Volume 9 Number 4 1981 Nucleic Acids Research

The conformation of chicken, rat and human U1A RNAs in solution

Christiane Branlant+, Alain Krol and Jean-Pierre Ebel
Institut de Biologie Moléculaire et Cellulaire du CNRS, Université Louis Pasteur de Strasbourg, 15, rue René Descartes, 67084 Strasbourg Cédex, and
Hélène Gallinaro, Eliane Lazar and Monique Jacob
Laboratoire de Génétique Moléculaire des Eucaryotes, Unité de Biologie Moléculaire et de Génie Génétique, 11, rue Humann, 67085 Strasbourg Cédex, France

Received 10 December 1980

ABSTRACT

Chicken, rat and human U1A RNAs in solution, were examined for secondary structure, using several methods including hydrolysis by various nucleases, hybridization to DNA oligomers and analysis of fragment interactions. The experimental results showed that the three U1A RNAs have the same structure, stable over a wide range of pH and ionic conditions. They allowed the selection of one out of several possible models constructed from the data of primary structure. This model is characterized by 4 hairpins and two single-stranded regions, the two hairpins from the 3' part of the molecule bearing very stable stems. In addition, the experimental results showed that in contrast to the 5' half of the molecule, the 3' half has a compact conformation probably stabilized by tertiary interactions. The 5' end of U1A RNA is accessible and free of base-pairing so that it might base-pair with regions of other RNA molecules, for instance, with the extremities of introns as has been recently proposed in a model of splicing.

INTRODUCTION

We recently determined the nucleotide sequence of U1A RNA from chicken, rat and man (1). Observations made in the course of this work suggested the existence of very stable base-pairings within the molecule. This paper reports an experimental study of the conformation of U1A RNA in solution. This study is based on different methods including hydrolysis by various nucleases, hybridization to DNA oligomers and analysis of fragment interactions. The results are interpretated on the basis of putative secondary structures constructed from chicken, rat and man U1A RNA primary structures. Implication of the data for the model of splicing in which U1A RNA align the extremities of introns (2,3,4) is discussed.

MATERIALS AND METHODS

1. Preparation of 3' end-labeled U1A RNA was previously described (1).

2. Limited digestions of 3' end-labeled RNA in denaturing conditions were performed with chemical reagents according to Peattie (5) or with ribonu-

cleases as previously described (6).

3. Limited digestions of 3' end-labeled U1A RNA in non-denaturing conditions. For each assay, 0.1 µg of U1A RNA mixed to 2.5 µg of carrier tRNA was used. Immediately after digestion, the RNA was loaded on polyacrylamide slab gels (900 x 300 x 0.5 mm) made up in Tris-borate-buffer.
Digestion with T1 RNase was in 6 µl of 10 mM Tris-HCl pH 7.5 buffer containing 1 or 10 mM $MgCl_2$. After a 10 min preincubation at 37°C the hydrolysis was performed for 2-10 min at the same temperature with 0.01 U of T1 RNase. It was stopped by addition of 1 µl of 400 mM NaH_2PO_4 pH 3.3 (7).
Digestion with S1 nuclease was in 6 µl 50 mM KCl, 1 mM $ZnCl_2$, 25 mM Na acetate pH 4.5 containing 1 or 10 mM $MgCl_2$. In some experiments this buffer was replaced by 25 mM Na citrate, 10 mM $MgCl_2$, 1 mM $ZnCl_2$ pH 6.3 buffer. After a 10 min preincubation at 37°C, the digestion was performed for 2-10 min at the same temperature with 2.5 U of S1 nuclease. It was stopped by addition of 5 µl of 10 mM ATP (7).
Digestion with *Naja oxiana* RNase (8) was carried out for 2-10 min at 0°C in 6 µl 10 mM $MgCl_2$, 10 mM Tris-HCl pH 7.5 buffer in the presence or absence of 350 mM KCl. It was stopped by addition of 2 µl 100 mM EDTA (8).

4. Hybridization of RNA to DNA oligomers and cleavage by RNase H. The procedure was that described by Donis-Keller (9). 1 µg of U1A RNA was hybridized with 2.5 µg of (dCp)6, (dAp)6 or (dGp)6 oligonucleotides in 4 mM $MgCl_2$, 1 mM dithiothreitol, 40 mM Tris-HCl pH 7.9 buffer. The reaction mixture was heated at 50°C for 3 min and then incubated at 30°C for 30 min. The hydrid was hydrolyzed with 1 U of RNase H for 30 min at 30°C. The resulting RNA fragments were then labeled with 5'-{^{32}P} pCp and RNA ligase (10) and fractionated by polyacrylamide gel electrophoresis.

5. Study of fragment interactions. 1 µg of U1A RNA was preincubated for 10 min at 37°C in 10 µl of 0.01 mM $MgCl_2$, 15 mM mercaptoethanol, 10 mM Tris-HCl pH 8. After addition of 0.01 U of T1 RNase, 200 µCi of highly {^{32}P}-labeled ATP and 2 U of T4 polynucleotide kinase, digestion and labeling were performed simultaneously for 10 min at 37°C. The labeled digest was immediately loaded on an 8% polyacrylamide gel made up in 40 mM Tris acetate buffer pH 8.4 (11). After electrophoresis, the strip of gel containing the fractionated labeled fragments was included at the top of a 15% polyacrylamide gel made up in Tris-borate buffer, pH 8.3 containing 8 M urea and reelectrophoresed.

RESULTS

1. Experimental study of the secondary structure of U1A RNA in solution

Three different approaches were used.

