Selected Fragment Analysis. A Method for Comparing Nucleotide Sequences in Ribonucleic Acid Molecules*

H. C. Birnboim

ABSTRACT: A method for comparing and contrasting the nucleotide sequence of two groups of ribonucleic acid molecules has been developed. To test the method, ^{3}H-labeled 28S ribosomal ribonucleic acid was compared with [^{14}C]18S RNA of HeLa cells. The ribosomal ribonucleic acids were mixed together and digested with pancreatic ribonuclease. A two-step procedure was adopted for examining the oligonucleotides formed. First, the digest was chromatographed at neutral pH, and hepta-, octa-, and nonanucleotides were separately precipitated. Then each was rechromatographed at low pH to separate some of the isomers. The [^{3}H]-oligonucleotides from 28S ribonucleic acid gave a pattern of oligonucleotide isomers that was different from the ^{14}C pattern (18S ribonucleic acid). This implied that significant differences in nucleotide sequence exist between these two ribosomal ribonucleic acid species.

The detailed study of cellular RNA frequently constitutes a formidable problem because of its great complexity. Whereas dramatic progress has been made in establishing the complete nucleotide sequence of some transfer and 5S RNA molecules (Holley *et al.*, 1965; Zachau *et al.*, 1966; Madison *et al.*, 1966; Raj Bhandary *et al.*, 1967; Brownlee *et al.*, 1967; Dube *et al.*, 1968), the prospect of establishing complete sequences of larger, more complex RNA molecules (such as ribosomal and DNA-like RNA) seems remote at the present. Because a need for distinguishing RNA molecules in complex mixtures is frequently encountered, alternative methods (less definitive than complete sequencing) are employed. Most frequently, measurements of sedimentation size and base composition are used, but it is clear that these alone may be inadequate for discriminating between different RNAs since molecules of the same sedimentation size and base composition may be very different in terms of their nucleotide sequence. Conversely, molecules of different sedimentation values may have sequences which are identical.

Other chemical methods have been utilized to supplement sedimentation and base composition measurements. Analyses of end groups (Lane and Tamaoki, 1967; Midgley and McIlreavy, 1966), terminal fragments (Midgley and McIlreavy, 1966; Lee and Gilham, 1966; Takanami, 1967; Sugiura and Takanami, 1967; Weith *et al.*, 1968), and 2′-*O*-methylribose dinucleotides (Lane and Tamaoki, 1967; Wagner *et al.*, 1967) have been carried out and differences between 18S (16S) and 28S (23S) rRNAs have been demonstrated. Another approach has been the analysis of oligonucleotides formed after digestion of the RNA with specific ribonucleases (Aronson, 1962, 1963; Sanger *et al.*, 1965; Amaldi and Attardi, 1968; Young, 1968). These authors, using methods which permit resolution of many of the isomers up to the tetranucleotide level, have demonstrated clear differences between the two rRNA species.

An alternative method for the analysis of oligonucleotides has been employed by Sinha *et al.* (1965), Robinson and Gilham (1967), and Thirion and Kaesberg (1968) for comparing RNA molecules. These authors used a two-stage analysis in which oligonucleotides were first separated according to chain length, then each peak was rechromatographed separately.

Complex RNA molecules have been compared using the technique of DNA–RNA hybridization competition but in the case of mammalian nucleic acids the interpretation of such results is complicated (*e.g.*, Birnboim *et al.*, 1967; Church and McCarthy, 1968). It seemed to us that a chemical method might prove to be a useful supplement to this kind of experiment. Our approach was to degrade the RNA molecules with a specific RNase and attempt to obtain a pattern of oligonucleotides which would be characteristic of the RNA molecules from which they were derived. The larger oligonucleotides were viewed as being potentially more informative, so these were selected and analyzed in a two-stage procedure similar to that used for the study of viral RNAs. To enable a direct comparison of the oligonucleotide patterns, a double-label technique was developed. In this way, the RNAs to be compared could be mixed, digested, chromatographed, and counted together. Our initial attempt to test the method has been to compare 18S with 28S rRNA of HeLa cells. In this report we show that significant differences exist.

