Human Homologs of TU Transposon Sequences: Polypurine/Polypyrimidine Sequence Elements That Can Alter DNA Conformation In Vitro and In Vivo

BARBARA HOFFMAN-LIEBERMANN,^{1†} DAN LIEBERMANN,^{1†} ANTHONY TROUTT,² LAURENCE H. KEDES,² AND STANLEY N. COHEN^{1,3*}

Departments of Genetics¹ and Medicine,³ Stanford University, Stanford, California 94305, and Veterans Administration Medical Center, Palo Alto, California 94304²

Received 28 April 1986/Accepted 15 July 1986

We previously have shown that homologs of the outer domain segment of the inverted repeat termini (IVR-OD) of the sea urchin TU transposons are conserved among multiple eucaryotic species, including humans. We report here that two cloned human DNA IVR-OD homologs, Hut2 and Hut17, consist of a series of tandem repeats of the trimer AGG/TCC, forming segments (313 and 221 base pairs in length, respectively) of polypurine/polypyrimidine (pPu/pPy or "Puppy") asymmetry in the two DNA strands; these are punctuated at certain sites with variant trimers, which are different for the two clones. Sequences homologous to the Hut2 pPu/pPy tract exist at multiple sites in the DNA of a wide variety of eucaryotes. Hybridization of human DNA with a Hut2 probe or with a previously described chicken DNA pPu/pPy sequence indicates that pPu/pPy sequences can be grouped into families distinguishable by the extent of their homology with each probe at different hybridization stringencies. Moreover, particular pPu/pPy tracts show species-specific differences in their distribution. Both the Hut2 and Hut17 pPu/pPy tracts are cleaved by S1 nuclease when tested on supercoiled plasmids. Most if not all of the 313-base-pair Hut2 pPu/pPy tract is also sensitive to S1 in its native location in HeLa cell chromatin, indicating that the sequence contains conformational information that can be expressed in vivo. This view is supported by evidence that exogenously derived Hut2 pPu/pPy tracts introduced into mouse L cells and integrated in chromatin can assume an S1-sensitive conformation.

Relatively little is known about the role of dispersed, moderately repetitive sequences in the eucaryotic genome. Suggested functions include a possible role in the coordinate regulation of gene expression (4, 9-11, 32, 64), in chromosome pairing and recombination (28, 32, 33, 37), in the processing of mRNA precursors (9, 11, 31, 32), and in DNA replication (31, 32, 65, 66).

The TU elements are a heterogeneous family of modularly structured transposons discovered originally in the genome of the sea urchin *Strongylocentrotus purpuratus* (29, 39). We have shown previously that sequences from the outer domain (OD) of the long terminal inverted repeat (IVR) segments of TU elements hybridize to DNA sequences dispersed at multiple sites in all eucaryotic organisms surveyed, including humans (38). The 15-base-pair (bp) tandemly repeating unit making up the TU element IVR-OD segment is similar to DNA sequence motifs found in the IVRs of transposable elements of *Drosophila* species and maize as well as to sequences located in the transcriptional control regions of some eucaryotic viral and cellular genes (38). In each case, segments of purine/pyrimidine asymmetry are seen in the two DNA strands.

We report here that certain IVR-OD homologs cloned from the human genome consist of long (200 to 300 bp) stretches of polypurine/polypyrimidine (pPu/pPy) asymmetry. Using these DNA segments and a previously described tract of pPu/pPy asymmetry cloned from the chicken genome (13) as probes, we show that a variety of experimen-

tally distinguishable pPu/pPy ("Puppy") sequences exist within the DNA of a wide range of eucaryotic organisms, exhibiting species-specific differences in their distribution. At least some of the pPu/pPy sequences found in the human genome can alter DNA conformation in vivo in chromatin, as well as in vitro in naked supercoiled DNA.

MATERIALS AND METHODS

General methods. Restriction endonuclease mapping was performed as described by Smith and Birnstiel (60). Single, double, and triple enzyme digestions were performed to confirm the end-labeling data. Restriction enzymes obtained from New England BioLabs, Inc., (Beverly, Mass.) or Bethesda Research Laboratories, Inc. (Gaithersburg, Md.) were used according to the instructions of the manufacturers, and DNA fragments were prepared as described previously (29). Both acrylamide and agarose gels were run in TBE buffer (0.089 M Tris borate [pH 8.3], 20 mM EDTA). Fragments of λ HindIII- and ϕ X174 replicative-form *Hae*III-digested DNA were used as molecular weight standards.

Screening of human λ bacteriophage library for IVR-OD-hybridizing clones. λ Hut clones were obtained by screening a bacteriophage λ library of human genomic DNA (see reference 19, library 1), using the IVR-OD fragment of TU1 (39) as a probe. Phage blots were preincubated in hybridization buffer (5 $\times$ SSC [1 $\times$ SSC is 0.15 M NaCl plus 0.015 M sodium citrate], 50 mM sodium phosphate [pH 7], 0.05% each of bovine serum albumin, Ficoll, and polyvinylpyrrolidone) for 3 to 6 hours at 60°C and hybridized overnight in the same solution with 100,000 cpm of the nick-translated (57) DNA fragment per ml. Plaque lifts were washed in 0.1 $\times$ SSC-0.1% sodium dodecyl sulfate-50 mM sodium phosphate (pH 7) at 50°C for 3 h and exposed as described previously (29). In all

* Corresponding author.

† Present address: Department of Biochemistry and Biophysics, School of Medicine, University of Pennsylvania, Philadelphia, PA 19104-6059.

other experiments with the IVR-OD fragment as the probe, prehybridization, hybridization, and washing conditions were as described above. Bacteriophage DNA was isolated as described by Maniatis et al. (43).

Recombinant plasmids: construction and identification. To obtain pHut2 and pHut17, λ Hut2 and λ Hut17 DNAs were digested with *Eco*RI and the fragments were then ligated to pBR322. For both psHut2 and ptk-Hut2-1, the recessed ends generated by the different restriction endonucleases used to digest pHut2 and the vectors pBR322 and ptk2 were filled in with *Escherichia coli* DNA polymerase Klenow fragment (Bethesda Research Laboratories), and the appropriate fragments were ligated to the appropriate vectors. In all cases the recombinant plasmids were introduced by transformation (8) into *E. coli* HB101. Transformants were selected by ampicillin resistance, and bacteria were lysed by the method of Thayer (67) and screened with the IVR-OD (see Fig. 1A and B) or fHut2 (see Fig. 2C and D) probe.

