Resolution of Aromatic Resonances in Proton Nuclear Magnetic Resonance Spectra of Transfer Ribonucleic Acid†

Paul G. Schmidt* and Ellen B. Edelheit

ABSTRACT: Purified tRNAs in D_2O yield a set of extremely narrow peaks in the carbon-bound aromatic proton region of ^1H nuclear magnetic resonance spectra. Partial transverse relaxation by a spin–echo train leaves 20–25 narrow peaks (all singlets) in the $tRNA^{Phe}$ (yeast) or $tRNA_1^{Val}$ (*E. coli*) spectra while the rest of the approximately 90 proton resonances between 6.5 and 9 ppm are relaxed away. The resulting spectra display resolution at least as good as that seen in hydrogen-bonded NH proton studies of tRNA in H_2O. Exchange of purine H8 for ^2H demonstrates that most narrow peaks come from H2 of adenosine while the rest are from H8 of A and G. Computer fitting of the complete aromatic spectrum in $tRNA_1^{Val}$ requires 90 Lorentzian peaks of full width at half-height ranging from 1 to 50 Hz. Approximately 30 peaks have experimental widths less than 10 Hz; 16 have widths less than 5 Hz. Similar results hold for $tRNA^{Phe}$ (yeast). Interproton distances, calculated from four different refinements of the $tRNA^{Phe}$ crystal structure, provide a straightforward rationalization of aromatic spectral characteristics. Essentially all H2 protons of A are relatively distant from other proton dipoles (>2.5 Å). All pyrimidine H6 protons are relatively close (1.7–2.5 Å) to other protons; most H8 protons of A and G are close, especially when the base is in a stacked, helical structure, while selected H8 protons are distant (often those at places where regular structure ends or turns). With the assumption of isotropic reorientation of interproton vectors and a single correlation time of 25 ns for $tRNA^{Phe}$, the trend of experimental line widths is correctly predicted by sums of inverse-sixth-power distances from the crystal structure. In particular, the very narrow peaks observed do not require that rapid internal motion be present, although limited internal motion is not precluded.

The discovery of hydrogen-bonded imino proton resonances in NMR[1] spectra of tRNA touched off a new era in nucleic acid structure studies (Kearns et al., 1971). It was later shown that both tertiary and secondary base pairs gave rise to peaks in the 10–15-ppm range (Reid et al., 1979). In recent years, much work has gone into assigning resonances and into extracting structural and dynamical information from the downfield spectra (Reid et al., 1979).

Along with successes in the use of H-bond peaks have emerged some limitations. Structural conclusions, and in many cases assignments, depend on correct interpretations of chemical shifts in terms of ring-current effects. But theories of ring-current anisotropies are still being refined, and there are uncertainties in the intrinsic shift positions of H-bonded tertiary, and even secondary structure, base pairs. Even when, or if, ring-current effects are well understood, the extrapolation of chemical shifts to give detailed structural information is a difficult task. The problem is woefully underdetermined since shifts of a proton due to even just one nearby base are dependent in general on three coordinates. In the face of irregular helix structure and more than one nearby base contributing to an observed shift, the extraction of detailed structural parameters is hopeless at present. Note that the observation of shifts from one or two more additional protons of a base pair would help greatly in calculating the proper relative orientation of nearby bases.

Another fundamental limitation is the lack of H-bond resonances from single-strand regions, in particular the 3′-end CCA sequence and the anticodon loop. These loci may be more important, mechanistically, than the rest of the structure. Methyl resonances of modified bases provide some probes for the anticodon loop (Kastrup & Schmidt, 1978) but not the CCA end.

In this report, we offer a partial remedy to the above-mentioned problems. Earlier we demonstrated that a portion of the tRNA spectrum, ordinarily thought to contain little or no resolution, actually contains some of the narrowest peaks (Schmidt & Kastrup, 1978). This is the aromatic carbon-bound proton region from 6.5 to 9 ppm, most easily studied in D_2O solutions. Using the examples of $tRNA^{Phe}$ (yeast) and $tRNA_1^{Val}$ (*E. coli*), we show in the present report that the resolution in this region can be made even better than that in the imino proton spectrum through use of a simple pulse scheme. A number of peaks display exceptionally narrow line widths ($W_{1/2}$ = 1–5 Hz) while widths of others range up to 50 Hz. By calculating expected *proton* coordinates from the $tRNA^{Phe}$ crystal structure, we are able to account in general for the widths of all peaks in the aromatic spectrum, based on a simple dipole–dipole relaxation model.

The aromatic resonances provide peaks from single-stranded regions as well as secondary and tertiary structure. Thus, structural changes throughout the macromolecule are relfected in these spectra. In addition, aromatic protons provide fertile ground for exploiting nuclear Overhauser effects, and examples have been published recently (Sanchez et al., 1980).

Materials and Methods

Isolation of tRNA. Unfractionated *Escherichia coli* tRNA was obtained from Plenum Scientific Research, Inc. Transfer RNA_1^{Val} was separated from the mixture by using a three-step protocol (Reid et al., 1977). Material from the last step, after

† From the Oklahoma Medical Research Foundation and the Department of Biochemistry and Molecular Biology, University of Oklahoma Health Sciences Center, Oklahoma City, Oklahoma 73104. *Received June 27, 1980.* This research was supported by National Institutes of Health Research Grant GM25261. P.G.S. is the recipient of Research Career Development Award AM00525 from the National Institutes of Health.

[1] Abbreviations used: NMR, nuclear magnetic resonance; DSS, sodium 4,4-dimethyl-4-silapentane-1-sulfonate; DHU, dihydrouridine; T, ribothymidine; m^7G, N^7-methylguanosine; Ψ, pseudouridine; tRNA, transfer ribonucleic acid; NOE, nuclear Overhauser effect; EDTA, ethylenediaminetetraacetic acid; P_i, inorganic phosphate.

