Structure of an Unmodified tRNA Molecule†

Kathleen B. Hall,‡ Jeffrey R. Sampson,§ Olke C. Uhlenbeck,§ and Alfred G. Redfield*,‡

Department of Biochemistry, Brandeis University, Waltham, Massachusetts 02254, and Department of Biochemistry and Chemistry, University of Colorado, Boulder, Colorado 80309

Received December 9, 1988; Revised Manuscript Received April 5, 1989

ABSTRACT: We have used NMR to study the structure of the yeast $tRNA^{Phe}$ sequence which was synthesized by using T7 RNA polymerase. Many resonances in the imino ^1H spectrum of the transcript have been assigned, including those of several tertiary interactions. When the Mg^{2+} concentration is high, the transcript appears to fold normally, and the spectral features of the transcript resemble those of $tRNA^{Phe}$. The transcript has been shown to be aminoacylated with kinetics similar to the modified $tRNA^{Phe}$ [Sampson, J. R., & Uhlenbeck, O. C. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 1033-1037], suggesting that the structure of the two molecules must be similar. In the absence of Mg^{2+} or at [tRNA]:[Mg^{2+}] ratios <0.2, the transcript does not adopt the native structure, as shown by both chemical shifts and NOE patterns. In these low Mg^{2+} conditions, a second GU base pair is found, suggesting a structural rearrangement of the transcript. NMR data indicate that the structure of a mutant having G20 changed to U20 is nearly identical with that of the normal sequence, suggesting that the low aminoacylation activity of this variant is not due to a substantially different conformation.

The structure of yeast $tRNA^{Phe}$ has been determined by several physical methods. In particular, crystal structures have been determined (Jack et al., 1976; Quigley et al., 1975; Stout et al., 1978; Sussman et al., 1978), and the solution structure as a function of solvent conditions has been studied by NMR (Reid, 1981; Roy & Redfield, 1982; Heerschap et al., 1982, 1983a,b). Those ^1H NMR experiments have observed the imino proton resonances of protons that are hydrogen-bonded in each base pair (Kearns & Shulman, 1974). Using nuclear Overhauser effect (NOE) methods (Noggle & Shirmer, 1971), with reference to the crystal structure, the resonances of the imino proton spectra of yeast $tRNA^{Phe}$ have been assigned to the individual base pairs of the molecule. A similar approach has been taken for the study of 5S RNA (Leontis & Moore, 1986).

Recently, experiments have shown that a yeast $tRNA^{Phe}$ transcript, synthesized in vitro by T7 RNA polymerase, is a good substrate for the cognate yeast phenylalanyl-tRNA synthetase (Sampson & Uhlenbeck, 1988), even though this transcript lacks all base modifications normally present on the tRNA. This result indicates that the modifications are not necessary for the aminoacylation reaction and suggests that the structure of the transcript must be similar to the native tRNA.

To study the structure of that tRNA transcript, we have used the NMR resonances of the imino protons to study the tRNA conformation as a function of temperature and $MgCl_2$ concentration. First, we describe the methods for synthesizing the large amount of RNA necessary for structural studies and then present the NMR results for the transcript of both the wild-type sequence of yeast $tRNA^{Phe}$ and a variant having a single G20-to-U20 substitution. A comparison of the NMR spectra of these two tRNA transcripts suggests an approach for structural comparisons of sequence variants of other large RNAs.

Materials and Methods

The construction of the plasmids p67YF0 and p67YF4 was described previously (Sampson & Uhlenbeck, 1988). When these plasmids are cleaved with the restriction enzyme *Bst*NI and transcribed with T7 RNA polymerase, the resulting transcripts begin with pppG1 of the yeast $tRNA^{Phe}$ sequence and end with the 3′-terminal CCA sequence.

Transcription Reactions. For each reaction, 0.6 mg of plasmid DNA was digested with *Bst*NI in 1.5 mL containing 10 mM Tris-HCl (pH 8 at 60 °C), 150 mM NaCl, 10 mM $MgCl_2$, and 500 units of *Bst*NI for 2 h at 60 °C. The DNA was phenol-extracted, ethanol-precipitated, and resuspended in 20 mM Tris/1 mM EDTA, pH 8.0.

For the 6-mL transcription reaction, 0.6 mg of the linearized DNA was incubated with 5 mM of each NTP (pH 8), 26 mM $MgCl_2$, 40 mM Tris, pH 8.1 at 37 °C, 5 mM DTT, 2 mM spermidine, 50 μg/mL BSA, and 0.36 mg of T7 RNA polymerase (1.6×10^5 units). The reaction continued for 4 h at 42 °C. The precipitate that appears was dissolved by the addition of 0.1 volume of 0.5 M EDTA, pH 8, before phenol extraction. The solution was then extracted with a solution of phenol/$CHCl_3$/isoamyl alcohol (25:24:1), and the phenol layer was then back-extracted with 0.6 volume of 20 mM Tris/1 mM EDTA, pH 8. The aqueous layers were combined (10 mL total), and 3 mL of 8 M NH_4OAc was added and mixed well. The nucleic acids were precipitated by the addition of 3 volumes of ethanol at −20 °C for 2 h. The precipitate was collected by centrifugation at 7500 rpm in a Sorvall swinging-bucket rotor for 30 min, washed with 70% ethanol, and dried. The pellet was resuspended in 1.5 mL of 50 mM NaOAc, 50 mM NaCl, and 1 mM EDTA, pH 7.0.

To separate the RNA from DNA, free nucleotides, and short abortive fragments, the resuspended RNA was run over a 15-mL Fractogel DEAE-Tsk 650S HPLC column. The 100-mL linear gradient began with 75% 50 mM NaOAc, 50 mM NaCl, and 1 mM EDTA, pH 7 (buffer A), and 25% 50 mM NaOAc, 800 mM NaCl, and 1 mM EDTA, pH 7 (buffer

† This work was supported by grants from the National Institutes of Health (O.C.U. and A.G.R.), the Damon Runyon-Walter Winchell Cancer Fund Fellowship, and the Lucille P. Markey Charitable Trust (K.B.H.).

‡ Brandeis University.
§ University of Colorado.

0006-2960/89/0428-5794$01.50/0

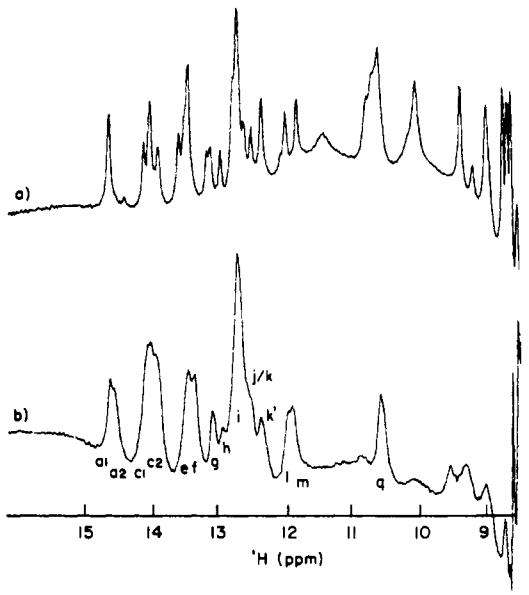

FIGURE 1: Imino ^{1}H NMR spectrum in H_2O of (a) modified yeast tRNA^{Phe} and (b) unmodified transcript at 30 °C in 100 mM NaCl, 10 mM sodium cacodylate, pH 7.0, and 5 mM free $MgCl_2$. The concentration of the modified tRNA is 1 mM; the concentration of the transcript is 2 mM. Spectra recorded at 500 MHz; ppm references to H_2O relative to DSS. The 2-1-4 pulse was used to suppress excitation of the H_2O resonance.