DNA sources. λ phage DNA was obtained from Bethesda Research Laboratories and *E. coli* DNA was from Worthington Diagnostics (Freehold, N.J.). DNA from chicken embryo, mouse L cells, monkey CV-1 cells, and human HeLa cells was prepared by the procedures described by Maniatis et al. (43). Sea urchin sperm DNA from *S. purpuratus* was prepared as described previously (38). DNAs from *Saccharomyces cerevisiae* (yeast), maize, *Drosophila melanogaster*, *Aplysia* species, and *Xenopus laevis* were obtained as generous gifts from J. Chenevert, T. Huynh, P. Beachy, R. Britten, and T. Mohun, respectively.

Hybridization of Southern and phage blots. The Southern blots (G1) shown in Fig. 4 were preincubated in hybridization buffer (4 $\times$ SSC, 50 mM sodium phosphate [pH 7], 0.05% each of bovine serum albumin, Ficoll, and polyvinylpyrrolidone) for 3 to 6 h at 68°C and then incubated overnight at the same temperature in hybridization buffer containing 0.5 $\times$ 10⁶ cpm of nick-translated probe per ml. Washing conditions and exposure for autoradiography were the same as for screening the λ phage library.

The phage blots shown in Fig. 5 and 6 were preincubated in hybridization buffer for 3 to 6 h and then incubated overnight in the presence of 100,000 cpm of the nick-translated probe per ml in the same buffer and at the same temperature. The hybridization conditions were as follows: for $\geq$ 90% sequence homology, 50% formamide, 4 $\times$ SSC, 50 mM sodium phosphate buffer (pH 7), 56°C; for $\geq$ 80% sequence homology, the same hybridization buffer as for 90%, but the hybridization was carried out at 40°C; for $\geq$ 70% sequence homology, the hybridization buffer contained 30% formamide but conditions were otherwise the same as for 80%. The washing conditions for the 90 and 80% lifts were the same as for screening the λ phage library; for the 70% lifts, the washing buffer contained 1 $\times$ SSC, but was otherwise unchanged.

Cell culture, transfections, preparation of isolated nuclei, and S1 nuclease digestions. Both Ltk⁻ and HeLa cells were grown in Dulbecco modified Eagle medium (GIBCO Laboratories, Grand Island, N.Y.) supplemented with 10% fetal calf serum. Transfections by the DNA-calcium phosphate precipitate method were performed as described by Wigler et al. (69). Stably transformed Tk⁺ cells were selected in hypoxanthine-aminopterin-thymidine medium. Isolation of nuclei and S1 nuclease treatment of nuclei were by the methods of Weintraub (68). S1 nuclease was obtained from Boehringer Mannheim Biochemicals (Indianapolis, Ind.), and 100 units of S1 per μ g of DNA was used for isolated nuclei, which were incubated at 37°C for 30 min. DNA from

nuclei was purified as described by Larsen and Weintraub (35). After restriction endonuclease digestion, 10 μ g of DNA was applied to each well. Hybridization and washing conditions were as described above for Fig. 4. S1 treatment of plasmid DNA was carried out by the method of Dybvig et al. (13).

DNA sequencing. DNA sequencing was performed by the technique of Maxam and Gilbert (45), using the four reactions G, GA, TC, and C. The sequence data were confirmed with the M13 sequencing system (47).

RESULTS

Isolation and characterization of Hut-containing clones. We screened a bacteriophage λ library of human genomic DNA (19) using the IVR-OD fragment of the TU element TU1 (39) as a probe. DNA from 10 phage isolates was digested with *Eco*RI, and fragments that hybridized to the IVR-OD probe were identified by Southern blot analysis (data not shown). In all 10 isolates (termed λ Hut2, -3, -4, -5, -10, -11, -12, -14, -15, -17), different sized fragments hybridized to the IVR-OD probe, showing that each isolate is a separate clone and that the IVR-OD-hybridizing sequences are dispersed in the human genome.

IVR-OD-hybridizing fragments from several of the λ clones were introduced into the pBR322 plasmid vector, yielding corresponding pHut subclones. Restriction endonuclease cleavage sites in pHut2 were mapped by the method of Smith and Birnstein (60), which involves partial digestion of end-labeled DNA fragments. Partial *Mn*II digestion of the larger of the two *Eco*RI-*Hpa*I fragments (Fig. 1A), end labeled at its *Eco*RI site, gave a smear on the gel in the region corresponding to the 0.9- to 1.25-kilobase size range (data not shown), indicating that multiple *Mn*II cleavage sites are present within the region located 0.9 to 1.25 kilobases from the *Eco*RI site. This segment of pHut2 DNA, which maps between the *Bst*NI and *Hin*I sites (Fig. 1A, thick line), is also the region of homology with the IVR-OD probe, as determined by Southern blot hybridization of the appropriate restriction digests (data not shown); digestion of pHut2 DNA with *Mn*II destroyed its ability to hybridize with the IVR-OD probe (Fig. 2).

No IVR-OD-hybridizing bands were seen after *Mn*II digestion of DNA from the pHut3, -4, -14, -5, -11, -12, -15, and -17 clones (Fig. 2), showing that their IVR-OD-hybridizing regions contain multiple *Mn*II cleavage sites. However, pHut10 DNA showed two hybridizing bands after *Mn*II digestion, indicating that *Mn*II cleaves only once within the IVR-OD-hybridizing region of this clone.

We isolated a fragment (termed fHut2) containing the IVR-OD-hybridizing region of pHut2 (indicated in Fig. 1A). All the plasmids except pHut10 hybridized to the fHut2 probe (data not shown). Thus, at least two variants exist among the cloned IVR-OD-hybridizing sequences; one is homologous to both the IVR-OD probe and to fHut2 and contains multiple *Mn*II recognition sites; the other (contained within pHut10) hybridizes to the IVR-OD probe but not to the fHut2 probe and lacks multiple *Mn*II sites.

Sequence determination for IVR-OD-hybridizing DNA. The DNA sequences of the IVR-OD-hybridizing regions of pHut2 (Fig. 1A) and pHut17 (Fig. 1B) show purine/pyrimidine asymmetry on the two DNA strands (Fig. 3), as found previously for the IVR-OD region of the TU element probe (39). The sequence on the polypyrimidine-rich strand consists in both cases of the trimer TCC tandemly repeated many times; however, different variations in this trimer-

FIG. 4. Southern blot analysis of genomic DNA isolated from the sources indicated. DNA was digested with *AluI* (A) or *MnII* (M), electrophoresed on 1.2% agarose gels, and hybridized to fHut2 after blotting onto filters. Amounts of DNA equal to the genome equivalents contained in 1.5 μ g of sea urchin DNA were used in all digestions, and bacteriophage λ DNA was included as an internal marker to indicate complete digestion. Dros., *D. melanogaster*.