Materials and Methods

Growth of Cells and Preparation of Labeled RNA. HeLa cells were grown in suspension culture as has been described by Eagle (1959). For the preparation

* From the Biology Branch, Biology and Health Physics Division, Atomic Energy of Canada Limited, Chalk River, Ontario, Canada. *Received August 30, 1968.*

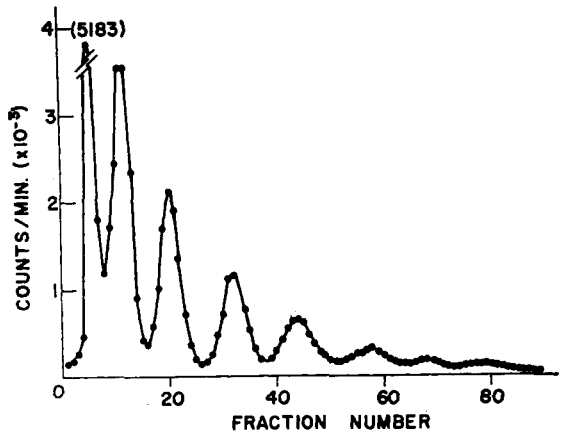

FIGURE 1: Chromatography at neutral pH of ribonuclease digest of rRNA. [^{3}H]28S RNA was mixed with [^{14}C]28S RNA and digested with pancreatic RNase as described in Results. The digest was applied to a DEAE-Sephadex column (0.9 × 20 cm) and eluted with a linear gradient of NaCl (0.18–0.32 M) in 8 M urea–0.01 M Tris-HCl (pH 7.5) at 0.5 ml/min. Shown on the ordinate is the sum of ^{3}H and ^{14}C counts.

of labeled RNA, two cultures of logarithmically growing cells at 1×10^5 cells/ml were used. To one culture [2,8-^{3}H]adenine (Nuclear-Chicago) was added to a concentration of 1.4 μCi/ml and to the second culture [8-^{14}C]adenine (Schwarz BioResearch, Inc.) was added to a final concentration of 0.1 μCi/ml. The final concentration of adenine was adjusted to 2×10^{-6} M with unlabeled base. The cultures were allowed to incorporate adenine for 40 hr. Total cell RNA was prepared, and the rRNA species was separated on sucrose gradients as has been previously described (Warner *et al.*, 1966). The 18S and 28S rRNA species were each separately precipitated by the addition of two volumes of ethanol at −20°. The precipitate was collected by centrifugation at 100,000*g* for 1 hr at 0°.

Digestion with Pancreatic Ribonuclease. Depending upon the particular experiment (see Results), [^{3}H]RNA was mixed together with [^{14}C]RNA in a ratio of three to five times the cpm. The solution was made 0.05 M in Tris-HCl (pH 7.5) and 0.01 M in EDTA. Pancreatic RNase (Worthington, previously heated to 90° for 10 min) was added at a concentration of 50 μg/mg of RNA. The mixture was incubated at 37° for 10 min; HCl was added to a final concentration of 0.1 N and the solution was further incubated for 1 hr at 37° (Sanger *et al.*, 1965). Following this step, the solution was neutralized with NaOH and diluted with suitable concentrated solutions to give a final concentration of 8 M urea and 0.02 M Tris-HCl (pH 7.5) in a volume of 10 ml.

Chromatography at Neutral pH on DEAE-Sephadex. DEAE-Sephadex A-25 (Pharmacia, Uppsala, Sweden) was prepared for use by cycling with acid and alkali as suggested by the manufacturer. The washed ion exchanger was equilibrated with 0.01 M Tris-HCl at pH 7.5 and 8 M urea. For chromatography, a column 0.9 × 20 cm was prepared and the RNase digest, prepared as described above, was applied to the column at a rate of 0.5 ml/min. Oligonucleotides were eluted from the column using a gradient of NaCl in 0.01 M Tris-HCl (pH 7.5) and 8 M urea. Further details are given in the legends to the figures.

Chromatography at Acid pH on DEAE-cellulose. Appropriate fractions (see legends to the figures) from the pH 7.5 chromatogram were pooled and the nucleotide material was precipitated with $Zn(OH)_2$ as described below. The precipitate was collected by low-speed centrifugation and the resulting pellet was dissolved in the minimum volume of 2 M ammonium formate (pH 3). This solution was diluted with water until the conductivity was less than 4500 μmhos. Rechromatography of isostichs, separated on the first column, was carried out on DEAE-cellulose (Whatman DE 32). Dimensions of the column were 0.9 × 20 cm. The exchanger was prepared by washing with acid and alkali and equilibrating it with 4 M urea–0.02 M ammonium formate (pH 3.3) (Whatman technical bulletin IE 2). The sample (20–30 ml) was applied to the column at a rate of 0.5 ml/min and the oligonucleotides were eluted with a lithium chloride gradient in 4 M urea–0.02 M ammonium formate (pH 3.3). Further details are given in the legends to the figures. The oligonucleotides were precipitated for counting as described below.

