Three-Dimensional Structure and Biological Function of Transfer RNA

Alexander Rich

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

Received February 22, 1977

Biological information is encoded in the nucleotide sequence of very long DNA molecules. However, biological information is expressed mostly in the synthesis of proteins, where the sequence of amino acids is determined ultimately by the nucleotide sequence in DNA through messenger RNA.

Transfer RNA (tRNA) is a small polynucleotide which plays a central role in this process. At one end of the molecule it interacts with messenger RNA in a specific manner; at the other end it is attached to a particular amino acid. In the ribosome, this is the structure upon which the polypeptide chain grows. We would like to understand how a particular amino acid becomes attached to a specific family of tRNA molecules, and further how the aminoacyl-tRNA interacts with the ribosomal machinery to assemble a protein.

A central theme in molecular biology is the attempt to understand the biological function of macromolecules through a study of their three-dimensional structure. It is rare that biological function is fully understood through an appreciation of structure alone, but structure provides an essential element. Transfer RNA molecules contain 75 to 90 nucleotides, and they are usually presented in the cloverleaf diagram, which seems to be a general feature.[1] Its pattern of stems, loops, and conserved bases only becomes comprehensible in terms of three-dimensional structure. The major method for determining structure is x-ray diffraction, and this requires crystallization of the macromolecule. Frequently, this is the rate-limiting step in solving the structure.

X-Ray Diffraction Studies

Transfer RNA molecules were first crystallized in 1968 independently in several laboratories.[2] A large number of species were found to form crystals, and this gave rise to optimism that the three-dimensional structure would be known in a relatively short time period. At that time, however, it was not generally appreciated that all of the crystal forms had a defect: they were disordered. Our own crystals of *E. coli* $tRNA^{Phe}$ had a resolution of only 20 Å[3] in the diffraction pattern, and they were virtually useless for the purpose of solving the three-dimensional structure of the molecule. In protein crystallography one must have an x-ray diffraction pattern with a resolution of between 2 and 3 Å in order to trace the polypeptide chain.

Alexander Rich is the William Thompson Sedgwick Professor of Biophysics at Massachusetts Institute of Technology. He received an A.B. degree from Harvard College and, in 1949, an M.D. from Harvard Medical School. While a postdoctoral fellow with Linus Pauling at Caltech, he began work on polynucleotides, which he continued after moving to the National Institutes of Health in 1954. Dr. Rich joined the MIT faculty in 1958. There he has worked on the physical chemistry and structural analysis of the nucleic acids and their derivatives, as well as on the molecular biology of the nucleic acids with some emphasis on the mechanism of protein synthesis.

The origin of the disorder in tRNA crystals is not readily apparent. A large part of it, however, must be due to the polyelectrolytic nature of the molecule. These molecules have from 75 to 90 negative charges, and the exact ordering of the tRNA molecules in the crystal lattice is very sensitive to the nature and number of cations present. It is likely that the polyelectrolytic nature of the molecule gives rise to frequent mistakes in the building of the lattice. Almost 100 tRNA crystals have now been reported, and they typically have a resolution between 6 and 8 Å.[4] In some cases resolution is slightly better, but in most cases the crystals are unsuitable for carrying out three-dimensional analysis to solve the conformation of the molecule.

In this laboratory we worked for 3 years before we discovered a suitable crystal form. In 1971 we reported that, if one added the oligocation spermine to yeast $tRNA^{Phe}$, an orthorhombic crystal could be formed which yielded an x-ray diffraction pattern with a resolution of nearly 2 Å.[5] Spermine will also stabilize other crystal lattices of yeast $tRNA^{Phe}$.[6,7] However, further investigation revealed that the addition of spermine to other tRNAs would not give rise to ordered lattices.

Another phenomenon associated with the crystallography of tRNA is the high degree of polymorphism, that is, the formation of several different crystal forms of the same molecule. Figure 1 shows microphotographs of four different types of yeast $tRNA^{Phe}$ crystals. Over a dozen different unit cell and space groups are now known for this tRNA species.[4] Polymorphic forms are usually created by simply varying crystallization conditions, mostly altering the type and number of cations in the crystallizing mixtures. Rather small changes can cause substantial changes in the type of lattice formed.

The peculiar role of spermine can be seen in a study of different crystal forms of yeast $tRNA^{Phe}$. In the crystals shown in Figure 1, the orthorhombic, cubic, and monoclinic crystal forms are stabilized by the addition of spermine and produce diffraction patterns with a resolution between 2 and 3 Å.[5-7] On the other hand, the hexagonal crystal form without spermine yields a diffraction pattern with a resolution near 15 Å.

We started to work on the orthorhombic crystal of yeast $tRNA^{Phe}$. It had one molecule in the asymmetric

(1) A. Rich and P. R. Schimmel, *Acc. Chem. Res.*, 10 (1977), first paper in this issue.
(2) The early crystallization work is reviewed by F. Cramer, *Prog. Nucleic Acid Res. Mol. Biol.*, 11, 391 (1971).
(3) S. H. Kim and A. Rich, *Science*, 162, 1381 (1968).
(4) For a recent review of tRNA sequences, molecular structure, chemical modifications, and solution chemistry, see A. Rich and U. RajBhandary, *Annu. Rev. Biochem.*, 45, 805 (1976).
(5) S. H. Kim, G. J. Quigley, F. L. Suddath, and A. Rich, *Proc. Natl. Acad. Sci. U.S.A.*, 68, 841 (1971).
(6) T. Ichikawa and M. Sundaralingam, *Nature (London), New Biol.*, 236, 174 (1972).
(7) S. H. Kim, G. Quigley, F. L. Suddath, A. McPherson, D. Sneden, J. J. Kim, J. J. Weinzierl, and A. Rich, *J. Mol. Biol.*, 75, 421 (1973).

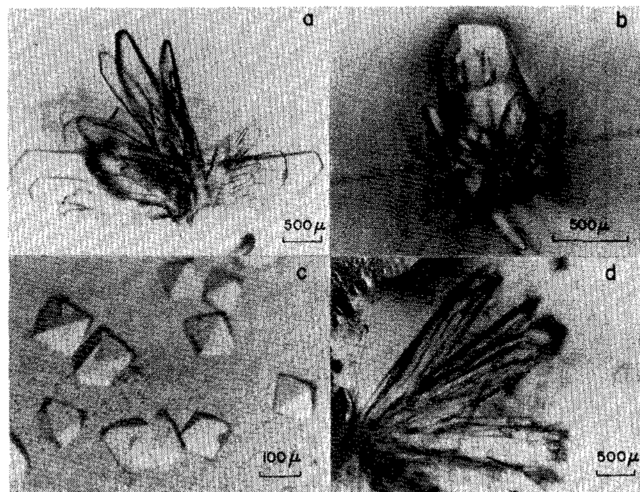

Figure 1. Photomicrographs of yeast phenylalanine tRNA crystals.[7] (a) Thin-lath clusters of orthorhombic crystals ($P2_12 2_1$); (b) hexagonal prisms ($P6_222$); (c) octahedral cubic forms ($I4_132$); (d) a cluster of striated thin laths of monoclinic crystals ($P2_1$).

unit and dimensions a = 33 Å, b = 56 Å, c = 161 Å. The diffraction pattern went out to a resolution of nearly 2 Å, and the crystal itself contained 70% water. However, the molecules slid together when the water content was reduced to produce a series of unit cells in which the c axis became successively shorter down to 109 Å without any appreciable change of the a and b directions.[8] This contraction was accomplished by reducing the size of aqueous channels which normally traverse the crystal in the fully hydrated state.

