Sequential Folding of Transfer RNA

A Nuclear Magnetic Resonance Study of Successively Longer tRNA Fragments with a Common 5' End

JOHN BOYLE

Department of Biochemistry
Mississippi State University
Mississippi State
Miss. 39762, U.S.A.

GEORGE T. ROBILLARD

Department of Physical Chemistry
University of Groningen
Groningen, The Netherlands

AND

SUNG-HOU KIM

Department of Chemistry
University of California
Berkeley, Calif. 94720, U.S.A.

(Received 9 May 1979, and in revised form 2 January 1980)

Most folding studies on proteins and nucleic acids have been addressed to the transition between the folded and unfolded states of an intact molecule, where an entire residue sequence is present during the folding event. However, since these polymers are synthesized sequentially from one terminus to the other in vivo, their folding pathways may be influenced greatly by the sequential appearance of the residues as a function of time.

The three-dimensional structure of yeast tRNA\textsuperscript{Phe} in the crystalline state is correlated with 360 MHz proton nuclear magnetic resonances from three fragments plus an intact molecule of the tRNA that share a common 5' end and are in a solution condition similar to that of the crystal structure. This has allowed identification of folded structures present in the fragments and presumably present in the growing tRNA molecule as it is being synthesized from the 5' end. The experiments show that only the correct stems are formed in the fragments; no additional or competing helical region is produced. This suggests that in the biosynthesis of this tRNA, correct folding of helical stems occurs before the entire molecule is formed. Further, some of the tertiary interactions (hydrogen bonds) found in the crystal structure are also probably present before the synthesis is completed. These findings are generalized to consider the precursor of the tRNA as well as other tRNAs.
1. Introduction

Most biologically functional macromolecules have definite conformations in solution. It is believed that each such conformation represents the global minimum energy state of the molecule, or a local energy minimum surrounded by high energy barriers. When one considers the biosyntheses of proteins and nucleic acids, one recognizes that there are two alternative overall mechanisms of folding: one in which the entire molecule is first synthesized as a "random" coil, then folds into a functional conformation; the other in which the molecule is sequentially folded as it is being synthesized. In both cases, intermediate folded structures may or may not rearrange as the folding progresses (Baldwin, 1975). Most of the studies on folding so far have been addressed to the first mechanism, i.e. the transition between the folded and unfolded states of an intact molecule, where entire sequences of amino acids or nucleotides are present during the folding event. We present here an n.m.r. study addressed to the second mechanism using a set of successively longer fragments of a transfer RNA with a common 5' end.

Transfer RNA is a small, globular nucleic acid that has proved to be amenable to detailed structural studies. The most exact method of structure determination, X-ray crystallography, has been used to provide an atomic picture of the conformation of a particular tRNA, yeast phenylalanine transfer RNA (tRNA^{Phe}) (Sussman et al.,

Fig. 1. Cloverleaf sequence of yeast tRNA^{Phe} and for tRNA fragments: (a) the 5' 1/4 fragment containing nucleotides 1 to 16; (b) the 5' 1/2 fragment containing nucleotides 1 to 36; (c) the 5' 3/5 fragment containing nucleotides 1 to 45; and (d) intact molecule containing nucleotides 1 to 76.

† Abbreviations used: n.m.r., nuclear magnetic resonance; p.p.m., parts per million; DSS, sodium 2,2-dimethyl-2-silopentane-5-sulfonate.
1978; Holbrook et al., 1978; Jack et al., 1976; Quigley et al., 1975; Stout et al., 1978). This provides a static as well as a limited dynamic (Kim, 1978) picture of the molecule in a crystalline state. However, recent work has shown that this tRNA and others in aqueous solution can be studied using high-resolution n.m.r. techniques that can provide rather detailed information on tRNA solution conformation (Robillard et al., 1976a, 1977a; Kan & Ts' o, 1977; Geerdes & Hilbers, 1977; Kan et al., 1977; Kearns, 1976; Crothers et al., 1974).

By using a set of successively longer fragments of tRNA that extend from the 5' end of the molecule as shown in Figure 1, it should be possible to mimic sequential folding, if any occurs during biosynthesis. While it is believed that all tRNAs have about the same overall conformation (Kim et al., 1974) it is essential for us to use the yeast tRNA^{Phe} as the experimental system because it is the only tRNA of which the three-dimensional structure is known in detail. Thus, any pattern of folding seen can be compared directly to the exact end-product of the folding. Comparison can then be made with the sequence of other tRNAs to see if the folding phenomena observed can be generalized.

Previous n.m.r. studies have shown that portions of several tRNAs that are expected to be helical according to the cloverleaf model have helical structures in fragments (Kearns, 1976; Robillard et al., 1976a; Rordoff, 1975; Lightfoot et al., 1973). But no n.m.r. study has looked at successively longer pieces of one tRNA with a common 5' end in conditions close to those used in the X-ray structural determination.

2. Materials and Methods

(a) Chemicals

Yeast phenylalanine tRNA was purchased from Boehringer Mannheim and used as a starting material without further purification. Use of different lots of the nucleic acid gave similar results. Aniline was reagent grade from Fisher Scientific Company. It originally had a brown color and had to be purified to remove oxidation products. It was distilled over powdered zinc metal and under reduced nitrogen pressure. This was repeated several times until all color was removed. Aniline was stored in a refrigerated, sealed, foil-covered flask and was redistilled whenever color began to appear. Sephadex G75 was obtained from Pharmacia Fine Chemicals. DEAE-cellulose (DE-52, preswollen) was a product of Whatman Ltd. Acrylamide for gel electrophoresis experiments was from E-C Apparatus Corp. and Eastman Kodak. It was a grade “suitable for u.v. scanning.” Bisacrylamide and Stains All (1-ethyl-2-[3-1-cthylnaphtho][1,2-d] thiazol-in-2-ylidene-2-methyl-propenyl] naptho [1,2-d] thiazolium bromide) was from E-C Apparatus Corp. All other chemicals were analytical reagent grade. Dialysis tubing with a molecular weight cutoff of 3500 was purchased from Arthur H. Thomas Co. This dialysis tubing was used in all experiments.

