bias in bioecology has been excessive [C. S. Holling, *Diversity and Stability in Ecological Systems* (report and symposium held 26–28 May 1969 at Brookhaven National Laboratory, Upton, N.Y.)].
21. A. Alland, *Evolution and Human Behavior* (Doubleday, Garden City, N.J., 1973).
22. G. Hardin, *Science* **162**, 1243 (1968).
23. H. G. Barnett, *Innovation: The Basis of Culture Change* (McGraw-Hill, New York, 1953).
24. R. Firth, *J. R. Anthropol. Inst.* **85**, 1 (1955).
25. For a comparable listing, see H. A. Selby, *Biennial Review of Anthropology*, B. Siegel and A. Beals, Eds. (Stanford Univ. Press, Stanford, Calif., 1972)].
26. R. Redfield, *The Primitive World and Its Transformations* (Cornell Univ. Press, Ithaca, N.Y., 1953).
27. A random selection of articles on instrumental cultural anthropology is as follows. F. Barth, *Am. Anthropol.* **69**, 661 (1967); M. D. Sahlins, in *Essays in Economic Anthropology*, J. Helm, Ed. (American Ethnological Society and Univ. of Washington Press, Seattle, 1964), p. 95; J. W. Bennett, *Southwest. J. Anthropol.* **24**, 276 (1968); R. E. Rhoades and S. I. Thompson, *Am. Ethnol.* **2**, 535 (1975); J. Westermeyer, *Am. Anthropol.* **75**, 123 (1973); R. M. Netting, in *Annual Review of Anthropology*, B. Siegel, Ed. (Annual Reviews, Palo Alto, Calif., 1974), p. 21; T. H. Hay, *Am. Anthropol.* **75**, 708 (1973); G. Britan and B. S. Denich, *Am. Ethnol.* **3**, 55 (1976); L. A. Despres, in *The New Ethnicity: Perspectives from Ethnology*, J. W. Bennett, Ed. (American Ethnological Society and West Publishing, St. Paul, Minn., 1975), p. 127; N. Whitten and D. Whitten, in *Annual Review of Anthropology*, B. Siegel, Ed. (Annual Reviews, Palo Alto, Calif., 1972), p. 247.
28. "Instrumental anthropology," with its concern for rational or purposive behavior, would appear to be concerned with phenomena similar to that contained in Daniel Bell's "techno-economic" domain; and my "interpretive anthropology" appears similar to that in Bell's "culture," or "expressive symbolic" domain of contemporary society. Bell's third domain, "polity," or the field of social control, is echoed in my emphasis on "policy" as a consequence of applying adaptational analysis to social behavior. However, I read Bell's book after completing this article and there has been no effort to bring concepts in line with his thesis [D. Bell, *The Cultural Contradictions of Capitalism* (Basic Books, New York, 1976)].
29. F. Boas, in *Freedom: Its Meaning*, R. N. Anshen, Ed. (Harcourt Brace, New York, 1940).
30. C. S. Belshaw, *The Sorcerer's Apprentice: An Anthropology of Public Policy* (Pergamon, New York, 1976).

# Three-Dimensional Structure of a Transfer RNA in Two Crystal Forms

## Analysis of three sets of atomic coordinates of yeast phenylalanine tRNA establishes common features.

Joel L. Sussman and Sung-Hou Kim

Transfer RNA (tRNA) plays a central role in decoding the genetic information in messenger RNA (mRNA) during protein biosynthesis. Recently, x-ray crystallographic studies on yeast phenylalanine tRNA ($tRNA^{Phe}$) revealed its intricate tertiary structure. Although the "complete" refinement of the yeast $tRNA^{Phe}$ structure will require a few more years, the three sets of currently published atomic coordinates of this tRNA in two different crystal forms are good enough to compare and draw conclusions about the structural features that are common.

The purpose of this article is, first, to critically analyze the three sets of published atomic coordinates, in order to determine the range of errors and the criteria used in defining structural features, especially hydrogen bonds, for each model; and second, to compare the three models so as to sort out those structural features that are present in all three models at high confidence level. The common structural features so obtained can provide a solid foundation for all studies on the structure-function relationship of tRNA.

There is a general tendency to freely accept x-ray crystallographic results of macromolecules despite the cautious statements investigators make. Such acceptance is usually safe for gross structural features such as backbone folding, secondary structures, and approximate coordination geometry around metal ions, but not for isolated features such as the existence or absence of a single hydrogen bond, small differences in conformational angles, or detail of coordination distances and symmetry. It is also not uncommon that the interpretation of electron density maps changes at successive stages of x-ray crystallographic studies. However, such changes become minor as the refinement proceeds.

These shortcomings can partially be overcome if one can compare several structures of the same molecule determined and refined by different groups, and then accept only those structural features that are common among them as reliable at high confidence level. Such is the case with yeast $tRNA^{Phe}$. For example, one can see in Fig. 1 a few changes in the assignment of tertiary base pairing at two different stages of refinement in each model. Although there may be more changes on further refinement, the gross differences among the three models have now disappeared. There are many minor differences (see below) at the present stage of refinement, but one should not take them as *real* until they are supported by other evidence. It is likely that most of these apparent differences will also disappear on further refinement of the three models.

