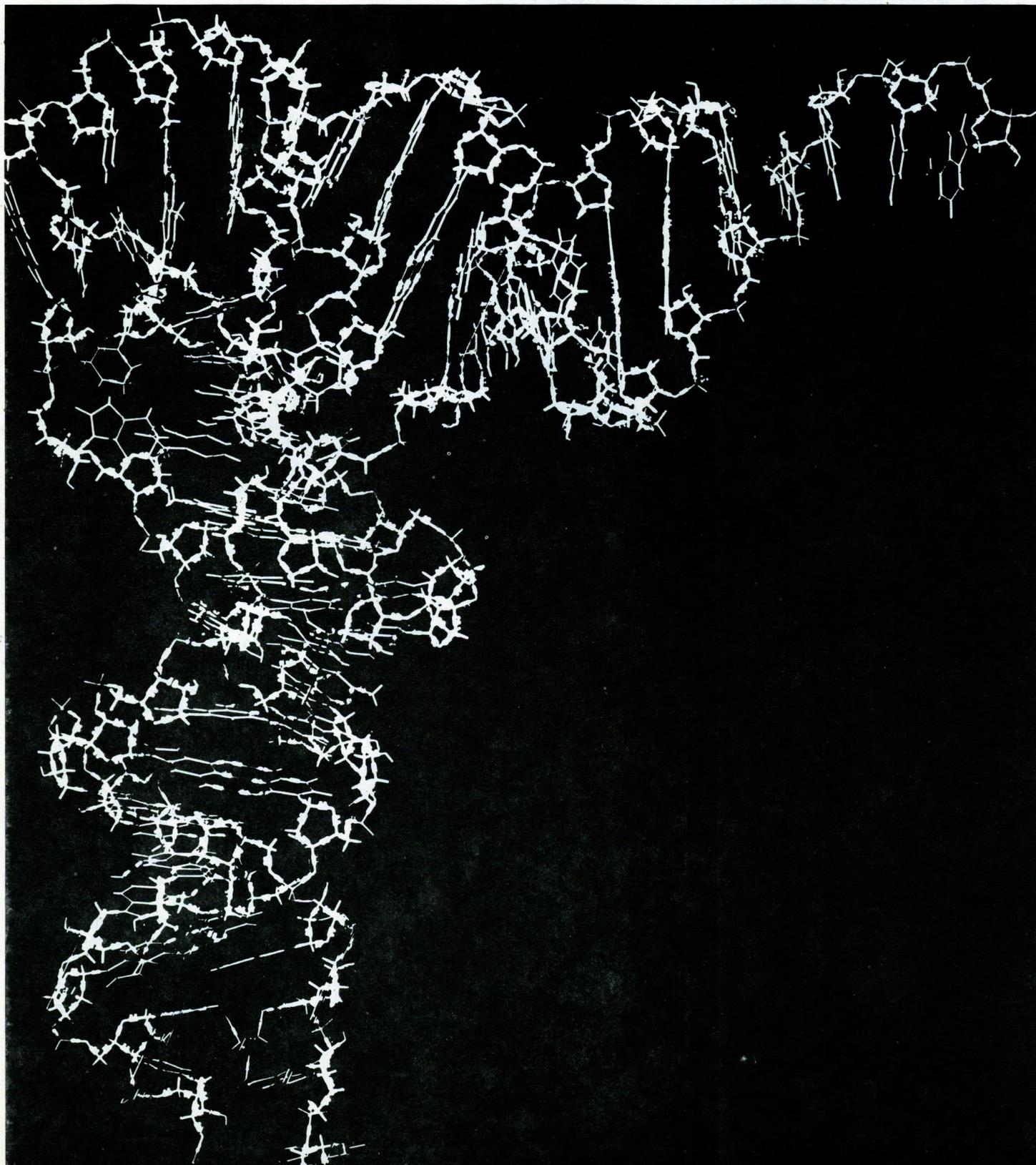


Three-Dimensional Tertiary Structure of Yeast Phenylalanine Transfer RNA

S. H. Kim, F. L. Suddath, G. J. Quigley, A. McPherson,
J. L. Sussman, A. H. J. Wang, N. C. Seeman and Alexander Rich



Reports

Three-Dimensional Tertiary Structure of Yeast Phenylalanine Transfer RNA

Abstract. *The 3-angstrom electron density map of crystalline yeast phenylalanine transfer RNA has provided us with a complete three-dimensional model which defines the positions of all of the nucleotide residues in the molecule. The overall features of the molecule are virtually the same as those seen at a resolution of 4 angstroms except that many additional details of tertiary structure are now visualized. Ten types of hydrogen bonding are identified which define the specificity of tertiary interactions. The molecule is also stabilized by considerable stacking of the planar purines and pyrimidines. This tertiary structure explains, in a simple and direct fashion, chemical modification studies of transfer RNA. Since most of the tertiary interactions involve nucleotides which are common to all transfer RNA's, it is likely that this three-dimensional structure provides a basic pattern of folding which may help to clarify the three-dimensional structure of all transfer RNA's.*

