Mature, closed circular mouse mitochondrial DNA contains a significant number of ribonucleotides throughout the genome. Previous studies have implicated the two origins of DNA replication as preferred sites of ribonucleotide retention. We have analyzed the site specificity of ribosubstitution by direct sizing of alkali-treated restriction fragments in comparison with the DNA sequence of untreated restriction fragments of cloned mouse mitochondrial DNA. These results have confirmed the observations that ribonucleotides are retained at the two origins of replication and are most likely remnants of RNA priming events associated with DNA replication. The map location of ribonucleotides at the light strand origin of replication has been refined to a triplex nucleotide (5'-CGG-3') in the light strand initiation region.

This approach has demonstrated that all four deoxyribonucleotides are subject to ribosubstitution and no single base (or subset of the four bases) predominates. An examination of selected regions of the mitochondrial DNA genome including the putative coding region for cytochrome oxidase subunit III and regions containing the genes for tRNA\(^{\text{CyB}},\) tRNA\(^{\text{CyC}},\) 12 S rRNA, and 16 S rRNA reveals preferred sites for ribosubstitution. These preferred sites do not relate in any obvious way to the functional aspects of these domains. In addition, the data indicate that every possible position of DNA sequences examined can be ribosubstituted at a very low frequency.

Closed circular mitochondrial DNA from a number of animal species is unique with respect to having ribonucleotides integrated throughout the genome. Previous studies have provided data on the kinetics of conversion from the supercoiled to the relaxed form after alkali or RNase H treatment and showed that the average number of integrated ribonucleotides per mtDNA molecule was between 10 and 30 for mouse LA9 mtDNA (1) and 10-18 for human HeLa mtDNA (2). Lonsdale and Jones (3) calculated the average number of ribonucleotides in rat liver mtDNA to be 13 by assaying the mean length of linear fragments generated through RNase H and S1 digestion in electron micrographs. These data were consistent with a random distribution of integration sites throughout the molecule. In contrast to mammalian mtDNA, Kolodner et al. (4) showed ribonucleotides integrated at preferred sites in chloroplast DNA from pea. The localization of alkali-labile sites at the H strand\(^1\) and L strand replication origins of mouse mtDNA allowed the first functional explanation of alkali-sensitive sites as remnants of primer ribonucleotides (5-7).

The experiments described here were designed to localize ribosubstitution sites at the nucleotide sequence level. The two origins of replication, O\(_L\) and O\(_H\), part of the COIII coding region, and the junctions of tRNA and rRNA genes were searched for ribonucleotides to determine whether these sites are dispersed truly at random or whether there exists a preference for certain nucleotides or functional domains in the mtDNA genome. Previous work from this laboratory has shown that a significant fraction of mouse mtDNA contains alkali-sensitive sites at the two origin regions to the exclusion of sites elsewhere in a large portion of the genome (7). This study also demonstrated that the frequency of origin-related or other alkali-sensitive sites was independent of the state of maturation of the molecule (i.e. parental strands of replicative intermediates, newly replicated relaxed closed circles, or fully mature closed circles) and that either strand of the genome (H or L) was digested to the same size distribution by alkali. This work did not, however, map alkali-sensitive sites at the nucleotide level in the origin regions or elsewhere in the molecule and the results did not eliminate the possibility that ribonucleotides were inserted at unique positions within any given molecule. The data reported here indicate that ribonucleotides are specifically retained at O\(_H\) in a region of less than six nucleotides. Although other regions of the genome have preferred sites for ribonucleotide retention, no other functionally definable nucleotides are maintained as ribonucleotides even at the H strand origin, O\(_L\).

**Experimental Procedures**

Mitochondrial DNA Isolation—Mouse LA9 cells were grown in suspension culture and mitochondria were isolated from cells as described by Bogenhagen and Clayton (8, 9). Closed circular mtDNA was obtained after purification through two cycles of ethidium bromide-CsCl buoyant density centrifugation with subsequent ethanol precipitation.

Isolation of Plasmid DNA—Escherichia coli SF-8-C600 r-m\(^+\) transformed with pBR322 vector containing either HR-B or HR-C inserts (the two largest HindIII + Eco RI restriction fragments of mouse mtDNA (10)) was grown without chloramphenicol amplification and plasmid DNA was isolated under P2-PE1 conditions as described by Battey and Clayton (11).