0006-2960/81/0420-0079$01.00/0

dialysis to lower the high concentration of $(NH_4)_2SO_4$, generally accepted 1.6 nmol of valine per A_{260} unit. Transfer RNA_1^{Val} prepared in this way yielded ^1H NMR spectra containing all the modified base resonances expected at the appropriate intensity ratios, and no other detectable modified base peaks.

$tRNA^{Phe}$ from brewer's yeast was obtained from Boehringer–Mannheim (lot 1418132) and used without further fractionation. NMR spectra gave no indication of any other tRNA species and were identical with spectra of highly purified $tRNA^{Phe}$ (Davanloo et al., 1979).

Preparation of NMR Samples. Solutions of tRNA at NMR concentrations (usually 0.5–1 mM) were dialyzed, prior to taking spectra, against a 100-fold excess of a buffer containing 0.15 M NaCl, 5 mM P_i, and 1 mM EDTA, pH 7, for 6 h with two changes at 22 °C followed by a similar treatment with EDTA. This procedure yields a solution containing one to two Mg^{2+} ions per tRNA and no detectable EDTA. Samples were lyophilized or evaporated several times from 98.8% D_2O with a final addition of 100 atom % D_2O (Aldrich). Paramagnetic ions are often present in untreated commercial D_2O. Effects of such ions on tRNA spectra are seen in broadening of resonances, but we have found that certain peak line widths are exquisitely sensitive to very low concentrations of Mn^{2+} at levels that do not perturb other peaks. The sensitive resonance serve as internal monitors of significant paramgnetic ion contamination.

^1H NMR Spectra. Spectra were obtained at 220 (Varian Associates, Nicolet Technology Fourier-transform conversion) and 270 MHz (Bruker Instruments magnet and probe; Nicolet Instruments 1180 computer, home-built quadrature phase-detection Fourier-transform spectrometer). The pulsed Fourier-transform mode was used in all cases; 90° pulses were used, and 100–512 acquisitions were time averaged. Recycle time was usually 6–10 s.

Chemical shifts were measured with respect to either the dilute *p*-dioxane in some samples or the residual water peak. Shifts were reported relative to DSS, with positive values being to low field of the standard.

Results

Proton NMR Spectra of $tRNA_1^{Val}$ (E. coli) and $tRNA^{Phe}$ (Yeast) in D_2O. Fourier-transform spectra at 270 MHz are shown in Figure 1. Modified base resonances, in the region 0–4 ppm, are clearly resolved and have served as very useful probes of structure for both species (Kastrup & Schmidt, 1978; Davanloo et al., 1979; Robillard et al., 1977; Kan et al., 1977). In Figure 1, the ribothymidine (T) methyl peak of $tRNA^{Val}$ is found at 1.0 ppm, with small fractions also appearing at 1.9 and 0.85 ppm. The 1-ppm position is appropriate for tRNA in a conformation similar to that of $tRNA^{Phe}$ (yeast) in the crystal (Kan et al., 1977; Robillard et al., 1977). The other positions for T represent altered or partially unfolded structures (Kastrup & Schmidt, 1978). We conclude that, under the solution conditions of Figure 1, $tRNA_1^{Val}$ exists predominantly (but not exclusively) in a tertiary form close to the crystal structure of $tRNA^{Phe}$. Resonance positions of the modified bases in $tRNA^{Phe}$ also indicate that this species is predominantly in a single, native conformation (Davanloo et al., 1979).

Our focus in this paper is not on modified bases but on regions of the spectrum not previously explored in detail. The spectrum divides rather naturally into sections: region I from 9 to 6.5 ppm generally contains peaks of H8 and H2 from A, H8 of G, and H6 of C and U; region II from 6.5 to 5 ppm consists largely of H5 of the pyrimidines and H1′ of all riboses; region III from 5 to 4 ppm contains the other ribose protons;

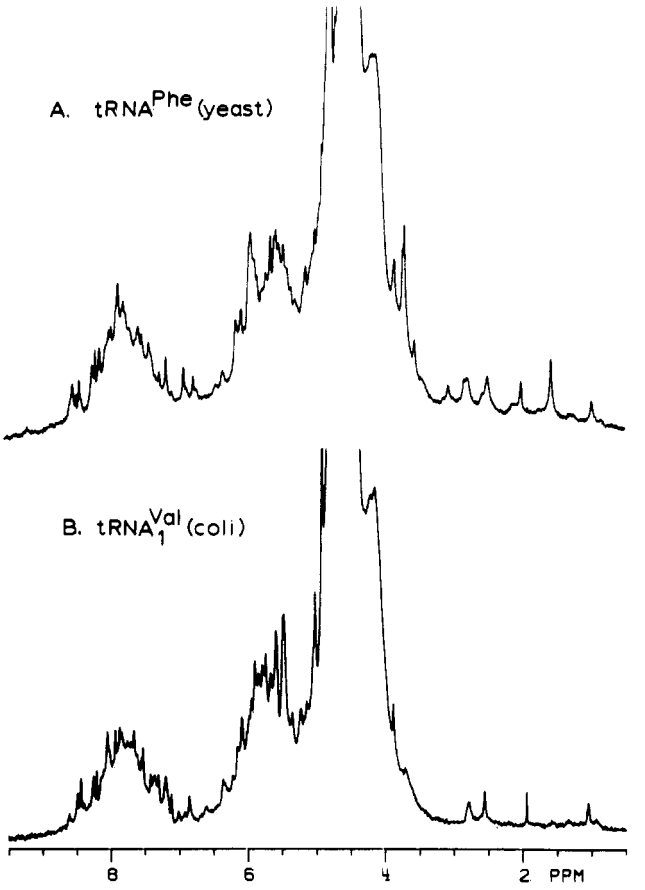

FIGURE 1: 270-MHz ^1H NMR spectra of $tRNA_1^{Val}$ and $tRNA^{Phe}$ in D_2O. Solutions were ca. 20 mg/mL in tRNA, 0.15 M NaCl, and 5 mM P_i, pH 7, 28 °C. Chemical shifts were measured from the residual D_2O peak and are referenced to DSS at 0 ppm after a correction for the temperature dependence of the water peak shift. The band of peaks between 6.5 and 9 ppm is denoted region I in the text.

