TRANSFER RNA: MOLECULAR STRUCTURE, SEQUENCE, AND PROPERTIES

Alexander Rich and U. L. RajBhandary
Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

CONTENTS
INTRODUCTION ............................................................................................................ 806
THE MULTIPLE BIOLOGICAL FUNCTIONS OF tRNA ................................................ 807
- tRNA Cycle in Protein Synthesis ........................................................................... 807
- tRNA and the Regulation of Enzyme Synthesis .................................................... 807
- Aminoacyl-tRNA Transferases ............................................................................ 808
- tRNA Participation in Polynucleotide Synthesis .................................................. 808
- tRNA as an Enzyme Inhibitor ............................................................................. 808
- tRNA Changes in Cells ....................................................................................... 809
NEWER METHODS FOR PURIFICATION AND SEQUENCE ANALYSIS OF tRNA ..... 809
- Purification of tRNAs ......................................................................................... 810
- Sequence Analysis of tRNA ............................................................................... 812
GENERAL FEATURES OF tRNA SEQUENCES ............................................................ 813
- Generalized Secondary Structure for tRNAs ....................................................... 815
- Invariant and Semi-invariant Nucleotides in tRNAs ......................................... 817
- Unique Features in Initiator tRNA Sequences .................................................... 818
MOLECULAR STRUCTURE OF NUCLEIC ACID COMPONENTS AND DOUBLE
- HELICAL NUCLEIC ACIDS .................................................................................. 819
CRYSTALLIZATION OF tRNA ................................................................................... 821
- High Resolution Crystals of Yeast tRNA™ ......................................................... 821
- Solution of X-ray Diffraction Patterns Using Heavy-Atom Derivatives ............... 823
SOLUTION OF THE YEAST tRNA™ STRUCTURE BY X-RAY DIFFRACTION ........ 825
- Folding of the Polynucleotide Chain at 4-Å Resolution—1973 ............................. 825
- Tertiary Interactions at 3-Å Resolution—1974 .................................................... 827
- Tertiary Interactions and Coordinates at 2.5-Å Resolution—1975 ....................... 828
THREE-DIMENSIONAL STRUCTURE OF YEAST tRNA™ ...................................... 829
- Acceptor Stem .................................................................................................... 829
- TVC Stem and Loop ............................................................................................ 830
- D Stem and Loop ................................................................................................. 835
- Anticodon Stem and Loop ................................................................................... 837
INTRODUCTION

Research in the field of transfer RNA (tRNA) has undergone revolutionary changes in the past few years. Although there has been a steady accumulation of chemical and biological information concerning this molecule for almost 20 years, until 1973 there was no firm information available about the three-dimensional structure of the molecule. Early in 1973, however, the polynucleotide chain of yeast tRNA⁰ was traced in a 4-Å X-ray diffraction analysis (1). Structural work has progressed rapidly since then to the point where atomic coordinates are now available as derived from 2.5-Å X-ray diffraction analyses from two different crystal forms of the same molecule (2–4). Knowledge of the detailed three-dimensional structure of the molecule makes a distinct change in the type of research that can be carried out. We are now in a position to ask many detailed questions concerning both the chemistry and the biological function of tRNA, using the structural information to guide our thinking.

The aim of this review is to describe in some detail the manner in which we have obtained knowledge of the three-dimensional structure of one tRNA species and to discuss the extent to which it explains and makes understandable various aspects of the chemistry and solution behavior of this and other tRNA species. We review tRNA sequences and the methods of obtaining them. We also try to direct attention toward unsolved problems associated with tRNA chemistry and point out various types of research that are beginning to lead us toward a more detailed molecular interpretation of tRNA biological function.

The major biological function of tRNA is related to its role in protein synthesis. The existence of a molecule-like tRNA is in a sense made necessary by the fact that although Nature encodes genetic information in the sequence of nucleotides in the nucleic acids, it generally expresses this biological information in the ordered sequence of amino acids in polypeptide structures. Transfer RNA has a fundamental biological role in acting at the interface between polynucleotides and polypeptides. It works in the ribosome by interacting with messenger RNA at one end while at the other end it contains the growing polypeptide chain. We do not know how this process occurs, but a detailed knowledge of the three-dimensional structure of one species of tRNA means that we are now in a position to ask intelligent questions about the molecular dynamics of this biological function.

Transfer RNA is involved in a large number of biological processes and it would be impossible to review adequately within the confines of any one article all of the
research going on in this field. We will of necessity be selective in this review. Fortunately, a number of excellent reviews dealing with various aspects of tRNA have been published recently. The review by Sigler (5) covers many of the aspects of structure determination. A comprehensive review of chemistry (6) is available and chemical modifications of tRNA are reviewed by Zachau (7) and Cramer & Gauss (8). Other reviews concern the role of tRNA in protein synthesis (9–11), biosynthesis of tRNA including the role of tRNA modifying enzymes, tRNA maturation enzymes and tRNA nucleotidyl transferase in this process (12–15), and the structure and function of modified nucleotides in tRNA (16).

THE MULTIPLE BIOLOGICAL FUNCTIONS OF tRNA

Although the role of tRNA in protein synthesis is usually emphasized, it is important to recognize that this molecule is involved in many other biological functions. They are outlined here; several of these specialized functions have been the subject of other recent review articles.

**tRNA Cycle in Protein Synthesis**

During protein synthesis tRNA interacts with a large number of different proteins that play an important role in its biological function. All tRNA molecules end in a common sequence, CCA, which is added by the nucleotidyl transferase enzyme to the 3'-end of the molecule. An important step in protein synthesis is the specific aminoacylation, which is carried out by means of 20 different tRNA-aminocacylating enzymes or aminoacyl tRNA synthetases. These enzymes recognize only a specific set of isoacceptor tRNA's as substrates and require ATP for the initial activation of the amino acid before it is transferred onto the tRNA. Although the amino acid is added to the 3'-terminal adenosine, it has been found recently that some of these enzymes aminoacylate on the 2' hydroxyl and some on the 3' hydroxyl groups (17, 18). There have been two recent reviews discussing the various aminoacyl-tRNA synthetases (19, 20).

The aminoacyl tRNA (aa-tRNA) is carried into the ribosome complexed with the transfer factor EF-Tu (21) in prokaryotes or EF1 in eukaryotes. It should be noted that the initiator tRNA_{Met} has its own factor for ribosomal insertion. Inside the ribosome tRNA interacts with a number of ribosomal proteins including the peptidyl transferase before it is finally released from the ribosome after its amino acid has been transferred to the growing polypeptide chain of an adjacent tRNA. Ribosomal processes have been reviewed in a recent volume (22). Although a fair amount is known about various aspects of tRNA biosynthesis and function during protein synthesis, virtually nothing is known about the manner in which tRNA molecules are degraded.

**tRNA and the Regulation of Enzyme Synthesis**

One of the remarkable features of aa-tRNA is the fact that it has been shown to play a role in regulating the transcription of messenger RNA for enzymes associated with biosynthesis of its amino acid. This was first discovered in the operon for histidine biosynthesis. The regulatory role of tRNA has been reviewed recently (23, 24).
Although most of the regulatory studies have been carried out on prokaryotic systems, it has recently been demonstrated that aa-tRNA in mammalian systems also regulates amino acid biosynthesis (25).

**Aminoacyl-tRNA Transferases**

Aminoacyl-tRNA transferases are a group of enzymes that catalyze the transfer of an amino acid from aa-tRNA to specific acceptor molecules without the participation of ribosomes or other kinds of nucleic acid. The acceptor molecules can be divided into three classes: (a) The acceptor can be an intact protein, in which case the amino acid is added to the N-terminus of the protein (26). (b) The acceptor may be a phosphatidyl glycerol molecule (27), in which case the enzyme catalyzes the formation of aminoacyl esters of phosphatidyl glycerol that are components of cell membranes. (c) The acceptor is an N-acetyl muramyl peptide, an intermediate in the synthesis of interpeptide bridges in bacterial cell walls (28). These are important links in cell wall biosynthesis, and somewhat specialized tRNAs are used for this (29). The aa-tRNA transferases have recently been reviewed by Soffer (30).

**tRNA Participation in Polynucleotide Synthesis**

Reverse transcriptase is an enzyme found in oncogenic viruses that is used for making a DNA copy of the viral RNA. It has been found that a particular species of tRNA is used as a primer in this process (31). Avian myeloblastosis reverse transcriptase uses tRNA\textsuperscript{Trp}, whereas the murine leukemia virus enzyme uses tRNA\textsuperscript{Pro} as a primer. Recent studies have further shown that the reverse transcriptase has a strong affinity for the tRNA primer (31a). An interesting finding that may bear some relationship to the above is the fact that many plant viral RNAs possess a “tRNA-like” structure at the 3'-end of the RNA. A number of plant viral RNAs (32) as well as an animal viral RNA (33) are found to act as substrates for aminoacylation by aa-tRNA synthetases. The work of Haenni and coworkers (33a) suggests that bacterial viral RNAs may also possess some features of “tRNA-like” structures, although not at the 3'-end. Furthermore, one of the proteins that binds specifically to aa-tRNA, the transfer factor EF-Tu (21), is also a component of the enzyme Qβ replicase (34), which is involved in the replication of the bacterial viral RNA. Whether these “tRNA-like” structures that appear to be present in many plant and bacterial viral RNAs play a role in the specific recognition of these RNAs by the corresponding RNA replicases is an interesting possibility that needs to be explored further.

**tRNA as an Enzyme Inhibitor**

tRNA is a potent inhibitor of *E. coli* endonuclease I. The work of Goebel & Helinski (35a) suggests that tRNA alters the mode of action of endonuclease I from that of double strand scission of DNA to a nicking activity.

A specific isoacceptor species of tRNA\textsuperscript{Tyr} in *Drosophila* has been found to act as an inhibitor to the enzyme tryptophan pyrrolase (35b), which is involved in the conversion of tryptophan to an intermediate in brown-pigment synthesis. In this case, an uncharged tRNA appears to act in a regulatory capacity by directly interfer-
ing with an individual enzymatic activity, although alternative explanations have been proposed recently (35c).

**tRNA Changes in Cells**

There is a large literature dealing with changes that have been observed in the cell content of tRNAs. Two review articles (23, 36) summarize a variety of results dealing with the changes of tRNA that occur in embryogenesis during various stages of development. It is not clear whether these changes reflect an expression of the role of tRNA in regulatory systems such as those discussed above or whether they are involved in the regulation or modification of other functions as well. In addition, there is a substantial literature reviewed in Cancer Research dealing with changes in tRNA during oncogenesis; an entire volume is devoted to this subject (37). The relationship of these changes to the changes observed during development is a subject that needs to be explored more fully in the future.

Why is tRNA used in such a large variety of biological functions? It is true that this class of molecules has been involved in the biochemistry of living organisms from the very onset of the evolutionary process and it may reflect the fact that Nature is opportunistic in using such molecules for other purposes; however, it is important to point out that we do not understand the rationale behind the multiplicity of functions carried out by tRNA molecules.

In a large number of biological functions, tRNA interacts with protein molecules in a highly specific manner. The nature of these interactions is largely unknown, but it is probable that the interactions involve the recognition of tRNA as distinct from other species of RNA by the three-dimensional folding of the molecule and the detection of specific nucleotides or nucleotide sequences in tRNA by many proteins. With our understanding of the three-dimensional conformation of one species of tRNA, we can now ask about the extent to which this molecular structure may serve as a useful guide for understanding the detailed manner in which tRNA interacts with a variety of proteins while carrying out a large number of different biological functions.

**NEWER METHODS FOR THE PURIFICATION AND SEQUENCE ANALYSIS OF tRNA**

The first tRNA molecule was sequenced in 1965 (38); the sequence of about 75 different tRNAs is now known. This wealth of sequence information has been invaluable both in understanding certain aspects of structure-function relationships (7, 39) and in establishing the generality of secondary structure of tRNAs. Now that the three-dimensional structure of a tRNA has been elucidated, the major aim in tRNA sequence studies in the future will be geared more toward understanding the role of tRNAs in regulation and control processes and in specific aspects of protein biosynthesis, rather than for the sole purpose of compiling tRNA sequences. These could include, for instance, sequence studies of eukaryotic suppressor tRNAs (40), tRNAs from eukaryotic organelles such as mitochondria and chloroplasts, tRNAs found specifically in tumor cells, tRNAs known to undergo changes during develop-
ment, and other tRNAs potentially involved in the regulation of protein synthesis and activity (23). Most of these tRNAs are expected to be available only in limited amounts. Consequently, the development of methods that allow the rapid purification and sequence analysis of tRNAs on a very small scale will play an important role in future work on tRNAs.

**Purification of tRNAs**

Following the earlier use of countercurrent distribution (42) in tRNA purifications, two of the most widely used methods in recent years have been chromatography on BD-cellulose (43) and on DEAE-Sephadex (44). These and other procedures more suitable for large-scale purification have been described elsewhere (45).

Kelmers and co-workers have recently developed two new high-pressure "reversed phase chromatography" systems, RPC-5 and RPC-6 (46). Of these two, RPC-5 has been the one most widely used. The principle behind the separation involves both ion exchange and hydrophobic interactions between the tRNAs and the coating material (47, 48). On the analytical scale (49), the RPC-5 system has been particularly useful for monitoring changes in tRNA isoacceptor patterns during development (50) and differences between normal and tumor-cell tRNAs (51, 52) and between tRNAs from quiescent cells and those from proliferative cells (53). Several reports have described large-scale purification of mammalian (54), *Escherichia coli* (55), and *Drosophila* (47, 56) tRNAs using RPC-5 chromatography.

Although initially described as a method for tRNA purification, RPC-5 has proved equally useful for the rapid separation of mononucleotides, oligonucleotides present in total T-R or pancreatic RNase digests of tRNA (55, 57, 58), large oligonucleotide fragments present in partial digests of tRNAs (55), homopolynucleotides (59), and even ribosomal RNAs (60). Using analogies of RPC-5 with anion-exchange polystyrene resins, Singhal (61) has developed Aminex-A28 as an alternative chromatographic support for tRNA separations. It is reported (62) that the resolution obtained on Aminex-A28 is superior to that on RPC-5, and B. Roe (personal communication) has used Aminex-A28 in the purification of tRNAs from mammalian sources.

Chromatography on Sepharose 4B has been used recently for the large-scale purification of *E. coli* tRNAs (63). The tRNAs are adsorbed to the Sepharose in the presence of a high concentration of ammonium sulphate at slightly acidic pH; elution of the tRNAs is then carried out with a linear negative gradient of ammonium sulphate. Holmes et al (63) have purified *E. coli* tRNA<sub>Leu</sub> in a simple two-step column chromatography using Sepharose 4B as the first step and RPC-5 as the second. Other workers have described the use of anion-exchange Sepharose 6B (64) and of various aminoalkyl derivatives of Sepharose 4B (65) in separation of tRNAs.

Another method applicable to the purification of specific tRNAs takes advantage of the fact that two tRNAs whose anticodon sequences are complementary form a 1:1 tRNA : tRNA complex. The association constant of complex formation between yeast tRNA<sup>Phe</sup> (anticodon sequence G<sub>m</sub>A<sub>A</sub>) and *E. coli* tRNA<sup>Glu</sup> (anticodon se-
sequence $\text{s}^2\text{UUC}$) is of the order of $10^7$ mole$^{-1}$ (66, 67). Grosjean et al (68) have immobilized yeast tRNA$^{\text{Phe}}$ by covalent linkage through its 3'-end to polyacrylamide (Biogel P20). Upon chromatography of crude E. coli tRNA through such a column, tRNA$^{\text{Glu}}$ is specifically retarded and a 19-fold enrichment of tRNA$^{\text{Glu}}$ is obtained after a single passage. Similarly, E. coli tRNA precursors have been purified by chromatography of a mixture of [$^32\text{P}$]tRNA precursors on columns containing the appropriate tRNAs immobilized onto them (69).

In another technique, the specificity of antigen-antibody interactions is exploited for the detection and purification of tRNA$^{\text{Phe}}$ species that contain the fluorescent nucleoside Y or its derivatives by immobilizing antibodies against Y nucleoside on columns (70, 71).

Several of the newer methods for tRNA purification involve aminoacylation of the desired tRNA with a specific amino acid as the first step in their purification. The most widely used procedure is that of Tener and co-workers (72), which in most cases includes the further derivatization of the amino group of aa-tRNA with an aromatic moiety. The chemically derivatized aa-tRNA is then selectively retarded on a column of BD-cellulose and thus separated from uncharged tRNA. In an example of this approach, aa-tRNA carrying a $p$-chloromercury phenyl group is separated from uncharged tRNA by chromatography on a column of Sepharose 4B containing reactive thiol groups (73). By this method, leucine, arginine, and tyrosine tRNAs from E. coli have been obtained in a high state of purity.