The backbone structure of yeast $tRNA^{Phe}$ was revealed from an x-ray crystallographic study on the orthorhombic form at a resolution of 4 Å (*1*), and the preliminary tertiary interactions within the structure have been described for both orthorhombic (*2*) and monoclinic (*3*) crystal forms based on 3 Å resolution data and recently reviewed (*4*). Preliminary comparison of the two crystal forms, based on the structure factor amplitudes at 4 Å (*5*) and on the general appearance of the electron density maps at 3 Å resolution (*6*), suggested the similarity of the structures in both crystal forms. In the last few months, three sets of atomic coordinates have been reported for this tRNA, so that it is now possible to make a detailed, objective comparison between the structures; two sets of atomic coordinates for this tRNA in an orthorhombic crystal form have been obtained by two different refinement procedures with the use of the same diffraction data (*7*, *8*), and one for the same tRNA in a monoclinic crystal form by another procedure (*9*). All three procedures are different but related to each other with the common goal that the model obtained should fit the experimental data and known stereochemistry.

For convenience, the structures in the orthorhombic crystal form refined by the Duke group (*7*) and the MIT (Massachusetts Institute of Technology) group (*8*) are called A and B, respectively, and that in the monoclinic form refined by the MRC (Medical Research Council of Great Britain) group (*9*) is called C.

Dr. Sussman is a fellow of the Arthritis Foundation, in the Department of Biochemistry, Duke University School of Medicine, Durham, North Carolina 27710. Dr. Kim is an associate professor in the Department of Biochemistry, Duke University School of Medicine.

### Method of Analysis and Comparison

On the basis of the three sets of published atomic coordinates, the models have been analyzed by calculating interatomic distances, bond angles, and dihedral angles as well as the planarity of each base, and by examining them on an interactive computer graphics system.

At the current resolution of the data, individual atoms are not resolved in the electron density map, but rather the groups of atoms—such as phosphates, riboses, and bases—appear as individually distinguishable peaks. Therefore, the positions of these group centers are closer to the experimental data than the atomic coordinates, which are in fact derived quantities based on the relative positions and the overall appearance of individual electron density peaks. This is the reason for not comparing individual torsion angles.

In matching a pair of structures, we arbitrarily defined the centers of a phosphate, a ribose, and a base to be the position of phosphorus, the unweighted center of the five-membered ring of a ribose moiety and the unweighted center of the six-membered ring of a base, respectively. A least-squares procedure was used to translate and rotate one model above another iteratively (*10*) to obtain a best match with the use of these group-center coordinates. A virtually identical match was obtained when atomic coordinates were used instead.

### State of Structure Refinement

The structure refinements are still far from complete (Table 1). The first of the three classes shown in Table 1 contains experimental quantities or those derived

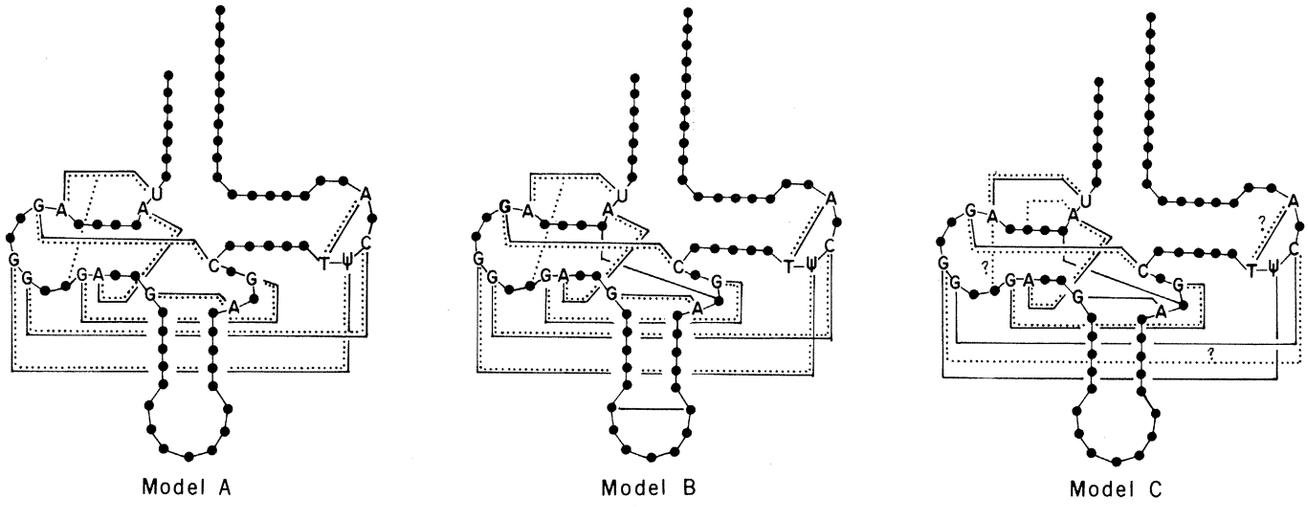


Fig. 1. Tertiary hydrogen bondings between bases as assigned from 3 Å resolution data for models A and B from (*2*, figure 4) and for model C from (*3*, figure 1) are shown in dotted lines. Those based on 2.7 Å data for models A (*7*) and B (*8*) and on 2.5 Å data for model C (*9*, *16*), are in solid lines. Nucleotide bases involved in the common tertiary H-bondings are labeled.