1.1 Limited digestion of 3' end-labeled U1A RNA with T1, S1 and *Naja oxiana* nucleases (8). As previously described for tRNA by Wrede et al. (7), 3' end-labeled U1A RNA was partially digested with T1 and S1 nucleases under conditions allowing the cleavage of single-stranded regions only. In addition, we also employed *Naja oxiana* nuclease under the conditions of cleavage of nucleotides involved in secondary or tertiary interactions (12). In all cases, the resulting fragments were electrophoresed. For the identification of the nucleotides preferentially cleaved in such mild conditions, U1A RNA was treated in denaturing conditions as for RNA sequencing and the fragments were electrophoresed in parallel. As apparent in Figure 1, differences of electrophoretic mobilities are observed for products having the same base residues but issued from different types of digestion. This is due to the fact that digestions with S1 and *Naja oxiana* nucleases release products bearing a phosphate at their 5'end whereas the fragments released by T1 RNase or boiling water lack this phosphate. The difference of mobility decreases relatively when the length of the fragment increases. As the localization of cuts by S1 and *Naja oxiana* nucleases became rough near the 5' end of the molecule, different gel concentrations had to be used for a complete analysis of the digestion patterns.

3' end-labeled U1A RNAs from chicken, rat and HeLa cells were partially digested with the 3 enzymes in the presence of 10 mM $MgCl_2$ which maximizes base-pairing. Furthermore, the buffer conditions were selected in order to maximize enzyme activities. S1 nuclease was employed at its optimum pH (4.5) and digestion with *Naja oxiana* RNase was performed in the presence of 350 mM KCl which increases its activity (12). The reactions were stopped by addition of NaH_2PO_4, pH 3.3 for T1 RNase (7), of ATP for S1 nuclease (7) and of EDTA for *Naja oxiana* RNase. The digests were immediately fractionated by gel electrophoresis. As can be observed in Figure 1, the patterns of digestion are the same for U1A RNA from chicken, rat and man, indicating that their secondary and probably tertiary structures are similar.

A variation of the duration of hydrolysis did not modify the patterns of digestion. (Fig. 1a). Only a reinforcement of certain bands was observed. This means that the base-pairs of U1A RNA are resistant to strong hydrolysis even by *Naja oxiana* RNase which specifically cleaves structured regions.

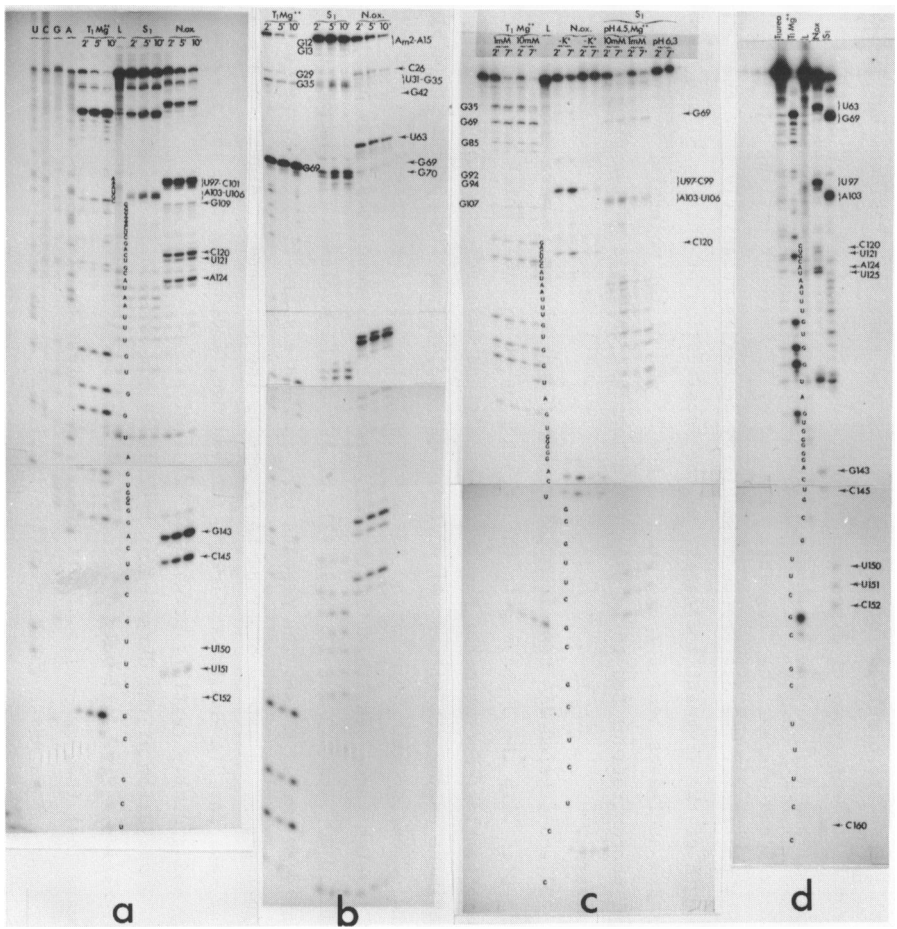


Figure 1. T1, S1 and *Naja oxiana* nucleases limited digestions of 3' end-labeled U1A RNAs.
a and b. Fractionation on 20% (a) and 15% (b) polyacrylamide gels of the digests from rat U1A RNA. RNA was digested under conditions maximizing base-pairing and enzymes activities as described in Methods. For each enzyme incubation was performed for 2,5 and 10 min. Positions of the cuts were determined by reference to statistical chemical digests (U, C, G, A).
c. Fractionation on a 20% polyacrylamide gel of the digests from chicken U1A RNA. RNA was digested under various ionic and pH conditions as described in Methods. T1 and S1 nuclease digestions were performed in the presence of 1 mM or 10 mM Mg^{++}, *Naja oxiana* nuclease digestion in the presence ($+K^+$) or absence ($-K^+$) of 0.35 M KCl. In each case the incubations were performed for 2 and 7 min. Positions of the cuts were determined by reference to a statistical digest in boiling water (H_2O) and with T1 RNase (G).
d. Fractionation on a 20% polyacrylamide gel of the digests from human U1A RNA. RNA was digested for 7 min under the same conditions as in Figure 1a and 1b and positions of the cuts were determined as in Figure 1c.