(b) Preparation of 5' 3/5 fragment

The methods of Wintembery & Zachau (1970) and of Simsek et al. (1973) for cleavage of a tRNA at the modified nucleoside m^7G were modified to give reduced side reactions and increased yield. A sample of tRNA was dissolved in a minimal amount of deionized water and was dialyzed overnight versus 1 l of 0.1 mm-EDTA. The sample was dried by evaporation with dry nitrogen gas. The tRNA was dissolved at a concen of 1 mg/ml in 50 mm-NaOH. This solution was incubated at 25°C for 15 min. The solution was then rapidly brought to pH 5-0 (taking care not to go below pH 5-0) by dropwise addition of 1 m-acetic acid. An equal volume of a pH 5-0 solution of 0-3 m-aniline (0-0275 ml aniline
in 1 ml water, pH adjusted by addition of 12 M-HCl) was then added to the sample and the resulting solution was incubated at 37°C for 4 h. At this point, the sample was either applied directly onto a heated G75 gel filtration column or frozen at -70°C for future use. The average yield was about 35%.

(c) Preparation of 5' 3/5 fragment

Pure 5' 3/5 fragment was dialyzed versus deionized water, lyophilized, and then dissolved to a concn of 1 mg/ml in 0.1 M-sodium phosphate, pH 2.9 (Li et al., 1973). The 3/5 fragment caused an increase in pH, and this was compensated for by addition of 0.1 M-H₃PO₄ to lower the pH to 2.9. This solution was incubated at 25°C overnight and at 37°C for 2 h on the next day. The pH was adjusted to 5.0 with 0.1 M-NaOH and 2 vol. 0.45 M-aniline, pH 5.0 (0.041 ml/ml water, pH adjusted with 12 M-HCl) were added. This preparation was incubated at 25°C for 5 h under a nitrogen atmosphere. The sample was then either applied to a heated G75 column or frozen at -70°C. The average yield was about 20%.

(d) Preparation of 5' 1/4 fragment

Yeast tRNA⁰⁰ was dissolved to a concn of 1.5 mg/ml in 50 mM-MgCl₂, 50 mM-Tris·HCl, pH 9.5 (the pH was checked at 50°C because the pH of Tris changes with temperature) and incubated at 50°C for 2.5 h (Wintermeyer & Zachau, 1973). The pH was then lowered to 5.0 with 1 M-acetic acid and the sample was either frozen at -70°C or applied to a heated G75 column. The average yield was about 30%.

(e) Fragment separation

All fragments prepared from the tRNA were purified by similar methods using gel filtration chromatography at elevated temperatures (Schmidt et al., 1970; Reid et al., 1972). Sephadex G75 in a 2.5 cm x 165 cm column was maintained at 60°C by circulating water at 63°C through a column jacket using a Haake FK waterbath. The column was eluted using 0.1 M-NaCl, 0.01 M-potassium phosphate, pH 7.2 (column buffer) (Reid et al., 1972). Buffer was boiled before use and continuously heated to approx. 80°C while being applied to the column. This prevented bubble formation in the gel. A minimum length of tubing between buffer bottle and column was also essential to prevent bubble formation. The column was eluted at a rate of approx. 110 ml/h and the effluent monitored by measuring absorbance at 254 nm using a Gilson Biochemical u.v. Monitor. The appropriate fractions were pooled and put onto a DEAE-cellulose column (DE-52, 1.0 cm x 2.0 cm, equilibrated with column buffer). The nucleic acid binds to the column allowing any foreign substance generated by the heat degradation of Sephadex to be washed off. This column also served to concentrate the tRNA fragments. The sample was removed from the DEAE column by eluting with 1 M-NaCl, 0.01 M-potassium phosphate (pH 7.2). The effluent was monitored and the peak collected and dialyzed against at least 3 changes of deionized water. It was then lyophilized. If necessary, gel filtration chromatography was repeated on fragments to obtain a pure sample. The 5' 3/5 fragment sequence and that of the 5' 1/2 fragment were confirmed by RNase T₁ digestion (Wintermeyer & Zachau, 1970, 1973). The presence in the high-field spectra of only those methyl resonances expected from modified bases within the specific fragments also helped confirm the fragment identities.

(f) Polyacrylamide gel electrophoresis

Denaturing gels were run to confirm the length and purity of the fragments (Philippsen & Zachau, 1972). Gels were stained with Stains All (Dahlberg et al., 1969), which allowed quantitation of bands (results not shown).

Rₑ values for gel bands were determined by measuring their migration distance and the migration distance of the tracking dye in each gel. This allowed direct comparison of
bands in different gels. Gels were scanned to allow quantitation. Typical gel patterns are presented in Fig. 2.

(g) Sample preparation for n.m.r. studies

Yeast tRNA$^\text{phe}$ from Boehringer was prepared for low-field n.m.r. studies (low-field is defined as the spectral region below $-9$ p.p.m., relative to DSS) by pressure dialysis against 10 mM-EDTA, 40 mM-sodium cacodylate (pH 6.8) and then against 0.1 mM-EDTA, 40 mM-sodium cacodylate (pH 6.8). The pressure dialysis was accomplished by pressurizing the dialysis tubing with nitrogen during dialysis. This allowed the sample volume to remain small. The final tRNA concn was 1 mM. The fragments of tRNA were all dissolved directly in 0.2 ml of 1 mM-EDTA, 40 mM-sodium cacodylate, pH 6.8 (or 40 mM-sodium phosphate (pH 7.0) in the case of the 5' 1/2 molecule). The final fragment concentrations were approximately 1 mM in each case. After initial low-field melting studies were completed, 5 µl of 1.6 M-MgCl$_2$ were added to each 0.2-ml sample, giving a final magnesium concn of 40 mM. The samples were again melted in the low field. No magnesium was used with the 5' 1/2 fragment. (The final salt concentrations in the n.m.r. samples will be slightly higher than the added salt solutions due to the Donnan effect in dialysis (Tanford, 1961).) Whenever we had to use solution conditions different from that of crystallization we chose them so that the effect on conformation was negligible as judged by n.m.r.
studies (Bolton & Kearns, 1977; Robillard et al., 1977a) or salt-jump experiments, where the absorbance change at 260 nm is monitored before and after rapid addition of a concentrated salt solution (Boyle et al., unpublished results).