B), and ended in 100% buffer B. The flow rate was 2.0 mL/min at 180–200 psi. A small aliquot of the collected RNA was removed for analysis by gel electrophoresis; the bulk of the RNA was precipitated with 2.5 volumes of ethanol overnight at −20 °C, centrifuged, washed with 70% ethanol, and resuspended in 350 μL of 50 mM EDTA, pH 8.0.

The RNA was dialyzed against two changes of 400 mL of 0.1 M EDTA, pH 8, to remove divalent cations and then twice against 450 mL of 10 mM sodium phosphate/1 mM EDTA, pH 7. An aliquot was removed for melting studies, and the rest was dried down and stored at 4 °C. The purified RNA was analyzed on a 10% polyacrylamide/8 M urea gel and stained with Stainsall (Sigma). Even when the gel was overloaded, there was no evidence of short abortive products contaminating the transcript which had been purified by chromatography and dialysis.

NMR Spectroscopy. The lyophilized tRNAs were first dissolved in H_2O for a final sodium phosphate concentration of 10 mM, followed by the addition of NaCl to 0.1 M and ^2H_2O to 5–10%. $MgCl_2$ was added in aliquots as necessary. For subsequent dialysis against 5 mM $MgCl_2$, the transcript was first dialyzed against two changes of 500 mL of 0.1 M EDTA, pH 8.0, to remove added $MgCl_2$ and then against 1 L of 10 mM sodium cacodylate, pH 7.0, 0.1 M NaCl, and 5 mM $MgCl_2$ followed by two changes of 500 mL. The final volume was adjusted to give 2 mM tRNA.

All spectra were taken on the 500-MHz NMR spectrometer at Brandeis University as described previously (Redfield, 1986). Data were referenced to the water signal assumed to be 4.8 ppm. The 2-1-4 pulse was used to observe the exchangeable imino protons in H_2O. Most spectra were taken with 8-KHz spectral width with offset in the imino region. NOEs were recorded as difference spectra, where the on-resonance spectra was subtracted from the interleaved off-resonance spectrum. For NOESY spectra, the 2-1-4 pulse was used to give all the 90° pulses; the mixing time was varied. The recycle time for both 1D and 2D experiments was usually 400 ms. Moderate weighting was applied to the free induction decay to improve resolution and diminish base-line curvature.

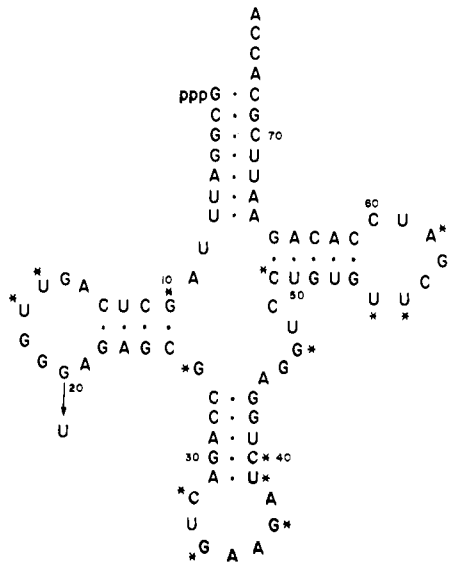

FIGURE 2: Cloverleaf structure of the yeast tRNA^{Phe} transcript. The positions of the modified bases in yeast tRNA^{Phe} are indicated by asterisks. Also shown is the sequence of the U20 variant transcript.

RESULTS

NMR Spectrum of the tRNA Transcript. Because yeast tRNA^{Phe} has been extensively studied by NMR [reviewed by Reid (1981)] and all the downfield imino resonances are assigned (Roy & Redfield, 1982; Heerschap et al., 1983b), the spectrum of the transcript was first compared to that of modified tRNA^{Phe} to facilitate assignment. The downfield spectra of the transcript and of modified yeast tRNA^{Phe} in 5 mM free Mg^{2+} are shown in Figure 1. While the spectra are similar, there is a loss of chemical shift dispersion in the spectrum of the transcript, with many overlapping resonances in the GC region of the spectrum. The differences are substantial enough to make it impossible to assign resonances by simple comparisons, so that it was necessary to assign the transcript resonances de novo.

Assignment of resonances in a molecule of this size is inherently difficult. The task is made more difficult in the present case by the lack of methylated residues which served as spectroscopic markers for the sequential assignment of the modified yeast tRNA^{Phe}. The only unambiguous resonances in the spectrum of the transcript are from the sole GU pair in the acceptor stem (see Figure 2), characterized by strong (30%) NOEs between the two downfield resonances of the two exchangeable imino protons. More ambiguous spectroscopic markers are provided by the relatively large number of AU pair resonances, identified both by chemical shift and by the sharp NOEs to the aromatic region resulting from the proximal uracil N3 and adenine C2 protons. These can be confused with NOEs of reverse-Hoogsteen AU pairs, which differ only by having a more downfield aromatic NOE.

We have used a combination of 1D NOE and 2D NOESY techniques to assign the spectrum of the transcript in buffers containing 10 mM sodium cacodylate and either 5 mM free $MgCl_2$ or 50 mM total $MgCl_2$. The NMR spectrum of the transcript in the latter solution is nearly identical with its spectrum in 5 mM free Mg^{2+}, except that an AU imino proton resonance at 13.6 ppm in 50 mM $MgCl_2$ (peak d in Figure 4b) shifts to 13.8 ppm in 5 mM free Mg^{2+} (peak c2 in Figure 4a). The other major difference is the loss of intensity in the region between 10.6 and 11.6 ppm when the transcript is dialyzed against Mg^{2+}. Since no NOEs are observed to or from this region, its origin is unclear. The assignments for each imino proton resonance are summarized in Table I.