The three-dimensional structure of transfer RNA (tRNA) is of considerable interest in that it may provide insight into its mode of action. Approximately a year and a half ago, we reported the overall backbone conformation of yeast phenylalanine tRNA as obtained from an x-ray diffraction study at a resolution of 4.0 Å (1). This showed that the molecule is L-shaped with the anticodon loop at one end of the L, the amino acid acceptor at the other end, and the TΨC (2) and D loops of the familiar tRNA cloverleaf structure meeting at the corner. We found that the hydrogen bonding implied in the cloverleaf structure (Fig. 1) was maintained in this three-dimensional structure. The nucleotide sequences of over 60 tRNA's have been reported and all of them maintain the same secondary cloverleaf structure (3). In addition, certain nucleotides at particular positions on the cloverleaf are retained in all tRNA molecules. Several months ago, we reported an electron density map at a resolution of 3 Å in which individual ribose groups, phosphate groups, and bases are visualized (4). This map confirmed the overall flow of the polynucleotide backbone which we had reported 1 year earlier. At that time, we considered alternative interpretations of the 3-Å electron density map which differed from each other in the identification of particular peaks of electron density with individual nucleotides. One of these interpretations was

described (4). However, further analysis and interpretation of the 3-Å electron density map has led us to conclude that an alternative assignment of nucleotides was more reasonable. Some of the detailed assignments of nucleotide residues to electron density peaks were in error, even though the overall tracing of the backbone was generally correct. We have now made a more comprehensive interpretation which enables us to define

the tertiary interactions in the molecule. The overall tracing of the 3-Å map remains in agreement with that described at 4-Å resolution. However, our present interpretation allows us to visualize a number of tertiary interactions which not only explain the structural role of most of the bases which are constant in tRNA's, but also allows us to understand in a direct and simple fashion the chemical modification data on tRNA. In addition, this pattern of tertiary interactions provides a basis for understanding the general three-dimensional folding of all tRNA molecules.

Crystallization conditions, heavy atom positions, and refinement statistics for the orthorhombic crystals of yeast phenylalanine tRNA have been reported previously (1, 4, 5). In addition, three sections of the electron density map at 3-Å resolution have been published which illustrate the extent to which ribose, phosphate, and base peaks are seen at this resolution. In interpreting the electron density map, Kendrew skeletal models at a scale of 2 cm per angstrom are used to build the polynucleotide chain in an optical comparator so that the model can be superimposed on the electron density map. In building the molecular model, the phosphates, riboses, and bases are put into positions with acceptable stereochemistry where there are peaks in the electron density map. As pointed out in our report on the 3-Å map (4),

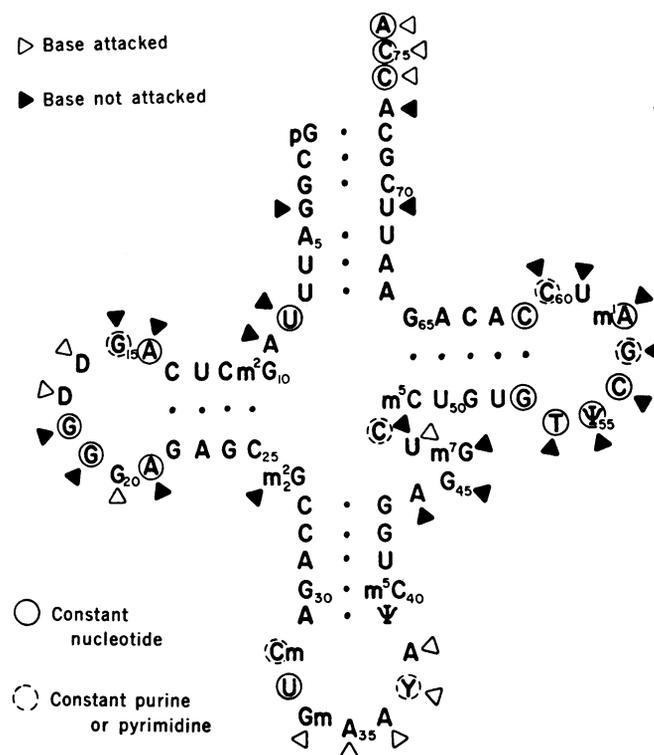


Fig. 1. Nucleotide sequence of yeast phenylalanine tRNA shown in the cloverleaf configuration (6). Circled bases are constant in all tRNA's, and dashed circles indicate positions which are occupied constantly by either purines or pyrimidines. Eukaryotic initiator tRNA's do not have the same constant nucleotides, however. The accessibility of various bases to chemical modification is indicated by open or closed triangles. The chemical modification data were obtained on various tRNA's (6).