X rays are scattered by electrons, and the scattering by a crystal is concentrated in a series of reflections, the intensity of which can be readily measured. The basic problem of x-ray crystallography is to determine the relative phases of the different reflections. With that information it is possible to reconstruct a three-dimensional electron density map which in turn can be interpreted in terms of the conformation of the folded polynucleotide chain. The method for determining the phases of the reflections involves placing various heavy atoms in the crystal lattice without disturbing the packing in the lattice. These modify the diffraction pattern, and from this the original phases can be determined.

However, this was the first macromolecular nucleic acid crystal to be studied and, unlike proteins, there was no literature concerning which substances could be used for heavy atom derivatives. This search took us almost 2 years, and we tried almost 200 derivatives to find suitable heavy-atom replacements. The most useful derivatives were the lanthanides, especially samarium chloride.[9] The samarium ion presumably replaces magnesium ions in the crystal lattice with minimal perturbation. Samarium ions appeared at several different sites and could be used not only as heavy atoms but also for their anomalous x-ray scattering component which helped to determined the phases as well as the handedness of the molecule. Another useful derivative, which went largely to a single site, was a bis(pyridyl)osmate diester which has the ability to bind to the adjacent cis hydroxyls at the 3′ end of the chain.[9] These derivatives were suggested by the work of Sigler and his colleagues.[10]

In early 1973 the folding of the polynucleotide chain was traced from an electron density map at 4-Å resolution.[11] The tracing was made possible at this resolution because the electron-dense phosphate groups could be seen even at this low resolution. The folding of the chain (Figure 2) was highly unusual and had not been anticipated by any of the model builders. This tracing demonstrated that the stem regions were in the form of RNA double helices as had been assumed from the sequence data. Furthermore, the stems had an unusual organization. The molecule was bent into an L shape with the acceptor stem and T-ψ-C stem forming one limb of the L while the D stem and anticodon stem formed the other limb. The 3′-terminal adenosine to which the amino acid was attached was at one end of the L, while the anticodon was at the other end, some 76 Å away. The molecule was fairly flat, about 20–25 Å in thickness. Although the coiling of the chain could be seen at 4-Å resolution, the details of the tertiary interactions had to await the results of a higher resolution analysis.

In 1974 the analysis was extended to 3 Å and it revealed a number of tertiary interactions (Figures 2 and 3). In particular, it showed how the T-ψ-C and D loops interacted in order to stabilize the corner of the molecule.[12] In 1974, Robertus et al.[13] also presented 3-Å results for the same spermine-stabilized yeast $tRNA^{Phe}$ in the monoclinic crystal lattice. This showed a virtually identical folding of the molecule. Although that analysis did not reveal the interactions of the T-ψ-C and D loops, further study at higher resolution showed that they were the same as in the orthorhombic crystal. A subsequent analysis of the monoclinic crystal[14] again confirmed that the folding of the molecule was virtually identical in the orthorhombic and monoclinic crystal lattices except for slight differences at the C-C-A end of the chain.

It can be seen in Figure 3 that many of the base–base tertiary hydrogen-bonding interactions involve the invariant or semiinvariant nucleotides. This suggests that the structure of yeast $tRNA^{Phe}$ may indeed be a generalized model for understanding the structure of all tRNAs, as discussed below.

At the present time the analysis has been extended to 2.5-Å resolution, and crystallographic refinement calculations now reveal a fair number of details concerning the tertiary interactions in the molecule.[15] These results are illustrated diagrammatically in Figures 2 and 3, while Figure 4 shows a stereoview of the molecule.

The RNA Double Helix

The major structural unit in yeast $tRNA^{Phe}$ is the RNA double helix which comes from the cloverleaf

(8) S. H. Kim, G. Quigley, F. L. Suddath, A. McPherson, D. Sneden, J. J. Kim, J. Weinzierl, and A. Rich, *J. Mol. Biol.*, **75**, 429 (1973).
(9) S. H. Kim, G. Quigley, F. L. Suddath, A. McPherson, D. Sneden, J. J. Weinzierl, P. Blattmann, and A. Rich, *Proc. Natl. Acad. Sci. U.S.A.*, **69**, 3746 (1972).
(10) R. W. Schevitz, M. A. Navia, D. A. Bantz, G. Cornick, J. J. Rosa, M. D. H. Rosa, and P. B. Sigler, *Science*, **177**, 429 (1972).
(11) S. H. Kim, G. J. Quigley, F. L. Suddath, A. McPherson, D. Sneden, J. J. Kim, J. Weinzierl, and A. Rich, *Science*, **179**, 285 (1973).
(12) S. H. Kim, F. L. Suddath, G. J. Quigley, A. McPherson, J. L. Sussman, A. Wang, N. C. Seeman, and A. Rich, *Science*, **185**, 435 (1974).
(13) J. D. Robertus, J. E. Ladner, J. T. Finch, D. Rhodes, R. S. Brown, B. F. C. Clark, and A. Klug, *Nature (London)*, **250**, 546 (1974).
(14) C. Stout, H. Mizuno, J. Rubin, T. Brennan, S. Rao, and M. Sundaralingam, *Nucleic Acid Res.*, **3**, 1111 (1976).
(15) G. J. Quigley and A. Rich, *Science*, **194**, 796 (1976).