Isolation of Mouse Liver DNA—Mitochondria from mouse liver (strain Akr/xm) were isolated after homogenization in a Dounce homogenizer (12), and the mtDNA was purified as described above.

5' End Labeling and DNA Sequencing—DNA was dephosphorylated by treatment with bacterial alkaline phosphatase in a 10-μl
Ribonucleotides in Mouse Mitochondrial DNA

RESULTS

Mapping Strategy for Ribosubstitution Sites—The apparent random distribution of ribonucleotides in mouse mtDNA has been attributed to their sporadic misincorporation during replication and subsequent incomplete removal from the newly synthesized strands (1). At the L strand origin, a region of eight nucleotides in mature, closed circular mtDNA has been shown by Martens and Clayton (5) to contain a cluster of alkali-labile sites considered to be primer ribonucleotide remnants. In the H strand origin region and at the 5' termini of D-loop DNA strands, similar groups of ribonucleotides have been found (6, 7). They have also been proposed as being residues of priming events for DNA replication in this part of the molecule.

In this study, a more precise mapping of these sites in closed circular mtDNA was undertaken. 5' End-labeled and alkali-treated restriction fragments were analyzed in parallel with the sequencing ladder of the same fragment isolated from cloned mtDNA, allowing the precise positioning of integrated ribonucleotides in the mtDNA. Cloned mtDNA does not contain ribonucleotides (16). In order to obtain a representative picture of alkali-labile sites, several functionally different regions were analyzed (Fig. 1). These regions include the two origins of replication (O₁ and O₂) and two separate and distinct coding sequences in the mtDNA genome.

Ribonucleotides in the Light Strand Origin Region—The location of integrated ribonucleotides for which an obvious functional explanation can be assigned is at O₁ (5). To precisely map these sites, cloned mtDNA containing the L strand origin region and mtDNA prepared from LA9 cells and fresh mouse liver tissue were individually cleaved with Hpa I (Fig. 1). The Hpa I digestion fragments were labeled at the 5' end, subsequently cleaved with Ava II, and the appropriate 5' end-labeled fragment was recovered after separation in an agarose gel. The lengths of the 5' end-labeled L strands obtained after alkali treatment of this fragment were determined on an 8% acrylamide and 50% urea sequencing gel (Fig. 2). The three major sites of ribosubstitution in LA9 mtDNA were at the nucleotides 5'-CGG-3' corresponding to positions 209–211 in the previously assigned cluster of eight nucleotides (5). The adjacent nucleotides 5'-TA-3' (Fig. 3) are also sites of alkali cleavage, but the frequency of ribosubstitution is <10% that of the three major sites as judged by the relative intensity of the species shown in Fig. 2. The excision of these primer ribonucleotides in vivo appears to be more effective in mouse liver cells when compared to cultured LA9 cells, as only the first two nucleotides (5'-CG-3') of the triplet show significant alkali sensitivity. The fact that there is not a large disparity in the relative abundance of fragments cleaved at each of the major ribosubstitution sites suggests that most individual molecules are ribosubstituted at only one position. If a significant percentage of the population contained two or more contiguous ribonucleotides at O₁, there would be a preponderance of fragments mapping at sites proximal to the labeled 5' end. This is because the assay does not detect cleavages which occur distal to the first ribosubstitution from the labeled end. The frequency of ribosubstitution in the sequences flanking O₁ is not elevated and, in fact, is somewhat lower than in the nonorigin-related regions that have been analyzed. However, there is a higher degree of ribosubstitution in liver mtDNA 5' to O₁ as compared to LA9 mtDNA. Upon overexposure of the gel shown in Fig. 2, one observes that the nearest detectable alkali-sensitive site is 24 nucleotides 3' to the substituted C in Fig. 3 (data not shown).

Ribonucleotides in the Heavy Strand Origin Region—The region of O₂, which serves as the template for synthesis of both D-loop and actual daughter H strands has been previously sequenced (6). It is known that the 5' ends of nascent D-loop strands contain ribonucleotides. In particular, the largest D-loop strand has been shown to contain from 1–10 ribonucleotides at the 5' end (6). We, therefore, hypothesized that this site on the genome might represent a defined H strand ribosubstitution region, in direct analogy with the major tri-nucleotide ribosubstitution site in the L strand at O₁.

