Proc. Nat. Acad. Sci. USA
Vol. 71, No. 4, pp. 1108–1112, April 1974

Interspersion of Repetitive and Single-Copy Sequences in Nuclear Ribonucleic Acid of High Molecular Weight

(RNA organization/RNA hybridization/ribosomal RNA)

DAVID S. HOLMES* AND JAMES BONNER

Division of Biology, California Institute of Technology, Pasadena, Calif. 91109

Contributed by James Bonner, December 7, 1973

ABSTRACT **Nuclear RNA of high molecular weight, approximately 15 to 30 × 10^3 nucleotides long, isolated from rat ascites cells, has been shown previously to contain both middle repetitive and single-copy sequences. However, it was not known whether these sequences were located on separate molecules or whether they were interspersed within the same molecule. The present communication distinguishes between these possibilities and shows that a large proportion of nuclear RNA of high molecular weight consists of interspersed middle repetitive and single-copy sequences.**

The occurrence of highly repetitive, middle repetitive, and single-copy sequences in the nuclear DNA of higher organisms is firmly established (1). Several studies have shown further that the majority of middle repetitive and single-copy DNA sequences are intimately interspersed (2–5) and that at least a portion of both these classes are transcribed into RNA (6). The question can be raised of whether nuclear RNA molecules of high-molecular-weight are transcribed from DNA organized with interspersed middle repetitive and single-copy sequences or from a minor fraction of the DNA with little or no interspersion. It is shown below that the majority of high-molecular-weight nuclear RNA (HnRNA) molecules, isolated from rat ascites, contain middle repetitive sequences interspersed with single-copy sequences.

MATERIALS AND METHODS

Preparation of DNA. Nuclear DNA was isolated from rat ascites as described (7). DNA was sheared to a number average length of 350 or 1500 bases (single-strand length judged by electron microscopy) by three passages through a Ribi-Sorvall cell fractionator at 50,000 or 10,000 lbs./inch^2, respectively. After shearing, DNA was re-extracted with phenol–chloroform (7).

Preparation of RNA. HnRNA (15 to 30 × 10^3 nucleotides) and 28S and 18S rRNA were prepared from rat ascites as described (7). For the preparation of HnRNA labeled *in vivo*, tumor-bearing rats were each given a 30-min intraperitoneal injection of 5 mCi of [^{3}H]uridine (New England Nuclear Corp., 180 Ci/mol). Alternatively RNA was labeled *in vitro* with [^{3}H]dimethylsulfate as described (11).

RNA Hybridization to excess DNA, monitored without ribonuclease treatment, was carried out essentially as described by Dina *et al.* (8). To determine the amount of [^{3}H]HnRNA that bound to filters without prior hybridization, 1 mg of rat nuclear DNA (1500 nucleotides, single-strand length) was denatured and reassociated in 1 ml of 3.7× SSC–47% formamide (SSC = 0.15 M NaCl–0.015 M Na_3 citrate) to various C_0ts. At the end of incubation, samples were cooled rapidly to 4° and 50 ml of ice-cold 6× SSC, containing pulse-labeled, unsheared [^{3}H]HnRNA, was added to each solution. Solutions were immediately passed through Millipore filters (47 mm, of HAWP) that had been soaked in 6× SSC. Each filter was washed with about 100 ml of 6× SSC, dried, incubated overnight at 37° in 3 ml of 0.1 M NaOH, and counted in 15 ml of Aquasol. Quenching was determined as follows. Filters were prepared with bound DNA as described above except that no [^{3}H]HnRNA was included in the solution. After drying, 5 μl of [^{3}H]HnRNA in water was added to each filter and the filters were allowed to dry, followed by incubation in 0.1 N NaOH and counting as described above. Quenching of radioactivity on these filters was determined with reference to 5 μl of [^{3}H]HnRNA counted in 15 ml of Aquasol. Quenching reduced the cpm to 50–60% of the reference value.

