Volume 4 Number 8 August 1977 Nucleic Acids Research

Mapping adenines, guanines, and pyrimidines in RNA

Helen Donis-Keller, Allan M. Maxam and Walter Gilbert

Department of Biochemistry and Molecular Biology, Harvard University, Cambridge, MA 02138, USA

Received 3 June 1977

ABSTRACT

The positions of adenines, guanines, and pyrimidines can be determined by partial nuclease digestion of a terminally labeled RNA molecule. In urea, at elevated temperatures, RNase T_1 generates a pattern reflecting cleavage at guanines while RNase U_2 cleaves only at adenine. A limited alkaline hydrolysis provides a continuum of fragments derived from breaks at every phosphodiester bond. The reaction products are electrophoretically fractionated by size in adjacent lanes of a polyacrylamide gel. An autoradiograph of the gel displays the sequence up to 100 nucleotides from the end of the molecule, although uracil cannot as yet be distinguished from cytosine. These techniques form the basis of an RNA sequencing method and are demonstrated on yeast 5.8S ribosomal RNA.

INTRODUCTION

We present a simple RNA sequencing technique that orders and locates all adenines and guanines within 100 nucleotides of one end of an RNA chain. This procedure also maps pyrimidines with respect to the purines, though it does not distinguish uracil from cytosine. The principle of the method is as follows: terminal labeling with ^{32}P first establishes a unique reference point at one end of the RNA; then partial digestion with ribonuclease T_1 or U_2 generates nested sets of labeled fragments, each extending from this reference point to an internal guanine or adenine. Polyacrylamide gel electrophoresis assorts the partial digestion products by size, and autoradiography finally displays those which are end-labeled. Limited alkaline hydrolysis yields another set of nested fragments, extending from the labeled

end to random bases in the RNA. When processed together, the two digests and this hydrolysate produce three parallel band patterns on the autoradiograph, one from cleavage at every guanine, another from cleavage at every adenine, and the third from cleavage at every base. Since the products of these cleavages are ordered by size, this autoradiograph alone determines the position of each adenine, guanine, and pyrimidine in the RNA sequence. We demonstrate this technique on an RNA of known sequence, yeast 5.8S ribosomal RNA.

MATERIALS

Bacterial alkaline phosphatase (Worthington) and ribonucleases T_1 and U_2 (Sankyo) were used without additional purification. Polynucleotide kinase was purified from phage T4 infected *E. coli* cells as described by Panet, *et al.* (1). Carrier tRNA (Swartz-Mann *E. coli* B tRNA) was further deproteinized by extraction with phenol. Ultra pure urea was obtained from R-plus Laboratories, Inc. and acrylamide from Bio-Rad. Reaction vessels were 1.5 ml Eppendorf conical polypropylene tubes with snap caps, treated with 5% (vol/vol) dimethyldichlorosilane in CCl_4 and rinsed with distilled water. Drawn-out capillaries were made from melting point tubes 1.5 mm x 150 mm.

METHODS

Isolation of Substrate RNA

5.8S ribosomal RNA from yeast (*Saccharomyces cerevisiae*) was isolated according to the method of Rubin (2). The RNA was fractionated on 9.7% acrylamide, 0.3% bis-acrylamide, 7 M urea, 50 mM Tris-borate, pH 8.3, 1 mM EDTA, preparative slab gels (0.6 cm x 20 cm x 20 cm); it was heated at 50°C for 3 min in 7 M urea, 5 mM Tris-borate, 1 mM EDTA, pH 8.3, 0.01% xylene cyanol, 0.01% bromophenol blue just prior to loading on the gel. Electrophoresis was carried out at room temperature approximately 11 hr at 200

volts; the reservoir buffer was 50 mM Tris-borate, pH 8.3, 1 mM EDTA. The 5.8S band was visualized by UV shadowing (3), cut out, and eluted from the gel (4).

[γ-^{32}P]ATP Exchange Synthesis

Specific activity routinely obtained is 1500 Ci/mmol and has been described elsewhere (4).