Then, we looked for the effect of mono and divalent cation concentrations on U1A RNA structure. The influence of divalent cations was tested with T1 and S1 nucleases. As can be seen in Figure 1b, when the Mg^{++} concentration was lowered from 10 to 1 mM, similar patterns of digestion by T1 and S1 nucleases were observed. All the major bands were detected but with slightly different yields. This might be related to a change in the tertiary structure of U1A RNA but not in its secondary structure. The influence of monovalent cations was tested with *Naja oxiana* nuclease. As shown by the lower intensity of the bands on the gel (Fig. 1b), the rate of hydrolysis decreases in the absence of KCl. This was expected from the requirement of monovalent cations for maximum enzyme activity. However, the relative yield of the bands was conserved indicating that the structure of U1A RNA was not modified in the presence or absence of 0.3 M KCl. Therefore, the structure of U1A RNA is stable over a wide range of ionic conditions. As the results of T1 and S1 nuclease digestions which were performed at pH 7.5 and 4.5 respectively, were in good agreement, the structure of U1A RNA should also be stable over a wide range of pH.

1.2 RNase H digestion of U1A RNA hybridized with deoxyoligonucleotides and study of the released fragments. This approach was used 1) in order to identify single-stranded regions owing to their ability to bind oligodeoxynucleotides, 2) in order to produce U1A RNA fragments whose structure can be analyzed using T1, S1 and *Naja oxiana* nucleases, as described above.

1.2.1 Identification of hybridized regions. U1A RNA was incubated in the presence of oligo (dGp)6, (dAp)6 and (dCp)6 under mild conditions preserving RNA secondary structure. For the detection and identification of RNA-DNA hybrids, the mixture was hydrolysed with RNase H which specifically digests DNA-RNA hybrids (9). The resulting products with free 3'-OH were 3' end-labeled by ligation with {5'-^{32}P}pCp. They were fractionated by polyacrylamide gel electrophoresis. No specific cleavage of U1A RNA occurred with oligo (dAp)6 and oligo (dCp)6. In contrast, specific cleavages were observed when U1A RNA was hybridized with oligo (dGp)6 (Fig. 2). Four fragments were released [1-4], the same for U1A RNA from chicken, rat and HeLa cells. These fragments were analyzed using the chemical method for RNA sequencing. The largest fragment extended from nucleotide 74 to the 3' end of the molecule (for nomenclature see Fig. 4). Fragments 2, 3 and 4 only differed by one nucleotide at their 3' end (73, 72, 71 respectively) and extended up to the 5' end of the molecule. Taking into account that at least 4 contiguous

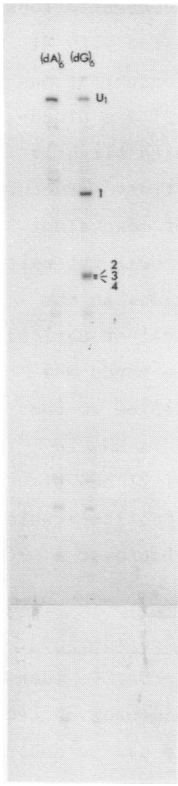


Figure 2. Fractionation of RNase H digests of hybrids between U1A RNA and deoxyoligonucleotides. Rat U1A RNA was hybridized with (dGp)6 and (dAp)6. The hybrids were digested with RNase H. The resulting products were labeled with {5'-^{32}P}pCp and fractionated on a 15% polyacrylamide gel.

base-pairs are necessary for recognition by RNase H (9), the results indicate that the sequence C-U-C-Cp at position 72-75 hybridized to oligo (dGp)6.

1.2.2 Enzymatic study of RNase H fragment 1. In order to check whether nucleotides from regions 1-73 and 74-165 were base-paired, fragment 1 and intact U1A RNA, labeled at their 3' end, were digested in parallel with T1, S1 and *Naja oxiana* nucleases. The resulting products were fractionated on the same gel (Fig. 3). Nucleotides 91-165 were cleaved in the same way in both cases, whereas nucleotides 76-83 were strongly cleaved by S1 nuclease in fragment 1 but not in U1A RNA. This indicates that nucleotides 91-165 have the same structure in fragment 1 and in U1A RNA and that nucleotides 76-83 are base-paired with nucleotides of the region 1-73 in U1A RNA.

1.3 Study of fragment association. U1A RNA was digested with T1 RNase under mild conditions maintaining base-pairing. The released fragments were 5' end labeled during digestion. They were fractionated by two-dimensional electro-