The restriction fragments containing the template region of O₂ were analyzed for ribosubstitution sites in both H and L strands (Fig. 4). The H strand clearly has one major and three less prominent, but distinct, sites. Surprisingly, these preferred sites lie outside O₂ and represent ribosubstitution at points 3' to the initiation of H strand replication within the initiation region (Fig. 4). Within O₂, there are many ribosubstitution sites with no evidence for a major site corresponding to the known map position of the 5' ribonucleotide end of the largest D-loop strand (6). The high frequency of ribosubstitution sites in the H strand is consistent with the observation that approximately one-half of the closed circular mtDNA population contains ribonucleotides in this region (7).

The L strand template at O₂ shows three prominent sites of alkali sensitivity. As in the case of H strand, these sites do not lie in the region of initiation. Since the L strand is initiated at O₁ (Fig. 1), the functional significance of these ribosubstitutions is unknown. Within the O₂ initiation region, the L

![Fig. 1. Map positions of the nucleotide sequences searched for ribonucleotides. The relative positions of the two rRNA genes (12S and 16S), two tRNA genes (Phe and Val), and the two origins of replication (O₁ and O₂) are shown. The boxed regions represent the searched sequences on the H strand (outer, thick circle) and L strand (inner, thin circle). The restriction map shown is limited to the regions of interest.](image-url)
Ribonucleotides in Mouse Mitochondrial DNA

Fig. 2. Ribonucleotides in the L strand sequence at O1. Cloned mtDNA and mtDNA isolated from both LA9 cells and mouse liver were 5' end-labeled at the Hpa I restriction site (Fig. 1). The L strand sequence is directed toward the Aua I cleavage site. Of the sequencing chemistries, only the G-specific cleaved fragments are shown (A). Shows the alkali-sensitive sites in mouse liver mtDNA, and B in mtDNA prepared from LA9 cells. The direction of elongation of L strand is from bottom to top (5'-3'). The two most frequently substituted nucleotides are the same in mouse liver and LA9 cells strand exhibits a single preferred G ribosubstitution located within a GC-rich portion of the sequence (Fig. 4).

Ribonucleotides in Defined Coding Regions—A coding region of the COIII gene was analyzed for a stretch of 150 nucleotides and exhibits a number of specific alkali-labile sites scattered apparently at random throughout the searched sequence (Fig. 5). No preference for any of the four nucleotides could be detected and each is found substituted by ribonucleotides. The frequency of ribosubstitution is clearly less than in the O1 and O2 regions. As predicted, alkali treatment of the cloned DNA sequence did not show any sites sensitive to alkali (Fig. 5).

From the distribution and relative amount of alkali-generated fragment, an estimate of the frequency of incorporated ribonucleotides at the more prominent sites can be approximated as 1 in 500 H strands being substituted by a ribonucleotide. This estimate is based on the relative intensity of the alkali-generated fragment as compared to the intensity of the corresponding fragment in the sequencing lanes in the same gel. Since the fragments generated by sequencing represent a standard amount of radioactivity, the relative amount of radioactivity in a corresponding alkali-generated fragment can be assessed and related to the total radioactivity of the alkali-treated sample loaded on the gel. For example, if an alkali-generated fragment is roughly equivalent to 50 cpm of isotope and the total alkali-treated sample contains 10,000 cpm of 32P end label, we estimate that approximately 1 in 200 molecules has been cleaved by alkali at that particular site. This type of assay and estimate of ribosubstitution frequency neglect the fact that if a given molecule contains a second alkali-labile site closer to the labeled 5' end, the distal site will not be scored. This is not a major problem given the fact that the overall ribosubstitution rate in this part of the genome is very low. Therefore, the number of molecules of the size of the restric-

(CG shown in the sequence in Fig. 3). These assignments are offset by one nucleotide in relationship to the sequence ladder due to the fact that the actual length of a fragment in a DNA sequence lane is one nucleotide less than a fragment containing the base which has been cleaved to generate that particular base-specific signal (14). mtDNA from LA9 cells is also substituted by ribonucleotides at high frequency at the adjacent 3' nucleotide and with diminishing frequency in the two following 3' positions.
Ribonucleotides in Mouse Mitochondrial DNA