Dephosphorylation

In a siliconized 1.5 ml Eppendorf tube, incubate RNA and an appropriate amount of alkaline phosphatase in 100 µl 100 mM Tris-HCl, pH 8.0, for 30 min at 37°C. Add 100 µl TE buffer (10 mM Tris-HCl, pH 7.4, 1 mM EDTA) and extract three times with 200 µl TE-saturated phenol and once with 500 µl ether. Then precipitate the RNA by adding 25 µl 3 M NaAc, pH 6.0, 0.1 M $MgAc_2$, 1 mM EDTA and 750 µl ethanol. Mix the solution by inverting the tube 4-5 times, then chill at -70°C (dry ice-ethanol bath) 5 min, sediment the RNA at 12,000 x g (Eppendorf microcentrifuge) 5 min, and remove the supernatant with a Pasteur pipette. Redissolve the pellet in 250 µl 0.3 M NaAc, add 750 µl ethanol, mix, chill, sediment the RNA, and remove the supernatant as described above. Rinse the pellet by adding 1 ml ethanol, chill, centrifuge, remove the supernatant as above, and remove residual ethanol under vacuum.

Labeling 5' Ends

Dissolve dephosphorylated RNA in 70 µl 10 mM Tris-HCl, pH 7.4, 1 mM spermidine, 0.1 mM EDTA. Heat at 50°C for 3 min, then chill in ice water. Add 10 µl 10X kinase buffer (500 mM Tris-HCl, pH 9.0, 100 mM $MgCl_2$, 50 mM dithiothreitol), 10 µl [γ-^{32}P]ATP in water (a minimum of 100 picomoles or an amount equivalent to the picomoles of RNA 5' ends present), and several units of polynucleotide kinase. The final reaction volume is 100 µl. Incubate at 37°C for 30 min, add 100 µl 4M NH_4Ac, mix, then add 2 µl carrier tRNA (10 mg/ml), 600 µl ethanol and mix well by inverting the

tube 4-5 times. Chill at -70°C (dry ice-ethanol bath) for 5 min, centrifuge at 12,000 x g for 5 min (Eppendorf microcentrifuge), remove the supernatant, rinse the pellet with ethanol, dry under vacuum, redissolve in loading buffer, and electrophorese.

Partial Digestion with Ribonucleases T_1 and U_2

Appropriate nuclease/RNA ratios ($T_1 = 10^{-4}$, $U_2 = 10^{-2}$ units RNase/µg RNA) for partial digestion are established by serial dilution of enzyme into a buffer containing 20 mM Na-citrate, pH 5.0, 7 M urea, 1 mM EDTA, 0.25 mg/ml tRNA, 0.025% xylene cyanol and bromophenol blue. Prepare 150 µl of this buffer just prior to use, and then add [5'-^{32}P] or [3'-^{32}P] RNA dissolved in a small volume of water (1-10 µl). Distribute the reaction mixture in six 20 µl portions into siliconized plastic Eppendorf tubes. Heat in a water bath at 50°C for 5 min, then quick chill the tubes on ice. Add 1 µl ribonuclease T_1 (Sankyo), 0.1 unit/µl, to the first of 3 tubes. Then make two successive tenfold serial dilutions using a fresh micropet each time and mixing. RNase U_2 (Sankyo) dilutions are made by adding 1 µl, 1 unit/µl enzyme, to one tube and diluting this tenfold into a second tube. One 20 µl portion is incubated without enzyme. Incubate all six tubes in a water bath for 15 min at 50°C. Chill, transfer into drawn-out capillaries, and load onto the gel.

Limited Alkaline Hydrolysis

Add 1-5 µl [5'-^{32}P] or [3'-^{32}P] RNA to 20 µl 50 mM $NaHCO_3/Na_2CO_3$, pH 9.0, 1 mM EDTA, 0.25 mg/ml tRNA at 0°C. Transfer the solution into a drawn-out capillary and seal both ends with a flame. Submerge the capillary in a 90°C water bath for 15 min. Rinse the contents of the tube into 20 µl 10 M urea, 0.05% xylene cyanol, 0.05% bromophenol blue and load onto the gel.