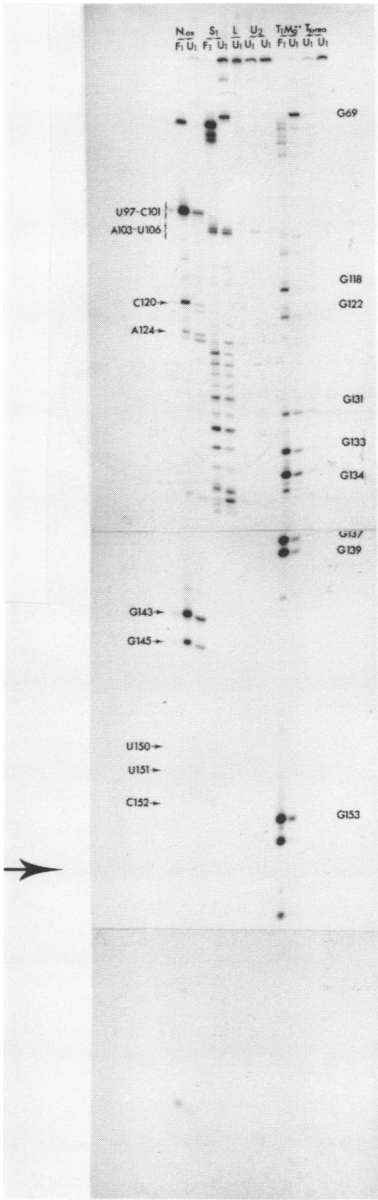
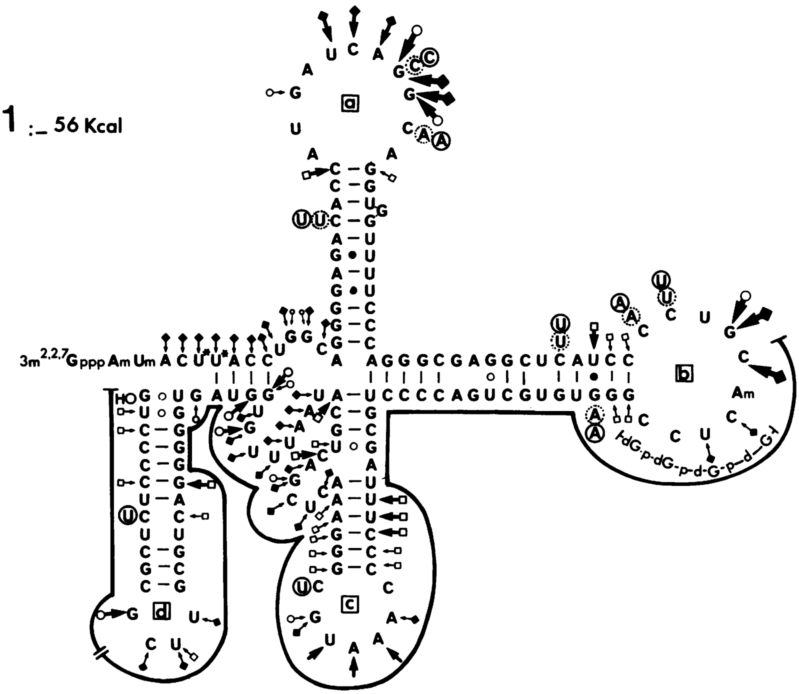


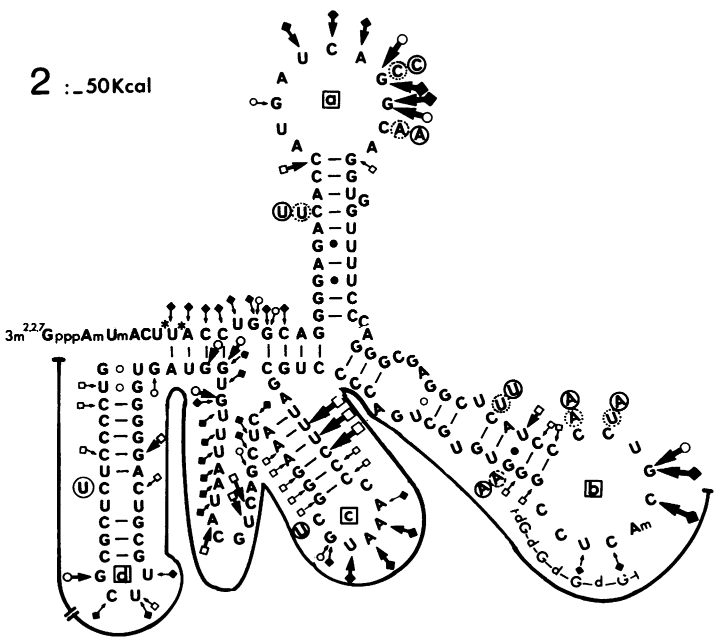
Figure 3. Comparison of T1, S1 and *Naja oxiana* nuclease digests from rat U1A RNA and from its fragment 1. Fragment 1 was obtained as described in Figure 2. Limited hydrolyses with T1, S1 and *Naja oxiana* nucleases were performed for 7 min under the conditions of Figure 1a.

phoresis on polyacrylamide gel. The buffers were choosen in order to maintain base-pairings in the first dimension and to destabilize them in the second dimension (11). Each fractionated labeled fragment was characterized by the