Table I: NOEs and Connectivities Observed from Imino Proton Resonances in 50 mM $MgCl_2$[a]

chemical shift (ppm)[b]	peak[c]	NOEs observed to: imino	NOEs observed to: aromatic[d]	assignment
12.5[e]	i	k,c2,e,f,d		C2G71
12.4[e]	j/k	l,q,i,c2		G3C70
10.46	q	l,k,c1		G4U69
11.86	l	c1,q,k		G4U69
13.98	c1	a1,l,q	7.98	A5U68
14.5	a1	e,c1	7.89	U6A67
13.31[e]	e	i,a1	7.3,7.01	U7A66
12.54	i	c2,e,f,k,d		C49G65
13.8[e]	c2	i,j,g,h	7.4	U50A64
12.4	j/k	c2,i		G51C63
13.3	e	i,j	7.0	U52A62
12.54	i	e,c2,f,k,d		G53C61
12.24	k′	f,a2		C48G15
14.37	a2	g,k′	8.32	U8A14
12.98	g	c2,a2		G22C13
13.8	c2	h,g,i,j	7.56[f]	A23U12
12.88	h	c2,i		G24C11(?)
12.54	i			C25G10(?)
9.3		m,d		U55
11.76	m	d,i		U55G18
13.6[g]	d	i,m[h]	8.28	T54A58
12.54	i	f,c2,e,k,d		C28G42
13.22	f	i,k′	7.01	A29U41
12.24	k′	f,a2		G30C40

[a] Gaps in the columns indicate the absence of further NOE connectivities from the adjacent resonances. [b] Chemical shifts are relative to DSS. [c] See Figure 7 for assigned spectrum. [d] NOEs to aromatic regions identify AU resonances. [e] Several peaks contain more than one resonance. See Figure 7 for the assigned spectrum. [f] Other aromatic NOEs observed to 8.86 and 9.3 ppm. See text for details. [g] Peak d moves to 13.8 ppm in 5 mM free Mg^{2+}. [h] Observed at or below 25 °C.

While some regions of the molecule appear reliably assigned, others are tentative, due in part to the number of overlapping resonances within single peaks or to the absence of a predicted NOE. Table I is arranged not only to present the definite NOEs observed but also to indicate how the NOEs appear to map the resonances onto the sequence of base pairs in the four helical stems. The details of this mapping and the degree of confidence that can be placed on each assignment will be discussed below. The minimum requirement of such mapping is that each entry in column three of Table I contain the resonances above and/or below it on the list. A result of this method of presentation is that several resonances are repeated on different lines of the table. The implicit assumption in commencing the sequential assignments is that the helical stems are intact in the $tRNA^{Phe}$ transcript. While the absence of anticipated NOEs is not a test of this assumption, the presence of NOE patterns not anticipated from the structure would indicate that the starting assumption is invalid.

Acceptor Stem. In modified yeast $tRNA^{Phe}$, the two resonances from the imino protons of the G4U69 base pair (see Figure 2) are found at 11.75 and 10.36 ppm in buffer with 5 mM free Mg^{2+} (Johnston & Redfield, 1978, 1981). In the spectrum of the transcript, peaks l and q are found at similar chemical shifts (11.86 and 10.46 ppm, respectively) and show strong (30%) mutual NOEs. Thus, peaks l and q were confidently assigned to the imino protons of G4U69 in the acceptor stem of the transcript. Using NOEs from these resonances, it was possible to extend the assignments of the acceptor stem. From peak q, there are additional NOEs to k and to c1, and to an upfield aromatic proton at 7.98 ppm which is indicative of an AH2 proton. Peak k is therefore assigned to G3C70, and peak c1 to A5U68, on the basis of their respective chemical shifts. As Table I indicates, peak c1 shows an NOE back to peaks l and q, as well as an NOE to peak a1, indicating that peak a1 should be assigned to U6A67. This resonance in turn shows a reciprocal NOE to peak c1 and other NOEs to peak e and to an aromatic proton at 7.89 ppm. Peak e, which shows NOEs to peak a1 and to several aromatic protons, can thus be assigned to the last AU of the acceptor stem, U7A66.

T-Loop and T-Stem. Several tertiary base pairs in the T-loop are expected to give distinct spectroscopic signals. Of these, a reverse-Hoogsteen base pair U54A58 in the transcript might be identified by its unique NOE pattern. As given in Table I, peak d shows NOEs to the aromatic region at 8.28 ppm, to peak i, and (at 25 °C) to peak m. The relatively downfield position of the 8.28 ppm aromatic NOE from peak d is characteristic of reverse-Hoogsteen pairs (Sanchez et al., 1984) and in turn suggests that peak d is U54A58. According to the crystal structure, the T54m^{1}A58 tertiary base pair should be flanked by G53C61 which is the terminal base pair of the T-stem, and another tertiary base pair from the T-loop, G18Ψ55. Of the two NOEs from peak d, peak i is assigned to G53C61 by virtue of its chemical shift. Since the chemical shift of peak m (11.76 ppm) is quite upfield for a GC pair, it is tentatively assigned to one of the imino protons of G18U55.

In modified $tRNA^{Phe}$, there are three imino proton resonances assigned to the G18Ψ55 tertiary base pair, 11.57 ppm (Ψ55 NH3), 10.36 ppm (Ψ55 NH1), and 9.8 ppm (G18 NH1) (Tropp & Redfield, 1981; Heerschap et al., 1982, 1983a,b), with NOEs observed from both the G18 NH1 and Ψ55 NH3 proton resonances to the adjacent T54 NH3 resonance, but not to or from that of Ψ55 NH1. In the transcript, which contains U instead of Ψ, peak m has been assigned to G18U55. NOEs from m to 8.28 ppm and to peak d, which are assigned to the U54 NH3 and A58 H8 protons, respectively, suggest that peak m corresponds to the NH1 proton of G18. The imino proton of U55 could correspond to a resonance at 9.3 ppm, which appears as a weak NOE from peak m, but there is no NOE between the resonance at 9.3 ppm and the U54A58 proton resonance. The weak NOE observed between m and d suggests either that the base pairs are quite separated or that the structure is flexible and transfer is not efficient. The temperature dependence of the intensity and chemical shift of peaks d and m, and the relatively low intensity of peak d, also suggest that this is an unstable or flexible region. In addition, peak d shifts from 13.6 ppm in 50 mM $MgCl_2$ to 13.8 ppm in 5 mM Mg^{2+}, indicating that this portion of the tRNA is very sensitive to Mg^{2+} ions. The shift to 13.8 ppm puts the resonance among other AU base pairs, but its presence is discernible by a new NOE to 8.32 ppm that is observed when the 13.8 ppm peak is irradiated.