The ability of aa-tRNAs to form a ternary complex with the E. coli protein synthesis elongation factor EF-Tu in the presence of GTP has been used by Klyde & Bernfeld (74) in the purification of chicken liver aa-tRNAs. The ternary complex is separated from any free aa-tRNA or uncharged tRNA by gel filtration on Sephadex G-100 (75). In the presence of limiting amounts of aa-tRNA, virtually all of the aa-tRNA forms the ternary complex. The procedure appears general and has led to the isolation of 90% pure tRNA$^{\text{Phe}}$ and highly purified preparations of tRNA$^{\text{Ser}}$, tRNA$^{\text{Leu}}$, and tRNA$^{\text{Lys}}$.

A major difference between aa-tRNAs and uncharged tRNA is that the latter contains a free 2',3'-diol end group at its 3'-terminal adenosine, whereas the former does not. This difference has been exploited by McKutchan et al (76) in a general procedure for the fractionation of aa-tRNAs from uncharged tRNAs using a column of DBAE-cellulose, which contains dihydroxyl boryl groups attached to aminoethyl cellulose. Uncharged tRNAs containing cis-diol groups form specific complexes with the dihydroxyl boryl groups and are retained on the column, whereas aa-tRNA is not retarded on the column (77–79).

Several groups (80–83) have described the use of two-dimensional gel electrophoresis on polyacrylamide for the simultaneous purification of different $^{32}\text{P}$-labeled small RNAs in a single step. Fradin et al (82) have used two-dimensional gel electrophoresis for the separation of yeast tRNA and yeast tRNA precursors. Several of the yeast tRNAs were shown to be homogeneous by fingerprint analyses (82). This technique has also been used more recently for the purification of $^{32}\text{P}$-labeled tRNAs isolated from HeLa cell mitochondria (J. D. Smith, personal communication).
Sequence Analysis of tRNA

The basic principles involved in the sequence analysis of tRNAs have been published by Brownlee (100). Techniques developed by Sanger and co-workers (84, 85) suitable for work on $^{32}$P-labeled tRNAs have greatly simplified both the separation and sequence analysis of tRNAs, and these account to a large extent for the dramatic increase in the knowledge of tRNA sequences, particularly from prokaryotic sources such as E. coli and Salmonella. In spite of these remarkable advances, sequence analysis of most eukaryotic tRNAs (notably from yeast, wheat germ, and mammalian sources) has still used the more classical procedure involving the identification of nucleotides by their ultraviolet absorption spectra, due to the problems involved in the labeling and subsequent purification of tRNAs with $^{32}$P, particularly from most higher eukaryotes. The latter procedure is more time-consuming and usually requires large amounts of purified tRNAs.

Several methods for the in vitro end-group labeling of oligonucleotides or tRNAs, which make possible sequence analysis of oligonucleotides on a small scale, have now been developed (86–89). These methods have also been used for the sequence analysis of tRNAs (90–92). It can be expected that further refinements in these techniques will eventually allow sequence analysis of nonradioactive tRNAs on as little as 25–100 μg of the tRNA.

3'-END-GROUP LABELING OF OLIGONUCLEOTIDES WITH $^3$H A general method for the specific labeling of 2',3'-diol end groups in RNAs and oligonucleotides and its use in sequence analysis was described previously (93, 94). It involves first oxidation of the 2',3'-diol end group with periodate followed by reduction of the 2',3'-dialdehyde end group with $[^3$H]sodium borohydride to yield a 3'-$^3$H-labeled dialcohol derivative of the tRNA. Randerath and his co-workers have now pioneered the application of this method in the sequence analysis of oligonucleotides (89) present in T$_1$- or pancreatic RNase digests of an RNA and have described the sequence analysis of a yeast leucine tRNA (90). Several of the new techniques introduced by Randerath for the separation of oligonucleotides by thin layer chromatography, detection of $^3$H on thin layer plates by fluorography, etc have now made this a relatively rapid and sensitive method for sequencing oligonucleotides (93, 95).

5'-END-GROUP LABELING OF OLIGONUCLEOTIDES WITH $^{32}$P An alternative procedure for sequence analysis of oligonucleotides on a small scale involves first the use of polynucleotide kinase for labeling oligonucleotides present in T$_1$- or pancreatic RNase digests of tRNA with $^{32}$P at the 5'-end (86, 96). The 5',32P-labeled oligonucleotides are separated (84) and partially digested with snake venom phosphodiesterase. These products are separated (85, 97) and the sequence of the oligonucleotide in question is deduced from the characteristic mobility shifts resulting from the successive removal of nucleotides from the 3'-end (85, 86, 91). This approach has been used to elucidate the cytoplasmic initiator tRNA sequence of salmon testes and liver (91), human placenta (92), and Neurospora crassa (A. Gillum, L. Hecker, W. Barnett, and U. L. RajBhandary, unpublished), the
tRNA\textsuperscript{Phe} from the chloroplasts of *Euglena gracilis* (92a), and lysine tRNAs of rabbit liver (H. Gross, M. Raba, K. Limburg, J. Heckman, and U. L. RajBhandary, unpublished).

3'-END-GROUP LABELING OF OLGONUCLEOTIDES WITH $^{32}$P Szeto & Söll (88) have developed a complementary method that uses polynucleotide phosphorylase to label the 3'-ends of oligonucleotides with $^{32}$P. The separation of the oligonucleotides and the principle behind their sequence analysis are similar to those for the 5'-labeled oligonucleotides except that the 5'-exonuclease used for partial digestion is spleen phosphodiesterase (98). Besides providing an alternate approach to the use of polynucleotide kinase for sequencing oligonucleotides, an important application of this method could well be in conjunction with polynucleotide kinase for sequencing long oligonucleotide fragments (15 or longer), which are occasionally found in total T\textsubscript{1}-digests of an RNA (99).

SEQUENCE ANALYSIS OF 5'- AND 3'-END LABELED RNAs A procedure for deriving the sequence of 20–25 nucleotides from each end of a tRNA and requiring no more than a few micrograms of tRNA has now been developed (M. Silberklang, A. Gillum and U. L. RajBhandary, in preparation). For the 5'-end, this involves labeling of the tRNA with $^{32}$P at the 5'-end with polynucleotide kinase followed by partial digestion of the 5'-labeled RNA with nuclease P\textsubscript{1}, a relatively nonspecific endonuclease from *Penicillium citrinum* (100a). The labeled oligonucleotides are separated by two-dimensional homochromatography and their sequence deduced as described previously (85, 86, 91). Exactly the same principle is used in the sequencing of the 3'-end except that the 3'-end is first labeled with $^{32}$P using tRNA nucleotidyl transferase (15).

GENERAL FEATURES OF tRNA SEQUENCES\textsuperscript{1} As of this writing, the sequences of about 75 different tRNAs are known (90, 91, 92, 101–114, 116, 117, 121; B. Dudock, personal communication; G. Dirheimer, personal communication; A. Gillum, L. Hecker, W. Barnett and U. L. RajBhandary, unpublished).\textsuperscript{2} This list includes tRNA sequences for all 20 amino acids except asparagine. While most of these are from yeast or *E. coli*, some of the more recent ones sequenced have been from *Bacillus stearothermophilus*, *Bacillus subtilis*, *Staphylococcus*, *N. crassa*, wheat germ, salmon, chick cells, mammals, and human placenta. In the case of tRNA\textsuperscript{Phe} and tRNA\textsuperscript{Met}, for which sequences from several

\textsuperscript{1}The nucleosides and bases are indicated by the usual symbols C, G, A, U, T, and $\Psi$ (pseudouridine). The molecular structure and the numbering system for the four major bases in tRNA are shown in Figure 1. Modifications are designated by symbols such as m\textsuperscript{7}G\textsubscript{46}, which indicates a methyl group on position 7 of guanine residue 46; m\textsuperscript{2}G\textsubscript{26} indicates two methyl groups on nitrogen 2 of guanine 26. Methylation of the 2'OH of ribose is indicated by an "m" after the symbol such as C\textsubscript{32}m. Watson-Crick base pairs are designated by a single dot, thus G\textsubscript{c}.

\textsuperscript{2}Corrected sequence of yeast tRNA\textsuperscript{Val} cited in Ref. 9.
mammalian sources are known, these have been found to be identical. It is, therefore, possible that the sequences of most if not all mammalian tRNAs may have been conserved. Similarly, the sequence of tRNA\textsuperscript{Trp}, which is used as a primer for DNA synthesis by avian myeloblastosis virus reverse transcriptase, may be identical to the corresponding tRNA from duck, mouse, rat, and human sources but different from \textit{E. coli} and from lower eukaryotes (31, 122). In the case of eukaryotic cytoplasmic initiator tRNAs, the sequences may be even more strongly conserved, since it has been shown that these tRNAs from salmon liver and testes (91) have essentially the same sequence as that from human placenta (92) and from rabbit, sheep, and mouse myeloma (123, 124).
Generalized Secondary Structure for tRNAs

The most striking aspect of all tRNAs that have been sequenced is that they can all be accommodated into the cloverleaf folding first proposed by Holley et al (38) as one of the possible secondary structures for tRNAs. The basic feature of this structure (Figure 2) is the folding back of the single polynucleotide chain upon itself with the formation of double helical stems and looped-out regions. Except for an

![Diagram of tRNA structure](image)

**Figure 2** A diagram of all tRNA sequences except for initiator tRNAs. The position of invariant and semi-invariant bases is shown. The numbering system is that of yeast tRNA^Phc_. Y stands for pyrimidine, R for purine, H for a hypermodified purine. R^{+}_{15} and Y^{+}_{48} are usually complementary. As noted in the text, positions 9 and 26 are usually purines, while position 10 is usually G or a modified G. The dotted regions α and β in the D loop and the variable loop contain different numbers of nucleotides in various tRNA sequences.
occasional G·U base pair or a mismatch (not shown in Figure 2), the stems are held together by Watson-Crick base pairs. The widespread occurrence of these stem regions led to the general assumption that their structural basis was an RNA double helix, which became evident with the tracing of the polynucleotide chain of yeast tRNA\textsuperscript{Phe} (1). All tRNAs contain four loops: dihydrouridine loop (D loop or loop I), anticodon loop (loop II), variable loop (loop III), and TψC loop (loop IV). Four of the stems are common to all tRNAs: acceptor stem, dihydrouridine stem (D stem), anticodon stem, and TψC stem; a fifth stem is present only in tRNAs that contain a long variable arm. For convenience, a loop and a stem are commonly referred to as an arm.

In the cloverleaf arrangement of tRNAs, the acceptor stem, the anticodon arm, and the TψC arm are constant in all tRNAs. The acceptor stem consists of seven base pairs and four nucleotides, including the 3'-terminal CCA sequence protruding at one end; the anticodon arm and the TψC are each made up of five base pairs and a loop of seven nucleotides. Thus, the large difference in the size of various tRNAs, which range from 73 to 93 nucleotides, is accounted for by variation in only two regions of the cloverleaf structure, the D arm and the variable arm. The D arm consists of 15–18 nucleotides, with three or four base pairs in the stem and 7–11 nucleotides in the loop. As discussed below, there is evidence that the fourth base pair in the D stem is stacked into the molecule and probably hydrogen-bonded even when the two bases do not form a Watson-Crick base pair. Accordingly, variation in the length of the D arm can be understood in terms of two regions in the D loop (a and β in Figure 2), which flank the two constant guanine residues and have variable numbers of nucleotides (125). These regions contain one to three nucleotides; most of them are pyrimidines with a high proportion of dihydrouracil residues. The variable arm is limited to two classes: (a) those which contain four or five bases in the loop with no helical stem or (b) those which contain a large variable arm consisting of 13–21 residues.

Three of the published tRNA sequences, yeast tRNA\textsuperscript{Gly} (126) and tRNA\textsuperscript{Val} from Torula yeast (127) and brewers’ yeast (128), contain only three nucleotides in the variable loop. The sequence of brewers’ yeast tRNA\textsuperscript{Val} has been recently reexamined and shown to contain five nucleotides in the variable loop (117). It is, therefore, possible that tRNA\textsuperscript{Val} from Torula yeast may also have five nucleotides in the variable loop. Folding of the polynucleotide chain determined for yeast tRNA\textsuperscript{Phe} (1) requires that the variable loop contain a minimum of four nucleotides (125, 129, 130, 131). In view of this, it would clearly be desirable to reexamine the sequence of yeast tRNA\textsuperscript{Gly} (126).

Based on the two variable regions of the cloverleaf structure, tRNAs sequenced to date can be fitted into three classes essentially similar to those proposed originally by Levitt (132). These include class I with four base pairs in the D stem and four or five bases in the variable loop (D\textsubscript{IV}V\textsubscript{4-5}); class II with three base pairs in the D stem and four or five base pairs in the variable loop (D\textsubscript{IV}V\textsubscript{4-5}); and class III with 3 base pairs in the D stem and a large variable arm (D\textsubscript{IV}V\textsubscript{M}). Since it appears not too important to differentiate three or four base pairs in the D stem (125), it is perhaps reasonable to use a simpler classification (131) based only on the size of the variable
arm, class 1 with 4 or 5 bases in the variable loop and class 2 with a large variable arm (13–21 bases).

**Invariant and Semi-invariant Nucleotides in tRNAs**

In addition to the general accommodation of all tRNAs into a common cloverleaf structure, tRNAs contain several invariant and semi-invariant residues located in the same relative position in all tRNAs. In Figure 2, these are indicated by the common nucleoside symbols A, C, U, G, T, ψ, etc for the invariant residues. Semi-invariant residues are indicated by R for purines, Y for pyrimidines, and H for a usually highly modified purine nucleoside located on the 3'-side of the anticodon. The numbering system used is that for yeast tRNA^Phe^, which belongs to class I and is 76 nucleotides long (133).

Except for initiator tRNA, which is discussed separately below, 15 of these invariant residues are present in almost all tRNAs that are active in protein synthesis. These are U₈, A₁₄, G₁₈, G₁₉, A₂₁, U₃₃, G₅₃, T₅₄, ψ₅₅, G₆₆, A₅₈, C₆₁, C₇₄, C₇₅, and A₇₆ at the acceptor end. U₈ may be s₄U in *E. coli* tRNAs, and A₅₈ is often m¹A in tRNA from eukaryotic sources; G₁₈ may be G₃ depending upon the individual tRNA, and more recent studies have shown that T₅₄ may be U, T₃, s²T, or ψ (130, 134–136). The eight semi-invariant residues present in almost every tRNA active in protein synthesis are Y₁₁, R₁₅, R₂₄, Y₃₂, H₃₇, Y₄₈, R₅₇, and Y₆₀. Most tRNAs contain a purine at position 9 (six exceptions), G or modified G at position 10 (three exceptions), and a purine at position 26 (four exceptions). Y₁₁ and R₂₄, noted recently as semi-invariant residues (137), are part of the D stem and form a Watson-Crick base pair; they are, therefore, correlated invariants. Thus, when Y₁₁ is C, R₂₄ is G and when Y₁₁ is U, R₂₄ is A. Besides prokaryotic initiator tRNAs (see below), the only exception to this is *E. coli* tRNA^Trp^, which has U₁₁ and G₂₄; it is worth noting that mutation of G₂₄ to A₂₄ enables this tRNA to suppress the terminator codon UGA without a concomitant change in the anticodon sequence of this tRNA (138). Another pair of correlated invariants first pointed out by Levitt (132) is R₁₅ and Y₄₈. As discussed below, we now know the structural role played by 20 of the 23 invariant and semi-invariant residues in maintaining the tertiary structure of tRNAs.

A few exceptions to the generalized cloverleaf structure and particularly the invariant and the semi-invariant residues in the structure do, however, exist. The most notable exception is provided by a class of glycine tRNAs (tRNA_Gly^species) from staphylococci (25) that are used for cell wall biosynthesis and are inactive in protein synthesis (139). While they do conform to the general folding scheme of the cloverleaf structure, several of the invariant or semi-invariant residues are missing in these tRNAs. Thus, G₁₈ and G₁₉ are both replaced by U residues, H₃₅ by either C or U and ψ₅₅ by G. In some strains of staphylococci, tRNA_Gly^ contains U in place of G₁₀ and also U₅₆ instead of C₅₆. Other tRNAs differ from the generalized structure of Figure 2 in a few minor respects; these include *E. coli* tRNA^His^ (49), *E. coli* tRNA^Leu^ (141–143), tRNA^Met^ from mouse myeloma and brewers' yeast (104, 107), the frame shift suppressor tRNA^Gly^ from Salmonella (144), and tRNA^Val^ from mouse myeloma (106).
Unique Features in Initiator tRNA Sequences

Both prokaryotic and eukaryotic initiator tRNAs conform to the general cloverleaf scheme of folding and contain almost all of the invariant and semi-invariant bases mentioned above (86, 116, 146–148). However, they possess certain unique features in their sequences that can be used to distinguish them as a class both from each other and from non-initiator tRNAs. The distinguishing feature of prokaryotic initiator tRNAs including those of *E. coli* (86), the blue-green alga *Anacystis nidulans* (146), *Streptococcus faecalis* (147), *B. subtilis* (116), mycoplasma (148), and *Thermus thermophilus* (S. Nishimura, personal communication) is that they all lack the Watson-Crick base pair at the end of the acceptor stem between the first nucleotide of the 5'-end to the fifth nucleotide from the 3'-end. In these six prokaryotic initiator tRNAs, the 5°-terminal nucleotide is C, whereas the nucleotide opposite it in the acceptor stem is C in *A. nidulans* (146) and A in the other five. The possible importance of this feature in the function of these prokaryotic initiator tRNAs is underscored by the fact that the change from 5'-terminal A to C in the case of *A. nidulans* initiator tRNA still preserves the lack of Watson-Crick base-pairing in this region.