Table 1. Crystal data and quality of refinement and stereochemistry.

| | Model A (*7*) | | Model B (*8*) | | Model C (*9*, *16*) | |
|---|---|---|---|---|---|---|
| Space group | $P2_12_12_1$ | | $P2_12_12_1$ | | $P2_1$ | |
| a, b, c (Å) | 33, 56, 161 | | 33, 56, 161 | | 56, 33, 63 | |
| $\alpha, \beta, \gamma$ (degrees) | 90, 90, 90 | | 90, 90, 90 | | 90, 90, 90 | |
| Resolution of data (Å) | 2.7 | 3 | 2.7 | 3 | 2.5 | 3 |
| Unique reflections used (No.) | 8427 | 6180 | 8448 | 6172 | 6139 | 4838 |
| Reflections used per atom (No.)* | 5.10 | 3.74 | 5.12 | 3.74 | 3.72 | 2.93 |
| Mean figure of merit | | 0.66 | | 0.66 | 0.59 | 0.67 |
| R factor (percent)† | 39 | | 35‡ | | 39 | 47 |
| Correlation coefficient§ | .74 | .60 | | | | .57 |
| Deviations from average bond lengths‖ | +0.1 Å; C(1′) to O(1′) of U50<br>−0.1 Å; C(5) to C(6) of A35 | | +0.2 Å; O(5′) to C(5′) of G18<br>−0.2 Å; N(9) to C(4) of G10 | | +0.4 Å; C(3′) to O(3′) of A21<br>−0.1 Å; C(2′) to C(3′) of T54 | |
| Deviations from average bond angles‖ | +10°; C(4′) to O(1′) to C(1′) of C56<br>−10°; C(3′) to C(2′) to O(2′) of U8 | | +8°; N(3) to C(2) to O(2) of C2<br>−8°; O(3′) to P to OP of A35 | | +11°; C(3′) to O(3′) to P of C75<br>−20°; C(3′) to O(3′) to P of G22 | |
| Shortest nonbonded contact # | 2.6Å; O(5′) of G19 to O(1′) of G19 | | 1.7 Å; O(2′) of U47 to OP of C48 | | 2.0 Å; O(2′) of A21 to O(6) of G46 | |
| r.m.s. of worst base plane | 0.1 Å; A58 | | 0.2 Å; Y37 | | 0.3 Å; Y37 | |
| Range of H bonds** | 2.6 Å; O(6) of G10 to N(4) of C25<br>3.2 Å; N(2) of G19 to O(2) of C56 | | 2.3 Å; O(2) of C32 to N(6) of A38<br>4.2 Å; O(1′) of G19 to N(2) of G57 | | 2.5 Å; N(6) of A21 to O(1′) of C48<br>4.1 Å; N(4) of C61 to OP of C60 | |
| Range of glycosyl bond torsion angles†† | −37°; C56<br>+91°; D16 | | −165°; A44<br>+119°; A9 | | −47°; C60<br>+107°; A9 | |

*Not counting metal ions and bound water oxygens. †Crystallographic discrepancy factor (R) is equal to $[\Sigma|F \text{ (observed)} - kF \text{ (calculated)}|]/[\Sigma F \text{ (observed)}]$. The summation is overall data and $F$'s are structure factors and $k$ is a scaling function or a scaling constant. ‡R factor calculated before the final stereochemical "idealization." §Correlation coefficient is less sensitive to the scaling constants and will be 1.0 for two identical data sets differing only by a scaling factor. The expression used is in (*22*). The values for 3 Å data are calculated before any refinement. ‖+ and − signs indicate the largest positive and the largest negative deviations from the average value of each bond type. Most deviating bonds and angles are also identified with the residue numbers. The numbering system used is the one in (*7*). The Y base was not included in the average. #Some of these may be classified as H bonds although they were not explicitly noted as such in published work. **Only for those H-bonds explicitly noted in published work. ††Residue 76 is excluded (see the text).

from them. High R factors (crystallographic discrepancy factors) are primarily due to the refinement being at its early stage, and partly due to (i) the exclusion of metal ions or bound water molecules in the structure factor calculations in all three models, and (ii) the "idealization" of the structure to minimize, for example, bad bond lengths, angles, and nonbonded contacts. The second class contains quantities that reflect the quality of stereochemistry, that is, deviations from average bond distances, bond angles, shortest nonbonded contacts, and the planarity of bases. The average deviations are about 0.04 Å and 4° for bond distances and bond angles for all three models, although the distributions are different (Table 1). The last class shows the ranges of hydrogen bond (H bond) distances and glycosyl bond torsion angles.

The existence of some bad nonbonded contacts and the relaxed criterion of defining H-bond distances are shown in Table 1. These indicate that the assignments of certain H bonds are still tenuous unless strongly supported by secondary structure, favorable stereochemical disposition, and other physical, chemical, and genetic evidence.

## Comparison of the Three Models

When the three models were compared as described above, all were found to be essentially the same. There is good agreement (the root mean square difference is 1.2 Å; the average difference is 1.0 Å) between models A and B, and this is not surprising because the models are obtained from the same diffraction data refined in two different ways. What is surprising is the near identity of model C, which is from a different crystal form, to models A and B (in both cases, the root mean square difference is 1.4 Å; the average difference is 1.1 Å). The ribose-backbone structures of the three models are shown in Fig. 2. The most obvious difference among the three models is found at the 3′ end, residue 76, where the electron densities are weak in all three cases. The group centers that differ by more than 3 Å in position between any two models after the least-squares fit are listed in Table 2.

Thus, as a first approximation, one can say that the overall structure of yeast $tRNA^{Phe}$ in the two different crystal forms is essentially the same, and that it is likely to be the structure of free tRNA in solution. [For a review of the correlation between crystal structure and solution studies see (*4*) and (*11*).]

