Volume 5 Number 12 December 1978 Nucleic Acids Research

RNA-RNA interactions in the binding site of protein L24 on 23S ribosomal RNA of E. coli II. Sequence analysis of the interacting fragments

A.Krol, M.A.Machatt, C.Branlant and J.P.Ebel

Institut de Biologie Moleculaire et Cellulaire du CNRS, Universite Louis Pasteur, 15, rue René Descartes, 67000 Strasbourg, France

Received 25 July 1978

ABSTRACT

A ribonucleoprotein complex containing several RNA subfragments from the 5' part of 23S RNA was recovered after digestion of the reconstituted complex between 23S RNA and protein L24. It was suggested in the preceding paper that the RNA subfragments 4B, 10A and 9, which are widely separated in the sequence, strongly interact. These subfragments were previously partially sequenced by the classical fingerprinting methods. Their sequences have now been completed with rapid new RNA sequencing methods. We propose here a base-pairing model showing how these subfragments may interact with one another.

INTRODUCTION

RNA subfragments encompassing the 500 nucleotides at the 5' end of 23S RNA were found associated with protein L24 after T1 ribonuclease digestion of the reconstituted protein L24-23S RNA complex[1]. The work described in the preceding paper[2] showed that an additional RNA subfragment 9, which is not part of these 500 nucleotides, was found in the ribonucleoprotein complex. Furthermore, a strong interaction between subfragments 4B, 9 and 10A was demonstrated.

Subfragments 4B and 10A were previously partially sequenced by the classical fingerprinting technique[1], and the T1 ribonuclease oligonucleotide composition of subfragment 9 was characterized[2]. The complete nucleotide sequence of the interacting regions is necessary for understanding their mode of association. Indeed, a likely explanation is that there are strong base-pairings which are stable in 4M urea. The primary structure of these regions was completed by the rapid new RNA sequencing technique using 5'- ^{32}p labelling and subsequent fractionation of the digests on polyacrylamide gels[3,4]. On the basis of the sequences of subfragments 4B, 9 and 10A, we present a base-pairing model which explains the interaction observed between these subfragments[2].

MATERIAL AND METHODS

1. Preparation of the L24 ribonucleoprotein complex

10 mg of uniformly ^{32}P labelled 23S RNA were complexed with protein L24 under the conditions described in the preceding paper[2]. The complex was digested with T1 ribonuclease at an enzyme/RNA ratio of 1/10 for 30 min. at 0°C. The digestion products were precipitated with ethanol and fractionated by electrophoresis on polyacrylamide gels[1].

2. Preparation of the RNA subfragments from the ribonucleoprotein complex

The ribonucleoprotein complex was dissociated as previously described by electrophoresis on a polyacrylamide gel containing 8M urea and 0.1 % Na dodecylsulfate[1]. The resulting RNA subfragments were identified by their characteristic electrophoretic mobility[1]. They were eluted from the gel using a modification of the procedure described by Ross and Brimacombe[5] : the gel bands were ground and shaken overnight in a solution containing 0.3 M NaCl and 0.1 % Na dodecylsulfate, in the presence of phenol saturated in the same solution. The aqueous phase was recovered by centrifugation through glass wool. The nucleic acid concentration of the solution was adjusted to 50 µg/ml by addition of calf thymus DNA. This DNA was prepared by sonication of commercial calf thymus DNA, followed by three phenol extractions. The nucleic acids were ethanol -precipitated twice and washed once with ethanol.

When the internal labelling of the RNA subfragments was too high, they were stored until the radioactivity had partially decayed before labelling with polynucleotide kinase.

3. 5'-end labelling of the RNA subfragments

The labelling was done using highly labelled γ- ^{32}P ATP (Amersham 3000 Ci/mmole) and T4-polynucleotide kinase prepared by Dr. G. Keith. A 2-fold molar excess of ATP relative to the 5' end of the RNA fragment was generally used. It was dried under vacuum and resuspended in 25 µl H_2O containing the RNA fragment to be labelled. 25 µl twice concentrated kinase buffer (20 mM $MgCl_2$, 12 mM mercaptoethanol, 20 mM Tris HCl pH 8) and 4 units polynucleotide kinase were added. The mixture was incubated for 30 minutes at 37°C. The RNA was ethanol -precipitated and the pellet dissolved in 20µl H_2O and 25 µl twice concentrated denaturation buffer (40 mM Na citrate pH5; 2 mM EDTA; 0.025% xylene cyanol, 0.025 % bromophenol blue, 10 M urea[3])were added. The solution was incubated at 55°C for 20 mn. This step was required to destabilise the RNA secondary structure and therefore to eliminate subfragments containing hidden breaks. The solution was then loaded onto a polyacrylamide gel containing 8 M urea and Tris-borate EDTA buffer. The purified 5'-labelled fragments were then

eluted as described above. They were ethanol-precipitated with carrier tRNA, three precipitations and careful washing of the pellet with ethanol were required to remove traces of residual phenol which would inhibit the subsequent enzymatic digestions.