B. Baumstark, S. T. Bayley, and U. L. RajBhandary (unpublished) have recently examined the terminal sequences of an initiator methionine tRNA from *Halobacterium cutirubrum*, a prokaryotic organism that is an exception to the general rule that all prokaryotic organisms utilize a formylated Met-tRNA for the initiation of protein synthesis (20, 150). In contrast to the other prokaryotic initiator tRNAs that use fMet-tRNA for initiation, *H. cutirubrum* initiator tRNA contains an A-U base pair at the end of the acceptor stem. This suggests that one of the functions of the unusual sequence feature of prokaryotic initiator tRNAs discussed above is related to their mode of utilization in vivo for protein synthesis (151). Additionally, it is interesting to note that all of the eukaryotic cytoplasmic initiator tRNAs, which like the *H. cutirubrum* initiator tRNA initiate protein synthesis with Met-tRNA but without formylation, contain an A-U base pair at the end of the acceptor stem. The functional significance of this unusual coincidence between the *Halobacterium* and eukaryotic initiator tRNAs is not known.

Another sequence feature unique to the prokaryotic initiators whose total sequences are known (86, 116, 146) is that they contain a A11-U24 base pair in the D stem in contrast to a Pyr11-Py4 Watson-Crick base pair found in all other tRNAs. The relationship, if any, of this feature to their function or to the unusual sequence feature at the end of the acceptor stem is unknown.

The most unusual feature of eukaryotic cytoplasmic initiator tRNAs is that they lack the invariant sequence Tψ and contain AU or AU* in the case of wheat germ tRNA. An additional difference from the general structure of Figure 2 is the presence of A at the end of the TψC loop instead of a pyrimidine nucleoside. In fact, the sequence of this entire loop—AU(U*)CGm1AAA—has been preserved in all of the eukaryotic cytoplasmic initiator tRNAs that have been examined, including those from yeast (152), wheat germ (153), *N. crassa* (A. Gillum, J. Hecker, A. Barnett, and U. L. RajBhandary, unpublished), salmon testes and liver, rabbit liver
(124), sheep mammary gland (124), mouse myeloma (123), and human placenta (92). The possible significance of this feature in the function of these eukaryotic initiator tRNAs has been discussed elsewhere (39, 58, 91).

Finally, another exceptional feature in the sequence of some, although not all (91, 92, 123, 124, 153), eukaryotic cytoplasmic initiator tRNAs is that the anticodon sequence CUA is preceded by C rather than by U as in all other tRNAs.

MOLECULAR STRUCTURE OF NUCLEIC ACID COMPONENTS AND DOUBLE HELICAL NUCLEIC ACIDS

Three types of X-ray diffraction studies that have been carried out on nucleic acids have yielded important structural information. These are single-crystal studies of nucleic acid components, polynucleotide fiber studies, and finally single-crystal analyses of macromolecular nucleic acids. These are interrelated in an important fashion, since information obtained from one type of study is used to interpret the results from another study.

During the last 25 years an impressive number of single-crystal analyses have been made of nucleic acid components so that we now have firm information about the molecular geometry of purines, pyrimidines, and nucleotides as well as their intermolecular complexes. In particular, these studies have given us information about the structural chemistry and potentialities for hydrogen bonding between the purines and the pyrimidines. Many types of hydrogen bonding are found in these crystal studies, including, but by no means confined to, the familiar Watson-Crick pairing found in double helical nucleic acids. These studies have been extensively reviewed (154–157). Bases are found joined to each other by one, two, or three hydrogen bonds and they are usually nearly coplanar.

Fiber diffraction studies provide other types of information, especially dealing with the conformation of the backbone and the types of hydrogen bonding that are consistent with periodic repeating structures. Studies of double helical RNA (158–160) and of its synthetic polynucleotide relatives (see reviews 155, 161–164) provide a background of information about the conformation of the ribose-phosphate backbone. These model systems can form two-, three-, or four-stranded helical complexes, the exact nature of which is determined by the hydrogen-bonding capabilities of the purine or pyrimidine side chains. Again, these studies underline the importance of other types of hydrogen bonding. For example, the first variant beyond Watson-Crick hydrogen bonding was described in 1957 for the three-stranded molecule consisting of one strand of poly(rA) and two strands of poly(rU) (165). It was pointed out that the second uracil residue could form H bonds with the amino group of adenine (N6) and the imidazole N7. This type of bonding was later confirmed in a single-crystal study by Hoogsteen (166) of the complex 9-methyl adenine and 1-methyl thymine. This is relevant because a form of this type of hydrogen bonding (reversed Hoogsteen pairing) is found in two places in the yeast tRNA_Phe structure (129, 130).

Further details of double helical organization have become available through studies of self-complementary dinucleoside phosphates, which form RNA double...
helical fragments in a crystalline lattice. The GpC (167, 168) and ApU (169) molecules form antiparallel right-handed double helices with Watson-Crick pairing between the complementary bases. Both of these structures were solved to atomic resolution and thus made it possible to obtain precise information not only about the geometry of the backbone, but also about the detailed organization of water in these heavily hydrated crystals. This was the first time that the Watson-Crick hydrogen bonding between adenine and uracil (or thymine) had been seen in a single-crystal analysis (169). Prior to that, only the Hoogsteen pairing (166) had been seen in single crystals. Another feature of the ApU single-crystal analysis was the presence of a sodium ion complexed in the minor groove of the double helix to the uracil carbonyl O2 atoms (169). Other dinucleoside phosphates have been crystallized in different conformations. This includes the protonated form of UpA (170–172) as well as ApU and UpA complexed to planar aromatic molecules (173, 174).

One of the remarkable features of the double helical ApU and GpC structures is the fact that they form a double helix with backbone torsional angles very close to those found in the polymeric double helical RNA (167). The stereochemistry of the polynucleotide chain has been studied (175–178), and it has become clear that the RNA backbone is far more constrained than the DNA backbone, with restricted rotation about the nucleotide residues (176).

The fact that the DNA backbone can adopt a number of conformations while the RNA backbone is limited to a rather narrow range of conformational angles is clearly an expression of the added bulkiness of the hydroxyl group attached to C2' in ribose, which stiffens the backbone. The RNA helix does not change very much when salt or water content is altered (154, 158–160, 179), in marked contrast to the many different forms of the DNA double helix. Because the characteristic RNA helical conformation is seen even with dinucleoside phosphates (167, 169), one could then expect to find somewhat similar conformations in the short stem regions of the tRNA molecule. This expectation was indeed borne out in the three-dimensional structure of yeast tRNA\(^{\text{Phe}}\), which shows torsion angles in the stem regions (2) that are very similar to those seen in the dinucleoside phosphates and in extended fibers of duplex RNA (154).

Most biochemists are familiar with the external form of the double helical DNA, which has a major and a minor groove. In the normal B form of DNA, the bases are intersected by the axis of the molecule, are stacked perpendicular to it, and form a central pillar around which the sugar phosphate chains are coiled. In duplex RNA no bases are found on the helical axis. Instead, the base pairs are tilted 14–15° from the helix axis, and are located away from the center (154). The RNA double helix has 11(A) or 12(A') base pairs per turn with a rise per residue of 2.8–3 Å. This has the effect of causing a marked difference between the major and the minor groove; the minor groove virtually disappears as the bases are close to the surface of the molecule, while the major groove is enormously deepened. If one looks down the axis of the RNA double helix (180), one sees a hole down the center of the molecule approximately 6 Å in diameter, which contains no material other than water. The
RNA double helix may thus be described as sort of a flat ribbon wound around a central region 6 Å in diameter. Similar geometry is found in the helical stems of tRNA.

CRYSTALLIZATION OF tRNA

The major method for determining the three-dimensional structure of large molecules is X-ray diffraction. The techniques and methodology of large-molecule diffraction studies have been developed during the last 20 years largely for application to crystalline proteins, and during this period about four dozen protein structures have been solved. However, prior to 1968 no macromolecular nucleic acid had been prepared in the form of a single crystal suitable for X-ray diffraction analysis. Nucleic acids and synthetic polynucleotides had been studied in oriented fibers, some of which had crystallized. However, these are not single crystals, and most of the techniques of single-crystal diffraction analysis could not be applied to them.

In 1968 five different groups reported the crystallization of tRNA (181–185), and three reported single crystals large enough for X-ray diffraction studies. Several different tRNAs formed single crystals, including *E. coli* tRNA[^Met] (182), *E. coli* tRNA[^Phe] (183), and yeast tRNA[^Phe] (184). Immediately there was a great surge of enthusiasm among workers in the field since they felt it would only be a short time before the structure of these crystals could be determined. Unfortunately, the best of these crystals barely diffracted to 6-Å resolution. Experience with crystalline proteins suggested that an electron-density map of 3-Å resolution was needed in order to accurately trace the polypeptide chain, although there was reason to believe that a polynucleotide chain could be traced at a somewhat lower resolution due to the electron-dense phosphate groups. However, there was little likelihood that studies at 6-Å resolution would be very useful in determining more than the overall size and packing of the molecules.

These early results stimulated an intensive study of the crystallization of tRNA (186–191). This work was implemented considerably by the availability in large quantity of several purified tRNA species (192). In addition, micro methods of crystal growing were developed and were useful in attempting to find suitable crystallization conditions that consumed only small amounts of tRNA (187). In the few years following the initial tRNA crystallization, a variety of crystal forms were reported involving several different tRNAs (193–197, 212). Two generalizations began to appear from the large accumulation of data. First, it was very difficult to obtain highly ordered crystals, i.e. crystals with a regularity in their lattice that produced a diffraction pattern higher than about 6-Å resolution. Secondly, polymorphism was very common.

The resolution in a diffraction pattern is related to the regularity in the crystal lattice. In crystals of small molecules this regularity extends to the sub-angstrom region. In normal X-ray diffraction work, X-rays are generated using a copper anode (\(\lambda = 1.54\ \text{Å}\)) and the limit of resolution frequently used in small-molecule, single-crystal analysis is 0.77 Å. An electron-density map reconstructed from this diffrac-
Diffraction patterns of good crystalline proteins generally extend to 3 Å, sometimes to 2 Å, and in a few cases to less than 2 Å. The electron-density map generated from this data does not show individual atoms, but rather groups of atoms. Thus the electron-density map has to be interpreted in terms of molecular models. The exact geometry of the monomeric components—bond angles and distances, possible conformations of the residues—is usually obtained from single-crystal studies. This is true in the interpretation of electron-density maps of nucleic acids as well as proteins.

Crystalline tRNA in general does not form a lattice with regularities extending beyond 6 Å. This is a frustrating situation because an electron-density map calculated at a resolution of 6 Å is not generally interpretable, since individual bases or ribose groups are not discernible on a map of this resolution. It is not altogether clear why crystalline tRNAs generally have such low resolution. It is probably related to the polyelectrolytic nature of the molecule. tRNAs have 73–93 negative charges, and in order for them to be packed in a regular lattice, the positioning of the cations is quite important. Indeed, in the search for adequate crystals of tRNA, the composition and concentration of cationic species is of central importance in addition to the purity of the tRNA species.

Polymorphism is another feature of tRNA crystals. Thus, a single tRNA species will form many different crystalline lattices. Although this phenomenon is not uncommon in protein crystals, it is very common in tRNA. For example, yeast tRNA^Phe_, which has been examined extensively, crystallizes in at least a dozen different unit cells (184, 197, 198, and A. Rich, unpublished observations). New polymorphic forms are discovered by simply altering the crystallization conditions. Polymorphism is also found in crystals of other tRNA species (187, 188, 196, 212) by altering the crystallization conditions.

Crystallization of tRNA suggested that the molecule has a stable conformation, and this stimulated a variety of proposals concerning the three-dimensional conformation of the molecule (132, 199–204, reviewed in 205). It would be difficult to find a better subject for a theoretical study of conformation. This arises out of the fact that all tRNA sequences fit in the cloverleaf diagram and have many invariant or semi-invariant base positions. If one assumes double helical stems and varies the loop regions of the cloverleaf diagram, there are only a finite number of plausible conformations, and many of these have been presented in the molecular models. Other constraints on model building arise from the molecular outline based on low-angle X-ray scattering (206), the limitations derived from the crystal lattice dimensions, and the interesting result of the photo-induced cross-linking between the s^4U_8 and C^13 in a number of E. coli tRNAs (207). This cross-linking has the remarkable feature of maintaining the molecule in a form such that it still has amino acid acceptance activity and can be used within the ribosome in protein synthesis. This suggested that positions 8 and 13 are near each other, and this was incorporated into some models. It is worth noting here that most models incorporated some features that were eventually found in the three-dimensional structure of tRNA,
since the cloverleaf was usually assumed as the starting point with its double helical stems. However, none of the models created a three-dimensional structure similar to that seen in the final structure analysis. In retrospect the failure to predict a useful model undoubtedly reflects the fact that not enough attention was focused on the invariant nucleotides, as almost all of them play a structural role in the three-dimensional structure. In addition, the model builders relied almost exclusively on Watson-Crick hydrogen bonding, although the actual molecule has many other types of tertiary interactions.

**High-Resolution Crystals of Yeast tRNA\textsuperscript{Phe}\textsuperscript{phe}**

The first big breakthrough in the preparation of crystals of tRNA with a high-resolution X-ray pattern occurred in 1971 (208) when a group at MIT working with Rich reported that it was possible to prepare crystals of yeast tRNA\textsuperscript{Phe}\textsuperscript{phe} with a resolution of 2.3 Å (the pattern actually extends out to nearly 2 Å). The crystal form was orthorhombic, P\textsubscript{2}1\textsubscript{2}1\textsubscript{2}, with four molecules in the unit cell and one in the asymmetric unit. The unusual feature that they introduced was the use of the polycationic spermine as a means of neutralizing some of the negative charges in the polynucleotide chain. Crystals were prepared in 10mM MgCl\textsubscript{2}, 10 mM cacodylate buffer at neutral pH and 1 mM spermine hydrochloride. The crystals were brought out of solution by vapor equilibration of 2-methyl-3,4-pentanediol or isopropanol. Although hexagonal crystals of yeast tRNA\textsuperscript{Phe}\textsuperscript{phe} had been reported earlier (184, 209), these yielded only low-resolution diffraction patterns. The addition of spermine apparently stabilized yeast tRNA\textsuperscript{Phe}\textsuperscript{phe} to produce a well-ordered crystalline lattice. Spermine-stabilized yeast tRNA\textsuperscript{Phe}\textsuperscript{phe} also forms high-resolution crystals in other lattices. Monoclinic crystals of spermine-stabilized yeast tRNA\textsuperscript{Phe}\textsuperscript{phe} have been formed under conditions very similar to those reported for orthorhombic crystallization (198, 210, 211), and they produce a high-resolution X-ray diffraction pattern. Good diffraction patterns are also obtained from spermine-stabilized yeast tRNA\textsuperscript{Phe}\textsuperscript{phe} in a cubic lattice (198). Removal of the CCA-terminus of yeast tRNA\textsuperscript{Phe}\textsuperscript{phe} still permits it to crystallize in the presence of spermine to produce orthorhombic crystals with a good diffraction pattern (198). Thus at least four different crystalline forms of spermine-stabilized yeast tRNA\textsuperscript{Phe}\textsuperscript{phe} have been reported, and the structures of two of these crystal forms have now been described in detail. This allows us to answer the question of what effect is produced by putting the same molecule in two different crystal lattices.

