Volume 9 Number 3 1981 Nucleic Acids Research

Nucleotide sequence of mouse satellite DNA

Wolfram Hörz and Werner Altenburger[1]

Institut für Physiologische Chemie, Physikalische Biochemie und Zellbiologie der Universität München, Goethestrasse 33, 8000 München 2, GFR

Received 24 November 1980

ABSTRACT

The nucleotide sequence of uncloned mouse satellite DNA has been determined by analyzing Sau96I restriction fragments that correspond to the repeat unit of the satellite DNA. An unambiguous sequence of 234 bp has been obtained. The sequence of the first 250 bases from dimeric satellite fragments present in Sau96I limit digests corresponds almost exactly to two tandemly arranged monomer sequences including a complete Sau96I site in the center. This is in agreement with the hypothesis that a low level of divergence which cannot be detected in sequence analyses of uncloned DNA is responsible for the appearance of dimeric fragments.Most of the sequence of the 5% fraction of Sau96 monomers that are susceptible to TaqI has also been determined and has been found to agree completely with the prototype sequence. The monomer sequence is internally repetitious being composed of eight diverged subrepeats. The divergence pattern has interesting implications for theories on the evolution of mouse satellite DNA.

INTRODUCTION

The discovery of satellite DNAs in eukaryotic organisms dates back twenty years when minor components were found in mouse and crab DNA by CsCl density gradient centrifugation (2,3, for a recent review, see ref. 4). In the mouse this fraction constitutes 5-10% of the total DNA, and it was recognized later that it differed from the rest of the mouse DNA by its very rapid rate of reassociation (5). This phenomenon was interpreted to mean that mouse satellite DNA is composed of an approximately 300 bp nucleotide sequence which is repeated a million fold per genome (5).

Detailed information on the structure of mouse satellite DNA became available when restriction nucleases were used to analyze the eukaryotic genome. It turned out that cleavage sites for a number of different restriction nucleases are spaced at highly

regular intervals in mouse satellite DNA (6-8). By the use of EcoRII it became apparent that essentially all of the mouse satellite consists of tandem repeats of approximately 240 bp and that there is evidence for subrepeats of 1/2, 1/4, and 1/8 this value. This picture of homogeneity was challenged by the finding that cleavage sites for many other restriction nucleases are spaced at the same periodicity but are clustered on only a fraction of the repeats (7-9). It was concluded that the satellite DNA is composed of different segments, each susceptible to a particular restriction nuclease. Evolutionary schemes were put forth to explain these features of the satellite structure (7,8).

Information on the fine structure of the satellite repeat units came from an analysis of oligopyrimidine isostichs and oligonucleotides from satellite transcripts (10-12). The data support the view that the DNA is composed of a basic repeating oligonucleotide and variants thereof. Direct sequence analysis by primed DNA synthesis of an EcoRI derived fragment cloned in a λ phage vector yielded a linear sequence of about 70 bp (13). Simultaneous sequence analysis of total satellite DNA primed at the EcoRII restriction site gave a sequence of the same length, which included regions of high ambiguity but could be aligned with the sequence of the cloned fragment (13).

In the course of our investigations on the structure of mouse satellite DNA and satellite DNA containing chromatin we reexamined the question of sequence analysis of uncloned mouse satellite DNA and could establish an unambiguous prototype sequence of 234 bp representative of the one million copies of the satellite repeat unit.

MATERIALS AND METHODS

Restriction nucleases and DNA. Sau96I and TaqI were prepared according to Greene et al. (14) and were gifts of R.E. Streeck. MnlI was from New England Biolabs, Beverly, MA. Mouse satellite DNA was prepared as previously described (8).

Terminal labeling, strand separation, and sequence analysis. Restriction fragments were isolated from polyacrylamide gels, labeled, and subjected to sequence analysis according to Maxam

and Gilbert (15) with the modifications described by Altenburger et al. (16). Strand separation of DNA fragments was carried out by heating the DNA in 50% dimethylsulfoxyde (15), 1 mM Tris-HCl, pH 7.4, 0.1 mM EDTA to 100° for 5 min followed by rapid cooling and electrophoresis in polyacrylamide gels (5% acrylamide, 0.25% N,N'-methylene bisacrylamide), run in 10 mM triethanolamine-HCl, pH 7.4, 2 mM EDTA. This buffer originally suggested by Varshavsky et al. (17) for the separation of deoxyribonucleoprotein particles largely prevented renaturation of denatured fragments due to its low ionic strength and gave good resolution of the two strands of mouse satellite DNA fragments.

