ON THE TERTIARY STRUCTURE OF TRANSFER RIBONUCLEIC ACID

Gerhard MELCHER

Service de Biochimie, Département de Biologie, Centre d'Etudes Nucléaires de Saclay, de Saclay, B.P. no. 2, 91, Gif-sur-Yvette, France, and Institut für Mikrobiologie und Biochemische Technologie der Technischen Hochschule, Getreidemarkt 9, Wien 6, Austria

Received 10 April 1969

1. Introduction

Many complete sequences of *t*-RNAs * have been published and a great number of models have been proposed [1–3] all of which have proven unsatisfactory for one or the other reason. The clover leaf [1] which describes the arrangement of the bases to each other in the plane will probably approach the structure which actually occurs in Nature.

Most physical measurements [2–4] indicate a very compact structure for *t*-RNA with a Stoke's radius of about 30 Å. Since the odd bases are formed only after the folding of the tertiary structure of the molecule they have to be on exposed non-helical sites in order to be approachable by the specific enzymes like methylases [5]. A model of the tertiary structure of *t*-RNA must take into account these facts and has to explain why selective chemical reactions specific for one base only [3] show only limited reaction with the bases in the structured molecule. In addition such a model has to explain the transition between *t*-RNA and aa-*t*-RNA [6], the differential degradation of *t*-RNA by polynucleotide phosphorylase [7] which implies two conformational states with different structure of the CCA-end, and the action of nucleases used for the elucidation of their sequences [1–3] and which can attack only at specific points which must be on the surface of the molecule. Finally, the role of *t*-RNA in protein synthesis has to be accounted for.

In the present communication a three dimensional model is developed which fulfills all the above mentioned constraints.

* Abbreviations used: A: adenine, G: guanine, U: uracil, C: cytosine, Ψ: pseudouridine, T: thymine, MG: methylguanine, M_2G: dimethylguanine, H_2U: dihydrouracil, MC: methylcytosine, *t*-RNA: transfer ribonucleic acid, aa-*t*-RNA: aminoacyl-*t*-RNA. Numering of the bases (superscript) starts from the 5′-phosphate end.

2. The model

The well known cloverleaf model consists of four or five helical regions and three loops. Fig. 1 shows the sequences for $t\text{-RNA}^{tyr}_{yeast}$ and $t\text{-RNA}^{val}_{yeast}$ [1]. In the present model the three loops are piled upon each other (fig. 2) and form a ring. The helical regions too are one above the other and are slightly turned (fig.3).

2.1. *Structure of the helical regions*

a) The helical region consists of two parellel stretched chains (fig. 3a). b) In fig. 3b the chains are also in one plane but the inner chain is slightly bent by folding the ribosephosphate chain in a zig-zag manner, with the phosphates on the extreme points. c) In the third case the inner chain is even more folded (fig. 3c). The riboses are piled upon each other and their hydrogen atoms can interlace and thus give more stability to the structure. The phosphates are here in such close contact, however, that the presence of bivalent metal ions would be desirable, perhaps even necessary to eliminate electrostatic repulsion [8].