DNA of average length of 1500 nucleotides was chosen in this experiment to provide sufficient single-stranded material, even after hybridization with middle repetitive sequence of RNA, to bind to the filters. The use of DNA of a larger size than this could cause viscosity problems at the concentrations used in subsequent hybridization experiments.

RESULTS

Hybridization of HnRNA to Excess DNA. The rate of hybridization of RNA to excess DNA is governed predominantly by the concentration of complementary DNA sequences if the latter are present in sufficient excess over RNA sequences (9, 10). Fig. 1*a* shows the result of hybridizing uniformly labeled, sheared, HnRNA or rRNA to excess DNA. After hybridization, the samples were subjected to mild ribonuclease treatment; the extent of hybridization was monitored by the method of Melli *et al.* (10). Twelve to 24% of HnRNA hybridizes with an average rate ($C_0t_{1/2}$ = 6) slightly slower than the major component of middle repetitive DNA ($C_0t_{1/2}$ = 1), and about 38–76% hybridizes at the same average

Abbreviations: SSC, standard saline–citrate solution (0.15 M sodium chloride–0.015 M sodium citrate); ×SSC means that the concentration of the solution used is times that of the standard saline–citrate solution; HnRNA, nuclear RNA of high molecular weight; C_0t, product of concentration of nucleotides of DNA and time of incubation.

* Present address: Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, Calif. 91109.

rate ($C_0t_{1/2}$ = 1.7 × 10^3) as single-copy DNA ($C_0t_{1/2}$ = 1.6 × 10^3). Since the hybridization reaction does not go to completion, the hybridization values have been expressed either as a percentage of total input RNA (lower value) or as a percentage of RNA hybridized (i.e., normalized to 100%). The actual value depends on which of the two kinetic classes, middle repetitive or single copy, is under-represented in this experiment. However, we favor the possibility that approximately 12%, rather than 24%, of HnRNA actually consists of middle repetitive sequences for reasons discussed below. Under the same conditions, rRNA hybridizes as a single kinetic component ($C_0t_{1/2}$ = 34). Several reasons for the observation that RNA hybridization does not go to completion in the presence of excess DNA have been discussed (11), and it should be noted that only 70–90% hybridization can be achieved with model systems (10, 19).

In contrast, when the extent of hybridization of unsheared, pulse-labeled, HnRNA to excess DNA is monitored without ribonuclease treatment, as shown in Fig. 1*b*, about 70% of the radioactivity of HnRNA hybridizes by a C_0t of 80. If it is assumed that this hybridization can be described by a single second-order reaction, then the $C_0t_{1/2}$ of the reaction is about 1.8. However, due to the scatter of points, the possibility that the hybridization could be better described by multiple second-order reaction curves cannot be eliminated. We interpret these results to mean that about 70–80% of HnRNA molecules contain at least one middle repetitive sequence, if it is assumed (see below) that the RNA is uniformly labeled. Since only approximately 12–24% of HnRNA consists of middle repetitive sequences, it is concluded that the majority of single-copy sequences are covalently attached to middle repetitive sequences. Thus, without ribonuclease treatment, the bulk of single-copy sequences are withdrawn into complexes by the formation of hybrids between middle repetitive RNA sequences and DNA. It is not possible to determine the sequence composition of the 20–30% of HnRNA that remains unhybridized at C_0t 80, although a portion of this is due to the subtraction of background (see *Methods*). Further incubation of the RNA to higher C_0ts, e.g., 10^4, results in a decrease in hybridization presumably due to RNA degradation.

