Proc. Nat. Acad. Sci. USA
Vol. 70, No. 7, pp. 2042–2045, July 1973

Determination of Secondary and Tertiary Structural Features of Transfer RNA Molecules in Solution by Nuclear Magnetic Resonance

(cloverleaf model/hydrogen bonding)

R. G. SHULMAN*, C. W. HILBERS*†, YENG P. WONG‡, K. LIM WONG‡, DONALD R. LIGHTFOOT§, BRIAN R. REID§, AND DAVID R. KEARNS‡

*Bell Laboratories, Murray Hill, New Jersey 07974; and ‡Departments of Chemistry§ and Biochemistry, University of California, Riverside, Calif. 92502

Communicated by Alexander Rich, April 20, 1973

ABSTRACT **High-resolution 300-MHz proton nuclear magnetic resonance spectra of the hydrogen-bonded protons in three different purified tRNA molecules are presented. The resonances in the region between −11 and −15 ppm from 2,2-dimethyl-2-silapentane-5-sulfonate (DSS) are assigned to the ring NH protons of specific base pairs by two approaches. First, intrinsic positions of −14.8 ppm and −13.7 ppm are taken for the AU and GC ring NH protons, respectively, and the spectra are calculated by including ring current shifts from the nearest neighbors. The spectra calculated in this way on the basis of the cloverleaf are in good agreement with the observed. Second, fragments of yeast tRNA^{Phe} were obtained, which helped in assignments of the spectrum of intact molecules. The close agreement strongly supports the cloverleaf model.**

Tertiary structural features were determined in a few cases where the ring currents at the terminal base pairs of helical regions depended upon stacking of the helices. In this way, we were able to show that in *Escherichia coli* tRNA^{Glu} the CCA stem forms a continuous helix with the TψC stem, which is in accord with the preliminary x-ray structure of yeast tRNA^{Phe}, suggesting that this stacking is observed in solution and may be a general property of different tRNA molecules. Similar reasoning suggests that in *E. coli* tRNA^{fMet} G-27 is stacked upon the dihydrouridine helix.

In order to understand in detail how tRNA functions in protein synthesis, the structure of tRNA in solution must be determined. The present note summarizes results of our recent high-resolution proton nuclear magnetic resonance (NMR) investigations (1–4) of tRNA in which we determined that the cloverleaf model is an accurate description of the hydrogen-bonded secondary structure of several purified tRNAs, including yeast tRNA^{Phe}, *Escherichia coli* tRNA^{Glu}, and *E. coli* tRNA^{fMet}. In addition, we obtained information about certain important tertiary structural features of each molecule. In this paper we emphasize these structural aspects to show how the NMR measurements in solution complement and supplement the crystal structure presently emerging from the x-ray studies (5).

The groundwork for the present study was developed in previous studies of the NMR spectra of tRNA molecules in solution where we showed that resonances observed in the region between −11 and −15 ppm from 2,2-dimethyl-2-silapentane-5-sulfonate (DSS) (referred to as the low-field region) could be ascribed to hydrogen-bonded ring NH protons in Watson–Crick base pairs (1–4). Since each base pair contains only one ring NH proton (i.e., U_3H or G_1H), each proton resonance detected in the low-field region corresponds to one hydrogen-bonded base pair in the molecule. From studies of various model systems, as well as tRNA molecules and their fragments as discussed below, we have deduced that the standard positions are $(AU)^\circ = -14.8$ ppm, $(GC)^\circ = -13.7$ ppm, and $(A\psi)^\circ = -13.5$ ppm. The major factor responsible for shifting the low-field AU and GC resonances from their standard positions appears to be ring current fields from adjacent bases (3). Since these shifts depend upon the positions of the nearest neighbors, if the secondary structure of a polynucleotide were known, to predict spectra one only needs the standard positions of the resonances for AU and GC base pairs and the spatial dependence of the ring current shift values of the four bases. The ring current shifts were based partly on calculations by