4. Partial digestion of 5' -^{32}P labelled RNA

4.1. Partial digestions with both T1 and U2 RNases were carried out using the same reaction mixture and incubation conditions as Donis-Keller et al.[3], with the following changes : an enzyme/RNA ratio between $2X10^{-3}$U/ g and $2X10^{-2}$U/µg was used for T1 RNase (Sankyo). The enzymatic/RNA ratio was 1U/µg in the case of U2 RNase (Sankyo), and incubation was for 30 minutes.

4.2 Partial digestion with PhyI RNase (generous gift of Dr. Bargetzi) was carried out using the same reaction mixture and incubation temperature as Simoncsits et al.[4], except that the enzyme/RNA ratio was $2X10^{-3}$U/µg and incubation for 2 to 3 minutes.

4.3 Partial digestion with RNase A was done using the same conditions as for U2 RNase except that the enzyme/RNA ratio was between 1.5 ng/µg and 4ng/µg.

4.4 Partial alkaline digestion: the conditions used were those described by Donis-Keller et al.[3] except that the incubation was done at 100°C. In addition, it was noticed that two times of incubation (15 minutes and 30 minutes) were required to obtain a good distribution of partial products.

5. Fractionation of the digestion products

The various digests described above were loaded onto a 20 % polyacrylamide slab gel, under the conditions described by Donis-Keller et al.[3]. The length of the gel was 40 cm or 90 cm depending on the size of the RNA subfragment to be analysed.

RESULTS

In order to use the new RNA sequencing techniques, the RNA subfragment were labelled at their 5'-end with γ-^{32}P ATP and T4-polynucleotide kinase. No dephosphorylation step was required since the RNA subfragments resulted from a T1 ribonuclease digest and therefore had a free 5'OH end. The efficiency of labelling was nevertheless variable depending on the subfragment (Fig.1). Small subfragments were more readily labelled than large ones, but among the larger ones, great differences were observed, for example subfragments 3B and 4B were well-labelled, whereas subfragments 3A and 4A were only poorly labelled. The former subfragments have an adenine residue at their 5' ends, while the latter ones have a uridine residue. The specificity of T4-polynucleotide kinase has been reported for 3' monophosphate nucleoside[6]: the Michaelis constant for rUp is higher than that of rAp, whilst the Vmax for rAp is higher

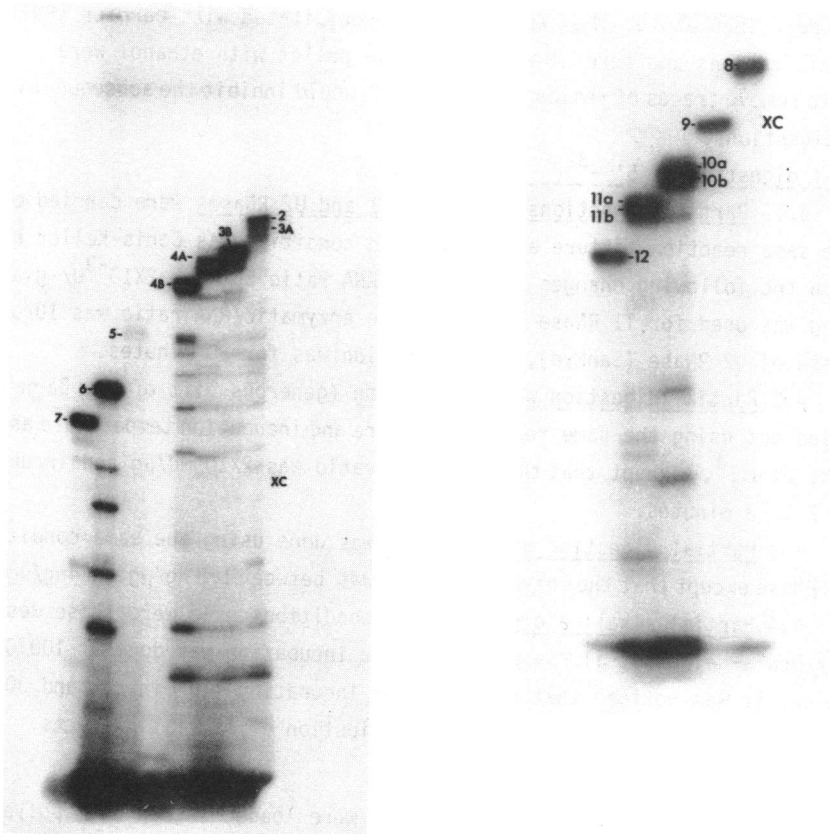

Figure 1 : Acrylamide gel electrophoresis in 7M urea of 5'- ^{32}P labelled RNA subfragments