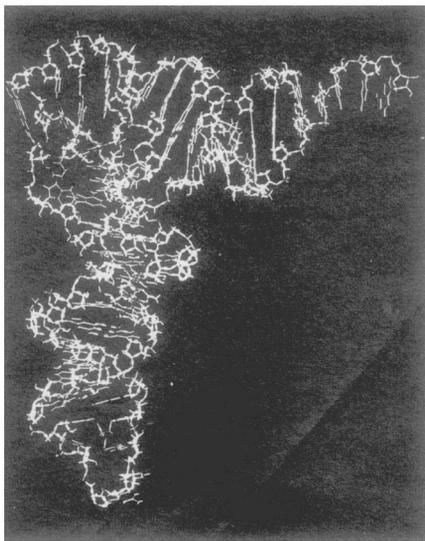


Fig. 2. Photograph of a molecular model of yeast phenylalanine tRNA built with Kendrew wire models at a scale of 2 cm per angstrom. The supporting rods for the model have been eliminated from the photograph and the molecule is oriented with the anticodon loop at the bottom, the acceptor stem at the upper right, and the TΨC loop at the upper left. The view is approximately perpendicular to the molecular plane.

it is relatively easy to trace positions of the polynucleotide chains in the double helical stems of the tRNA molecule; however, it is more difficult to trace their course in the loop regions. A nucleotide assignment was previously made to individual peaks in the electron density map of a stem region, which we subsequently found to be incorrect by one nucleotide. Although this did not modify the overall tracing of the chain, it nonetheless had significant effects in modifying the fine details of the molecule, such as the orientation of the hydrogen-bonded base pairs in the D stem, the assignment of residues in the anticodon loop, and the position of the Y base. Ambiguities in tracing the

chain, especially in the loop regions, were sometimes resolved by using a "partial structure method" in which structure factors were calculated for segments of the molecule whose position was unambiguously identified, such as those in the helical stem regions. In carrying out these calculations, simplified scattering factors were used for ribose, phosphate, and base residues. Calculations of this type together with a careful study of the electron density map have now led us to a complete detailed tracing of the polynucleotide chain.

Form of the molecule. Figure 1 shows the nucleotide sequence of yeast phenylalanine tRNA (6). The circled nucleotides are those which are constant in most tRNA's that have been sequenced. The triangles next to the nucleotides in the loop regions indicate whether or not the base is modified when tRNA is subjected to chemical modification experiments (7). Figure 2 is a photograph of the skeletal model of yeast phenylalanine tRNA. As

pointed out in our 4-Å analysis (1), the molecule is somewhat flattened and has an L-shaped conformation. The photograph in Fig. 2 was taken perpendicular to the molecular plane with the anticodon at the bottom, the CCA acceptor end at the upper right, and the TΨC loop at the upper left corner of the molecule. To aid in the interpretation of this photograph, a schematic drawing is presented in Fig. 3 with the molecule in approximately the same orientation. The bars between the helical backbones represent the hydrogen-bonding elements of secondary and tertiary structure. Tertiary interactions are shown as black rods. The numbers refer to the numbering system of the residues shown in Fig. 1. The overall tracing of the polynucleotide chain shown in Fig. 3 is virtually the same as that illustrated for the 4-Å analysis (1).

The most striking feature of the molecule as seen in Fig. 2 is the fact that each arm of the L contains an elongated column of bases which are stacked roughly perpendicular to the arm. Part of this stacking is due to the fact that the helical stem regions are oriented approximately along these arms, but the remainder is due to contributions from the loop regions. As pointed out earlier (3), the helical stems are not exactly colinear. The acceptor and TΨC stems diverge from each other by about 17° while the anticodon and D stems diverge by about 24°. However, the acceptor stem is approximately at right angles to the anticodon stem.