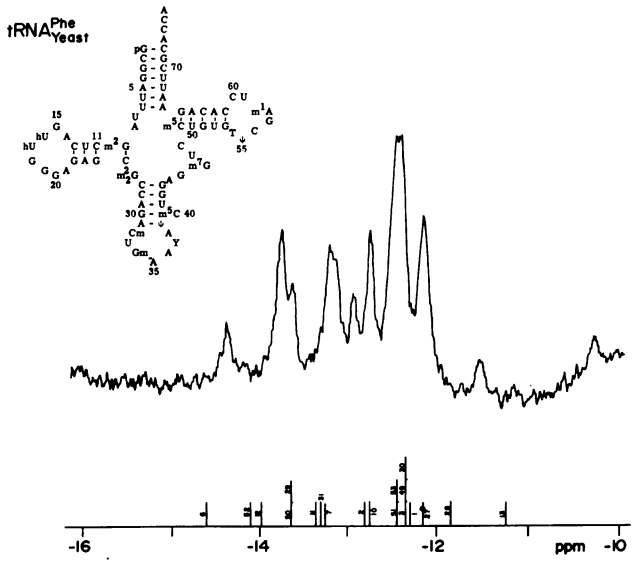

FIG. 1. The 300-MHz proton NMR spectrum of yeast tRNA^{Phe} at pH 7.0, in 0.01 Mg^{2+}–0.1 M NaCl–[tRNA] about 50 mg/ml, and temperature about 35°. The cloverleaf form of the primary sequence (7) is shown as well as the calculated spectrum and its assignments to particular base pairs (3).

Abbreviations: NMR, nuclear magnetic resonance; DSS, 2,2-dimethyl-2-silapentane-5-sulfonate.

† On leave from the Katholieke Universiteit, Nijmegen, The Netherlands.

Giessner–Prettre and Pullman (6) and partly on a preliminary best fit to all the available NMR spectra upon the assumption of regular helical regions. Consequently, we have a set of intrinsic positions and ring current shifts needed to predict the spectra of the hydrogen-bonded ring NH protons. We now discuss the assignment of the three different tRNA spectra mentioned above, showing what the NMR spectra tell us about the secondary and tertiary structure of these molecules in solution.

RESULTS

Yeast $tRNA^{Phe}$

We computed the positions expected of all hydrogen-bonded ring NH proton resonances in yeast $tRNA^{Phe}$, assuming the cloverleaf model for the known sequence (7). These results are compared with the observed spectrum in Fig. 1. Although the calculated and observed spectra agree rather well, there were uncertainties in some assignments. In order to help resolve these uncertainties, we prepared and examined the NMR spectra of different fragments of $tRNA^{Phe}$. The spectrum of the 5′ half shown in Fig. 2 consists of three sharp resonances and one very broad resonance extending from about −12.5 ppm to −11.5 ppm. The positions of the three sharp resonances agree very well with the ring current calculations for GC-10, GC-11, and UA-12, but the broad resonance arising from GC-13 is anomalous. This resonance is predicted to be at −11.5 ppm, providing A-14 is stacked upon C-13 as in a regular helix so that it exerts its full ring current shift of +1.3 ppm. However, if A-14 is not stacked upon C-13, then its ring current shift of the GC-13 ring NH could vanish. Since the broad GC-13 resonance did not become narrow at lower temperatures, we suggest that its breadth is due to very slow interconversion between several conformations of the dihydro-U loop that produce a set of different environments for GC-13. The behavior of this resonance illustrates that resonances associated with terminal base pairs of a helix may be sensitive to the conformation of neighboring bases beyond that helical region. As will be seen below, this is one way in which the NMR spectra can provide information about the tertiary structure of tRNA in solution.

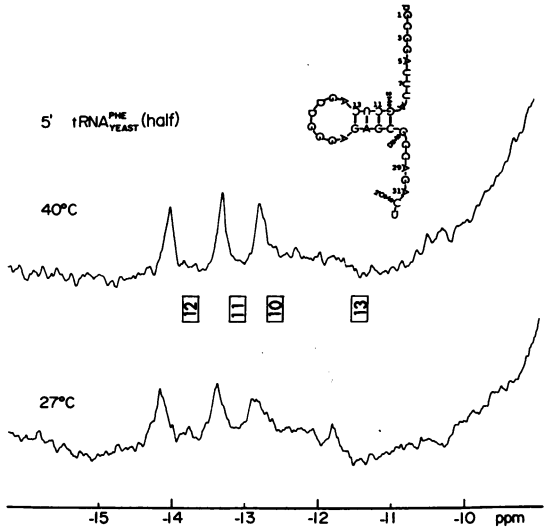

FIG. 2. Proton NMR spectrum at 300 MHz of the 5′ half of a molecule of yeast $tRNA^{Phe}$ (see *insert*) at 40° and 27°. The calculated positions of resonances associated with specific base pairs are indicated by the numbered boxes.

