Volume 9 Number 21 1981 Nucleic Acids Research

Primary and secondary structure of U2 snRNA

Ramachandra Reddy, Dale Henning, Paul Epstein and Harris Busch

Department of Pharmacology, Baylor College of Medicine, Houston, TX 77030, USA

Received 18 May 1981

ABSTRACT

With the improved rapid sequencing techniques, the earlier sequence of U2 RNA of Novikoff hepatoma (Shibata et al, J. Biol. Chem. 250, 3909-3920, 1975) was reanalyzed and modified. The improved sequence of U2 RNA is 188 (or 189) nucleotides long and is in register with a characterized U2 RNA pseudogene (Denison et al, PNAS 78, 810-814, 1981) except for an 11 nucleotide sequence (nucleotides 147-157) which is absent from the pseudogene. From these results, a secondary structure of U2 RNA is proposed which is supported by the preferred cleavage sites with T_1-RNase, RNase A and S1 nuclease. Isolated U2 RNA was cleaved by T_1-RNase preferentially at positions 64 and 164, whereas U2 RNA in U2-snRNP was cleaved only at position 64, indicating that position 164 is protected in U2-snRNP. As with U1 RNA (Epstein et al, PNAS 78, 1562-1566, 1981) the 5'-end of isolated U2 RNA was not preferentially cleaved by T_1-RNase.

INTRODUCTION

U-snRNAs (1) in snRNPs have been implicated in processing of hnRNAs (2,3). The sequences of the snRNAs in these RNPs appear to be important to their function. The 5'-end sequence of U1 RNA was shown to be complementary to several intron/exon splice junctions (2,3). The conformation of U1 RNA in solution (4,5) and in U1-snRNP (4) were analyzed earlier. The results showed that the conformation of U1 RNA was similar in solution and in RNP form (4) and also that 5'-end 12 nucleotides of U1 RNA may be accessible to binding splice junctions of hnRNA (4,5). In the present study, the secondary structure of U2 RNA was analyzed by enzymatic digestions.

In addition, the sequence of U2 RNA was reevaluated with ladder sequencing methods. The sequences of U1 RNA (6) and U2 RNA (7,8) had been determined with uniformly labeled RNA.

The original sequence of U1 RNA (6) was corrected using improved sequencing techniques (9-11). The sequences of U3 (12, 13), U4 (14,15,16), U5 (1,17,18) and U6 (19,20) RNA were determined with these improved sequencing methods. The modified sequence of U2 RNA was found to be in register with a U2 RNA pseudogene isolated from human placenta (21), except the sequence of nucleotides 147-157 was missing in the pseudogene. With these modifications of the U2 RNA sequence, the primary sequences of all known capped U-snRNAs have been completed with the ladder sequencing methods. The sequences now provide a more satisfactory basis for elucidating the structure and function of the U-snRNPs.

MATERIALS AND METHODS

Preparation of ^{32}P-labeled U2 RNA from Novikoff hepatoma ascites cells. The harvested Novikoff hepatoma ascites cells were incubated overnight in culture medium in the presence of $[^{32}P]$ orthophosphate as described previously (22). Preparation of citric acid nuclei, isolation of RNA, and fractionation of 4 to 8S RNA were done as reported earlier (6,22,23). Nuclear 4 to 8S RNA was subjected to electrophoresis on 10% acrylamide 7 M urea gels at pH 8.3 (12,24). U2 RNA separated well from other RNAs.

RNases T_1, T_2, U2, Phy M and RNase A were obtained as noted earlier (12). [γ-^{32}P] ATP and 5' [^{32}P] pCp were used to label termini (24,25).

The U2 RNA was labeled on its 3'-end using 5'[^{32}P] pCp and RNA ligase (25) and the digestions were carried out by the chemical method (29). Identification of Ψp residues from Up residues was carried out by chromatography on 3 MM paper using isopropanol: HCl:water (680:176:144) system (30).

Preparation of NP-40 nuclei, digestion of RNPs with T_1-RNase, precipitation of RNPs using anti-Sm antibodies and analysis of RNA fragments was done as described earlier (4).