Continuing sequential assignments by NOEs from the T-loop or acceptor stem resonances into the T-stem was not possible, as both terminal CG pairs of the T-stem are likely to lie in peak i which contains several GC resonances. C49G65 (assigned to peak i) at the T-stem/acceptor stem junction is flanked by two AU base pairs, one of which, U7A66, has already been assigned. To continue T-stem assignments, we looked for NOEs from other AU imino protons to peak i in order to identify the sequence U7A66-C49G65-U50A64: such NOEs were observed from peaks c2, e, and f. NOESY spectra (not shown) show at least three AU resonances in peak e/f distinguished by their aromatic proton resonances, two of which, e1 and f, show aromatic NOEs at 7.0 ppm at 30 °C. In agreement with NOE experiments, cross-peaks are observed in NOESY experiments from e2-to-a1, e2-to-i, f-to-k′, and f-to-i, but there is regrettably no cross-peak from e1, using

Table II: Chemical Shifts[a] of Modified and Unmodified tRNA^{Phe}

base pair	peak	unmodified[b]	modified[c]
acceptor stem			
G1C72			12.85
C2G71	i	12.46	12.5
G3C70	k	12.32	12.25
G4U69	l	11.82	11.75
	q	10.42	10.36
A5U68	c1	13.94	13.86
U6A67	a1	14.46	14.39
U7A66	e	13.25	13.25
T-stem and loop			
C49G65	i	12.54	12.5
U50A64	c2	13.8	13.19
G51C63	j	12.44	12.35
U52A62	e	13.3	13.8
G53C61	i	12.5	12.46
U54A58	d	13.6	12.46
U55G18	m	11.8	11.57
			10.47
		9.3	9.8
C56G19			12.46
D-stem			
C48G15	k′	12.24	12.1
U8A14	a2	14.37	14.39
G22C13	g	12.98	12.91
G46			13.3
A23U12	c2	13.78	13.79
G24C11	h	12.88	13.63
C25G10	i	12.56	12.72
anticodon stem			
G26A44			10.36
C27G43			12.53
C28G42	i	12.5	12.1
A29U41	f	13.2	13.2
G30C40	k′	12.2	12.53
A31U39			13.19
			10.54

[a] Chemical shifts are at 24 °C, relative to DSS. [b] 50 mM $MgCl_2$. [c] See Heerschap et al. (1983b) for a summary of the assignments.

mixing times from 50 to 120 ms (data not shown). Peak e2 has already been assigned to the terminal U7A66 of the acceptor stem. Irradiating the composite peaks i and j/k at several frequencies partially separates the contributing resonances, although the interpretation must be considered tentative, given the number of GC imino proton resonances found at this frequency. By selectively irradiating at 30-Hz increments through peaks i, j, and k, it is possible to observe NOEs to peaks c2 and e2 from a single frequency in i, another to peak f, to peaks c2 and e1 from j, and to peaks l and q from k. Peak k′ is more resolved from the larger peak and shows a k′-to-f NOE. This pattern allows the assignment of the T-stem as shown in Table I, with c2 assigned to U50A64, j to G51C63, and e1 to U52A62. The chemical shifts of the two AU pairs do not correspond well with those in the T-stem of the modified tRNA (see Table II).

D-Stem. Near the inner end of the D-stem in the modified tRNA, there is a tertiary reverse-Hoogsteen AU base pair, U8A14, which is identifiable in the NMR spectrum of the transcript by its NOE pattern. Peak a2 shows an NOE to a downfield aromatic resonance at 8.32 ppm, characteristic of an AH8 proton, leading to its assignment as U8A14. Table I summarizes the network of connectivities involving peak a2 and indicates the identification of the sequence k′-a2-g-c2-h as the D-stem, with a2 corresponding to U8A14 and g to C13G22. Peak k′ is assigned to another tertiary base pair, G15C48, at the end of the D-stem. These assignments agree well with those for the modified tRNA (see Table II).

In the crystal structure, two base pairs of the D-stem participate in two base triples. One of these, the base triple A9-A23U12, is distinguished by the stable amino protons of

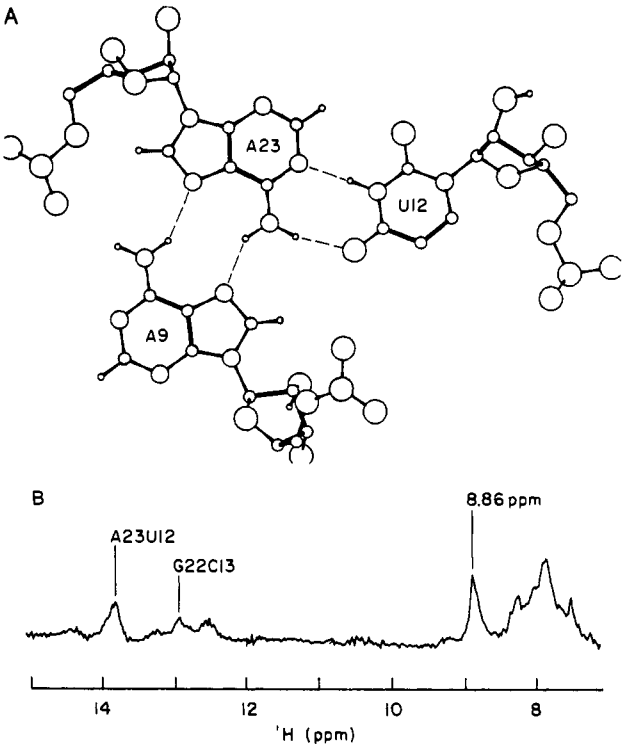

FIGURE 3: (A) Structure of the base triple A9-A23U12 from the crystal structure, showing the hydrogen-bonding pattern expected from the A23 amino protons to the adjacent bases. (b) NOEs seen from the resonance at 8.86 ppm assigned to the amino proton of A23. NOEs to peaks c2 (A23U12) and g (G22C13) are indicated.

A23 which are hydrogen-bonded to the carbonyl on U12 and the nitrogen on A9, diagrammed in Figure 3. Although normally the amino protons are obscured by overlap with other protons, and are not seen by NOEs, this group should be visible since both hydrogen-bonded protons should be relatively solvent-inaccessible. In fact, in the transcript, there are NOEs from c2 and g to a broad but strong peak at 8.86 ppm, and the reverse NOEs to c2 and g are also observed (Figure 3). This peak is thus tentatively assigned to an A23 amino proton. Two additional NOEs are observed upon irradiation at 8.86 ppm. One of these, at 8.08 ppm, is not observed when any AU imino proton resonance is irradiated and is therefore assigned to the H8 proton of A9, as this proton is close to the A23 amino group. The other resonance, at 9.2 ppm, is also observed upon irradiation of c2 and g. It is possible that this is the second proton of the A23 amino group. These chemical shifts differ from those of Choi and Redfield (1985), who found the A23 amino group at 9.8 ppm and the A9H8 at 8.4 ppm in the modified tRNA.

The A23 amino group is expected to be close to both adjacent base pairs C13G22 and C11G24, and should therefore show NOEs to the imino protons of both base pairs. An NOE is observed from 8.86 ppm to peak g which was assigned to C13G22. The assignment of C11G24 is not clear, however, as there are NOEs to other peaks, including i and a weak NOE to peak h. On the basis of the 8.86 ppm-to-h NOE and the imino proton NOE from c2, peak h is tentatively assigned to C11G24. In modified tRNA^{Phe}, C11G24 is found at 13.63 ppm, which is downfield for a GC pair, and is 0.8 ppm downfield from our assignment of the corresponding pair in the transcript. This shift could be due to the absence of the methyl group normally found on the adjacent m^2G10C25 (see Figure 2) as the modified base will influence the chemical shift of both C11G24 and G10C25. The G10 methyl modification was one of the spectroscopic markers that allowed the NMR

assignment of these last base pairs of the D-stem in the modified tRNA. The absence of methyl groups and the greater spectral overlap in the CG region of the transcript make the data for these last assignments ambiguous.