The RNA subfragments were labelled as described in "Material and Methods". Large subfragments were loaded onto a 12 % polyacrylamide gel (1a) and smaller ones onto a 18 % polyacrylamide gel (1b). A main 5'-labelled fragment appears in each lane except for subfragments 3A and 4A. The minor bands correspond to hidden breaks within the RNA subfragments which have been revealed by the pretreatment at 55°C (see Material and Methods).

than that of rUp. Nevertheless, it seems unlikely that the difference observed only resulted from the specificity of the polynucleotide kinase. It is more likely a question of accessibility of the 5' end of these subframents. From previous results[1,2], subfragments 3A and 4A have a compact structure, which is not the case for subfragments 3B and 4B. Since the use of the new RNA sequencing procedures requires 5'-end labelling, the applicability of the technique will strongly depend upon the conformation of the RNA fragment studied. Surprisingly, subfragments which were present in very low yield and

therefore were not characterized previously have been well-labelled, for example part of subfragment 7A (sub 7'A) and one of subfragment 7B (sub 7'B)[1]. Fortunately, subfragments 4B, 9 and 10A whose sequences are particularly interesting for the study of RNA-RNA interactions within the binding site of protein L24[2] were easily labelled.

We studied subfragments 4B, 6B and its product 10A, subfragments 7'A, 7'B and 9. To sequence them, we adapted the limited digestion conditions for A, T1, U2 and Phy I RNases established by Donis-Keller et al.[3] and Simoncsits et al.[4] T1, U2 and Phy I RNases gave satisfactory results (Fig. 2-4). Digests with RNase A were less clear since this enzyme has a strong specificity for the nucleotide following the pyrimidine residue (7) and is very sensitive to remaining secondary structure.

In spite of non-enzymatic cleavages which occurred in some cases, it was generally possible to deduce the nucleotide at the position of the cut by comparing the intensity of the bands in the various digests. These cleavages, which probably resulted from heating to 50°C in the presence of urea, always occurred after pyrimidine residues preceding an adenine residue, like the bonds broken preferentially by alkali digestion at elevated temperature. Problems of interpretation also arose when the secondary structure of the fragment was not completely destroyed, under the conditions used, in which case some of the bands were weak. Furthermore, we observed in some cases that RNA pieces differing by one or two nucleotides had the same electrophoretic mobility on the gel.

Gel sequencing enabled us to correct some errors made previously using the classical fingerprinting techniques. However our previous results have been very useful in several cases to confirm the present data. The results we obtained are summarized in fig.5 : the nucleotide sequence of the 60 nucleotides at the 5' end of subfragment 4B was completed as well as the sequences of subfragments 6B and 7'B. The sequence of subfragment 9 has been determined, and corrections were made in that of subfragment 7'A. The subfragment 9 is part of the 13S fragment of 23S RNA (1300 nucleotides at the 5' end of the molecule) because the sequence GUACAA was found in it. This sequence is unique to the 13S RNA fragment and comes from the T1 oligonucleotide UACAAG (unpublished results). It is not possible to have a more precise localization of subfragment 9 within the 13S fragment, since the characteristic T1 RNase digestion products of this subfragment contain only one uridine residue and since these oligonucleotides have not yet been localized within the different sections of the 13S fragment[8]. The sequence of the 17 nucleotides

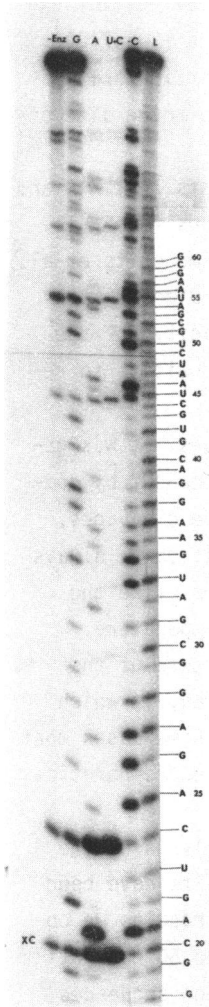

Figure 2: Sequence analysis of subfragment 4B

- Enz corresponds to the fragment incubated without enzyme.

G corresponds to the T1 RNase digest, A to that with U2 RNase, U + C to that with RNase A.