## Overall Structure, Symmetry, and "Hole"

The overall structure of yeast $tRNA^{Phe}$ in the three models has the double helical stems implicit in the cloverleaf diagrams, but it also folds into an L shape with the amino acid acceptor stem at one end and the anticodon loop at the other end as shown in Fig. 3. The amino acid acceptor stem and the TψC (*12*) stem are approximately colinear along one arm of the L while the dihydrouridine stem and the anticodon stem are roughly at right angles to it, along the other arm of the L. The dihydrouridine and TψC loops come together at the corner of the L. A summary of the structural features common in all three models is given below. Additional comments should be made about the presence of symmetry and a "hole" in this structure. There is a pseudo twofold symmetry axis relating the long helix (of the amino acid acceptor and TψC stems) to the other long helix (of the dihydrouridine and anticodon stems) of the L (Fig. 4). This pseudo twofold symmetry is the key feature in a proposed symmetry-matching process by which a tRNA molecule and an aminoacyl-tRNA synthetase molecule prealign

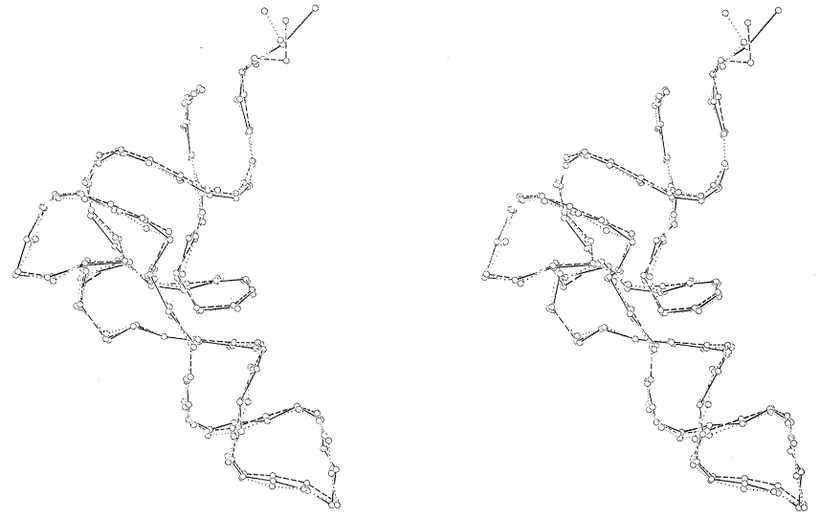

Fig. 2. Stereo drawings of the backbones of the three yeast phenylalanine tRNA ($tRNA^{Phe}$) models (model A in solid lines, model B in dashed lines, model C in dotted lines). Circles indicate the centers of the ribose rings.

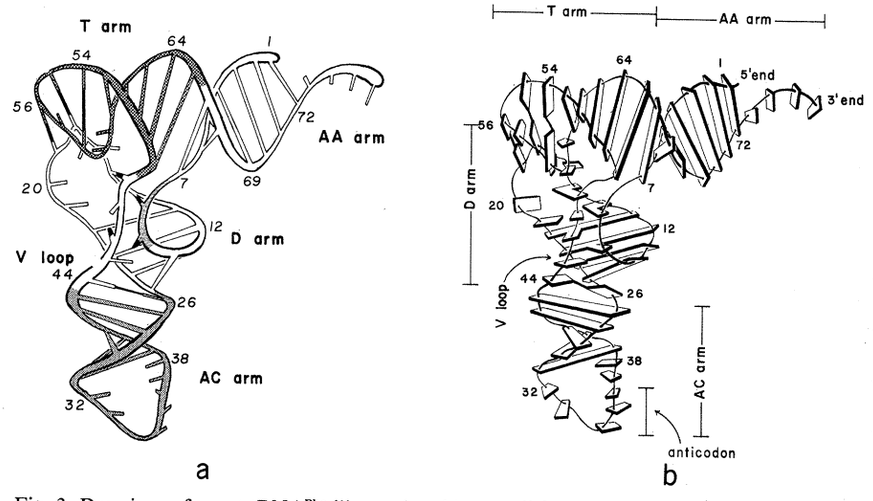


Fig. 3. Drawings of yeast $tRNA^{Phe}$ illustrating the overall features common in all three models. (a) The backbone is shown as a continuous cylinder with bars indicating H-bonded base pairs. The eight tertiary base pairs are indicated by black rods while single bases are shown as short bars. The TψC arm is heavily stippled, the anticodon (AC) arm is lightly stippled, interstem residues 8, 9, and 26 are shaded by horizontal lines, and the amino acid acceptor (AA) and dihydrouridine (D) arms as well as the V loop are unshaded. (b) The backbone is shown as a continuous thin line with bases indicated by slabs. This figure shows the extensive base stacking along two major axes of the molecule. The eight tertiary base pairs are indicated by either fusing two slabs together at an angle or connecting two slabs with dark bars. The scale at right is 10 Å per division.

themselves before the specific recognition between a cognate pair can take place (*13*).

Another interesting feature of this structure is the "hole" (approximately 10 Å by 6 Å across) in the molecule surrounded by the TψC stem and loop, and the dihydrouridine loop and stem (Fig. 5). This hole could be a site where ribosomal 5*S* RNA interacts to open up the tertiary H bonds between the TψC and dihydrouridine loops to form a new complex between the TψC loop and 5*S* RNA, as suggested indirectly by a competitive binding experiment of tRNA and TψCG sequence to ribosomes (*14*).

It is also clear from Fig. 5 (see also *7*, figure 1) that the yeast $tRNA^{Phe}$ structure has one deep groove and one shallow groove winding around each of the two axes of the L.