Anticodon Stem. The assignments of the resonances of the anticodon stem must be considered the most tentative. As shown in Table I, there are i-to-f and f-to-k′ NOEs, showing that there are GCs (i and k′) flanking this AU (f). A29U41 is assigned to peak f. The resonance k′ could correspond either to C28G42 or to G30C40; it is assigned to G30C40 because the chemical shift of peak k′ and its relative sensitivity to temperature suggest that its environment is different than many of the GC pairs in the molecule. In the anticodon stem, the C28G42 base pair, flanked by an AU and another GC pair, is in the middle of a helix, whereas G30C40 is flanked by two AU pairs, one of which, the terminal A31U39, will probably be unstable in the transcript. In the structure of the modified tRNA, in the presence of Mg^{2+}, both imino protons from Ψ39 are observed, showing that neither is freely exchanging with solvent. The substitution of a U for the Ψ and the absence of modified bases in the loop will change the properties of the anticodon stem/loop and may destabilize the end of the helix.

The tertiary G26A44 base pair and the adjacent C27G43 are unassigned. In the modified tRNA, the m_2^2G26A44 imino proton resonance is found at 10.38 ppm and was identified in part by its NOE to the A44H2 proton, and in part by use of the methyl groups. It is not found there in the transcript. The frequency of the G26A44 imino proton resonance in the transcript should be shifted from its position in the modified tRNA, however, since the absence of the m_2^2G26 and adjacent m^2G10 are expected to cause structural changes in the junction between the D-stem and anticodon stem of the transcript. Leontis and Moore (1986) found a GA base pair in one of the 5S RNA helices at 12.5 ppm. In the case of the transcript, the chemical shift of the G26A44 proton might be expected to be similar to the 5S GA resonance, which would place it in the midst of many GC resonances in the spectrum where it would be impossible to identify.

Magnesium Dependence of the Downfield Spectrum. There are several specific binding sites for Mg^{2+} ions in the tRNA structure (Holbrook et al., 1977; Jack et al., 1977; Quigley et al., 1978) and probably more nonspecific sites. The NMR spectrum of modified $tRNA^{Phe}$ in 0.1 M NaCl shows several changes when Mg^{2+} is added; the largest changes are the 0.3 ppm downfield shift of the U8A14 resonance and the 0.5 ppm upfield shift of the Ψ55 NH1, but the chemical shift differences are otherwise not pronounced (Heerschap et al., 1983b).

The NMR spectra of the transcript as a function of Mg^{2+} concentration are shown in Figure 4. In contrast to the case for the modified $tRNA^{Phe}$, it is obvious that Mg^{2+} induces significant changes in the spectrum of the transcript. For example, in the absence of Mg^{2+} or at low $MgCl_2$ concentrations, the chemical shift dispersion is worse, resonances are broad, and many peaks present in the spectrum in 50 mM $MgCl_2$ have been lost. As the temperature is increased to 40 °C (data not shown), there is considerable sharpening of most resonances, due to decreased viscosity of the solution and also in part to the shift and loss of resonances. In particular, peak e (see Figure 4c) is lost at 40 °C, while a resonance in c disappears between 40 and 50 °C. The resonances in peaks i, j, and k also shift to become more resolved at 40 °C.

Due to these spectral changes of the transcript in the absence of Mg^{2+}, it was necessary to reassign the resonances (see Table III). Again, the imino resonances of G4U69 in the acceptor stem were identified, and the other acceptor stem assignments

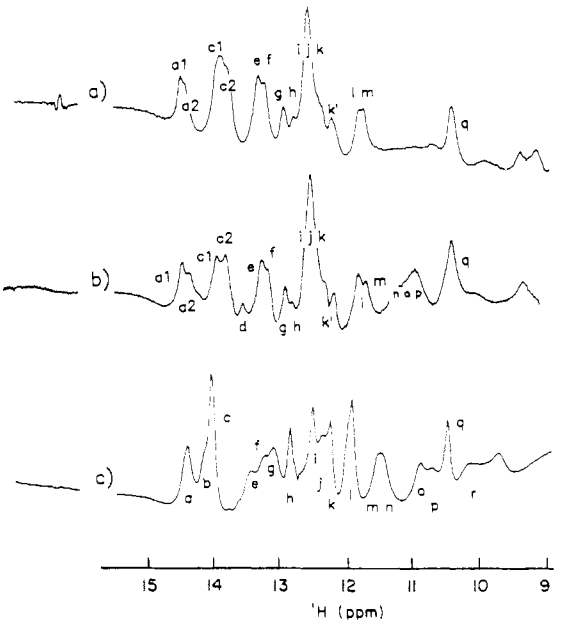

FIGURE 4: ^{1}H NMR spectra of the imino region in H_2O of the transcript in (a) 5 mM free Mg^{2+} (dialyzed), (b) 50 mM total $MgCl_2$, and (c) 0–10 mM total $MgCl_2$ at 30 °C.

Table III: NOEs[a] Observed in the Transcript in the Absence of Mg

resonance	peak[b]	NOEs observed: imino	NOEs observed: aromatic	assignment[c]
14.36	a	f,i,m,l	7.29	
14.17	b	j	7.76	
14.07	c	l,q,k,e	7.79, 8.0	A5U68
13.47	e	c,m/n,i	7.33	U6A67/U7A65
13.26	f	a,i	7.79	
13.10	g	l,m,o,i	7.5	
12.85	h	i,k,l		G2G71
12.53	i	h,a		G1C72
12.83	j	b		
12.27	k	q,l,h,c		G3C70
11.95	l	q,m,k,g,c		G4U69
11.56	m/n	l,g,a,o		
11.47	n	f,p		
10.91	o	g,m		
10.76	p	l,n		
10.51	q	l,k,c		G4U69
10.16	r	i		

[a] NOEs observed at 30 °C. [b] The letters given to the resonances in this spectrum do not correspond to those in the spectrum of the transcript in high Mg^{2+} conditions. [c] See Figure 6 for labeled spectrum.

followed. When the intensity of the NOEs resulting from a long irradiation of peak q (G4) is used as an indication of their relative distance from the GU base pair, peaks i, h, k, l and q, and c are assigned to G1C72, C2G71, G3C70, G4U69, and A5U68, respectively, of the acceptor stem. Irradiating peak c did give an NOE to peak e at 13.3 ppm at 22 °C, where the reverse e-to-c NOE and a sharp NOE from peak e to 7.26 ppm (7.33 at 30 °C) were also present. Therefore, this seemed to be an AU/AU NOE, but because it did not appear as a secondary NOE from peak c during long irradiations of peak q, and because NOEs to the aromatic region indicated that this peak consisted of several resonances, it was not clear which AU/AU sequence it represented. The AU pairs at the end of the stem appear to be labile, a feature shared by the modified tRNA in similar buffer conditions (Heerschap et al., 1982; Roy & Redfield, 1983).