For high-field studies (high field is defined as the spectral region between -5 and 0 p.p.m. relative to DSS), the 5' 3/5 fragment used in the low-field work was diluted with deionized water and concentrated using an Amicon UM05 membrane. This was repeated 3 times and the sample was then lyophilized. A buffer of 40 mM-sodium phosphate (pH 6·8) in 2H2O was added to the powder. (Salts used in the high-field buffers were all the deuterated forms.) Lyophilization was repeated and 2H2O was added. This final solution was used in the high-field studies. The 5' 1/4 fragments were not retained by a UM05 filter, so a Sephadex G25 column (1·0 cm × 2·0 cm) was used to exchange the cacodylate buffer with the phosphate buffer. Again, the fragment was lyophilized and dissolved in the

40 mM-sodium phosphate (pH 6·8) buffer and lyophilized. The pH values stated for deuterated buffers are uncorrected pH meter readings.

The high-field spectrum of the 3/5 fragment was also recorded in the presence of magnesium. Deuterated magnesium chloride was added up to 40 mM by addition of 5 µl of a 1·6 mM solution to the 0·2 ml samples.

The high-field melting studies on the fragments were complicated by the fact that the fragments had been melted in the presence of magnesium in the low field. This may cause chain scission. The spectra for the 5' 3/5 and 5' 1/4 fragments were checked by high-field melts on samples that had not been studied in the low field. The samples, 0·3 mM in fragment, gave resonances that were identical to the previously melted samples (whose concentrations were somewhat less than 1 mM due to loss in handling).

Cylindrical n.m.r. microcells (Wilmad) were used in all studies. A volume of 0·15 to 0·20 ml was needed to fill them.

FIG. 3. Illustration of the extensive base-stacking in the crystal structure of yeast tRNA\textsuperscript{Phe}. The Watson-Crick base-pairs in the stems are shown as long slabs and unpaired bases as short slabs. The tertiary base-pairs are shown either as bent slabs or slabs connected by dark rods. a.a., amino acid; a.c., anticodon.
(b) n.m.r. system

N.m.r. spectra were recorded on a Bruker HXS-360 spectrometer equipped with a Nicolet NC-80 data system. An identical Nicolet NC-80 data system was used for off-line data manipulation. The spectra were obtained by correlation spectroscopy (Dadok & Sprecher, 1974).

The temperature of the sample was constant to within ± 1 deg. C. Both low and high-field spectra were collected for each sample, starting at or below room temperature and continuing until the spectrum for each was devoid of resonances (in the low field) or had resonances in the denatured positions (in the high field). The sample was allowed to equilibrate at each new temperature by observing whether the water resonance had stopped "moving" with respect to its previous position. Equilibration normally required 15 to 30 min. After melting, each sample was cooled slowly and magnesium was added. Then a melting study in the presence of magnesium was carried out.

Areas of resonances were measured from the plotted spectra by counting blocks on ruled paper. The areas of isolated resonances such as those at 14.4 p.p.m. were considered as belonging to one proton (Robillard et al., 1976).

3. Results

(a) Identification of low-field n.m.r. resonances in the whole molecule

Since the pioneering studies by Kearns et al. (1971), it has been well-established that each Watson-Crick base-pair in a double helix produces one low-field n.m.r. resonance in aqueous solution due to a hydrogen-bonded proton from the ring nitrogen. The relative positions of these resonances are a direct consequence of different environments (for a review, see Kearns, 1976).

A number of recent theoretical (Robillard et al., 1976b; Kearns, 1976; Geerdes & Hilbers, 1977; Kan & Ts'o, 1977) and experimental studies (Robillard et al., 1976a; Römer & Voradi, 1977) have been executed in attempting to assign the resonances in the low-field spectrum of intact yeast tRNA^Phe. These studies show good agreement as regards the assignment of 20 proton resonances (from 20 Watson-Crick base-pairs in the stems) to seven spectral regions indicated in Table 1. Detailed assignments within each region, however, differ among research groups due to the different assumptions that enter into calculations and interpretations. For our purpose, we used mostly the assignment of spectral regions. For further discussions, the specific assignments used here for proton resonances from Watson-Crick hydrogen bonds are those determined previously (Robillard et al., 1977a) by a combination of temperature-jump relaxation data and n.m.r. melting studies on the high and low-field spectral regions.