**Solution of X-ray Diffraction Patterns Using Heavy-Atom Derivatives**

Macromolecular structures are generally solved through the method of multiple isomorphous replacement. Several different sets of diffraction data are collected from the same crystalline form where one crystal has only the macromolecule in it while the others have additional heavy atoms in the lattice. Ideally the heavy atoms should not distort the lattice, so that the crystals remain isomorphous. The heavy atoms introduce small changes in the intensity of the diffraction patterns, and from these the position of the heavy-atom derivatives can be determined. In this way it is possible to determine the phase of the individual diffracted rays of the native
A number of different methods for obtaining isomorphous derivatives have been attempted in many laboratories. The simplest method is that of diffusing into the hydrated crystal lattice a compound containing a heavy atom. For tRNA work, the atom should have at least 70 electrons and a high enough binding constant for particular sites in the molecule to give a reasonably high occupancy. One of the interesting limitations in this regard is the fact that it is relatively easy to interpret a single heavy atom, but much more difficult to interpret multiple heavy atoms, which may occupy four or five sites in the molecule. The discovery of the first heavy-atom derivative is thus of great importance because it provides rough phase information that facilitates the discovery of subsequent heavy atoms. Heavy atoms can also be introduced directly into the covalent structure of tRNA. This can be done, for example, by reacting heavy atoms with side groups such as the sulfur atoms that occur in various tP~NAs (213). Other possibilities include the introduction of derivatives in the CCA~end of the molecule. These can be chemically or enzymatically iodinated (214–216). Mercurated compounds (217) or the introduction of thiolated nucleotides (197, 218, 219) can also be used.

The first useful heavy-atom derivative of tRNA was developed by Schevitz (220) in an attempt to react a molecule with the 3'-terminus of tRNA where a cis diol group is present that is a potentially reactive site for osmium derivatives. An osmium bis pyridine derivative reacted with crystals of yeast tRNA$_{Met}$ and produced a 1:1 complex at a single site that could be located crystallographically. These crystals were analyzed biochemically, and it was found that the osmium was not reacting at the 3'-terminus but was reacting with a cytosine near the base of the anticodon stem (221). The MIT group tried a variant of this procedure using a bis-pyridyl osmate diester of ATP. The ATP osmium bis pyridine complex was diffused into the crystal and was shown to be lodged primarily in one site in the orthorhombic crystal (222) near the 3'-OH end (1). Subsequent analysis revealed that although there was one major site, there were two other minor sites that bound the osmium derivatives (129, 223). The same ATP osmium bis pyridine also provided a multiple-site derivative for the monoclinic crystal form of yeast tRNA$_{Ph^*}$ (130). The molecular structure of the bis pyridine osmate ester of adenosine has been determined, and the osmium is linked to both O2' and O3' (224).

The first isomorphous osmium derivative helped the MIT group discover the second important class of isomorphous derivatives, the lanthanides (222). Trivalent lanthanides are known to be effective substitutes for the magnesium ion in renaturing tRNA (225). The high degree of isomorphism found in the lanthanide derivatives is undoubtedly due to the fact that they replace individual magnesium ions in the lattice with only a minimum of distortion in the molecular packing. Lanthanides have an additional advantage for crystallographic studies in that they have a strong anomalous scattering component, which helps to improve the phases and simplifies the choice of the handedness of the enantiomorphs. Of the lanthanides, samarium has the largest anomalous component, and it was selected for use with the ortho-
rhombic crystals (222) to obtain both normal and anomalous phasing information in the orthorhombic crystal. It is interesting that lanthanides can also be used as spectral probes since they have fluorescent properties that are useful for energy transfer studies (226). In the orthorhombic lattice, samarium occupied four different sites (223). A number of other derivatives were found for the orthorhombic lattice including Pt(II) (222) and Au(III) (A. Rich, unreported observations).

In the spermine-stabilized monoclinic crystal of yeast tRNA_{Phe}, Robertus et al (130) initially used the same ATP-Os-bis pyridine complex and lanthanides [Lu(III) as well as Sm(III)] as were used in the orthorhombic crystals (222) plus a trans PtCl₂(NH₃)₂ derivative that was bound covalently to the anticodon end of the molecule (227). Subsequently a mercurial derivative (hydroxy mercuri-hydroquinone-OO-diacetate) was also used (137).

SOLUTION OF THE YEAST tRNA_{Phe} STRUCTURE BY X-RAY DIFFRACTION

Folding of the Polynucleotide Chain at 4-Å Resolution—1973

Using osmium, samarium, and platinum derivatives, the MIT group produced a three-dimensional electron-density map at 4-Å resolution in early 1973 (1). Although segments of the polynucleotide chain could be seen in an earlier 5.5-Å resolution map (222), it was impossible at that stage to trace the chain. At 5.5-Å resolution, large areas in the lattice were seen in which the aqueous solvent was sharply delineated from the tRNA molecule as a whole. Part of the molecular outline could be discerned, although it was impossible to separate the molecules especially around the twofold screw axis. However, at 4.0-Å resolution more detail could be seen and an envelope of nearly zero electron density could be seen surrounding most of each individual molecule. The molecule that had seemed elongated at 5.5-Å resolution (222); was now clearly seen in a bent, L-shaped form. There were about 80 prominent peaks seen in the electron-density map, and since the chain had 76 nucleotides, it was surmised that all of the electron-dense phosphate groups of the nucleotides were seen in the map. A number of features made it possible for the chain to be traced. Several sections of the electron-density map showed two chains winding around each other in the form of a right-handed double helix with weaker connecting regions of electron density (1). These were interpreted to be the four stem regions of the cloverleaf. At one end of the molecule, four peaks in a row extended out from the body of the molecule, which was believed to be the 3'-ACCA-end of the polynucleotide chain. This interpretation was strengthened by the fact that the osmium derivative appeared about 7 Å from the terminal residue, a position that it would occupy if it were complexed to the cis diol of the terminal ribose. The molecule was found to be somewhat flattened about 20–25 Å thick, and the two limbs of the L were oriented more or less at right angles to each other. Most of the chain tracing was unambiguous in that the electron-dense phosphate groups were seen to be an average of 5.8 Å apart, very close to that which is anticipated in an RNA double helix (154). The acceptor stem and the ÍyC stem were found to be virtually colinear, forming one limb of the L with 12 base pairs. The other limb
contained the D stem and anticodon stem, but they were not quite colinear. The
anticodon was found at the end of that limb. A perspective diagram of the chain
tracing is shown in Figure 3, illustrating the folding of the polynucleotide chain seen
at 4-Å resolution. An unusual coiling was found at the corner of the molecule where
the D loop overlapped the T\psi C loop. The polynucleotide chain was found to have
a very sharp bend in the vicinity of residues 9, 10, and 11. This had the net effect
of bringing residue 8 rather close to residue 13, which was in agreement with the
carlier studies on photo-induced cross-linking of residues \( s^4 U_8 \) and \( C_{13} \) (207). It was
surmised that bases 8 and 13 were close enough to form the photodimer. This folding
of the polynucleotide chain had not been anticipated by any of the model builders,
and it has been verified by higher-resolution analysis in both the orthorhombic (129)
and the monoclinic lattice (130).

Although most of the chain tracing was unambiguous, there were a few regions
where the chains came close enough together so that alternative tracings were
possible at this resolution; however, only one of the possible chain tracings was
compatible with the cloverleaf diagram.

It was pointed out that the electron density spanned by the five nucleotides in the
extra loop had a somewhat erratic course and covered a distance that could be
spanned by as few as four nucleotides (1). In addition, since the variable loop was
at the surface of the molecule, it could of course accommodate a much larger extra

---

**Figure 3** The folding of the polynucleotide chain of yeast tRNA\(^{Phe}\) as revealed by the 4-Å
electron-density map (1). In this perspective view the horizontal part of the L-shaped molecule
is rotated slightly toward the reader so that the acceptor stem is closer. It can be seen that
the D loop covers part of the T\psi C loop near the corner of the molecule.
loop. Even at that stage, the suggestion was clear that this was a folding of the molecule that could serve as a model for all tRNA structures.

An interesting feature of the orthorhombic crystals is the fact that they are unstable along one axis. The $a$ axis ($33 \, \text{Å}$) and $b$ axis ($56 \, \text{Å}$) are stable to a slight loss of water, but the $c$ axis ($161 \, \text{Å}$) is unstable and decreases in steps to $128 \, \text{Å}$, $117 \, \text{Å}$, and finally $109 \, \text{Å}$ (228). Since the diffraction pattern changed only slightly other than the change in spacings, this was interpreted as indicating that the molecules could slide over each other. In the initial analyses (1, 222) large aqueous channels found passing through the crystal parallel to the $a$ axis measured approximately $30 \times 40 \, \text{Å}$. These channels are gradually obliterated during the cell shrinkage, associated with a sliding of the molecules.

**Tertiary Interactions at 3-Å Resolution—1974**

Tertiary interactions are taken to mean the hydrogen bonds that occur between bases, between bases and backbone, and between backbone residues, except for the interactions in the double helical stem regions, which are considered secondary.

During 1974, 3-Å resolution analyses were published for yeast tRNA$^{\text{Phe}}$ in two different crystal forms, the orthorhombic (129, 223) from which the polynucleotide chain had been traced at 4-Å resolution and a monoclinic form (130). These results were very similar, but not identical. We describe the differences first, and then discuss the general structure of the molecule as defined by the more recent 2.5-Å analyses of both crystal forms.

Two papers were published in 1974 describing the 3-Å electron-density map of the orthorhombic lattice of yeast tRNA$^{\text{Phe}}$. The first was a preliminary paper (223), which essentially reinforced the general disposition of the parts of the polynucleotide chain that had been initially traced at 4-Å resolution (1). The tertiary structure was not described in detail, but some errors were subsequently found (129) in the tentative interpretation. In particular, incorrect residue assignments were made in the D stem and in the position of the Y base (223). These were corrected in the comprehensive interpretation of the electron-density map in the second paper (129).

A number of tertiary interactions were described involving the nucleotides in the loop regions that serve to stabilize the L-shaped form of the molecule (129). Several interactions were found involving bases hydrogen-bonded to the wide groove of the D stem, and the other interactions were found between bases hydrogen-bonded on either end of the stem. In addition, a series of interactions were found where the D loop was near the TΨC loop. An interaction that was subsequently modified on further inspection of the 3-Å map (125) was A$_{21}$, which was in the plane of residues of U$_8$ and A$_{14}$ and initially thought to be hydrogen-bonded to them. Further inspection showed that it was hydrogen-bonded to nearby ribose 8. The most striking feature of the tertiary interactions was the extent to which they involved many of the bases that are constant in all tRNA sequences (Figure 2). This made it likely that the structure of yeast tRNA$^{\text{Phe}}$ could serve as a useful model for understanding the three-dimensional structure of all tRNA structures (125).

At the same time in mid-1974, a 3-Å analysis of monoclinic crystals was reported by a group working with Klug (130), at the MRC Laboratory in Cambridge,
England. They used the same spermine-stabilized yeast tRNA\textsuperscript{Phe} as in the orthorhombic analysis. The method for preparing the monoclinic crystals (113, 114) is very similar to that used for crystallizing the orthorhombic form. In addition, since two of the cell dimensions were the same (33 Å, 56 Å), it suggested that the structures would have elements of similarity. Both crystal forms have 2\textsubscript{1} screw axes, with the major differences due to a head-to-head, tail-to-tail packing in the orthorhombic lattice, as opposed to a head-to-tail packing in the monoclinic lattice (229, 230). The overall analysis was very similar; however, several important differences were reported. The electron-density map of the monoclinic crystal could not be resolved completely. In particular, the region at the corner of the molecule where the D loop and T\PsiC loops came close together could not be interpreted at 3-Å resolution. Some differences were reported relative to the orthorhombic lattice; an important one concerned the interaction of T\textsubscript{54} with A\textsubscript{58} in the T\PsiC loop. The orthorhombic analysis clearly showed a reversed Hoogsteen pairing (129), while Robertus et al (130) reported a Hoogsteen interaction in the monoclinic form. This suggested that there might be a significant difference in the conformation of the T\PsiC loop and therefore a possible difference in the interaction of the T\PsiC loop and the D loop at the corner of the molecule. Another important difference was found in the region connecting the D stem with the anticodon stem. Analysis of the orthorhombic crystals (129) led to a hydrogen-bonding interaction between A\textsubscript{44} and m\textsuperscript{2} G\textsubscript{26}; Robertus et al (130) described residue m\textsuperscript{2} G\textsubscript{26} intercalating between A\textsubscript{44} and G\textsubscript{45}. Thus it appeared that there were significant differences between the form of the molecule in the two lattices.

**Tertiary Interactions and Coordinates at 2.5-Å Resolution—1975**

The results of a 2.5-Å analysis of yeast tRNA\textsuperscript{Phe} were published in 1975 for both the orthorhombic (231) and monoclinic (137) crystal forms, and atomic coordinates were reported for both forms (2–4). Thus we can compare in detail the structure of yeast tRNA\textsuperscript{Phe} in the two different lattices. In the higher-resolution analysis, further details of the molecular structure became visible. A number of interactions were found between the bases and the ribose-phosphate backbone as well as between various segments of the backbone. Preliminary atomic coordinates obtained from analysis of the multiple isomorphous replacement map were subjected to refinement calculations to varying extents. These calculations are designed to optimize the assignment of coordinates to produce normal bond angles and distances and at the same time to improve the fit of the molecule to the observed electron-density map.

At 2.5-Å resolution it is not possible to visualize atoms in the electron-density map, but it is possible to visualize clearly individual peaks associated with bases, sugars, and phosphate residues. Because of this, assignments can be made concerning the conformation of the sugar residues. Even though most of the ribose residues are in the normal 3'-endo conformation, a significant number are found to be in the 2'-endo conformation, particularly in those regions in which the polynucleotide chain is elongated or undergoes sharp bends (2–4).

The higher-resolution analysis of the orthorhombic crystals (231) generally reinforced the interpretations of the tertiary interactions seen at 3-Å resolution, and a
number of additional hydrogen-bonding interactions were described as discussed below.

The results of the monoclinic analysis at 2.5-Å resolution (137) also yielded a number of additional interactions. Furthermore, those regions of the electron-density map involving the interaction of the D and TΨC loops that had not been interpreted at 3-Å resolution could now be interpreted. Ladner et al (137) confirmed the interpretation that had been described for the orthorhombic lattice at 3-Å resolution (129) in terms of the hydrogen-bonding interactions between the D and the TΨC loop. In addition, they revised their interpretation of the interaction between T_{54} and A_{38} (4, 137), making it a reversed Hoogsteen pairing in agreement with that seen in the orthorhombic analysis (129). Finally, the region between the D stem and the anticodon stem was also revised in both the hydrogen-bonding and the stacking interactions in that region so that it now agreed with the results of the orthorhombic analysis (129). Thus the apparent differences between the structures in the two lattices that had been suggested at 3-Å resolution disappeared in the higher-resolution analysis. From a comparison of the atomic coordinates (2) it could be seen that only minor differences persisted in the conformation of the 3'-terminal residues C_{75} and A_{76}.

THREE-DIMENSIONAL STRUCTURE OF YEAST tRNA^Ph

In view of the virtually identical conformation of the molecule in both the orthorhombic and monoclinic lattice (2), this description applies to both studies. However, appropriate references will indicate the areas where differences have been reported.

Studies at 3-Å and 2.5-Å resolution showed more details of the somewhat flattened L-shaped molecule, with the acceptor and TΨC stems forming one limb while the D stem and anticodon stems formed the other. The tertiary hydrogen-bonding interactions between bases are shown on the cloverleaf diagram of Figure 4, which also indicates which of the bases are invariant or semi-invariant in chain-elongating tRNAs. Figure 5 is a diagram of both sides of the molecule, where the backbone is represented as a coiled tube and solid bars indicate base-base tertiary interactions. The details of the hydrogen bonding are shown more fully in Figure 6 and Table 1. The base-base hydrogen-bonding interactions involve one, two, or three hydrogen bonds, and in general they form a network that maintains virtually all of the bases of the molecule in two stacking domains corresponding to the two limbs of the bent molecule. As shown in Figure 4, there are ten tertiary interactions between bases, eight of which were visualized in the 3-Å analysis of the molecule in the orthorhombic lattice (129).

**Acceptor Stem**

The acceptor stem takes the form of an RNA A helix with nucleotides 73–75 at the 3'-ends in a conformation in which the bases are slightly stacked upon each other, especially at the 3'-end. The electron density at the 3'-end of the molecule is not as great as that found elsewhere in the molecule in the orthorhombic form (2, 3, 129); this may be the result of some disordering or thermal motion at this point.
Figure 4 The nucleotide sequence of yeast tRNA⁰⁰ (41). Tertiary base-base hydrogen-bonding interactions are shown by solid lines, which indicate one, two, or three hydrogen bonds (2, 129, 137). The invariant and semi-invariant positions are indicated by solid and dashed boxes around the bases. Y₃₇ is a hypermodified base.

In the orthorhombic crystal, the 3'-terminal A₇₆ is not stacked on C₇₅. In the monoclinic cell, residues 75 and 76 seem to be in a somewhat more extended form (2, 4). There is a slight perturbation in the acceptor stem where the base pair G₆₉•U₄ is held together by two hydrogen bonds in a wobble pairing (232) as shown in Figure 7. The nature of the perturbation may be due to a change of a torsion angle around ribose 4 (137), but the detailed description of this region will have to await further refinement.