While the acceptor stem appears to be intact, the lack of extended sequential imino proton NOEs suggests that other sequences of the molecule are adopting structures that are not as stable. Only seven AU base pairs are identified in the

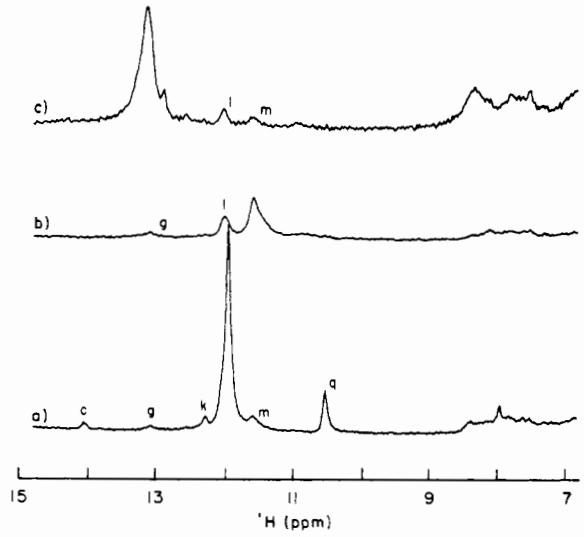

FIGURE 5: NOE patterns of the second GU base pair in the transcript in the absence of $MgCl_2$. (a) Irradiating peak l (11.98 ppm) shows NOEs to resonances assigned to the acceptor stem: peak q at 10.5 ppm assigned to G4U69, to peak c at 14.07 ppm assigned to A5U68, and to peak k at 12.27 assigned to G3C70. NOEs to several other resonances: to peak m at 11.56 ppm and to peak g at 13.10 ppm. (b) Irradiating peak m gives NOEs to peak l and to peak g. (c) Irradiating peak g gives NOEs to several peaks, among them m and l. The m to l NOE is real, not the result of spillover from the irradiation.

transcript in the absence of Mg^{2+}, where there should be eight Watson–Crick and two reverse-Hoogsteen base pairs. It is possible that the transcript is in a semidenatured state, where most of the acceptor stem is stable but other stems and tertiary interactions are not.

A Second GU Pair. Peak l gives a strong NOE to q (G4U69), as discussed above, as well as weaker NOEs to peaks k, c, and g. However, it also gives a strong NOE to peak m/n (15%) (Figure 5a). Irradiating peak m/n gives a strong NOE back to l (Figure 5b), other NOEs to g, o/p, and f, and a weaker NOE to a. The strong NOE from l-to-m is not the result of spillover; this was confirmed by the NOESY spectrum of the transcript which clearly shows that there is a strong cross-peak between these two resonances (data not shown). The magnitude of the m-to-l NOE and their chemical shifts suggest that it is a GU base pair.

The location of this extra GU base pair in the structure of the transcript is not clear. The modified tRNA^{Phe} structure contains only one Watson–Crick GU base pair, in the acceptor stem (G4U69, see Figure 2), but has a second, parallel tertiary GU base pair formed by G18Ψ55. It is possible that the apparent second GU base pair in the unmodified tRNA structure may be this tertiary G18U55 pair, with its conformation altered from that in the modified tRNA to now allow an NOE to be observed between the two imino protons. The observation of reciprocal NOEs between peaks m, l, and g, and the appearance of an NOE from g to 7.5 ppm, suggests that this putative GU pair is flanked by an AU pair. This AU might correspond to the tertiary base pair T54A58 normally found adjacent to the G18Ψ55 in the tRNA^{Phe} structure.

Alternatively, the extra GU pair could come from a rearrangement of the stem structures. It is possible that there are two resonances in peak g, one corresponding to an AU base pair with an aromatic AH2 proton at 7.5 ppm and another to a GC base pair. The absence of an NOE to 7.5 ppm upon irradiation of peak mn suggests that the GU pair at m,l would be flanked by a GC pair. The NOEs from peak o to peaks m and g are puzzling, as resonances found at the chemical shifts of peaks o and p are often ascribed to non-hydrogen-bonded but relatively solvent-inaccessible imino protons (Heerschap et al., 1983b). The structural feature of the transcript that gives rise to these resonances is unknown.

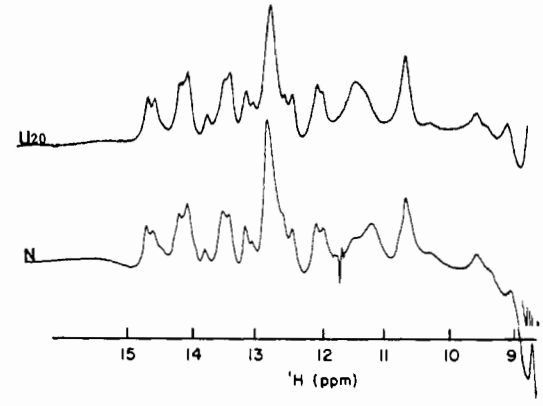

FIGURE 6: Comparison of the ^1H NMR spectra in H_2O of the imino protons of the U20 variant and the normal sequence (N) in 50 mM $MgCl_2$ at 30 °C. The difference in intensity in the two spectra in the region between 10.6 and 11.6 ppm is due in part to base-line roll. No NOEs are observed to or from this region in either molecule.

The apparent conformational transition of the transcript as a function of Mg^{2+} concentration determined from both the imino proton chemical shifts and the NOEs observed illustrates a structural flexibility or instability that was unexpected. The presence of an extra GU pair suggests that there is a structural rearrangement rather than a destabilized cloverleaf conformation. This is in contrast to the structure of the modified tRNA^{Phe} and could account for the Mg^{2+} concentration dependence of the aminoacylation activity of the transcript (Sampson & Uhlenbeck, 1988).

Structure of the U20 Variant. The spectrum of the normal transcript was compared to that of a transcript containing a U20 in place of the G20. This variant was of interest because its rate of aminoacylation was decreased 14-fold from that of the normal transcript (Sampson & Uhlenbeck, 1988). The G at this position is present only in tRNA^{Phe} (Sprinzl & Gauss, 1984), which suggests that it is a site-specific contact point for the synthetase. Alternatively, it was possible that the substitution created some structural change that could be responsible for the decrease in activity of the variant, even though the crystal structure predicted that this substitution should be isomorphic. In the crystal, G20 extends out into solution and does not participate in interactions with other bases in the tRNA. NMR provides a means to determine if any structural perturbations have in fact occurred.

The NMR spectrum of the G20-to-U20 transcript shown in Figure 6 is virtually identical with the normal tRNA spectrum both with and without $MgCl_2$. The NOE patterns observed in the normal spectra are also present in the spectra of this variant, including the extra GU base pair found in semidenaturing conditions. From the similarity of the spectra, and implied structural similarity of the transcripts, we conclude that a major structural perturbation is not the cause of the low aminoacylation activity of this variant. Rather, the lack of large structural changes and the unique appearance of a guanosine at position 20 in yeast tRNA^{Phe} suggest that G20 is an essential contact point for the synthetase.