Integration of the spectrum of yeast tRNA^Phe in the presence of magnesium at 25 to 35°C gives an area indicative of 25 or 26 protons between -11.5 and -14.5 p.p.m. with an estimated error of about 1 proton (Robillard et al., 1977a; Kan & Ts'o 1977; Geerdes & Hilbers, 1977). This work, although the solution conditions are different from those in the above studies, also gave an area representative of 26 ± 1 protons in the native molecule. Since there should be only 20 resonances from Watson-Crick base-pairs in the stems to be accounted for in this region, the remaining area can be assigned to resonances resulting from certain tertiary base-pairs, i.e. tertiary hydrogen bonds involving hydrogens of ring nitrogens. These include the tertiary Watson-Crick pair G19·C56, reversed Hoogstein pairs U8·A14 and T54·A58, a
### Table 1

*Spectral regions (in p.p.m.) and their assignments for intact yeast tRNA<sub>Pho</sub>*

<table>
<thead>
<tr>
<th>Spectral region</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acceptor stem</td>
<td>—</td>
<td>5, 6, 7</td>
<td>—</td>
<td>—</td>
<td>1, 2, 3</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>D stem</td>
<td>—</td>
<td>12</td>
<td>11, 13</td>
<td>—</td>
<td>19</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Anticodon stem</td>
<td>—</td>
<td>—</td>
<td>29, 31</td>
<td>—</td>
<td>27, 30</td>
<td>—</td>
<td>28</td>
</tr>
<tr>
<td>T stem</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>50, 52</td>
<td>53</td>
<td>49, 51</td>
<td>—</td>
</tr>
<tr>
<td>Tertiary base-pairs</td>
<td>54.58, 8.14</td>
<td>—</td>
<td>—</td>
<td>19.56, 22.46</td>
<td>—</td>
<td>26.44</td>
<td>(15.48)</td>
</tr>
</tbody>
</table>

Each base-pair is indicated by a single number corresponding to the sequence number of the first of the 2 bases in the base-pair, for example 5 represents the base-pair 5-68. Tertiary base-pairs are denoted by 2 sequence numbers joined by a dot.

Minus signs in the spectral regions are omitted for simplicity.
base-pair in the 13-22-46 triplet, namely G22·m7G46, and base-pair m2G26·A44. Other resonances that may be expected in the low-field regions include tertiary pairs G15·C48 and G4·U69, both of which contain ring nitrogen hydrogens hydrogen-bonded to exocyclic oxygens. Except for the two tertiary interactions m2G26·A44 and G22·m7G46, the assignments of resonances to protons in tertiary hydrogen bonds are from those proposed by Robillard et al. (1977a). The assignment of the proton from the 26·44 base-pair has been shifted to region F, because n.m.r. spectra of three class I tRNAs having a purine 26·purine 44 interaction show a resonance here, while two tRNAs lacking this interaction have no peak in this region (Reid & Hurd, 1977). The location of the proton resonance from G22·m7G46 is uncertain; chemical modification experiments in different laboratories generate contradictory results (for a review, see Robillard, 1978). However, this interaction is not important in the present fragment studies; therefore, its assignment will not be discussed further.

The assignments used in this study are presented in Table 1. To help follow the discussion, a drawing emphasizing the base-stacking feature of the crystal structure of yeast tRNA\textsuperscript{Phe} is shown in Figure 3.

(b) Low-field n.m.r. study on the 5' 1/4 fragment

This fragment shows little or no n.m.r. pattern in the absence of magnesium, even at very low temperatures. However, in the presence of the divalent ion and at the lowest temperatures, there are low, broad peaks centered around −14.15, −13.2 and −12.9, and in the −10.8 to −10.5 region. All of those are reasonably clear at 7.5°C (Fig. 4). The location of the peaks suggests that any structure involved should have both A·U (below −12.7 p.p.m.) and G·C pairs (above −12.7 p.p.m.). The peak in the −10.8 to −10.5 regions can represent many different hydrogen bonds such as G·U pairs and those between amino groups and exocyclic oxygens. Examination of all possible base-pairings reveal only two reasonable intramolecular base-paired structures.

\[
\begin{align*}
5' & \text{G·C·G·G·A·U·U·U·U·A·G} \\
3' & \text{D·G·A·C·U·C·G·A}
\end{align*}
\]

Using the results of Gralla & Crothers (1973a,b) we find that both possible helices would have melting temperatures of approximately −60°C at the salt concentrations used in these experiments. However, there are two possible bimolecular structures with approximate melting temperatures of −17 and −8°C:

\[
\begin{align*}
5' & \text{G·C·G·G·A·U·U·U·A·G·C·U·C·A·G·D} \\
3' & \text{D·G·A·C·U·C·G·A·U·U·U·A·G·G·C·G}
\end{align*}
\]

and

\[
\begin{align*}
5' & \text{G·C·G·G·A·U·U·U·A·G·C·U·C·A·G·D} \\
3' & \text{D·G·A·C·U·C·G·A·U·U·U·A·G·G·C·G}
\end{align*}
\]
These helices could be responsible for the resonances observed in Figure 4. However, since they appear to result from intermolecular aggregation at the high concentrations employed, they are irrelevant for the folding of a single chain. In summary, there is no stable intramolecular structure formed for this fragment under the experimental conditions.

(c) Low-field n.m.r. of the 5' 1/2 fragment

The temperature dependence of the 5' 1/2 fragment in 0.04 M-sodium phosphate (pH 7.0) is shown in Figure 5. To a first approximation, the lowest temperature spectrum shows the four resonances expected for the D stem. These are the peaks in regions B (-14.1 p.p.m.), C (-13.27 p.p.m.), E (-12.7 p.p.m.) and G (-11.7 p.p.m.). The areas under these peaks yield a ratio of about 1:1:1:1. Examination of all possible base-paired structures for this fragment shows that there are four structures that have three or more consecutive base-pairs:

3'...C-U-C... 3'...C-U-C-G... 3'...U-U-U-A-G... and 3'...U-C-A-G...

Of these, only the second one, corresponding to the D stem, has an estimated melting temperature close to that observed by a temperature-jump measurement (Boyle, Cole

Fig. 4. 360 MHz proton n.m.r. spectra of the low-field region of the 5' 1/4 fragment as a function of temperature in 40 mM-sodium cacodylate, 1 mM-EDTA, 40 mM-MgCl₂ (pH 6.8). The regions designated by letters are those to which protons are assigned (see Table 1). The chemical shift is in p.p.m. with respect to DSS.
& Kim, unpublished data) and is compatible with the n.m.r. spectra in Figure 5. Also, the resonances show that the base-pairs melt in unison, as would be expected from the presence of only one structure.