Tertiary interactions. We have been able to construct a model fitting the electron density map which contains ten different tertiary interactions involving hydrogen bonding. Nine are illustrated diagrammatically in Fig. 4. Figure 4a shows the sequence of yeast phenylalanine tRNA drawn in the cloverleaf conformation. The enclosed bases are connected by solid lines indicating hydrogen bonding between the bases. The bonding involves two or three hydrogen bonds, but may be single hydrogen bonds in some cases as discussed below. The cloverleaf sequence has been redrawn in Fig. 4b to illustrate schematically the manner in which the D and TΨC loops are close to each other at the corner of the molecule, with the tertiary interactions as indicated.

TΨC loop. The most striking characteristic of this loop is the extent to

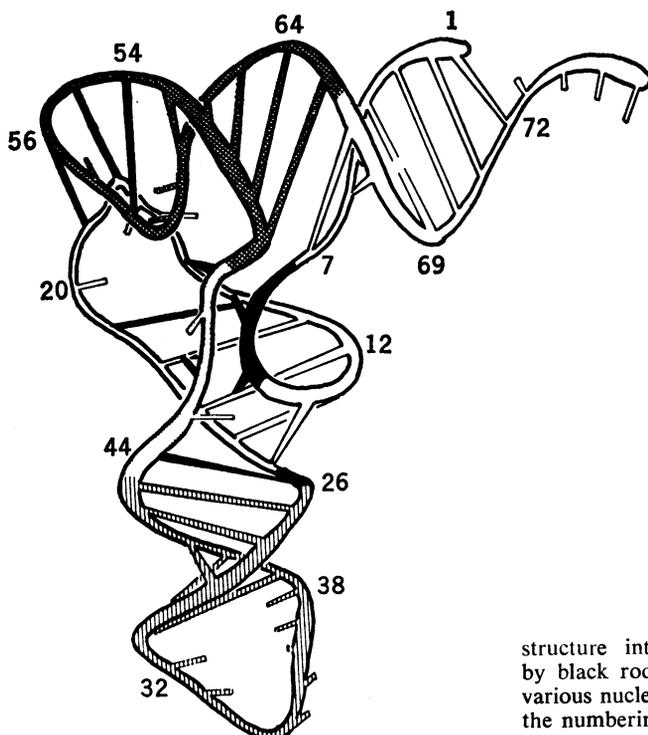


Fig. 3. Schematic model of yeast phenylalanine tRNA drawn from the coordinates of the molecule shown in Fig. 2. The ribose phosphate backbone is drawn as a continuous cylinder with bars to indicate hydrogen-bonded base pairs. The positions of single bases are indicated by rods which are intentionally shortened. The TΨC arm is heavily stippled, and the anticodon arm is marked by vertical lines. The black segments of the backbone include residues 8 and 9 as well as 26. Tertiary structure interactions are illustrated by black rods. The numbers indicate various nucleotides in accordance with the numbering system in Fig. 1.

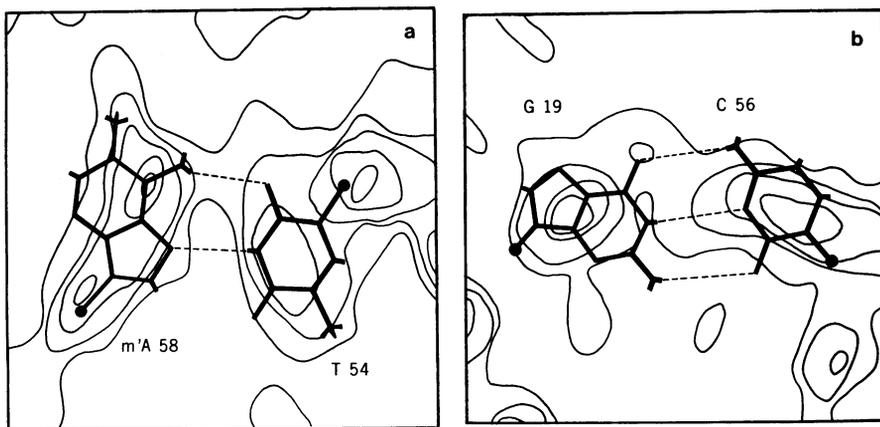


Fig. 5. Sections of the electron density map showing the interaction of two of the base pairs involved in tertiary interactions. These include the hydrogen bonding between (a) 1-methyladenine 58 and thymine 54 and (b) guanine 19 and cytosine 56. Solid circles indicate C1' atoms.