1 :_ 56 Kcal



2 :_50Kcal



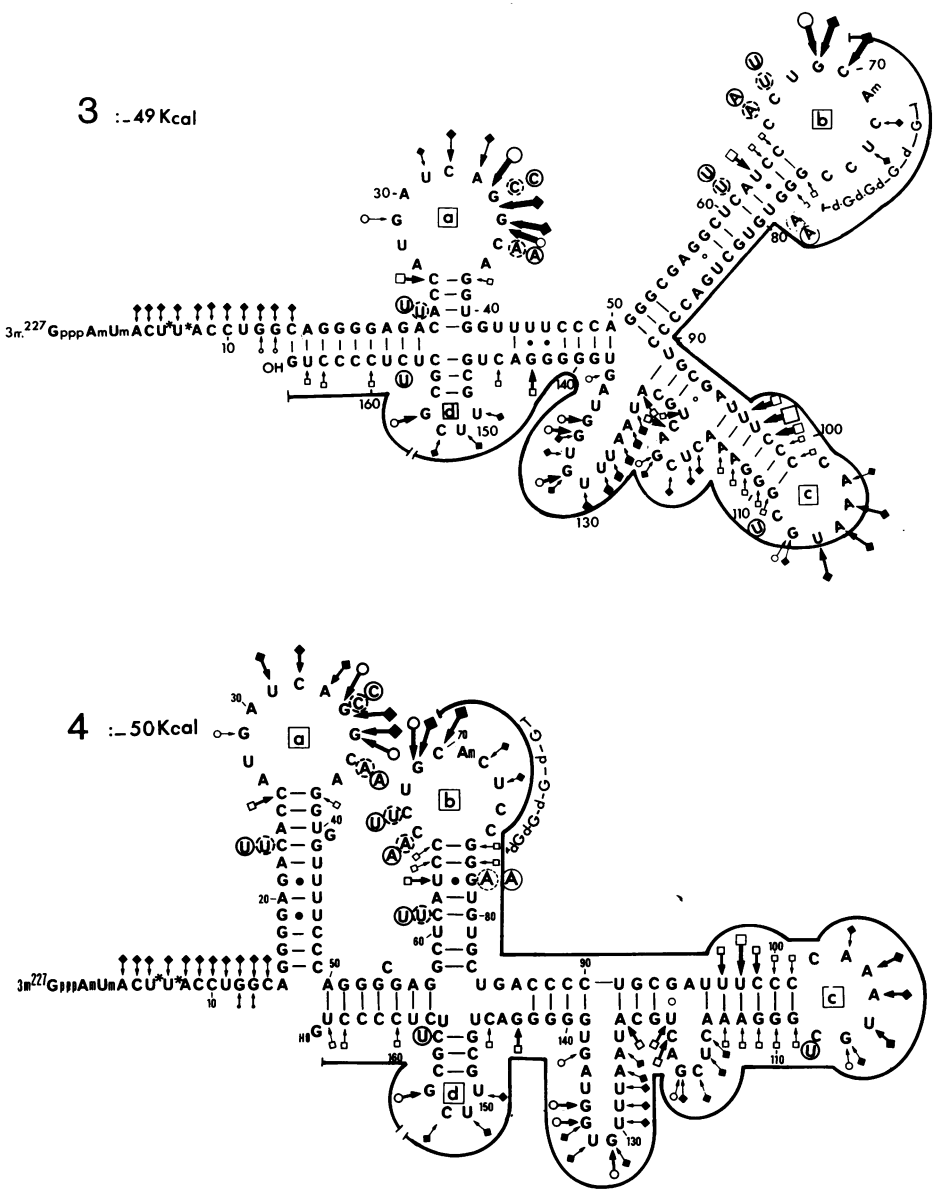


Figure 4. Possible base-pairing schemes for U1A RNA. The sequence is that of chicken U1A RNA (1). The base substitutions found for rat and man U1A RNAs are indicated ((◌) rat, ◯ man). The positions of enzymatic cleavages are those of : ◀—○ T1 RNase, ◀—◆ S1 nuclease, ◀—□ *Naja oxiana* RNase. The position of hybridization with (dGp)6 is shown, as well as that of the two fragments found associated under non-denaturing conditions (——). The free energies of the corresponding structures is indicated.

enzymatic method for RNA sequencing. Due to steric hindrances, all the fragments produced by T1 RNase digestion were not labeled by the large molecule of polynucleotide kinase (140 000 daltons). This complicated somewhat the interpretation of the results. Nevertheless, we determined that fragments 70-153 and 154-165 were strongly associated (for nomenclature see Fig. 4).

2. The secondary structure of U1A RNA in solution

2.1 U1A RNA putative secondary structures. On the basis of the primary structure of chicken, rat and man U1A RNAs, several base-pairing schemes can be proposed for this RNA. The four most stable ones (1 to 4) are represented in Figure 4. All of them have the same hairpin-loops (a to d). They only differ in their base-paired regions. Their thermodynamic stabilities are indicated in Figure 4, but the values must be considered as rough. Indeed, whereas the free energy corresponding to the formation of base-pairs within a classical hairpin are rather well defined (13, 14), no experimental data allow the calculation of destabilizing effect due to the existence of branched structures. Pipas and Mc Mahon (15) established a computer program for determining RNA secondary structure. In this program, they assign zero energy to every multibranched internal loop. As Salzer (16), we feel that this may not correspond to reality. But the problem is how to estimate the corresponding destabilizing energy. We considered a branched structure as resulting from a stem with an interior or a bulge loop which is itself structured. Thus, the corresponding free energy was the sum of the free energies due to base-pairings in the stem and in the structured loop, diminished by the destabilizing effect due to the presence of this loop. The value taken for this effect was that established by Tinocco et al. (14) for a non structured loop of the same length. However, this way of calculation does not take into account the steric hindrance that may exist at the level of a branched loop. For instance, on a steric point of view, the branched structure proposed in model 2 might be difficult to form, and the stability of model 1 compared to those of models 2, 3 and 4 might be higher than expected from the calculated values.

2.2 The most probable structure. These 4 models may now be discussed on the basis of the experimental results described above.

Hybridization experiments are in agreement with the 4 models. Thus all the runs of 5 guanines which might have hybridized to (dCp)6 are base-paired. Among the runs of uridines which might have hybridized with (dAp)6, only U-U-U-G-Up [128-132] is free in the models but is probably not long enough to form a stable hybrid (17). Most of the long stretches of pyrimidines which would be good candidates for hybridization with (dGp)6 are base-paired in

the models and only C-U-C-Cp [72-75] which indeed hybridized with this oligonucleotide is single-stranded (hairpin-loop b). These experiments strongly suggest that U1A RNA bears the features common to the 4 models, in particular hairpin-loop b.

The existence of this hairpin-loop b and of the three other ones is in good agreement with the sensitivity of the corresponding parts of the molecule to T1 and S1 nucleases. This is also true for the single-stranded region 125-137. The results of enzymatic digestion of the region 118-121 are not in such good agreement with the single-stranded structure proposed for this region in all the models, but this may be accounted for by the existence of tertiary interactions or of steric hindrance as will be discussed below. An additional single-stranded region [120-124] is proposed in model 2 but may be eliminated in view of hydrolysis by *Naja oxiana* nuclease and of absence of hydrolysis by T1 and S1 nucleases. On the other hand, as the 5' terminal region [1-14] is sensitive to S1 nuclease, it should be single-stranded. This is only the case in model 4, but it should be pointed out that in model 3 only cytidine 14 is base-paired with the 3' terminal guanine, the other 13 nucleotides being free, and in model 1 depending upon the modification of nucleotide 7, the base-pairing proposed for the segment 7-10 may be unstable. Therefore, we will now introduce a model 1' (Fig. 5) which only differ from model 1 by the opening of the base-pairing between segments 7-10 and 133-136. Model 2 can be eliminated on the basis of enzymatic digestion results but the choice between models 1', 3 and 4 requires examination of the other experimental results described above.