### Tertiary Hydrogen Bonds Between Bases

One of the most interesting findings from the tertiary structure of yeast $tRNA^{Phe}$ is that most bases which are invariant or semi-invariant (constant purines or constant pyrimidines) in all known tRNA sequences form tertiary H bonds among themselves, presumably maintaining the specific tertiary structure that is common to all tRNA's (*15, 16*).

Eight such tertiary interactions that are common in both the orthorhombic form (models A and B) and the monoclinic crystal form (model C) of yeast $tRNA^{Phe}$ are shown in Fig. 6 on the cloverleaf and "L diagrams," and in Fig. 3 on schematic three-dimensional structures. They are six base pairs (U8·A14, G15·C48, G18·ψ55, G19·C56, G26·A44, T54·A58) and two base triples (U12·A23·A9, and C13·G22·G46). The details of the H-bonding schemes have been described (*15–17*).

Table 2. Distances in group-center positions greater than 3 Å between two models calculated, on the basis of the atomic coordinates from references (*7–9*). A, B, and C refer to the three models (Fig. 1); P, R, and B refer to the centers of phosphates, riboses, and bases as defined in the text.

| | A vs. B | A vs. C | B vs. C |
|---|---|---|---|
| P1 | 2.3 Å | 4.0 Å | 2.0 Å |
| B16 | 3.1 | 3.3 | 1.2 |
| B17 | 1.9 | 1.6 | 3.6 |
| P47 | 3.3 | 1.5 | 3.2 |
| B47 | 3.7 | 2.0 | 2.0 |
| R75 | 2.7 | 1.1 | 3.8 |
| B75 | 2.3 | 1.5 | 3.7 |
| P76 | 3.0 | 2.5 | 5.1 |
| R76 | 5.4 | 8.7 | 4.1 |
| B76 | 5.8 | 5.4 | 9.6 |

The earlier ambiguities in the complex region of the TψC and dihydrouridine loops in the monoclinic crystal form (*3*) have now been resolved (*16*), and the current interpretation agrees with that of the orthorhombic crystal form (*2*). With regard to the H-bonded tertiary interactions between bases, the two remaining discrepancies among the three models involve interaction of G10 with G45 and interaction of C32 with A38 (see solid lines in Fig. 1). At the current stage of refinement of model A, although bases G10 and G45 are near one another, they are too far apart to be H-bonded.

In addition, there is a minor difference in the H-bond scheme for G18·ψ55. In models A and B, N(2) of G18 makes an H bond to O(4) of ψ55. But in model C, both N(2) and N(1) of G18 make H bonds to O(4) of ψ55 (*16*), which requires a distorted H-bond geometry. These discrepancies will likely be resolved on further refinement of the structure.

All eight of these H-bonded tertiary interactions are between invariant or semi-invariant bases of the class of tRNA's that have four base pairs in the dihydrouridine stem and five residues in the variable (V) loop (class D4V5, which makes up more than 50 percent of all the known tRNA sequences (*15, 18*), thus explaining the reason for their invariance in the sequence.

### Base Stacking Interaction

Besides the secondary structure, very extensive base stacking along the central core of both molecular axes appears to be the primary stabilizing force of the structure in both crystal forms (Fig. 5) (*7*, figure 1). This base stacking is analogous to the hydrophobic core observed in many protein structures. About a year and a half ago, in the interpretation of the electron density maps at 3 Å resolution (*2, 3*), both the orthorhombic and the monoclinic structures appeared to have approximately the same base stacking, except for one possibly significant region between the dihydrouridine stem and the anticodon stem. In the orthorhombic form, G26 forms a tertiary base pair with A44 (*2*), while in the monoclinic form, G26 was partially intercalating between A44 and G45 (*3*), thus tilting the whole anticodon arm relative to the dihydrouridine stem more than in the orthorhombic form. At that time, such an apparent difference implied that two functional states of the orientation of the anticodon stem relative to the rest of the molecule might have been "frozen" in the two different crystal forms. Reinterpretation of this region in the monoclinic form based on a 2.5 Å resolution map, however, revealed that G26 is not intercalated between A44 and G45, but forms a tertiary base pair to A44 (*16*) as in the orthorhombic crystal form (*2*). The overall base stacking in both crystal forms is now the same and is shown schematically in Fig. 3b. In all three models, all but five bases (D16, D17, G20, U47, and A76) are stacked along two major axes of the molecule.

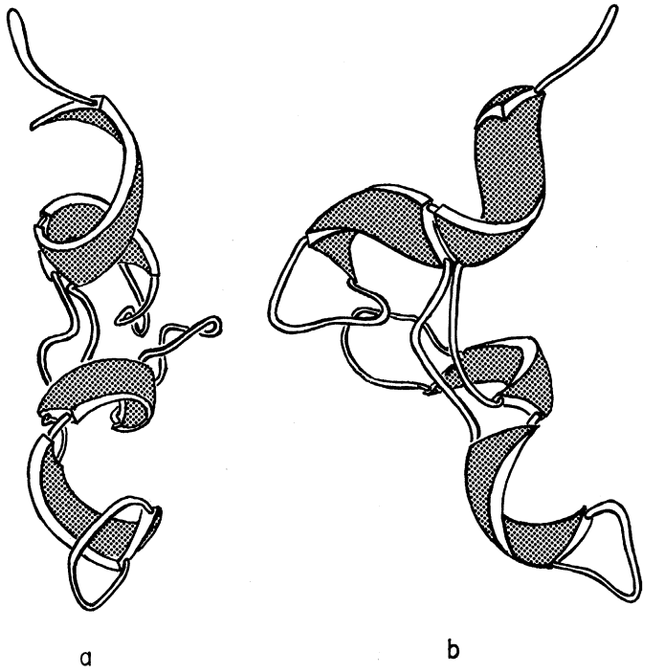


Fig. 4. Drawing of the yeast $tRNA^{Phe}$ structure, showing the double helical stems as shaded ribbons. (a) A view looking down the pseudo twofold axis which relates the amino acid acceptor (AA) stem–TψC (T) stem helix to the dihydrouridine (D) stem-anticodon (AC) stem helix. (b) A view [rotated 90° from (a)] perpendicular to the plane of the molecule showing the two major helical axes at approximately right angles.