The hybridization of uniformly labeled, unsheared rRNA to excess DNA, determined without ribonuclease treatment, is also shown in Fig. 1*b*. Again, the scatter of points renders it difficult to determine with accuracy the kinetics of hybridization. However, assuming that the rate can be described by a single second-order reaction, then the $C_0t_{1/2}$ is approximately 19. The increase in hybridization rate of unsheared, and hence larger, HnRNA and rRNA compared to the respective sheared RNA is essentially consistent with the effect of length on reassociation rate (12). However, interpretation of the kinetics of these hybridization experiments is complicated by parameters that have not been evaluated at present, such as the effect of residual secondary structure of HnRNA and rRNA on hybridization rate and the possible presence of a spectrum of reiteration frequencies in HnRNA of families of middle repetitive sequences.

Interspersion of Middle Repetitive and Single-Copy Sequences in HnRNA. The above result shows that the majority of HnRNA molecules of 15 to 30 × 10^3 nucleotides contain at least one middle repetitive sequence. However, two major models of more detailed sequence organization can be proposed. Either middle repetitive and single-copy sequences are located predominantly on separate molecules or they are interspersed within the same molecule. If they are on sepa-

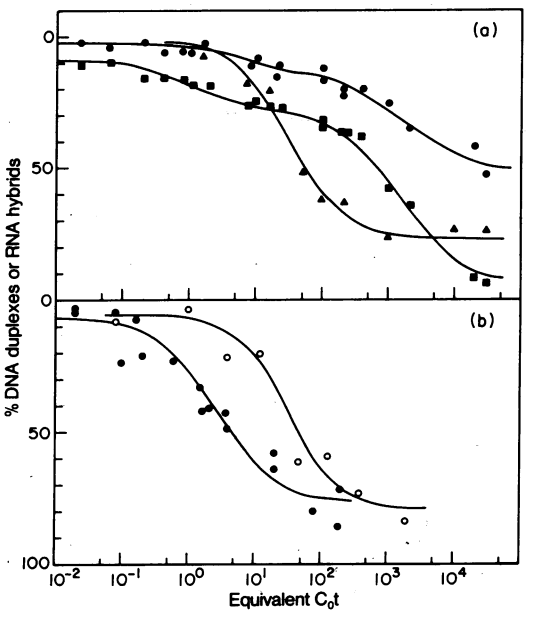

FIG. 1. Hybridization of radioactive RNA to excess unlabeled total rat DNA. (*a*) (■) DNA reassociation; (▲) hybridization of uniformly labeled, rat ascites [^{3}H]rRNA (specific activity = 110,000 cpm/μg); (●) hybridization of uniformly labeled [^{3}H]-HnRNA (specific activity = 76,000 cpm/μg), prepared by sedimentation through dimethylsulfoxide and selection of RNA of 15 to 30 × 10^3 nucleotides long (11). Mass ratio of DNA was 170,000 to HnRNA and 319,000 to rRNA. RNA was sheared by sonication to 6–8 S in size and, after heat denaturation, hybridized in solution with excess DNA (350 nucleotides single-strand length) in either 0.12 M or 0.48 M phosphate buffer, pH 6.8, at 62° or 69°, respectively (11). At the end of incubation the reaction mixtures were cooled rapidly to 4°, made 0.24 M in phosphate buffer, pH 6.8, and incubated for 15 min at 37° with 20 μg/ml of ribonuclease A and 20 units/ml of ribonuclease T1. Ribonuclease-resistant RNA was precipitated onto filters with cold 10% trichloroacetic acid as described by Melli *et al.* (10). Second-order reaction curves have been fitted to the data by a computer program least-squares fit. Root mean square of curve describing DNA reassociation is 0.02, that describing the HnRNA or rRNA hybridization is 0.03. None of the data has been normalized. (*b*) (○) Uniformly labeled 28S [^{3}H]rRNA hybridization (specific activity = 170,000 cpm/μg); (●) pulse-labeled [^{3}H]rRNA hybridization (specific activity = 120,000 cpm/μg), prepared by pooling fractions 12–18 of a dimethylsulfoxide gradient similar to one shown in Fig. 3*a*. Unsheared RNA was heat-denatured and hybridized to excess DNA in solution (1500 nucleotides single-strand length) in 3.7× SSC–47% formamide at 37°. Mass ratio of DNA was 1000 to HnRNA and 100,000 to rRNA. At the end of incubation the reaction mixes were adjusted to 6× SSC by the addition of 50 volumes of 6× SSC, and the extent of hybridization was monitored by the filter binding method of Dina *et al.* (8). Five to 15% of the RNA bound to filters without hybridization (see *Methods*) and has been subtracted as background. Quenching due to the high DNA load per filter was monitored and radioactivity was adjusted accordingly (see *Methods*). Second-order reaction curves have been fitted to the data by a computer program least-squares fit. Root mean square of curve describing either rRNA or HnRNA hybridization is 0.07. HnRNA was reduced from 15 to 30 × 10^3 nucleotides to 10 to 15 × 10^3 nucleotides at the end of incubation to DNA C_0t 80, as judged by sedimentation in a dimethylsulfoxide gradient.