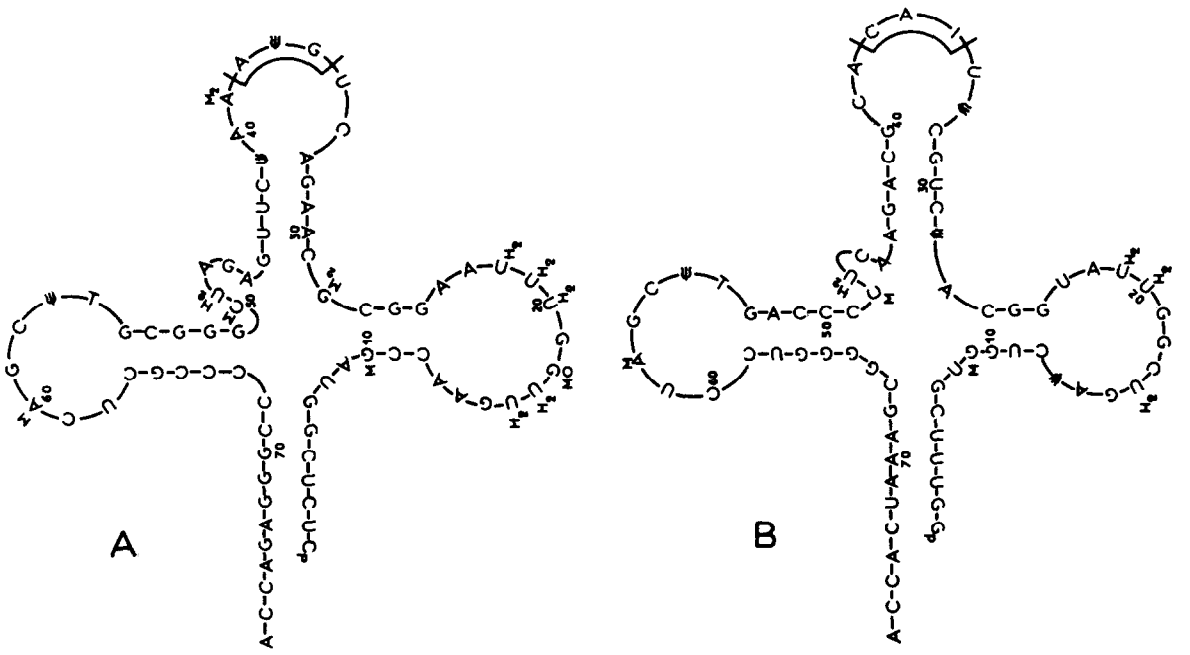

Fig. 1. Cloverleaf models of a) $t\text{-RNA}^{tyr}_{yeast}$ and b) $t\text{-RNA}^{val}_{yeast}$ [1].

2.2. *Structure of the loops*

The structure of the ribosephosphate chain is close to that found in the A-form of DNA. As shown in fig. 4 the bases in the loops are turned inwards. The piling up of the three loops can lead to the following interactions: a) The bases remain in the arrangement of fig. 4. b) The bases form hydrogen bonds between the loops. c) The basic amino groups and nitrogens of the bases form ionic bonds with the phosphate groups of the neighbouring chains. All three possibilities can and shall exist, as shall be shown by the existence of certain odd bases on specific sites.

2.3. *Structure of t-RNA*

The folding of the tertiary structure of *t*-RNA shall

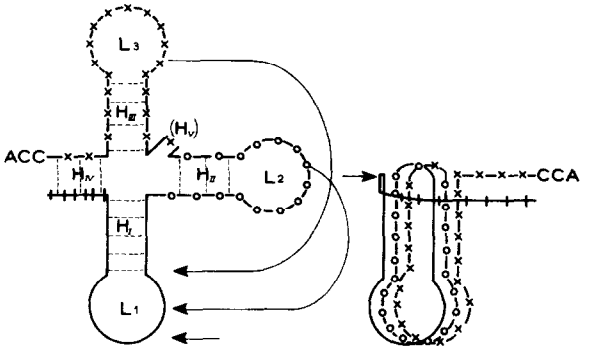

Fig. 2. Folding scheme of the clover leaf to form the proposed structure. The helical regions are numbered with roman numbers (H-I, etc.) and the loops with arabic numbers (L-1, etc.) and correspond to those in the text.

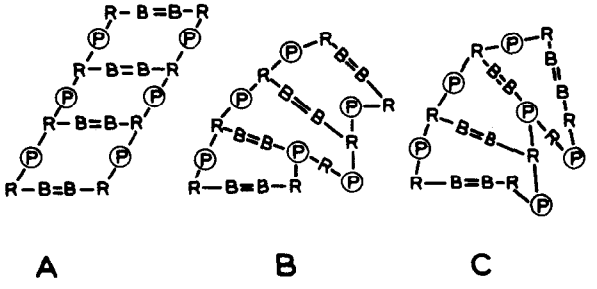

Fig. 3. Three possible arrangements of the ribose phosphate chains of a helical region (see text).