RESULTS

The pattern obtained upon digestion of mouse satellite DNA with the restriction nuclease Sau96I closely resembles the one obtained previously with EcoRII (7,8) which we termed a type A pattern. All of the DNA is cleaved into a series of fragments, the predominant one being a unit length fragment of about 240 bp. In addition there are small amounts of fragments of 2,3,4 times, $^1/_2$, $1^1/_2$, $2^1/_2$ etc. times and $^1/_4$, $^3/_4$, $1^1/_4$, $1^3/_4$ etc. the length of the unit size fragment. A number of other restriction nucleases yield what we called type B patterns. They cleave only a small fraction of the satellite DNA into fragments of the same periodicity, but with a size distribution that is skewed towards the higher multiples.

In order to obtain a representative nucleotide sequence, we first analyzed the Sau96 monomer fragments which constitute 60-70% of the total satellite DNA digest. This fraction was isolated from a preparative polyacrylamide gel and labeled at the 5'-ends with polynucleotide kinase. Since further cleavage of the labeled fragment with EcoRII was unsuccessful (see below) we subjected the Sau96 monomer fraction to strand separation. The two strands were well separated (not shown) and both used for the sequencing reactions.

An unambiguous sequence of 234 bases could be established from both strands of the Sau96 monomer fragment (see Fig. 1) with hardly any indication of alternative bases throughout the entire sequence. The 234 bp can be arranged in four homologous

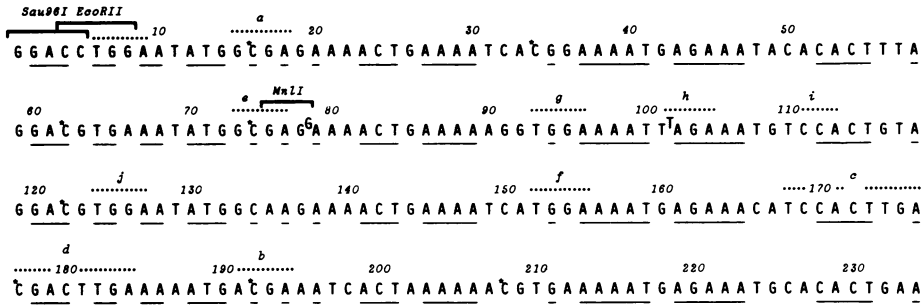

Figure 1. Prototype sequence of the repeat unit of mouse satellite DNA. The nucleotide sequence of both strands of Sau96 monomer fragments was determined and the A-rich strand is shown. Nucleotide homologies in the four subrepeats are underlined. Restriction nuclease cleavage sites are indicated by brackets. Asterisks denote methylation of C residues as deduced from the presence of a gap or only a faint band in the sequencing gel and a band corresponding to a G residue in the complementary strand (18,19). Methylation of C is extensive in all of the 8 CpG pairs on both strands of the monomer sequence except in the CpG dinucleotide at position 208/209 which is only about 50% methylated on either strand. All tetranucleotide sequences which differ in only one nucleotide from the TaqI site (TCGA) are indicated by little circles. The small letters suggest a possible correlation between these sites and small amounts of fragments present after TaqI cleavage of the Sau96 monomer as shown in Fig. 2.

subrepeats, three times 58 bp and one time 60 bp long. It is evident that the Sau96 and EcoRII sites partially overlap which explains why Sau96 monomers fail to be digested by EcoRII and vice versa. Because of this overlap a difference of a single nucleotide in the third subrepeat (position 123) as compared to the first one destroys both, the Sau96 and the EcoRII cleavage site.