lengthened by the addition of two base pairs as described above (G15 · C48 and U8 · A14). In addition, the stacking is complemented by three bases which form ternary hydrogen-bonding systems in the major groove of the helix. These additional bases are largely to the left of the vertical stem in Fig. 2. The adenine residue of A21 is near the ribose of U8 and is coplanar with the base pair U8 · A14. The electron density map clearly shows A21 in a position where it can hydrogen bond, but at present we are unable to be certain whether this involves either one or two hydrogen bonds with the U8 · A14 base pair. The next ternary complex involves hydrogen bonding between m⁷G46 with the pair G22 · C13 of the D stem (see Fig. 3). The amino group and N1 of m⁷G46 may be hydrogen bonded to the O6 and the N7 of G22. In the base pair next to this, the adenine of A9 is hydrogen bonded to the adenine of A23 which is itself paired to U12 in the D stem. This hydrogen bonding between two adenines and the uracil residue is probably similar to that reported in the crystal complex containing the uracil derivative phenobarbital, which is hydrogen bonded to two molecules of 9-ethyladenine (8).

At the bottom of the D stem, as shown in Fig. 2, are two bases which lie almost in the same plane. One of these, the dimethylated guanine (m₂²G26) is intercalated between the base pair m²G10 · C25 of the D stem and C27 · G43 at the top of the anticodon stem. Furthermore, m₂²G26 appears to be hydrogen bonding to A44 of the variable loop, which is approximately parallel to it, even though the

two purine planes are twisted in a propeller fashion relative to each other. This close relation may involve one or two hydrogen bonds.

The most striking feature of this center part of the molecule is the extent to which the bases of the variable loop occupy the major groove of the D stem. The only exception to this is residue U47, which protrudes away from the stem so that the uracil is completely external and not stacked. Residue G45 has its guanine stacked into the center of the stem although it is tilted relative to the two bases on either side, A44 and A9 (see Fig. 3).

D loop. This is on the outside of the molecule where it overlaps and covers part of the T Ψ C loop. This feature of the molecule was apparent in the tracing of the chain at 4-Å resolution (1). We describe the loop sequentially. As mentioned above, residue A14 of the D loop is hydrogen bonded to U8, while G15 is hydrogen bonded to C48. The two dihydrouracil nucleotides D16 and D17 appear as a kind of bulge on the exterior as the polynucleotide chain loops out in such a way as to leave both dihydrouracil rings on the outside of the molecule. The chain then abruptly turns into the molecule where the guanine of G18 is intercalated between the adenine of m¹A58 and G57 of the T Ψ C loop (see Fig. 3). Guanine of G19 is hydrogen bonded to C56 as described above, while the guanine of G20 is in a region that loops out so that the guanine base is exposed. The adenine of A21 is hydrogen bonded as described above. The D loop is rather elongated with most of the bases extending into the molecule where they are either stacked or hydrogen bonded

to other bases, except the two dihydrouracil residues D16 and D17 and the guanine of G20 as has already been noted.

Anticodon stem. In comparison with the D stem and its considerable number of ternary interactions, the anticodon stem is remarkably simple. It has no tertiary interactions, but only a double helical stacking of the five base pairs leading up to the anticodon loop. The bases in the anticodon loop are largely stacked with residues A38 and Y37 approximately parallel to the base pair A31 · Ψ 39 at the very end of the anticodon stem. The anticodon bases themselves are found at the tip of the anticodon loop as shown in Fig. 3. These are involved in the interactions associated with the twofold rotation symmetry axis in the crystal; the fine details of these interactions will be described elsewhere.

An interesting feature is found at the junction of the D stem and the anticodon stem. Although the dimethyl G26 is clearly stacked between the ends of the D and anticodon stems, the base of A44 does not appear to be in van der Waals contact with G45 but instead some empty space is found between them. This suggests the possibility of some movement of the anticodon stem and loop relative to the remainder of the molecule when tRNA is not in the crystal lattice.

T Ψ C and acceptor stems. There are relatively few interactions between the T Ψ C and acceptor stems and the rest of the molecule, in contrast to the numerous interactions involving the D stem. Both the major and minor grooves of both these stems are accessible to solvent. As mentioned previously (4), the electron density associated with the base pair G4 · U69 lies between G3 · C70 and A5 · U68 in a stacking interaction. At the present time, we cannot say with certainty whether this involves one or two hydrogen bonds. At the end of the acceptor stem, the base of A73 is stacked on the base pair G1 · C72, while the remaining residues 74 through 76 are approximately an extension of the helical stem.

To facilitate understanding of the molecule, we have reproduced in Fig. 6 a computer drawing which shows the relative positions of the phosphate and ribose groups together with dotted lines indicating the hydrogen bonding associated with secondary interactions. This diagram, together with the schematic diagram in Fig. 3, may help in visual-

izing the overall structure of the molecule.