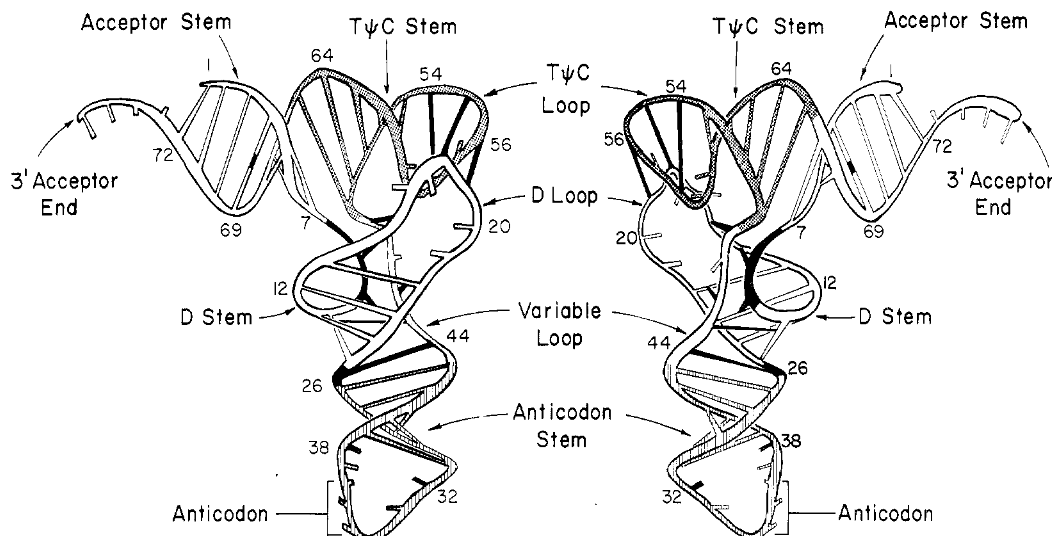

Figure 2. A schematic diagram showing two side views of yeast $tRNA^{Phe}$. The ribose phosphate backbone is depicted as a coiled tube, and the numbers refer to nucleotide residues in the sequence. Shading is different in different parts of the molecules, with residues 8 and 9 in black. Hydrogen-bonding interactions between bases are shown as solid black rungs, which indicate either one, two, or three hydrogen bonds between them as described in the text. Those bases that are not involved in hydrogen bonding to other bases are shown as shortened rods attached to the coiled backbone.

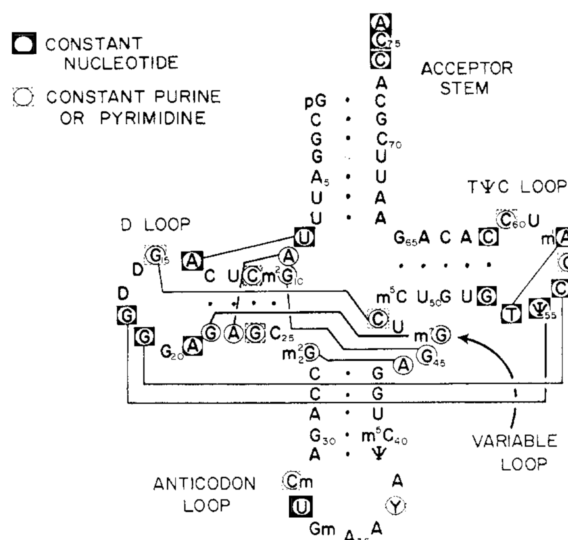

Figure 3. Cloverleaf diagram of the nucleotide sequence of yeast $tRNA^{Phe}$. Tertiary base–base hydrogen-bonding interactions are shown by solid lines, which indicate one, two, or three hydrogen bonds. The invariant and semiinvariant positions are indicated by solid and dashed boxes around the bases. Y_{37} is a hypermodified base.

stems. These are very close to an RNA helix with approximately eleven residues per turn. This conformation is maintained due to the presence of the ribose 2′-OH group in the backbone. DNA lacks this group, and its double helix is substantially different.

The presence of a 2′-hydroxyl group in the ribose phosphate backbone contributes to the stabilization of the standard ribose 3′-endo conformation which is found generally in RNA molecules. Its absence in DNA leads to a stable deoxyribose 2′-endo conformation. This conformation difference is responsible for the differences in the overall shape and form of the RNA vs. DNA double helix. The RNA double helix differs from the familiar DNA double helix in that the base pairs are not perpendicular to the helix axis but rather are tilted about 15°. Furthermore, they are not found on the axis of the helix but rather are displaced away from the center. There is a hole approximately 6 Å in diameter which goes down the center of the RNA double helix so that the molecule looks more like a band wrapped around an imaginary cylinder rather than a double helical twisted molecule which fills its axis. In consequence, the deep groove of the RNA double helix is extremely deep, whereas the narrow groove is extremely shallow. This is seen in the stereoview of the molecule in Figure 4.

The helix containing the acceptor stem and T-ψ-C stem seems almost uninterrupted. The sequence of yeast $tRNA^{Phe}$ (Figure 3) shows a G·U pair in the acceptor stem. This introduces only a slight perturbation in the helix. Examination of the electron density map at 2.5 Å (Figure 5a) shows these bases are held together by two hydrogen bonds in a typical "wobble" pairing.[16]

Base Stacking and Hydrogen Bonding

An outstanding feature of yeast $tRNA^{Phe}$ is the fact that most of the bases are involved in base stacking as well as hydrogen-bonding interactions. That this involves the bases in the double-helical stem regions was of no great surprise, but it was of considerable interest to find base-stacking interactions involving bases in the nonhelical loop regions. Virtually all of the bases in the molecule are organized into two large base stacking domains along each of the two limbs of the L-shaped molecule (see Figures 2 and 4). The horizontal stacking domain in Figure 2 includes bases of the acceptor and T-ψ-C stems as well as many of the bases of the T-ψ-C loop and some from the D loop. Most of the remaining bases are involved in the vertical stacking domain which extends down to and includes bases in the anticodon loop. Only 5 of the 76 bases are not involved in stacking interactions.

At the end, the 3′-terminal adenosine-A76 is unstacked. The remaining four unstacked residues are found in the regions of the cloverleaf diagram which contain variable numbers of nucleotides in different tRNA species.[1] The two dihydrouracil residues, D16

(16) F. H. C. Crick, *J. Mol. Biol.*, 19, 458 (1966).