The prototype sequence contains also an MnlI restriction site, GAGG (20), extending from position 75 to 78 (Fig. 1). The CTCC C-residue complementary to G at position 75 is extensively methylated, however, and apparently methylation at this C is sufficient to prevent MnlI cleavage since only a very weak partial digest was obtained upon digestion of the satellite DNA with MnlI (not shown).

Minor components of mouse satellite DNA also conform to the prototype sequence. In a Sau96I or EcoRII limit digest approxi-

mately 20% of all repeat units are present as dimeric fragments (7,8). This is compatible with the assumption that some 3% mutations had eliminated cleavage sites on these repeat units. The dimeric fragments would therefore be expected to have a sequence very closely related to two tandemly arranged monomer sequences. In order to test this assumption and establish the sequence across the region of the Sau96 cleavage site we determined the first 250 bases on the T-rich strand of dimeric fragments from a Sau96 limit digest. The sequence was almost identical to the prototype sequence including the complete Sau96 restriction site at position 235-239. The only difference was that at position 6 a mixture of T and A and at position 37 a mixture of T and C was detected rather than just T at both positions as in the monomer fragment. This corresponds to the presence of both A and G at position 202 in the sequence as written in Fig. 1 and A and T at position 233.

In a further effort to shed light on sequence variants we turned to nucleases which yield type B digestion patterns. Cleavage sites for these nucleases are present on only a few percent of the repeat units. These repeats are clustered thus giving rise to the characteristic segments of the satellite DNA (7,8). TaqI has been shown by Brown and Dover to generate a very strong type B pattern (9). When the Sau96 monomer fraction is digested with TaqI, 5-10% of the molecules are susceptible and yield two major fragments of about 70 bp and 165 bp length (Fig. 2). In addition many faint bands are present in the digestion pattern. As seen from the sequencing gels the smaller fragment corresponds to the A-rich strand and comprises 72 bases which could be read unambiguously up to position 68. The first 152 bases of the T-rich strand in the large fragment were determined which in the numbering of the complementary strand in Fig. 1 covers the region from position 83 to position 234. The entire sequence that could be read was identical to the prototype sequence. The Taq segment has apparently arisen from a mutation at position 73 changing the sequence GCGA to TCGA and a subsequent spreading of repeat units containing the newly acquired Taq site. The fact that the C residue is most likely methylated in these repeats does not interfere with Taq cleav-

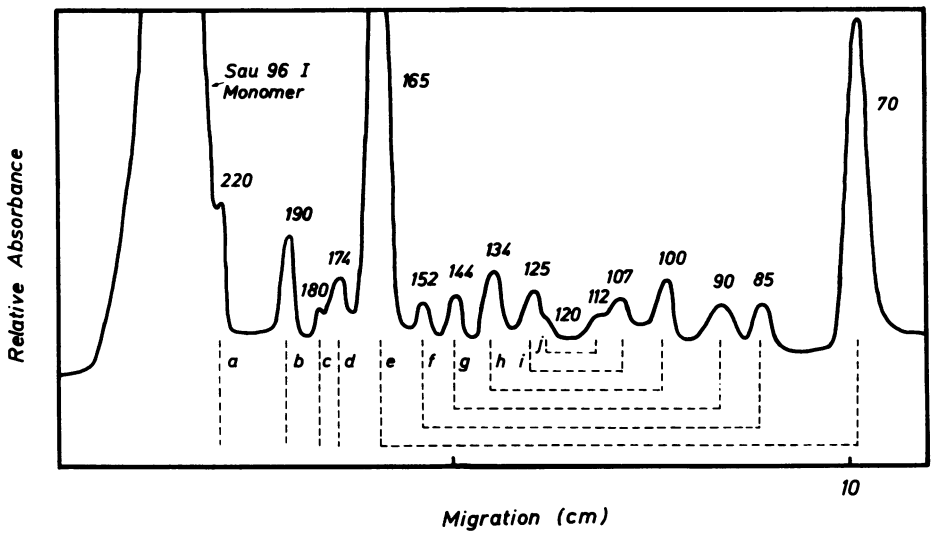

Figure 2. TaqI sites on variants of the satellite repeat unit. Sau96 monomer fragments were digested to completion with TaqI and the digest analyzed by electrophoresis in a 5% polyacrylamide gel. The gel was stained with ethidium bromide, photographed, and the negative scanned as described (21). The sizes of the fragments are given in bp as determined from their electrophoretic mobilities relative to satellite monomer fragments, half and quarter molecules. The ten largest fragments resulting from Taq digestion are assigned letters a to j. The remaining six fragments complement fragments e to j. Digestion of pBR322 with TaqI in control experiments gave no evidence of overdigestion indicating that cleavage only occurs at complete TaqI sites under the conditions used.

age (22).