Precipitation of Oligonucleotides. Oligonucleotides were precipitated from dilute solutions at 90–100% efficiency without carrier using $Zn(OH)_2$ (Daneck *et al.*, 1965; Bock, 1967).[1] For the concentration of materials from the first column prior to application to a second column, the procedure was as follows. Zinc sulfate (0.1 ml of 0.5 M) was added per 10 ml of dilute nucleotide solution. The pH was adjusted to approximately 9 with NaOH. The precipitate which formed was allowed to age for approximately 1 hr and then collected by centrifugation at 2000*g* for 5 min. For concentrating the ^{3}H- and ^{14}C-labeled oligonucleotides recovered from the low pH column, a similar procedure was used. Zinc sulfate (0.5 ml of 0.05 M) was added to each of approximately 100 10-ml fractions using a syringe equipped with a valve and a no. 21 needle. The zinc sulfate solution was vigorously injected into the fraction tubes. This adequately mixed the sample and avoided the necessity of additional mixing. The pH of the solution was adjusted to approximately 8 by the addition of 0.8–1 ml of 0.2 N NaOH using a syringe and needle. This method of addition of the $ZnSO_4$ and NaOH gave a flocculent precipitate using a minimum of zinc. The precipitate was allowed to age for 30 min following which it was collected on glass fiber filter disks (Whatman GF/A, 2.4-cm diameter).

Counting Procedures for Radioactive Nucleotide Material. (a) For counting of ^{3}H and ^{14}C in solutions

[1] A variety of compounds have been found to affect the precipitation of Zn^{2+} as $Zn(OH)_2$, and the attendant precipitation of oligonucleotides. EDTA and high concentrations of formic acid inhibit; at low concentrations (0.01–0.1 M) NH_4^+ facilitates precipitation, whereas at higher concentration it inhibits; concentrated urea solutions inhibit slightly but as this seemed to vary with the preparation of urea, this may be due to the presence of a contaminant such as ammonia. The difficulty in the latter case was circumvented by adjusting the amount of $ZnSO_4$ added.

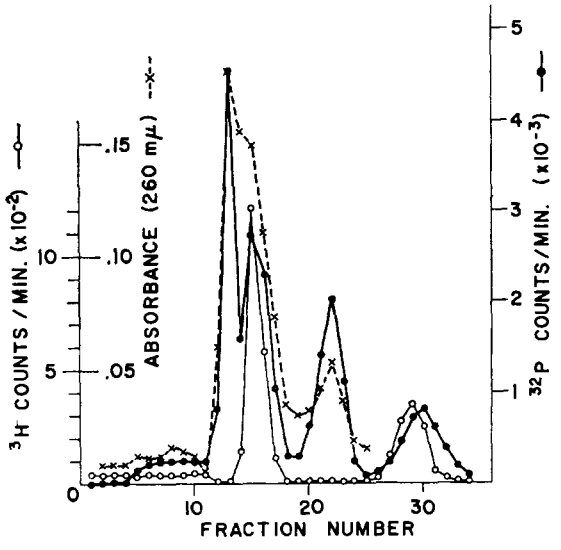

FIGURE 2: Chromatography at neutral pH of [^{32}P]- and [^{3}H]-oligonucleotides to establish chain length. [^{32}P] RNA was digested with pancreatic RNase, then ^{3}H-labeled nucleotide material from the first and third peaks of the chromatogram shown in Figure 1 was added. The mixture was chromatographed as described in the legend of Figure 1. Only the first 32 fractions are shown in this figure. ^{3}H and ^{32}P counts were determined by counting 0.3 ml of each fraction in 10 ml of dioxane-based scintillator. [^{3}H]Oligonucleotides derived from the first and third peaks of Figure 1 show correspondence with the second and fourth ^{32}P peaks.

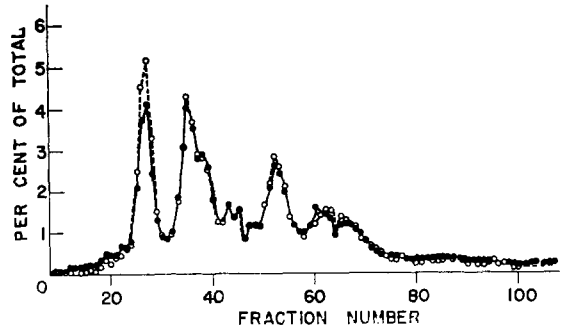

FIGURE 3: Rechromatography at acid pH of heptanucleotides. Fractions corresponding to the sixth peak (heptanucleotides) of Figure 1 were pooled and the concentrated oligonucleotides applied to a column of DEAE-cellulose (0.9 × 20 cm) equilibrated with 4 M urea–0.02 M ammonium formate (pH 3.3). The column was eluted with a 1-l. linear gradient of LiCl (0.12–0.20 M) in the same buffer; the nucleotide material in each 10-ml fraction was precipitated and counted as described in Methods. The results have been normalized in order to compare the ^{3}H and ^{14}C elution profiles. This was done by summing separately the ^{3}H and ^{14}C counts recovered from the column, then expressing the counts in each fraction as a percent of the total for each isotope. Total ^{3}H counts per minute recovered were 15,500 and ^{14}C counts per minute recovered were 3650. (●) [^{3}H]28S RNA; (○) [^{14}C]28S RNA.