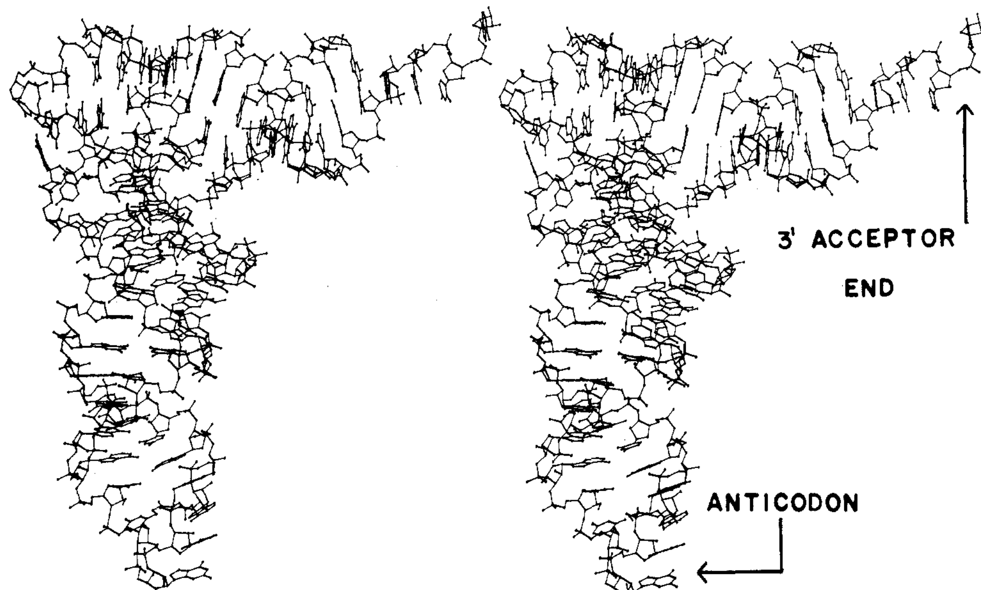

Figure 4. Stereoscopic diagram of yeast $tRNA^{Phe}$.[15] This diagram can be seen in three dimensions by using stereoscopic glasses. However, the three-dimensionality of the diagram can be seen without glasses if the reader simply lets the eye muscles relax and allows the eyes to diverge slightly, so that the two images are superimposed. Reprinted with permission from ref 15. Copyright 1976 by the American Association for the Advancement of Science.

and 17 of the α regions and G20 of the β region, are not stacked, nor is U47 in the variable loop region. The α and β regions are in the D loop.[1]

It is interesting that the bases not involved in stacking interactions are mostly found in segments of the polynucleotide chain which do not have constant numbers of nucleotides in all tRNA molecules. This suggests that there is a structural explanation for this variability. The bases in the α and β regions[1] as well as one of the bases in the variable loop do not extend into the molecule, but rather project outward. They also occur in regions in which there is a bulging out or arching of the backbone polynucleotide chain. This arching makes it possible to accommodate variable numbers of nucleotides in different tRNA molecules; a larger arch is likely to be found for those containing larger numbers of nucleotides or a smaller arch when there are fewer. In the case of variable loop, it is most likely that U47 is eliminated for those tRNAs containing four nucleotides in the variable loop. Thus the structure provides information relative to the question of whether the yeast $tRNA^{Phe}$ structure can serve as a model for all tRNAs. In this case, something special in the structure is found in regions of varying nucleotide number.

The Corner of the Molecule

There are two regions in the yeast $tRNA^{Phe}$ molecule in which there is moderate structural complexity. One is at the corner of the molecule where the T-ψ-C and D loops interact to stabilize the L-shaped conformation. The other is in the core of the molecule near the D stem in which there is considerable additional hydrogen bonding.

The T-ψ-C loop is organized in such a manner as to stabilize its interaction with the D loop and thereby maintain the two limbs of the molecule approximately at right angles to each other. This is accomplished using both detailed hydrogen bonding and base-stacking interactions. All of the bases in the T-ψ-C loop are stacked parallel to the bases in the T-ψ-C stem with the exception of U59 and C60 which are excluded from the stacking interactions by a number of hydrogen bonds. However, the bases U59 and C60 are oriented at right angles to the rest of the bases in the loop where they serve to nucleate the stacking interactions on which the vertical stacking domain in Figure 4 is built.

Hydrogen bonding in the T-ψ-C loop has many unusual features, as seen in the electron-density sections of Figure 5. T54 forms a hydrogen-bonded pair with m^1A58, but this pair does not use conventional Watson–Crick pairing as in the double helical segments; rather, the pairing involves the imidazole nitrogen (N.7) of the adenine residue (Figure 5b). This type of pairing has been observed for several purine–pyrimidine complexes,[17] initially by Hoogsteen.[18] The base plane next to this has a complex interaction in which ψ55 is hydrogen bonded both to G18 and to the phosphate group P58 (Figure 5c). Oxygen-4 of ψ55 is within hydrogen-bonding distances of both N.1 and the amino group N.2 of G18. The hydrogen-bonding interaction between ψ55 and P58 stabilizes the sharp turn of the polynucleotide chain in this corner of the molecule.

There is an unusual stacking of three guanine residues with G57 inserted between G18 and G19. G19 forms a tertiary hydrogen-bonding interaction with C56. This is the only Watson–Crick interaction among the nine tertiary base–base hydrogen-bonding interactions. As shown in Figure 5d, G57 has several hydrogen-bonding interactions with the backbone, two of which involve interactions with the ribose-18 and ribose-19. N7 of G18 receives a hydrogen bond from the 2′-hydroxyl group of ribose-55. Position 57 is normally occupied by a purine, and hence the latter hydrogen-bonding interaction can be preserved in other tRNA species. However, the major stabilization which accrues from the presence of a purine in position 57 is due to the enhanced stacking interaction when it is intercalated between the two constant guanine residues at

(17) D. Voet and A. Rich, *Prog. Nucleic Acid Res. Mol. Biol.*, **10**, 183 (1970).
(18) K. Hoogsteen, *Acta Crystallogr.*, **16**, 907 (1968).

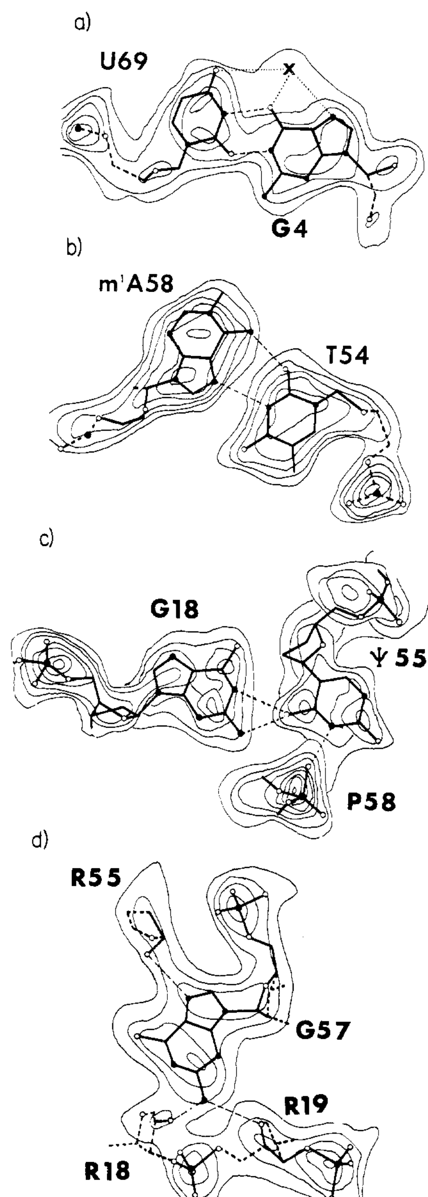

Figure 5. Electron density sections from the acceptor stem and the corner of yeast tRNA^{Phe}.[15] Curved lines represent contours of electron density; heavier solid lines represent the molecular model as defined by the refined atomic coordinates. Hydrogen bonds between the bases are shown as thin, dashed lines. Heavier dashed lines are used to show segments of the polynucleotide chain which are outside the plane of the section. Oxygen atoms are shown as small open circles, nitrogen atoms as small solid circles. The phosphorus atoms are slightly larger solid circles. Some diagrams show more than one section, as the segments of the molecule that are illustrated do not all lie in one plane. In these cases, parallel sections 1 Å apart are used. The position marked X in diagram a corresponds to an ion or solvent molecule which may be coordinating with the bases. Reprinted with permission from ref 15. Copyright 1976 by the American Association for the Advancement of Science.

positions 18 and 19. This intercalation is accomplished by having the ribose phosphate backbone along 18 and 19 in an unusual 2′-endo conformation. That has the effect of separating the bases from each other, in contrast to the 3′-endo conformation which generally prevails.