TΨC Stem and Loop

The TΨC stem is stacked on the acceptor in a continuation of the RNA A helix within 12° of being colinear. There are some unusual conformations found in the TΨC loop region. The loop is stabilized by several interactions that elongate the loop and have the effect of bringing two parts of the polynucleotide chains closer together.
STRUCTURE OF TRANSFER RNA

Figure 5 A schematic diagram showing two side views of yeast tRNA\textsuperscript{Phe} (2, 231). 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 cross-rungs. Tertiary 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 and Table 1. Those bases that are not involved in hydrogen bonding to other bases are shown as shortened rods attached to the coiled backbone.

in the loop than they are in the double helical stem. This conformation is stabilized by a hydrogen-bonding interaction between C\textsubscript{61} (N4) and phosphate P\textsubscript{60} (2, 137). This interaction may account for the constant GC base pair that is found at the end of the T\textsuperscript{ψ}C stem in all tRNA sequences. Stacked on the pair C\textsubscript{61}G\textsubscript{53} is a reversed Hoogsteen pairing between T\textsubscript{54} and m\textsuperscript{1}A\textsubscript{58}. This accounts for the fact that a uracil derivative is present here in all tRNAs involved in polypeptide chain elongation. However, it does not explain why the methyl group of thymine is found here, since the uracil could work just as well in terms of hydrogen bonding. It is interesting that a mutant has been found that has no thymine in its tRNA (233) and seems to function normally in protein synthesis. Next to the T\textsubscript{54}A\textsubscript{58} pair is an interesting interaction between ψ\textsubscript{55} and G\textsubscript{18}. As shown in Figure 6 and described in Table 1, the O4 of ψ\textsubscript{55} appears to be within hydrogen-bonding distance of both N2 and N1 of G\textsubscript{18}, giving rise to the possibility of two hydrogen-bonding interactions with the same oxygen atom. Split hydrogen-bonding of this type has been seen in single-crystal studies of uracil derivatives (234). In addition, however, the N1 of ψ\textsubscript{55} forms a hydrogen bond to the phosphate oxygen of P\textsubscript{58} and in this manner stabilizes the rather tight turn of T\textsuperscript{ψ}C loop in this region. The stacking of ψ\textsubscript{55} is terminated or capped by the phosphate group P\textsubscript{57}. These interactions can be seen in the stereo diagram of the molecule shown in Figure 8.

The next base at this corner of the molecule is G\textsubscript{57}, which is stacked between G\textsubscript{18} and G\textsubscript{19} of the D loop. The furthermost corner of the molecule is formed by residues C\textsubscript{56} and G\textsubscript{19}, which form a Watson-Crick hydrogen-bonded pair (129, 137).
Figure 6  Tertiary hydrogen bonding in yeast tRNA\(^{\text{Phe}}\) (2, 3, 129, 137). Five-membered furanose rings are shown attached to the bases, and the circle with a dot indicates that the ribose-phosphate chain is coming toward the reader, while the circle with a cross indicates that it is going away from the reader.
## Table 1  Tertiary hydrogen bonding of bases in yeast phenylalanine tRNA\(^a\)

<table>
<thead>
<tr>
<th>Interaction</th>
<th>Description</th>
<th>Structural role</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>A. Base-Base Interactions</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. U(_8)-A(_14)</td>
<td>Reversed Hoogsteen pairing explains constant U(_8) and A(_14)</td>
<td>Stabilizes sharp bend near residues 9-10, maintains orientation of acceptor stem</td>
</tr>
<tr>
<td>2. A(_9)-A(_23)</td>
<td>Poly(rA) pairing of A(_9) with A(_23) in major groove of D stem</td>
<td>Stabilizes sharp bend near residues 9-10</td>
</tr>
<tr>
<td>3. m(^2)G(<em>{10})-G(</em>{45})</td>
<td>Single H-bond from G(<em>{44})N2 onto m(^2)G(</em>{10})O6 in major groove of D stem</td>
<td>Maintains interaction between variable loop and D stem</td>
</tr>
<tr>
<td>4. G(<em>{15})-C(</em>{48})</td>
<td>Trans pairing with two H-bonds, explains constant R(<em>{15}) and Y(</em>{48})</td>
<td>Stabilizes joining D stem with T(\Psi)C stem by stacking and hydrogen bonding with trans pairing required by parallel chain directions</td>
</tr>
<tr>
<td>5. G(<em>{18})-(\Psi)(</em>{55})</td>
<td>G(<em>{18})N2 and N3 both within H-bonding distance of (\Psi)(</em>{55}) may partially explain role of constant G(<em>{18}) and (\Psi)(</em>{55})</td>
<td>Maintains interaction of D loop and T(\Psi)C loop, key to possible interloop opening mechanism in protein synthesis</td>
</tr>
<tr>
<td>6. G(<em>{19})-C(</em>{56})</td>
<td>Watson-Crick pair (the only tertiary one), explains constant G(<em>{19}) and C(</em>{56})</td>
<td>Forms outermost corner of molecule, stabilizes interaction of D and T(\Psi)C loops</td>
</tr>
<tr>
<td>7. G(<em>{22})-m(^7)G(</em>{46})</td>
<td>Protonated m(^7)G(<em>{46}) donates two H-bonds to G(</em>{22}) in major groove of D stem</td>
<td>Interaction of extra loop and D stem also stabilized by electrostatic bond due to charged m(^7)G(_{46})</td>
</tr>
<tr>
<td>8. m(^2)G(<em>{26})-A(</em>{44})</td>
<td>Two H-bonds; involve m(^2)G(<em>{26})N1 and O6 with A(</em>{44})N1 and N6</td>
<td>Stabilizes the continuity of interactions from the D stem to the anticodon stem</td>
</tr>
<tr>
<td>9. C(<em>{m32})-A(</em>{38})</td>
<td>One H-bond from A(<em>{38})N6 to C(</em>{m32})O2, may explain constant Y(_{32}) in anticodon loop</td>
<td>Stabilizes the anticodon loop</td>
</tr>
<tr>
<td>10. T(<em>{54})-m(^1)A(</em>{58})</td>
<td>Reversed Hoogsteen pairing, partially explains constant T(<em>{54}) and role of constant A(</em>{58})</td>
<td>Maintains sharp bend in T(\Psi)C loop at residues 55-57</td>
</tr>
<tr>
<td><strong>B. Base-Backbone Interactions</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11. C(<em>{11})N4-S(</em>{9})O2(^')</td>
<td>H-bond in major groove of D stem, may explain constant Y(<em>{11})-R(</em>{24}) in D stem (137)</td>
<td>Stabilizes sharp turn near residues 9-10</td>
</tr>
<tr>
<td>12. G(<em>{18})N2-S(</em>{5})O2(^')</td>
<td>One H-bond to furanose ring O</td>
<td>Stabilizes interaction of D and T(\Psi)C loops</td>
</tr>
<tr>
<td>13. A(<em>{21})N1-S(</em>{6})O2(^')</td>
<td>A(_{21})N1 and N6 face R(_8), explains constant A</td>
<td>Stabilizes D loop and folding of chain</td>
</tr>
<tr>
<td>14. U(<em>{33})N3-P(</em>{3})O2</td>
<td>Explains constant U(_{33})</td>
<td>Stabilizes sharp turn in anticodon loop</td>
</tr>
<tr>
<td>15. (\Psi)(<em>{53})N3-P(</em>{5})O2</td>
<td>Partially explains constant (\Psi)(_{55})</td>
<td>Stabilizes sharp turn of T(\Psi)C loop</td>
</tr>
<tr>
<td>16. G(<em>{57})N7-S(</em>{5})O2(^')</td>
<td>Explains constant R(_{57}) (2)</td>
<td>Stabilizes interaction of D and T(\Psi)C loops by augmenting stacking interaction</td>
</tr>
<tr>
<td>17. G(<em>{57})N2-S(</em>{1})O2(^')</td>
<td>G(<em>{57}) amino group is near S(</em>{18})</td>
<td>Stabilizes interaction of D and T(\Psi)C loops by augmenting stacking interaction</td>
</tr>
<tr>
<td>18. G(<em>{57})N2-S(</em>{1})O2(^')</td>
<td>G(<em>{57}) amino group is near S(</em>{19})</td>
<td>Stabilizes interaction of D and T(\Psi)C loops</td>
</tr>
<tr>
<td>19. C(<em>{61})N4-P(</em>{6})O1</td>
<td>Explains constant G(<em>{53})-C(</em>{61})</td>
<td>Stabilizes T(\Psi)C loop</td>
</tr>
</tbody>
</table>

\(^a\)Standard symbols are used for referring to bases and to their atoms. R stands for a purine and Y stands for a pyrimidine. The symbol S is used for the ribose residue, and it is generally followed by an atom designation such as O2'. P stands for the phosphorus atom, and the atoms O1 and O2 are phosphate oxygens.

It is interesting that two of the residues in the T\(\Psi\)C loop, U\(_{59}\) and C\(_{60}\), are not involved in the stacking interactions of the rest of the loop, but rather are oriented almost at right angles to this and nucleate the base-stacking interactions, which extend through the D stem, the anticodon stem, and into the anticodon loop (125). C\(_{60}\) is buried in the molecule, and it has been suggested that there is insufficient room for a purine at that point (2).

In the T\(\Psi\)C loop there is a hydrogen bond between O2' of ribose 55 and N7 of G\(_{57}\) (2). G\(_{57}\) has two other hydrogen bonds between the amino group N2 and two other oxygen atoms of ribose 18 and 19 (2). A slightly different set of hydrogen bonds for G\(_{57}\) have been described for the monoclinic lattice (137), and the final details will have to await the results of further refinement.
Figure 7  The base pairing found in the acceptor stem between G4 and U69. This is the same as that proposed in the "wobble" pairing between codon and anticodon (232).

Figure 8  Stereoscopic view of yeast tRNA^{Phe}. This can best be viewed with stereoscopic glasses, which will fuse the two images. However, the two images can be fused without the use of glasses by simply relaxing the eye muscles until the three-dimensional image appears. This is the same view as that shown on the right side of figure 5, with the anticodon at the bottom and the 3'-acceptor end at the upper right. This model was made from a slightly refined version of the orthorhombic coordinates (2).
**D Stem and Loop**

In describing this part of the molecule it is useful to describe the D loop and its interactions with the TψC loop first and then describe it in descending order relative to the diagram in Figure 5. The two constant G residues, G18 and G19, are involved in hydrogen-bonding interactions with the TψC loop as discussed above. As mentioned earlier, these are flanked by residues in the two variable segments α and β of the D loop, containing D16, D17, and G20, respectively. The backbone in the region of D16 and D17 arches away from the molecule to form an enlarged protuberance with the D residues on either side of the loop (Figure 4). These are not involved in the stacking interactions of that limb and the same is true for the residue G20, which lies more or less in the center of the D loop with the base at right angles to the rest of the other bases in the stem.

One of the remarkable features of the D stem is the way in which it is integrated in a series of stacking interactions that extend from the TψC loop all the way down to the anticodon. As mentioned above, this stacking is initiated by two of the bases in the TψC loop; C60 is stacked above U59 and U59 in turn is stacked on the base pair G15oC48. Stacking interactions for the entire molecule are shown schematically in Figure 9. G15 from the D loop has trans pairing with C48 of the extra loop (Figure 6). This is related to the fact that the polynucleotide chains are running parallel to each other at this point (235). On the plane immediately below this are three bases: U8 and A14, which are hydrogen-bonded in a reversed Hoogsteen interaction and A21, which is hydrogen-bonded to the backbone at ribose 8.

The U8•A14 pair lies immediately above the base pair C13•G22; the relative position of these bases is shown in Figure 10. In 1969 it was shown that s4U8 and C13 formed a photodimer and that the molecule was still biologically active (207). In a photochemical study, Bergstrom & Leonard (236) predicted that the C4-C5 bond of C13 should be parallel to the carbon sulfur bond of s4U. This prediction is amply borne out as shown in Figure 10, where s4U is drawn instead of the uracil that is found in yeast tRNA^phe^.

One of the unusual features of the D stem is the fact that there are hydrogen-bonding interactions with all of the base pairs in the major groove. Two of these interactions are shown in Figure 6; the protonated m7G46 bonds to G22 and immediately below that A9 bonds to A23. The A9•A23 pairing is the same as found in double helical poly(rA) (237), while the total complex of U12, A9, and A23 is analogous to the pairing of phenobarbital with 9-ethyl adenine (238). In the base pair immediately below there is a hydrogen-bonding interaction between O2' of ribose 9 with N4 of C11 reported for the monoclinic crystal (137). Since the hydroxyl group can either donate or receive a hydrogen bond, this may account for the constant pyrimidine-purine pair that is seen in this position in the D stem. Finally, the bottom base pair of the D stem (Figure 5) has a single hydrogen-bonding interaction between O6 of G10 and N2 of G45 from the extra loop. Immediately below this (Figure 5) there is a propeller-like orientation of the m7G26 and A44, which are held together by two hydrogen bonds (Figure 6).
Figure 9 A diagram illustrating the hydrophobic stacking interactions between various nucleotides in yeast tRNA<sup>Phe</sup>. The nucleotides are represented by letters, and the short solid lines represent the bases. The boxes between the nucleotides represent stacking interactions where both full and partial stacking is shown. If nucleotides are adjacent to each other in the polynucleotide chain, they are connected by a thin solid line. The connectivity of the molecule is indicated by the numbering only. All of the nucleotides involved in stacking interactions are shown in the diagram.
The net result of these interactions is to stabilize the variable loop in the center of the molecule with a series of four different interactions involving four different bases. The only base not involved in a hydrogen-bonding and stacking interaction is U₄₇, which projects away from the rest of the molecule, with a slight protuberance of the polynucleotide chain at that point.

**Anticodon Stem and Loop**

In contrast to the many interactions of the D stem, the anticodon stem is relatively simple. It is not colinear with the D stem, and the axes may deviate by as much as 25°. The anticodon loop has a conformation that is similar in some respects to the TΨC loop. Immediately below the bottom base pair A₃₁-U₃₉ (Figure 5) there is a continuation of stacking interactions that involve residues A₃₈ through G₃₄ so that one side of the anticodon loop is stacked while the other is not. This is somewhat similar, but not identical, to the conformation suggested by Fuller & Hodgson (239). The stacked bases in the anticodon loop do not continue the form of a double helix; the loop itself is rather narrow. There appears to be a single hydrogen-bonding interaction between the N₆ of A₃₈ and O₂ of C₃₂ (2). The hypermodified Y base is stacked below the A₃₆; however, the side chain of the Y base has a lower electron density (2, 3, 137) and this may be associated with some mobility in the lattice. The three anticodon bases A₃₆, A₃₅, and G₃₄ are stacked upon each other in a manner that is greater than would be found in one strand of an RNA double helix. The three bases are readily accessible for hydrogen bonding. A view of the anticodon and the Y base is shown in Figure 11. Immediately after G₃₄ the polynucleotide chain takes a sharp bend. This bend is stabilized by the hydrogen-bonding interaction from N₁ of U₃₃ to P₃₆ (2, 4), and the phosphate group of P₃₅ is stacked upon U₃₃. It should be pointed out that the participation of U₃₃ in hydrogen bonding with P₃₆ is quite analogous to the role played by ψ₅₅ in which N₁ is hydrogen-bonded to P₅₈. These two residues, which are constant in all chain-elongating tRNAs, appear to have similar roles in stabilizing a sharp turn of the polynucleotide chain.
GENERAL STRUCTURE OF OTHER tRNA MOLECULES

One of the striking features of tRNA sequences is the fact that they can all be arranged in a cloverleaf configuration; the physical basis for this was understood with the discovery that the stem regions form double helical RNA segments in the L-shaped molecule (1). In the sequence of yeast tRNAPhe there are 23 positions that are invariant or semi-invariant and are occupied by constant bases, or constant purines or pyrimidines. As described above and in Figure 5 and Table 1, a structural reason for 20 of the 23 invariant position is now understood. Only the terminal CCA invariants have no apparent structural role, and they are undoubtedly involved in the ribosomal and synthetase interactions. Since these invariants are found in all tRNA molecules that elongate polypeptide chains, this clearly suggests that the structure of yeast tRNA\textsuperscript{Phe} is a good model for understanding the three-dimensional structure of all these tRNAs (125).

As shown in Figure 2, there are three regions in tRNA sequences that have variable numbers of nucleotides. These are the regions labeled $\alpha$ and $\beta$, which flank the two constant guanines ($G_{18}$ and $G_{19}$). The structure of tRNA\textsuperscript{Phe} provides a clue to the manner in which variable numbers of nucleotides in the $\alpha$- and $\beta$-regions of the D loop can be accommodated within the structure. Regions $\alpha$ and $\beta$ are found to contain from one to three nucleotides in different sequences (125). Yeast tRNA\textsuperscript{Phe} has two D residues, 16 and 17, in region $\alpha$ and they are accommodated into the structure by an arching out or bulging of the polynucleotide chain in this region. It may be assumed that this bulge is even greater when three nucleotides are present and virtually disappears when only one nucleotide is present. An analogous situation is likely to be found in the $\beta$-region. These two regions of variability are

![Diagram](https://example.com/diagram.png)

*Figure 11* A diagramatic view of the anticodon as seen from the bottom of the molecule in Figures 5 and 8, and perpendicular to the base A35. Four levels of bases are shown with Gm34 on the top, A35 below it, followed by A36. Finally, the dashed Y37 is furthest away. The three anticodon bases are in the form of a right-handed helix.
very close to each other and surround the constant residues G_{18} and G_{19}, which are used in the interaction between the D loop and T\psiC loops.