RESULTS

Figure 1 shows the T_1-RNase fingerprint of uniformly labeled U2 RNA. The sequences, derived from analysis of the

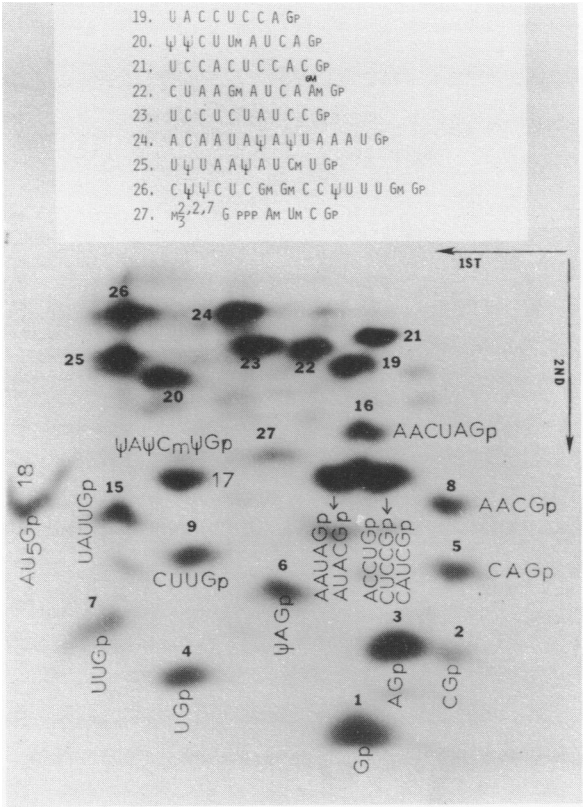

Fig. 1 Autoradiograph of a two-dimensional fractionation of a T_1-RNase digest of [^{32}P]-labeled U2 RNA of Novikoff hepatoma. The first dimension was on a cellulose acetate strip at pH 3.5 and the second dimension was homochromatography on thin layers of polyethyleneimine (Cell 300, Brinkmann).

products obtained after T_1-RNase oligonucleotides were digested with U2 RNase or RNase A, were essentially the same as reported earlier (8). The 5' or 3'-end labeled large oligonucleotides were partially digested; the resulting fragments were analyzed by mobility shift analysis (Fig. 2). From these results and from the digestion products of uniformly labeled RNA the sequences of oligonucleotides 19 to 26 were determined.

For example, the sequence of oligonucleotide T-26, a hexadecanucleotide was deduced from the following data. 25 µg of

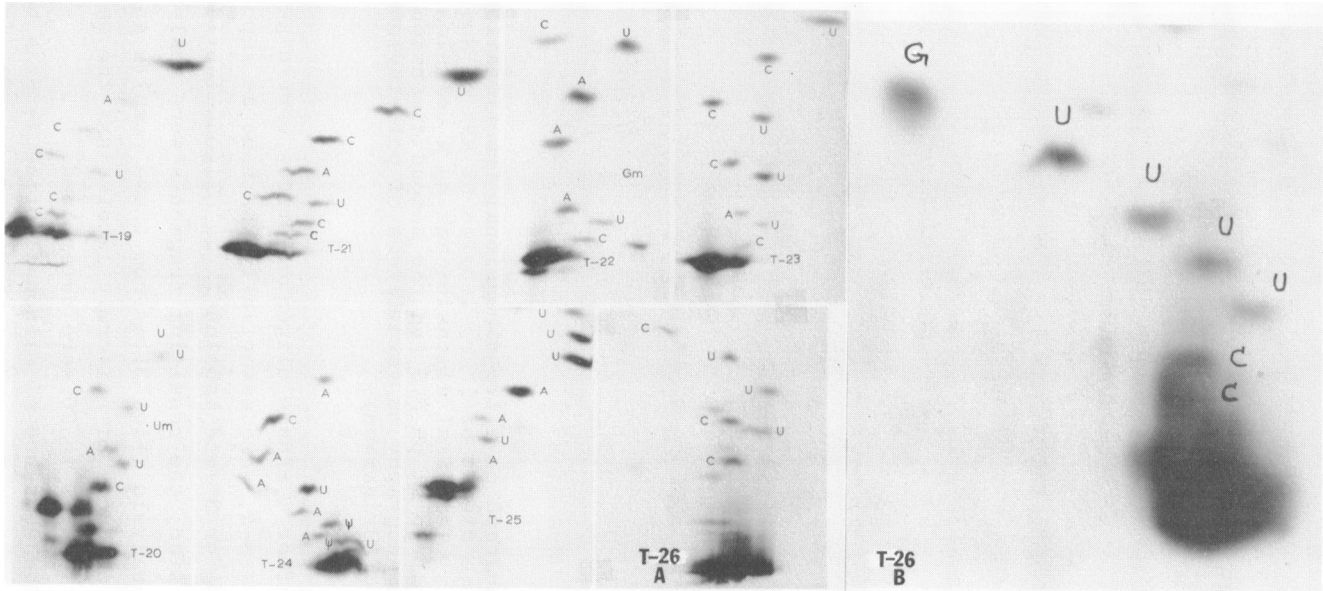

Fig. 2 Autoradiograph of mobility shift analyses of large T_1-RNase fragments. Complete T_1-RNase fragments were end-labeled and partially digested with formamide at 100°C for 30 min. and separated on a two-dimensional fractionation system. T-26A was labeled on the 5'- end and T-26B was labeled on the 3'-end. For other details, see Fig. 1.