The sizes of the other fragments present in very small amounts in the Taq digest of the Sau96 monomer (Fig. 2) were determined from their electrophoretic mobilities. As indicated in Fig. 2 the twelve smaller fragments add up pairwise to approximately the size of the monomer fragment. Fragments complementary in size to the four larger ones were not resolved in the gel. There are 13 sites, indicated in Fig. 1, that differ in only one base pair from the TaqI recognition site. The sizes of the fragments in Fig. 2 correspond very well to those predicted from these sites consistent with the view that in a small number of repeats divergence has generated TaqI sites in at least 10 of

the 13 positions. Cleavage at the first site (position 6-9) cannot be monitored since the product would not be resolved from undigested monomer, and it cannot be decided if fragments c and d should be assigned to only two or more of the four successive sites between position 168 and 184.

DISCUSSION

The prototype sequence of mouse satellite DNA. The demonstration that the mouse satellite DNA repeat unit is 234 bp long sets an end to a long lasting uncertainty. Previous determinations based on the electrophoretic mobility of restriction fragments of this satellite have yielded rather divergent values (6-8, 23-26) due to an anomalous behavior of mouse satellite DNA in electrophoresis. Undigested mouse satellite DNA has been found to migrate more slowly in polyacrylamide gels than main band DNA (27). The effects of base composition are particularly pronounced in polyacrylamide gels (28). This explains why in these gels the repeat size of the satellite is determined to be 255 bp (not shown), but even in agarose gels a value of 240±2 bp was obtained (26). The GC content of mouse satellite DNA as calculated from the sequence is 36.8% in good agreement with values of 35.2% (29) and 34.2% (30) derived from nucleoside analysis. The anomalous mobility in gel electrophoresis is, however, probably not just due to the low GC content but is, like buoyant density and melting point anomalies, a result of the unusual arrangement of nucleotides in this sequence. Particularly striking is the unequal distribution of A and T residues in the two strands, the basis for strand separation by alkaline CsCl density gradient centrifugation (30) and gel electrophoresis (31). The heavy strand in CsCl, which is the slow strand in gel electrophoresis, contains 45.7% T and only 17.5% A.

The CpG content of the prototype sequence is 3.4%, roughly the percentage expected on a statistical basis and clearly higher than the 1.1% reported for total mouse DNA (32). The sequence analysis indicates that almost all CpG pairs are extensively methylated, in agreement with previous estimates of 3-4% for the m^5C content of mouse satellite DNA (33,34). We have no evidence for C-methylation except in CpG pairs even

though low levels of m^5C at other positions would not be detected by the sequence analysis. On the other hand, methylation of C in only the CpG pairs of the prototype sequence can explain the m^5C containing pyrimidine oligonucleotides perfectly which have been detected in a direct analysis of the distribution of m^5C in mouse L-cell satellite DNA (35).

The prototype sequence contains restriction sites for EcoRII, Sau96I, and MnlI, although the MnlI site is largely methylated and the satellite therefore almost resistant to cleavage by MnlI. The site cleaved by Sau96I belongs to the subset of Sau96I sites (GGNCC) also cleaved by AvaII ($GG\begin{smallmatrix}A\\T\end{smallmatrix}CC$). It was this enzyme which was originally reported (36) to give a type A cleavage pattern with mouse satellite DNA like EcoRII. Sites of dyad symmetry in the sequence which are potential candidates for restriction nucleases still to be discovered are ATAT, ACGT, TGCA, and CATG.