The unusual orientation of residues U59 and C60 is maintained by two different hydrogen-bonding interactions. One of these involves a hydrogen bond between the N.4 amino group of C61 and the phosphate group of C60. This is undoubtedly the reason for the constant G53·C61 base pair at the end of the T-ψ-C stem. The second interaction is a hydrogen bond between the 2′-hydroxyl group of ribose 58 and the phosphate group of P60. In this conformation it stabilizes the looping out or arching of residue C60 away from the polynucleotide chain so that it is in a position where the pyrimidine ring can stack along the vertical axis of Figure 4 rather than with the other bases of the T-ψ-C loop.

In this corner of the molecule many features have the overall effect of bringing about a rather intricate fitting together of different components of the molecule so as to stabilize the unusual conformation. Broadly, the double helix has been engineered to make a right angle turn.

Core of the Molecule and the Anticodon

The core of the molecule immediately beneath the T-ψ-C stem in Figure 2 includes some of the most complex hydrogen-bonding interactions. As seen in Figure 2, four segments of polynucleotide chains come into this region; two of the chains are part of the D stem and the others are from the variable loop and the section of chain joining the acceptor stem to the D stem (residues 8 and 9). We can examine the hydrogen bonding from the top down, as seen in Figure 2, by looking at sections which are generally horizontal. Uppermost is the pairing between G18 and C48 which stacks immediately beneath the residue U59. As shown in Figure 6a, these bases are hydrogen bonded together through a trans pairing rather than the cis pairing which is normally found in C·C base pairs. This pairing is imposed because the polynucleotide chains are running in the same direction in this section of the molecule.

Immediately below that is the hydrogen bonding of U8 and A14 as shown in Figure 6b. This is a reversed Hoogsteen hydrogen bonding,[17,18] and residue A21 is located nearby where it interacts with the 2′-hydroxyl of ribose-8. Immediately below this is a complex system of hydrogen bonding in which the C13·G22 base pair of the D stem has hydrogen-bonding interactions with two other elements: one involving m^{7}G46 of the variable loop and the other involving phosphate-9, as shown in Figure 6c. Four polynucleotide chains are seen in this section, and they are all hydrogen bonded together. At this point the molecule is stabilized by the positive charge which is found on the m^{7}G46 base due to methylation in position 7. This positive charge makes for stronger hydrogen bonding by N.1 and N.2 of the guanine residue, and it also stabilizes the molecule electrostatically because of the close proximity of phosphates-9 and -22.

In the section immediately below (Figure 6d), the base pair U12·A23 of the D stem has residue A9 hydrogen bonding into its major groove. The hydrogen bonding is the same as between two adenine residues in double-helical polyriboadenylic acid.[19]

Other base–base tertiary interaction in this region includes hydrogen bonding between G45 and the C25·m^{2}G10 base pair of the D stem (Figure 6e). Immediately below this is an unusual hydrogen bonding between A44 and m$_2$^{2}G26 which are paired, as shown

(19) A. Rich, D. R. Davies, F. H. C. Crick, and J. D. Watson, *J. Mol. Biol.*, 3, 71 (1961).

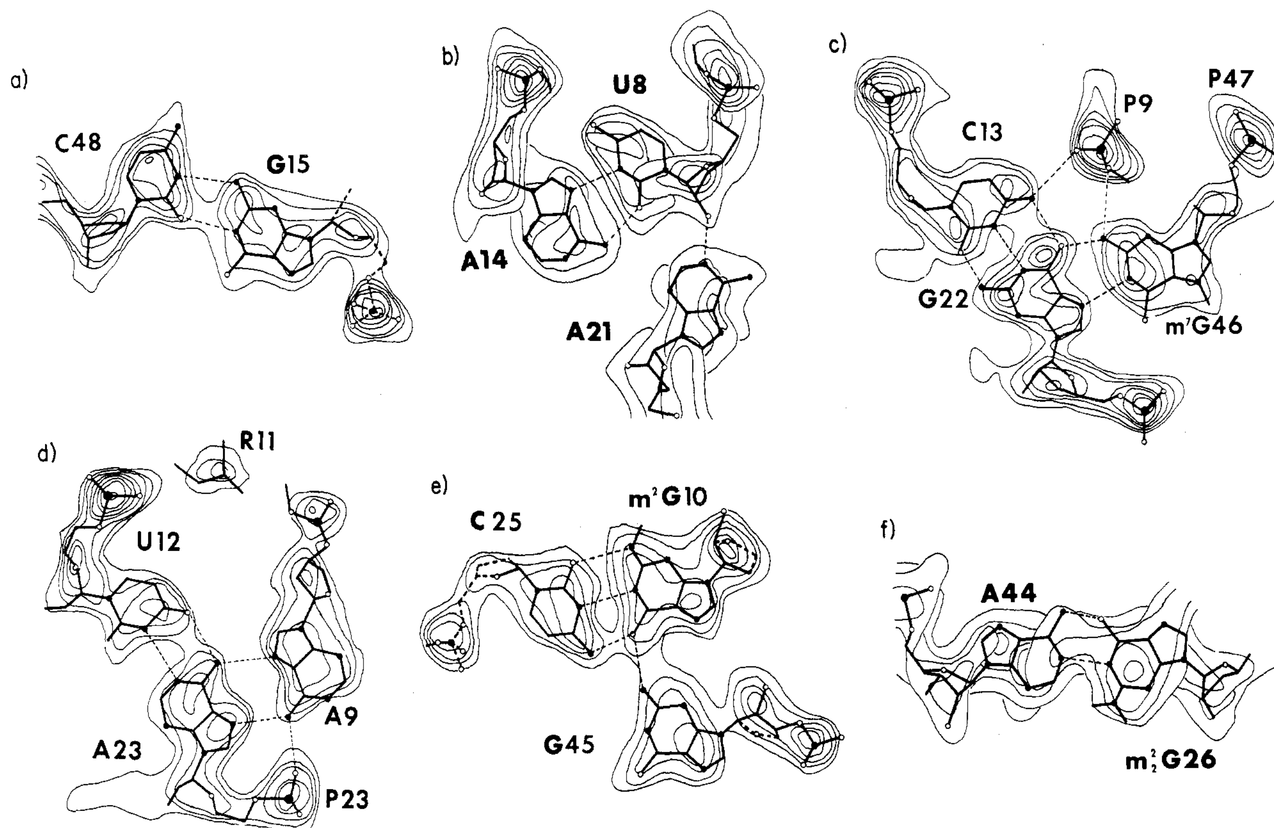

Figure 6. Electron density sections from the core region of yeast $tRNA^{Phe}$.[15] See legend of Figure 5 for description.