unlabeled U2 RNA mixed with 50,000 cpm of uniformly labeled U2 RNA was subjected to fingerprinting after complete T_1-RNase digestion (8). Oligonucleotide T-26 was eluted from DEAE-paper and was labeled at its 5'-end (24). The labeled fragment was subjected to mobility shift analysis after partial digestion (28). Oligonucleotide T-26 was also labeled on its 3'-end (after removal of 3'-phosphate) and subjected to mobility shift analysis. The sequence from the 5'-end was deduced to be C-U-U-C-U-C- and the sequence from 3'-end was deduced to be -C-C-U-U-U-U-GmGp (Fig. 2). These digests were also run on a sequencing gel and the positions corresponding to Gm-Gm were found to be cleaved only to a small extent. The identity of Gm-Gm-Cp and GmGp was established in this study and was consistent with previous results. Similar data (Fig. 2) was obtained for all large T_1-RNase fragments. The differences (Fig. 1) from the sequences determined earlier were primarily in the order of pyrimidine residues.

Partial digestion fragments. Uniformly labeled U2 RNA was mixed with 100 µg of unlabeled U2 RNA and subjected to partial digestion with T_1-RNase or S1 nuclease. The resulting fragments were separated on an acrylamide gel (Fig. 3A) and aliquots of each fragment were fingerprinted. The fingerprints of partial T_1-RNase digestion fragments 1-64, 1-164 and 65-164 are shown in Fig. 4.

Under the partial T_1-RNase digestion conditions used over 95% of the U2 RNA was recovered in fragments larger than 25 nucleotides. The main cleavages occurred at nucleotide positions 164, 64, 37 and 107. This result shows that the most susceptible regions to cleavage by T_1-RNase of U2 RNA to be nucleotides 164>64>37>107, nucleotide 164 and 64 being the most preferred sites.

Sequencing gels. Partial digestion fragments obtained by T_1-RNase or S1 nuclease were end-labeled and subjected to sequencing gel analysis. Fig. 3B shows a sequencing gel corresponding to nucleotides 96-164 of U2 RNA. The U2 RNA was also labeled at its 3'-end and the sequence of U2 RNA was analyzed using ladder sequencing methods (25). Using the chemical digestion (29), the positions of A and G residues were confirmed from nucleotide