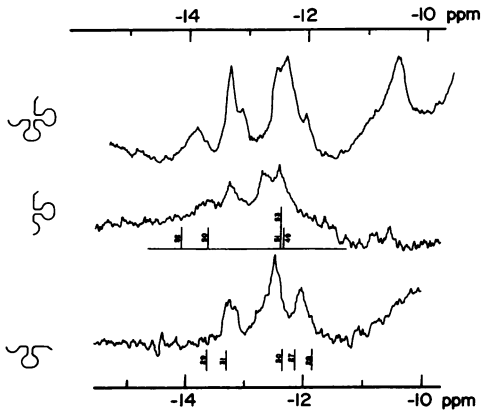

FIG. 3. Proton NMR spectrum at 300 MHz of the 3′ three-quarters of a molecule of yeast $tRNA^{Phe}$ (*top*), the 3′ half of a molecule (*middle*), and the anticodon half molecule (*bottom*), along with the predicted spectra for the two halves.

The spectrum of the anticodon half-molecule (Fig. 3) at about 50° contains three peaks with relative areas of 1:2:1, while on cooling an additional resonance appears around −13.3 ppm. The temperature-dependent resonance is assigned to Aψ-31 because of its location at the end of the helix and because the other resonance positions agree so well with the calculations. It is from this assignment that an intrinsic position of −13.5 ppm was obtained for an Aψ pair.

The spectrum of the 3′ half is not as well resolved, but it is evident from internal and external intensity comparisons and comparison with the calculated spectrum that between four and six resonances are present, in agreement with the five expected on the basis of the primary sequence of this fragment.

The NMR spectrum of the 3′ three-quarters of a molecule is shown at the top of Fig. 3; allowing for small shifts (usually less than 0.2 ppm), it is the sum of the NMR spectra of its component parts, i.e., the anticodon hairpin and the 3′ half. The spectrum of the 3′ three-quarter fragment should have an intensity corresponding to 10 protons. Assuming that the resonances near −13.2 ppm correspond to three protons, then the peak at −13.9 ppm corresponds to one proton and the resonances around −12.3 ppm correspond to six protons for a total integrated intensity of 10. The resonances between −10 and −11 ppm, which appear in the 3′ three-quarter molecule and the intact tRNA, have not been definitely assigned, but similar peaks are observed in the intact tRNAs; they probably are due to slowly exchanging, nonhydrogen-bonded, ring NH protons of inaccessible bases.

The spectrum of the intact molecule should include all the resonances seen in the fragments as well as those from the aminoacid acceptor stem. It is assumed that the GU-4 base pair does not contribute a resonance in the low-field region since the previous evidence supporting GU base pairing (8) is in doubt (9). Therefore, we expect from the cloverleaf model a total of 20 resonances in the low-field region of the spectrum of $tRNA^{Phe}$. However, because of the GU mismatch the helix might be destabilized in this region, and an analysis (4) of the NMR spectrum suggested that possibly the neighboring base pair AU-5 is opened, which would give 19 base pairs. 19 Resonances would be consistent with the observa-