containing urea, the following procedure was used. A solution of radioactive nucleotides in 8 M urea (up to 0.4 ml) was diluted with 0.8 ml of water, then 10 ml of the following scintillator was added: 6 g of Omnifluor (New England Nuclear Corp.) and 90 g of naphthalene to 1 l. of dioxane. Under these conditions the sample was homogeneous and ^{3}H was counted at 16–18% efficiency.[2] (b) For counting of ^{3}H and ^{14}C samples collected as the $Zn(OH)_2$ precipitate on glass fiber filters, the air-dried filters were immersed in 10 ml of scintillator prepared as follows: 6 g of Omnifluor and 155 g of naphthalene to 1 l. of dioxane. The efficiency of ^{3}H counting after an initial 1–2-hr period was 25%, and this value was constant for at least 48 hr. Unquenched samples counted ^{3}H at 32–33% efficiency. Solubilization of the zinc hydroxide gel with EDTA or NH_4OH led to a marked decrease in efficiency.

For counting of mixed ^{3}H and ^{14}C samples, a Nuclear-Chicago Mark I counter was used. The windows were adjusted to give 10% ^{14}C cross-contribution in the ^{3}H channel (13% ^{3}H efficiency) and 0.5% ^{3}H cross-contribution in the ^{14}C channel. Uniform quenching was assured by careful addition of the zinc sulfate, and was checked by using an external standard. Further evidence for the uniformity of counting conditions is provided by the control experiments of Figures 3–5. Since the profiles for the two labels were expected to be identical, the differences between them provide a measure of the maximum error due to nonuniform sample preparation. (c) For those experiments in which ^{32}P was used, aqueous samples were counted in a liquid scintillation counter without added scintillator (Clausen, 1968). Efficiency under these conditions was approximately 40% in a Nuclear-Chicago Mark I counter.[3] One advantage of this method is that samples could readily be recovered after counting when necessary for further analysis. The use of Cerenkov counting for a similar purpose has recently been described (Matthews, 1968).

Determination of Oligonucleotide Chain Length. Mono- and oligonucleotide peaks separated by chromatography at neutral pH were assigned values for chain length on the basis of the ratio of inorganic to total phosphate after phosphatase treatment. [^{32}P]RNA was used to prepare the oligonucleotides, and the conditions of phosphatase treatment and separation of P_i were according to Furlong (1965).

Base Composition Analysis. Conditions used for alkaline hydrolysis were 0.1 N KOH at 100° for 20 min (Bock, 1967). Up to 100 μl of sample (not neutralized) was applied to Whatman No. 3MM paper as 2-cm lines. Nucleotides were separated by electrophoresis in 0.05 M ammonium formate (pH 3.6) at 50 V/cm for 1.75 hr.

[2] It has been found that if the scintillator described in b was used under similar conditions, a gross precipitate was formed. Despite the turbid appearance, [^{3}H]H_2O was counted at 23% efficiency, whereas [^{3}H]RNA was counted at 8–9% efficiency. The difference is presumed to be due to precipitation of the RNA.

[3] In a Packard Tri-Carb scintillation counter the corresponding efficiency was 16%. Tritium in dioxane scintillator was counted at nearly equal efficiencies in the two counters. The difference was presumed to be due to differences in the sensitivity of the respective phototubes to the wavelength of the emitted light.

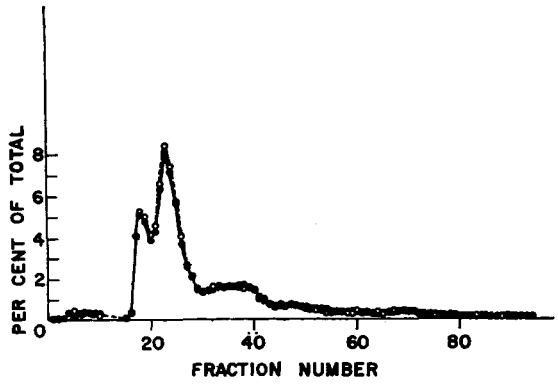

FIGURE 4: Rechromatography at acid pH of octanucleotides. The seventh peak (octanucleotides) from the experiment of Figure 1 was rechromatographed as described in the legend to Figure 3. The linear LiCl gradient used extended from 0.16 to 0.21 M. Fractions 11–14 were accidentally lost. Total ^{3}H counts per minute recovered was 21,600 and total ^{14}C counts per minute was 5100. (●) [^{3}H]28S RNA; (○) [^{14}C]28S RNA.