Discussion

We have used NMR to study unmodified yeast tRNA^{Phe} and to understand how modifications contribute to structure. The NMR data show that in the presence of high concentrations of $MgCl_2$ (50 mM total of 5 mM free), the structure

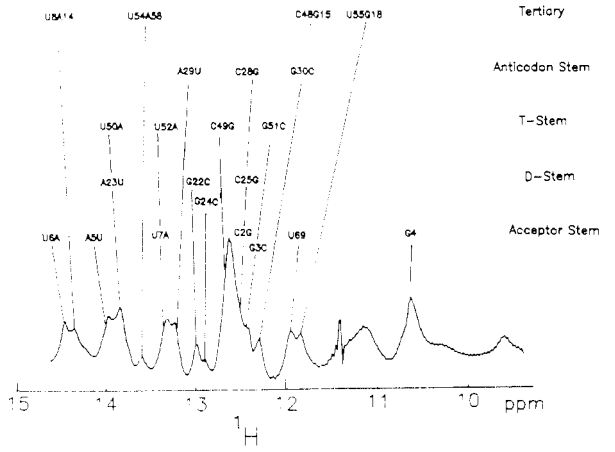

FIGURE 7: Imino proton spectrum of the transcript with the assignments indicated.

of the transcript is similar to that of the modified $tRNA^{Phe}$. In the spectrum of the transcript in 50 mM $MgCl_2$, we have assigned most of the base pairs in the stems and five of the tertiary interactions, including the base triple A9A23U12. Figure 7 summarizes the assignments, and Table II compares the chemical shifts of the resonances of the transcript with those of the modified tRNA. There is excellent agreement in the chemical shifts of the base pairs assigned to the acceptor stem of the two molecules, indicating that the RNA geometry of this helix is nearly identical. This is significant, for this stem alone contains no modified bases.

However, there are several spectral features of the transcript that differ from those of the modified $tRNA^{Phe}$, and may be correlated with the proximity of modifications. In particular, the assignments of the protons of the U54A58 base pair and of the adjacent G18U55 are the most different for the modified and unmodified tRNA. In the modified tRNA, the m^1A58T54 imino proton resonance is found at 12.4 ppm (Roy et al., 1984), far upfield for an AU. Its position in the transcript, at 13.6 ppm (or 13.8 ppm in 5 mM free Mg^{2+}), is typical of an AU imino proton and may be a result of the lack of modifications on the base pair itself and on the adjacent G18Ψ55. This difference could be in part a ring current effect, but it is more likely that the modifications contribute to the structure of this region either through hydrogen bonding with solvent or through counterion interactions, as the positive charge on the m^1A58 (Agris et al., 1986) would create a peculiar electrostatic environment. Similarly, the absence of modifications at G10 and G26 could also explain the C11G24 chemical shift differences (see Table II) In this case, the methylations of these residues will change the structure of the junction between the D-stem and the anticodon stem. In the transcript, the chemical shift of C11G24 is more usual for a GC pair, indicating that its environment is more standard. A similar difference in chemical shifts would be expected between m_2^2G26A44 and the unmodified G26A44 in the transcript, but we have not been able to assign the G26A44 imino proton. These chemical shift changes may illustrate a general property of small structural changes caused by base modifications. Even though the transcript appears to be folding normally, the overlap in the spectrum is greater than it is for the modified tRNA. The local structural perturbations may produce distinct magnetic environments which result in greater chemical shift dispersion in the NMR spectrum; the absence of the modifications may result in more regular helices with more equivalent chemical and magnetic environments.

The conformational transition of the transcript that we observe is dependent on the Mg^{2+} concentration. No corresponding transition has been observed for the modified yeast $tRNA^{Phe}$ either by NMR (Heerschap et al., 1983a) or by nuclease mapping (Wrede et al., 1979) under conditions in which pH, ionic strength, and Mg^{2+} concentration were varied. In contrast, nuclease mapping of *Escherichia coli* $tRNA^{Glu_2}$, which contains only five modified bases, indicated that it adopted two conformations as a function of the Mg^{2+} concentration (Eisinger & Gross, 1975; Wrede et al., 1979). These results show that Mg^{2+} can stabilize a particular conformation and imply that modified bases can do the same, although presumably by a different mechanism. In the case of yeast $tRNA^{Phe}$, the modified bases are found in the junctions between stems and in the loops, where it is possible that they stabilize stems and tertiary interactions by increasing stacking interactions. The mechanism of structural stabilization by Mg^{2+} is also unknown, although the crystal structure identifies four specific binding sites (Holbrook et al., 1977; Jack et al., 1977; Quigley et al., 1978). It is interesting to note that, with the exception of the anticodon loop, none of these Mg^{2+} ions is coordinated to a modified base. Additional evidence for a separation of effects comes from measurement of the melting temperature of the two yeast $tRNA^{Phe}$ molecules. Under conditions where the $MgCl_2$ concentration would exceed the NMR conditions of 5 mM free Mg^{2+}, the melting temperature of the transcript is still 5 °C lower than that of the fully modified $tRNA^{Phe}$ (Sampson & Uhlenbeck, 1988).

The presence of an extra GU pair in the low-Mg^{2+} form of the transcript suggests that this structure is an alternative conformation that is not accessible to the fully modified molecule. The temperature stability of this structure, and its persistence in the presence of low concentrations of Mg^{2+}, indicates that it is a stable conformation. As such, it may not be recognizable by the aminoacyl-tRNA synthetase, which would explain the low aminoacylation activity of the transcript at low Mg^{2+} concentrations.

The comparison of the structure of the U20 variant and the normal tRNA transcript indicates how NMR of these large molecules can be used to relate structure to function. This variant, by NMR criteria of both chemical shifts and NOE patterns, adopts the structure of the normal sequence. Thus, a structural change caused by the G20 to U20 substitution cannot be the cause of the low aminoacylation activity of this variant, but rather the isomorphic nature of the substitution indicates that the normal G20 is a specific contact for the synthetase. This kind of NMR comparison has great potential for the determination of structural variations of site-specific mutations.

REFERENCES

Agris, P. F., Sierzputowska-Gracz, H., & Smith, C. (1986) *Biochemistry 25*, 5126–5131.

Choi, B. S., & Redfield, A. G. (1985) *Nucleic Acids Res. 13*, 5249–5254.

Eisinger, J., & Gross, N. (1975) *Biochemistry 14*, 4031–4041.

Griffey, R. H., Poulter, C. D., Yamaizumi, Z., Nishimura, S., & Hurd, R. E. (1982) *J. Am. Chem. Soc. 104*, 5811–5813.

Heerschap, A., Haasnoot, C. A. G., & Hilbers, C. W. (1982) *Nucleic Acids Res. 10*, 6981–7000.

Heerschap, A., Haasnoot, C. A. G., & Hilbers, C. W. (1983a) *Nucleic Acids Res. 11*, 4483–4499.

Heerschap, A., Haasnoot, C. A. G., & Hilbers, C. W. (1983b) *Nucleic Acids Res. 11*, 4501–4520.