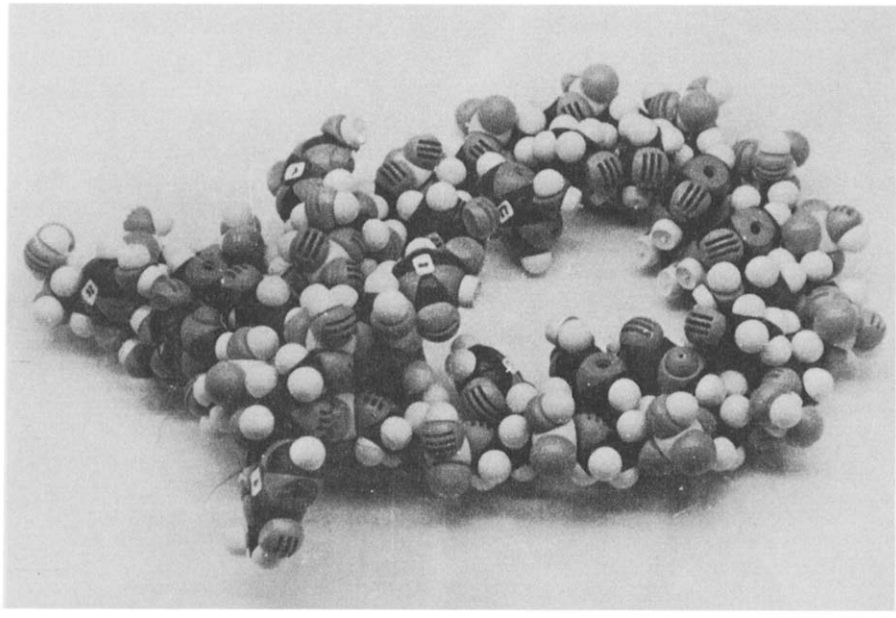

Fig. 4. Helical region (H-I) and first loop (L-1) of $t\text{-RNA}^{val}_{yeast}$.

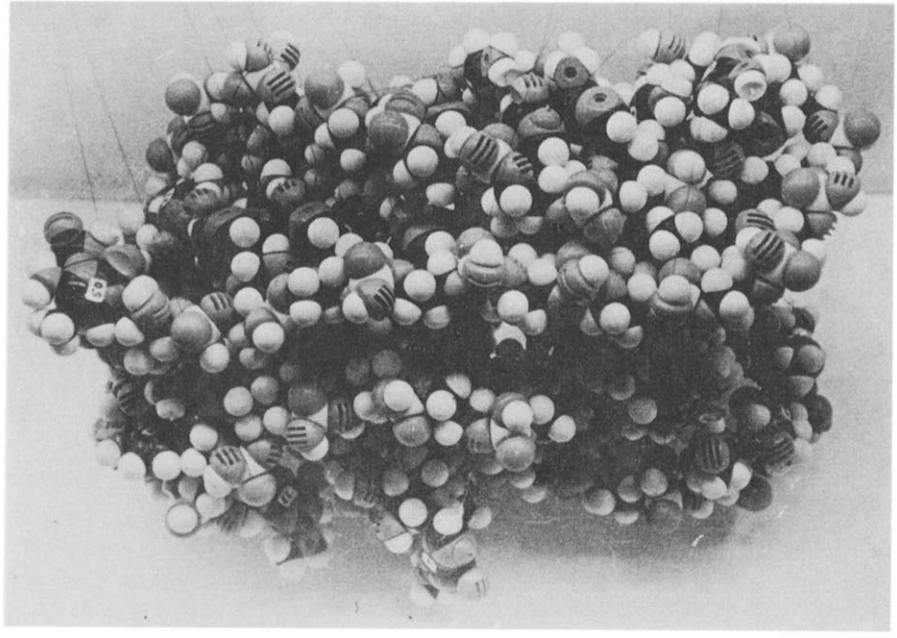

Fig. 5. Complete structure of $t\text{-RNA}^{tyr}_{yeast}$ built with CPK-models.

follow the above mentioned rules. All presently known sequences [1] follow the same principle. Fig. 4 shows bases 7 to 26 of t-RNA$^{val}_{yeast}$, (H-I and L-1). The helical region in the first plane is highly folded (helix form, fig. 3c) (not visible because covered by the ribose of MG^9) so that the chain in the back is forced to turn around. The immediately preceeding bases 7, 8 and 9 run over the first helical region backwards. The segment is quite important since it contains the generally methylated base no.9 (or no.10 in other t-RNAs). From the model in fig. 4 one can see that base no.9 has to be methylated since it will be pushed outwards because of the bent of the chain. The zig-zag structure of H-I would permit that the amino group of MG^9 could form an ionic bond with a phosphate residue. This would imply that a methylase could attack only from one direction in order to avoid further methylation. It is important that U^8 is situated just above H-I and directed into the loop. This structural particularity was found in all the eight structures built. Base no.7 is already hidden by the helical region and with bases nos.6 to 1 can pair with the corresponding bases next to the CCA-end to form H-IV which in turn imposes the structure described upon the bases 7 to 12.

Fig. 5 shows the complete structure of t-RNA$^{tyr}_{yeast}$ (fig. 1a). Base no.26 (on the far left of the picture in fig. 4) is pushed forward in the complete structure and with the next bases will make the connection with L-2 which is sitting on top of L-1. M_2G^{28} is sticking out without possibility of an ionic bond with a phosphate (in contrast to MG^{10}). This base is therefore dimethylated. (In several other sequences this base is already in H-II and unmethylated.)