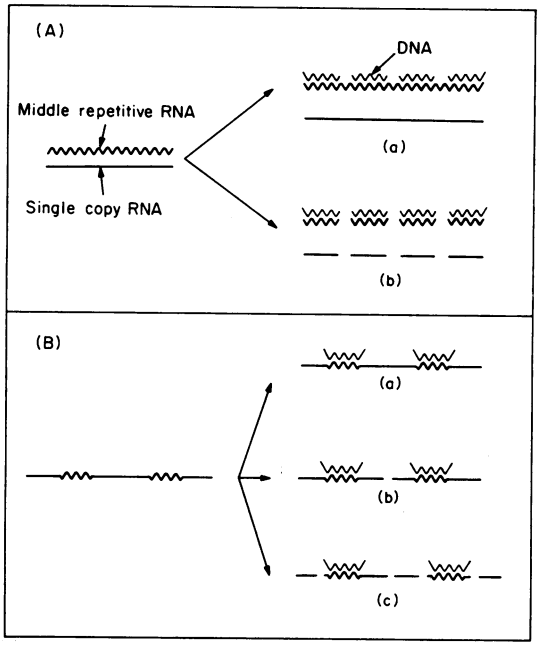

FIG. 2. Two major models for the detailed sequence organization of HnRNA are shown. The symbols used in (*B*) are the same as those in (*A*). See *text* for further details.

rate molecules, then the fraction of RNA recorded as containing at least one repetitive sequence, after hybridization to excess DNA at a middle repetitive DNA C_0t and without ribonuclease treatment, should be the same without shearing of the RNA (Fig. 2*A*,*a*) as it is with shearing (Fig. 2*A*,*b*). On the other hand, if they are interspersed on the same molecule then a moderate amount of shearing (Fig. 2*B*,*b*) should yield the same value of hybridization as no shearing (Fig. 2*B*,*a*), but further shearing to fragment sizes less than the distance between repetitive sequences should produce a concomitant reduction in the amount of RNA recorded as hybrid (Fig. 2*B*,*c*).

To distinguish between these two models the following experiment was carried out. Unsheared (Fig. 3*a*) and sheared (Fig. 3*b*) pulse-labeled [^{3}H]HnRNA were prepared by sedimentation through dimethylsulfoxide gradients. HnRNA was isolated from appropriate fractions of the gradients, and molecular weights were determined by reference to [^{14}C]rRNA included in the gradient. In some instances several fractions were pooled to provide sufficient HnRNA for subsequent hybridization experiments. Pooled fractions were assigned number average molecular weights. After hybridization to excess DNA, to DNA C_0t 80, hybrids were determined without ribonuclease treatment essentially by the method of Dina *et al.* (8). The results of this experiment are shown in Fig. 4*a*. Fig. 4*b* is an enlargement of part of Fig. 4*a* and includes the equivalent hybridization of sheared and unsheared, uniformly labeled, 28S rRNA to excess DNA. Clearly there is general interspersion of middle repetitive and single-copy sequences in at least the majority of HnRNA molecules, consistent with the model shown in Fig. 2*B*. For example, there is only a moderate reduction in hybridization with about a twofold decrease in size from 20 to 30 $\times$ 10^3 nucleotides to 10 to 15 $\times$ 10^3 nucleotides. Furthermore, 20–40% of HnRNA of length 1200 nucleotides contains at least one repetitive sequence. About 9–13% of HnRNA, sheared to 200–300 nucleotides, is capable of hybridization. This is consistent with the view that about 12% of HnRNA consists of middle repetitive se-