Starting from the left in Figure 5, the peak in region B is from base-pair 12. It is shifted downfield from the predicted location in the native molecule. This can be explained by assuming that the U12•A23•A9 base-triplet is not formed, although all three bases are present in the fragment. When the base-triplet melts in the native structure, base-pair 12 suffers a similar downfield shift (Robillard et al., 1977a). The peak in region C is from base-pair 11. Its location is in very good agreement with the position predicted from the crystallographic structure.

Fig. 5. 360 MHz proton n.m.r. spectra of the low-field region of the 5' 1/2 fragment as a function of temperature in 40 mM-NaH₂PO₄ (pH 7.0). The lettered regions are as in Table 1. Resonance assignments for the base-pairs in the D stem of the intact tRNA are in regions B, C and E for base-pairs 12; 11 and 13; and 10, respectively, and shown here for comparison.
The intensity at -12.7 p.p.m. in Figure 5 can be attributed to the proton of base-pair 10. In the 5' 1/2 fragment, m2G26 is not fully stacked on this base-pair, resulting in a reduced upfield shift relative to its location in the intact molecule. The position of base 26 will definitely exert an effect on base-pair 10 in the fragment, because it has the largest contribution of this base-pair's shift in the native molecule (Robillard et al., 1977b). Without some stacking by base 26 and its subsequent contribution to the shift, the resonance from base-pair 10 would be far downfield of the position seen. There is an apparent partial resonance in regions D and E that may result from an alternative environment for this base-pair. This peak is seen most clearly in the 31°C spectrum. Here the sum of the two apparent peaks in regions D and E yields one proton. The presence of two resonances from one base-pair could be explained by a slow interconversion between two states with the -12.7 p.p.m. peak having the higher population (James, 1975). This base-pair is on the end of a helix and so could experience changing environments.

The peak in region G is the broadest of the resonances reported here. Its position is in good agreement with that expected from G·C 13 if tertiary base-pair 8·14 is not present. When the tertiary pair is present in the native molecule, base 8 is removed from a stacking position over base-pair 13. However, when U8 is not hydrogen-bonded to A14, A14 will stack over base-pair 13 and so cause an upfield shift to the position seen in this fragment and seen in intact tRNA without magnesium (Robillard et al., 1977a). Therefore, tertiary pair 8·14 is not present in this fragment.

The broad nature of this resonance is seen in other work on similar 1/2 fragments (Rordoff, 1975; Lightfoot et al., 1973). It may be due to the presence of different environments as was possible with base-pair 10, because G·C 13 is at the end of a helix. As in the case of G·C 10, there may be a shoulder on this resonance; it is on the downfield side. It is seen at 31°C most clearly. As before, this may be explained by a slow interconversion between two (or more) states.

In summary, the 5' 1/2 fragment contains all the base-pairs in the D stem, but two tertiary base-pairs, 8·14 and 9·12·23, found in this portion of the intact molecule are absent.

(d) Low-field n.m.r. of the 5' 3/5 fragment

This fragment contains bases that can make up the complete D and anticodon helices and loops. The examination of all possible base-paired structures of this fragment shows that there are seven such structures that can form three or more consecutive base-pairs with reasonable loop size (Gralla & Crothers, 1973a,b; Borer et al., 1974).

\[
\begin{align*}
5'\ldots & G\cdot A\cdot G \ldots & 5'\ldots & G\cdot A\cdot G\cdot C \ldots & 5'\ldots & A\cdot A\cdot G\cdot U\cdot C \ldots & 5'\ldots & A\cdot G\cdot U\cdot C \ldots \\
3'\ldots & C\cdot U\cdot C \ldots & 3'\ldots & C\cdot U\cdot C\cdot G \ldots & 3'\ldots & U\cdot U\cdot U\cdot A\cdot G \ldots & 3'\ldots & U\cdot C\cdot A\cdot G \ldots \\
5'\ldots & G\cdot U\cdot C \ldots & 5'\ldots & G\cdot A\cdot C \ldots & 5'\ldots & G\cdot G\cdot U\cdot C\cdot U \ldots \\
3'\ldots & C\cdot A\cdot G \ldots & 3'\ldots & C\cdot U\cdot C \ldots & 3'\ldots & C\cdot C\cdot A\cdot G\cdot A \ldots
\end{align*}
\]

Out of these, only the second and seventh, corresponding to the D and anticodon stems, have estimated melting temperatures close to those observed experimentally
Fig. 6. 360 MHz proton spectra of the low-field region of the 5' 3'/5 fragment as a function of temperature in 40 mM-sodium cacodylate, 1 mM-EDTA (pH 6.8). The regions are as in Table 1. Resonance assignments for the base-pairs in the D stem and anticodon stem of the intact tRNA are in regions B, C, E and G for base-pairs 12; 11, 13, 29 and 31; 10, 27 and 30; and 28, respectively, and they are shown here for comparison.
by temperature-jump measurements (Boyle, Cole & Kim, unpublished data). Those two structures can also account for the distribution of the resonance peaks according to the assignments in Table 1.

To a first approximation, it appears that the expected resonances from the D and the anticodon stem are present in the n.m.r. spectra of the 5' 3/5 fragment. Figure 6 shows the spectra as a function of temperature in the absence of magnesium. Proceeding through the spectra from left to right, we can assign peaks and explain melting behavior (the relationships between n.m.r. and optical melting have been described by Crothers et al., 1974).

The lowest-field peak is most likely the resonance from U12·A23 that has undergone a downfield shift. This shift could be explained by assuming that the U12·A23·A9 base triplet is absent, just as in the 5' 1/2 fragment. Even in this larger fragment, this tertiary bond appears not to be formed.