and region IV from 4 to 0 ppm is made up of modified base peaks. As with ring NH protons of tRNA in H_2O solution, the spread of chemical shifts in region I is due mainly to ring-current effects. Most resonances are seen upfield of their nominal positions which are close to the shifts of the corresponding 5′-mononucleotides: H8, A = 8.51; H2, A = 8.26 H6, C = 8.04; H6, U = 8.01; H8, G = 8.13 ppm (Borer et al., 1975).

The most striking feature of Figure 1 is the appearance of numerous narrow peaks in regions I and II. Such resulution had not been observed in D_2O spectra of tRNAs previously published. We have focused on region I in an effort to understand the origin of the remarkably narrow lines found and to determine whether these peaks can serve as probes of tRNA structure.

Figure 2 shows region I of a 270-MHz spectrum of 1 mM $tRNA_1^{Val}$ in 0.15 M NaCl. The solution contained 1–2 Mg^{2+} ions per tRNA molecule. Figure 2a is the normal Fourier-transform spectrum where the free induction decay (FID) has been weighted to increase the signal to noise ratio, resulting in an extra line broadening of 1 Hz. It appears from this spectrum that region I consists of a number of relatively narrow peaks arranged on a background of broad resonances. This is shown more clearly in Figure 2b which is a convolution differences spectrum (Campbell et al., 1973) and most dramatically in a partially transverse relaxed spectrum (Figure 2c) (Campbell et al., 1975) utilizing a spin–echo pulse train. In the latter trace, only 15–20 protons contribute out of a possible 89. Time between the 90° pulse and the start of data

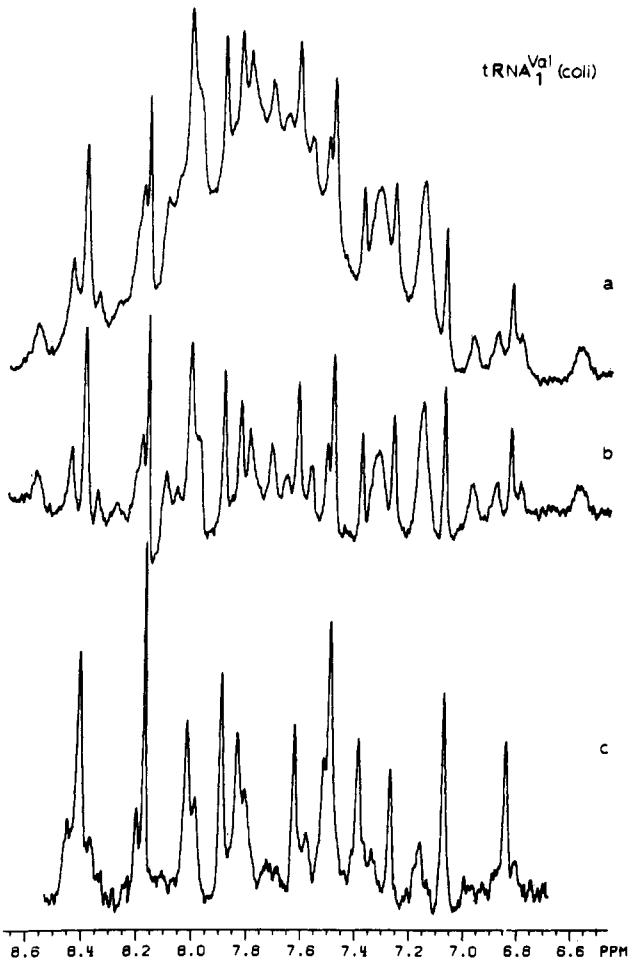

FIGURE 2: Region I of tRNA$_1^{Val}$. (a) Normal Fourier-tranform spectrum, 1-Hz line broadening. (b) Convolution difference spectrum, 15-Hz line broadening subtracted from 1-Hz broadening. (c) Partially relaxed spin–echo spectrum; pulse delay (2τ) between echoes = 1.0 ms; free induction decay collected 42 ms after start of pulse train.

acquisition was 42 ms (21 echoes); peaks with line widths at half-height [$W_{1/2} = 1/(\pi T_2)$] of 20 Hz or more would be essentially eliminated (less than 8% of their amplitude would remain). On the other hand, peaks whose widths were less than 5 Hz would lose less than half of their amplitude. Judging from the relatively flat base line in Figure 2c, it appears that most peaks which make up the broad underlying hump of region I must have line widths in excess of 20 Hz.

Region I should contain singlets from H2 of adenosine and H8 of adenosine and guanosine, plus doublets due to H6 of the pyrimidines. In tRNA$_1^{Val}$, the H6 of T54, 5-oxyacetic acid, U34, and Ψ55 should be found in region I; because of substitutions at the 5 position, their H6 resonances are singlets. Of 33 possible doublets with $J_{5,6} \simeq 8$ Hz from C or U, there are none resolved in the tRNA$_1^{Val}$ spectrum. This implies that all H6 protons of unmodified pyrimidines have line widths greater than 10–15 Hz. The peak at 6.6 ppm in Figure 2a,b, while not resolved, does have a line shape appropriate for a pyrimidine doublet. It could belong to an H6 proton shifted far upfield of its nominal position or an H5 shifted downfield.