FIG. 4. Positions of preferred ribonucleotide retention at O_H. At O_H, the alkali-sensitive sites in the H strand and the L strand were determined. For the H strand sequence, mtDNA from LA9 cells and cloned plasmid DNA were 5′ end labeled at the Sau 3A/Mbo I site (Fig. 1). Alkali-treated LA9 mtDNA in A shows multiple ribonucleotide integration sites at a much higher substitution frequency when compared to the coding sequences. The region denoted O_H encompasses nucleotides 32–109 of the sequence reported by Gillum and Clayton (6). B shows the base-specific cleavages in order from left to right: G, G + A, A > C, C + T, and C. The L strand analysis was done from the 5′ end-labeled Xba I site as shown in Fig. 1. As in the H strand sequence, A shows the alkali-generated fragments from LA9 cell mtDNA; the order of the sequencing chemistries in B is reversed for easier recognition of the overlapping sequences (L strand chemistries in B from left to right: C, C + T, G + A, and G).

Fig. 5. Distribution of alkali-labile sites in the COIII region. mtDNA from LA9 cells and cloned plasmid DNA were 5′ end labeled at the HindIII restriction site (Fig. 1), and the H strand sequence was determined. Shown are two exposures of this same sequencing gel. A shows four base-specific cleavages of cloned sequence in the following order from left to right: G, G + A, C + T, and C. B displays the alkali-labile sites in mtDNA from LA9 cells. Overexposure of the gel shows that every nucleotide in this region can be replaced at very low frequency by a ribonucleotide. C is mtDNA from LA9 cells incubated in sterile 50 mM Tris-HCl (pH 8.5) and 10 mM EDTA for 24 h at room temperature as a control for the alkali treatment. D shows alkali-treated cloned mtDNA and E the same DNA treated as described for C. As expected, no alkali-sensitive sites are detected in the cloned mtDNA sequences even on long exposure.

FIG. 4. Positions of preferred ribonucleotide retention at O_H. At O_H, the alkali-sensitive sites in the H strand and the L strand were determined. For the H strand sequence, mtDNA from LA9 cells and cloned plasmid DNA were 5′ end labeled at the Sau 3A/Mbo I site (Fig. 1). Alkali-treated LA9 mtDNA in A shows multiple ribonucleotide integration sites at a much higher substitution frequency when compared to the coding sequences. The region denoted O_H encompasses nucleotides 32–109 of the sequence reported by Gillum and Clayton (6). B shows the base-specific cleavages in order from left to right: G, G + A, A > C, C + T, and C. The L strand analysis was done from the 5′ end-labeled Xba I site as shown in Fig. 1. As in the H strand sequence, A shows the alkali-generated fragments from LA9 cell mtDNA; the order of the sequencing chemistries in B is reversed for easier recognition of the overlapping sequences (L strand chemistries in B from left to right: C, C + T, G + A, and G).

Fig. 5. Distribution of alkali-labile sites in the COIII region. mtDNA from LA9 cells and cloned plasmid DNA were 5′ end labeled at the HindIII restriction site (Fig. 1), and the H strand sequence was determined. Shown are two exposures of this same sequencing gel. A shows four base-specific cleavages of cloned sequence in the following order from left to right: G, G + A, C + T, and C. B displays the alkali-labile sites in mtDNA from LA9 cells. Overexposure of the gel shows that every nucleotide in this region can be replaced at very low frequency by a ribonucleotide. C is mtDNA from LA9 cells incubated in sterile 50 mM Tris-HCl (pH 8.5) and 10 mM EDTA for 24 h at room temperature as a control for the alkali treatment. D shows alkali-treated cloned mtDNA and E the same DNA treated as described for C. As expected, no alkali-sensitive sites are detected in the cloned mtDNA sequences even on long exposure.

DISCUSSION

Selected Regions for Analysis of Ribosubstitution—From the sizing analysis of mtDNA fragments obtained through RNase H and subsequent S1 digestion of rat liver mtDNA (3), it was clear that the ribonucleotide integration sites are not limited to a few readily discernible locations on the mitochondrial genome, but that they are distributed in a more complex pattern. The specific regions to be searched for ribosubstitution were selected to obtain a representative analysis of the mouse mtDNA genome. In addition to the two origin regions previously implicated as ribosubstitution sites, the COIII coding strand was chosen as a structural gene sequence distinct
from the two origins of replication and, thus, presumably was unaffected by features of topology or functional events associated with these origins. The COIII gene-coding strand (H strand) also represents a segment of the genome that is replicated by leading strand displacement synthesis within a duplex region (5). In contrast, the L strand of the rRNA region represents a region of the molecule in which DNA synthesis is a "gap-filling" event on a daughter molecule segregant (5). Although it is not known whether different elongation proteins are utilized in this final stage of L strand replication, the topology of the DNA primer-template is clearly very different than for H strand elongation.