Gel Electrophoresis

The sequencing reaction mixtures are fractionated in adjacent lanes of a polyacrylamide gel (slab dimensions 0.15 cm x 33 cm x 40 cm, sample well dimensions 1 cm x 1 cm x 0.15 cm) containing 20% acrylamide, 0.67% methylene bis-acrylamide, 7 M urea, 50 mM Tris-borate, pH 8.3, 1 mM EDTA, 0.07% ammonium persulfate. Pre-electrophorese the gel at 800 volts for at least 2 hrs prior to loading samples, and then electrophorese at 800-1000 volts, regulating voltage or wattage (power). The gel should be slightly warm to keep the larger cleavage products denatured, and constant power will maintain this heat throughout the electrophoretic run.

RNA fragments ∿25 nucleotides long run with the xylene cyanol marker dye and fragments ∿8 nucleotides long run with the bromophenol blue. Sequences more than 25 nucleotides distant from the terminal ^{32}P label can be derived by running the dye markers far down or off the gel. By loading a portion of each set of degradation products whenever the xylene cyanol has migrated half way down the gel, all regions of the sequence out to about 100 nucleotides can be expanded. For sequences less than 10 nucleotides distant from the end label, run the cleavage products on a tighter polyacrylamide gel (25% acrylamide, 0.83% bis-acrylamide) to insure fractionation by size and not charge.

After electrophoresis, remove one glass plate from the polyacrylamide slab, cover the gel with Saran Wrap, and mark the positions of the electrophoresis dyes on the wrap with ^{14}C ink. Autoradiograph at -20°C for exposure times longer than 5 hours to prevent diffusion, or at room temperature for shorter exposures.

RESULTS

To develop methods to elicit partial, base-specific cleavage products we needed substantial amounts of homogeneous, end-labeled RNA of known sequence. Yeast 5.8S ribosomal RNA is monophosphorylated at the 5' end and con-

tains 158 nucleotides, sequenced by Rubin (5). We purified 600 µg by phenolizing intact yeast cells, precipitating the RNA with ethanol, and electrophoresing on a preparative polyacrylamide gel. After removing the 5' phosphate with alkaline phosphatase, we labeled 10-15 µg portions of this RNA with [γ-^{32}P] and isolated the product from a 10% polyacrylamide, 7M urea slab gel (Figure 1). This provided about 50 µCi of ^{32}P-end-labeled RNA, enough for 100-150 trial digests with ribonucleases.

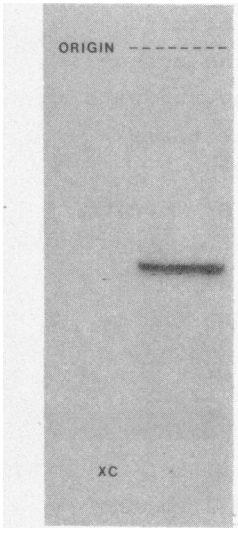

Figure 1. Autoradiograph of [5'-^{32}P] end-labeled yeast 5.8S ribosomal RNA after electrophoresis on a 10% polyacrylamide, 7 M urea slab gel. 15 µg of 5.8S rRNA was end-labeled as described under MATERIALS AND METHODS.

In a limit digest at neutral pH, RNase T_1 cleaves all GpX phosphodiester bonds in single-stranded RNA, leaving the phosphates on the guanosines (6). In 5 or 7 M urea the enzyme not only retains its activity and base specificity, but exhibits greater apparent activity at a new pH optimum, pH 4.5 (7). Hoping to attack every GpX bond within 100 nucleotides of the labeled 5' end of the RNA, including those in regions which could have secondary structure, we digested with RNase T_1 in 7 M urea at pH 5.0 and 50°C. Three tubes containing ^{32}P-end-labeled 5.8S rRNA, 5 µg carrier RNA, pH 5.0 buffer, and urea, but no enzyme, were

heated at 50°C and chilled. Then 0.1 unit RNase T_1 was added to the first, and serially diluted by factors of 10 down through the remaining two. All three reaction mixtures were then heated at 50°C for 15 min, layered on a 20% polyacrylamide, 7 M urea slab gel (4), electrophoresed, and the gel autoradiographed. This produced the patterns shown in Figure 2.