There are many regions in the prototype sequence where an exchange of one base pair would generate a restriction nuclease cleavage site. Such exchanges with subsequent local amplification have been invoked to explain the segmental structure of the satellite which is responsible for the type B restriction nuclease digestion patterns (7,8). The sites for a number of such nucleases have been mapped relative to the EcoRII site (8). HindII and Bsu sites were both found at 47/53% of the EcoRII monomer that is around position 115 of the sequence as numbered in Fig. 1. There is indeed a sequence GTCCAC extending from position 109-114 that can give rise to a Bsu site by a T $\longrightarrow$ G or a Hind II site by a C $\longrightarrow$ purine transversion. Similarly, there is a potential AluI site at 42% (AACT, 140-143) and the two potential RI/RI* sites at 16% and 38% (GAAATC, 193-198; AAAATC, 145-150). The site generating the Taq segment which Brown and Dover (9) mapped approximately 80 bp away from the EcoRII site is GCGA at position 73-76. As is obvious for TaqI, the mere presence of a potential site does not necessarily lead to a sufficient number of repeats with complete sites to generate a segment. This is true also for the other nucleases: there are 2 potential sites for HindII, 7 for Bsu and 5 for AluI. It cannot be decided at present if in each case mutations at only

one of the potential sites were that much more frequent or if the corresponding repeats were spread around to a greater extent subsequent to the mutations.

The satellite repeat unit is internally repetitious. It is immediately apparent that the repeat unit of the satellite DNA consists of four subrepeats as shown in Fig. 1. There is, however, significant internal repetition of an even shorter periodicity equivalent to $^1/_8$ of the repeat size. As shown in Fig. 3a, units of 28 bp (designated α) alternate with units of 30 bp (designated ß). If the sequence is arranged as to obtain maximum homology all ß units contain a one nucleotide gap around the center and an extra trinucleotide on the right side. Even though these trinucleotides differ in each ß unit (ACA,GTC,ATC,GCA) they are clearly related. The only irregularities in the lengths of the subrepeats are a G and a T "insertion" at positions 78 and 102, respectively.

The consensus sequence derived from the 8 subrepeats is shown in Fig. 3b). Omitting the extra C in all α subunits and the additional trinucleotides in the ß units the consensus sequence can be derived from a hypothetical ancestral sequence consisting of three related nonanucleotides, GA_5TGA, GA_6CT, GA_5CGT, as shown in Fig. 3c). These oligonucleotides are identical to or a variation of those found originally in RNA transcripts of mouse satellite DNA (12). The other three related sequences occurring with a high frequency in the cRNA transcript (GA_4CTGA, GA_4TGA, GA_3TG (12)) are each found at least twice in the prototype sequence.

The finding of four prominent oligonucleotides in RNA transcripts together with the results of restriction nuclease digestion experiments led Southern in 1975 to propose a scheme for the evolution of mouse satellite DNA (7). This scheme entailed a series of multiplication steps with intermittent periods of divergence to create successively 60, 120, and 240 bp periodicities starting with a sequence 9-18 bp long. From analyses by restriction nucleases we could demonstrate a weak periodicity also at the 30 bp level (8) which is confirmed by the sequence analysis. The present day prototype sequence is in complete agreement with the general outline of the scheme proposed by