- C to that with Phy I RNase and L to the ladder obtained by alkaline digestion. The digestion conditions are given in "Material and Methods". The digests were loaded onto a 90 cm long polyacrylamide slab gel. The electrophoresis was carried out for a long time: 70 hours at 100 V/cm. The oligonucleotides migrating faster than the xylene cyanol blue were lost, but a very good fractionation of oligonucleotides between 17 and 60 long was obtained. The sequence of the first 17 nucleotides has been deduced using a shorter electrophoresis. Some non-enzymatic cleavages occured at positions 20, 24, 45 and 55 during the various treatments of the subfragemnt. Nevertheless, in each case it was possible to deduce the corresponding nucleotide from the intensity of the bands and from previous data concerning the sequence. RNase A gave very faint bands at certain positions: 23, 42 and 48. In this case the interpretation was facilitated by the previous results.

G and C were sometimes cut by U2 RNase (positions 22 and 30) but it was a very low degree and did not interfere with interpretation of the results. Phy I enzyme did not cleave after C when followed by U, C or G, nor after U, A or G when followed by U. Cleavage after G or A when followed by C was poor. By taking into account these particularities of Phy I cleavage, and by comparison of the different lanes, it was easy to distinghish between U and C.

preceding subfragment 4B and corresponding to the 5' end of 23S RNA has been recently verified and is indicated on Figure 5.

Examination of the sequences in Figure 5 reveals the existence of several complementary regions in subfragments 4B, 9, 10A and 12. The most stable base-paired structures are represented on Figure 6.

Subfragment 12 can be base-paired with subfragment 10A (structure 9) or with subfragment 4B (structure 4). The first structure is very unstable and it is not surprising that it was not found in the presence of 2M urea[2]. On the other hand, the second one could be stable in the presence of 2M urea. Since

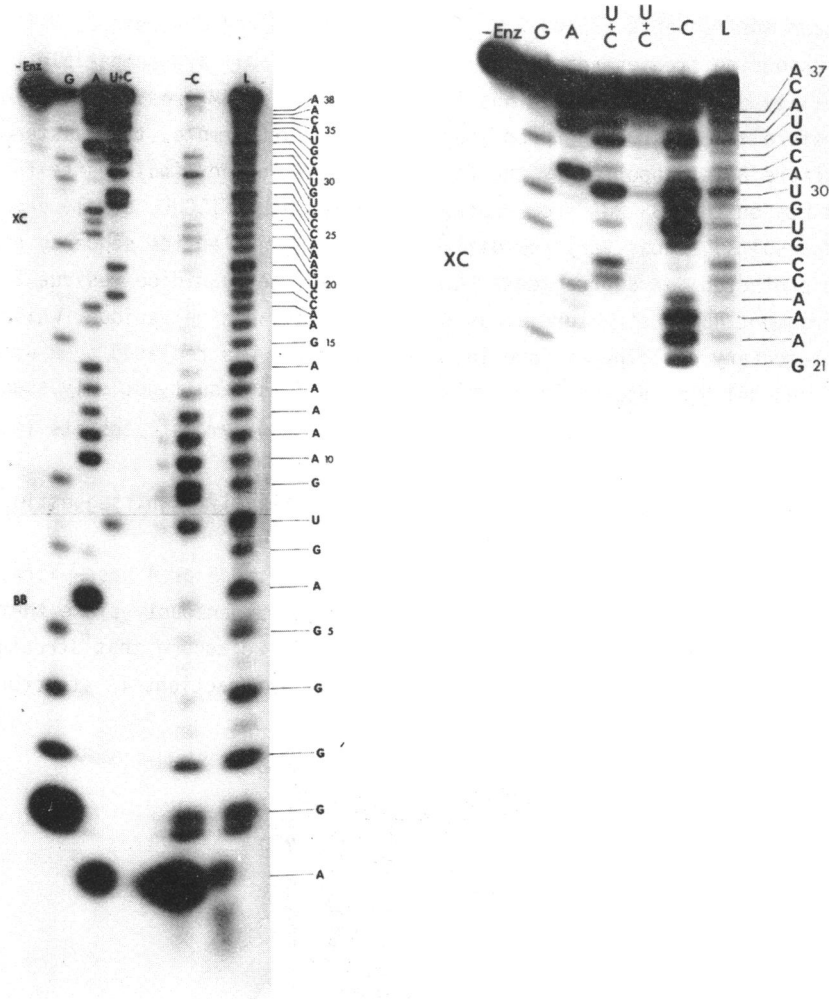

Figure 3 : Sequence analysis of subfragment 9

The digestion products were loaded onto a 90 cm long polyacrylamide gel, and electrophoresis was done for 48 hours at 100 V/cm (a). The nucleotide sequence of the 5 nucleotides at the 3' end has been determined on another gel which was run for a longer time (b).

is was not found[2], it is unlikely to occur. Since subfragment 12 was found to interact with subfragment 4A[2], a base-pairing more stable than in structure 4 may occur between these two subfragments. This cannot be checked at present since the nucleotide sequence of subfragment 4B is not yet completed.