In the study of fragment 1 released by RNase H digestion, we observed that nucleotides from the region 1 to 73 should be hybridized to nucleotides from region 76-83 but not to those of region 91-165. This result is compatible only with model 1'. On the other hand, we observed a strong association between fragment 70-153 and 154-165. This association might well be explained by the existence of the stable 3' terminal hairpin in model 1'. In models 3 and 4, the fragment 154-165 would be strongly associated with fragments containing the sequences 14-22 and 50-57, respectively, a fact which is not compatible with the experimental results. Therefore, these two types of experiments strongly suggest that U1A RNA in solution has the structure 1'.

Observations, previously made in the course of U1A RNA primary structure determination, reinforce this idea. Indeed, the major difference between model 1' and models 3 and 4 is the existence of the stable 3' terminal hairpin in model 1' only. When the enzymatic method of RNA sequencing was applied to 3' or 5' end-labeled U1A RNA, nucleotides 139 to 165 were hardly cleaved by the

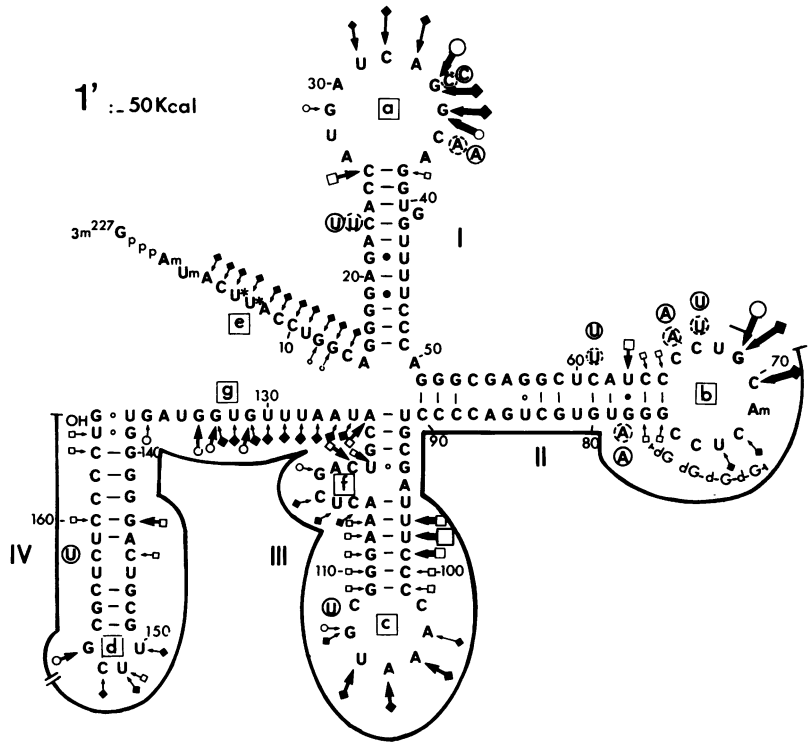


Figure 5. The secondary structure of U1A RNA in solution. Model 1'. Same symbols as in Figure 4.

nucleases, with the exception of nucleotides 153, 157, 158 located in the interior and top loops of the 3' terminal hairpin (1) (Fig. 6). These results are difficult to explain with models 3 and 5. When the chemical methods of RNA sequencing were used with 3' end-labeled U1A,RNA nucleotides 139-165 were cut by the chemical reagents under strongly denaturing conditions but the bands corresponding to the cleavage of guanines 140-143 were compressed on the sequencing gels (Fig. 6). This confirms the existence of the 3' terminal hairpin of model 1'. Let us remark that similarly, band compression and absence of enzymatic cleavage also confirmed the existence of the base-paired region 96-101 and 109-114 (1).

The same structure could be deduced from experiments with 3' and 5' end-labeled U1A RNA showing that the cytidine added during 3' end-labeling did not change the RNA structure. In addition, it should be mentioned that the

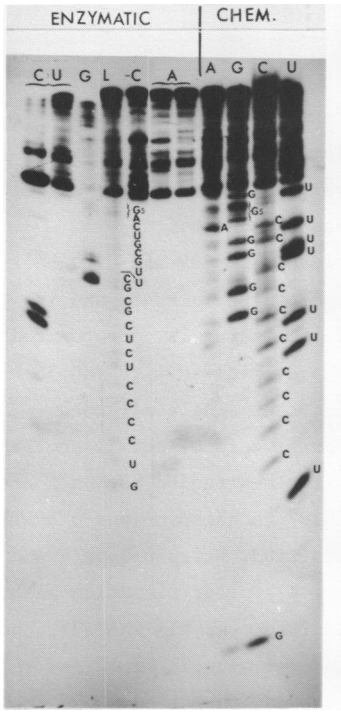


Figure 6. Sequencing gel of 3' end-labeled U1A RNA from chicken. 3' end-labeled U1A RNA was analyzed by both the enzymatic and the chemical methods for RNA sequencing. In the enzymatic method, we used T1 (G), pancreatic (U + C), Phy I (-C) and U_2 (A) RNases hydrolyses and digestion in boiling water (L). The various digests were fractionated on a 25% polyacrylamide gel.

base substitutions occurring during evolution from chicken to man either occurred in single-stranded regions or did not alter base-pairing (6 and 3 substitutions, respectively).

Thus, several lines of evidence are in favor of structure 1' for U1A RNA in solution. For further discussion, the hairpins of model 1' will be designated from I to IV and the single-stranded regions from a to g.