Hilbers, C. W., Heerschap, A., Haasnoot, C. A. G., & Walters, J. A. L. I. (1983) *J. Biomol. Struct. Dyn. 1*, 183–207.

Holbrook, S. R., Sussman, J. L., Warrant, R. W., Church, G. M., & Kim, S. H. (1977) *Nucleic Acids Res. 4*, 2811–2820.
Jack, A., Ladner, J. E., & Klug, A. (1976) *J. Mol. Biol. 108*, 619–649.
Jack, A., Ladner, J. E., Rhodes, D., Brown, R. S., & Klug, A. (1977) *J. Mol. Biol. 111*, 315–328.
Johnston, P. D., & Redfield, A. G. (1978) *Nucleic Acids Res. 5*, 3913–3927.
Johnston, P. D., & Redfield, A. G. (1981) *Biochemistry 20*, 1147–1156.
Kearns, D. R., & Shulman, R. G. (1974) *Acc. Chem. Res. 7*, 33–39.
Leontis, N. B., & Moore, P. B. (1986) *Biochemistry 25*, 3916–3925.
Noggle, J. H., & Shirmer, R. E. (1971) *The Nuclear Overhauser Effect*, Academic Press, New York.
Quigley, G. L., Wang, A. H. J., Seeman, N. C., Suddath, F. L., Rich, A., Sussman, J. L., & Kim, S. H. (1975) *Proc. Natl. Acad. Sci. U.S.A. 72*, 4866–4870.
Quigley, G., Teeter, M., & Rich, A. (1978) *Proc. Natl. Acad. Sci. U.S.A. 75*, 64–68.
Redfield, A. G. (1986) *NATO ASI Ser. 107*, 1.
Reid, B. R. (1981) *Annu. Rev. Biochem. 50*, 969–996.
Rich, A., & RajBhandary, U. L. (1976) *Annu. Rev. Biochem. 45*, 805–860.
Roy, S., & Redfield, A. G. (1983) *Biochemistry 22*, 1386–1390.
Roy, S., Papastravros, M. Z., & Redfield, A. G. (1982) *Nucleic Acids Res. 10*, 8341–8349.
Roy, S., Papastravros, M. Z., Sanchez, V., & Redfield, A. G. (1984) *Biochemistry 23*, 4395–4400.
Sampson, J. R., & Uhlenbeck, O. C. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 1033–1037.
Sanchez, V., Redfield, A. G., Johnson, P. D., & Tropp, J. (1980) *Proc. Natl. Acad. Sci. U.S.A. 77*, 5959–5662.
Sprinzl, M., & Gauss, D. (1984) *Nucleic Acids Res. 12*, r1–r131.
Sundaralingam, M. (1978) *Acta Crystallogr. B34*, 1529.
Sussman, J. L., Holbrook, S. R., Wade-Warrant, R., Church, G. M., & Kim, S. H. (1978) *J. Mol. Biol. 123*, 607–630.
Tropp, J., & Redfield, A. G. (1981) *Biochemistry 20*, 2133–2140.
Wrede, P., Wurst, R., Vournakis, J., & Rich, A. (1979) *J. Biol. Chem. 254*, 9608–9616.

Secondary Structure of DNA Modified by Monofunctional Psoralen Derivatives†

F. Gaboriau,*,‡ P. Vigny,‡ and J. Moron§

Laboratoire de Physique et Chimie Biomoléculaire (CNRS UA 198), Section de Physique et Chimie, Institut Curie, and Université Paris VI, 11 rue P. et M. Curie, 75231 Paris Cedex 05, France, and Laboratoire de Synthèse Organique (CNRS UA 533), Section de Biologie, Institut Curie, Bat. 110, 91045 Orsay Cedex, France

Received June 29, 1988; Revised Manuscript Received February 16, 1989

ABSTRACT: Monofunctional psoralens such as 3-carbethoxypsoralen (3-CPs), 7-methylpyrido[3,4-*c*]psoralen (MePyPs), or 7-methylpyrido[4,3-*c*]psoralen (2N-MePyPs) are generally less genotoxic than the bifunctional ones presently used in the photochemotherapy associated with UV-A light (PUVA therapy). In spite of the structural similarities of their respective monoadducts, these compounds widely differ in their photobiological effects. In this paper we compare the local structure of the alterations that are respectively induced within DNA by 3-CPs, MePyPs, and 2N-MePyPs in a first attempt to explain these photobiological differences from a conformational point of view. The internal location of the monoadducts with respect to the helix does not depend on the chemical structure of the derivatives and is consistent with a cis stereochemistry. In contrast, sterical effects related to the chemical structure of the bound residues seem to play a major role in the DNA secondary structure in the vicinity of the alterations. The single-strand-specific S_1 endonuclease, which is used as a probe of the local denaturation around each monoadduct in DNA, does not hydrolyze the MePyPs-induced alterations. In contrast, the photobinding of one 3-CPs or one 2N-MePyPs molecule locally induces the destruction of about 7 base pairs as detected by the sensitivity of the respective modified DNA toward the S_1 endonuclease. Such disruption suggests a possible classification of the monofunctional psoralen derivatives related to the DNA conformation around their monoadducts.

The photochemotherapy of some skin diseases (psoriasis, micosis fungoid, vitiligo) using psoralens such as 8-methoxypsoralen (8-MOP)[1] or 4,5′,8-trimethylpsoralen (TMP) associated with UV-A light (PUVA) has been widely stimulated for the past two decades [for a review, see Pathak et al. (1981)]. A large number of photochemical studies have been developed to understand, at the molecular level, the mechanism of action of these photosensitizing derivatives [for a review, see Song and Tapley (1979) and Averbeck (1984)]. The photobiological properties of psoralens have been partly related

† This work was supported by the Ligue National Française Contre Le Cancer, by INSERM (Grant PRC 121036), and by CNRS (Grant RCP 080572 and ARI interface Chimie-Biologie).

* Author to whom correspondence should be addressed.
‡ Laboratoire de Physique et Chimie Biomoléculaire.
§ Laboratoire de Synthèse Organique.

[1] Abbreviations: 3-CPs, 3-carbethoxypsoralen; MePyPs, 7-methylpyrido[3,4-*c*]psoralen; 2N-MePyPs, 7-methylpyrido[4,3-*c*]psoralen; 8-MOP, 8-methoxypsoralen; TMP, 4,5′,8-trimethylpsoralen; 5-MOP, 5-methoxypsoralen; 4′,5′-DHCPs, 4′,5′-dihydro-3-carbethoxypsoralen; Thy, thymine; dThd, thymidine; dThd<$^{5'4'}_{5'3'}$>3-CPs, C_4 cycloadduct of 3-CPs to dThd involving the 4′,5′-furan ethylenic bond and the 5,6 pyrimidine bond of Thy; UV-A, ultraviolet light of class A (320–400 nm); bp, base pairs; CL, interstrand DNA cross-links; RFI, relative fluorescence intensity.

0006-2960/89/0428-5801$01.50/0