Structures 1 and 2 both correspond to very stable possible associations between subfragments 4B and 9 and subfragments 4B and 10A respectively. The corresponding free energies (-26 Kcal and -18.5 Kcal) are compatible with a stability of these base-pairings in the presence of 4M urea. They could therefore account for the preceding results[2]. Furthermore, the existence of structure 2 is supported by the following result : the binding site for protein L24 on the 23S RNA from *Proteus Vulgaris* (J.L. FISCHEL et al. unpublished results), shows very few differences in the nucleotide sequence of the most protected RNA subfragments. Nevertheless, the guanidine residue 22 of subfragment 4B in structure 2 was replaced by an adenine residue, while the complementary cytidine residue in subfragment 10A was replaced by a uridine residue, which maintains the complementarity. This result not only suggests the existence of this base-pairing scheme but also demonstrates its importance in 23S RNA function.

The free energies of structures 3 and 5 are relatively high, suggesting that the corresponding base-pairing may occur.

Three other interactions, 6, 7 and 8, only involving 3 or 4 base-pairs, are possible. But structures 7 and 8 cannot occur simultaneously since they involve the same sequence from subfragment 9. If we assume that structures 1 and 2 exist, the strains created mean that the interactions in structures 6

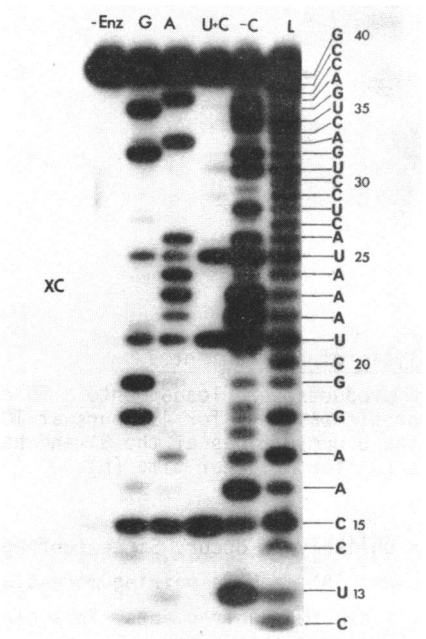

Fig.4a

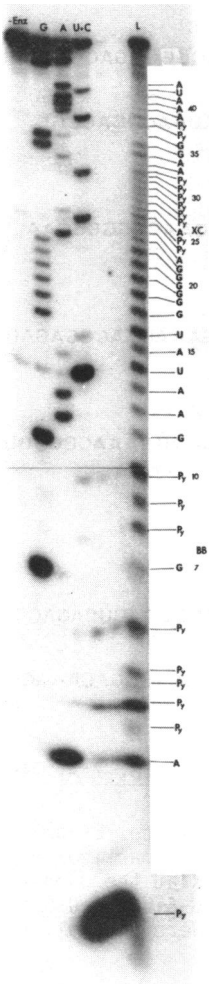

Fig.4b (1)

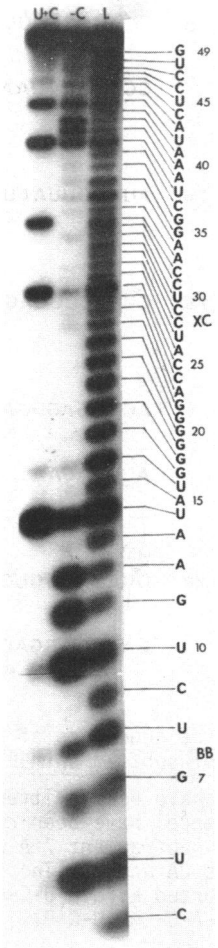

Fig.4b (2)

Figure 4 : Sequence analysis of subfragments 6B and 10A

4a) subfragment 10A :

The digests were loaded onto a 40 cm long polyacrylamide gel. Electrophoresis was carried out for 18 hours at 100 V/cm.

4b) Subfragment 6B

Since Phy I RNase conditions were too strong on the first gel, (1), a second run was made with this enzyme together with pancreatic and alkaline digestions (2).