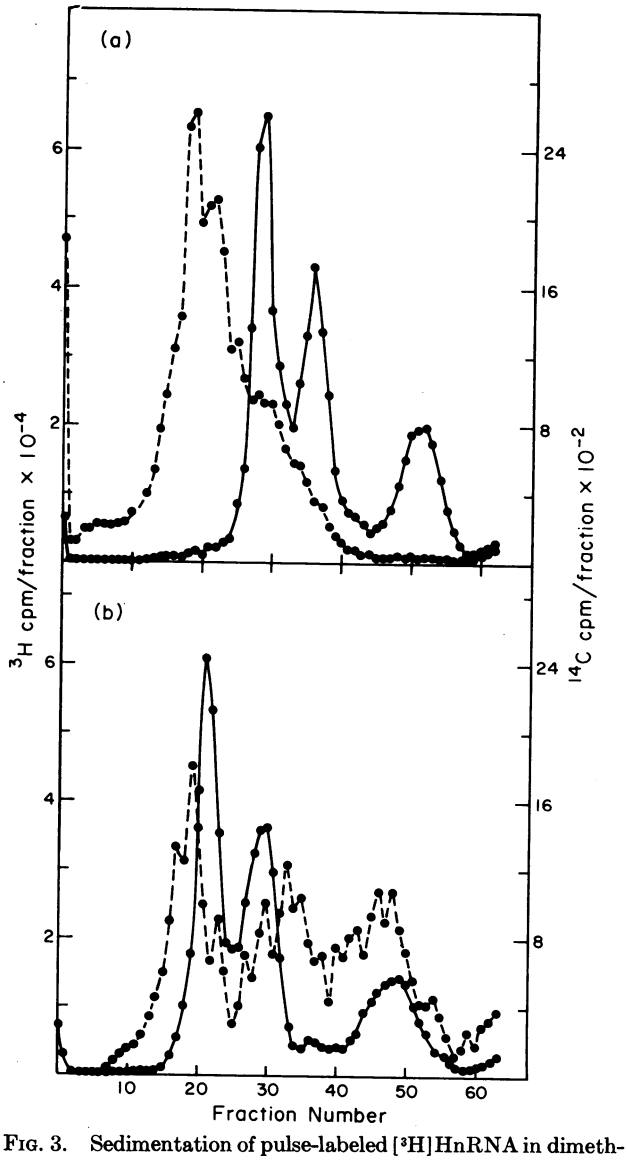

FIG. 3. Sedimentation of pulse-labeled [^{3}H]HnRNA in dimethylsulfoxide gradients (7). (*a*) (●——●) Unsheared [^{3}H]HnRNA; (●——●) BHK[^{14}C]rRNA [28 S = 1.9 $\times$ 10^6 and 18 S = 0.75 $\times$ 10^6 daltons (20)]. BHK rRNA was a gift from Lloyd Smith. Sedimentation (from right to left) was for 8 hr. (*b*) (●- - -●) Sheared [^{3}H]HnRNA prepared as follows: fractions 12–18 of several gradients such as the one shown in (*a*) were pooled, divided into two equal parts, and precipitated overnight at −18° with 2 volumes of 95% ethanol + 0.1 volume of 25% sodium acetate, pH 6.0. Precipitated RNA was pelleted by centrifugation at 30,000 rpm for 30 min at −15° in a Spinco 50.1 rotor. RNA from the two parts was separately resuspended in 3 ml of degassed 1 mM EDTA (pH 8.0) and sonicated as described (7). One part was sonicated for 5 sec and the other part for 2 min. After sonication, RNA was extracted with phenol–chloroform (7) and precipitated and suspended as above. The two parts were then combined and run on a dimethyl gradient (7) as shown here. (●——●) BHK rRNA. Sedimentation (from right to left) was for 10 hr. ^{3}H and ^{14}C cpm have been multiplied by 10^{-4} and 10^{-2}, respectively.