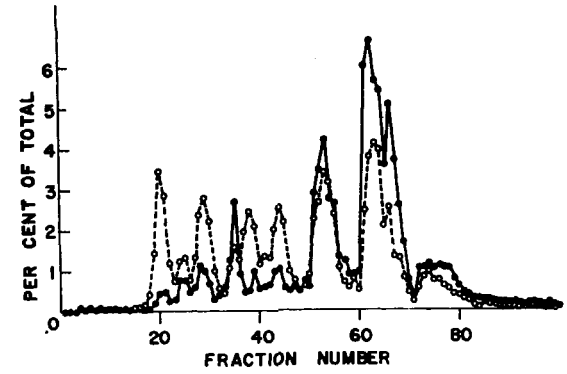

FIGURE 6: Rechromatography at acid pH of heptanucleotides. [^{3}H]28S RNA and [^{14}C]18S RNA were mixed together and digested as described in Results. The digest was chromatographed at neutral pH under conditions described in Figure 1, with similar results. The sixth peak (heptanucleotides) was rechromatographed under conditions described in Figure 3. The linear LiCl gradient used extended from 0.12 to 0.20 M. Total ^{3}H counts per minute recovered was 11,300 and the total ^{14}C counts per minute was 3300. (●) [^{3}H]28S RNA; (○) [^{14}C]28S RNA.

Results

Labeled 18S and 28S rRNA was prepared as described in Methods. Under the conditions used for their isolation, *i.e.*, a single sedimentation, it was estimated that each rRNA species was cross-contaminated with 5–10% of other species. Of the label incorporated into RNA, 90% was in adenine and 10% in guanine.

A control experiment to test the labeling, isolation, and counting procedures was performed. Approximately 60 μg of [^{3}H]28S RNA (1×10^{6} cpm) was mixed with an equal weight of [^{14}C]28S RNA (3×10^{5} cpm) and the mixture was incubated with pancreatic ribonuclease as described in Methods. The oligonucleotides formed were separated according to chain length as described in Methods and the elution profile is shown in Figure 1.

Chain length was established by comparing on a column some of the material from the first and third peaks of the chromatogram shown in Figure 1 with a pancreatic RNase digest of [^{32}P]RNA. The mixture was chromatographed at neutral pH and the results

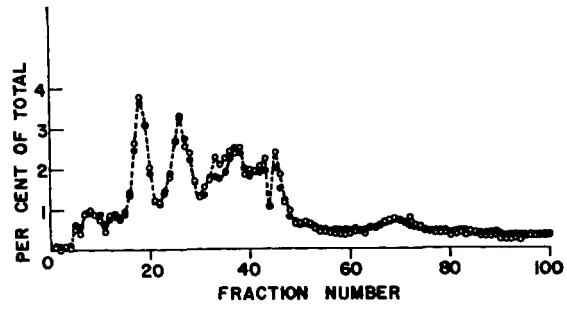

FIGURE 5: Rechromatography at acid pH of nonanucleotides. The eighth peak (nonanucleotides) from the experiment of Figure 1 was rechromatographed as described in the legend to Figure 3. The LiCl gradient used for elution extended from 0.18 to 0.23 M. Total ^{3}H counts per minute recovered from the column was 16,250 and total ^{14}C counts per minute recovered was 3700. (●) [^{3}H]28S RNA; (○) [^{14}C]28S RNA.

are shown in Figure 2. The first ^{3}H peak corresponds to the second ^{32}P peak (dinucleotide). This was expected since the ^{3}H label was in a purine and only pyrimidine mononucleotides are released by pancreatic ribonuclease. The first and second ^{32}P peaks correspond to mono- and dinucleotides, respectively. This was shown by the correspondence of the first absorbance peak with the first ^{32}P peak (which ruled out P_i) and by the finding that these peaks gave close to the expected values for the ratio of terminal to total phosphate (a measure of oligonucleotide chain length) (see Methods). Subsequent peaks were assumed to be trinucleotides, tetranucleotides, etc.

Fractions corresponding to the hepta-, octa-, and nonanucleotides were each precipitated separately and rechromatographed at pH 3.3 as described in Methods. The elution profiles are shown in Figures 3, 4, and 5, respectively. The counts for each isotope have been normalized by expressing each fraction as a per cent of total counts of that isotope recovered from the column. These diagrams show that there is very close correspondence between the ^{3}H and ^{14}C counts. Since these oligonucleotides are derived from the same kind of RNA, *i.e.*, 28S RNA, these results established the adequacy of our labeling, isolation, and counting techniques.