H8 of m^{7}G, while carbon bound, is especially labile and exchanges readily for ^{2}H in D_2O (Hurd & Reid, 1979; Johnston et al., 1979). We have observed the m^{7}G H8 resonance of tRNA$_1^{Val}$ at 9.2 ppm, in freshly prepared D_2O solutions at 28 °C. Within 0.5 h, the peak is no longer detectable.

Region I of tRNA^{Phe} (yeast) similarly displays numerous resolved peaks (Figure 3A). When the free induction decay is collected after a train of echoes in the spin–echo experiment,

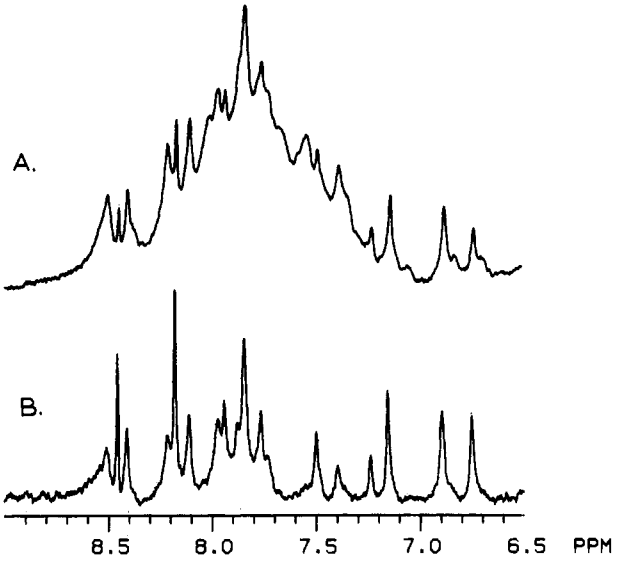

FIGURE 3: Region I of tRNA^{Phe} (yeast). (A) Normal 270-MHz Fourier-transform spectrum. (B) Partially relaxed spin–echo spectrum; $2\tau = 1.0$ ms; free induction decay collected 62 ms after start of pulse sequence. The vertical gain in (B) is greater than in (A).

the resulting frequency spectrum contains only the narrowest peaks (Figure 3B). In this case, acquisition was started 62 ms after the 90° pulse, with a delay, 2τ, between echoes of 1 ms. As with tRNA$_1^{Val}$ spectra, only singlets are resolved in region I of tRNA^{Phe} (yeast). Potential singlets are H2 of A, H8 of A, G, and Y, and H6 of T54, m^{5}C40, m^{5}C49, Ψ39, and Ψ55, for a total of 64. Obviously, most of these do not contribute significantly to the spin–echo spectrum of Figure 3B, which contains 18 resolved peaks, representing about 22–25 protons. Note that, because an individual peak area depends on its T_2 value, it is difficult to calculate the number of protons contributing to unresolved resonances. But it is clear that a limited number of region I peaks, all singlets, are especially narrow.

^1H–^2H Exchange. When a sample of tRNA$_1^{Val}$ was examined after a series of high-temperature runs (several hours at 60–85 °C), much of the original intensity in region I had been lost due to exchange of purine H8 for ^{2}H solvent (Gamble et al., 1976). Initially, the postheating spectrum corresponded to a denatured form of tRNA$_1^{Val}$. $MgCl_2$ was added to give 10 mM, and the sample was incubated at 60 °C for 5 min and then rapidly cooled. Spectra after this treatment were identical in the methyl region with those of tRNA$_1^{Val}$ before heating, indicating restoration of native structure. Attempts to renature tRNA^{Phe} (yeast) after similar heating in D_2O were not successful in restoring the native spectrum. Figure 4 shows convolution difference spectra at 28 °C before and after the temperature runs. Peaks from which intensity disappeared are marked by an x. Note that there are also some chemical-shift changes between the two spectra due to the addition of Mg(II) to the posttemperature-run sample. In another study (Kastrup, 1976), the effects of Mg^{2+} addition to tRNA$_1^{Val}$ solutions were monitored. These data allow us to trace peaks back from the postheated spectrum to the pretemperature run.

Substantial intensity has been lost in peaks throughout the spectrum. In particular, downfield of 8.2 ppm, the area drops to 26% of that seen before heating. The H8 of unperturbed adenosine is farthest downfield of all and falls in this region of the spectrum. Since most shifts due to stacking of bases are upfield, we can tentatively assign peaks between 8.2 and 8.6 ppm to H8 of A.

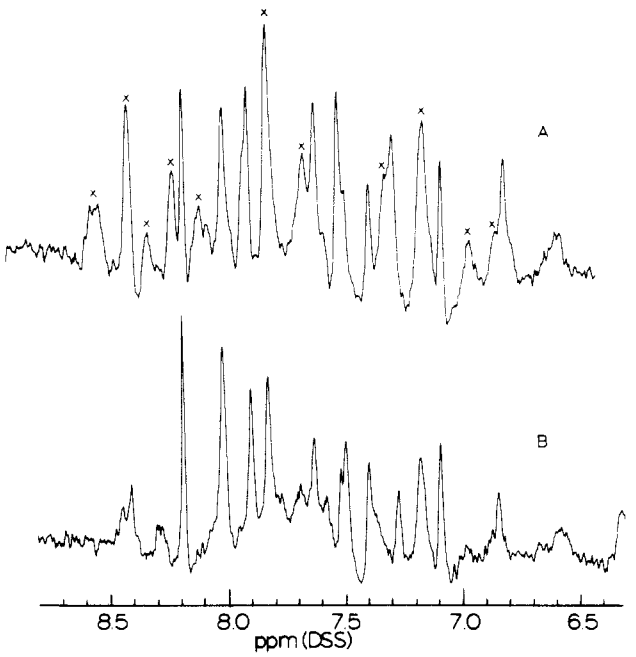

FIGURE 4: Loss of H8 peaks after heating in D_2O. (A) Convolution difference spectrum of $tRNA_1^{Val}$ in D_2O and 0.25 M NaCl; 220 MHz, 28 °C. (B) Spectral conditions as in (A) except sample had been heated at temperatures up to 85 °C for several hours prior to running at 28 °C. $MgCl_2$ (10 mM) was added after heating to promote return of native conformation (see text). Peaks marked with an x in (A) show substantial loss of intensity in (B).