Presence of Hut2-hybridizing sequences in the DNA of a wide variety of eucaryotic organisms. Probes made from DNA fragments (1, 2, 3, and 4 in Fig. 1A) containing the segments that flank the λ Hut2 and pHut2 pPu/pPy sequences were hybridized to Southern blots of restriction digests of all the λ Hut clones as well as to *EcoRI*-digested human DNA. These probes did not hybridize to any of the Hut cloned

DNAs other than Hut2 and appear to be single copy within the human genome (data not shown). However, sequences annealing with Hut2 under hybridization conditions that detect 80% homology are highly repeated in the human genome, as shown by the broad smear seen upon hybridization of the fHut2 probe to Southern blots of *AluI*-digested human DNA (Fig. 4). When human DNA was digested with

TABLE 1. Tracts of purine/pyrimidine asymmetry in various animal species^a

Sequence	Species (references)
1. (AGCTGCCCCCTCCTGCTGCTCCAGGTGCTCC) ₃₆	Human (26)
2. (TGCTCCCGCATCTTCTCCTCC) ₄₆	Human (26)
3. (TAGCTTGCCCCCTGCTCCTTC) ₁₅	Rat (24)
4. (ACCCGCGTCCCCCTCCTGC) ₅₅	Human (26)
5. (CCCCTCCCCACTGT) ₁₂₋₃₉	Human (52)
6. (CCCCACACCCCTGT) ₂₆₋₂₀₉	Human (1)
7. (AGCTCTCACCTCCC) ₉	Mouse (48)
8. (PyCCTTTCTPyCTTT) ₅	Human (23)
9. (TCTCCC) ₁₂	Mouse (20,56)
10. (CCCTGC) ₁₇	Human (26)
11. (TCTCC) ₃₂₋₃₆	Mouse (2,12), chick (13,44)
12. (CTTCC) ₂₇	Chick (44)
13. (TTTCC) ₅₅	Chick (44)
14. (TCCC) ₈	Chick (18)
15. (TTTC) ₁₁₋₄₃	Human (50), mouse (20)
16. (CCCG) ₃₅₋₅₂	Human (52)
17. (TCC) ₁₀₋₄₁₆	Human, mouse (7,26), rat (54)
18. (TC) ₂₇₋₁₅₀	Mouse (2,20,21,56), sea urchin (63)
19. pPu/pPy without apparent DRU.....	Human (6,41,42), chick (25), drosophila (62)

^a In all cases the sequence of the pyrimidine-rich strand is shown. The subscript next to sequences indicates the number of direct repeating units (DRUs) in the region of purine/pyrimidine asymmetry. The segments that lack a DRU (number 19) are >100 bp long and show greater than 90% purine/pyrimidine asymmetry. All the DRUs of sequence 8, contained within 100 bp of the human β -interferon promoter, are not contiguous. Sequence elements 1, 2, 4, 5, 7, 9, 11, 14, and 17 contain multiple *MnII* recognition sites (-CCTC-), whereas single base-pair changes found in some of the DRUs of the sequences indicated in lines 10, 12, and 13 generate *MnII* recognition sites.

FIG. 5. (A) Plaque lifts of λ Hut2 or λ Hut17 hybridized with fHut2 under conditions that form stable hybrids at $\geq 90\%$ sequence homology. (B) Triplicate plaque lifts of phage from a human genomic library hybridized to nick-translated fHut2 under conditions that give stable hybrids at $\geq 90\%$ sequence homology, $\geq 80\%$ sequence homology, and $\geq 70\%$ sequence homology. There were 2,000 to 3,000 plaques on each plate. The arrow indicates a plaque that hybridizes strongly at 90% homology, as well as at 70% and 80% homology. Stars were used for orientation of multiple plaque lifts.

MnII, such hybridization was almost completely obliterated, indicating that nearly all the Hut2-hybridizing sequences in the human genome contain multiple *MnII* recognition sites (CCTC). Hut2-hybridizing sequences were present also in DNA samples from all other eucaryotic organisms tested, ranging from yeasts to monkeys; in each case the hybridization diminished considerably when the DNA was digested with *MnII* before Southern blot analysis. No hybridization with either *E. coli* DNA or λ DNA was seen (Fig. 4).

Existence of a variety of experimentally distinguishable types of pPu/pPy sequences. Multiple stretches of pPu/pPy asymmetry greater than 100 bp in length are known to exist in the genomes of a wide variety of eucaryotes (29) (Table 1). Since the hybridization conditions used for the Southern blots shown in Fig. 4 can detect DNA sequences having less than complete homology with Hut2 and Hut17, the hybridizing sequences potentially could include a multiplicity of previously identified pPu/pPy sequences such as the tandemly repeated pentamer $\begin{smallmatrix} \text{TCTCC} \\ \text{AGAGG} \end{smallmatrix}$ found in high copy number in both the mouse and chicken genomes (2, 13). The extent of variability of the pPu/pPy sequences detected with the Hut probes was therefore examined by hybridizing triplicate plaque lifts from a human genomic library to the fHut2 probe, employing increasing stringencies of hybridization so that stable hybrids would form at 70, 80, or 90% sequence homology. The plaque number surveyed represented 1/10th of the DNA content of the human genome. As

controls, plaque lifts of λ Hut2 and λ Hut17 were hybridized to the fHut2 probe (Fig. 5A); under hybridization conditions that detect signals at 90% homology, DNA from λ Hut17 plaques hybridized to the fHut2 probe, but gave weaker signals than λ Hut2 plaques hybridized to the same probe. Thus, even sequences as similar to each other as Hut2 and Hut17 (Fig. 3) are distinguishable.

With increasingly stringent conditions, most hybridization signals identified with the fHut probe disappeared, but some remained and maintained their intensity (Fig. 5B). The plaque identified by an arrow in Fig. 5B is an example of one that still hybridized strongly under conditions that require at least 90% homology. Such clones contain sequences that appear to be either identical to Hut2 or more similar to Hut2 than is Hut17; we infer that the sequences showing weak hybridization with the fHut2 probe at the highest stringency are nevertheless as homologous to Hut2 as is Hut17 (Fig. 5). The frequency of plaques that hybridized strongly under conditions requiring a high degree of homology indicates that such sequences are present about 70 times in the haploid human genome.

