the fact that spc DNA is a nonspecific side product of DNA degradation. The prevalence of repetitive sequences in the chromosome would ensure their association with spc DNA, although several indirect lines of evidence argue against this possibility. First, more than 80% of *Drosophila* spc DNA appears to share sequence homologies with the chromosomal middle repetitive component (Stanfield & Lengyel, 1979), which constitutes only 16–17% of the genome (Young, 1979), implying that spc DNA formation in this species is a directed phenomenon. Second, data presented in this paper reveal a striking periodicity in the sizes of spc DNA molecules hybridizing to a consensus α sequence probe. This periodicity approximately corresponds to that expected for an oligomeric series of circular DNAs with size increments roughly corresponding to the α monomeric sequence. It is difficult to reconcile such periodicities with a nonspecific phenomenon. It is more likely that such circular DNAs are derived from chromosomal tandem repeats as a result of homologous recombination, leaving a deletion in the chromosome and creating an extrachromosomal form of the repeat which may be free to amplify and/or to translocate to other intra- or interchromosomal locations as envisaged by the saltatory replication hypothesis. In the case of dispersed repeated chromosomal sequences (e.g., the *Alu* sequence), an excised circle would carry "passenger" sequences originally located between two members of the repeated element. Such phenomena could be analogous to, but distinct from "true" transposition according to bacterial models (Shapiro, 1979; Calos & Miller, 1980) and might be involved in many transposition events including the generation of orphons (Childs et al., 1981). Some or all of spc DNA might thus represent extrachromosomal intermediates of mobile genetic elements.

One could also consider some spc DNA molecules to be not so much intermediates as byproducts in genetic rearrangements. Thus, these DNAs might arise by excision processes suggested above to bring two chromosomal sequences together, very much like what is known to happen in certain immunoglobulin gene rearrangements (Early & Hood, 1981; Molgaard, 1980). Once such rearrangement has been accomplished by homologous recombination between two repeat DNA sequences flanking the sequence to be deleted, the resultant DNA circle would have no further function. Genomic rearrangements by DNA loss may also occur at random without a specific purpose. Schmookler-Reis & Goldstein (1980) have recently found that human repetitive DNA sequences (specifically the *Eco*RI family, which is the human equivalent of the monkey α sequence; Wu & Manuelidis, 1980) are selectively lost from primary human diploid fibroblast cultures. Perhaps the excision processes involved in this phenomenon generate small heterogeneous circular DNA species.

Whether any of the above hypotheses are correct must await, at the very least, the complete nucleotide sequence determination of several spc DNA circles and the chromosomal regions from which they were derived. In the meantime, our preliminary data indicate that spc DNA isolated from a human cell line (HeLa) shares several organizational features with the monkey DNA (L. M. Refolo and M. Z. Humayun, unpublished experiments).

Acknowledgments

We thank Horace Lozina for technical assistance, Heinz Annus for help in the preparation of figures, and Helen Beale-Holcombe for preparation of the manuscript. Data processing facilities made available by SUMEX and the Stanford Molgen Projects are gratefully acknowledged.