On this gel ^{32}P-labeled fragments differing in chain length by only one nucleotide are resolved (4) and appear as bands on the film. The weakest bands derive from fragments which extend from the labeled end to more-or-less random breaks in the RNA used for this experiment (see lane A), while the intense ones vary with the level of RNase T_1. The RNase/RNA ratio decreases by factors of ten in lanes B through E on the Figure 2 autoradiogram, and in the middle lane C a distinct pattern emerges. By comparing in order all weak and intense bands on the autoradiograph of Figure 2 with nucleotides in the 5.8S sequence in Figure 3, it can be seen that the strong partial products found in the pattern are just those predicted by the specificity of the enzyme (these are numbered on the film and under the sequence in the two figures). Hence RNase T_1 cleaves the RNA faithfully and almost exclusively at guanines in a partial digest.

RNase U_2 cleaves both GpX and ApX phosphodiester bonds in an extensive digest of RNA, although the ApA linkage is said to be somewhat resistant (8). However, in a partial digest achieved by dilution of the enzyme into the same buffer used for RNase T_1 and reaction at 50°C for 15 min, as above, we find that U_2 cleaves only ApX, and it does so quite uniformly. Figure 4 shows RNase U_2 and T_1 partial digests of the end-labeled 5.8S RNA. The products of partial cleavage at guanines, produced by T_1, are absent from the U_2 lanes, and correlation of bands on the autoradiograph with the nucleotide sequence shows that RNase U_2 cleaves after every adenine.

Alkali, 0.3 N KOH for 16 hours at 37°C, will hydrolyze all RNA phosphodiester bonds between unmodified sugars (9).

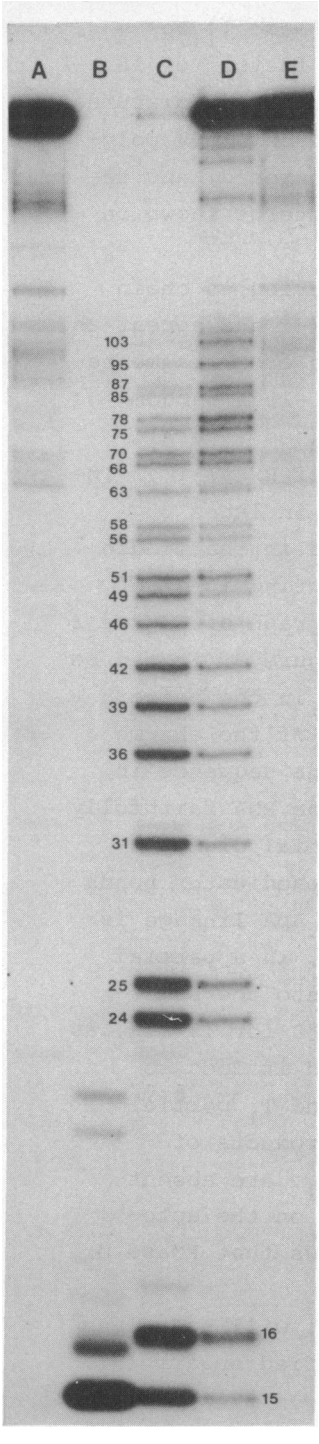

Figure 2. Autoradiograph of partial RNase T_1 digest products from 5' end-labeled yeast 5.8S rRNA, separated by size on a polyacrylamide slab gel. Five 20 µl samples, each containing [5'-^{32}P] 5.8S RNA, 5 µg unlabeled tRNA, 20 mM Na-citrate, pH 5.0, 1 mM EDTA, 7 M urea, 0.02% xylene cyanol, and 0.02% bromophenol blue, were heated at 50°C for 5 min, and chilled. Then 0.1 unit RNase T_1 was introduced into sample B and serially diluted through C, D, and E by factors of 10. All samples, including one to which no enzyme was added, were then heated at 50°C for 15 min, chilled, and layered on a 20% polyacrylamide, 7 M urea gel, and electrophoresed for 20 hr at 900 volts, and the gel exposed to X-ray film for 8 hr at -20°C.