Region C has four predicted resonances in the intact native molecule that would arise from the D and anticodon helices. Integration shows an intensity of three protons. This is easily reconcilable when the case of base-pair 13 is considered. As in the 5' 1/2 fragment, because tertiary base-pair 8·14 is lacking, the resonance for base-pair 13 is shifted far upfield. This region is the first to lose a resonance on heating. By 32.5°C, the area amounting to one proton is lost. This is probably from base-pair A·Ψ31. This pair is thought to be less stable than an A·U pair and is expected to exchange at a lower temperature (Lightfoot et al., 1973; Robillard et al., 1977a; Reid & Hurd, 1977). The next base-pair to melt in region C is 29. This is an A·U pair and its melting occurs between 32.5 and 45°C. The last to melt in this region is base-pair 11, consistent with its being in the D stem, the last stem to melt in the intact tRNA (Robillard et al., 1977a).

Region D is predicted to have no resonance from the D and anticodon helices, yet one or two resonances are present. The area for one proton can be explained by the downfield shift of base-pair 10. In the intact tRNA, G45 is very near base-pair 10; it is close to being coplanar with it. Since base 46 is not present and tertiary base-pair 22·46 cannot form, base 45 may no longer be constrained to stay near base-pair 10. Also base-pair 10 may not stack as well on base 26 in the fragment, since the 13·22·46 base triplet is not present, and this might allow the D stem to shift a little away from the helix axis of the anticodon stem. The position of this resonance in this fragment is identical to that in the smaller 5' 1/2 fragment, arguing for identical environments in both sequences. The extra area in region D may result from the presence of the tertiary pair m5G26·A44, which is an extension of the anticodon helix, and stacked under the D stem (see Fig. 3).

Region E has an area equivalent to two protons, as expected if the resonance of base-pair 10 is shifted downfield to region D. They represent base-pairs 27 and 30 from the anticodon stem and, along with base-pair 28 in region G, they are the last pairs to melt in this stem. It is assumed that base-pair 27 is the first to melt in this region, since it is at the top of the stem. It is lost by 40°C. The fact that the resonance of base-pair 27 is in this region is good evidence for the existence of tertiary pair 26·44. If the tertiary pair were gone, base 44 would not stack as well on base-pair 27, and its resonance would be expected to be shifted into region D, leaving E with only one proton.
Region F contains a resonance from base-pair 13, which has been shifted upfield by the stacking of A14. It is in the same position in the spectrum of the 5' 1/2 fragment. Region G contains the resonance of base-pair 28. This proton is strongly shifted in the native tRNA by the closeness of the ring centers of G43 and A29 in the helix (Robillard et al., 1977a). In the fragment, if tertiary pair 26-44 is present, the proximity of the rings should be the same. The peak in G is seen at $-11.8$ p.p.m., whereas in the whole molecule it is at $-11.6$ p.p.m.

In summary, the 5' 3/5 fragment contains all the base-pairs in the D and anticodon stems plus one tertiary base-pair, G26·A44. However, two remaining tertiary base-pairs, U8·A14 and U12·A23·A9, are still not formed in the fragment.

The spectra in the presence of magnesium (Fig. 7) are similar to those in its absence (Fig. 6). Close examination of intensities points out several differences. The largest change is above $-11.5$ p.p.m.; here the relative intensities due to hydrogen bonds between the amino nitrogen and exocyclic oxygen have increased in the presence of magnesium. This probably reflects a more stable structure allowing less free exchange of these protons. In addition, magnesium acts to stabilize the tRNA with respect to melting by specifically stabilizing tertiary structure (Stein & Crothers, 1976; Holbrook et al., 1977). The elevation of melting temperature is far more greatly enhanced in the intact molecule than in this fragment; magnesium ions raise the melting temperature (as judged by hydrogen exchange) of the intact molecules by $20 \sim 30$ deg. C. However, the corresponding increase for the 5' 3/5 fragment is, at the most, $10$ deg. C.

Thus, magnesium seems to be more important for maintaining the stability of the intact nucleic acid than for the correct folding of it. As seen in all fragments and the intact molecule, folding occurs accurately even in the absence of magnesium with a stability that is lower but comparable to that in its presence.

(e) Low-field n.m.r. of the intact molecule

Spectra of the intact tRNA were recorded under the buffer conditions used for the fragments in this study. The spectra in Figure 8 demonstrate that, both in the presence and absence of Mg$^{2+}$, the spectra under the present buffer conditions, which are close to that of the crystal structure, are essentially identical to those recorded in previous studies, under different conditions, by Robillard et al. (1977a). Thus, the procedure employed here, of using previous spectral assignments to interpret spectra recorded under our conditions, is justified.

(f) High-field n.m.r. studies on tRNA fragments

Bases in tRNA are frequently modified by the addition of methyl groups. The methyl protons (and methylene protons from dihydrouridine) will resonate between 0 and $-5$ p.p.m. These protons provide useful reporter groups that help in elucidating molecular structure changes (Kan et al., 1974; Robillard et al., 1977a; Kan & Ts'o, 1977). In the case of the fragments, the study of their resonances provides additional support for interpretations inferred in the low-field n.m.r. studies.