For each set of hybridization conditions, we infer that signals of equivalent intensity represent sequences equally different from the fHut2 probe sequence. Based on the variety of types of signals observed, we conclude that multiple families of pPu/pPy sequences are present in the human genome; we calculate that Hut2 hybridizes to about

FIG. 6. Duplicate plaque lifts from a human genomic library were hybridized to either fHut2 or 21H probes under conditions in which stable hybrids form at $\geq 70\%$ sequence homology, $\geq 80\%$ sequence homology, or $\geq 90\%$ sequence homology. Circles on the lift hybridized with fHut2 at $\geq 90\%$ homology indicate positions of plaques hybridizing to 21H probe at the equivalent stringency. There were 1,500 to 3,000 plaques on each plate. The p21H plasmid was obtained from M. J. Schlesinger, and the 21H DNA fragment, containing the pPu/pPy tract, used as a probe for hybridizations was the *HincII-HaeIII* fragment described previously (13).

35,000 sequence elements per haploid genome under conditions that allow for formation of stable hybrids at 70% sequence homology.

To detect other pPu/pPy sequences in the human genome and to further distinguish among the ones already detected by hybridization of the fHut2 probe to multiple plaque lifts, we used a probe obtained from a chicken DNA clone, 21H, which is present in multiple copies in the chicken genome and which contains the sequence $\begin{smallmatrix} \text{TCTCC} \\ \text{AGAGG} \end{smallmatrix}$ tandemly repeated 32 times (13). Duplicate plaque lifts from a human genomic library were hybridized either to fHut2 or to the 21H probe; the results for one plate are shown in Fig. 6 for each set of annealing conditions used. At a level of stringency that detects 90% homology, no signals common to these chicken and human DNA probes were observed. As the hybridization stringency was lowered, an overlap in the hybridization signals for the fHut2 and 21H probes was seen; the amount of overlap increased as the stringency decreased. Under conditions that detect 70% sequence homology, half of the plaques hybridizing with the 21H probe also hybridized with fHut2. Taking into account this overlap, we estimate from our data that at least 46,000 pPu/pPy elements are present in the human genome.

The above experiments show that the pPu/pPy sequences in the human genome can be grouped by the extent of their homology with specific DNA probes. Those sequences that hybridize at high stringency with fHut2 (hybrids of 90% sequence homology) make up one family, and those that hybridize with the 21H probe under the same conditions

comprise another family. If, as our data imply, sequences hybridizing with one or the other of these probes only at lower stringency are members of still other sequence element families, we can infer the existence of a superfamily of pPu/pPy sequences that includes the various families identified by these experiments.

Species differences in distribution of different pPu/pPy families. Genomic DNA from a variety of eucaryotic organisms was digested with *AluI* and subjected to Southern blot analysis with either fHut2 (Fig. 7A) or a probe containing the 21H sequence (Fig. 7B) under hybridization conditions expected to detect 90% homology. The results of this experiment indicate that different pPu/pPy families are widely distributed among eucaryotic organisms but that the size of these families varies from species to species. Neither the yeast nor *Xenopus* genomes appear to contain members of the 21H family, as defined by hybridization with the chicken 21H probe under high-stringency conditions (Fig. 7B). In the case of maize, yeasts, *D. melanogaster*, and *X. laevis*, little or no hybridization with the fHut2 probe was detected at conditions of high stringency. However, when the stringency was reduced such that stable hybrids could form at 80% homology, many copies of hybridizing sequences were detected for maize, *D. melanogaster*, and *X. laevis*, and a few copies were detected for yeasts (Fig. 4). These observations suggest that few if any members of the Hut2 family are present in the genomes of these organisms but that other pPu/pPy sequences exist in those genomes.

S1 nuclease sensitivity of Hut2 and Hut17 sequences in

supercoiled plasmids. S1 nuclease cleaves specific DNA sequences near DNase I-hypersensitive sites known to be associated with active chromatin (35); such S1 sensitivity has been interpreted as indicating the presence of an altered DNA conformation in the region of cleavage. Using supercoiled naked DNA, mapping of S1-sensitive sites located near, and 5' to, various cloned genes has demonstrated that the S1 sensitivity *in vitro* resides in short stretches of pPu/pPy asymmetry that range from less than 10 to 42 bp in

FIG. 7. Genomic DNA from the species indicated was digested with *AluI*, electrophoresed on 1.2% agarose gels, blotted onto filters, and hybridized with fHut2 probe (A) or 21H probe (B) under conditions which give stable hybrids at $\geq 90\%$ sequence homology. The hybridization conditions were the same as described for the experiments presented in Fig. 5 except that 0.5×10^6 cpm of nick-translated probe per ml were used. Ma, Maize; Ye, yeast; Dr, *D. melanogaster*; Ap, *Aplysia* species; S.U., sea urchin (*S. purpuratus*); Xe, *X. laevis*; Ch, chick; Mo, mouse; Mon, monkey; Hu, human.

length (17, 18, 30, 40, 51, 58, 59). The pPu/pPy tract found in pHut2 was similarly tested for S1 sensitivity by digesting supercoiled plasmid DNA from psHut2 (fHut2 DNA subcloned in pBR322; Fig. 1C) with S1 nuclease and subsequently with both *EcoRI* and *HindIII*. In an analogous experiment, S1-treated pHut17 DNA (Fig. 1B) was digested with *EcoRI* and *BamHI* (Fig. 8A). For psHut2, the 390-bp cleavage fragment, and for pHut17 the 950-bp cleavage fragment (both containing the pPu/pPy sequences of the plasmid), disappeared as a result of S1 digestion. A total of 313 bp of the 357-bp insert subcloned in psHut2 consists of a pPu/pPy sequence, making it highly probable that the S1-sensitive sites in the fragment reside in the pPu/pPy tract. No sites in the pBR322 vector were cleaved by S1 nuclease under the conditions we used (data not shown).