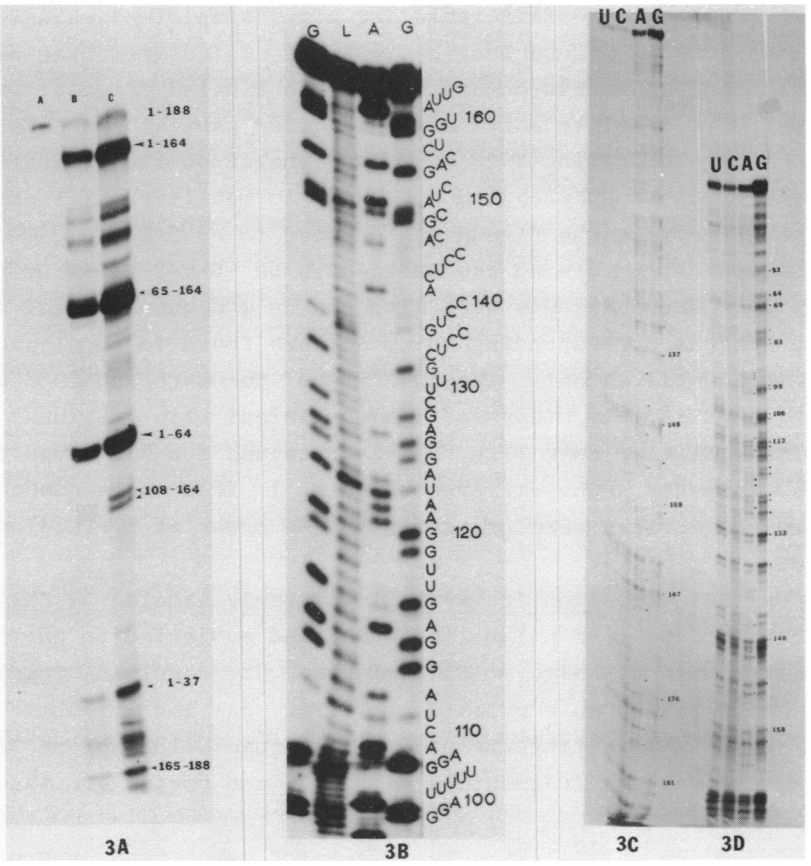

Fig. 3 A. Autoradiograph of a U2 RNA digest separated on a 12% acrylamide gel. Lane A is undigested U2 RNA and lanes B and C, U2 RNA digested with T_1-RNase (Enzyme to substrate ratio 1:1000) in 0.01 M Tris, 1 mM $MgCl_2$, pH 7.5 for 5 min. at 0°C. The reaction was stopped by SDS-phenol extraction.

B. Autoradiograph of a sequencing gel of 5'-end labeled fragment (U2 RNA 65-164). The enzyme digestions and running of a 8% sequencing gel were as described by Donis-Keller et al (24).

C. Autoradiograph of a sequencing gel of 3'-end labeled U2 RNA. The chemical method of Peattie (29) was employed for digestion of RNA.

D. Same as in Fig. 3C, except that the electrophoresis run for longer time to visualize longer oligonucleotides.

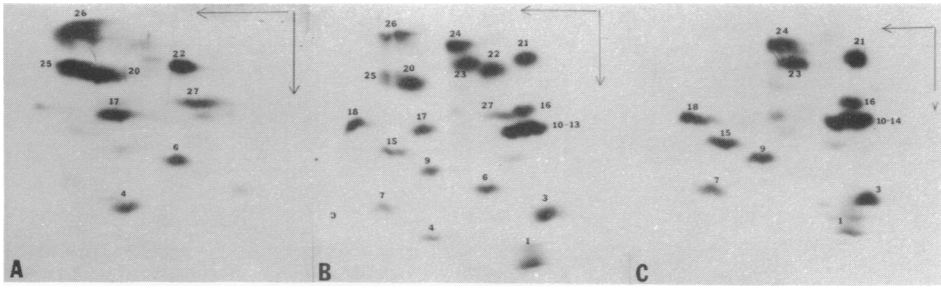

Fig. 4 Autoradiographs of fingerprints of partial T_1-RNase fragments. The fragments obtained by partial T_1-RNase digestion of U2 RNA (as described in Fig. 3A) were fingerprinted after complete T_1-RNase digestion. Fragments of U2 RNA 1-64, A; 1-164, B; and 65-164, C.

53-188 (Fig. 3C and 3D). Fingerprinting of repurified U2 RNA did not show C-Gp residue (Fig. 5) that was found earlier (7,8) and the sequencing gels did not show gel bands corresponding to C-Gp between oligonucleotide T-15 and T-5 (Fig. 3C). Therefore C-Gp was not used in this U2 RNA. Data obtained in the present study and earlier studies (7,8) provided the 188 or 189 nucleotide sequence of U2 RNA shown in Fig. 6.

Modified nucleotides. U2 RNA contains many modified nucleotides (31). All these modified nucleotides were shown to be present in the 5'-third of U2 RNA (8). The present study confirmed this interesting observation. In oligonucleotides T-6 and T-17, ψ residues were the only U residues (Fig. 7). When RNase A and U2 RNase products obtained from T-24 were analyzed, ψp was found in Aψ and (ψU) Ap. Therefore, position 92 was designated ψp. Another Up in T-24 was found to be a partial conversion to ψp at position 90. T-25 contained two ψp, showing that position 59 is ψp; the other ψp was found in U2 RNase fragment (Uψ U)Ap. When 5'-end labeled T-25 was digested with nuclease P1 and the labeled oligonucleotide analyzed, it was found to be pU, localizing the ψp to either position 55 or to 56. The ψp at position 55 was based on earlier results (7,8). Oligonucleotide T-20 contained two ψp residues and both were found in U2 RNase fragment ψψCUUA obtained from T-20. Analysis of the