Most resolved peaks of the postheating spectrum must belong to H2 of A since this position does not exchange with the solvent, and the peaks are essentially all singlets. As pointed out above, H6 of T, Ψ, and 5-oxyacetic acid U could also appear. It is interesting to note that those peaks remaining in the D_2O-exchanged spectrum are also the ones with the longest T_2 values as seen by their persistence in the partially relaxed Carr–Purcell spectrum (Figure 2c). One notable exception is the adenosine H8 at 8.4 ppm, which is exchanged for ^2H, yet which has a long T_2.

Spectral Simulation. We used the Nicolet 1180 curve analysis program to fit region I spectra. The result for $tRNA_1^{Val}$ is displayed in Figure 5, where the computer-generated spectrum is shown above the experimental trace. The fit is quite satisfactory.

As one check on the validity of our simulation, we compared the area of the narrow peak at 7.06 ppm to that of the total between 6.7 and 8.6 ppm. This particular peak is known to contain one proton; its chemical shift is strongly temperature dependent, and we have shown that no other resonance lies underneath (Kastrup, 1976). If the 7.06-ppm peak has an area of 1.0 proton, then the whole band contains 90 ± 3 protons. A total of 89 protons are expected in region I, assuming none is shifted more than 1.8 ppm upfield. From studies of the temperature dependence over the range 22–85 °C, we conclude that no more than one or possibly two protons belonging to region I are shifted further upfield than 6.7 ppm (P. G. Schmidt and R. V. Kastrup, unpublished experiments), so our computer simulation appears to give the correct relative area.

A detailed look at the results of this spectral simulation reveals that several peaks have less than unit area. Examples include resonances at 6.82 (0.7 protons), 7.5 (0.8 protons), and the very narrow peak at 8.18 ppm (0.8 protons). The latter has been assigned to the 3′-terminal adenosine H2. Because of its mobility relative to the rest of the macromolecule, and the distance of the H2 proton from other nuclear dipoles, this resonance has a uniquely long T_1 value (~7 s). Under the pulse-repetition conditions employed for the spectrum of Figure 5, the peak area is selectively suppressed.

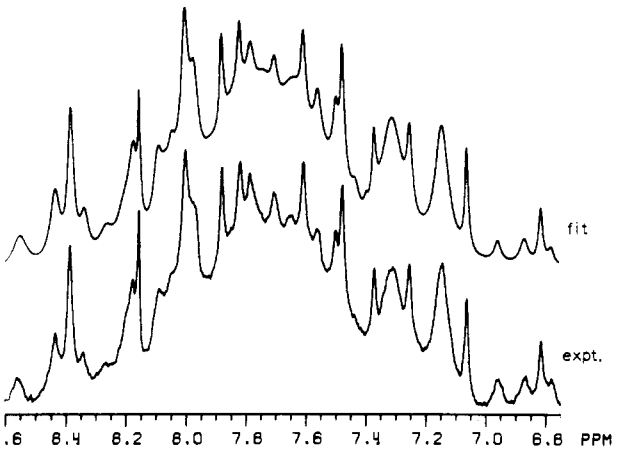

FIGURE 5: Computer simulation of $tRNA_1^{Val}$ aromatic peaks. The lower tracing, labeled expt., is the experimental spectrum of Figure 2a. The upper plot, labeled fit, is a computer-generated simulation (Nicolet CAP program), representing the sum of 47 Lorenztian peaks ranging in width from 1 to 49 Hz. Several of the broader peaks contained more than unit area.

The other peaks with less than unit area are generally from protons on bases involved in tertiary structure or whose shift is determined by tertiary structure nucleotides. $tRNA_1^{Val}$ at 28 °C exists predominantly in a tertiary structure that resembles the crystal structure of $tRNA^{Phe}$ (yeast). But, as revealed in the pattern of modified base resonances, there is a fraction of the molecules in other forms (Kastrup & Schmidt, 1978). Typically, for $tRNA_1^{Val}$, this fraction ranges from 20% to <5%, depending on the solution conditions and pretreatment of the sample.

Out of the computed-aided fit to region I we have extracted 89 protons contained in 41 peaks. Five individual resonances were found to have less than 0.5 proton and were not used. A total of 25 peaks were assigned to single protons. Several had less than unit intensity due to nonhomogeneous tertiary structure but were ascribed to 1 proton so as to more nearly approximate the spectrum of a single tertiary structure.

While the precision of the fit for narrow, single-proton lines is very good (estimated ± 10% for the line width), it is more difficult to accurately fit the broad, unresolved peaks which make up the "hump". In particular, there could be a few peaks with extremely large line widths (>100 Hz) underneath. However, we were able to make an excellent fit with peaks having a maximum width of 49 Hz.

After 1 Hz was substracted for digital broadening, the $tRNA_1^{Val}$ region I spectrum contains 16 peaks whose full width at half-height is 5 Hz or less and 29 with widths less than 10 Hz. Those are remarkably narrow lines for a macromolecule of mass 25 000 daltons. Note that, for example, the imino protons of tRNA in H_2O solution are approximately 25–30 Hz wide near 30 °C (Reid et al., 1979).