Hydrogen-bonding and stacking interactions. We have defined hydrogen-bonding tertiary interactions as those which arise where there are electron density peaks associated with purines and pyrimidines in positions close enough to assume that the bases will be connected by hydrogen bonds. Examples are shown in the electron density maps of Fig. 5. In describing hydrogen-bonding interactions, we have indicated that some of the tertiary interactions may be associated with either one or two hydrogen bonds. The uncertainty arises from the fact that in some cases the electron density associated with the base does not define its position adequately for discrimination. An extension of this study to higher resolution and the introduction of multiple phasing methods should lead to the resolution of these ambiguities. Most of the bases involved in the tertiary interactions are stacked parallel to the bases in stem regions. Thus, the tertiary interactions which promote the stability of the molecule include stacking interactions as well as hydrogen-bonding networks. The latter provide a key element of specificity in the three-dimensional folding of the molecule.

In Fig. 1 we have circled the bases which are common to most tRNA molecules. If we ignore the bases in the CCA end of the molecule and in the anticodon loop, a total of 15 nucleotides are circled either as constant bases or constant purines or pyrimidines. It is interesting to note that 2 of these are involved in secondary structure (G53 · C61) and 13 are involved in hydrogen-bonding interactions which contribute to the specificity of tertiary structure, if we include the possible interaction between C60 and C61 discussed above. Only one base (G57) is involved in stacking interactions without forming any apparent hydrogen bonds to other nucleotides.

We have described ten hydrogen-bonded tertiary interactions, and most of these involve bases which are constant features of tRNA structures (3). Some of these interactions have been proposed by previous investigators who have tried to assemble molecular models of tRNA tertiary structure. Levitt (9), for example, postulated 11 tertiary interactions, and 5 of those are among the 10 which we observe in the present structure.

Figure 1 summarizes a number of

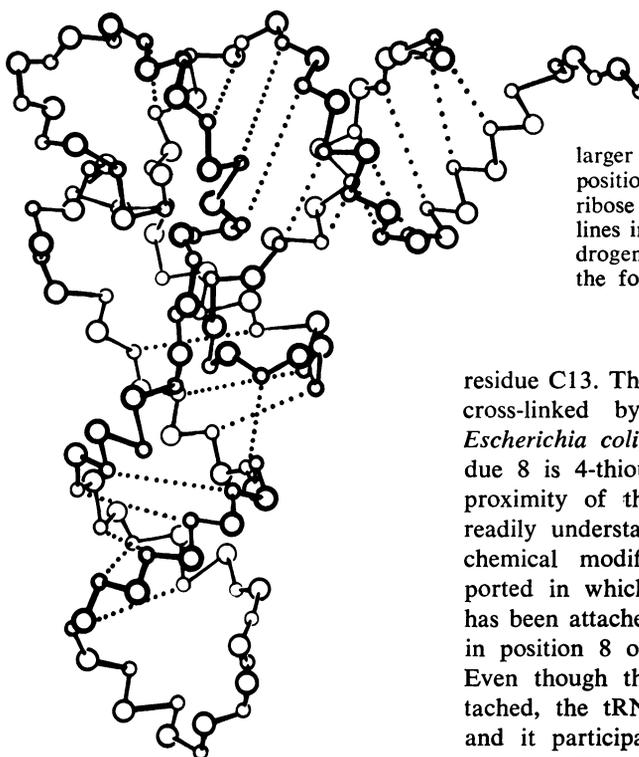


Fig. 6. Computer drawing of the ribose phosphate backbone of yeast phenylalanine tRNA. The larger circles indicate phosphate positions and the smaller ones ribose positions. The dotted lines indicate the secondary hydrogen-bonding interactions in the four stem regions.

chemical modification studies which have been carried out in yeast phenylalanine tRNA as well as other tRNA's (7). The most important data are nucleotides which are not modified chemically, since those are less prone to artifactual unfolding of the molecule and subsequent modification. There is a very good correlation between the chemical modification of bases and their accessibility in the structure. In particular, residues D16, D17, and G20 in the D loop are all exposed and available for chemical modification. Similarly, none of the residues in the TΨC loop are exposed and none are chemically modified. Residues 8 and 9 are likewise blocked. In the variable loop, residue U47 is the only one not involved either in hydrogen bonding or in an intercalative stacking interaction. That residue sits outside the molecule and is readily modified chemically. Most of the residues in the anticodon loop are available for chemical modification. A possible exception to this may be associated with a stacking interaction which we observe between residues C32 and U33 and which suggests that they may not be fully exposed. Residue A73 is stacked on the end of the acceptor stem overlapping the base pair G1 · C72. Perhaps the stability of this stacking interaction is enough to prevent ready accessibility for chemical modification. In this model, residue U8 is parallel to and almost stacked over