A No enzyme

B 10^{-1} unit RNase T_1 / 5 µg RNA

C 10^{-2} unit RNase T_1 / 5 µg RNA

D 10^{-3} unit RNase T_1 / 5 µg RNA

E 10^{-4} unit RNase T_1 / 5 µg RNA

Compare the numbered bands with identically numbered partial sequences in Figure 3.

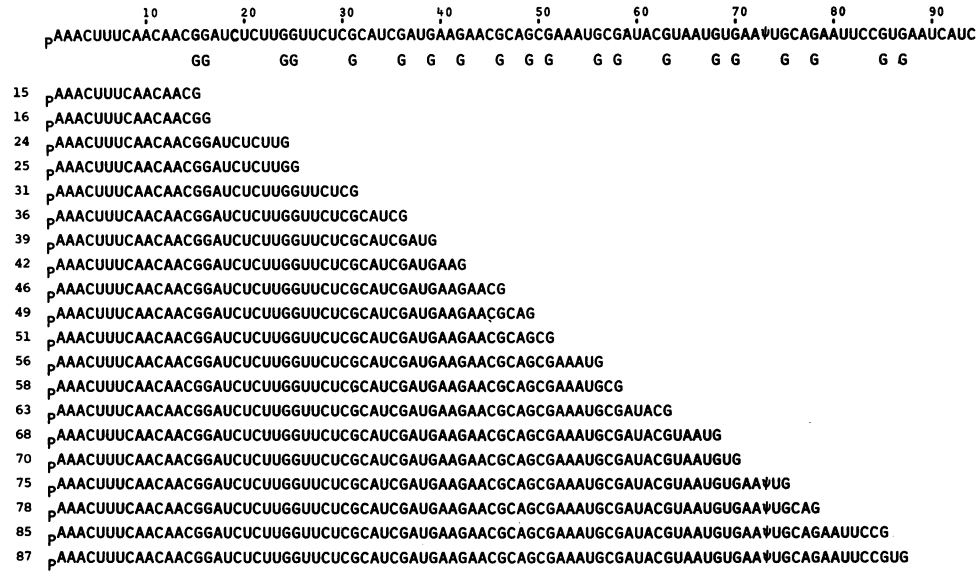

Figure 3. Nucleotide sequence at the 5' end of Saccharomyces cerevisiae 5.8S ribosomal RNA as determined by Rubin (5), with a set of RNase T_1 partial digestion products extending from the end to internal guanines underneath. Each of the latter is designated by the position of its 3'-terminal guanylic acid residue in the 5.8S RNA sequence; these numbered segments of the RNA should be compared with identically numbered bands on the autoradiograph in Figure 2.

Although a limited KOH reaction can hydrolyze about 1 in 100 bonds, which this sequencing technique requires, we find that a reaction mixture buffered at pH 9.0 with sodium bicarbonate does so more reproducibly. When 5.8S RNA is heated at 90°C for 15 minutes in 50 m$\underline{M}$ $NaHCO_3/Na_2CO_3$, pH 9.0, it is randomly hydrolyzed, to a limited extent, at every phosphodiester bond. The end-labeled products of this limited hydrolysis fill the interpurine spaces when fractionated next to RNase T_1 and U_2 partial digests, indicating the positions of pyrimidines in the sequence (Figure 4). We note that the stronger bands in the Figure 4 alkaline hydrolysis pattern derive from a certain specificity of the reaction, wherein XpG phosphodiester bonds break more readily than others.