When the supercoiled psHut2 and pHut17 plasmids were first digested with restriction enzymes and subsequently digested with S1 nuclease, the fragments containing the pPu/pPy sequence were no longer cleaved by S1 nuclease (Fig. 8A, lanes 3 and 6), indicating that the S1 sensitivity observed for these plasmids is abolished by linearization of the DNA. Analogous findings have been reported with regard to previously studied sequences that show S1 sensitivity *in vitro* (13, 18, 35, 51).

pPu/pPy sequences of Hut2 form S1 nuclease-sensitive structures *in vivo* in chromatin after transfection. To determine whether the pPu/pPy sequences we cloned specify the capacity to yield S1-sensitive sites when introduced *in vivo* and stably integrated in the genome, mouse L cells deficient in thymidine kinase production (*Ltk⁻*) were transfected with ptk-Hut2-1 (Fig. 1D), a plasmid containing the pPu/pPy tract from pHut2, plus some of its flanking sequences, inserted upstream from the herpes simplex virus thymidine kinase gene promoter. Cloned transformants were selected by survival in hypoxanthine-aminopterin-thymidine medium. Isolated nuclei from four cloned cell lines were treated with S1 nuclease, DNA purified from the nuclei was cleaved with *BamHI* and fractionated on an agarose gel, and the Southern blot was hybridized with the nick-translated *BamHI*-*AccI* fragment, which flanks the pPu/pPy sequence in pHut2 (Fig. 1A and D). In all cases the same results were obtained.

The data for the ptk-Hut2-1 plasmid and for one cell line are shown in Fig. 8B. As seen in the figure, the 1,350-bp hybridizing band containing the pPu/pPy tract of the plasmid was decreased sharply in amount after treatment with S1, and a hybridizing smear of fragments ranging from 400 to 700 bp was seen, demonstrating that the Hut2 sequence is cleaved at multiple sites that span the pPu/pPy tract. For S1 nuclease-treated isolated nuclei (Fig. 8B, lane 4), the 1,350-bp band was less intense than for untreated nuclei, and there was a hybridizing smear containing fragments that comigrate with the most intensely hybridizing region within the smear in lane 2. These data show that at least some of the exogenously derived pPu/pPy tracts are S1 sensitive *in vivo* when integrated into chromosomes. The failure of S1 nuclease to completely eliminate the 1,350-bp band (lane 4) may indicate that some of the integrated Hut2 sequences are in a conformation that is not sensitive to S1. Thus, while pPu/pPy nucleotide sequences impart the potential for altering DNA conformation in chromatin, other factors may be involved in determining whether this conformation actually is assumed *in vivo* by a particular pPu/pPy sequence.

S1 sensitivity of endogenous pPu/pPy sequences. Using the *HinfI*-*BamHI* fragment generated from pHut2 as a hybridization probe, we found that HeLa cell DNA digested with both *SsrI* and *HpaI* contains a 1,130-bp fragment that

FIG. 8. (A) S1 nuclease digestion of psHut2 and pHut17. Supercoiled plasmid DNA (0.5 μ g) was digested with S1 (0 units, lanes 1 and 4; 50 units, lanes 2 and 5), and subsequently psHut2 was digested with *Eco*RI and *Hind*III and pHut17 was digested with *Eco*RI and *Bam*HI. Plasmids which were first digested with restriction enzymes and subsequently digested with S1 nuclease are shown in lanes 3 and 6. A portion of each sample was electrophoresed on a 5% acrylamide gel and stained with ethidium bromide. The lane marked M contains *Hae*III-digested ϕ X174 (replicative form) fragments. (B) Some pPu/pPy tracts in ptk-Hut2-1 (Fig. 1D) are S1 nuclease sensitive in isolated nuclei from stably transfected Ltk⁻ cells. Isolated nuclei from Ltk⁻ cells stably transfected with ptk-Hut2-1, as well as the naked supercoiled plasmid, were digested with S1 nuclease. The purified DNA was digested with *Bam*HI, and after electrophoreses on a 1% agarose gel, the Southern blot was hybridized to the *Bam*HI-*Acc*I fragment. Lanes 1 and 2 contain plasmid ptk2-Hut2-1 DNA, untreated (lane 1) or S1 treated (lane 2). Lanes 3 and 4 contain DNA from transfected Ltk⁻ cells whose nuclei were untreated (lane 3) or S1 nuclease treated (lane 4). (C) S1 sensitivity of endogenous pPu/pPy sequences. Plasmid Hut2 DNA (lanes 1 to 4) was not treated (lanes 1 and 3) or treated (lanes 2 and 4) with S1 nuclease. Isolated nuclei from HeLa cells (lanes 5 to 8) were not treated (lanes 5 and 7) or treated with S1 nuclease (lanes 6 and 8). Purified DNA was digested with either *Sst*I and *Hpa*I (lanes 1, 2, 5, and 6) or *Hinf*I and *Hpa*I (lanes 3, 4, 7, and 8) followed by fractionation on a 1% agarose gel. The Southern blot was hybridized to the *Hinf*I-*Bam*HI fragment from pHut2 (Fig. 1A).

comigrates with the *Sst*I-*Hpa*I fragment (Fig. 8C, lanes 1 and 5) containing the pPu/pPy tract of the pHut2 plasmid. Moreover, the comigrating fragment was the only hybridizing band detected in HeLa cell DNA. These findings enabled us to investigate whether in isolated nuclei from HeLa cells the endogenous pPu/pPy tract flanking the *Hinf*I-*Hpa*I fragment (Fig. 1A) contains any S1-susceptible sites.

Purified DNA samples from untreated or S1 nuclease-treated pHut2 DNA and isolated HeLa nuclei were digested with either *Sst*I and *Hpa*I or *Hinf*I and *Hpa*I. For the untreated plasmid DNA and HeLa nuclei, the *Hinf*I-*Hpa*I fragments comigrated (Fig. 8C, lanes 3 and 7), as did the previously noted *Sst*I-*Hpa*I fragments (Fig. 8C, lanes 1 and 5). Treatment of plasmid pHut2 DNA with S1 nuclease and subsequently with both *Sst*I and *Hpa*I generated, in addition to a weakly hybridizing 1,130-bp *Sst*I-*Hpa*I band, a hybridizing smear located at a position corresponding to the size range 550 to 870 bp (Fig. 8C, lane 2), demonstrating again that multiple S1 nuclease-sensitive sites occur within the pPu/pPy tract of pHut2. In contrast, digestion of the same S1-treated plasmid DNA with *Hinf*I plus *Hpa*I yielded a hybridizing band of the same size seen with plasmid DNA that was not subjected to S1 treatment, showing that there

are no S1 nuclease-sensitive sites within the *Hinf*I-*Hpa*I region flanking the pPu/pPy tract. Similar results were obtained with the S1 nuclease-treated nuclei isolated from HeLa cells: the *Sst*I-*Hpa*I fragment almost disappeared and a hybridizing smear was seen between gel positions 600 and 900 bp (Fig. 8C, lane 6). In addition, the hybridizing fragment for *Hinf*I-*Hpa*I-digested DNA from S1-treated nuclei comigrated with the band seen in DNA from untreated nuclei (Fig. 8C, lanes 7 and 8). The S1 sensitivity of the entire endogenous pPu/pPy tract flanking the *Hinf*I-*Hpa*I (single-copy) fragment in the HeLa cell genome indicates that the conformation of this 300-bp region is altered *in vivo* in chromatin.