3. Informations on the tertiary structure of U1A RNA in solution. As seen in Figure 1 and schematized in Figure 4, the rates of cleavage of the various regions of U1A RNA by each of the three nucleases differ widely. Comparisons of these rates may provide information on the tertiary structure of U1A RNA. Several considerations should be taken into account for such investigation. 1) A weak enzymatic cleavage may be the consequence of steric hindrance or of the involvement of the concerned RNA region in tertiary interaction. 2) S1 nuclease being a larger molecule than T1 or *Naja oxiana* nucleases (350 000 instead of 11 000 and 15 000 daltons, respectively), steric hindrances will be more marked for S1 than for the other nucleases. 3) When two enzymatic

cleavages occur at the same rate, that proximal to the labeled end will be visualized as a stronger band on the sequencing gel.

3.1 Comparison of the behaviour of the various single-stranded regions. Taking point 3 into account, as nucleotide 69 is extremely sensitive to both T1 and S1 nucleases, and as the labeling was at the 3' end, in the experiments of Figure 1, a comparison of rates of cleavage can only be made on the single-stranded regions located between nucleotides 69 and 165.

According to high sensitivity to T1 and S1 nucleases of G69 and C70, hairpin-loop b should be very accessible at the surface of the molecule and is probably not involved in tertiary interactions. This finding is corroborated by the fact that the sequence C-U-C-Cp of this loop hybridized with (dGp)6 and that the hybrid was hydrolyzed by RNase H. The other single-stranded regions towards the 3' end of U1A RNA are less sensitive to S1 and T1 nucleases than hairpin-loop b. Hairpin-loop c and the long single-stranded region g are cleaved at a moderate rate. The slow cleavage of G107 in hairpin-loop c by T1 RNase might be accounted for by a base-pairing with C102. Hairpin-loop d and the bulge-loop f are hardly cleaved by S1 nuclease. A similar situation was found for $tRNA^{Phe}$ when analyzed under the same conditions as U1A RNA (7), the D and T loops were not cleaved by S1 nuclease. Wrede et al. (7) explained this result by the existence of tertiary interactions between these loops and this may also be the case for the d and f loops of U1A RNA. This view is also supported by the fact that these 2 regions are slowly attacked by *Naja oxiana* nuclease which is expected to cleave nucleotides involved in tertiary interactions in addition to base-paired nucleotides (12) .

Though the rates of enzymatic cleavage in region 1-69 cannot be compared to those in region 70-165, it may nevertheless be observed that guanines 34 and 35 of loop a are rather sensitive to S1 and T1 nucleases suggesting that hairpin-loop a is also accessible.

The conclusions of this comparative study of the sensitivities of the various single-stranded regions to T1 and S1 nucleases attack are in good agreement with observations made during fragment association study. As described above, in such experiment, U1A RNA was digested with T1 RNase under conditions maintaining base-pairing, and the resulting fragments were simultaneously labeled with $\{\gamma\text{-}^{32}P\}$ATP and polynucleotide kinase. We found that, though T1 RNase cleaved most of the single-stranded regions, only fragments whose 5' end was located in hairpin-loops a and b were strongly labeled. Poor labeling at other positions probably reflects steric hindrance to the large molecule of T4 polynucleotide kinase (140 000 daltons). This observation confirms that

loops a and b are accessible at the surface of the molecule whereas the single-stranded regions of the 3' half of the molecule are buried.

3.2 Comparison of the behaviour of the double-stranded regions. Since U98 is very rapidly cleaved by *Naja oxiana* nuclease, it is difficult to make a quantitative comparison of the rate of cleavage of the various base-paired regions. Nevertheless, a remark can be made on the peculiar way this enzyme cuts hairpins I, II and III. In each case, it primarily cleaved the parts of the 5' strand close to the hairpin loop. This might be accounted for by enzyme specificity, but such feature was never described. An alternative is that these 5' parts of the hairpin-stems are located on the surface of the molecule and are the most accessible to the enzyme. The high rate of cleavage of the region U97-C101 suggests, in any case, that this region is particularly exposed. The behaviour of the stem of hairpin IV is different as no such specific site of cleavage was observed.

DISCUSSION

The experimental studies described above show that chicken, rat and human U1A RNAs have the same secondary structure, in solution. This structure is characterized by 4 hairpins and two single-stranded regions. It is stable over a wide range of ionic and pH conditions. This is probably due to the presence of stretches of G-C pairs in each of the 4 hairpin-stems. The stability of the hairpins from the 3' region of the molecule is particularly high, since part of hairpin III and hairpin IV even resist 8 M urea at acidic pH. Hairpin IV might protect the 3' end of the RNA from exonuclease attack, as the cap structure might do for the 5' end. In addition, our data show that the 3' part of the molecule has a compact conformation which is probably due to the existence of tertiary interactions, whereas the 5' part of the molecule has a looser and less stable structure. The hairpin-loops a and b are very accessible to nucleases and the stem of hairpins I and II are less stable than those from hairpins III and IV (Fig. 5).

The compact conformation of the 3' end of the molecule may explain the lower yield of 3' end-labeling of U1A RNA compared to those observed for 5S and tRNAs (unpublished observations). It may also explain the absence of base-pairing between segments 7-10 and 133-136.

As mentioned above, calculation of the free energy of the various secondary structure models was difficult due to the presence of branched structures. In our calculation, we penalized such branched structures and, in this way, the

experimentally determined structure 1' had a higher stability than structures 2, 3 and 4. In the absence of such penalization as proposed by Pipas and Mc Mahon (15), a higher stability of structures 3 and 4 as compared to structure 1' would have been found. It must also be remembered that tertiary interactions which are quite marked in the case of structure 1' have a stabilizing effect and were not taken into account in the calculations of free energy.