in Figure 6f, with the bases tilted rather sharply in a propeller-like conformation. This has the effect of maximizing the stacking interaction of $m_2{}^2G26$ with base-pair C25·m^2G10 immediately above it (in Figure 2) while the opposite tilting of A26 maximizes its stacking with the base pair below it, C27·G43, which is the uppermost pair of the anticodon stem. In this particular instance, dimethylation of the amino group of guanine contributes to the propeller-like twisting of the two hydrogen-bonded bases; in turn, this stabilizes the stacking interaction of the two paired bases, each with a structurally different part of the molecule.

The anticodon stem is an RNA double helix with about 11 residues per turn. The conformation of the bases in the anticodon loop is somewhat similar to that suggested by Fuller and Hodgson[20] in that the three anticodon bases are at the end of a stacked series of bases at the 3′ end of the loop. The constant residue U33 plays an interesting role in maintaining the conformation of the loop since it is hydrogen bonded through N.3 to the phosphate group of P36. Indeed, there is great similarity between the conformation of the polynucleotide chain in the region of the T-ψ-C loop and the conformation in the anticodon loop.[15] In both places, the polynucleotide chain makes a sharp bend and a uridine residue (U33 or ψ55) plays a key role in stabilizing it through the formation of a hydrogen bond to a phosphate residue on the other side of the loop (see Figure 5c). We have termed this common structure the uridine turn or U turn of the polynucleotide chain. In both cases, the hydrophobic stacking of the uridine residue is terminated by a phosphate group which is in van der Waal's contact with the uridine ring. This suggests the possibility of additional stabilization due to an ion–induced-dipole interaction.

(20) W. Fuller and A. Hodgson, *Nature (London)*, 215, 817 (1967).

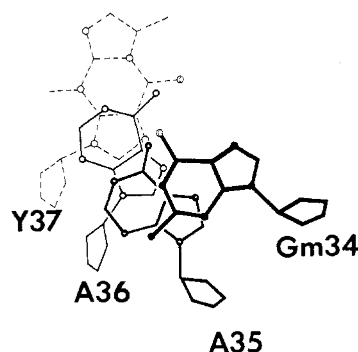

Figure 7. Diagram illustrating the stacking of anticodon bases 34 to 36 on the hypermodified Y base as viewed from the exterior of the molecule. Oxygen atoms are stippled circles, and nitrogen atoms are unstippled circles. Not all of the side chain of the Y base is included. It can be seen that the anticodon takes the form of a right-handed helix.

The configuration of the anticodon bases is shown in Figure 7 as viewed from the bottom of the molecule. It can be seen that the three anticodon bases have the form of a right-handed helix with approximately eight residues per turn. They are in a conformation such that they can form hydrogen-bonding interactions with the codon. However, it is not clear whether the anticodon conformation seen in the crystal is maintained when it interacts with messenger RNA.

Role of the Ribose 2′-Hydroxyl Group

Unlike DNA, RNA has a systematic 2′-OH group on its sugar backbone. In the double-helical regions of the yeast $tRNA^{Phe}$ molecule, the 2′-hydroxyl groups are not involved systematically in hydrogen-bonding interac-

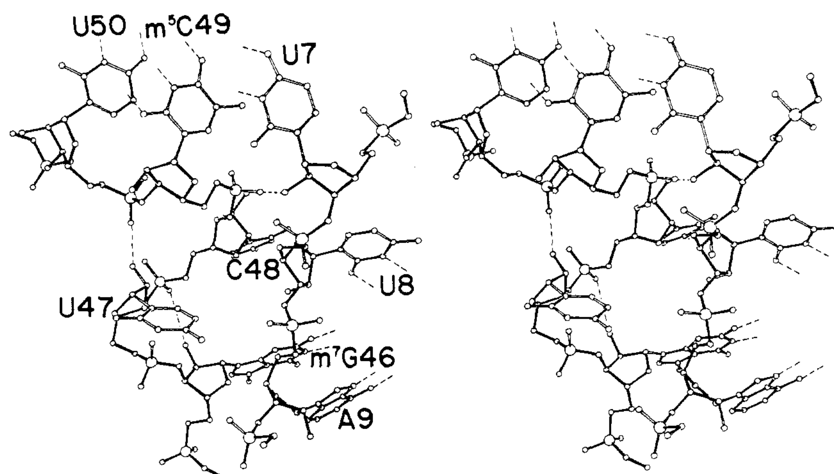

Figure 8. Stereoscopic view of the juncture between the acceptor and T-ψ-C stems and the variable loop near the core of the molecule. This region is near the inside of the L-shaped molecule where it turns the corner. The axis of the acceptor stem and T-ψ-C stem is horizontal across the top of the diagram. The molecule is roughly in the same orientation as that shown in Figure 4 except that the acceptor stem is rotated toward the reader. The coiling of the polynucleotide chain is stabilized by several hydrogen bonds between 2′-hydroxyl groups and phosphate residues, as described in the text. Dashed lines attached to the bases imply that they are hydrogen bonding to other bases, which are not shown.

tions. However, about 50% of the 3′-hydroxyl groups in the nonhelical segments are involved in hydrogen bonding.[15] Some of these participate in base–backbone hydrogen bonding while others are involved in backbone–backbone hydrogen bonding. In the latter group, an interesting class of interactions are those in which the 2′-hydroxyl group bonds to the oxygen of a phosphate group. This has the effect of stabilizing an arch conformation in which the polynucleotide chain loops out and then returns. An example is the conformation around residue C60, discussed above.

As mentioned, there are several places where the polynucleotide chain loops out in such a way that bases project away from the remainder of the molecule. This loop or arch conformation is often stabilized by a hydrogen bond between the ribose 2′-hydroxyl group at one side of the arch and the phosphate group at the other. Three examples are seen in Figure 8. This shows a stereoview of the juncture between the acceptor and the T-ψ-C stems and the variable loop in the core of the molecule. One hydrogen bond is between the 2′-hydroxyl of ribose 7 and the phosphate of m^{5}C49. This hydrogen bond in effect maintains the continuity of the double helix by joining one of the double-helical backbones connecting the acceptor and T-ψ-C stems. What arches out at that juncture is the remainder of the molecule, forming the entire vertical stacking domains of Figures 2 and 4.