DISCUSSION

Our data indicate that the human genome contains multiple experimentally distinguishable families of pPu/pPy sequences and suggest that these can be grouped into what we have termed the pPu/pPy ("Puppy") superfamily. From results obtained with representatives of two different families of pPu/pPy sequences as hybridization probes (the human Hut2 sequence and the chicken 21H sequence), we

estimate that at least 46,000 pPu/pPy elements exist per haploid genome in human DNA. Given the previously identified stretches of purine/pyrimidine asymmetry in DNA (Table 1) and assuming that most of these are present and repetitive in the human genome, a conservative extrapolation is that 100,000 or more pPu/pPy elements may be present per haploid human genome. Interestingly, most of the human pPu/pPy sequence elements we analyzed were in close physical proximity to an *Alu* family sequence (our unpublished data), as has been found also for some *KpnI* family members (49).

Distinguishable pPu/pPy families also occur widely in other eucaryotic species, although species-specific differences in their distribution were seen. The species conservation of the human Hut2 and chicken 21H sequence families, together with previous observations indicating that sequences showing purine/pyrimidine asymmetry occur widely in DNA from a variety of higher organisms (3) (Table 1), supports the notion that a superfamily of pPu/pPy sequences may exist in the DNAs of all eucaryotic organisms.

It has been proposed, most notably by Britten and Davidson (4, 9, 10), that dispersed repetitive sequences are involved in gene regulation. Sequences homologous to the Epstein-Barr virus IR3 region, which consists of stretches of purine/pyrimidine asymmetry (26), are found on every chromosome except Y (27), which contains only a short segment capable of carrying functional genes (16) and shows only low levels of transcriptional activity (34). Earlier data have shown that the short tracts of purine/pyrimidine asymmetry (10 to 42 bp in length) found near the 5' ends of certain cloned genes are associated with S1 sensitivity in naked supercoiled DNA (17, 18, 30, 40, 51, 58, 59). For the α - and β -globin genes, S1-sensitive sites have been found in vivo, and their transcription has been shown to be developmentally regulated (35). S1-sensitive sites have also been identified in the cloned DNAs of histone gene spacer regions (22, 28), mapping within a tract of GA sequences.

The Hut2 and Hut17 human DNA sequences are sensitive to S1 nuclease when they are contained on supercoiled plasmids. The multiple sites of S1 sensitivity in both the Hut2 and Hut17 pPu/pPy tracts implies that alteration of DNA structure is occurring over the entire region of pPu/pPy asymmetry and not just at one or two sites. The S1 nuclease sensitivity of the Hut2 sequence in nuclei isolated from HeLa cells indicates that the endogenous Hut2 pPu/pPy sequence can assume an altered conformation in vivo in chromatin, as well as in vitro in supercoiled plasmid DNA. The single-copy region flanking Hut2 is transcribed (unpublished data). However, we do not know at this time whether transcription of the flanking DNA and the S1 sensitivity of the nearby Hut2 sequence are functionally related.

The structural conformation resulting in the observed S1 sensitivity of the pPu/pPy sequences is not single-stranded DNA (5, 53) but may be left-handed helices (5) or triplexes (36). Particular types of DNA sequence tracts are known to be able to adopt specific secondary structures in appropriate environments, and various biological roles for non-B DNA have been suggested (55). In some instances proteins that may be involved in the generation, stabilization, or recognition of S1-sensitive DNA regions have been described (14, 15).

The ability of the Hut2 pPu/pPy element to alter DNA conformation over a long tract in chromatin indicates that this pPu/pPy sequence contains conformational information. Many previously identified stretches of purine/pyrimidine asymmetry are greater than 200 bp long and some extend up

to several kilobases in length (26) (Table 1). It is possible that some or all of these segments also have the potential to assume an altered DNA conformation in vivo, although such alterations may not occur in all cells, or even at all times in a given cell, for a particular pPu/pPy sequence. The multiplicity of dispersed pPu/pPy sequence elements in the eucaryotic genome and evidence that at least some pPu/pPy elements have the potential for assuming an altered DNA conformation in vivo over long stretches leads one to speculate that such sequences may significantly affect overall chromatin structure and possibly also gene expression. The sequence of a particular pPu/pPy tract, its length, and its location relative to a gene potentially might be relevant to a putative biological role.

ACKNOWLEDGMENTS

These studies were supported by American Cancer Society grant MV44 to S.N.C. and by grants from the National Institutes of Health (Public Health Service grant) GM17995 and the U.S. Veterans Administration to L.H.K..

LITERATURE CITED

1. Bell, G. I., M. J. Selby, and W. J. Rutter. 1982. The highly polymorphic region near the human insulin gene is composed of simple tandemly repeating sequences. *Nature (London)* **295**:31-35.
2. Bennett, K. L., R. E. Hill, D. F. Pietras, M. Woodworth-Gutai, C. Kane-Haas, J. M. Houston, J. K. Heath, and N. D. Hastie. 1984. Most highly repeated dispersed DNA families in the mouse genome. *Mol. Cell. Biol.* **4**:1561-1571.
3. Birnboim, H. C., R. R. Sederoff, and M. C. Paterson. 1979. Distribution of polypyrimidine polypurine segments in DNA from diverse organisms. *Eur. J. Biochem.* **98**:301-307.
4. Britten, R. J., and E. H. Davidson. 1969. Gene regulation for higher cells: a theory. *Science* **165**:349-357.
5. Cantor, C. R., and A. Efstratiadis. 1984. Possible structures of homopurine-homopyrimidine S1-hypersensitive sites. *Nucleic Acids Res.* **12**:8059-8072.
6. Christophe, D., B. Cabrer, A. Bacolla, H. Targovnik, V. Pohl, and G. Vassart. 1985. An unusually long poly(purine)-poly(pyrimidine) sequence is located upstream from the human thyroglobulin gene. *Nucleic Acids Res.* **13**:5127-5143.
7. Cohen, J. B., K. Effron, G. Rechavi, Y. Ben-Neriah, R. Zakut, and D. Givol. 1982. Simple DNA sequences in homologous flanking regions near immunoglobulin V_H genes: a role in gene interaction? *Nucleic Acids Res.* **11**:3353-3369.
8. Cohen, S. N., A. C. Y. Chang, and L. Hsu. 1972. Nonchromosomal antibiotic resistance in bacteria: genetic transformation of *Escherichia coli* by R-factor DNA. *Proc. Natl. Acad. Sci. USA* **69**:2110-2114.
9. Davidson, E. H., and R. J. Britten. 1979. Regulation of gene expression: possible role of repetitive sequences. *Science* **204**:1052-1059.
10. Davidson, E. H., H. T. Jacobs, and R. J. Britten. 1983. Very short repeats and coordinate induction of genes. *Nature (London)* **301**:468-470.
11. Davidson, E. H., and J. W. Posakony. 1982. Repetitive sequence transcripts in development. *Nature (London)* **297**:633-635.
12. Davis, M. M., S. K. Kim, and L. E. Hood. 1980. DNA-sequences mediating class switching in α -immunoglobulins. *Science* **209**:1360-1365.
13. Dybvig, K., C. D. Clark, G. Aliperti, and M. J. Schlesinger. 1983. A chicken repetitive DNA sequence that is highly sensitive to single strand specific endonucleases. *Nucleic Acids Res.* **11**:8495-8508.
14. Emerson, B. M., and G. Felsenfeld. 1984. Specific factor conferring nuclease hypersensitivity at the 5' end of the chicken adult β -globin gene. *Proc. Natl. Acad. Sci. USA* **81**:95-99.
15. Emerson, B. M., C. D. Lewis, and G. Felsenfeld. 1985. Interaction of specific nuclear factors with the nuclease-hypersensitive region of the chicken adult β -globin gene: nature of the binding