The temperature dependence of the methyl resonance of the 5' 3/5 fragment in 40 mM-Na$_2$PO$_4$ (pD 6.8) with no Mg$^{2+}$ is presented in Figure 9. The 5' 3/5 fragment's high temperature spectrum contains all of the resonances expected from the methyl or
FIG. 7. 360 MHz proton spectra of the low-field region of the 5' 3/5 fragment as a function of temperature in 40 mM-MgCl$_2$ (pH 6.8). The regions are as in Table 1. Resonance assignments for the base-pairs in the D stem and anticodon stem of the intact tRNA are in regions B, C, E and G for base-pairs 12; 11, 13, 29 and 31; 10, 27 and 30; and 28 respectively, and they are shown here for comparison.
Fig. 8. 360 MHz proton spectra of the low-field region of yeast tRNA^Phe as a function of temperature in 40 mM-sodium cacodylate, 1 mM-EDTA (pH 6.8), + 40 mM-MgCl$_2$ (a); no MgCl$_2$ (b). The regions are as in Table 1.
Fig. 9. 360 MHz proton spectra of the high-field regions of the 5' 3/5 fragment in 40 mM NaH₂PO₄ (pH 6.8). Impurity peaks are indicated by asterisks. The resonances have been assigned to the various modified bases as indicated on the highest temperature spectrum. These resonances are then followed down through the temperatures used, by the broken lines. The DSS resonance is also labeled. There are two resonances from the D stem residues in this region.
methylene groups present. The resonance positions at high temperatures match those from the totally denatured tRNA or corresponding mononucleotides (Kan et al., 1977; Robillard et al., 1977a). This is true in both the presence and absence of magnesium. Chemical shifts obtained from these studies are shown in Figure 10. Small changes in position are seen for m5C40, Y and D. These changes are too small to decide if they result from the melting of a helix or some other rearrangement. m5C40 is in a helix but its methyl group is not near bases from the opposite strand. Therefore, its shifting could simply reflect the unstacking of Y39 from a position near the methyl group. Similarly, the shifts of the resonances of Y and D could result from unstacking in loops.

The largest effect of temperature on any methyl proton resonance is the effect on the methyl protons of m2G26. In intact tRNA, in the absence of Mg2+, the resonances from this group shift smoothly from one position to the other (Robillard et al., 1977a). This behavior is clearly dependent on solvent conditions, however. In this fragment, intensity broadens at one position and simultaneously appears at another (compare the 68 and 61.5°C spectra). The shifting in the native tRNA is indicative of a fast transition between two states (James, 1975). A simultaneous disappearance and

![Figure 10](image)

**Fig. 10.** Plot of the chemical shift as a function of temperature for the high-field resonances of the 5'-3' fragment in 40 mM-NaH2PO4 (pD 6.8). The resonance identities are on the right. The points are from 2 separate sets of data. The Δ points are to emphasize the drastic change of m2G10 resonance as a function of temperature.
reappearance of the resonance in these fragment spectra reflects an intermediate or slow transition similar to that seen in the intact tRNA for the T methyl group (Robillard et al., 1977a).

It is perhaps significant that at the lower temperatures the resonances from the 5' 3/5 fragment are at the same location as in the intact molecule. This argues strongly for identical environments for methylene and methyl groups in both species and, thus, the same structures for this portion in the native tRNA and in the fragment.

The small fragment, 5' 1/4, was examined in a limited way. The only resonance expected and the only one present was from m2G10. The position seen was that of denatured tRNA and of mononucleotides.

4. Conclusions and Discussion

(a) Correct secondary structures are formed but very few tertiary hydrogen bonds are established in the "folding intermediates"

The n.m.r. studies described here show that secondary structure (base-paired double-helical stems) of yeast tRNA\textsuperscript{Phe} can be built up as the molecule is being synthesized; the entire sequence need not be present for correct parts of the final structure to be formed. When the nucleotides available for formation of the D stem are present, this secondary structure will form. Likewise, when the additional nucleotides needed for formation of the anticodon stem are present, this secondary structure will form (see Fig. 11).

Furthermore, some correct tertiary hydrogen bonds (between non-complementary bases) are made at an intermediate stage before the entire sequence is available: in the 5' 1/2 fragment it was seen that the tertiary base-pair 8-14 and the tertiary triplet 12-23-9 do not form, despite the fact that all the necessary residues are present in this fragment. However, in the 5' 3/5 fragment, although the tertiary base-pairs 8-14 and 12-23-9 remain unformed, the 26-44 tertiary base-pair appears to be present. Since the latter is at the junction between the stems, the presence of the 26-44 tertiary pair suggests that the D stem is stacking over the anticodon stem (as in the crystal structure) even in this incomplete molecule. Thus, the D stem–anticodon stem “domain” may be stabilized even early in its synthesis. The lack of the 8-14 and 12-23-9 tertiary bonds indicates that there is no rigid form present in the 5' tail of the D stem, and thus no predominant tertiary structure as a folding intermediate. This single-stranded tail will later form the acceptor stem and then 8-14 and 12-23-9 tertiary interactions.

In protein folding, there are two major hypotheses: one in which the hydrophobic core formation is considered as the primary event of the protein folding (Kauzmann, 1959); the other emphasizing the formation of stable hydrogen-bonded secondary structures, such as \(\alpha\) helices, which in turn interact among themselves to form a tertiary structure (Ptitsyn & Rashin, 1975). In the case of tRNA folding, both the hydrophobic core and the hydrogen bonded secondary structure formation are achieved simultaneously by the base-paired double-helical stems, in which hydrogen-bonded base-pairs are stacked to form the long hydrophobic “core” of the double helices. The high stability of these stems probably prevent any major reshuffling of the folding intermediate structures of tRNA.
FIG. 11. Schematic representation of the sequential folding of yeast $tRNA^{Phe}$. The broken line on the 5' end represents possible precursor sequences. (a), (b) and (c) are fragments studied in this work. (c), (d) and (e) contain a possible intervening sequence in the anticodon loop. (a) contains no secondary or tertiary structure; (b) contains the D stem; (c) contains the D stem, the anticodon stem, and a tertiary interaction pair 26-44; (d) contains all secondary and tertiary interactions except for the complete acceptor stem; (e) is a completed precursor tRNA; (f) is the final product of the biosynthetic pathway. Base-pairs are indicated by long open rods and tertiary base-pairs by solid rods.
Magnesium ions are not essential for correct folding, but enhance the stability

Magnesium or other divalent metals are known to perform a crucial role in stabilizing tRNA structure (Fresco et al., 1966). But this, as well as earlier n.m.r. studies, shows that correct secondary and tertiary structures are able to form without any divalent metal ions; i.e. they are stabilizing factors but not essential for the correct folding. A similar role was found for spermine (Bolton & Kearns, 1977).