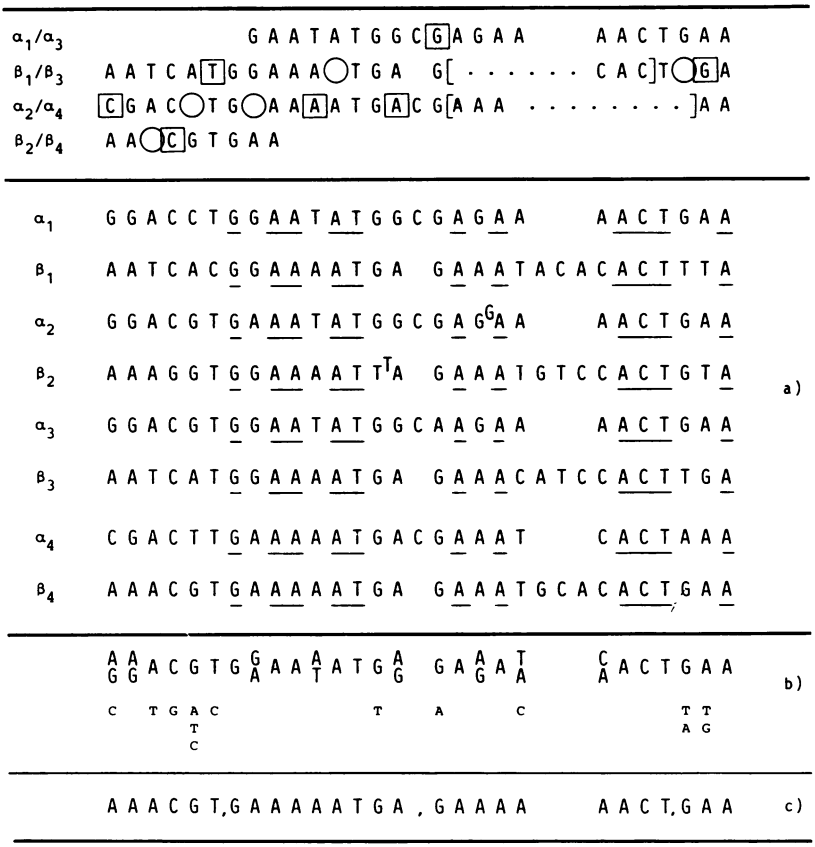

Figure 3. Internal repetitivity of the mouse satellite DNA repeat unit. a) The 234 bp of the repeat unit are arranged in 8 subunits such as to obtain maximum homology. The subunits are numbered α_1 to β_4 as indicated on the left. Nucleotides identical in all 8 repeats are underlined. The consensus sequence derived from the eight repeats is shown in b) with nucleotides occurring only once or twice in a given position set in small print. A possible ancestral sequence of mouse satellite is shown in c). It consists of three tandemly repeated nonanucleotides related to each other: GA_5TGA, GA_6CT, GA_5CGT. At the top the most-likely sequence of a 78 bp region for total mouse satellite DNA as taken from Shmookler Reis and Biro (13) is shown. Brackets enclose ambiguous loci. This sequence corresponds to either a region from α_1 to β_2 or α_3 to β_4 of the prototype sequence written below. The nucleotide determinations of Shmookler Reis and Biro in the seven positions where $\alpha_1 \longrightarrow \beta_2$ differs from $\alpha_3 \longrightarrow \beta_4$ are framed. Except for the G residue in the top line they all correspond to $\alpha_3 \longrightarrow \beta_4$. Circles denote nucleotides missing in their sequence as compared to the sequence in a).

Southern (7). The closer homology between $ß_1$ and $ß_3$ and likewise between $ß_2$ and $ß_4$ than between $ß_1$ and $ß_2$ or $ß_4$ (Fig. 4) is in accord with an amplification of the sequence in several stages interrupted by periods of divergence. It is unexpected, however, that the first three α units are much more conserved and that $α_4$ is so different from the other α units; instead it resembles more closely a ß unit. Specific restrictions concerning mutational drift in different parts of the sequence might provide some explanation. On the other hand, there might be processes leading to a conversion of the nucleotide sequence especially in adjacent areas as shown for the globin gene family (37).

The existence of internal repeats is a serious obstacle for sequence studies by primed DNA synthesis as performed by Shmookler Reis and Biro (13). Their most likely sequence for total mouse satellite DNA on the basis of primed synthesis and oligonucleotides from RNA transcripts is shown in Fig. 3 at the top. Their sequence matches the first half of the prototype sequence ($α_1$ through $ß_2$) almost as well as the second half ($α_3 \rightarrow ß_4$) because of the internal repetition. Wherever there are differences between the two tracts (indicated by framed symbols in Fig. 3 at the top) their assignments agree with the second half except for the G residue in the top line. Their major difficulty