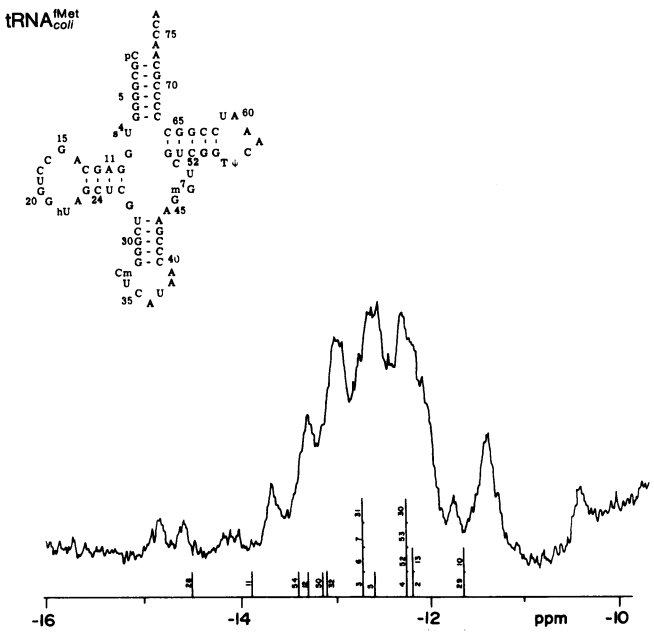

FIG. 4. 300-MHz spectrum of *E. coli* tRNA^{fMet} under the same conditions as Fig. 1.

tions; they were assigned as shown in Fig. 1. Several conclusions follow: First, spectra computed from the cloverleaf model fit the observed spectra both in total number of base pairs and in their dependence upon the primary structure; second, there is better agreement between the calculated positions and with those observed in the fragments than with those observed in the intact molecule, suggesting perhaps that there are larger deviations in the latter from the regular A′ RNA helices assumed for the calculations.

The NMR spectra provide *no* evidence for additional tertiary-structure Watson–Crick base pairs in yeast tRNA^{Phe} beyond the cloverleaf, as reported previously (2). While one additional GC pair could be hidden in our spectra, providing there were compensating losses from the cloverleaf, there is no support for those models of tertiary structure that predict several additional Watson–Crick pairs (10). Similar and supplementary information is obtained from the other two tRNA molecules.

E. coli tRNA^{fMet}

The spectrum of tRNA^{fMet} is shown in Fig. 4 along with the line positions calculated from the primary sequence given by S. K. Dube *et al.* (11). This molecule contains only two AU pairs in the cloverleaf structure, and these are predicted to give resonances at −14.0 (AU-11) and −14.6 to −13.4 (AU-28), depending upon the orientation of A-45 relative to AU-28. Several resonances are found at very low field between −15 and −14 ppm with a total integrated intensity of one proton, and these are assigned to AU-28. By analogy with the behavior of GC-13 in the dihydro-U loop of yeast tRNA^{Phe}, we suggest that the splitting of AU-28 is caused by different orientations of the nonhelical region, probably by A-45 taking different conformations relative to the anticodon helix.

Because of the high GC content, most of the resonances are grouped together in the region around −12.7 ppm, and while the calculated lines account well for the intensity distribution, it is difficult to be confident of the assignment of any particular resonance. However, the isolated peak of intensity two at −11.5 ppm can be assigned to GC-10 and GC-29, since both of these are calculated to occur at −11.7 ppm and no other GC pairs in the molecule can be shifted so far upfield. Since GC-29 is subjected to an upfield shift by an A residue and a G residue, *the identical shifts suggest that the nonpaired G-27 is stacked above GC-10.* We had previously done an approximate integration of the spectrum of tRNA^{fMet} and found 22 ± 3 protons. From the 300-MHz spectrum (Fig. 4), upon the assumption that the peak at −11.5 ppm corresponds to two protons, we obtain a value of 18 ± 2 protons, which agrees very well with the 19 expected from the cloverleaf model (neglecting the GU-51 pair) and also is consistent with the previous determination.

E. coli tRNA^{Glu}

The 300-MHz spectrum of the ring NH hydrogen-bonded protons of *E. coli* tRNA^{Glu} is presented in Fig. 5 along with the predicted positions of the resonances based on the sequence determined by Munninger and Chang (12). The resonances in the region from −15 to −13.3 ppm with an integrated relative intensity of four protons correspond to the four AU base pairs, two of which are predicted to occur at −14.5 ppm and two at −13.5 ppm. The remaining 16 GC resonances are predicted to occur in the region between −13.3 ppm and −12.0 ppm, and the intensity in this region does in fact correspond to 16 protons on the assumption that the four lower-field resonances correspond to a total of four protons. Thus, once again the NMR spectra in the −12 to −15 ppm region correspond almost exactly to the total number and type of base pairs expected from the cloverleaf model. Both AU-2 and AU-11 are predicted to have the same position, at −14.5 ppm, whereas we observe two lines at −14.7 ppm and −14.3 ppm that correspond to these protons. While the calculated positions correspond quite well with the average of the experimental positions, the splitting is not ex-