For purposes of comparing the oligonucleotide pattern of two different kinds of RNA, the following experiment was performed. [^{3}H]28S RNA was mixed with [^{14}C]18S RNA and the mixture was digested with ribonuclease as described in Methods. The elution pattern was very similar to that shown in Figure 1. The hepta- and octanucleotides were recovered and rechromatographed at low pH. The resultant elution profiles are shown in Figures 6 and 7. By contrast with the results of the control experiment, oligonucleotides derived from *different* kinds of RNA gave rise to significantly different isomers. This was also true of rechromatographed heptanucleotides derived from a

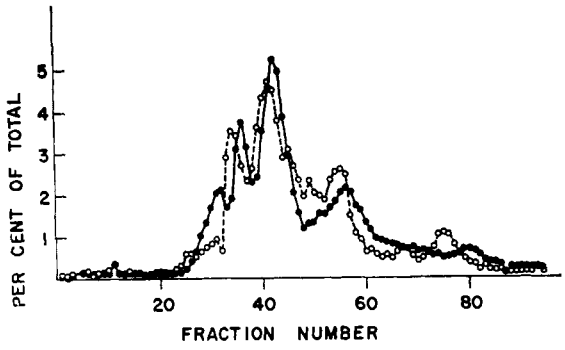

FIGURE 7: Rechromatography at acid pH of octanucleotides. The octanucleotide peak from the experiment described in Figure 6 was rechromatographed as in Figure 3. The LiCl gradient used extended from 0.14 to 0.21 M. Total ^3H counts per minute recovered was 16,400 and total ^{14}C counts per minute was 4270. (●) [^3H]18S RNA; (○) [^{14}C]28S RNA.

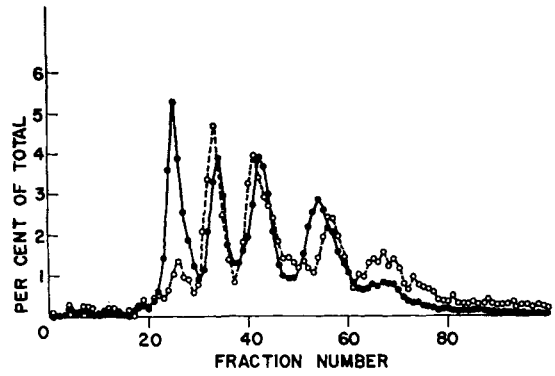

FIGURE 8: Rechromatography at acid pH of heptanucleotides. [^3H]18S RNA was mixed with [^{14}C]28S, digested, and chromatographed as described in the legend to Figure 1. The sixth peak (heptanucleotides) was rechromatographed as in Figure 3. The LiCl gradient used extended from 0.13 to 0.20 M. Total ^3H counts per minute recovered was 9200 and total ^{14}C counts per minute was 1537. (●) [^3H]18S RNA; (○) [^{14}C]28S RNA.

separate experiment in which the labels were reversed, *i.e.*, [^{14}C]28S RNA and [^3H]18S RNA (Figure 8).

One experiment was performed in which [^{32}P]28S RNA was used. This RNA was digested with ribonuclease and chromatographed at neutral pH as described above. The pentanucleotide peak was rechromatographed at low pH and various peaks were analysed for their nucleotide composition. The data showed that elution at pH 3.3 was in the expected order, that is, oligonucleotides rich in guanine and uracil tended to be eluted after oligonucleotides rich in adenine and cytosine. This is in accord with the extensive analyses of Thirion and Kaesberg (1968) of oligonucleotides chromatographed at low pH.

Discussion

In this report we have described a method which is intended to supplement existing procedures for comparing and contrasting mixtures of RNA molecules. In accord with a large body of chemical and DNA–RNA hybridization data, the method has permitted the demonstration of differences between 18S and 28S rRNA species. The determination of these differences was based upon a comparison of the pattern of selected oligonucleotides derived from each of the RNAs. Since the oligonucleotides examined probably originated from positions throughout the length of the molecules, the patterns are thought to reflect significant differences in the nucleotide sequences of the RNAs.

The use of a two-step procedure for the analysis of oligonucleotides is based upon consideration of different kinds and amounts of oligonucleotides which would be expected to occur in pancreatic or T1 RNase digests (Table I). These were calculated assuming the RNA is composed of the four common nucleotides present in equimolar amounts and randomly distributed. It may be seen that the probability that a fragment of a specified composition or sequence will *not* occur in a RNA molecule solely by chance increases markedly with chain length. Thus, larger oligonucleotides are more likely to be distinctive and may be used with greater certainty to identify the RNA molecules from which they were derived. A two-step procedure was therefore adopted in order to separate the more informative large fragments from the less informative but more plentiful smaller fragments.

We propose that the method may be widely applicable to the study of other large, complex RNA molecules. A similar approach was sensitive enough to detect small differences between viral RNA molecules (Sinha *et al.*, 1965; Robinson and Gilham, 1967; Thirion and Kaesberg, 1968). Even more complex molecules may be amenable to this approach by choosing appropriately large oligonucleotides for analysis.