References

Britten, R. J., & Kohne, D. E. (1968) *Science (Washington, D.C.) 161*, 529–540.
Brutlag, D. L. (1980) *Annu. Rev. Genet. 14*, 121–144.
Brutlag, D. L., Fry, K., Nelson, T., & Hung, P. (1977) *Cell (Cambridge, Mass.) 10*, 509–519.
Calos, M. P., & Miller, J. H. (1980) *Cell (Cambridge, Mass.) 20*, 579–595.
Chan, C. S. M., & Tye, B. (1980) *Proc. Natl. Acad. Sci. U.S.A. 77*, 6329–6333.
Childs, G., Maxson, R., Cohn, R. H., & Kedes, L. (1981) *Cell (Cambridge, Mass.) 23*, 651–663.
Clewell, D. B. (1972) *J. Bacteriol. 110*, 667–676.
Colman, A., Byers, M. J., Primrose, S. B., & Lyons, A. (1978) *Eur. J. Biochem. 91*, 303–310.
Davis, R. W., Simon, M., & Davidson, N. (1971) *Methods Enzymol. 21D*, 413–428.
Delap, R. J., & Rush, M. G. (1978) *Proc. Natl. Acad. Sci. U.S.A. 75*, 5855–5859.
Delap, R. J., Zouzias, D., & Rush, M. G. (1976) *Biochim. Biophys. Acta 418*, 257–265.
Delap, R. J., Rush, M. G., Zouzias, D., & Khan, S. (1978) *Plasmid 1*, 508–521.
Denhardt, D. T. (1966) *Biochem. Biophys. Res. Commun. 23*, 641–652.
Early, P., & Hood, L. (1981) *Cell (Cambridge, Mass.) 24*, 1–3.
Grunstein, M., & Wallis, J. (1979) *Methods Enzymol. 68*, 379–389.
Hirt, B. (1967) *J. Mol. Biol. 26*, 365–373.
Hourcade, D., Dressler, D., & Wolfson, J. (1973) *Cold Spring Harbor Symp. Quant. Biol. 38*, 537–550.
Humayun, M. Z., & Chambers, R. W. (1978) *Proc. Natl. Acad. Sci. U.S.A. 75*, 774–778.
Humayun, M. Z., Jeffrey, A., & Ptashne, M. (1977) *J. Mol. Biol. 112*, 262–277.
Jelinek, W. R., Toomey, T. P., Leinwand, L., Duncan, C. H., Biro, P. A., Choudary, P. V., Weissman, S. M., Rubin, C. M., Houck, C. M., Deininger, P. L., & Schmid, C. W. (1980) *Proc. Natl. Acad. Sci. U.S.A. 77*, 1398–1402.
Kahn, M., Kolter, R., Thomas, C., Figurski, D., Meyer, R., Remaut, E., & Helinski, D. R. (1979) *Methods Enzymol. 68*, 268–280.
Kelly, T. J., Jr., & Nathans, D. (1977) *Adv. Virus Res. 21*, 85–173.
Krowlewski, J. J., Murphy, E., Novick, R. P., & Rush, M. G. (1981) *J. Mol. Biol. 159*, 19–33.
Krowlewski, J. J., Bertelsen, A. H., Humayun, M. Z., & Rush, M. G. (1982) *J. Mol. Biol. 154*, 399–415.
Lewin, B. (1980) *Gene Expression*, Vol. 2, pp 503–569, Wiley, New York.
Maniatis, T., Jeffrey, A., & Kleid, D. G. (1975) *Proc. Natl. Acad. Sci. U.S.A. 72*, 1184–1188.
Maxam, A. M., & Gilbert, W. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 560–564.
McCutchan, T., Singer, M., & Rosenberg, M. (1979) *J. Biol. Chem. 254*, 3592–3597.
Molgaard, H. V. (1980) *Nature (London) 286*, 657–659.
Morrison, D. A. (1979) *Methods Enzymol. 68*, 327–331.
Potter, S. S., Brorein, W. J., Dunsmuir, P., & Ruben, G. M. (1979) *Cell (Cambridge, Mass.) 17*, 415–427.
Radloff, R., Bauer, W., & Vinograd, J. (1967) *Proc. Natl. Acad. Sci. U.S.A. 57*, 1514.
Rosenberg, H., Singer, M., & Rosenberg, M. (1978) *Science*

(*Washington, D.C.*) *200*, 394–402.
Schmid, C. W., & Jelinek, W. R. (1981) *Science* (*Washington, D.C.*) (in press).
Schmookler-Reis, R. J., & Goldstein, S. (1980) *Cell* (*Cambridge, Mass.*) *21*, 739–746.
Shapiro, J. A. (1979) *Proc. Natl. Acad. Sci. U.S.A. 76*, 1933–1937.
Smith, C. A., & Vinograd, J. (1972) *J. Mol. Biol. 69*, 163–178.
Southern, E. M. (1975) *J. Mol. Biol. 98*, 503–517.
Stanfield, S., & Helinski, D. R. (1976) *Cell* (*Cambridge, Mass.*) *9*, 333–345.
Stanfield, S. W., & Lengyel, J. A. (1979) *Proc. Natl. Acad. Sci. U.S.A. 76*, 6142–6146.
Thayer, R. E., Singer, M. F., & McCutchan, T. F. (1981) *Nucleic Acids Res. 9*, 169–181.
Wakamuya, T., McCutchan, T., Rosenberg, M., & Singer, M. (1979) *J. Biol. Chem. 254*, 3584–3591.
Wu, J. C., & Manuelidis, L. (1980) *J. Mol. Biol. 142*, 363–386.
Young, M. W. (1979) *Proc. Natl. Acad. Sci. U.S.A. 76*, 6274–6278.