quences (Fig. 1*a*). Technical problems make it difficult to determine if a further reduction in size would produce a concomitant decrease in hybridization.

When sheared or unsheared 28S rRNA is reacted with excess DNA to C_0t 80, about 50–70% of the RNA hybridizes virtually independent of size (Fig. 4*b*). This value is that predicted from inspection of Fig. 1*b*. We are not certain whether the slight reduction in hybridization of 28S rRNA below about 1000 nucleotides is evidence that extensive shearing by sonication, *per se*, affects the ability of RNA to hybridize or whether the result is due to random scattering of the points. However, even if real, it is clear that shearing artifacts alone could not account for the sharp decrease in hybridization of HnRNA of lengths below 1000 nucleotides.

We have arbitrarily drawn a set of least-squares fitted curves through the HnRNA hybridization data in Fig. 4 based on the sollowing assumptions. (*1*) On a mass basis, 10% of HnRNA consists of middle repetitive sequences and 90% consists of single-copy sequences. (*2*) Twenty-eight percent of HnRNA is organized as middle repetitive sequences, 200 nucleotides long, interspersed with single-copy sequences of length 1200 nucleotides. This component is responsible for the lines labeled "1" in Fig. 4. (*3*) An additional 43% of HnRNA is organized so that middle repetitive sequences, of length 200 nucleotides, are interspersed with single-copy sequences 9800 nucleotides in length. This component is responsible for the lines labeled "2" in Fig. 4. No statement can be made about the remaining 29% of HnRNA. Since the HnRNA that was used to generate the sheared RNA fragments consists of a heterodisperse collection of molecules ranging from 15 to 30 $\times$ 10^3 nucleotides in length, and possibly consists of molecules with various sequence organizations at unknown relative concentrations, this particular model should be considered as only one of many possible organizations. This does not invalidate the principal point of this experiment, that middle repetitive and single-copy sequences are finely interspersed in HnRNA as demonstrated by the fact that 20–40% of RNA fragments 1200 nucleotides long contain at least one middle repetitive sequence.

It is unlikely that the data reflect the hybridization of the RNA molecules with various specific activities resulting from differential labeling *in vivo* because we have previously shown that the distributions of radioactivity and mass of pulse-labeled HnRNA, determined under denaturing conditions, are virtually superimposable (11). Hence, it is concluded that a considerable proportion of HnRNA is organized with interspersed middle repetitive and single-copy sequences, but that the exact numerical analysis of the organization requires further work.