### Ribose Conformation

X-ray studies on nucleotides (*19*) have indicated that the most common conformation for the ribose moiety is 3′ *endo*. All three models agree on this general tendency. All the riboses in both crystal forms are in the 3′-*endo* conformations except for nucleotide residues 7, 9, 17, 19, 21, 46, 48, and 60 in model A (*7*), residues 7, 9, 18, 19, 48, 58, and 60 in model B (*8*), and residues 7, 9, 17, 18, 19, 21, 46, 48, 58, and 60 in model C (*9*); for these, 2′-*endo* conformation has been assigned. It is expected that the discrepancies in the above lists will be reduced as the refinements progress further.

### Glycosyl Bond Conformation

So far, all the known crystal structures of 5′-ribonucleotides display the *anti* conformation (*19*) for the glycosyl bond connecting the base and the ribose. This appears to be the case in all three models except in a few ambiguous residues.

In all residues, the glycosyl bond was classified as being in the *anti* conformation for model C (*9*); in all except residue 44 in model B (*8*); and in all except residues 17, 19, 60, and 76 of model A, where unambiguous assignments are considered to be impossible at the present time (*7*). The glycosyl bond conformation for 2′-*endo* nucleosides are particularly difficult to assign at the resolution now available.

### G·U Base Pair

Yeast tRNA$^{Phe}$ has G4 and U69 in the amino acid acceptor stem in position to form at least one H bond. At 3 Å resolution, G4·U69 was interpreted to form a "wobble" pair (*20*) in the monoclinic form (*3*), but was ambiguous in the orthorhombic form (*2*). At higher resolution, it appears that the "wobble" pair is more likely in the orthorhombic form as well (*7, 8*). It is unambiguous, however, that this unusual base pair in the helical stem disrupts the regularity of the helix significantly. This distortion of the helical backbone may have a role in one of the recognition processes involving this tRNA.

### Other Tertiary Hydrogen Bonds

As is clear from Table 1, the range of assigned H-bond distances is wide in all three models, and there are some close nonbonded contacts in models B and C. Therefore, assignment of H bonds is considered tenuous unless supported by other evidence. The H bonds between the bases are relatively easy to assign because the stereochemical requirements to form such base pairs are rather strict. However, assignment of H bonds involving riboses or phosphates is much more difficult and somewhat arbitrary at the present time. For example, in model B, the longest potential H bond assigned is 4.2 Å between O(1′) of G19 and N(2) of G57; in model C, 4.1 Å between N(4) of C61 and a phosphate oxygen of C60.

Because of the loose stereochemical restrictions of H bonding, discrepancies between the models on H-bonded tertiary interaction other than those between the bases are numerous. At the current state of refinement, 8 hydrogen bonds of this category have been assigned in model B (*8*) and 13 in model C (*16*). Among many potential H bonds in model A, only the following five can be assigned with some confidence at the present time (OP refers to phosphate oxygen): O(2′) of U7 is H-bonded either to OP or O(1′) of C49; O(2′) of U8 is H-bonded to N(1) of A21; O(2′) of A21 is H-bonded to O(6) of G46; O(2′) of C48 is H-bonded to O(2′) of U59; and OP of C60 to N(4) of C61.

Among those published (*8, 9*) and listed above, only two agree in all three models. They are H bonds between O(2′) of U8 and N(1) of A21, and between OP of C60 and N(4) of C61. This is a strong indication that the assignment of H bonds involving riboses and phosphates is subjective and tenuous at the current state of refinement.

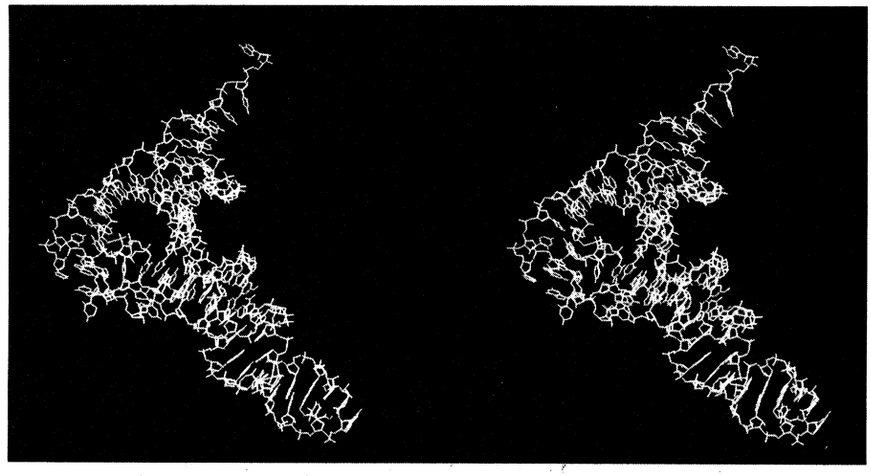

Fig. 5. Stereo view of yeast tRNA$^{Phe}$ showing the "hole." Another good stereo view of this molecule is in (*7*).