residue C13. These residues are readily cross-linked by ultraviolet light in *Escherichia coli* tRNA, in which residue 8 is 4-thiouracil (10). The close proximity of these bases makes that readily understandable. An interesting chemical modification has been reported in which a photoaffinity label has been attached to S4 of 4-thiouracil in position 8 of *E. coli* tRNA (11). Even though this affinity label is attached, the tRNA is fully chargeable and it participates readily in protein synthesis. This is understandable from inspection of the model, since the hydrogen bonding between U8 and A14 involves the uracil O2 rather than O4. Oxygen 4 is available to the solvent since it lies at the bottom of a U-shaped trough on the surface of the molecule.

Since most of the constant bases in all tRNA sequences are involved in these tertiary interactions, the three-dimensional structure of phenylalanine tRNA may serve as a model for understanding the structure of all tRNA's. In particular, it may serve as a very good guide for understanding the tertiary interactions of tRNA species which have four base pairs in the D stem and five nucleotides in the variable loop. However, it is likely that some additional tertiary interactions will be found as we explore the structure of other tRNA molecules, especially those with large loops.

S. H. KIM

*Department of Biochemistry,
Duke University School of Medicine,
Durham, North Carolina 27710*

F. L. SUDDATH, G. J. QUIGLEY
A. MCPHERSON

*Department of Biology,
Massachusetts Institute of Technology,
Cambridge 02139*

J. L. SUSSMAN

*Department of Biochemistry,
Duke University School of Medicine*

A. H. J. WANG, N. C. SEEMAN

ALEXANDER RICH

*Department of Biology,
Massachusetts Institute of Technology*

References and Notes

1. S. H. Kim, G. J. Quigley, F. L. Suddath, A. McPherson, D. Sneden, J. J. Kim, J. Weinzierl, A. Rich, *Science* **179**, 285 (1973).
2. The abbreviations used and notations are as follows: A, adenosine; T, thymidine; G, guanosine; C, cytidine; U, uridine; D, dihydrouridine; Ψ , pseudouridine; Y, a purine nucleoside; and m_1 , methyl, and m_2 , dimethyl (positions of these are noted by superscripts). Positions of the bases are indicated by the numbers placed after the name of the base.
3. A number of tRNA sequences are found in R. Holmquist, T. H. Jukes, S. Pangburn, *J. Mol. Biol.* **78**, 91 (1973).
4. F. L. Suddath *et al.*, *Nature (Lond.)* **248**, 20 (1974).
5. S. H. Kim, G. Quigley, F. L. Suddath, A. McPherson, D. Sneden, J. J. Kim, J. Weinzierl, A. Rich, *J. Mol. Biol.* **75**, 421 (1973); S. H. Kim, G. J. Quigley, F. L. Suddath, A. McPherson, D. Sneden, J. J. Kim, J. Weinzierl, P. Blattmann, A. Rich, *Proc. Natl. Acad. Sci. U.S.A.* **69**, 3746 (1972).
6. U. L. Rajbhandary and S. H. Chang, *J. Biol. Chem.* **243**, 598 (1968).
7. Chemical modifications are reviewed by: F. Cramer, *Prog. Nucleic Acid Res. Mol. Biol.* **11**, 391 (1971); S. E. Chang and D. Ish-Horowitz, *J. Mol. Biol.* **84**, 375 (1974).
8. S. H. Kim and A. Rich, *Proc. Natl. Acad. Sci. U.S.A.* **60**, 402 (1968).
9. M. Levitt, *Nature (Lond.)* **224**, 759 (1969).
10. M. Yaniv, A. Favre, B. G. Barrel, *ibid.* **223**, 1331 (1969).
11. I. Schwartz and J. Offengand, *Proc. Natl. Acad. Sci. U.S.A.*, in press.
12. We are indebted to George Harris, Lisa Eunson, and B. Rubin for technical assistance. This research was supported by grants from the National Institutes of Health (CA04186 and CA15802) and the National Science Foundation (GB30688 and GB40814), the National Aeronautics and Space Administration, and the American Cancer Society. G.J.Q. is a fellow of the Medical Foundation, J.L.S. is a fellow of the Arthritis Foundation, N.C.S. is a fellow of the National Institutes of Health, and A.H.J.W. is supported by grant CA 14015 from the National Cancer Institute.

3 July 1974

we found that the bottom waters of the lagoon were uncomfortably hot (44° to 47°C) and were overlain by progressively cooler surface waters. Thus part of the study was directed toward this unusual thermal layering phenomenon.