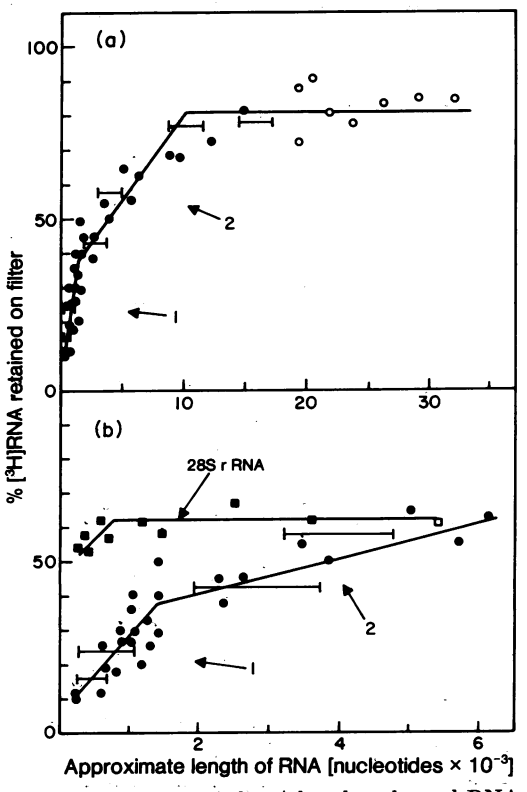

FIG. 4. Hybridization of sheared and unsheared RNA to excess DNA to a DNA C_0t of 80. (*a*) Hybridization of sheared (●) and of unsheared (○) pulse-labeled [³H]HnRNA to excess total rat DNA. (*b*) Hybridization of sheared (●) pulse-labeled [³H]-HnRNA (this is an enlargement of part of Fig. 4*a*) and of sheared (■) and unsheared (□) uniformly labeled rat ascites 28S [³H]-rRNA. The hybridization reactions and the determination of hybrids are the same as those described in the legend to Fig. 1*b*. Sheared and unsheared HnRNA were prepared from dimethylsulfoxide gradients as described in the legend to Fig. 3. Sheared and unsheared 28S rRNA were prepared in a similar fashion. Curves have been fitted to the data as described in the *text*. Molecular weights were estimated with reference to BHK rRNA (Fig. 3), assuming that sedimentation was a function of the logarithm of the molecular weight (11). Molecular weights have been converted to number of nucleotides, assuming 3.3 $\times$ 10^2 daltons/nucleotide. |——| indicates the hybridization of HnRNA from pooled fractions of dimethylsulfoxide gradients.

DISCUSSION

The middle repetitive sequences of rat DNA are, on average, about 100–300 base pairs long and are interspersed on a fine scale with single-copy sequences of average length 1000–2000 base pairs, although some single-copy sequences are up to 16,000 base pairs in length. Interspersed sequences account for about 80% of the rat genome. The remaining 20% includes highly repetitive sequences, such as satellites, and DNA containing single-copy or middle repetitive sequences that are not finely interspersed (5).

We have previously shown that rat ascites, rapidly labeled HnRNA of 5 to 10 $\times$ 10^6 daltons, isolated under denaturing conditions, contains both middle repetitive and single-copy sequences (11). Therefore, it is of interest to determine whether these sequences are interspersed in HnRNA, as in the majority of rat DNA, or whether HnRNA is transcribed from a minor fraction of the DNA with no interspersion. Since HnRNA from rat ascites hybridizes, in an RNA excess hybridization, to only about 6% of total rat DNA, the latter case was considered a real possibility. However, the experiments described in this paper strongly argue for intimate interspersion of middle repetitive and single-copy sequences, with at least 70% of HnRNA containing a minimum of one middle repetitive sequence. Even after HnRNA is sheared to about 1200 nucleotides, 20–40% contains at least one middle repetitive sequence. However, we cannot eliminate the possibility that a small fraction ($\leqslant$ 30%) of HnRNA consists of sequences not organized in this way.

The biological significance of the interspersion of sequences in primary RNA transcripts is unknown. Several possibilities might be considered, none of which is mutually exclusive. (*1*) Interspersed middle repetitive sequences might be required for the proper post-transcriptional processing and packaging of the RNA; for example, by conferring on the RNA the ability to assume particular secondary structures and/or providing appropriate recognition elements for specific nucleases. (*2*) They might represent the transcriptional product of repetitive DNA sequences involved in transcriptional regulation (13). (*3*) They might reflect the transcription of sequences involved in DNA replication or recombination. (*4*) They might be involved in translational control. (*5*) They might consist of similar but not identical sequences, and should be considered as parts of single-copy sequences that have regions of sufficient homology to form stable duplexes even under relatively stringent conditions of reassociation. (*6*) Some single-copy sequences might be spacers between repetitive genes, such as those coding for histone mRNA.