In surveying sequences, it is found that the fourth base pair in the D stem is complementary in some cases and not complementary in others. The prominent stabilizing feature in yeast tRNA^{Phe} is the stacking interactions which stabilize the two limbs of the L-shaped molecule; it is likely that this is preserved in other structures. Thus, even when there is not a Watson-Crick complementary pair, it is likely that the noncomplementary nucleotides in positions 13 and 22 could pair with each other, possibly in a \textit{trans} orientation (125).

A number of variations of the hydrogen bonding found in yeast tRNA^{Phe} are likely to be found in other sequences (125, 235). For example, most tRNAs with five nucleotides in the extra loop contain m\textsuperscript{7}G\textsubscript{46}; however, some of them contain adenine in that position, and it is likely that these will involve a slightly modified form of hydrogen bonding from that shown in Figure 6 (125, 235). In a similar way, it is likely that there will be other changes in hydrogen bonding found in the major groove of the D stem. In many cases a guanine residue is found in position 9, and it is likely that this interacts with residues 12 and 23 somewhat differently from that found for A\textsubscript{9} in yeast tRNA^{Phe}. Levitt pointed out (132) that the purine in position 15 usually has a complementary pyrimidine in position 48. These could both form a \textit{trans} pairing similar to the G\textsubscript{15}C\textsubscript{48} pair in Figure 6, except that the A\textendashU pair would have a slight lateral displacement (125, 235). In general, the changes in the sequences can be accommodated in the yeast tRNA^{Phe} structure, except that there may be a slight displacement of about 2 \AA{} in the relative position of the bases, with altered hydrogen bonding.

About 80\% of the tRNAs that have been sequenced contain either four or five nucleotides in the extra loop. It is likely that the tRNAs with four nucleotides are formed by the simple omission of a nucleotide analogous to U\textsubscript{47} in yeast tRNA^{Phe} (125). That nucleotide forms a bulge in the extra chain loop, and the uracil residue itself is not involved in stacking interactions. Thus, the structure of yeast tRNA^{Phe} is a very good pattern for the structure of tRNAs that contain either four or five nucleotides in the extra loop and that also contain three or four base pairs in the D stem. Only small adjustments in the structure are necessary to accommodate these sequences.

However, the structure of those tRNAs with very large extra loops containing 13–21 nucleotides is unknown at the present time. A number of observations have been made regarding the sequences in the series (125). For example, position 26 is always a purine whereas position 44 is always a pyrimidine, and they are always anti-complementary, that is either G\textendashU or A\textendashC pairs. These are structurally very similar if there is \textit{trans} pairing between them. However, more structural analysis will have to be carried out before we know the structure of tRNAs with large extra loops.

Most of the comments made up to this point refer to the structure of tRNAs involved in polypeptide chain elongation. There are separate considerations associated with the tRNA\textsuperscript{Met}_{f} that is involved in the initiation of the polypeptide chain. As discussed above, the sequence of the prokaryotic initiator is quite similar to that
of chain elongation tRNAs, but there are substantial sequence differences in the eukaryotic initiators. The greatest difference is found in the TψC loop in which a sequence AU is found instead of Tψ. It is likely that this could still form a TψC loop, but one modified with an A•A pairing rather than a T•A pairing of the type found in the interaction between residues T54 and A58 (5). It will be necessary to carry out structural studies on eukaryotic initiators in order to confirm these structural differences. Significant progress has been made by Sigler and his colleagues (220, 240) on the yeast initiator tRNA, but it is possible that higher-resolution diffraction data will be necessary before a definitive statement can be made about its three-dimensional structure.

**General Observations Regarding tRNA Structure**

With our present knowledge of the detailed structure of yeast tRNA\textsuperscript{Phe} and the inferences that can be made from sequence data concerning other species of tRNA, a number of general observations can be made regarding the structure of tRNA.

1. Extensive stacking interactions of the molecule seem to dominate the structure. These make the largest contribution to the stability of the molecule.

2. The stem regions appear to be largely normal double helical RNA in the A conformation with only a slight perturbation due to G•U pairs.

3. In the loop regions a variety of tertiary hydrogen-bonding interactions are found. Some of these involve hydrogen bonds between bases and others are between bases and backbone. Many different types of base-base interactions are found including only one example of Watson-Crick hydrogen bonding. The result of this bonding leads to the comprehensive stacking interactions described above.

4. Judging from other sequences, it seems as if the molecule can accommodate a number of structural changes and still function as a tRNA. These changes may involve alterations of hydrogen bonding between bases that involve a dislocation relative to the yeast tRNA\textsuperscript{Phe} structure of 2–3 Å in the plane of the base, or variations in the type of hydrogen bonding. In addition, the molecule seems to be able to accommodate distortions such as those induced by the photo cross-linking (207) between the residues 8 and 13, which are still consistent with biological activity.

5. The structure is able to accommodate variations in numbers of nucleotides both in the variable sections of the D loop (α and β) and in the extra loop. The biological role of these variations is not known.

6. Sharp bends in the polynucleotide chain appear to be stabilized by interactions between the bases and the backbone as seen in the similar role played by ψ\textsubscript{55} and U\textsubscript{33} in the sharp turning of the chain in the TψC and anticodon loops.

7. The molecule has a functional design. Conformation of the anticodon loop seems well designed to leave the anticodon bases available for hydrogen bonding with messenger RNA. The CCA-end of the molecule remains single-stranded in a potentially flexible conformation, which may be used in peptide bond formation. Finally, the interactions between the D loop and the TψC loop are not very extensive, thereby reinforcing the idea that they may undergo a conformational opening during protein synthesis (241, 242).
Future Work on Yeast tRNA<sub>Phe</sub>

Data on yeast tRNA<sub>Phe</sub> are now collected to 2.5-Å resolution (137, 231). It is likely that some additional X-ray diffraction data can be obtained, but the amount is not very great. The diffraction patterns for both orthorhombic and monoclinic lattices go out to nearly 2 Å, but the intensity of the data in this region is very low. However, further refinements in the calculations will be carried out, which will probably result in slight movements within the molecule.

The refinement analysis of the orthorhombic data is being carried out in two different laboratories, yielding two different but closely related sets of coordinates (2, 3). At present these have a mean deviation from each other of 0.96 Å; however, the analogous differences between the coordinates from the orthorhombic (2) and monoclinic (4) structures differ by 0.99 Å. The molecule thus seems largely indistinguishable in these two lattices at the present stage of refinement analysis.

Refinement calculations are carried out by a number of methods and are designed to produce a better fit between the observed intensity data and the calculated coordinates. The agreement between these is measured by an error residual factor R, which decreases as the agreement increases. At present, the residual factor R has a value of 0.33 (2) and 0.39 (3) for the two independent analyses of the orthorhombic data, and a value of 0.39 (4) for the monoclinic data. All of these numbers will decrease with further calculations, and it is probable that the coordinates will converge on a common solution, perhaps with the exception of some small differences that may persist between the orthorhombic and monoclinic lattices, especially in the conformation of C<sub>75</sub> and A<sub>76</sub> at the 3'-end of the molecule. At present there are some differences in the assignment of 2'-endo conformation in different residues (2–4), but they are expected to disappear on further refinement. It is also anticipated that refinement will result in an increase in the number of tertiary hydrogen-bonding contacts, especially those involved in hydrogen bonding of the backbone. At present, approximately 35 tertiary hydrogen bonds are described in the structure; it is likely that in the completed structural refinement over 50 hydrogen-bonding interactions will be seen. These interactions together with the interactions of various ions in the structure should lead to a better understanding of the overall molecular stability.

SOLUTION STUDIES OF tRNA

One of the central problems associated with the determination of a biological structure by X-ray crystallography is whether the structure in the crystal is the same as the structure in solution where it is biologically active. Macromolecular crystals are generally heavily hydrated; in the case of yeast tRNA<sub>Phe</sub> the hydration is 71% in the orthorhombic and 63% in the monoclinic lattice (198), and the conditions of crystallization are not very drastic. In the case of crystalline enzymes, the question of biological activity can often be resolved by measuring enzymatic activity in the crystal. However, the biological functions of tRNA are more complex,
so other more indirect methods must be adopted to answer the important question of crystal versus solution conformation.

Many studies have been carried out, and results from a wide variety of probes of tRNA conformation in solution are consistent with the three-dimensional structure for yeast tRNA\textsuperscript{Phe} determined in the crystal. These studies include the use of chemical modifications (243), the study of susceptibility of tRNAs to nucleases (244), and the exchange of tritiated water (\textsuperscript{3}H) with the hydrogen atoms of purine C\textsubscript{8}-H of tRNAs (245) to probe exposed or hindered regions. Other solution studies include the binding of oligonucleotides to tRNAs as a probe for single-stranded regions (246) and NMR spectroscopy of tRNAs in water to study hydrogen-bonded protons. Several fluorescent labels have been introduced into yeast tRNA\textsuperscript{Phe}. Energy transfer has been used to measure the distance between the labels, and the results are in good agreement with the X-ray structure (247). A number of low angle X-ray studies have been carried out on tRNA solutions (205). A study of yeast tRNA\textsuperscript{Phe} by Kratky and co-workers in 1970 (251) led them to a molecular envelope that looks very similar to the form of the molecule seen in the crystal lattice.

Laser Raman spectroscopic measurements are a sensitive index of vibrational modes of the molecule and these have been applied to yeast tRNA\textsuperscript{Phe}. Denaturing tRNA by heat or by lowering the ionic strength changes the spectrum (248). Spectra have been taken of tRNA\textsuperscript{Phe} directly in the orthorhombic crystals and in a solution in which the molecule is biologically active. The two spectra are identical (248).

**Chemical Modification Studies on Yeast tRNA\textsuperscript{Phe}**

Requirements for the use of chemical modification as a probe of tRNA structure have been discussed by Cramer (205) and by Brown (6). The most important requirement is that conditions used for the reaction permit the three-dimensional structure of tRNA to be preserved throughout the reaction. The underlying principle in the use of chemical modification studies is that the exposed bases in the tRNA will be reactive, whereas the buried ones will react at a much slower rate, if at all. It should be noted that "exposed" or "buried" refer to those atoms of the base specifically involved in the chemical reaction. Thus, depending upon the reagent used, the mechanism of the reaction, and the site of reaction, a particular base may be reactive toward one reagent and not toward the other.

Except for the cross-linking of s\textsubscript{4}U\textsubscript{5} to C\textsubscript{13} described above (236), the information gained from chemical modification studies cannot in most cases be used as a direct proof of any structure. With the postulation of a three-dimensional structure it is, however, possible to now ask (a) whether the evidence from the use of a wide variety of chemical reagents is in agreement with the proposed structure and (b) whether the same pattern of reactivity by chemical modification is also observed in other tRNAs and, therefore, whether the same basic structure is also likely to hold for other tRNAs as well.

The results obtained with at least six different reagents on yeast tRNA\textsuperscript{Phe} in solution are entirely in agreement with that expected on the basis of the three-dimensional structure observed in the crystal (129, 243, 249). Reagents included
those that could potentially react with all of the bases: (a) kethoxal, which reacts specifically with G residues at N1 and the N2 amino groups (250), (b) perphthallic acid, which reacts with A at the N1 position (251), (c) methoxyamine, which reacts with C residues (252), (d) carbodiimide, which reacts with G and U residues in the N1 and N3 positions, respectively (253, 254), (e) sodium borohydride, which reacts with D, Y, m7G, and m1A residues (255, 256), and (f) I2/TCI3, which iodinates C residues to yield 5-iodoC (257, 258). The basic result of all these studies for yeast tRNA\textsuperscript{phe} is that those nucleotides exposed in the structure, for instance D\textsubscript{16}, D\textsubscript{17}, G\textsubscript{20}, G\textsubscript{34}, A\textsubscript{35}, A\textsubscript{36}, U\textsubscript{47}, and the 3'-terminal CCA, are all reactive. A\textsubscript{38} is quite reactive toward perphthallic acid (251), although it is stacked and N6 is hydrogen-bonded to C\textsubscript{32} (40), probably because N1 of A\textsubscript{38}, which is the site of reaction, appears quite exposed; furthermore, this reaction probably has "in-plane" stereochemistry (6). C\textsubscript{32} reacts only slightly and U\textsubscript{33} reacts partly. This is probably correlated with the single hydrogen bonds C\textsubscript{32}(O2) to A\textsubscript{38}(N6); U\textsubscript{33}(N3) to P\textsubscript{36}(O1) (2). None of the nucleosides in the T\psiC loop react with any of the above reagents to any significant extent, a result in accord with the tight structure of this loop. The reaction of tRNA\textsuperscript{Phe} with sodium borohydride results in modification of D\textsubscript{16}, D\textsubscript{17}, G\textsubscript{20}, G\textsubscript{34}, A\textsubscript{35}, A\textsubscript{36}, U\textsubscript{47}, and the 3'-terminal CCA, are all reactive. A\textsubscript{38} is quite reactive toward perphthallic acid (251), although it is stacked and N6 is hydrogen-bonded to C\textsubscript{32} (40), probably because N1 of A\textsubscript{38}, which is the site of reaction, appears quite exposed; furthermore, this reaction probably has "in-plane" stereochemistry (6). C\textsubscript{32} reacts only slightly and U\textsubscript{33} reacts partly. This is probably correlated with the single hydrogen bonds C\textsubscript{32}(O2) to A\textsubscript{38}(N6); U\textsubscript{33}(N3) to P\textsubscript{36}(O1) (2). None of the nucleosides in the T\psiC loop react with any of the above reagents to any significant extent, a result in accord with the tight structure of this loop. The reaction of tRNA\textsuperscript{Phe} with sodium borohydride results in modification of D\textsubscript{16}, D\textsubscript{17}, and Y, but no reaction with either m7G\textsubscript{46} or m1A\textsubscript{58} (255, 256). The latter residues are, however, fully reactive when the 3'-half of the tRNA\textsuperscript{Phe} obtained by cleavages in the anticodon is treated with sodium borohydride. Addition of the 5'-half to the 3'-half was found to decrease the reaction of m1A\textsubscript{58} to approximately 40% of theoretical (256). These results suggest that the lack of reactivity of m1A and also probably m7G is not a consequence of the conformation of the T\psiC and the variable loops alone, but rather is a consequence of the constraints imposed by the three-dimensional structure of tRNA\textsuperscript{Phe}. It may also suggest that the intra-loop interaction between T\textsubscript{54} and m1A\textsubscript{58} exists only on the intact RNA when the T\psiC and D loops are interacting. Of the six G and one Gm residues present in the loops of the cloverleaf structure of the tRNA, only G\textsubscript{20} and G\textsubscript{34} react with kethoxal or with carbodiimide, a result in perfect agreement with that predicted by the structure (249, 250, 259). The only base in the variable loop that is exposed in the three-dimensional structure is U\textsubscript{47}, and this is totally reactive (249).

The m7G\textsubscript{46} residue is positively charged and is near the negatively charged phosphate 9, which suggests that yeast tRNA\textsuperscript{Phe} may be stabilized by an electrostatic interaction in that region (125). The reduction of m7G\textsubscript{46} by borohydride was studied by Wintermeyer & Zachau as a function of ionic strength (260). At low ionic strength, the reduction was rather slow, while at higher ionic strength, the reduction was greatly accelerated. This result is direct evidence for an electrostatic interaction of m7G\textsubscript{46}.

**Chemical Modification Studies on the Other tRNAs**

The results of most of the chemical modification work on other tRNAs are basically similar to that of yeast tRNA\textsuperscript{Phe}. The tRNAs studied include those with a large variable loop (261, 262, 264) and those with four or five nucleotides in the variable loop. In addition, both prokaryotic and eukaryotic initiation tRNAs have been studied (263, 261). In general, the \(\alpha\)- and \(\beta\)-regions of the D loop are reactive, and
the anticodon loop is reactive with a decreasing gradient of reactivity toward the 5'-side of the anticodon. In the tRNAs that contain a long variable arm, the stem regions are unreactive but the looped residues of the variable arm are reactive. In these tRNAs as well, the residues equivalent to G15 and C48 of yeast tRNA_Phe are unreactive. These tRNAs are also likely to contain the same trans-pairing interaction as between the G15 and C48 of yeast tRNA_Phe. However, in a mutant of *E. coli* tRNA_{Tyr} in which G15 has mutated to A15, the C48 is now susceptible to chemical modification. Generally speaking, none of the residues of the TψC loop react to any significant extent, whereas the 3'-terminal CCA is fully reactive. In all cases studied, the nucleotide preceding the CCA sequence has been found to be unreactive, and this is presumably because it is stacked on top of the acceptor arm helix. In a few instances, partial reactivity of bases that normally do not react in yeast tRNA_Phe have been reported (265, 266). It is not clear whether this is due to some unfolding of the tRNA during the chemical modification, to the presence of some denatured tRNAs in the sample, or to subtle differences in the three-dimensional structures of some of these tRNAs.