Region I for a sample of $tRNA^{Phe}$ (yeast) in 0.2 M NaCl and no added Mg^{2+} was also fit by computer simulation (Figure 6). Memory-size limitations prevented including the peaks at 6.67, 6.7, and 6.76 ppm in the calculated trace; these were measured by hand. The details of the fitting procedure were very similar to those of $tRNA_1^{Val}$ spectral simulation. Again, approximately 90 protons contributed to the region I spectrum of $tRNA^{Phe}$, with line widths varying from less than 1 Hz to 50 Hz. All peaks were treated as single Lorentzian lines since, as with the $tRNA_1^{Val}$ spectrum, no doublets were resolved. We also fit a region I spectrum of a $tRNA^{Phe}$ sample

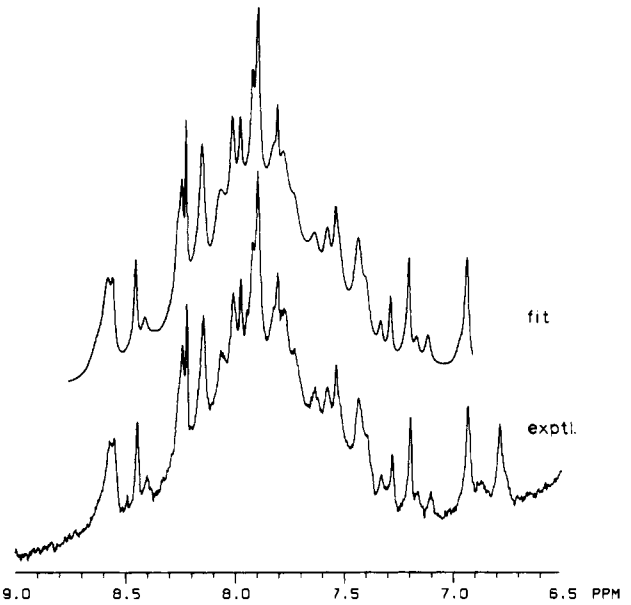

FIGURE 6: Computer simulation of $tRNA^{Phe}$ (yeast) aromatic spectrum. Details of the fitting procedure were similar to those of Figure 5 for $tRNA_1^{Val}$.

containing 10 mM $MgCl_2$. While the chemical shifts of many resolved peaks changed somewhat upon addition of Mg^{2+}, the line widths stayed close to their original values.

Interproton Distances Calculated from Crystal Structures. Carbon-Bound Protons. For tRNA in D_2O, ^1H NMR line widths are determined by proton–proton, dipole–dipole interactions modulated by rotational reorientation of the internuclear vectors. The width at half-height of a resonance can be approximated by

$$W_{i,1/2} \cong C(\sum_{i \neq j} r_{ij}^{-6}) f_{ij}(\tau_R) \quad (1)$$

where C collects constants of the dipolar interaction energy, r_{ij} is the distance between the proton of interest and other protons, and $f_{ij}(\tau_R)$ is a function of the rotational correlation time of the internuclear vectors. Because of the inverse-sixth-power dependence on distance, the line widths are exquisitely sensitive to r_{ij}.

We used the published coordinates of all C, N, and O atoms in $tRNA^{Phe}$ (yeast) to generate to coordinates of all protons. We recognized that the results might be significantly different when using the models refined in the laboratories at the Medical Research Council (MRC) (Ladner et al., 1975), Massachusetts Institute of Technology (MIT) (Quigley et al., 1975), or Duke University (Duke I) (Sussman & Kim, 1976), (Duke II) (Sussman et al., 1978), so we combined and compared results by using four sets of refined coordinates. The MIT data were generously supplied to us by Dr. A. H.-J. Wang as a computer printout; the MRC and both sets of Duke data were on tape supplied by the Protein Data Bank of Brookhaven National Laboratory (Bernstein et al., 1977).

Proton coordinates were calculated with the aid of the Fortran program HYGEN, kindly supplied to us by Professor I. C. Paul, University of Illinois. Mark Weber in our laboratory modified the HYGEN program for more facile use of the tRNA data. This program requires as input the coordinates of the carbon atom to which the proton is attached and coordinates of other bonded C, N, or O atoms. A C–H bond distance of 1.08 Å was specified. The HYGEN routine then sought the best-fit position for each proton.

Using those calculated coordinates, we then determined interproton distances for each of the four models. The program printed out spans less than 7 Å and generated a sum of inverse-sixth-power distances. Proton–proton distances greater than 7 Å were found to contribute negligibly to the line width.

Table I presents the sum of inverse-sixth-power distances for carbon-bound protons of region I of $tRNA^{Phe}$ (yeast). The mean from data of the MIT, MRC, Duke I, and Duke II laboratories and the standard deviation are shown. Also listed is the closest neighboring atom, and its distance, based on the latest Duke tRNA structure only (an arbitrary choice but quite representative). In a few cases, an interproton distance was found to be less than the sum of van der Waals radii. When a data set contained $r_{ij} < 1.7$ Å, we simply chose not to average that $\sum r_{ij}^{-6}$ value with the other values.

Several aromatic protons are adjacent to methyl groups of the same base. We included these interactions by measuring average distances for three different rotamers. Methyl proton contributions to $\sum r_{ij}^{-6}$ were not particularly sensitive to the choice of rotamers or their relative populations; an average value of 6×10^{-3} Å^{-6} was found for 6-membered rings and 4×10^{-3} Å^{-6} for the 5-membered ring of m^7G. In a few cases, methyl groups are located near aromatic protons of other bases. For these, the average interproton distance was approximated by using the methyl carbon coordinates. The result is therefore only a qualitative check of proximity. Rapid internal rotation of methyl groups undoubtedly diminishes their effect on line widths. But it is interesting to note that substantial nuclear Overhauser effects are seen between modified base methyl resonances and nearby aromatic protons (Sanchez et al., 1980).