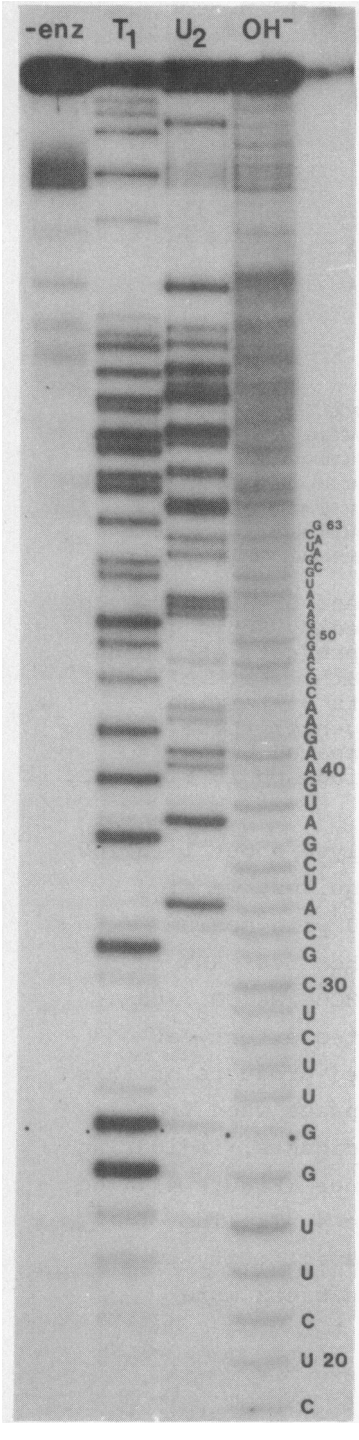

Figure 4. Autoradiograph of [5'-^{32}P] labeled yeast 5.8S RNA partially cleaved at every guanine (T_1), adenine (U_2), and each base (OH-).

DISCUSSION

These partial hydrolyses with RNase T_1, RNase U_2, and alkaline buffer provide half the information necessary for a complete sequence. Nonetheless the ability to determine the positions of the adenines, guanines, and pyrimidines in extended regions of a sequence has proved useful. The purine distribution at the 5' end of a messenger RNA or an *in vitro* transcript provides enough information to locate the initiation point for that transcript in a known DNA sequence, and this has been done with two messenger RNAs synthesized *in vitro* from phage T_7 (D. McConnell, personal communication, and U. Siebenlist, personal communication). Furthermore, this method can be used to identify and to order the T_1 fragments in conventional RNA sequencing, and to order adenine residues within the T_1 fragments.

As yet, pancreatic RNase does not yield patterns that determine the pyrimidines. In our hands it is too sequence-specific to generate useful partials. We are currently seeking other enzymatic methods which will distinguish uracils from cytosines.

ACKNOWLEDGMENTS

This work was supported by National Institute of General Medical Sciences Grant GM09541. W.G. is an American Cancer Society Professor of Molecular Biology.

REFERENCES

1 Panet, A., van de Sande, J.H., Loewen, P.C., Khorana, H.G., Raae, A.J., Lillehaug, J.L., and Kleppe, K. (1973) Biochem. 12, 5045-5050
2 Rubin, G. (1975) in Methods in Cell Biology, Vol. 12, pp.52-53. Academic Press, New York
3 Hassur, S.M. & Whitlock, H.W., Jr. (1974) Anal. Biochem. 59, 162-164
4 Maxam, A.M. & Gilbert, W. (1977) Proc. Natl. Acad. Sci. (USA) 560-564
5 Rubin, G. (1973) J. Bio. Chem. 248, 3860-3875
6 Uchida, T. & Egami, F. (1966) in Procedures in Nucleic Acid Research, Cantoni, G.L. and Davies, D.R., Eds., pp.3-13. Harper and Row, New York

7 Warrington, R.C. (1974) Biochim. Biophys. Acta. 353, 63-68
8 Brownlee, G.G. (1972) Determination of Sequences in RNA, pp. 199-200. North-Holland, Amsterdam
9 Bock, R.M. (1967) in Methods in Enzymology, Grossman, L. and Moldave, K., Eds., Vol. 12A, pp.218-221. Academic Press, New York