- domain. *Cell* 41:21-30.
16. Emery, A. E. H., and D. L. Rimoim, (ed.). 1983. Principles and practice of medical genetics, vol. 1, p. 58-59. Churchill Livingstone, Ltd., Edinburgh.
 17. Evans, T., E. Schon, G. Gora-Maslak, J. Patterson, and A. Efstradiadis. 1984. S1-hypersensitive sites in eukaryotic promoter regions. *Nucleic Acids Res.* 12:8043-8058.
 18. Finer, M. H., E. J. Fodor, H. Boedtker, and P. Doty. 1984. Endonuclease S1-sensitive site in chicken pro- $\alpha 2(I)$ collagen 5' flanking gene region. *Proc. Natl. Acad. Sci. USA* 81:1659-1663.
 19. Fritsch, E. F., R. M. Lawn, and T. Maniatis. 1980. Molecular cloning and characterization of the human β -like globin gene cluster. *Cell* 19:959-972.
 20. Gilliam, A. C., A. Shen, J. E. Richards, F. R. Blattner, J. F. Mushinski, and P. W. Tucker. 1984. Illegitimate recombination generates a class switch from C μ to C δ in an IgD-secreting plasmacytoma. *Proc. Natl. Acad. Sci. USA* 81:4164-4168.
 21. Gillies, S. D., V. Folsom, and S. Tonegawa. 1984. Cell type-specific enhancer element associated with a mouse MHC gene, E β . *Nature (London)* 310:594-597.
 22. Glikin, G. C., G. Gargiulo, L. Rena-Descalzi, and A. Worcel. 1983. *Escherichia coli* single-strand binding protein stabilizes specific denatured sites in superhelical DNA. *Nature (London)* 303:770-774.
 23. Goodbourn, S., K. Zinn, and T. Maniatis. 1985. Human β -interferon gene expression is regulated by an inducible enhancer element. *Cell* 41:509-520.
 24. Harris, S. E., P. Mansson, D. B. Tully, and B. Burkhart. 1983. Seminal vesicle secretion IV gene: allelic difference due to a series of 20 base-pair direct repeats within an intron. *Proc. Natl. Acad. Sci. USA* 80:6460-6464.
 25. Heilig, R., R. Muraskowsky, and J. Mandel. 1982. The ovalbumin gene family. The 5' end region of the X and Y genes. *J. Mol. Biol.* 156:1-19.
 26. Heller, M., E. Flemington, E. Kieff, and P. Deininger. 1985. Repeat arrays in cellular DNA related to the Epstein-Barr virus IR3 repeat. *Mol. Cell. Biol.* 5:457-465.
 27. Heller, M., A. Henderson, and E. Kieff. 1982. Repeat array in Epstein-Barr virus DNA is related to cell DNA sequences interspersed on human chromosomes. *Proc. Natl. Acad. Sci. USA* 79:5916-5920.
 28. Hentschel, C. C. 1982. Homocopolymer sequences in the spacer of a sea urchin histone gene repeat are sensitive to S1 nuclease. *Nature (London)* 295:714-716.
 29. Hoffman-Liebermann, B., D. Liebermann, L. H. Kedes, and S. N. Cohen. 1985. TU elements: a heterogeneous family of modularly structured eucaryotic transposons. *Mol. Cell. Biol.* 5:991-1001.
 30. Htun, H., E. Lund, and J. E. Dahlberg. 1984. Human U1 RNA genes contain an unusually sensitive nuclease S1 cleavage site within the conserved 3' flanking region. *Proc. Natl. Acad. Sci. USA* 81:7288-7292.
 31. Jelinek, W. R., T. P. Toomey, L. Leinwand, C. H. Duncan, P. A. Biro, P. V. Choudary, S. M. Weissman, C. M. Rubin, C. M. Houck, P. L. Deininger, and C. W. Schmid. 1980. Ubiquitous, interspersed repeated sequences in mammalian genomes. *Proc. Natl. Acad. Sci. USA* 77:1398-1402.
 32. Kao, F. 1985. Human genome structure. *Int. Rev. Cytol.* 96:51-88.
 33. Kedes, L. H. 1979. Histone genes and histone messengers. *Annu. Rev. Biochem.* 48:837-870.
 34. Kuo, M. T., and W. Plunkett. 1985. Nick-translation of metaphase chromosomes: *in vitro* labeling of nuclease-hypersensitive regions in chromosomes. *Proc. Natl. Acad. Sci. USA* 82:854-858.
 35. Larsen, A., and H. Weintraub. 1982. An altered DNA conformation detected by S1 nuclease occurs at specific regions in active chick globin chromatin. *Cell* 29:609-622.
 36. Lee, J. S., M. L. Woodsworth, L. J. P. Latimer, and A. R. Morgan. 1984. Poly(pyrimidine) · poly(purine) synthetic DNAs containing 5-methylcytosine form stable triplexes at neutral pH. *Nucleic Acids Res.* 12:6603-6614.
 37. Lehrman, M. A., W. J. Schneider, T. C. Sudhof, M. S. Brown, J. L. Goldstein, and D. W. Russell. 1985. Mutation in LDL receptor: Alu-Alu recombination deletes exons encoding transmembrane and cytoplasmic domains. *Science* 227:140-146.
 38. Liebermann, D., B. Hoffman-Liebermann, A. B. Troutt, L. H. Kedes, and S. N. Cohen. 1986. Sequences from the sea urchin TU transposons are conserved among multiple eucaryotic species, including humans. *Mol. Cell. Biol.* 6:218-226.
 39. Liebermann, D., B. Hoffman-Liebermann, J. Weinthal, G. Childs, R. Maxson, A. Mauron, S. N. Cohen, and L. Kedes. 1983. An unusual transposon with long terminal inverted repeats in the sea urchin *Strongylocentrotus purpuratus*. *Nature (London)* 306:342-347.
 40. Mace, H. A. F., H. R. B. Pelham, and A. A. Travers. 1983. Association of an S1 nuclease-sensitive structure with short direct repeats 5' of *Drosophila* heat shock genes. *Nature (London)* 304:555-557.
 41. Maeda, N., J. B. Bliska, and O. Smithies. 1983. Recombination and balanced chromosome polymorphism suggested by DNA sequences 5' to the human δ -globin gene. *Proc. Natl. Acad. Sci. USA* 80:5012-5016.
 42. Maeda, N., F. Yang, D. R. Barnett, B. H. Bowman, and O. Smithies. 1984. Duplication within the haptoglobin Hp² gene. *Nature (London)* 309:131-135.
 43. Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. Molecular cloning, a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
 44. Maroteaux, L., R. Heilig, D. Dupret, and J. L. Mandel. 1983. Repetitive satellite-like sequences are present within or upstream from 3 avian protein-coding genes. *Nucleic Acids Res.* 11:1227-1243.
 45. Maxam, A., and W. Gilbert. 1977. A new method for sequencing DNA. *Proc. Natl. Acad. Sci. USA* 74:560-564.
 46. McKnight, S. L., and E. R. Gavis. 1980. Expression of the herpes thymidine kinase gene in *Xenopus laevis* oocytes: an assay for the study of deletion mutants constructed *in vitro*. *Nucleic Acids Res.* 8:5931-5948.
 47. Messing, J. 1983. New M13 vectors for cloning. *Methods Enzymol.* 101:20-78.
 48. Miyake, S., T. Yokota, F. Lee, and K. Arai. 1985. Structure of the chromosomal gene for murine interleukin 3. *Proc. Natl. Acad. Sci. USA* 82:316-320.
 49. Miyake, T., K. Migita, and Y. Sakaki. 1983. Some KpnI family members are associated with the Alu family in the human genome. *Nucleic Acids Res.* 11:6837-6846.
 50. Moos, M., and D. Gallwitz. 1983. Structure of two human β -actin-related processed genes one of which is located next to a simple repetitive sequence. *EMBO J.* 2:757-761.
 51. Nickol, J. M., and G. Felsenfeld. 1983. DNA conformation at the 5' end of the chicken adult β -globin gene. *Cell* 35:467-477.
 52. Proudfoot, N. J., A. Gil, and T. Maniatis. 1982. The structure of the human zeta-globin gene and a closely linked nearly identical pseudogene. *Cell* 31:553-563.
 53. Pulleyblank, D. E., D. B. Haniford, and A. R. Morgan. 1985. A structural basis for S1 nuclease sensitivity of double-stranded DNA. *Cell* 42:271-280.
 54. Qasba, P. K., and S. K. Safaya. 1984. Similarity of the nucleotide sequences of rat α -lactalbumin and chicken lysozyme genes. *Nature (London)* 308:377-380.
 55. Rich, A., A. Nordheim, and A. H. J. Wang. 1984. The chemistry and biology of left-handed Z-DNA. *Annu. Rev. Biochem.* 53:791-846.
 56. Richards, J. E., A. C. Gilliam, A. Shen, P. W. Tucker, and F. R. Blattner. 1983. Unusual sequences in the murine immunoglobulin μ - δ heavy-chain region. *Nature (London)* 306:483-487.
 57. Rigby, P. W., M. Dieckmann, C. Rhodes, and P. Berg. 1977. Labeling deoxyribonucleic acid to high specific activity *in vitro* by nick translation with DNA polymerase I. *J. Mol. Biol.* 113:237-251.
 58. Ruiz-Carrillo, A. 1984. The histone H5 gene is flanked by S1 hypersensitive structures. *Nucleic Acids Res.* 12:6473-6492.
 59. Schon, E., T. Evans, J. Welsh and A. Efstradiadis. 1983. Conformation of promoter DNA: fine mapping of S1-hypersensitive sites. *Cell* 35:837-848.