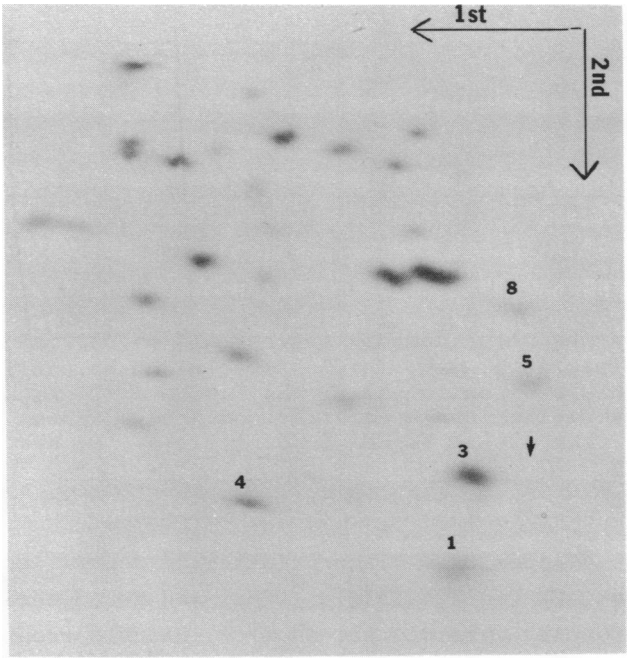

Fig. 5 Autoradiograph of a two-dimensional fractionation of a T_1-RNase digest of [^{32}P]-labeled repurified U2 RNA. U2 RNA obtained as described in "Materials and Methods" was repurified on a 12% polyacrylamide gel. The arrow points to the position of C-Gp spot, not found in repurified U2 RNA. Other details as in Fig. 1.

5'-end labeled T-20 showed the nucleotide at position 44 to be ψp. The ψp at position 45 was based on earlier results (7,8). There were three ψp residues at positions 7,8 and 16 based on reexamination of earlier data (7,8) to fit into the sequence derived from end labeled RNA fragments. Therefore, U2 RNA contains 12-13 ψ residues and the ψp residues at positions 35, 38, 40, 42, 44, 59, 92 (90 partially) were confirmed in this study. The ψp residues at 7, 8, 16, 45, 55 may need confirmation by other methods (9).

Secondary structure of U2 RNA. A possible secondary structure for the U2 RNA molecule is shown in Fig. 8 . This structure was developed with the aid of the computer program of Korn

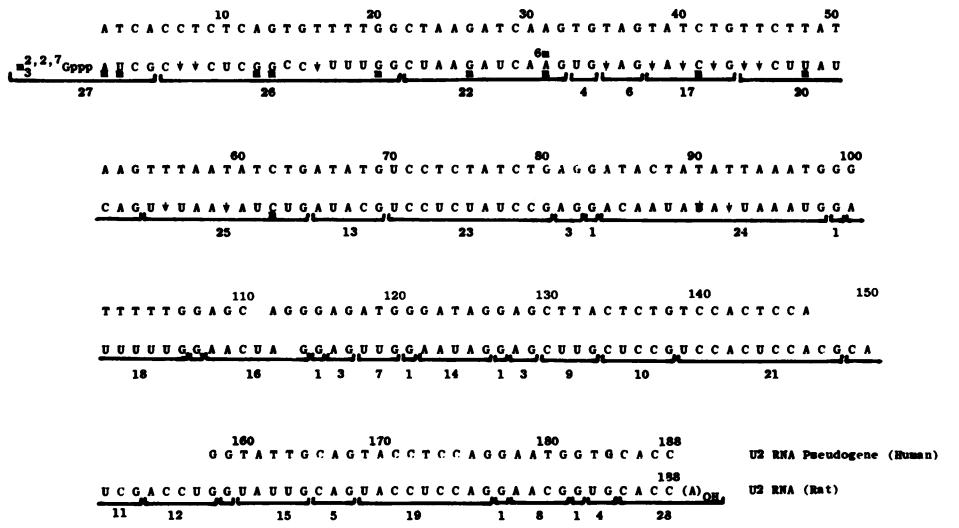

Fig. 6 Comparison of U2 RNA sequence of Novikoff hepatoma with a human U2 RNA pseudogene (21). The numbers below the U2 RNA sequence correspond to the T_1-RNase oligonucleotides shown in Fig. 1.

et al (32) to establish maximum base-pairing. It is consistent with the preferential cleavage sites of T_1-RNase, RNase A, and S1 nuclease.