In several tRNAs the fourth base pair in the D stem is not complementary, especially in those tRNAs with large variable loops. These noncomplementary fourth base pairs are not reactive chemically, in keeping with the suggestion (125) that these bases have an alternative hydrogen bonding and preserve the stacking in this part of the molecule.

An important class of chemical modification are the photo-induced cross-links between s^4Us and C13, which were observed in *E. coli* tRNA by Favre and colleagues (207). As explained above (see Figure 10), this is directly understandable from the three-dimensional structure of yeast tRNA\_Phe. The cross-linking occurs only in the native tRNA. A further study has been carried out on ten purified tRNA species from *E. coli* (267). They all form photo-cross-links, but the rates vary, which suggests that the superposition of these residues may differ slightly due to differences in tRNA sequence among these species.

Chemical modification studies on *E. coli* and mouse myeloma initiator tRNAs also show basically the same pattern of reactivity as in yeast tRNA_Phe (254, 263). In contrast to most tRNAs, the *E. coli* tRNA\_Met contains a 5'-terminal nucleotide that is not base-paired to the fifth nucleotide from the 3'-end. This 5'-terminal C is, therefore, reactive with methoxyamine and is deaminated by bisulfite. Although the myeloma initiator tRNA similar to all other eukaryotic initiator tRNAs has a unique sequence in the TψC loop, the residue equivalent to C56 of yeast tRNA_Phe is also unaffected by methoxyamine (263). Other studies show that the AUGC sequence in the myeloma initiator tRNA, which replaces the TψCG(A) sequence present in chain elongation tRNAs, is also buried in this tRNA. Thus, it is likely that the basic folding pattern of initiator tRNAs may also be quite similar to that of yeast tRNA_Phe.

Litt & Greenspan (266) have studied the modification of *E. coli* tRNA_{Val} and tRNA\_Phe with kethoxal. *E. coli* tRNA\_Phe lacks the residue equivalent to G20 of yeast tRNA_Phe, and none of the G residues of the *E. coli* tRNA_Phe, including those equivalent to the G18 and G19 of yeast tRNA_Phe, react with kethoxal except for the one G in the anticodon. With *E. coli* tRNA_{Val}, which does not contain a G in the
anticodon, the major site of reaction is G$_{20}$ in the D loop as in yeast tRNA$^{\text{Phe}}$. Litt, Chang, and their coworkers (262, 266) have also compared the kinetics and sites of kethoxal modifications of native and denatured tRNAs. It was found that in the denatured form of \textit{E. coli} tRNA$^{\text{Trp}}$, G residues normally not reactive in the native form are now reactive. Similar studies on the denatured form of yeast tRNA$^{\text{Leu}}_3$ show that while the reaction of kethoxal with native tRNA$^{\text{Leu}}_3$ occurs slowly and reaches a plateau at the level of two G residues modified per tRNA, in the denatured form as many as 13 G residues react (265). Since the two G residues that react in the native tRNA$^{\text{Leu}}_3$ are analogous to G$_{18}$ and G$_{19}$ of yeast tRNA$^{\text{Phe}}$, which do not normally react in the tRNA$^{\text{Phe}}$, it is possible that tRNA$^{\text{Leu}}_3$ is susceptible to partial denaturation and part of this denaturation process may well be the exposure of G$_{18}$ and G$_{19}$.

In those instances where the reactions of a tRNA with methoxyamine and bisulfite, two cytidine-specific reagents, have been compared, the results obtained are in almost complete agreement. All of the reactive cytidines are located either in the $\alpha$- or $\beta$-region of the D loop, the anticodon loop, and the 3'-terminal CCA (269, 270). An exception is the reaction of \textit{E. coli} tRNA$^{\text{Glu}}$ with bisulfite (271). In this case, the three cytidine residues present in the D loop are totally unreactive, whereas the residue equivalent to C$_{26}$ of yeast tRNA$^{\text{Phe}}$ is reactive. A possible explanation for this is that under the reaction conditions used, tRNA$^{\text{Glu}}$ exists in a conformation different from the native one. Evidence for such an alternative conformation of \textit{E. coli} tRNA$^{\text{Glu}}$ has been recently provided (67).

Weinstein and co-workers have used N-acetoxy-2-acetyl aminofluorene as a reagent for the specific modification of guanine residues at the C8 position (272, 273). Reaction with \textit{E. coli} tRNA$^{\text{Met}}$ occurs almost exclusively on the guanine residue analogous to G$_{19}$ of yeast tRNA$^{\text{Phe}}$. Although G$_{19}$ forms a Watson-Crick base pair to C$_{56}$ in the structure of yeast tRNA$^{\text{Phe}}$, the C8 position of this guanine residue is nonetheless exposed at the corner of the molecule. Thus, the reactivity of G$_{19}$ toward this reagent is entirely consistent with the X-ray structure. Similar studies have been carried out on yeast tRNA$^{\text{Tyr}}$ (271). Modification of the tRNA was less extensive and more specific when carried out in the presence of 3mM MgCl$_2$, 0.1 M KCl, conditions that are more likely to preserve the three-dimensional structure of the tRNA, rather than 3 mM EDTA. Approximately 1 mole of the reagent reacted per tRNA molecule. Partial modification of the tRNA occurred at three different sites, the G residues equivalent to G$_{18}$ and G$_{19}$ of yeast tRNA$^{\text{Phe}}$ and the G residue in the anticodon.

\textbf{Use of NMR Spectroscopy for the Analysis of tRNA Structure in Solution}

Significant advances have been made recently in the use of NMR spectroscopy as a probe for hydrogen-bonded hydrogen atoms in tRNAs in solution (274). The important finding making this possible was that hydrogen-bonded ring N-H hydrogen atoms, although exchangeable with water, possessed adequately long helix lifetimes, and the resonances from these protons were shifted downfield far enough so they could be separated from the large peak due to water protons (275, 276). On the basis of model studies with synthetic oligonucleotides (277) and several well-characterized fragments of yeast tRNA$^{\text{Phe}}$, Shulman et al and Lightfoot et al were

\textit{Annual Reviews}

\textit{www.annualreviews.org/aronline}

\textit{Annu. Rev. Biochem. 1976.45:805-860. Downloaded from arjournals.annualreviews.org by Columbia University on 01/24/07. For personal use only.
able to generate a set of "ring current shift" rules and to assign the resonances present in the high-resolution NMR spectra of yeast tRNA\textsuperscript{Phe} to specific Watson-Crick base pairs in the cloverleaf model of the tRNA (278, 279). Details of this work and a general discussion of the use of NMR spectroscopy in tRNA structure analysis are reviewed by Kearns & Schulman (274) and Sigler (5).

More recent results suggest that NMR spectroscopy may be also useful in elucidating additional elements of tRNA structure. Reid and co-workers (280, 281) and Daniel & Cohn (282) have obtained high-resolution NMR spectra of several purified tRNAs, including \textit{E. coli} tRNA\textsuperscript{Val}, tRNA\textsuperscript{Arg}, tRNA\textsuperscript{Phe}, tRNA\textsuperscript{Met}, and yeast tRNA\textsuperscript{Asp}. The first three \textit{E. coli} tRNAs, which contain one G\textsuperscript{U} base pair each in the cloverleaf stem structure, have in their NMR spectra a predominant peak in the \(-10\) to \(-11\) ppm region corresponding to an intensity of about two protons. This signal is amplified about four to five times in yeast tRNA\textsuperscript{Asp}, which contains three "G\textsuperscript{U}" and one "G\textsuperscript{\psi}" base pairs. Thus, in contrast to earlier interpretations (274, 279), G\textsuperscript{U} base pairs can be detected by NMR spectroscopy (281).

In both \textit{E. coli} tRNA\textsuperscript{Val} and yeast tRNA\textsuperscript{Phe}, Reid & Robillard (280) report the presence of resonances due to 26\textsuperscript{\pm} 1 N-H hydrogen bonds. Approximately the same number (27) was also observed in \textit{E. coli} tRNA\textsuperscript{Met} by Daniel & Cohn (282). Since the number of base pairs in the cloverleaf model of these tRNAs is 20–21, it is suggested that these tRNAs contain at least six tertiary base pairs detected by NMR spectroscopy, which is in general agreement with the crystallographic three-dimensional structure.

If tertiary-structure base pairs are common to all or most tRNAs, resonances from such interactions should be discernible from the broad background of overlapped and unresolved resonances due to the secondary-structure Watson-Crick type of base pairs. With this rationale, Wong & Kearns (283) have examined the NMR spectrum of unfractionated \textit{E. coli} tRNA and have assigned the anomalously low-field resonance at \(-14.8\) ppm (280, 282) to a s\textsuperscript{4}U-A tertiary interaction. This resonance is seen only in unfractionated \textit{E. coli} tRNA, which has s\textsuperscript{4}U, and not in yeast tRNA, which does not. Furthermore, chemical conversion of s\textsuperscript{4}U to U shifts the peak at \(-14.8\) ppm to \(-14.3\) ppm (283). Removal of the N-H proton in s\textsuperscript{4}U by photo-cross-linking to C\textsubscript{13} (207) or by conversion to uridine-4-thiocyanate (284) results in the disappearance of the peak at \(-14.8\) ppm (285). These findings suggest that the NH proton of s\textsuperscript{4}U is involved in hydrogen bond formation and, based on the location of the peak, with an adenine residue, most probably A\textsubscript{14}. More recently, Reid and co-workers (280, 281) and Wong et al (285) have examined the NMR spectra of several purified \textit{E. coli} tRNAs with essentially the same result. Daniel & Cohn (282) have studied the NMR spectrum of \textit{E. coli} tRNA\textsuperscript{Met}, which had been spin-labeled in the s\textsuperscript{4}U residue. The spin-labeling modification results in a s\textsuperscript{4}U ring structure that is now incapable of donating a hydrogen for any tertiary H-bond interaction, and the low-field resonance at \(-14.8\) ppm is now missing. Thus, since crude tRNA, purified tRNAs, and the \textit{E. coli} initiator tRNA all show this, the s\textsuperscript{4}U\textsubscript{8}-A\textsubscript{14} interaction is most probably a feature common to all \textit{E. coli} tRNAs.

Some evidence that tRNAs contain common tertiary interactions besides the s\textsuperscript{4}U\textsubscript{8}-A\textsubscript{14} pair discussed above may be obtained from the work of Bolton & Kearns
STRUCTURE OF TRANSFER RNA

(286), who have investigated the effect of salt and temperature on the NMR spectra of unfractionated \textit{E. coli} and yeast tRNAs. It was found that, in the temperature range of 40–50°C, the presence or absence of Mg\textsuperscript{2+} had a significant effect on the NMR spectrum. The different NMR spectra at 50°C with and without Mg\textsuperscript{2+} showed three pronounced peaks at -13.8, -13.0, and -11.8 ppm. It was concluded that these spectral changes arise from tertiary structural base pairs, which are common to most or all tRNAs, and that these are the weakest base pairs in the structure (286). By means of the ring-current-shift calculations (274), two of these peaks (-13.8 and 13.0 ppm) have been tentatively assigned to the T\textsubscript{54}A\textsubscript{58} and G\textsubscript{19}C\textsubscript{56} base pairs known to be present in the crystallographic three-dimensional structure of yeast tRNA\textsubscript{Phe}. Similar resonances are also present in the NMR spectra of purified \textit{E. coli} tRNA\textsubscript{Val} and have been assigned to the same two tertiary base pairs by Reid & Robillard (280).

\textit{E. coli} tRNA\textsubscript{Glu} contains four A•U base pairs. Two of these are at the ends of the acceptor arm helix (UA\textsubscript{7}) and the TΨC helix (AU\textsubscript{49}). If the two helices were to stack, forming a continuous linear helix, both of these base pairs would contain an A residue stacked on the 3'-side of the base pair and both of these resonances would be subjected to strong upfield shifts (274, 278). A strong upfield shift would be -13.5, while no upfield shift would predict a resonance at -14.2. The resonances from both UA\textsubscript{7} and AU\textsubscript{49} occur within 0.1 ppm of -13.5 (5), thereby suggesting that the helical arm of the acceptor stem and the TΨC stem are arranged in a continuous helix.

Thus the NMR studies present a variety of evidence to support the conclusion that the structure of the tRNA molecule in solution is the same as that in the crystal. Furthermore, the evidence points to many similarities of structure between different tRNA species.

Susceptibility of tRNAs toward Nucleases

The first indication that certain sites in tRNA are more susceptible to nucleases than others came during the sequence analysis of yeast tRNA\textsubscript{Ala}. Using conditions of limited digestion with T\textsubscript{1}-RNase, Penswick & Holley (287) were able to cleave tRNA\textsubscript{Ala} specifically in the anticodon to yield two half fragments. Since then similar studies using limited digestion with both T\textsubscript{1}-RNase and pancreatic RNase have been carried out on yeast tRNA\textsubscript{Ser} (288), tRNA\textsubscript{Tyr} (289), tRNA\textsubscript{Phe} (290), tRNA\textsubscript{Val} (291, 128), \textit{E. coli} tRNA\textsubscript{Tyr} (292), tRNA\textsubscript{Met} (293), and several other tRNAs. The general finding that has emerged from all of these studies can be summarized as follows: (a) The stem regions of the tRNA are relatively resistant toward nucleases. (b) The anticodon loop and the 3'-terminal CCA-end are most susceptible to nucleases, whereas the TΨC loop and the variable loop are by and large resistant. (c) The D loop is also relatively resistant except for the variable α- and β-regions within this loop, which are readily cleaved by nucleases. In general, the above results are in good agreement with those obtained from studying the sites on tRNA most reactive toward chemical modifications.

Schmidt et al (294) have examined the action of T\textsubscript{1}-RNase on yeast tRNA\textsubscript{Phe}. The anticodon loop of this tRNA contains a modified G residue (Gm) that is not susceptible to hydrolysis by T\textsubscript{1}-RNase. Consequently, under the conditions of par-
tial digestion used, initial cleavage of the tRNA occurred exclusively in the D loop to yield the 5'-terminal one quarter fragment pG1-G_{18} and the 3'-terminal three quarter fragment A_{21}-A_{76}. Upon prolonged incubation, the latter fragment was subsequently cleaved in the Ti\&C loop to yield a 37-nucleotide fragment A_{21}-G_{57} and the 3'-terminal quarter fragment m\textsuperscript{1}A_{38}-A_{76}. Even with a large excess of enzyme and at 37°C, cleavage in the presence of Mg\textsuperscript{2+} occurred preferentially in the D loop. Since G_{19} and G_{20} were not present in any of the fragments, it is not known with certainty whether the initial site of attack by T\textsubscript{1}-RNase was on G_{18}, G_{19}, or G_{20}. However, since in most tRNAs the common residues G_{18} and G_{19} are relatively resistant to T\textsubscript{1}-RNase (292), and since G_{20} is also the most readily available G residue in the D loop to chemical modifications, it is most likely that G_{20} was the primary cleavage site on yeast tRNA\textsubscript{phe} toward T\textsubscript{1}-RNase. No cleavage on G_{45} of the variable loop was observed, suggesting that G_{45} is protected in the three-dimensional structure of the tRNA.

Essentially similar results were obtained by Samuelson & Keller (295), who observed quantitative cleavage of yeast tRNA\textsubscript{phe} in the D loop and at a much slower rate in the Ti\&C loop. Under conditions when complete cleavage occurred at G_{18} and G_{20}, no cleavage at all was found in the G_{15} of the same loop. The above results taken together with those of Schmidt et al (294) suggest that, of the G residues present in the cloverleaf structure of yeast tRNA\textsubscript{phe}, G_{15}, G_{18}, G_{19}, G_{45}, and G_{57} are most probably shielded against nucleolytic attack in the tertiary structure of this tRNA.

An extensive survey of sites on yeast tRNA\textsubscript{phe} and tRNA\textsubscript{ser} susceptible toward several other nucleases has also been carried out by Harbers et al (244). The results with T\textsubscript{1}-RNase and Neurospora endonuclease are basically similar to those described above. With pancreatic RNase, the two most susceptible sites were the anticodon loop and the CCA-end in agreement with earlier findings (290). With T\textsubscript{2}-RNase, splitting occurred in the D loop between D_{16} and D_{17}, and in the anticodon loop and the CCA-end.