A number of generalizations are readily made from the data of Table I. The H2 protons of adenosine lie relatively distant from other protons. Except when a methyl group is close, the average value of $\sum r_{ij}^{-6}$ is 3×10^{-3} Å^{-6} or less. According to the Duke II refinement, H2 atoms have proton neighbors no closer than 2.5 Å away. This aloofness results from the position of H2 on the inside of helices and other stacked structures, generally away from the carbon-bound protons of the ribose rings. By way of contrast, H6 atoms of the pyrimidines all have relatively close neighbors. Most possess an H5 proton about 2.4 Å away or methyl group protons, and the generally adopted trans orientation of the base places H6 near ribose protons. As a result, the *smallest* value of $\sum r_{ij}^{-6}$ for H6 is 8×10^{-3} Å^{-6} (for C56, a tertiary structure base), and the average is 21×10^{-3} Å^{-6}. With H8 of G and A, the pattern is more diverse. Except for m^7G, there is no nearby proton on the base so that, in every case but one, ribose protons are the nearest neighbors. When a purine is involved in the regular structure, such as a helical arm or a base-stacked single-strand section, the H8 is almost always relatively close to ribose protons (within 1.7–3 Å), likewise, the H6 of stacked pyrimidines. But in most cases, it is the ribose protons of the nucleotide one removed of the 5′side that are closest. Thus, in the anticodon arm, for example, C27 H6 is nearest m_2^2G26 H2′; C28 H6 and C27 H2′; A29 H8, C28 H2′; and so forth. The scheme breaks down where the tRNA structure goes off in a new direction such as at A9–m^2G10 or U33–G_m34 and, of course, at the 5′ end of the molecule.

In summary, H2 protons of all adenosines are relatively distant from other protons in tRNA, all pyrimidine H6 protons are relatively close, and H8 protons of A and G range from very distant to very close to other protons. The distant H8 protons often occur on bases at interesting junctions in the structure.

Interproton Distances and NMR Line Widths. According to eq 1, there ought to be a linear correlation betweed $\sum r_{ij}^{-6}$ and observed ^1H NMR line widths in tRNA if $f_{ij}(\tau_R)$ is similar

Table I: Interproton Distances in $tRNA^{Phe}$ (Yeast) (D_2O)[a]