Secondary structure of U2 RNA in U2 RNP. Analysis of the T_1-RNase digestion fragments of U2 RNP as done for U1 RNA in U1 RNP (4), showed there was a major cleavage site at nucleotide 64, one of the positions preferentially cleaved by T_1-RNase in isolated RNA. This result indicates that the conformation of U2 RNA in isolated RNA and in the U2 RNP have similarities as noted earlier for U1 RNA and U1 RNP (4).

DISCUSSION

This study presents a secondary structure of U2 RNA in solution and in RNP form. Like U1 RNA (4), U2 RNA appears to have a stable secondary structure in solution, it may have a similar conformation in RNP form as suggested by the cleavage shown with T_1-RNase. Since the 5'-end of U1 RNA was shown to be complementary to splice junctions of hnRNA (2,3) the function

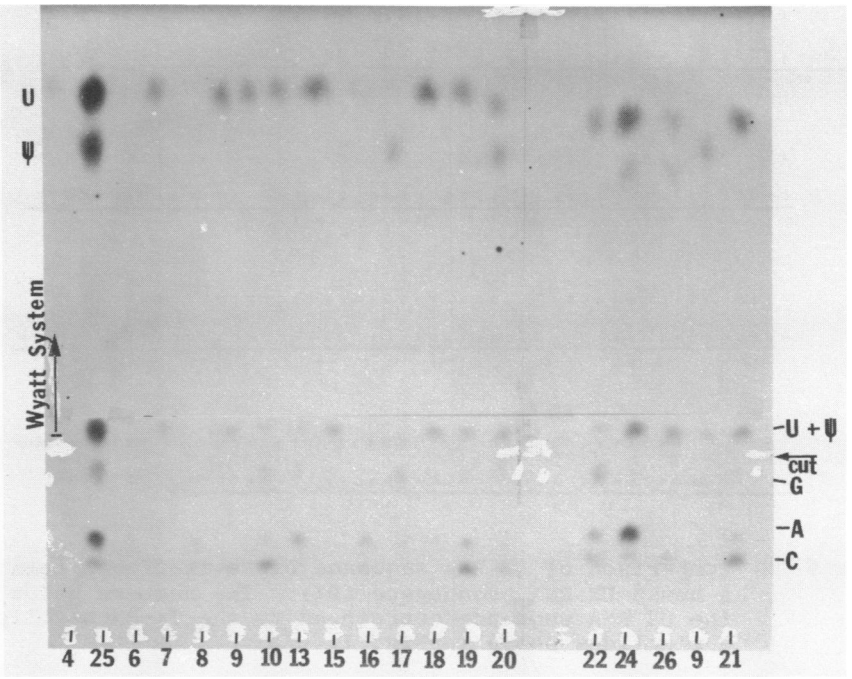

Fig. 7 Analysis of U2 RNA fragments for pseudouridine residues. The T_1-RNase oligonucleotides were digested with T_2-RNase and fractionated on a 3 MM paper at pH 3.5 by electrophoresis. The 3 MM paper was cut between Gp spots and Up spots and the mixture of Up and Ψp spots were separated using the Wyatt system (30). The numbers of fragments correspond to fragments in Fig. 1. T-9 was digested with nuclease P1, instead of T_2-RNase.

of RNAs in snRNPs may be two-fold. First they may provide a stable skeleton to bind protein(s) and second, they may recognize splice junctions to confer fidelity in splicing reactions. Although the 5'-end of free U2 RNA was accessible to T_1-RNase, the cleavage at nucleotide 5 was not a high yield cleavage site, compared to positions 64 and 164 in free U2 RNA.

This study also presents a modified sequence of U2 RNA which differs from the previous sequence (8) mainly in the sequence of pyrimidine residues in large T_1-RNase oligonucleotides. Among snRNAs of Novikoff hepatoma heterogeneity was