The data presented have been interpreted as demonstrating *differences* between molecules, but the method is limited with respect to the information that it gives about *similarities* between molecules. Such information would require greatly improved techniques for resolving sequence isomers, or might be approached by introducing a three-step analysis. A further limitation is the ability to make only qualitative, not quantitative, statements about similarities or differences. For example, the present data rule out the possibility that 28S RNA is a dimer of 18S RNA, or that it is a dimer with a segment rich in guanine and cytosine. However, the molecules could share common sequences over 30–40% of their length. Such information is frequently adequate for many experiments.

Finally, the principle of comparing mixtures of macromolecules by prior selection of appropriate fractions for a second-step analysis may also prove useful in the study of mixtures of protein or DNA molecules.

Acknowledgment

The skillful technical assistance of John Martin is gratefully acknowledged. I thank D. K. Myers for many helpful suggestions during preparation of this manuscript and A. Teich for assistance in working out the theoretical distribution of oligonucleotide fragments.

TABLE I: Theoretical Distribution of Oligonucleotides in a Ribonuclease Digest.[a]

Chain Length (L)	Yield (Y) (%)	No. of Compositional Isomers (I_c)[b]	No. of Sequence Isomers (I_s)[c]	Q_c	Q_s
			a. Pancreatic Ribonuclease Digest		
1	25	2	2	$<10^{-10}$	$<10^{-10}$
2	25	4	4	10^{-10}	10^{-10}
3	18.8	6	8	2.83×10^{-5}	3.92×10^{-4}
4	12.5	8	16	0.0198	0.141
5	7.80	10	32	0.210	0.614
6	4.68	12	64	0.522	0.885
7	2.73	14	128	0.757	0.970
8	1.56	16	256	0.885	0.992
9	0.873	18	512	0.948	0.998
10	0.480	20	1024	0.976	0.9995
			b. Ribonuclease T1 Digest		
1	6.25	1	1	$<10^{-10}$	$<10^{-10}$
2	9.38	3	3	1.44×10^{-7}	1.44×10^{-7}
3	10.56	6	9	2.78×10^{-3}	0.0199
4	10.56	10	27	0.071	0.376
5	9.90	15	81	0.267	0.783
6	8.88	21	243	0.494	0.941
7	7.77	28	729	0.673	0.985
8	6.40	36	2,187	0.801	0.996
9	5.67	45	6,561	0.869	0.9990
10	4.70	55	19,683	0.918	0.9998

[a] The following assumptions have been made: the RNA molecules are large and are composed of the four common bases in equimolar amounts, randomly distributed; the digestion is carried to completion; pancreatic RNase is specific for pyrimidines and RNase T1 is specific for guanine. [b] The number of possible compositional isomers (*i.e.*, without regard to sequence) was calculated from an empirical relationship. [c] For a pancreatic RNase digest, the number of possible sequence isomers is 2^L. For RNase T1, the number is $3^{(L-1)}$. Q_c = approximate probability that an oligonucleotide of a specified composition will *not* occur solely by chance in a molecule 1000 nucleotides long.

$$Q_c = \left[1 - \frac{Y \cdot 10^{-2} \cdot 10^3}{I_c \cdot L \cdot 10^3}\right]^{1000}$$

$$Q_s = \left[1 - \frac{Y \cdot 10^{-2} \cdot 10^3}{I_s \cdot L \cdot 10^3}\right]^{1000}$$

References

Amaldi, F., and Attardi, G. (1968), *J. Mol. Biol. 32*, 737.

Aronson, A. I. (1962), *J. Mol. Biol. 5*, 453.

Aronson, A. I. (1963), *Biochim. Biophys. Acta 72*, 176.

Birnboim, H. C., Pene, J. J., and Darnell, J. E. (1967), *Proc. Natl. Acad. Sci. U. S. 58*, 320.

Bock, R. M. (1967), *Methods Enzymol. 12*, 224.

Brownlee, G. G., Sanger, F., and Barrell, B. G. (1967), *Nature 215*, 735.

Church, R. B., and McCarthy, B. J. (1968), *Biochem. Genetics 2*, 55.

Clausen, T. (1968), *Anal. Biochem. 22*, 70.

Daneck, K., Weimann, G., and Cramer, F. (1965)' *Experientia 21*, 417.

Dube, S. K., Marker, K. A., Clark, B. F. C., and Cory, S. (1968), *Nature 218*, 232.

Eagle, H. (1959), *Science 130*, 432.

Furlong, N. B. (1965), *Anal. Biochem. 12*, 349.

Holley, R. W., Apgar, J., Everett, G. A., Madison, J. T., Marquisee, M., Merrill, S. H., Penswick, J. R., and Zamir, A. (1965), *Science 147*, 1462.

Lane, B. G., and Tamaoki, T. (1967), *J. Mol. Biol. 27*, 335.

Lee, J. C., and Gilham, P. T. (1966), *J. Am. Chem. Soc. 88*, 5685.