In Vitro Synthesis of the Membrane-Bound D-Lactate Dehydrogenase of *Escherichia coli*†

Eugenio Santos,‡ Hsiang-fu Kung,§ Ian G. Young, and H. Ronald Kaback*

ABSTRACT: Synthesis of the membrane-bound, flavin-linked D-lactate dehydrogenase of *Escherichia coli* has been studied by using a recombinant plasmid containing the *dld* gene [Young, I. G., Jaworowski, A., & Poulis, M. (1982) *Biochemistry* (following paper in this issue)]. Expression of the cloned *dld* gene was achieved either in vivo with transformed minicells or in vitro with a fractionated transcription/translation system. In both instances, a product is observed that is specifically immunoprecipitated by γ-globulin prepared against the purified enzyme and comigrates with authentic D-lactate dehydrogenase on sodium dodecyl sulfate–polyacrylamide gel electrophoresis. Furthermore, the product is catalytically active and binds to membrane vesicles during or after synthesis. Thus, it seems likely that the protein is synthesized in mature form and binds to the membrane without a leader peptide sequence. Interestingly, addition of flavin adenine dinucleotide to the in vitro reaction mixtures causes a 2-fold increase in the synthesis of the enzyme, suggesting that the cofactor plays a regulatory role in the synthesis of the apoprotein. Finally, L factor, a protein involved in regulation of protein elongation, has an inhibitory effect on the expression of the *dld* gene and a stimulatory effect on the expression of the *ndh* gene (encoding NADH dehydrogenase).

Substrate oxidation via a membrane-bound respiratory chain drives the active transport of a wide variety of solutes in right-side-out membrane vesicles prepared from many bacteria (Kaback, 1972, 1974; Owen & Kaback, 1978, 1979a,b) by mechanisms in which an electrochemical gradient of protons ($\Delta\bar{\mu}_{H^+}$,[1] interior negative and alkaline) is the immediate driving force (Mitchell, 1963, 1968, 1973; Kaback, 1976; Harold, 1977; Konings & Boonstra, 1977). Furthermore, for reasons that are not fully understood, although the vesicles have the capacity to oxidize a number of substrates, generation of $\Delta\bar{\mu}_{H^+}$ is relatively substrate specific (Kaback, 1972, 1974). Thus, in vesicles from *Escherichia coli* and *Salmonella typhimurium*, D-lactate serves as the most effective physiological electron donor for the generation of $\Delta\bar{\mu}_{H^+}$, even though the vesicles respire at higher rates in the presence of other substrates (Barnes & Kaback, 1971; Schuldiner & Kaback, 1975; Stroobant & Kaback, 1975; Ramos et al., 1976).

Incubation of *E. coli* vesicles with radioactive D-lactate results in stoichiometric conversion to pyruvate (Kaback & Milner, 1970; Barnes & Kaback, 1970), and the D-lactate dehydrogenase (D-LDH) that catalyzes this reaction has been solubilized from the membrane and purified to homogeneity (Kohn & Kaback, 1973; Futai, 1973; Pratt et al., 1979). The enzyme has a molecular weight of 70K ± 10%, is composed of a single polypeptide chain containing 1 mol of flavin adenine dinucleotide (FAD) per mol of protein, exhibits a high degree of specificity for D(−)-lactate, and is specifically inactivated by the "suicide substrate" 2-hydroxy-3-butynoic acid (HBA) (Walsh et al., 1972). Moreover, antibody inhibition (Short et al., 1975a,b) and immunoadsorption studies (Owen & Kaback, 1978, 1979a,b) demonstrate clearly that the protein is associated with the cytoplasmic surface of the plasma membrane.

Vesicles prepared from mutants specifically defective in D-LDH do not catalyze D-lactate oxidation or D-lactate-dependent active transport (Hong & Kaback, 1972). However, the vesicles can be reconstituted with D-LDH by exposure to crude (Reeves et al., 1973) or purified (Short et al., 1974) enzyme preparations, and in the reconstituted system, it is apparent that the enzyme is bound to the outer surface of the vesicle membrane (Short et al., 1975a,b; Futai, 1975). Remarkably, despite the abnormal location of the enzyme, oxidation of D-lactate by reconstituted *dld*⁻ vesicles leads to the generation of a $\Delta\bar{\mu}_{H^+}$ that is indistinguishable in polarity and

† From the Laboratory of Membrane Biochemistry, Roche Institute of Molecular Biology, Nutley, New Jersey 07110 (E.S., H.K., and H.R.K.), and the Department of Biochemistry, John Curtin School of Medical Research, Australian National University, Canberra City, A.C.T. 2601, Australia (I.G.Y.). *Received September 22, 1981.*

‡ Postdoctoral fellow of the Spanish–North American Joint Committee for Scientific and Technological Cooperation.

§ Present address: Department of Molecular Genetics, Hoffmann-La Roche, Nutley, NJ.

[1] Abbreviations: $\Delta\bar{\mu}_{H^+}$, electrochemical gradient of protons; D-LDH, D-lactate dehydrogenase; FAD, flavin adenine dinucleotide; HBA, 2-hydroxy-3-butynoic acid; NDH, NADH dehydrogenase; Cl_3CCOOH, trichloroacetic acid; PBS, phosphate-buffered saline; $NaDodSO_4$, sodium dodecyl sulfate; PMS, phenazine methosulfate; Tris, tris(hydroxymethyl)aminomethane.

0006-2960/82/0421-2085$01.25/0