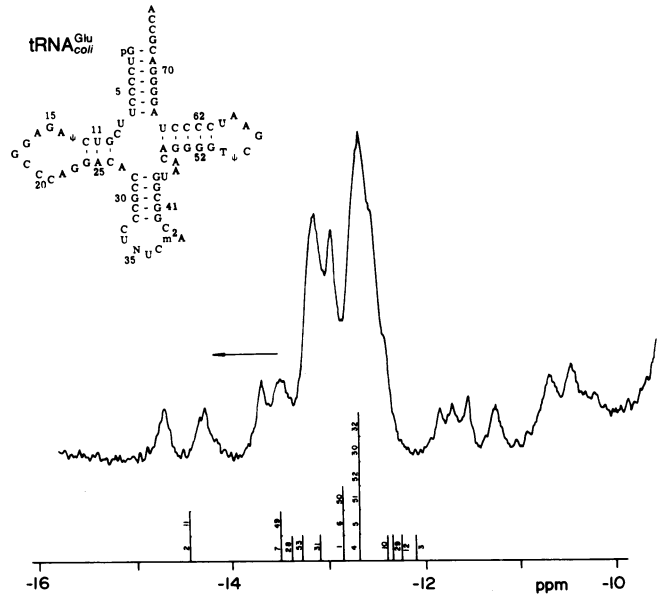

FIG. 5. 300-MHz spectrum of *E. coli* tRNA^{Glu} under the same conditions as Fig. 1. The *arrow* indicates that resonances assigned to AU-7 and AU-49 base pair would be calculated to fall 0.7 ppm further downfield if they were not assumed to be stacked upon each other in a regular helix.

plained. Since the nearest neighbors are identical and since our present calculation only includes nearest neighbor contributions, it is clear that the difference of 0.4 ppm represents a structural difference not included in the calculation. Perhaps the pitch of the helix is different in the dihydro-U stem that only contains three base pairs. This finding obviously requires further investigation.

Valuable information about the tertiary structure of the molecule is provided by the observation that the other two AU resonances assigned to AU-7 and AU-49 occur at −13.6 and −13.4 ppm, respectively. According to the ring current shift calculations, both of these are predicted to occur at −13.5 ppm only *if it is assumed that the acceptor stem and the TψC stem form a continuous helix so that A-66 is stacked on U-65*. Without this resulting shift both resonances would be located at −14.2 ppm. Thus, the observation that resonances from AU-7 and AU-49 both occur within 0.1 ppm of −13.5 ppm is consistent with the assumption that the helical arms of the acceptor and TψC stems are arranged in a continuous helix.

CONCLUSIONS

From our analysis of the NMR spectra of these and other tRNA molecules, we can draw the following conclusions about the structures of these molecules in solution. (*i*) The secondary structures in solution of yeast $tRNA^{Phe}$, *E. coli* $tRNA^{fMet}$, and *E. coli* $tRNA^{Glu}$ are well described by the cloverleaf model. In all three cases, plus a few others discussed elsewhere (3), the total number of base pairs and the AU/GC composition of the cloverleaf agree with those observed. In these three cases we have also shown that the resolved resonances can be assigned by the calculations to specific base pairs whose resonance positions are very sensitive to their neighboring bases and the assumption that the helical regions are all right handed (13). (*ii*) Because the low-field resonances can be accounted for so well by the cloverleaf model, we find no evidence for tertiary-structure Watson–Crick base pairs such as are proposed in several models (10); certainly there cannot be four or five additional pairs. (*iii*) The acceptor stem and TψC stem in $tRNA^{Glu}$ appear to form a continuous helix in which AU-7 of the acceptor stem is stacked on AU-49 base pair at the terminus of the TψC stem. (*iv*) In $tRNA^{fMet}$, the G-27 residue located between the dihydro-U stem and the anticodon stem is stacked on GC-10, and A-45 of the minor loop is, in various conformations of the molecule, differentially stacked on the AU-28 terminal base pair of the anticodon helix. (*v*) In yeast $tRNA^{Phe}$, the A-14 residue in the dihydro-U loop is stacked in at least two different ways upon the terminal GC-13 base pair of the dihydro-U arm.