Most polysomal mRNAs, with known coding functions, consist predominantly of single-copy sequences (14–18). It is also known that the majority of the nucleotides of nuclear RNA never reach the cytoplasm. Hence, the question arises of whether an HnRNA molecule contains just one structural gene sequence with the bulk of the molecule having some other function or whether a fraction of HnRNA molecules consists of polycistronic structural genes with interspersed middle repetitive sequences while the remaining molecules have some other function. The possibility that none, or a very few, of the HnRNA molecules contains structural genes should also be considered; in this instance the majority of polysomal mRNAs could be generated from nuclear RNA molecules of lower molecular weight. While the present work cannot distinguish between these possibilities it does point to a high degree of order in the sequence organization of large primary RNA transcripts, suggesting that the sequence organization has an important biological function(s).

This work was supported by U. S. Public Health Service Grant GM 13762 and the Lucy Mason Clark Fund of the California Institute of Technology.

1. Britten, R. J. & Kohne, D. E. (1966) *Carnegie Inst. Washington Yearbk.* **65**, 78–106.
2. Britten, R. J. & Smith, J. (1970) *Carnegie Inst. Washington Yearbk.* **68**, 378–386.
3. Wu, J.-R., Hurn, J. & Bonner, J. (1972) *J. Mol. Biol.* **64**, 211–219.
4. Davidson, E. H., Hough, B. R., Amenson, C. S. & Britten, R. J. (1973) *J. Mol. Biol.* **77**, 1–23.
5. Bonner, J., Garrard, W., Gottesfeld, J., Holmes, D. S., Sevall, J. S. & Wilkes, M. M. (1973) *Cold Spring Harbor Symp. Quant. Biol.*, **38**, in press.
6. Davidson, E. H. & Britten, R. J. (1974) *Quart. Rev. Biol.*, in press.
7. Holmes, D. S. & Bonner, J. (1973) *Biochemistry* **12**, 2330–2338.
8. Dina, D., Crippa, M. & Beccari, E. (1973) *Nature New Biol.*, **242**, 101–105.
9. Gelderman, A. H., Rake, A. V. & Britten, R. J. (1971) *Proc. Nat. Acad. Sci. USA* **68**, 172–176.
10. Melli, M. C., Whitfield, K. V., Rao, V., Richardson, M. & Bishop, J. O. (1971) *Nature New Biol.* **231**, 8–12.
11. Holmes, D. S. & Bonner, J. (1974) *Biochemistry* **13**, 841–848.
12. Wetmur, J. & Davidson, N. (1968) *J. Mol. Biol.* **31**, 349–369.
13. Britten, R. J. & Davidson, E. H. (1969) *Science* **165**, 349–357.
14. Greenberg, J. R. & Perry, R. P. (1971) *J. Cell Biol.* **50**, 774–786.
15. Bishop, J. O., Pemberton, R. & Baglioni, C. (1972) *Nature New Biol.* **235**, 231–234.
16. Harrison, P. R., Hell, A., Birnie, G. D. & Paul, J. (1972) *Nature* **239**, 219–221.
17. Suzuki, Y., Gage, L. P. & Brown, D. D. (1972) *J. Mol. Biol.* **70**, 637–649.
18. Goldberg, R. B., Galau, G. A., Britten, R. J. & Davidson, E. H. (1973) *Proc. Nat. Acad. Sci. USA*, **70**, 3516–3520.
19. Straus, N. A. & Bonner, T. (1972) *Biochim. Biophys. Acta* **277**, 87–95.
20. McConkey, E. H. & Hopkins, J. W. (1969) *J. Mol. Biol.* **39**, 545–550.