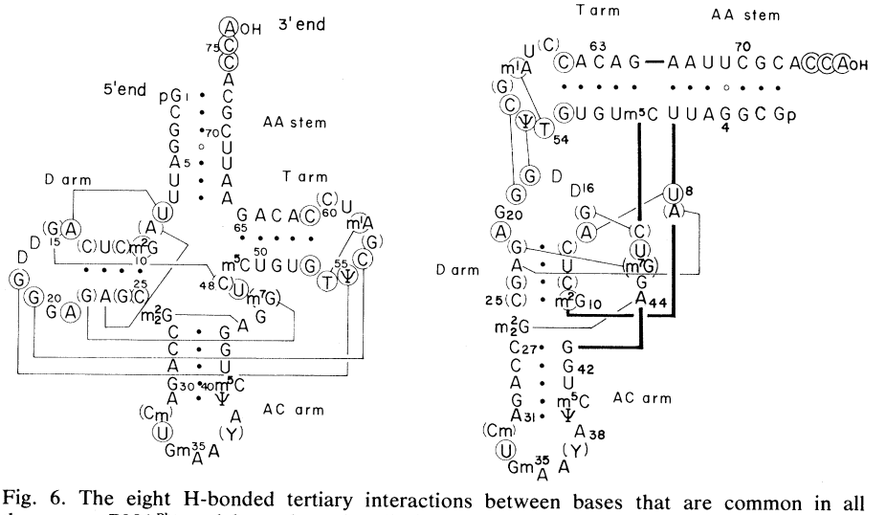


Fig. 6. The eight H-bonded tertiary interactions between bases that are common in all three yeast tRNA$^{Phe}$ models are shown in solid lines in the cloverleaf (left) and "L diagram" (right). Circles indicate invariant and parentheses semi-invariant bases (constant purines or constant pyrimidines) of the class of tRNA's that have four base pairs in the dihydrouridine stem and five residues in the V loop (D4V5). More than 50 percent of all the known tRNA sequences belong to this class.

### Irregular Double Helical Stems

Studies at 3 Å resolution showed that all double helical stems appeared as "irregular" RNA-A type (*21*) in the orthorhombic crystal form (*15*), and considerable disturbance in the helical backbone was observed around the G4·U69 pair (*22*). However, in the monoclinic crystal form, at the same resolution, the G4·U69 pair did not seem to introduce major disturbance to the amino acid stem, nor did other helical stems appear to be irregular (*3*). Helical parameters calculated by a least-square procedure (*23*), on the basis of the latest atomic coordinates of all three models, show considerable irregularity in the helicity of all four double helical stems in both crystal forms. The degree of irregularity is apparent from the r.m.s. errors of about 6° in rotational angles and 0.6 Å in rise per residue. In all three models the base pairs are tilted with respect to the helical axes, the amino acid acceptor, anticodon, and TψC stems are close to the 11-fold and the dihydrouridine stem is closer to the 10-fold helix. However, owing to the irregularity of these helices, it is more appropriate to describe them only as of genus type A (*24*). Irregularities of the individual helix may be a reflection of the variability of helical parameters observed in many different RNA fibers (*24*).

### Conclusion

In earlier 3 Å resolution studies (*2, 3*) of this tRNA, considerable differences seemed to exist between the structures in two crystal forms. However, they were primarily due to difficulties in the interpretation of a complex region between the TψC and dihydrouridine loops, and the apparent misinterpretation of a region near G26·A44 in the monoclinic crystal form (*3*).

Although the two crystal forms are related (see Table 1), the molecular packing is different, and the amounts of mother liquor present in these two crystals are drastically different, about 75 percent in the orthorhombic crystal and 55 percent in the monoclinic crystal by volume. However, the comparison of the atomic coordinates shows that the three-dimensional structures of yeast $tRNA^{Phe}$ in both crystal forms are essentially the same except for a few minor differences. These are at the 3′ end, the 5′ end, and two looped-out regions around D16, D17, and U47. There are a few other differences in the assignment of conformation of riboses and glycosyl bonds, as was discussed earlier. In addition, discrepancies are found in the assignment of H bonds, involving riboses and phosphates. Our analysis indicates that most of these discrepancies are likely to be artifacts because of the loose criteria used in defining this class of H bonds at the current resolution. Besides, they appear to play a supplementary role in maintaining the complex three-dimensional structure rather than the principal role, which is played by base pairs (secondary and tertiary) requiring more stringent stereospecificity and base stacking, as discussed above. It is likely that most of the remaining few differences discussed above, such as ribose conformations, glycosyl conformations, and H bonds between the bases, will disappear as the refinement of the structures in both crystal forms continues.

The structural features common in all three models at the present time can be summarized as follows:

1) The molecule has an overall shape of a letter L.

2) The secondary structure predicted from the cloverleaf model is present.

3) The amino acid arm and the TψC arm make one long helix; the dihydrouridine arm and the anticodon arm make the other long helix. These two long helices are related by a pseudo twofold axis and form an L.

4) The 3′ terminus is at one end and the anticodon is at the other end of the L. The TψC and dihydrouridine loops are at the corner of the L.

5) The molecule is about 22 Å thick and the lengths of the two axes of the L are about 73 and 70 Å, respectively.

6) All bases except five are stacked along the two axes of the molecule. They are D16, D17, G20, U47, and A76.

7) All the double helical stems are of the genus A type.

8) There are eight H-bonded tertiary interactions between bases, which are either invariant or semi-invariant in most tRNA's. These are U8·A14, G15·C48, G18·ψ55, G19·C56, G26·A44, T54·A58, and there are two base triples, that is, U12·A23·A9 and C13·G22·G46.