Thus, the NMR spectra of these tRNA molecules provide conclusive evidence about their secondary structure and information about certain features of their tertiary structure in solution. Recently Kim *et al.* have reported (5) substantial progress in determining the crystal structure of yeast $tRNA^{Phe}$ to 4 Å. Although important questions remain as to the relation between the structure of these molecules in solution and the structure found in the crystalline state, the following points are worth noting. The 4-Å resolution picture indicates that the acceptor stem and the TψC stem form a continuous double helix. While the positions of the resonances in yeast $tRNA^{Phe}$ did not permit us to investigate this particular structural feature in $tRNA^{Phe}$, we do find that these two stems in *E. coli* $tRNA^{Glu}$ appear to form a continuous helix. Furthermore, the x-ray structure suggests (5) that the dihydro-U stem forms a continous helix with the anticodon stem, the purine between them being stacked on both. Evidence for this stacking of G-27 on the dihydro-U stem of *E. coli* $tRNA^{fMet}$ is presented above. These results suggest that both stackings observed in the crystal structure of yeast $tRNA^{Phe}$ may be common structural features of different tRNAs and do appear to be responsible for the NMR spectra in solution. The present work definitely has shown that these molecules maintain a cloverleaf secondary structure in solution and that there is no evidence for the large number of additional Watson–Crick base pairs postulated in certain models of the tertiary structure. When higher resolution x-ray data become available and as our assignments of the NMR spectra improve it will be interesting to compare additional structural features of these molecules in solution and in the crystalline state.

We are indebted to Lewis Cary and Varian Associates for aid in obtaining the 300-MHz NMR spectra. We thank Dr. A. D. Kelmer of Oak Ridge National Laboratory for supplying samples of $tRNA^{fMet}_{coli}$ and $tRNA^{Glu}_{coli}$. The support of the U.S. Public Health Service (Grant GM 19313 to D.R.K. and CA 11697 to B.R.R.) for that part of the work carried out at the University of California, Riverside, is most gratefully acknowledged. The partial support of the Niels Stensen Stichting to C.W.H. is gratefully acknowledged.

1. Kearns, D. R., Patel, D. J. & Shulman, R. G. (1971) *Nature* **229**, 338–339.
2. Wong, Y. P., Kearns, D. R., Reid, B. R. & Shulman, R. G. (1972) *J. Mol. Biol.* **72**, 725–740.
3. Shulman, R. G., Hilbers, C. W., Kearns, D. R., Reid, B. R. & Wong, Y. P. (1973) *J. Mol. Biol.*, in press.
4. Lightfoot, D. R., Wong, K. L., Kearns, D. R., Reid, B. R. & Shulman, R. G. (1973) *J. Mol. Biol.*, in press.
5. Kim, S. H., Quigley, G. J., Suddath, F. L., McPherson, A., Sneden, D., Kim, J. J., Weinzierl, J. & Rich, A. (1973) *Science* **179**, 285–288.
6. Giessner-Prettre, C. & Pullman, B. (1970) *J. Theor. Biol.* **27**, 87–95.
7. RajBhandary, U. L., Chang, S. H., Stuart, A., Faulkner, R. D., Hoskinson, R. M. & Khorana, H. G. (1967) *Proc. Nat. Acad. Sci USA* **57**, 751–758.
8. Lee, G. C. Y. & Chan, S. I. (1972) *J. Amer. Chem. Soc.* **94**, 3218–3229.
9. Wong, Y. P., Wong, K. L. & Kearns, D. R. (1972) *Biochem. Biophys. Res. Commun.* **49**, 1580–1587.
10. Levitt, M. (1969) *Nature* **224**, 759–763.
11. Dube, S. K., Marcker, K. A., Clark, B. F. C. & Cory, S. (1969) *Eur. J. Biochem.* **8**, 244–255.
12. Munninger, K. O. & Chang, S. H. (1972) *Biochem. Biophys. Res. Commun.* **46**, 1837–1842.
13. Arnott, S. (1971) *Progr. Biophys. Mol. Biol.* **22**, 181–213.