The Preferred Sites of Ribosubstitution—The clearly localized ribosubstitution sites in mouse LA9 cell mtDNA are at the two replication origins. Martens and Clayton (5) had established a stretch of eight nucleotides being accessible to alkali nicking in the L strand sequence at the beginning of L strand replication. In LA9 cells, the experiments described here show the three nucleotides 5'-CGG-3' ribosubstituted with a frequency of approximately 1 in 300 molecules as judged from the relative intensity of bands on the sequencing gels. Since these three species are of roughly equal abundance, approximately 1 in 100 molecules is ribosubstituted at one of these three nucleotides.

The region of the genome with the highest frequency of ribosubstitutions is at the H strand origin. The priming of DNA replication at OH is complex with a family of D-loop strands containing ribonucleotides at the 5' ends (6). The major alkali-labile sites are not coincident with the mapped ends of these D-loop DNAs, but the ribonucleotide primers at the 5' ends of the different D-loop species, when aligned with the L strand sequence (6), generally do match with one of the less prominent ribosubstitutions in the H strand sequence. The significance of the total population of sites is unclear due to the number of ribosubstitution points in this part of the molecule. In particular, the major sites of ribosubstitution in this region lie outside the H strand origin and are not preferentially retained in the H strand. Therefore, the majority of ribonucleotides present cannot be implicated as RNA primer remnants of DNA synthesis.

General Properties of Ribosubstitution—An initial hypothesis with regard to nucleotide specificity of incorporated ribonucleotides was that ribosubstitution throughout the genome was exclusively adenine. The basis of this idea was the well known fact that mitochondria contain a high level of riboadenosine phosphates that might well result in a higher level of misincorporation of ATP as compared to other ribonucleoside triphosphates. The observation that there is no obvious preference for any of the four ribonucleotides suggests that this hypothesis is invalid.

In the L strand of the coding region for tRNA\textsuperscript{Val}, adjacent to the 12 S rRNA gene, and tRNA\textsuperscript{His}, adjacent to the 12 S and 16 S rRNA genes (10), ribonucleotide integration sites are far less frequent than in the H strand origin region. As at OH and OL, each of the four different nucleotides can be substituted by a ribonucleotide, and there is no preference for any subset of the four nucleotides. Attempts to correlate the substitution sites in the cloverleaf secondary structure of the two tRNAs (10) do not show a matching of the sites in the two structures. Thus, the position of a nucleotide in such a folded structure is not obviously relevant to a ribonucleotide misincorporation in the DNA sequence.

The experiments described here indicate that incorporation of ribonucleotides in mouse mtDNA may be due to different factors. Ribonucleotides at the two replication origins appear to be remnant sequences from priming events and the alkali-labile sites in the coding genes may be due to ribonucleotides misincorporated during replication, similar to the ribonucleotide-containing CoEI plasmids grown in the presence of chloramphenicol (18). Alternatively, if priming of replication occurs throughout the entire genome at very low frequency, these scattered ribonucleotides might be primer remnants similar to the ribonucleotide clusters at OH and OL. However, there is no evidence for discontinuous synthesis of mouse mtDNA and multiple priming events are not required due to the asynchronous mode of mtDNA replication (5).

We note that the relative frequency of ribosubstitution in different regions of the genome is variable. As much as one-half of the total closed circular mtDNA is ribosubstituted at OH (7), whereas approximately 1 in 100 molecules is ribosubstituted at OL. Roughly 1 in 500 molecules is ribosubstituted in the coding regions examined (e.g. 1 in 500 strands of the 150-nucleotide H strand of COIII). If the same frequency prevails throughout the entire ~15,000 base pair coding portion of the duplex genome, approximately 40% of the closed circular mtDNA population would contain at least one ribosubstitution site outside of the OH and OL regions. Although these frequencies are not precise, they have been observed consistently and suggest that different mechanisms are responsible for the overall pattern of ribosubstitution. Further support for the concept that ribosubstitution at the origins and coding regions is not linked is the fact that <10% of the total closed circular mtDNA population is totally devoid of ribonucleotides (7).

Acknowledgment—We thank D. P. Tapper for helpful discussions of experimental strategy.

REFERENCES