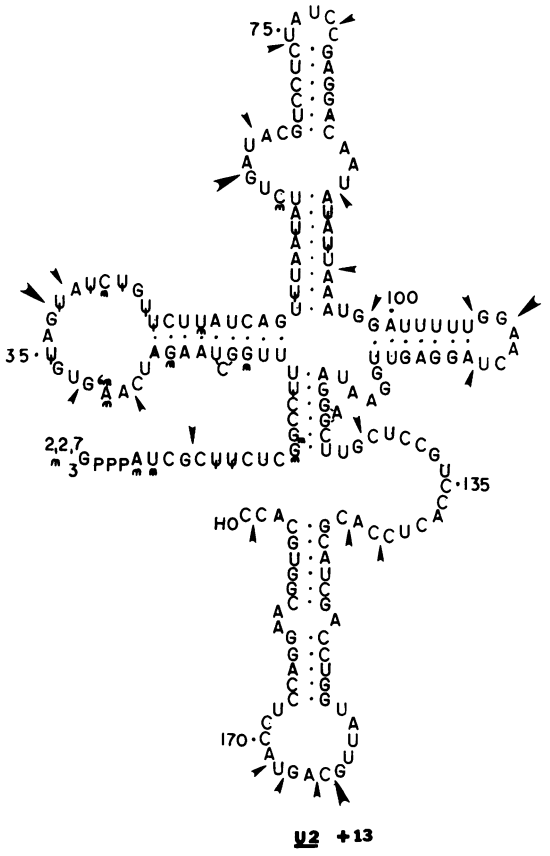

Fig. 8 One of the possible secondary structures of U2 RNA. The large arrowheads point to most susceptible regions in U2 RNA to T1-RNase cleavage. The cleavages at positions 164, 64, 37 and 107 (large arrowheads) occurred with frequencies of 75%, 60%, 10% and 8%, respectively. The small arrowheads show the positions cleaved by RNase A or T_1-RNase or S1 nuclease. The stability number calculated according to Tinoco et al (38) was + 13.

found in U3 RNA (13), U4 RNA (14) and in U6 RNA (1,15). In all these cases, the heterogeneity was confined to the 3'-ends; the 5'-ends were conserved. Heterogeneity of U2 RNA was also found at the 3'-end; in 30% of the U2 RNA, an additional A residue was at position 189 (7,8).

U2 RNA is the second most abundant U-snRNA (33) and was shown to be associated with hnRNP (34-37). Alignment of the 5'-end of U2 RNA to several intron/exon splice junctions (2) showed that U2 RNA had a less satisfactory complementarity to splice junctions compared to U1 RNA.

Conservation of U2 RNA sequence during evolution. U2 RNA of HeLa cells (40), IMR-90 cells (40) and Ehrlich ascites cells (41) was analysed by fingerprinting and analysis of oligonucleotides by further enzymatic digestions. The results indicated that the sequences of U2 RNA of mouse rat and human are very similar. The human U2 RNA (40) was shown to contain A at position 117 and mouse U2 RNA (41) appears to contain G at position 110 in a fraction of mouse U2 RNA population. Since these sequence studies on human or mouse U2 RNA are not complete, nucleotide substitutions at other positions cannot be ruled out.

The modified sequence of U2 RNA was compared with known sequences of mouse U2 RNA gene (42) and human U2 RNA pseudogenes (21,39). The Novikoff hepatoma 188 nucleotide long U2 RNA sequence was identical to the mouse U2 RNA gene sequence (except for nucleotide 110 which was G in mouse DNA and RNA). The U2 RNA sequence was also in perfect register with human U2 RNA pseudogenes (21,39; Fig. 6).

ACKNOWLEDGEMENTS

These studies were supported by the Cancer Research Grant CA-10893, P.4, awarded by the National Cancer Institute, Department of Health, Education and Welfare. We thank Mrs. Rose K. Busch for supply tumor-bearing rats and Dr. Alan M. Weiner for useful discussions.

REFERENCES

1. Reddy, R. and Busch, H. (1981) Cell Nucleus, Vol. 8, Ed. H. Busch, Academic Press, pp. 161-308.
2. Lerner, M. R., Boyle, J. A., Mount, S. M., Wolin, S. L. and Steitz, J. A. (1980) Nature 283, 220-224.
3. Rogers, J. and Wall, R. (1980) Proc. Natl. Acad. Sci. USA 77, 1877-1879.
4. Epstein, P., Reddy, R. and Busch, H. (1981) Proc. Natl. Acad. Sci. USA 78, 1562-1566.
5. Branlant, C., Krol, A., Ebel, J., Lazar, E., and Jacob, M. (1981) Nucleic Acid Res. 9, 841-858.