If structure 1' is the most likely for U1A RNA in solution because of its stability and of its compatibility with experimental results, we cannot exclude that structures 3 or 4 exist *in vivo* or that transitions between these structures occur according to the physiological state. But even if mediated by proteins, the probability of such transition is low in view of the high stability of structure 1' over a wide range of pH and ionic environment. The conformation of U1A RNA *in vivo* is currently investigated.

Like the other snRNAs, U1A RNA was found to be associated with proteins in free snRNPs and in hnRNPs (18-21). It is also directly associated with hnRNA probably by hydrogen bonds (22, 23). The very exposed hairpin-loops a and b are good candidates for binding proteins in hnRNPs or snRNPs, whereas the two single-stranded regions e and f which are not closed by base-pairing within the molecule of U1A RNA are good candidates for association with hnRNA. Regions e and f do not seem to be very exposed at the surface of the molecule, but RNA-RNA interactions are less prevented by steric hindrance than RNA-protein interactions.

According to a certain complementarity between the consensus sequence of premessenger RNAs at the extremities of introns, the 5' end of U1A RNA was proposed to insure the proper alignment of these sequences for correct splicing (2, 3). The authors of this hypothesis considered that the large region from nucleotides 4 to 22 was involved in the mechanism. Within isolated U1A RNA, the segment 4 to 14 is the single-stranded region e, the segment 19-22 is base-paired with the segment 42-45 but due to the presence of two G.U pairs the interaction is not very stable. On the other hand, the three guanines 16-18 are strongly base-paired with cytosines 47-49, and it seems very unlikely that such base-pairing would be displaced in favor of the G.U rich base-pairs made with the consensus sequence (4). Thus, the study of the secondary structure of U1A RNA shows that only segments 4-14 and 19-22 are accessible for hybridization.

On the other hand, by examining possible base-pairing between the sequence 4-22 of U1A RNA and 69 intron sequences around the splice point, we observed that no stable hybrid could be expected between nucleotides 13-22 from U1A RNA and the

premessenger RNA sequence (4). Only the segment A-C-C-Up [8-11] of U1A RNA was complementary to the 2 nucleotides at each intron extremity (U-G, G-A) in all cases. As suggested by sequence examination the stability of such hybrid may possibly be increased by hybridization of adjacent nucleotides in segment 4-12 to adjacent premessenger RNA nucleotides. The experimental data reported here show that the segment 8-11 of U1A RNA in solution is in the middle of a long single-stranded region [1-14] accessible to enzymes. Even if the conformation of U1A RNA changes *in vivo* (to stable models 3 and 4), this segment would remain accessible. Therefore, on the basis of primary and secondary structures of U1A RNA, no objection can be made to the model of splicing assuming that fragment 4-12 of U1A RNA participates in the alignment of the intron sequences of premessenger RNAs.

ACKNOWLEDGMENTS

This work was supported by grants from the "Centre National de la Recherche Scientifique", the "Institut National de la Santé et de la Recherche Médicale" and the "Délégation Générale à la Recherche Scientifique et Technique". We are very grateful to Dr. P. Carbon and S. Vassilenko for their generous gift of *Naja oxiana* nuclease. The excellent technical assistance of B. Muller and L. Kister is acknowledged.

+To whom all correspondence should be sent.

REFERENCES

1. BRANLANT, C., KROL, A., EBEL, J.P., LAZAR , E., GALLINARO, H., JACOB, M., SRI WIDADA, J. and JEANTEUR, P., (1980), Nucl. Acids Res., 8, 4143-4155.
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. GALLINARO, H., LAZAR , E., JACOB, M., KROL, A. and BRANLANT, C., (1980), Mol. Biol. Rep., in the press.
5. PEATTIE, D., (1979), Proc. Natl. Acad. Sci. USA, 79, 1760-1764.
6. BRANLANT, C., KROL, A., MACHATT, M. and EBEL, J.P., (1979), FEBS Letters, 107, 177-181.
7. WREDE, P., WURST, R., VOURNAKIS, J. and RICH, A., (1979), J. Biol. Chem., 254, 9608-9616.
8. VASSILENKO, S.K. and BABKINA, G.T., (1965), Biokhymiya, 30, 705-712.
9. DONIS-KELLER, H., (1979), Nucl. Acids Res., 7, 179-192.
10. ENGLAND, T. and UHLENBECK, O., (1978), Nature, 275, 560-561.
11. VIGNE, R. and JORDAN, B., (1971), Biochimie, 53, 981-986.
12. VASSILENKO, S.K. and RYTE, V., (1975), Biokhimiya, 40, 578-582.
13. BORER, P., DENGLER, B., TINOCO, I. and UHLENBECK, O., (1974), J. Mol. Biol., 86, 843-853.

14. TINOCO, I., BORER, P., DENGLER, B. and LEVINE, M., (1973), Nature New Biology, 246, 40-41.
15. PIPAS, J. and Mc. MAHON, J., (1975), Proc. Natl. Acad. Sci. USA, 72, 2017-2020.
16. SALZER, W., (1977), Cold Spring Harbor Symp. Quant. Biol., 42, 985-1002.
17. MARTIN, F. and TINOCO, I., (1980), Nucl. Acids Res., 8, 2295-2299.
18. RAJ, N.B.K., RO-CHOI, T.S. and BUSCH, H., (1975), Biochemistry, 14, 4380-4385.
19. DEIMEL, B., LOUIS, Ch. and SEKERIS, C., (1977), FEBS Letters, 73, 80-84.
20. NORTHEMANN, W., SCHEURLEN, M., GROSS, V. and HEINRICH, P., (1977), Biochem. Biophys. Res. Commun., 76, 1130-1137.
21. GALLINARO, H. and JACOB, M., (1979), FEBS Letters, 104, 176-182.
22. FLYTZANIS, C., ALONSO, A., LOUIS, Ch., KRIEG, L. and SEKERIS, C., (1978), FEBS Letters, 96, 201-206.
23. JELINEK, W. and LEINWAND, L., (1978), Cell., 15, 205-214.