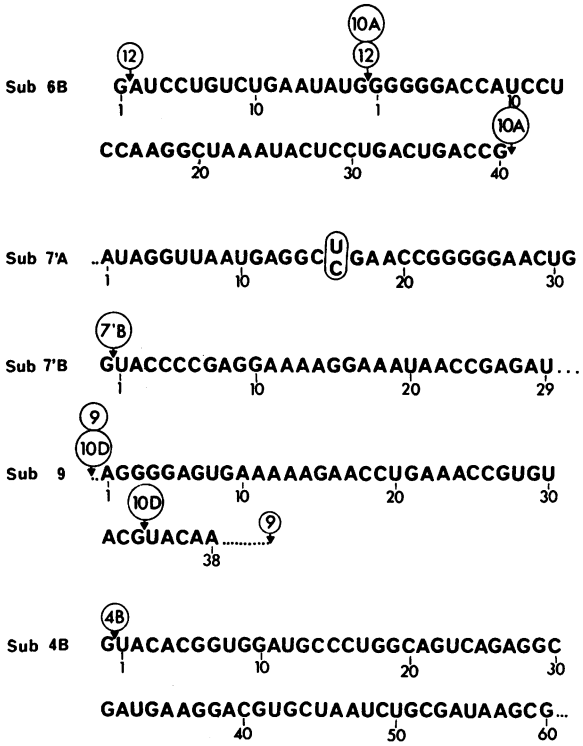

Figure 5 : Nucleotide sequences of subfragments 6B, 7'A, 7'B, 9 and of the 5' end of subfragment 4B

Hyphens have been omitted to save space. Errors made in the previous partial sequence[1] have been corrected. In particular, the sequence A-G-G-C-$\frac{U}{C}$-G within subfragment 7'A was placed at the wrong position at the 5' end of subfragments 2A and 2B. The sequence U-A-U-C-C-U-G in subfragment 6B was previously reported as U-A-U-C-C-C-U-G. Herr and Noller have already found that it is U-A-U-C-C-U-G[15].

and 7 or 8, if they exist, could be somewhat like codon-anticodon interactions. They would participate in the tertiary structure of the molecule. An interesting point is that the sequences from subfragment 9, involved in the double-stranded structures 6 and 8, are close in the primary structure. Those in subfragment 10A involved in the same structures are also close to each other in subfragment 10A. One could imagine that subfragments 9 and 10A are in proximity permitting interactions 6 and 8 to occur. It should be noted that three stretches of adenine residues (underlined in figure 7a), would be in the same vicinity in this model, allowing the possible formation of a triple helix. No stable base-pairing can be proposed for the region between nucleotides 29

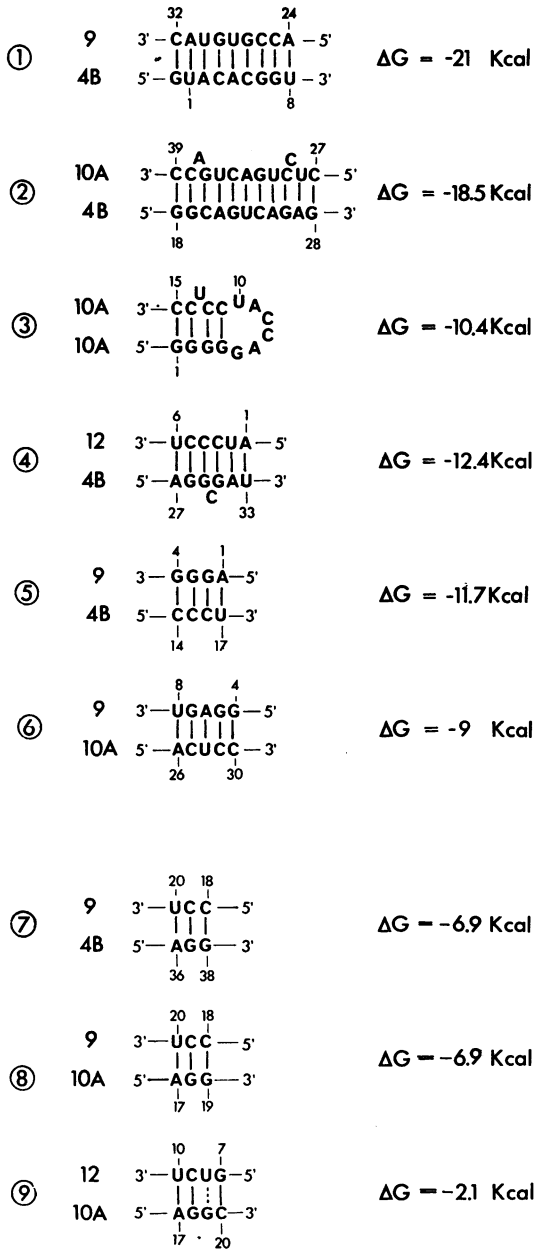

Figure 6 : Complementary sequences within subfragments 9, 10A, 12, and the 5' end of subfragment 4B

The numbering of the nucleotides is the same as in figure 5.