6. Reddy, R., Ro-Choi, T. S., Henning, D. and Busch, H. (1974) J. Biol. Chem. 249, 6486-6494.
7. Shibata, H., Reddy, R., Henning, D., Ro-Choi, T. S. and Busch, H. (1974) Molec. Cell Biol. 4, 3-19.
8. Shibata, H., Ro-Choi, T. S., Reddy, R., Choi, Y. C., Henning, D. and Busch, H. (1975) J. Biol. Chem. 250, 3909-3920.
9. Branlant, C., Krol, A., Ebel, J., Lazar, E., Gallinaro, H., Jacob, M., Sri-Widada, J. and Jeanteur, P. (1980) Nucleic Acids Res. 8, 4143-4154.
10. Roop, D. R., Kristo, P., Stumph, W. E., Tsai, M. J. and O'Malley, B. W. (1981) Cell, in press.
11. Reddy, R., Henning, D. and Busch, H. (1981) Biochem. Biophys. Res. Commun. 98, 1076-1083.
12. Reddy, R., Henning, D. and Busch, H. (1979) J. Biol. Chem. 254, 11097-11105.
13. Reddy, R., Henning, D. and Busch, H. (1980) J. Biol. Chem. 255-7029-7033.
14. Reddy, R., Henning, D. and Busch, H. (1981) J. Biol. Chem. 256, 3532-3538.
15. Kato, N. and Harada, F. (1981) Biochem. Biophys. Res. Commun. 99, 1477-1485.
16. Krol, A., Branlant, C., Lazar, E., Gallinaro, H. and Jacob, M. (1981) Nucleic Acids Res., in press.
17. Kato, N. and Harada, F. (1981) Biochim. Biophys. Res. Commun. 99, 1468-1476.
18. Krol, A., Gallinaro, H., Lazar, E., Jacob, M. and Branlant, C. (1981) Nucl. Acids Res. 9, 769-788
19. Epstein, P., Reddy, R., Henning, D. and Busch, H. (1980) J. Biol. Chem. 255, 8901-8906.
20. Harada, F., Kato, N. and Nishimura, S. (1980) Biochem. Biophys. Res. Commun. 95, 1332-1340
21. Denison, R. A., Van Arsdell, S., Berstein, L. B. and Weiner, A. (1981) Proc. Natl. Acad. Sci. USA 78, 810-814.
22. Mauritzen, C. M., Choi, Y. C., and Busch, H. (1970) Methods in Cancer Res. 6, 253-282.
23. Ro-Choi, T. S., Moriyama, Y., Choi, Y. C., and Busch, H. (1970) J. Biol. Chem. 245, 1970-1977.
24. Donis-Keller, H., Maxam, A. M. and Gilbert, W. (1977) Nucleic Acids Res. 4, 2527-2538.
25. England, T. E. and Uhlenbeck, O. C. (1978) Nature 275, 560-561.
26. Raj Bhandary, U. L. (1968) Methods Enzymol. 12B, 317.
27. Brownlee, G. G., Sanger, F. and Barrell, B. G. (1968) J. Mol. Biol. 34, 379-412.
28. Fuke, M. and Busch, H. (1977) Nucleic Acids Res. 4, 339-352.
29. Peattie, D. A. (1979) Proc. Natl. Acad. Sci. U.S.A. 76, 1760-1764.
30. Wyatt, G. R. (1951) Biochem. J. 48, 584-590.
31. Reddy, R., Ro-Choi, T. S., Henning, D., Shibata, H., Choi, Y. C. and Busch, H. (1972) J. Biol. Chem. 247, 7245-7250.
32. Korn, L. J., Queen, C. L. and Wegman, N. M. (1977) Proc. Natl. Acad. Sci. USA 74, 4401-4405.
33. Weinberg, R. A. and Penman, S. (1968) J. Mol. Biol. 38, 289-304.

34. Sekeris, C. E. and Niessing, J. (1975) Biochem. Biophys. Res. Commun. 62, 642-650.
35. Northemann, W., Scheurlen, M., Gross, V. and Heinrich, P. C. (1977) Biochem. Biophys. Res. Commun. 76, 1130-1137.
36. Guimont-Ducamp, C., Sri-Widada, J. and Jeanteur, P. (1977) Biochimie 59, 755-758.
37. Gallinaro, H. and Jacob, M. (1979) FEBS Lett. 104, 176-182.
38. Tinoco, I., Uhlenbeck, O. C. and Levine, M. D. (1971) Nature (London) 230, 362-366.
39. Van Arsdell, S., Denison, R., Bernstein, L., Weiner, A., Manser, T. and Gesteland, R. (1981) Nucleic Acids Res. (in press)
40. Nohga, K., Reddy, R. and Busch, H. (1981) Cancer Res. 41, 2215-2220.
41. Lerner, M. and Steitz, J. A. (1979) Proc. Natl. Acad. Sci. 76, 5495-5499.
42. Huang, R. C. Personal communication.