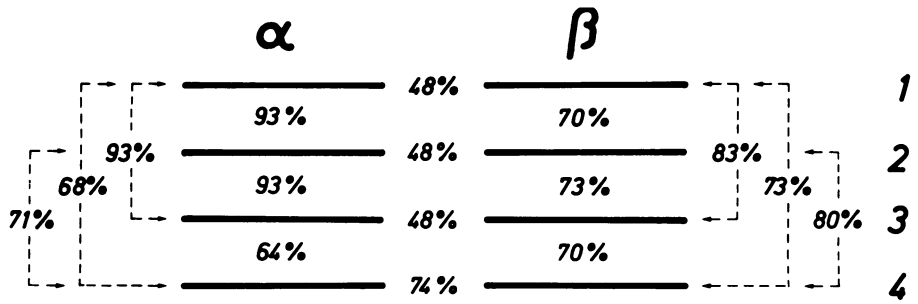

Figure 4. Homologies within α and ß subunits of mouse satellite DNA. Nucleotide sequences of all α and ß subunits as defined in Fig. 3 are compared and homologies indicated. The extra G and T residues at positions 78 and 102 are not included in the calculations. In the comparison between α and ß units a gap is taken to be a mismatch.

was in the region where ß-units have the extra trinucleotide. Their data did not allow definite sequence assignments indicating that primed synthesis had proceeded in several registers.

Variations of the prototype sequence. The prototype sequence determined from Sau96 generated monomer fragments is representative of 60-70% of the total satellite DNA. There are no ambiguities in the nucleotide assignments, and there is no evidence for minor alternative nucleotides at any position. Still it is clear from the sequence analysis that the Sau96 monomer is not a unique sequence. Comparison of the pyrimidine oligonucleotides in the prototype sequence with those previously determined (10-12) shows that most of the major tracts that were found at the time are contained in the prototype sequence but that there are a number of minor ones occurring every 300 to 1000 bp which must be derived from diverged variants.

A sequence of 250 bases from dimeric fragments turned out to be essentially identical to the prototype sequence even including the Sau96 site in the center of the molecule. This is surprising at first sight since the dimer fragment was derived from a limit digest and therefore refractory to Sau96 cleavage. It proves, however, that it is not a single base exchange in all of the molecules which is responsible for Sau96 resistance but instead divergence distributed across the five positions of the Sau96 site. In the sequence analysis, a combination of, for example, 80% G, 7% A, 7% C, and 7% T would still show up as an unambiguous G.

Variant repeats from mouse satellite DNA are conveniently accessible through the use of restriction nucleases that yield type B digestion patterns. For example only about 5% of all repeats carry a TaqI site at position 73 to 76. It would be conceivable therefore that the repeat units with TaqI sites differ from the prototype sequence also at other positions. This was not found to be the case, however. The explanation proposed for the type B digestion patterns (7,8) is that restriction sites were introduced by divergence into some repeat units which were then spread by unequal crossover between sister chromatids (7) or other mechanisms (8). Eventually such new variants can, in theory, be spread through the entire array. Monomers

with the Taq site, however, are clustered in a segment taking up about 35% of the total satellite DNA, and within this segment, on the average, one out of seven units carries the Taq site. Any other sequence variation present within this segment would presumably have been homogenized to a similar degree. It would therefore be detectable only if its presence were closely linked to the mutation generating the TaqI site since repeat units were selected in the sequence analysis only for the presence of this site.

We conclude therefore that although the monomers susceptible to TaqI constitute only a 5% minority of the total population of a diverged set, it is not surprising that their sequence conforms to the prototype sequence aside from the exchange from G to T at position 73.

Concluding Remarks. The elucidation of the nucleotide sequence of mouse satellite DNA has concluded a chapter of structural analysis at the molecular level which began twenty years ago. It has given a clear answer to the question of the internal architecture of the satellite and provided a basis for detailed evolutionary considerations. It also should set the stage for investigations of satellite DNA containing chromatin as it is organized within the cell nucleus which might give new insights into the still enigmatic question as to the function of this DNA.

ACKNOWLEDGMENT

We thank R. Gerl and H. Rothe for expert technical assistance. This work was supported by Deutsche Forschungsgemeinschaft.