Madison, J. T., Everett, G. A., and Kung, H. (1966), *Science 153*, 531.

Matthews, H. R. (1968), *J. Chromatog. 36*, 302.
Midgley, J. E. M., and McIlreavy, D. J. (1966), *Biochem. J. 101*, 32P.
RajBhandary, U. L., Chang, S. H., Stuart, A., Faulkner, R. D., Hoskinson, R. M., and Khorana, H. G. (1967), *Proc. Natl. Acad. Sci. U. S. 57*, 751.
Robinson, W. E., and Gilham, P. T. (1967), *7th Intern. Congr. Biochem.*, p 660.
Sanger, F., Brownlee, G. G., and Barrell, B. G. (1965), *J. Mol. Biol. 13*, 373.
Sinha, N. K., Enger, M. D., and Kaesberg, P. (1965), *J. Mol. Biol. 12*, 299.
Sugiura, M., and Takanami, M. (1967), *Proc. Natl. Acad. Sci. U. S. 58*, 1595.
Takanami, M. (1967), *J. Mol. Biol. 29*, 323.
Thirion, J.-P., and Kaesberg, P. (1968), *J. Mol. Biol. 33*, 379.
Wagner, E. K., Penman, S., and Ingram, V. M. (1967), *J. Mol. Biol. 29*, 371.
Warner, J. R., Soiero, R., Birnboim, H. C., Girard, M., and Darnell, J. E. (1966), *J. Mol. Biol. 19*, 349.
Weith, H. L., Asteriadis, G. T., and Gilham, P. T. (1968), *Science 160*, 1459.
Young, R. J. (1968), *Biochemistry 7*, 2263.
Zachau, H. G., Dutting, D., and Feldmann, H. (1966), *Z. Physiol. Chem. 347*, 212.

Incorporation into Polypeptide and Charging on Transfer Ribonucleic Acid of the Amino Acid Analog 5′,5′,5′-Trifluoroleucine by Leucine Auxotrophs of *Escherichia coli**

Eric D. Fenster† and H. S. Anker

ABSTRACT: Some leucine auxotrophs of *Escherichia coli* were previously shown to be adaptable to grow on 5′,5′,5′-trifluoroleucine in a chemostat. In these experiments bacteria *naive* to trifluoroleucine were tested for their ability to carry out various reactions when confronted with this analog in the absence of leucine. A crude extract of aminoacyl synthetases charged about two to three times as much trifluoroleucine as leucine to transfer ribonucleic acid. The trifluoroleucine charging was inhibited, in decreasing order of effectiveness, by leucine, isoleucine, phenylalanine, and cysteine.

Intact cells incorporated trifluoroleucine into polypeptide for about an hour, whereupon incorporation ceased. The initial rate was somewhat greater than the initial rate of leucine incorporation. *In vivo* incorporation of trifluoroleucine was inhibited by leucine, isoleucine, and phenylalanine. When histidine was used as an indicator of protein synthesis *in vivo* the ratio of trifluoroleucine to histidine incorporation was about twice the leucine to histidine ratio. The trifluoroleucine to histidine ratio was reduced in the presence of isoleucine or phenylalanine; the leucine to histidine ratio was unaffected. A combination of isoleucine and phenylalanine decreased the trifluoroleucine to histidine ratio below the leucine to histidine ratio. Ribonucleic acid synthesis, as measured by uracil incorporation, was markedly reduced by incubation in trifluoroleucine. Synthesis was fully restored by 1 μg/ml of chloramphenicol, an amount too low to relax ribonucleic acid synthesis in the absence of amino acids.

The adaptation of some leucine auxotrophs of *Escherichia coli* to growth in 5′,5′,5′-trifluoroleucine was reported by Rennert and Anker (1963). The organisms were grown in a chemostat in which the leucine of the medium was gradually replaced by trifluoroleucine. Amino acid analysis of the total protein of adapted bacteria showed that trifluoroleucine was present in nearly the same relative concentration as leucine was in unadapted bacteria. No leucine was found in the fully adapted bacteria. In the experiments described below some responses of unadapted bacteria to trifluoroleucine were studied. Trifluoroleucine was found to be recognized as leucine, isoleucine, and phenylalanine both in the tRNA-charging reaction *in vitro* and in polypeptide synthesis *in vivo*.

* From the Department of Biochemistry, University of Chicago, Chicago, Illinois 60637. *Received June 28, 1968.* Supported by a grant from the American Cancer Society. E. D. F. was supported by a U. S. Public Health Service Training Grant to the department.

† Taken from a thesis, submitted by E. D. F., in partial fulfillment of the Ph.D. requirements in the Department of Biochemistry.

Present address: Space Sciences Laboratory, University of California, Berkeley, Calif. 94720.