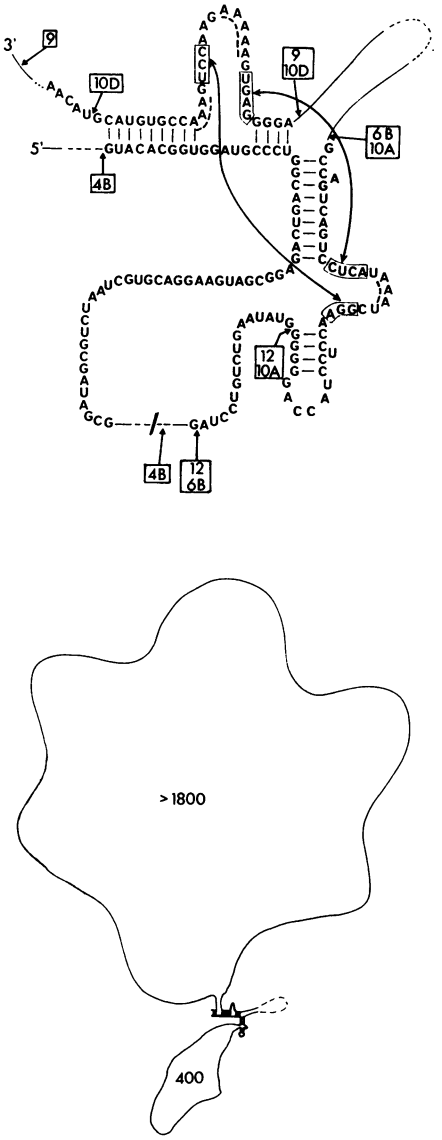

Figure 7: a) Model of interaction between subfragments 4B, 9 and 10A

b) Representation of these possible base-pairings on the scale of 23S RNA

We have represented the base-pairing between the two ends of 23S RNA that we proposed previously[14]. The numbers represent the length of the loops.

and 60 in subfragment 4B. This sequence probably interacts with the 3' half of subfragment 4B, whose sequence is not yet completed.
In figure 7a, a model is proposed to explain the interaction between subfragments 4B, 9 and 10A.

DISCUSSION

The RNA sequencing techniques developed by Donis-Keller et al.[3] and Simoncsits et al.[4], were applied to the RNA subfragments present in the L24 binding site. The first step of these methods consists in the 5' end labelling of the RNA fragment. RNA subfragments with 5' ends involved in base-pairing were found to be labelled with low efficiency. Fortunately the subfragments in which we were interested displayed a low degree of internal base-pairing and were well-labelled. Use of these new techniques enabled us to complete the partial sequence we previously obtained using the fingerprinting technique of Sanger et al.[9].

In addition, gel sequencing allowed us to correct errors made previously with the classical fingerprinting technique. Three types of errors were difficult to avoid using the fingerprinting technique : i) errors in the positioning of the T1 RNase digestion products C-G and U-G (subfragment 4B), ii) errors in the ordering of uridine and cytidine residues in long stretches of pyrimidines (subfragment 10A), iii) errors in the estimation of the number of guanine residues within large pancreatic RNase digestion products (subfragments 7'A and 10A). In addition, two inversions were made in the ordering of T1 RNase digestion products because they were preceded by the same pancreatic RNase digestion product. The sequence determination of a molecule of 3000 nucleotides is about the limit compatible with the use of the fingerprinting technique, but a combination of the results obtained previously and of those obtained by the rapid new gel sequencing technique allowed us to obtain results which are far more reliable. Indeed, it must be pointed out that the new methodology for RNA sequencing in its present state is not accurate enough to be used alone. It is more and more obvious that nucleic acid sequences should be determined using two parallel approaches. Even in the case of DNA, the determination of the sequences of the two strands is required[10].
Nevertheless, the results of the fingerprinting technique were required for the determination of the relative order of the RNA subfragments, which is a prerequisite for determination of the complete nucleotide sequence of the RNA region[1]. And this technique is still quite useful to rapidly identify RNA subfragments as was done in the preceding paper[2].