REFERENCES

1. Present address: Institut für Molekularbiologie der Universität Zürich, Hönggerberg, 8093 Zürich, Switzerland
2. Kit, S. (1961) Fed. Proc. 20, 354
3. Sueoka, N. (1961) J. Mol. Biol. 3, 31-40
4. Brutlag, D.L. Ann. Rev. Genet., in press
5. Waring, M. and Britten, R.J. (1966) Science 154, 791-794
6. Hörz, W., Hess. I., and Zachau, H.G. (1974) Eur. J. Biochem. 45, 501-512
7. Southern, E.M. (1975) J. Mol. Biol. 94, 51-69

8. Hörz, W. and Zachau, H.G. (1977) Eur. J. Biochem. 73, 383-392
9. Brown, S.D.M. and Dover, G.A. (1980) Nucl. Acids Res. 8, 781-792
10. Harbers, K. and Spencer, J.H. (1974) Biochemistry 13, 1094-1101
11. Harbers, K., Harbers, B., and Spencer, J.H. (1974) Biochem. Biophys. Res. Commun. 58, 814-821
12. Biro, P.A., Carr-Brown, A., Southern, E.M., and Walker, P.M.B. (1975) J. Mol. Biol. 94, 71-86
13. Shmookler Reis, R.J. and Biro, P.A. (1978) J. Mol. Biol. 121, 357-374
14. Greene,P.J., Heyneker, H.L., Bolivar, F., Rodriguez, R.L., Betlach, M.C., Covarrubias, A.A., Backman, K., Russel,D.J., Tait, R., and Boyer, H.W. (1978) Nucl. Acids Res. 5, 2373-2380
15. Maxam, A.M. and Gilbert, W. (1980) Meth. Enzym. 65, 499-560
16. Altenburger, W., Steinmetz, M., and Zachau, H.G. (1980) Nature 287, 603-607
17. Varshavsky, A.J., Bakayev, V.V., and Georgiev, G.P. (1976) Nucl. Acids Res. 3, 477-492
18. Ohmori, H., Tomizawa, J., and Maxam, A.M. (1978) Nucl. Acids Res. 5, 1479-1485
19. Miller, J.R., Cartwright, E.M., Brownlee, G.G., Fedoroff, N.V., and Brown, D.D. (1978) Cell 13, 717-725
20. Roberts, R.J. (1980) Nucl. Acids Res. 8, r63-r80
21. Hörz, W. and Zachau, H.G. (1980) J. Mol. Biol., in press
22. Streeck, R.E., Gene, in press
23. Bellard, M., Oudet, P., Germond, J.-E., and Chambon, P. (1976) Eur. J. Biochem. 70, 543-553
24. Noll, M. and Kornberg, R.D. (1977) J. Mol. Biol. 109, 393-404
25. Maio, J.J. Brown, F.L., and Musich, P.R. (1977) J. Mol. Biol. 117, 637-655
26. Streeck, R.E. and Zachau, H.G. (1978) Eur. J. Biochem. 89, 267-279
27. Zeiger, R.S., Salomon, R., and Peacock, A.C. (1971) Biochemistry 10, 4219-4223
28. Thomas, M. and Davis, R.W. (1975) J. Mol. Biol. 91, 315-328
29. Corneo, G., Ginelli, E., Soave, C., and Bernardi, G. (1968) Biochem. 7, 4373-4379
30. Flamm, W.G., McCallum, M., and Walker, P.M.B. (1967) Proc. Natl. Acad. Sci. USA 57, 1729-1734
31. Zeiger, R.S., Salomon, R., Dingman, C.W., and Peacock, A.C. (1974) Biochemistry 13, 3388-3393
32. Setlow, P. (1975) in Handbook of Biochemistry and Molecular Biology, 3rd Edition, Nucleic Acids Vol. II (Fasman, G.D. ed.) pp. 312-318, CRC Press, Cleveland
33. Salomon, R., Kaye, A.M., and Herzberg, M. (1969) J. Mol. Biol. 43, 581-592
34. Bond, H.E., Flamm, W.G., Burr, H.E., and Bond, S.B. (1967) J. Mol. Biol. 27, 289-302
35. Harbers, K., Harbers, B., and Spencer, J.H. (1975) Biochem. Biophys. Res. Commun. 66, 738-746
36. Dover, G. (1978) Nature 272, 123-124
37. Slighton, J.L., Blechl, A.E., and Smithies, O. (1980) Cell 21, 627-638