60. Smith, H. O., and M. L. Birnstiel. 1976. A simple method for DNA restriction site mapping. *Nucleic Acids Res.* **3**:2387-2399.
61. Southern, E. M. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* **98**:503-517.
62. Southgate, R., A. Ayme, and R. Voellmy. 1983. Nucleotide sequence analysis of the *Drosophila* small heat shock gene cluster at locus 67B. *J. Mol. Biol.* **165**:35-57.
63. Sures, I., J. Lowry, and L. H. Kedes. 1978. The DNA sequence of sea urchin (*S. purpuratus*) H2A, H2B and H3 histone coding and spacer regions. *Cell* **15**:1033-1044.
64. Sutcliffe, J. G., R. J. Milner, J. M. Gottesfeld, and R. A. Lerner. 1984. Identifier sequences are transcribed specifically in brain. *Nature (London)* **308**:237-241.
65. Szostak, J. W., and E. H. Blackburn. 1982. Cloning yeast telomeres on linear plasmid vectors. *Cell* **29**:245-255.
66. Taylor, J. H., and S. Watanabe. 1981. Eukaryotic origins: studies of a cloned segment from *Xenopus laevis* and comparisons with human blur clones, p. 597-606. In D. S. Ray (ed.), *The initiation of DNA replication*. Academic Press Inc., New York.
67. Thayer, R. E. 1979. An improved method for detecting foreign DNA in plasmids of *E. coli*. *Anal. Biochem.* **98**:60-63.
68. Weintraub, H. 1983. A dominant role for DNA secondary structure in forming hypersensitive structures in chromatin. *Cell* **32**:1191-1203.
69. Wigler, M., R. Sweet, G. K. Sim, B. Wold, A. Pellicer, E. Lacy, T. Maniatis, S. Silverstein, and R. Axel. 1979. Transformation of mammalian cells with genes from procaryotes and eucaryotes. *Cell* **16**:777-785.