H-II starts with C^{29} (just above M_2G^{28} in fig. 5). H-II can be straight or a slightly folded structure (fig. 3a or b). L-2 contains the anticodon bases G^{36}-Ψ^{37}-A^{38}. They are on the back right of the molecule and are covered by L-3. The chain returns via H-II (on the back of the molecule) to the left. The bases A^{46}-G^{47}-A^{48} (fig. 1a) are partly covered by H-IV, while the two following bases are pushed outwards (fig. 5) and are modified: H_2U^{49}-MC^{50}.

L-3 is sitting on top of L-2 and contains the typical sequence G-T-Ψ-C-G [1]. H-III (again as in fig. 3a or 3b) is above H-II and is slightly tilted backwards so that the last bases approach the bases 7 to 1. This latter chain comes out of the back of the molecule and will form H-IV with the last bases of the sequence (CCA-end). H-IV can be seen in fig. 5 to run from the left back to the right front of the molecule. It thus brings the CCA-end (unbound) into the vicinity of the ring of loops. H-IV has the additional effect of tying the structure together.

After a detailed study of the model the following catalytic mechanism for the sequence G-T-Ψ-C-G could be proposed: the terminal A^{78} can form hydrogen bonds with T^{56} and in a successive step with Ψ^{57}. The terminal A^{78} will have to switch from *anti* to *syn*. Thus the amino group of C^{58} can bind with the phosphate of A^{78} breaking the Ψ^{57}:A^{78} bond. In the next step G^{59} can replace C^{58}. By these means the CCA-end slips into the ring formed by the three loops since the terminal A^{78} in *syn* conformation will take much less space than in the *anti* conformation and can now bind perfectly well with U^8 which is on the bottom of the molecule. This pairing will in turn keep the 3′-OH of the terminal adenosine in close vicinity of the anticodon bases which are sticking into the ring in the middle loop. The aminoacylation can be pictured in such a manner that the activated amino acid will be pushed by the aminoacyl synthetase into the ring from the bottom and the aminoacylation will take place in presence of the anticodon bases [8]. Aminoacylation would dissociate the CCA-aa end which can thus move out of the ring. The rather weak H-IV (always contains a G:U pair) will break up if the CCA-aa end is now freely moving around [9]. The middle loop can now switch over that the anticodon bases are on the surface of the molecule. This form can now serve as the active aa-t-RNA in protein synthesis.

It should be pointed out that all the prerequisites asked to be fulfilled by an operative model of t-RNA discussed in the introduction fit with the structure proposed. Also several other physico-chemical findings, like helical content of the t-RNA and aa-t-RNA [2, 3] are compatible with the present model. In cases where an additional helical region H-V exists [1], this comes to lay between H-II and H-III. The U at the end of this branch sticks into the ring just above U^8. In these cases the terminal A has two binding possibilities in the ring. It should be mentioned that formylmethionyl-t-RNA [10] contains a large hole in the surface of the molecule.

Acknowledgements

The model building studies were performed in Saclay, in spring 1968 during the tenure of a special fellowship of the Commissariat à l'Energie Atomique. The author is particularly grateful to Dr. P.Fromageot for having made possible this stay in his laboratory. The author also acknowledges the extensive discussions and help with the model building of Dr.W. Guschlbauer (Saclay). Professor R.Brunner, Dr. G. Högenauer and Dr. M.Röhr (Vienna) gave much encouragement to this work.

References

[1] FEBS-Symposium on Structure and Function of *t*-RNA and 5*S* RNA, eds. Laland and Fröholm (Academic Press, New York and Universitetsforlaget, Oslo, 1968).
[2] G.L.Brown, Progr. Nucleic Acid Res. 2 (1963) 259.
[3] K.I.Miura, Progr. Nucleic Acid Res. 6 (1967) 39.
[4] A.Adams, T.Lindahl and J.R.Fresco, Proc. Natl. Acad. Sci. U.S. 57 (1967) 1684.
[5] B.C.Baguley and M.Staehelin, Biochem. 7 (1968) 45.
[6] P.S.Sarin and P.C.Zamecnik, Biochem. Biophys. Res. Comm. 20 (1965) 400.
[7] M.N.Thang, W.Guschlbauer, H.G.Zachau and M.Grunberg-Manago, J. Mol. Biol. 26 (1967) 403.
[8] G.Melcher, manuscript in preparation.
[9] S.Englander and J.J.Englander, Proc. Natl. Acad. Sci. U.S. 51 (1964) 570.
[10] S.K.Dube, K.A.Marcker, B.F.C.Clark and S.Cory, Nature 218 (1968) 232.