The small village of Pueblo Gran Roque extends along the west coast of the island. The northern part of Gran Roque consists, in general, of a narrow ridge of metadiabase and metalamprophyre intruded by quartz diorite, granitic aplite, and pegmatite, which rises up to 115 m above sea level. The southeastern part of the island is flat and contains several connected small lagoons. The westernmost lagoon in the series, Lago Pueblo, contains the hot brines.

Lago Pueblo is roughly circular in outline and, on the average, less than 1 m deep. The unusual feature of the lagoon is a deep pool near its southwestern edge which was probed to a depth of 5.25 m. The water in the lagoon is derived to the largest extent from seawater inflow from the adjacent lagoons to the east but also in part from a freshwater stream draining the ridge to the northwest (1).

We measured the temperature of the water along a traverse from the eastern shore of the lagoon toward the deep by placing a thermometer at the desired level, allowing it to equilibrate, and then reading the temperature in place by means of a snorkel mask. Temperature readings at depths greater than 75 cm were difficult because of (i) the blurring of the water caused by thermal mixing, (ii) the extreme discomfort occasioned by the hot brine in the deeper parts, and (iii) the buoyancy effect of the dense brine. Samples of the water were taken at specific depths for the determination of densities and refractive indices.

The plot of temperature versus depth (Fig. 2) shows that the lagoon is divided into three distinct temperature regimes. The water from the surface to a depth of 27 cm was relatively cool, and the temperature increased gradually from 25° to 26.7°C . At 30 cm a thermocline was encountered in which the temperature gradient increases by 0.594°C per centimeter up to a maximum temperature of 43.3°C , encountered at depths over 53 cm. Both thermocline surfaces were visually observed at about 25 and 52 cm below the water surface and appeared as sharp, uneven planes in cross section.

Hot Brines on Los Roques, Venezuela

Abstract. *Solar heating of dense natural brines to 47°C has been observed in a small lagoon in the Venezuelan Antilles and in an artificial brine reservoir in New York State. Calculations indicate that approximately 90 percent of the solar energy can be trapped in the brine by a combination of refraction and absorption. Brine ponds should be considered as solar energy collectors.*

Natural hot brines (as opposed to geothermal brines) have been found in a restricted shallow island lagoon in the Venezuelan Antilles. The dense brines result from the normal evaporation of seawater. They are covered by less dense, cooler natural seawater and freshwater.

The lagoon is located on Gran Roque or El Roque, the principal island in a group of islands known as Los Roques (see Fig. 1). This island

group is located approximately 150 km north of Caracas, Venezuela. We visited the island in the early part of December 1973, about a week after the rainy season, when water levels were unusually high and the lagoons enlarged. Even the central part of the island, normally dry, was inundated with brackish water. The original purpose of our visit was to study the formation of lagoonal gypsum and of associated algal carbonates. However,

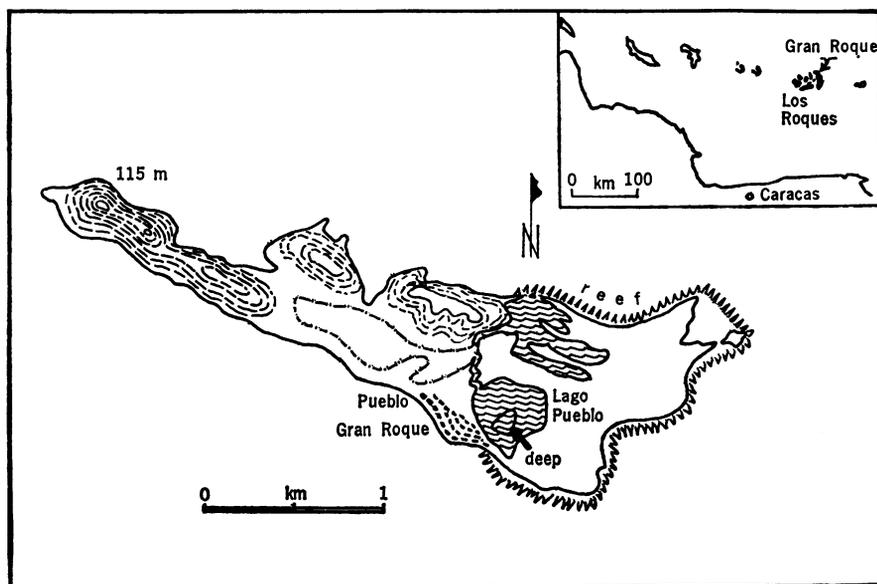


Fig. 1. Location map: outline of Gran Roque and of Lago Pueblo from aerial photographs.