Streeck & Zachau (297, 298) have also compared the patterns of degradation obtained from native yeast tRNA\textsubscript{phe} and tRNA\textsubscript{ser} with those obtained from corresponding denatured forms of these tRNAs. Detailed studies using T\textsubscript{1}-RNase, pancreatic RNase, T\textsubscript{2}-RNase, sheep kidney nuclease, and hog spleen acid RNase revealed characteristically different partial fragmentation patterns for the native and denatured forms. These findings support the assumption implicit in these studies that the relative resistance of most of the nucleotides in the tRNAs toward partial digestion with nucleases is due to the shielding of these nucleotides in the tertiary structure of tRNAs.

**Oligonucleotide Binding Experiments**

A simple and direct method for probing polynucleotide structures is the introduction of oligonucleotides, usually trimers or larger, to see if they will bind to the polynucleotide. It became apparent from the initial studies of Uhlenbeck (310) and Högenauer (300) that this would be a useful probe for tRNA structure in solution. The binding is generally measured by equilibrium dialysis, and the assump-
tion is usually made that Watson-Crick pairing will account for the observed results. However, it is important to realize that other effects can be important, such as the influence of nucleotides on the base stacking of neighboring nucleotides in an oligomer, as this can influence binding constants. Furthermore, other types of hydrogen bonding may occur in addition to Watson-Crick pairing.

Some generalizations are possible from the results of studies on several different tRNAs, including yeast tRNA\textsuperscript{Phe}. Almost all tRNAs can bind oligomers complementary to the last three or four bases at the 3’-end of the molecule. Furthermore, they all bind oligomers complementary to the anticodon. No binding is found for the stem regions of the cloverleaf, and virtually none for the T\textsubscript{Ψ}C loops, even for eukaryotic initiator tRNAs with an altered nucleotide sequence in the loop (301). Some binding, usually weaker, is found for the D loop and variable loop in some cases.

Experiments with yeast tRNA\textsuperscript{Phe} have been carried out by many investigators (302–308). Pongs and co-workers (302, 303, 308) showed that two regions were fully accessible to binding, the 3’-terminal ACCA for binding by complementary tri- or tetra-nucleotides and the anticodon region. The binding of the anticodon (UUC) was considerably augmented by the addition of a fourth purine on the 3’-side, either A or G, but less by the addition of U (308). Strong binding of larger oligomers in this region has also been seen by other investigators (305, 307). Eisinger & Spahr (305) reported that the pentamer UUCAG is bound to yeast tRNA\textsuperscript{Phe}. These results raise the question of whether or not residues 32 and 33 in the yeast tRNA\textsuperscript{Phe} sequence are available for binding in solution. As pointed out above, they are unavailable in the crystal lattice. However, the conformation of the anticodon could change on binding, giving rise to an altered conformation. Another explanation has been offered in recent experiments comparing the binding of UUCA, UUCG, and UUC-purine (R. Bald and O. Pongs, unpublished data). All of these have high binding constants to yeast tRNA\textsuperscript{Phe}, even though the purine residue cannot form two hydrogen bonds to U\textsubscript{33}. It suggests that the 3’-purine, as with terminal A or G residues, acts more to stabilize the stacking of the oligonucleotide to produce a higher binding constant than to engage in hydrogen bonding.

In contrast to the strong binding to the 3’-end and at the anticodon, binding to the D and T\textsubscript{Ψ}C loops and the variable loop of yeast tRNA\textsuperscript{Phe} is weaker (302, 308). Using slightly different conditions, Cameron & Uhlenbeck (306) found somewhat higher binding to the D loop. How do we interpret binding to regions that may be inaccessible in the crystal structure? Two explanations come to mind. The tRNA solution may contain some denatured species, for it has been shown that the pattern of oligonucleotide binding changes radically in the stable denatured form of yeast tRNA\textsuperscript{Leu}\textsubscript{3} (309). Alternatively, the crystal structure is a static view of what is obviously a dynamic structure, and the molecule might open up, especially with competition of an oligonucleotide that can bind. An explanation of this type has been suggested by Uhlenbeck (310).

The results from other tRNAs are broadly similar. \textit{E. coli} tRNA\textsuperscript{Tyr} (310), yeast tRNA\textsuperscript{Leu}\textsubscript{3} (309), \textit{E. coli} (299, 300, 310, 311) and yeast tRNA\textsuperscript{Met} (301), and yeast tRNA\textsuperscript{His} (312) have the anticodon as well as the 3’-end of the molecule available for
binding. Some binding has been reported in the D loop and variable loop of *E. coli* tRNA^Tyr^ (310), yeast tRNA^Leu^ (309), and *E. coli* tRNA^Met^ (310), but this was thought to be competing with the native structure. Little or no binding of oligomers to the D loop was found for yeast tRNA^Ile^ (312) or yeast tRNA^Met^ (301). As in other tRNAs, the wobble codon (232) is also efficiently bound to the anticodon of the initiator tRNA (301). It is interesting that the codon AUG binds to both *E. coli* tRNA^Met^ and tRNA^Met^, but three times more efficiently to the former species, which may be associated with base modifications (311).

The pattern of oligonucleotide binding is changed somewhat when the Y base is removed from yeast tRNA^Phe^. There is a decreased binding to the anticodon (303, 306), which might be expected since the Y base is adjacent and its removal would allow increased flexibility in the anticodon. Somewhat less obvious is the result of a substantial change in the binding of oligonucleotides to the D loop (306) induced by removal of the Y base some 40–50 Å away.

In general, the three-dimensional structure of yeast tRNA^Phe^, which is obtained from the crystal data, is broadly compatible with the results of oligonucleotide binding studies, although further work will be necessary to interpret some of these findings.

On reviewing the results of solution studies on tRNA with the crystal structure analysis, we find striking agreement on the whole. The structural model can be used to interpret a wide variety of investigations that probe many different aspects of the molecule. Future work will be even more sharply focused on the question of comparing fine details of structure obtained in the crystallographic studies with those derived from solution studies. Indeed, the availability of the three-dimensional structure should make possible even more rigorous tests in the future. The molecule is dynamic, even though the crystallographic structure results are static. In the future, solution studies should provide important access to the detailed nature of the tRNA molecular movements.

### tRNA CONFORMATIONAL CHANGES AND BIOLOGICAL FUNCTION

With the structure of yeast tRNA^Phe^ known in one conformation, a crucial question is whether the molecule changes its conformation during biological function. While the evidence at present is not conclusive, it suggests strongly that changes do occur. The type of conformation changes that the molecule can undergo has certain constraints. The photo-cross-linking experiments tying s^4^U_8^ with C$_{13}$ (207) yield a molecule capable of aminoacylation and protein synthesis, which suggests that this part of the molecule is not likely to unfold during these processes. Furthermore, it is possible to attach an affinity label onto the sulfur atom of s^4^U_8^ in *E. coli* tRNA^Phe^, and the molecule functions normally (313). This label fills the groove between the D stem and the TuC stem, so that space must not be intruded into during aminoacylation or ribosomal passage.

Does the tRNA molecule change conformation after interaction with the aminoa- cyl synthetases? An NMR study of yeast tRNA^Phe^ (314) showed no change upon aminoacylation. However, these experiments are difficult because of the lability of
the aminoacyl ester bond at the 3'-terminal ribose. Experiments by Kan et al. (315) have utilized tRNA that has an NH$_2$ group in place of the 3'-OH and forms a stable amide linkage between the amino acid and the tRNA (316). An NMR study of amide-linked phenylalanyl-tRNA$^{Phe}$ showed two peaks in the low-field spectrum that were altered (315). It is not clear which hydrogen-bonded interactions are involved, but these results suggest a small change in conformation upon aminoacylation. Further work will be needed to resolve this question.

What happens when the tRNA goes into the ribosome? The experiments of Erdmann and his colleagues (241, 242, 319) strongly suggest that the T$_{\Psi}C$ loop disengages from the D loop and opens up so that the sequence T$_{\Psi}C$ is now available for binding, possibly to a complementary sequence GAA on the ribosomal 5S RNA. They have shown that the tetranucleotide T$_{\Psi}CG$ will prevent the binding of aa-tRNA to the A site on the ribosome. Furthermore, T$_{\Psi}CG$ would not prevent the binding of initiator tRNA to the P site (320), reinforcing the idea that the function of this loop in initiator tRNA is somewhat different from that found in chain-elongating tRNAs. These experiments have been extended by Gassen and co-workers (317, 318), who have shown that CGAA is bound, presumably to the T$_{\Psi}CG$ of tRNA$^{Phe}$, only when the tRNA combines with the codon. These experiments were carried out either with ribosomal subunits and an oligouridylicate (U$_{7\ldots8}$) mRNA fragment or with U$_{7\ldots8}$ alone. The suggestion is that the tRNA molecule is capable of undergoing a conformational change in the region of the T$_{\Psi}C$ loop when an interaction occurs at the anticodon loop approximately 60–70 Å away. These results bring to mind the oligonucleotide binding experiments showing that removal of the Y base from the anticodon loop of yeast tRNA$^{ahc}$ led to altered binding of an oligonucleotide complementary to the D loop (306). These both suggest that modification in the anticodon loop can bring about a change in the molecule that can be detected about 50 Å away. A similar phenomenon has been reported by Wells and his colleagues (115, 119), who have measured the stabilization of one region of a DNA helix by an adjacent region that may be 15 base pairs (~50 Å) removed. This phenomenon was termed telestability. What is observed in the tRNA molecule may be a more complex version of the same phenomenon.

This may be an expression of the fact that the tRNA molecule exhibits long-range order, and therefore modifications in one part of the molecule may be expressed by changes in properties at another part of the molecule. As mentioned above, the tRNA molecule is a dynamic system, and these effects may be an expression of this property.

A long-distance interaction, such as the one described linking conformational changes in the anticodon with changes in the D loop, may help to explain the otherwise puzzling observations concerning *E. coli* tRNA$^{Trp}$. Mutation of the G$_{24}$ of the D stem to A$_{24}$ enables this tRNA to alter its anticodon function so that it acts to suppress termination without changing the anticodon sequence (138). This might suggest that a conformational change is induced in the anticodon as a consequence of an alteration in the D stem.

It is quite clear that tRNA can be made to undergo conformational changes as, for example, in thermal denaturation, which has been studied in a variety of investigations (120, 140, 145, 309). A detailed and comprehensive study of the unfolding
of tRNA_Met has been carried out by Crothers et al (145). A conformational change has also been observed at room temperature by measuring the diffusion constant of tRNA as a function of ionic strength (149). In 1 mM Mg^{2+} and at ionic strength 0.1, the diffusion constant of both unfractionated E. coli tRNA and pure yeast tRNA_Phe was found to rise sharply, indicating that the molecule had folded into a more compact form. Raising or lowering the ionic strength resulted in a decreased diffusion constant. This has been interpreted as an environment in which the T\psi C and D stems disengage, and the two limbs of the L-shaped molecule fold together. The stability of tRNA conformation is thus a sensitive function of the salt concentration, and it remains to be seen whether this is related to functional changes.

Many of the future studies of tRNA conformation will undoubtedly be directed toward the goal of understanding what happens when it carries out its biological functions. Movement of tRNA clearly occurs inside the ribosome, and the relation between the tRNA and the ribosomal A site and P site is yet to be understood. A suggestion has been made that movement between these two sites may have a rotatory component, perhaps associated with a turning of the mRNA when it is being read (296). In another model, there is no movement of the tRNA between A and P sites (299). However, we do not have enough data at present to evaluate ideas of this type.

The next phase of work on tRNA is likely to involve a more direct structural approach to the biological function of the molecule. In this work, it is likely that knowledge of the three-dimensional structure will play an essential role.

**BIOLOGICAL MYSTERIES OF TRANSFER RNA**

There are many unknown features associated with the tRNA molecule. Knowing the three-dimensional structure in a sense helps us gain perspective in that we can sometimes separate things that we understand from things we do not understand. The three-dimensional structure of yeast tRNA_Phe yields structural reasons for most of the invariant and semi-invariant bases in this tRNA species and, by inference, in most other tRNA species as well. In a sense this provides us with a very good framework for understanding a large part of the information imbedded in the tRNA sequences. Most of the functional aspects of tRNA are not known at present, but we have some understanding about how to proceed in our research to learn more about them. From the comparative information embodied in the tRNA sequence data, however, we learn of three major mysteries—base modification, variations in the \( \alpha- \) and \( \beta- \)regions of the D loop, and the tRNAs with large variable loops.

What is the biological role of the base modifications, which are varied and great in number? Are they associated with the role of tRNA in protein synthesis? It is likely that some of them are; for example, N2 methylation of a non-cognate tRNA substrate for yeast phenylalanine synthetase at position G_10 results in faster aminoacylation by the enzyme (59). Since the normal yeast tRNA_Phe substrate has a methyl group in that position, it is reasonable to believe that this modification may be involved in specificity of aminoacylation. Another example, but not involving protein synthesis, is found in the repressor activity of Salmonella histidyl-tRNA_His.
In a mutant tRNA where the conversion of a U to $\psi$ in the anticodon stem of tRNA$^{\text{His}}$ is blocked, there is no repression of the transcription of the histidine operon (23, 49). In this case, the base modification is necessary for the tRNA regulatory activity. However, these examples are very rare, since the biological role of the large bulk of modified bases in tRNA is unknown.

A second mystery concerns the D loop and the large number of variations in tRNA sequence in the $\alpha$- and $\beta$-regions. They can have one to three nucleotides; most of them are pyrimidines and 70% of the nucleotides are dihydrouracil (125). In the three-dimensional structure, the bases of the $\alpha$- and $\beta$-regions are near each other on the surface of the molecule where they can interact with other molecules. It has been suggested that they could be a recognition region for synthetases (137), but that is unlikely since some isoacceptor tRNA species have different numbers of nucleotides in the $\alpha$- and $\beta$-regions, even though they are all recognized by the same enzyme. The biological function of the variations in the $\alpha$- and $\beta$-region is not obvious.

The third mystery surrounds the large variable loops (13–21 nucleotides), which account for approximately 20% of the known tRNA sequences. They are confined to three species in prokaryotes—serine, leucine, and tyrosine—but eukaryotic tyrosine tRNA does not have a large variable loop. Is this phenomenon related to the role of tRNA in protein synthesis or are other functions involved?

At present, we do not have answers to these questions, but stating the questions sometimes facilitates and directs our research activities toward their solution.

Acknowledgments

This work was supported in part by grants from the National Institutes of Health, the National Science Foundation, the American Cancer Society, and the National Aeronautics and Space Administration.

Literature Cited


118. Reference deleted in proof


176. Sundaralingam, M. 1972. See Ref. 175, pp. 417–56
STRUCTURE OF TRANSFER RNA
CONTENTS

AN AUTOBIOGRAPHY, Edward A. Doisy 1

THE BIOCHEMISTRY OF MUTAGENESIS, John W. Drake and Richard H. Baltz 11

TRANSLATIONAL CONTROL OF PROTEIN SYNTHESIS, Harvey F. Lodish 39

ANGIOTENSIN-CONVERTING ENZYME AND THE REGULATION OF VASOACTIVE PEPTIDES, Richard L. Soffer 73

THE ROLE OF POLYPRENOI-LINKED SUGARS IN GLYCOPROTEIN SYNTHESIS, Charles J. Waechter and William J. Lennarz 95

BIOCHEMISTRY OF POLYISOPRENOID BIOSYNTHESIS, Enrique D. Beytia and John W. Porter 113

MULTIFUNCTIONAL PROTEINS, Kasper Kirchner and Hans Bisswanger 143

THE CONTROL OF GLYCOGEN METABOLISM IN THE LIVER, H. G. Hers 167

SOLUBLE FACTORS REQUIRED FOR EUKARYOTIC PROTEIN SYNTHESIS, Herbert Weissbach and Severo Ochoa 191

COMPARATIVE ASPECTS OF GLYCOPROTEIN STRUCTURE, Rosalind Kornfeld and Stuart Kornfeld 217

CALCIUM-BINDING PROTEINS, Robert H. Kretsinger 239

PHOTOREGULATION OF BIOLOGICALLY ACTIVE MACROMOLECULES, Bernard F. Erlanger 267

1,4-DIAMINOBUTANE (PUTRESCINE), SPERMIDINE, AND SPERMINE, Celia White Tabor and Herbert Tabor 285

CONCEPTS AND PERSPECTIVES IN ENZYME STEREOCHEMISTRY, Kenneth R. Hanson 307

SOME PERTINENT ASPECTS OF MECHANISM AS DETERMINED WITH SMALL MOLECULES, Thomas C. Bruice 331

THE REOVIRUS REPLICATIVE CYCLE, Samuel C. Silverstein, Judith K. Christman, and George Acs 375

NITROGENASE, Harry C. Winter and R. H. Burris 409

PROTEINS OF CONTRACTILE SYSTEMS, H. G. Mannherz and R. S. Goody 427

THE BIOLOGICAL ORIGIN OF ANTIBODY DIVERSITY, A. R. Williamson 467