Another example in Figure 8 is the stabilizing hydrogen bond between ribose-46 and phosphate-48. This has the effect of neatly excluding nucleotide-47 from the remainder of the molecule so that its base is not involved in stacking interactions and is readily accessible. Probably this nucleotide is not present in those tRNAs which have only four nucleotides in the variable loop instead of the five which are found in yeast tRNA^{Phe}.

A third example in Figure 8 is the hydrogen bond between the ribose 2′-hydroxyl of U47 and the phosphate-50. This stabilizes the rather sharp arch of the polynucleotide chain between the T-ψ-C stem and the variable loop. There are also other examples of this type of arch conformation in the molecule in which the 2′-hydroxyl group plays a role in stabilizing a nonhelical portion of the polynucleotide chain.[15]

Thus the ribose 2′-hydroxyl group has special structural roles in the nonhelical or irregular coiling of the polynucleotide chain. It contributes to overall stabilization of the molecule either by stabilizing the arch conformation or by contributing to a variety of backbone–base hydrogen-bonding interactions.

Structure of the Molecule in Solution Compared to the Crystal

The orthorhombic crystal of yeast tRNA^{Phe} contains approximately 70% water. This suggests that the transformation from the crystalline state to aqueous solution is not likely to have enormous structural consequences. Available evidence from a variety of experimental techniques strongly supports this interpretation. The structure appears to be the same in solution as it is in the crystalline state.[4] It should be noted, however, that not all of the techniques which have been used are capable of making precise comparisons.

An important technique for comparison is chemical modification, in which reagents are used to modify individual purines or pyrimidines, but only if they are accessible, i.e., not hydrogen bonded or stacked. The reactions are frequently nucleotide specific and, as such, they make possible a precise correlation between structure and chemical reactivity. In general there is an excellent correlation between solution chemical modification studies on yeast tRNA^{Phe} and the observed three-dimensional structure.[4] The bases not involved in stacking interactions are strongly reactive. Some bases which are partially obscured or partially inaccessible through hydrogen bonding are somewhat less reactive. These studies have also been carried out on a variety of other tRNA species, and the general pattern of reactivity is similar to that in yeast tRNA^{Phe}. This, of course, strengthens the probability that the three-dimensional structure of other tRNA molecules is very similar to that for yeast tRNA^{Phe}.[4]

Another important technique is that of high-resolution NMR studies.[21] The results obtained by

high-resolution NMR study of dissolved tRNA correlate very well with the structure from x-ray crystallography. Simulated high-resolution NMR spectra generated by calculations starting from refined atomic coordinates closely resemble the observed spectra, not only for yeast $tRNA^{Phe}$, but for other species as well. This further suggests that the structure is the same in solution as in the crystalline state, especially because these calculations are very sensitive to small changes in the overlapping geometry of the bases.

Other techniques which have been used include the binding of oligonucleotides to tRNA, laser Raman spectroscopic analysis, and the susceptibility to a variety of nucleases.[4] These experiments also show good correlation between the structure observed in the crystal and the results obtained by solution studies. It should be pointed out, however, that some conformational features of the molecule in solution might not be detectable by some of the techniques.

Generalized Structure of tRNA Molecules

As mentioned, there is good reason to believe that the molecular structure of yeast $tRNA^{Phe}$ can be used as a model for understanding the molecular structure of all of the tRNAs.[21] Differences in the number of nucleotides in different tRNA sequences can be accommodated by variable sized arches looping out of the molecule in the α and the β region of the D loop[1] or in the variable loop for those tRNAs which have four or five nucleotides. However, some uncertainty exists about the conformation of those tRNAs that have very large extra loops containing 13 to 21 nucleotides. These undoubtedly form a stem and loop structure which projects away from the diagonal region of the molecule, as seen in Figure 2. Further work will be necessary before the details are known.

An interesting question concerns the conformation of initiator tRNAs. These have slightly different nucleotides in the invariant or semiinvariant positions, especially in the T-ψ-C loop.[4] Probably these sequence differences cause small differences in the conformation of the loop. The differences can not be very large in view of the close similarity between patterns of chemical modification for initiator tRNAs and chain-elongating tRNAs.

There are likely to be some altered forms of hydrogen bonding which are found in other sequences.[22] Those hydrogen bonds in yeast $tRNA^{Phe}$ which do not involve invariant or semiinvariant positions are probably altered in other tRNA molecules. In several cases it is possible to guess the detailed nature of these modifications, but many will need to be determined by further experimentation.

The tRNA Molecule in Protein Synthesis

The aminoacylating enzymes each add an amino acid to a selective group of tRNAs, yet all of these aminoacyl tRNAs pass through the same ribosomal apparatus. What is the mechanism for differentiation between different tRNA species during aminoacylation and how do all of the aminoacyl tRNAs pass through the same ribosomal apparatus? In short, where are the components of uniqueness and the components of commonality?

There has been considerable effort to understand the mechanism of aminoacylation and to find those regions of the tRNA molecule which may be recognized by synthetase enzymes. Different workers have suggested the acceptor stem, the D stem, or the anticodon.[23] Much experimental work on synthetase-tRNA interactions is in accord with the suggestion that synthetases recognize varying aspects of the tRNA structure along the diagonal side of the molecule,[24,25] i.e., the outer edges in Figure 2. The synthetases are probably all different, and there is not likely to be a common recognition system even if they all approach the same side of the tRNA molecule.

We now know in detail the three-dimensional structure of one synthetase,[26] yet it is still not possible to describe the molecular basis of recognition. For example, the tRNA or the synthetase may undergo a conformational change on forming the complex. Thus one has to have an open mind about the possible recognition mode. It is likely, however, that this recognition takes two parts. One is a recognition of the ribose phosphate chain which would be sensitive to the folding of the tRNA molecules. Secondly, there must be a recognition by the protein of specific bases in the double-helical stems or among the unpaired segments of the molecule. There are several ways in which proteins can recognize nucleic acid sequences.[27] It is likely that the basis of specificity resides in this detailed sequence recognition rather than in any conformational differences between tRNA molecules. Furthermore, there is an adequate basis for specificity in these sequences. The number of nucleotides or base pairs recognized by the enzyme need not be very large in order to obtain the requisite specificity.

Far more puzzling is the question of what goes on in the ribosome. Our information in this field is scanty. For example, there are two sites within the ribosome, one of which is occupied by the peptidyl-tRNA and the other by the aminoacyl-tRNA. There are indications that these two sites are physically distinct. However, it is not known with certainty whether the adjacent codons at these two sites are both occupied by tRNA molecules at the same time. This knowledge would of course influence our interpretation of tRNA movement from one site to the other.