The common structural features described in this article are compatible with and explain vast amounts of physical, chemical, and genetic data on tRNA [see (*4*)]. There is little doubt that the structure of yeast $tRNA^{Phe}$ in crystals is the same as that in solution as a free, unbound molecule. This structure also provides a basis for understanding all other tRNA structures (*15*). The three-dimensional structure determined by x-ray crystallographic methods is necessarily a static one, and no dynamic conclusions can be drawn directly. Nevertheless, it provides the most solid foundation and starting point for understanding the functional roles of tRNA, be it static or dynamic.

*Note added in proof*: A fourth set of atomic coordinates of this same tRNA in the monoclinic form has been described by Stout *et al.* (*25*) who state that the average discrepancy between their model and the three discussed here (models A, B, and C) is about 1.5 Å while that between models A, B, and C is about 1.0 Å. The stereochemistry of model B has been improved on further refinement (*26*).

#### 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. S. H. Kim, F. L. Suddath, G. J. Quigley, A. McPherson, J. L. Sussman, A. Wang, N. C. Seeman, A. Rich, *ibid.* **185**, 435 (1974).
3. J. D. Robertus, J. E. Ladner, J. T. Finch, D. Rhodes, R. S. Brown, B. F. C. Clark, A. Klug, *Nature (London)* **250**, 546 (1974).
4. S. H. Kim, *Prog. Nucleic Acid Res. Mol. Biol.* **17**, 181 (1976); P. B. Sigler, *Ann. Rev. Biophys. Bioeng.* **4**, 477 (1975).
5. G. J. Quigley, F. L. Suddath, A. McPherson, J. J. Kim, D. Sneden, A. Rich, *Proc. Natl. Acad. Sci. U.S.A.* **71**, 2146 (1974).
6. A. Klug, J. D. Robertus, J. E. Ladner, R. S. Brown, J. T. Finch, *ibid.*, p. 3711.
7. J. L. Sussman and S. H. Kim, *Biochem. Biophys. Res. Commun.* **68**, 89 (1976).
8. G. J. Quigley, N. C. Seeman, A. H. J. Wang, A. Rich, *Nucleic Acid Res.* **2**, 2329 (1975).
9. J. E. Ladner, A. Jack, J. D. Robertus, R. S. Brown, D. Rhodes, B. F. C. Clark, A. Klug, *ibid.*, p. 1629.
10. S. C. Nyburg, *Acta Crystallogr. Sect. B* **30**, 251 (1974).
11. R. Langlois, S. H. Kim, C. R. Cantor, *Biochemistry* **14**, 2554 (1975).
12. The notations for the nucleotide bases are as follows: A, adenosine; T, thymidine; C, cytidine; U, uridine; G, guanosine; ψ, pseudouridine; and D, dihydrouridine; Y, highly modified purine.
13. S. H. Kim, *Nature (London)* **256**, 679 (1975).
14. V. A. Erdman, M. Sprinzl, O. Pongs, *Biochem. Biophys. Res. Commun.* **54**, 942 (1973).
15. S. H. Kim, J. L. Sussman, F. L. Suddath, G. J. Quigley, A. McPherson, A. H. J. Wang, N. C. Seeman, A. Rich, *Proc. Natl. Acad. Sci. U.S.A.* **71**, 4970 (1974).
16. J. E. Ladner, A. Jack, J. D. Robertus, R. S. Brown, D. Rhodes, B. F. C. Clark, A. Klug, *Proc. Natl. Acad. Sci. U.S.A.* **72**, 4414 (1975).
17. G. J. Quigley, A. H. J. Wang, N. C. Seeman, F. L. Suddath, A. Rich, J. L. Sussman, S. H. Kim, *ibid.*, p. 4866.
18. For a convenient compilation, see B. G. Barrell and B. F. C. Clark, *Handbook of Nucleic Acid Sequences* (Joynson-Bruvvers, Oxford, 1974).
19. M. Sundaralingam, *Jerusalem Symp. Quantum Chem. Biochem.* **5**, 417 (1973).
20. F. H. C. Crick, *J. Mol. Biol.* **19**, 548 (1966).
21. S. Arnott, D. W. L. Hukins, S. D. Dover, *Biochem. Biophys. Res. Commun.* **48**, 1392 (1972).
22. J. L. Sussman and S. H. Kim, *Jerusalem Symp. Quantum Chem. Biochem.* **8**, 535 (1976).
23. J. M. Rosenberg, thesis, Massachusetts Institute of Technology (1973).
24. S. Arnott, R. Chandrasekaran, E. Selsing, in *Structure and Conformation of Nucleic Acids and Protein–Nucleic Acid Interactions*, M. Sundaralingam and S. Rao, Eds. (University Park Press, Baltimore, 1975), p. 577.
25. C. Stout, H. Mizuno, J. Rubin, T. Brennan, S. Rao, M. Sundaralingam, *Nucleic Acid Res.* **3**, 1111 (1976).
26. G. Quigley and A. Rich, personal communication.


27. Supported by NIH grants CA 15802 and RR 00898 and NSF grant GB 40814. We thank W. Warrant, S. Holbrook, and G. Church for their help, J. Rosenberg for his computer program, and the members of the molecular graphics group at the University of North Carolina, Chapel Hill, for a collaborative effort in implementing the crystallographic display into their graphics system. We also thank C. Goldsmith for technical assistance. J.L.S. is supported by a fellowship from the Arthritis Foundation.