Extra sequences in tRNA precursors probably do not affect the folding

Some qualifications must be applied to the conclusions reached in this work. The tRNA fragments used as models for nucleotide sequences at intermediate stages of tRNA biosynthesis were fully modified, whereas the true intermediates are unmodified when they are synthesized. However, examination of all the possible base-paired structures shows that none of the possible helical structures contain those modified bases that might be expected to hinder base-pair formation. A more serious consideration involves the precursor forms of tRNA. It has been found in eukaryotes (Garber et al., 1978) and prokaryotes (Altman & Smith, 1971; Seidman et al., 1975) that a tRNA is synthesized as a precursor containing more nucleotides at one or both termini than in the final form.

To examine possible additional structure in the 5'-end portion of a precursor, we took the known DNA sequence starting 29 nucleotides before the mature 5' end and ending at the end of the D stem of yeast tRNA\textsuperscript{phe} gene (Valenzuela et al., 1978), and constructed all possible base-paired structures for a hypothetical tRNA precursor intermediate. (There is, at present, no way of estimating how many of the added bases at the 5' end are really present in the precursor, but some points can still be made.) Among these, none of the extra bases in the "hypothetical precursor intermediate" was found to interfere with formation of the D stem by forming a more stable stem. The next most stable structure is a short stem that can be formed using three of the bases of the D stem and the first base preceding the mature 5' end plus the first two bases in the tRNA. It is definitely less stable than the D stem and, if it does exist in a precursor, might be expected to give its bases to form a D stem by the time the complete D stem sequence has been synthesized. Besides, this second most stable stem can only form in two out of the three sequences known for yeast tRNA\textsuperscript{phe} genes (Valenzuela et al., 1978). There is additional experimental evidence: the chemical modification of a tRNA precursor from Escherichia coli showed the same protection pattern as the mature tRNA, indicating that its structure is extremely similar to the final processed product (Chang & Smith, 1973). These observations give strong credence to the notion that, even in tRNA precursors, the mature portion of the precursor is folded correctly just as in the mature tRNA.

Recently, intervening sequences have been found in three genes for yeast tRNA\textsuperscript{phe} and in semi-matured precursors of the tRNA (Valenzuela et al., 1978; Knapp et al., 1978). In all cases the extra sequences are located immediately after the anticodon. Such added nucleotides can form alternate base-paired structures in this region. Thus, the anticodon stem, as such, may not be formed in the precursor. However, all of the stable structures postulated (Abelson, 1979) maintain at least three base-
pairs at the top of the anticodon stem. Therefore, even for the precursors, the stacking of the D stem on top of the anticodon stem mediated by the 26-44 tertiary base-pair (as seen in the 5' 3/5 fragment) seems likely. Also, since the anticodon stem points away from the rest of the molecule and interacts with it only at its top (Holbrook et al., 1978), the intervening sequence structure will not be likely to interfere with the rest of the structure. Hence, the folding pattern seen will hold even in the presence of intervening nucleotides. Further, not all of the tRNA gene sequences known from eukaryotes have inserted sequences. The genes for tRNA^Arg and tRNA^Asp in yeast reveal no inserted nucleotides (G. Knapp, personal communication). Thus, a general mode of folding, we propose, should also account for these tRNAs, and the most general model is one in which the intervening sequence does not play an important role for the overall structure and mode of folding.

(d) Sequential folding to limit the number of folding pathways

How important is sequential folding to the final tRNA structure? It is believed that all tRNAs have the same overall secondary and tertiary structure (Kim et al., 1974; Kim, 1978) representing the free energy minimum or one of several minima of all tRNAs. If an entire sequence for a tRNA is present, there will undoubtedly be a large number of folding pathways to one or more final stable structures. However, by folding sequentially from the 5’ end, the number of folding pathways is severely limited and the correct folding for the tRNA population is readily ensured: it is likely that the correct folding of the D stem at an early stage of tRNA biosynthesis leads to one or a limited number of smooth pathways to the final correct tertiary folding.

There may exist some evolutionary pressure to maintain a proper sequence of folding, the D stem folding in particular. Using the tRNA sequences available (Sprinzl et al., 1978), all possible base-paired structures for the nucleotide sequences up to the ends of the D stems were examined. The energies for any stems formed were calculated by the methods of Gralla & Crothers (1973a,b) and Tinoco et al. (1973), and compared to that of the D stem. Of all of the class I tRNAs (4 base-pairs in the D stem), only one forms a more stable structure than the D stem. Even this structure is more stable by only 0.9 kcal/mol. (This difference could be readily accounted for by errors in the energy calculation procedures.) This suggests that there is a strong requirement for the correct formation of the D stem, presumably in order to limit the number of folding pathways among the vast number of possible pathways.

In summary, the sequential appearance of nucleotides during tRNA biosynthesis may play an important role in an efficient folding to achieve the correct functional form. A large number of possible folding pathways can certainly be reduced drastically to one or a few choices by the sequential appearance of secondary structure and, consequently, in some cases, closely probable structures can be eliminated by the sequential folding.

This work was supported by grants from the National Science Foundation (PCM76-04248) and the National Institutes of Health (CA-15802), and by a travel grant to one of us (J. B.) from Duke University. We also wish to acknowledge the ZWD (Netherlands
Foundation for Pure Research) for their support of the 360 MHz n.m.r. facility at the University of Groningen. Most of this work was completed while two authors (J. B. and S. H. K.) were in the Department of Biochemistry, Duke University, Durham, N.C., U.S.A.

REFERENCES