Perhaps both sites are occupied at the same time. For this reason the tRNA molecule may have been designed to have the form of a double helix that turns a corner. The L shape of the molecule may make it possible for two adjacent tRNAs to come close together at one end near their anticodons while they are interacting with adjacent codons of the messenger RNA. At the same time the C-C-A acceptor ends may also be able to come close together due to the L shape. These two C-C-A ends must come close together in order to allow the ribosomal peptidyl transferase to transfer the

(21) R. B. Reid and R. E. Hurd, *Acc. Chem. Res.*, **10** (1977), third paper in this issue.

(22) S. H. Kim, J. L. Sussman, F. L. Suddath, G. J. Quigley, A. McPherson, and A. Rich, *Proc. Natl. Acad. Sci. U.S.A.*, **71**, 4970 (1974).

(23) L. L. Kisselev and O. O. Favorova, *Adv. Enzymol.*, **40**, 141 (1974).

(24) A. Rich and P. R. Schimmel, *Nucleic Acids Res.*, **4**, 1649 (1977).

(25) P. R. Schimmel, *Acc. Chem. Res.*, **10**, (1977), fifth paper in this issue.

(26) M. J. Irwin, J. Nyborg, B. R. Reid, and D. M. Blow, *J. Mol. Biol.*, **105**, 577 (1976).

(27) N. C. Seeman, J. M. Rosenberg, and A. Rich, *Proc. Natl. Acad. Sci. U.S.A.*, **73**, 804 (1976).

peptide chain from one tRNA molecule to the other. It is not at all clear how this is accomplished, although it has been suggested that the two codons become unstacked during the reading process so that in effect the messenger "turns a corner" while it is read.[28] This remains an exciting area for further research work.

Important work by Erdmann and his colleagues[29] has suggested that the T-ψ-C loop may become disengaged from the D loop inside the ribosome so that it can then hydrogen bond with the 5S RNA of the ribosome. This interaction may be an important component in the translocation of tRNA from aminoacyl site to the peptidyl site. If this is true, the tRNA molecule undergoes a conformational change within the ribosome. Determination of the nature of these events remains an important research goal. It is possible that this conformational change is triggered by codon–anticodon interactions, and this may be a consequence of the fact that the tRNA molecule as a whole exhibits long-range order.[4] Thus interactions at one end may give rise to a change in reactivity or in conformation at a more remote part of the molecule.

The end product of biological crystallography is the determination of three-dimensional structure. However, the major problem to be solved is that of understanding biological function. In this regard, structural information is particularly valuable in generating reasonable models which can be explored to explain aspects of function. In tRNA research we now have a firm view of the three-dimensional structure of the molecule. What remains to be understood is how this three-dimensional structure is used to carry out the various functions of tRNA in biological systems.

Elucidation of the structure of tRNA involved the work of many colleagues as cited in the bibliography, and I wish to acknowledge the important role which they played in this effort. I thank especially S. H. Kim, G. Quigley, and F. L. Suddath, who worked for many years on this project. Research for this work was supported by grants from the National Institutes of Health, National Science Foundation, National Aeronautics and Space Administration, and the American Cancer Society.

(28) A. Rich in "Ribosomes", M. Nomura, A. Tissières, and P. Lengyel, Ed., Cold Spring Harbor Laboratory, New York, N.Y., 1974, p 171.

(29) V. A. Erdmann, M. Sprinzl, and O. Pongs, *Biochem. Biophys. Res. Commun.*, 54, 942 (1973); D. Richter, V. A. Erdmann, and M. Sprinzl, *Proc. Natl. Acad. Sci. U.S.A.*, 71, 3326 (1974).

Application of High-Resolution Nuclear Magnetic Resonance Spectroscopy in the Study of Base Pairing and the Solution Structure of Transfer RNA

Brian R. Reid* and Ralph E. Hurd

Biochemistry Department, University of California, Riverside, California 92521

Received February 17, 1977

The functional aspects of tRNA and its structure as revealed by x-ray crystallography have already been described in this issue and elsewhere.[1-3] The crystallographic approach yields a static structure of tRNA in the crystalline solid state with unrivaled precision. However, the molecule functions biologically in the liquid state, and it is important to ask if the solution conformation (or conformations) is related to the crystal structure and whether there are important dynamic conformational equilibria in the solution structure. We have addressed ourselves to these questions using high-resolution nuclear magnetic resonance (NMR) as a spectroscopic probe of the structure of tRNA in solution.

Although there are several classes of protons in the various regions of the magnetic resonance spectrum of nucleic acids, perhaps the most informative spectral window is the extreme low-field region between −11 and −15 ppm downfield from the standard DSS (2,2-dimethylsilapentane-5-sulfonate) reference. We shall refer to this region as the low-field spectrum. Early NMR studies on model nucleosides in aprotic solvents by Katz and Penman[4] revealed that the ring NH protons of guanosine (N1H) and uridine (N3H) are highly deshielded and move even further downfield in the low-field spectrum upon forming hydrogen-bonded base pairs with their complementary nucleosides.

Although the ring NH protons of free monomers are rapidly exchangeable with water (and thus time-average into the large H_2O peaks at −4.65 ppm in aqueous solvents), in polymers such as tRNA the base–base hydrogen bonds have long lifetimes (seconds) in the interior of the water-inaccessible helices.[5]

These considerations led Kearns and Shulman and their collaborators to attempt the first low-field NMR

Brian R. Reid received his B.A. degree from Cambridge University and his Ph.D. from the University of California, Berkeley. He then held a Jane Coffin Childs postdoctoral fellowship at Dartmouth Medical School and in 1966 joined the Biochemistry Department at the University of California, Riverside, where he is Professor of Biochemistry. His research interests include structural studies on transfer RNA and aminoacyl-tRNA synthetase enzymes using a variety of physical methods, and the mechanism of protein–nucleic acid recognition.

Ralph E. Hurd received his B.A. degree from California State College, San Bernardino, in 1973. He is currently carrying out doctoral thesis research in the Biochemistry Department, University of California, Riverside, on the solution structure and conformational dynamics of regulatory and nonregulatory mutant tRNAs.

(1) S. H. Kim, F. L. Suddath, G. J. Quigley, A. McPherson, J. L. Sussman, A. H. J. Wang, N. C. Seeman, and A. Rich, *Science*, 185, 435 (1974).

(2) J. D. Robertus, J. E. Ladner, J. T. Finch, D. Rhodes, R. S. Brown, B. F. C. Clark, and A. Klug, *Nature (London)*, 250, 546 (1974).

(3) J. L. Sussman and S. H. Kim, *Science*, 192, 853 (1976).

(4) L. Katz and S. Penman, *J. Mol. Biol.*, 15, 220 (1966).

(5) S. W. Englander and J. J. Englander, *Proc. Natl. Acad. Sci. U.S.A.*, 53, 370 (1965).