Figure 7a represents the model of secondary structure that can be proposed when one looks for maximum base-pairing within the sequences of subfragments 9 and 10A and of the 5' half of subfragment 4B. This model is in agreement with the following experimental data : the electrophoretic mobility of subfragment 4B, its sensitivity to T1 RNase, its accessibility to T4-polynucleotide kinase, imply that this subfragment has a low degree of internal secondary structure. But it should be base-paired to other 23S RNA fragments since it is very resistant to T1 RNase in the L24-23S RNA complex[1] as well as in isolated 23S RNA[11]. Finally base-pairing between subfragments 9, 10A and 4B is very likely because of the strong interactions between these subfragments[2].

The incompletely sequenced 3' part of subfragment 4B could also have sequences complementary to subfragments 9 and 10A, but the probability of the existence of more stable structures than those in the model is very low. Furthermore, the positions of T1 RNase cleavages (fig. 7a) fit the proposed model and the observation that the complementarity between subfragments 4B and 10A is preserved during evolution is a good argument for the proposed base-pairing scheme.

In our model, sequences widely separated in the molecule are base-paired : the interacting regions in subfragments 4B and 10A are separated by 400 nucleotides and those of subfragments 4B and 9 by many more. Such long-range interactions have also been proposed in the 5' end region of 16S RNA[12,13]. The synthetic S4-16S RNA complex, seen by electron microscopy by Cole et al.[13], is strikingly similar to our model. Figure 7b represents the interactions reported in this paper within total 23S RNA.

We have no experimental data concerning the possible less stable interactions we propose between subfragments 9 and 10A. The conservation of these interactions in the structure of 23S RNA from different bacterial species would be the best proof of their existence. This study reveals once more how difficult it is to determine the secondary and tertiary structures of an RNA molecule even when its primary structure is known.

The model we propose results from a study on the reconstituted L24-23S RNA complex, and so one could object that 23S RNA may have a different conformation within the 50S subunit. However, previous experiments[8] have shown that the 5' end region in 23S RNA within the 50S subunit has a very compact conformation, which would be in agreement with the present data.

ACKNOWLEDGMENTS

We are indebted to Dr. G. Keith for his generous gift of T4-polynucleotide kinase, to Dr.J.P. Bargetzi for providing us with Phy I RNase and to Profes-

sor H.G. Wittmann for his generous gift of protein L24. The skilled technical assistance of Miss B. Muller is acknowledged. This work was financially supported by the "Centre National de la Recherche Scientifique", the "Délégation Générale à la Recherche Scientifique et Technique" and the "Commissariat à l'Energie Atomique".

REFERENCES

1. BRANLANT, C., SRI WIDADA, J., KROL, A. and EBEL, J.P. (1977) Eur. J. Biochem., 74, 155-170

2. SLOOF, P., HUNTER, J., GARRETT, R. and BRANLANT, C. (1978) Nucl. Acid. Res. 5,3503-3514

3. DONIS-KELLER, H., MAXAM, A. and GILBERT, W. (1977) Nucl. Acid Res.,4, 2527-2538

4. SIMONCSITS, A., BROWNLEE, G.G., BROWN, R., RUBIN, J. and GUILLEY, H. (1977) Nature, 269, 833-836

5. ROSS, A. and BRIMACOMBE, R. (1978) Nucl. Acid. Res., 5, 241-256

6. LILLEHAUG, J. and KLEPPE, K. (1975) Biochemistry, 14, 1221-1229

7. WITZEL, H. (1963) Progr. Nucl. Acid. Res., 2, 221-258

8. BRANLANT, C., SRI WIDADA, J., KROL, A. and EBEL, J.P., (1977) Nucl. Acid. Res., 4, 4323-4345

9. SANGER, F., BROWNLEE, G.G. and BARRELL, B.G. (1965) J. Mol. Biol., 13, 373-398

10. Mc REYNOLDS, L., O'MALLEY, A., NISBET, A., FORTHERGILL, J., GIVOL, D., FIELDS, S., ROBERTSON, M. and BROWNLEE, G.G. (1978) Nature, 273, 723-728

11. BRANLANT, C., SRI WIDADA, J., KROL, A., FELLNER, P. and EBEL, J.P. (1975) Biochimie, 57, 175-225

12. UNGEWICKELL, E., EHRESMANN, C., STIEGLER, P. and GARRETT, R. (1975) Nucl. Acid. Res., 2, 1867-1888

13. COLE, M., BEER, M., KOLLER, T., STRYCHARZ, W. and NOMURA, M. (1978) Proc. Natl. Acad. Sci. USA, 75, 270-274

14. BRANLANT, C., SRI WIDADA, J., KROL, A. and EBEL, J.P. (1976) Nucl. Acid. Res., 3, 1671-1687

15. HERR, W. and NOLLER, H. (1978) Biochemistry, 17, 307-315