base	proton	$\Sigma r_{ij}^{-6} \times 10^3$ (Å^{-6})[b] MIT	MRC	Duke I	Duke II	av (σ)	nearest[c] neighbor	r_{ij} (Å)[d]	base	proton	$\Sigma r_{ij}^{-6} \times 10^3$ (Å^{-6})[b] MIT	MRC	Duke I	Duke II	av (σ)	nearest[c] neighbor	r_{ij} (Å)[d]
G1	H8	2	2	1	4	2 (1)	G1H5′	3.0	G42	H8	15	16	14	13	14 (1)	U41H2′	2.2
C2	H6	32	31	10	59	33 (20)	C2H5′	1.8	G43	H8	12	13	7	26	14 (8)	G42H2′	1.9
G3	H8	15	12	75*[e]	22	16 (5)	C2H3′	1.9	A44	H8	39	22	21	12	24 (11)	G43H2′	2.3
G4	H8	80*	42	6	13	20 (19)	G4H5′	2.3		H2	1	1	1	1	1[f] (–)	G45H1′	3.9[f]
A5	H8	150*	13	9	27	16 (9)	G4H2′	2.1	G45	H8	7	6	14	12	10 (4)	A44H3′	2.3
	H2	3	3	2	3	3 (0.5)	U6H1′	2.9	m^7G46	H8	43	21	10	29	26 (14)	m^7G46H5′	2.0
U6	H6	21	15	102*	21	19 (3)	A5H3′	2.3	U47	H6	42	6	19	20	22 (15)	U47H5′	2.0
U7	H6	60	16	13	14	26 (23)	U7H5	2.4	C48	H6	121*	9	9	13	10 (2)	C48H5	2.4
U8	H6	11	11	13	12	12 (1)	U8H5	2.4	m^5C49	H6	9	11	10	12	10 (1)	m^5C49H5′	2.7
A9	H8	35	31	5	11	20 (15)	A9H2′	2.2	U50	H6	21	31	13	85*	22 (9)	m^5C49H2′	1.5
	H2	1	1	1	1	1 (–)	G22H5′	4.1	G51	H8	17	18	10	17	16 (4)	U50H3′	2.1
m^2G10	H8	2	2	2	6	3 (2)	m^2G10H5′	2.5	U52	H6	69*	25	26	35	29 (6)	G51H2′	1.8
C11	H6	37	20	15	73*	24 (12)	m^2G10H2′	1.6	G53	H8	16	9	6	8	10 (4)	U52H2′	2.6
U12	H6	15	16	24	29	17 (8)	C11H2′	1.9	T54	H6	20	18	16	15	17 (2)	G53H2′	2.6
C13	H6	24	10	10	9	13 (7)	C13H5	2.4	Ψ55	H6	11	11	13	14	12 (2)	T54H2′	2.3
A14	H8	20	5	6	7	10 (7)	A14H5′	2.6	C56	H6	9	7	7	8	8 (1)	C56H5	2.4
	H2	3	2	2	3	2 (0.6)	G22H1′	2.9	G57	H8	15	14	41	>200*	28 (15)	G47H5′	1.0
G15	H8	38	10	5	24	19 (15)	A14H3′	2.0	m^1A58	H8	56*	26	11	9	15 (9)	m^1A58H2′	2.2
G18	H8	39	13	3	7	16 (16)	G18H2′	2.4		H2	10	13	9	13	11 (2)	C60H2′	2.5
G19	H8	14	15	4	12	11 (5)	G19H2′	2.2	U59	H6	10	7	12	>200*	10 (3)	U59H5′	1.2
G20	H8	4	4	4	5	4 (1)	G19H3′	2.7	C60	H6	28	33	21	18	25 (7)	C60H2′	2.4
A21	H8	16	19	3	10	12 (7)	A21H5′	2.3	C61	H6	12	17	20	37	22 (11)	C61H5′	2.0
	H2	4	2	3	2	3 (1)	U8H1′	3.0	A62	H8	15	34	31	40	30 (11)	C61H2′	1.7
G22	H8	8	3	3	13	7 (5)	G22H5′	2.3		H2	5	3	3	2	3 (1)	G53H1′	3.0
A23	H8	11	12	5	37	16 (14)	G22H2′	1.8	C63	H6	25	21	14	16	19 (5)	A62H2′	2.4
	H2	3	1	3	1	2 (1)	C13H1′	3.4	A64	H8	19	25	15	39	24 (11)	C63H2′	1.7
G24	H8	23	21	24	67*	23 (2)	A23H2′	1.6		H2	2	1	3	3	2 (1)	G65H1′	2.7
C25	H6	20	16	11	36	21 (11)	G24H2′	1.9	G65	H8	15	13	4	7	10 (5)	A64H2′	2.6
m_2^2G26	H8	13	7	13	88*	11 (3)	C25H3′	1.5	A66	H8	88*	36	8	16	20 (14)	G65H2′	2.2
C27	H6	136*	29	42	127*	36 (9)	m_2^2G26H2′	1.4		H2	3	2	1	2	2 (1)	A67H1′	3.5
C28	H6	24	14	18	28	21 (6)	C27H2′	1.9	A67	H8	16	21	8	18	13 (6)	A66H2′	2.1
A29	H8	8	28	7	28	18 (12)	C28H2′	1.9		H2	1	1	1	1	1 (–)	U68H1′	3.8
	H2	3	3	1	5	3 (2)	G30H1′	2.5	U68	H6	45	20	134*	16	27 (16)	U68H5	2.4
G30	H8	9	9	9	5	8 (2)	A29H3′	2.9	U69	H6	21	26	19	28	24 (4)	U68H2′	2.0
A31	H8	7	6	9	14	9 (4)	G30H3′	2.3	C70	H6	86*	30	24	60*	27 (4)	U69H2′	1.6
	H2	1	1	1	1	1 (–)	C_m32H1′	4.4	G71	H8	108*	19	18	139*	18 (1)	C70H2′	1.4
C_m32	H6	46	19	18	60	36 (21)	A31H3′	1.7	C72	H6	72*	13	13	17	14 (2)	G71H2′	2.3
U33	H6	36	50	11	10	27 (20)	U33H5	2.4	A73	H8	8	20	9	20	14 (7)	C72H2′	2.1
G_m34	H8	3	2	2	3	2 (1)	G_m34H5′	3.1		H2	2	1	2	2	2 (0.5)	C74H1′	2.9
A35	H8	13	19	13	23	17 (5)	G_m34H2′	2.0	C74	H6	26	14	12	35	22 (11)	A73H2′	1.8
	H2	1	1	1	3	2 (1)	A36H1′	2.8	C75	H6	22	141*	13	18	18 (5)	C74H3′	2.2
A36	H8	10	11	4	4	7 (4)	A36H5′	3.2	A76	H8	6	3	4	5	4 (1)	A76H2′	2.7
	H2	1	1	1	2	1 (0.5)	Y37H1′	3.2		H2	0	1	1	1	1 (0.5)	A76H5′	3.4
Y37	H8	209*	34	10	16	20 (12)	A36H2′	2.3									
A38	H8	28	39	5	23	24 (14)	Y37H2′	2.1									
	H2	2	1	1	1	1 (0.5)	Ψ39H1′	3.4									
Ψ39	H6	70*	28	12	47	29 (18)	A38H3′	1.7									
m^5C40	H6	216*	30	46	38	38 (8)	Ψ39H2′	1.9									
U41	H6	27	15	13	24	20 (7)	m^5C40H2′	2.1									

[a] Includes carbon-bound protons only. [b] Sum of inverse-sixth-power distances. [c] From data of Sussman et al. (1978) (Duke II). [d] Distance of nearest neighbor to relevant proton. [e] The asterisk indicates that Σr_{ij}^{-6} contains an $r_{ij} < 1.7$ Å. [f] Values do not include contribution of protons on a methyl group of m_2^2G26 approximately 2.7 Å from A44 H2.

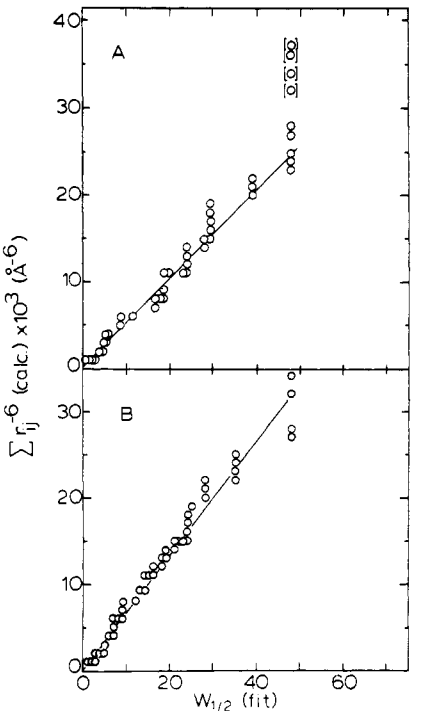

FIGURE 7: Plots of calculated inverse-sixth-power distances vs. experimental line widths. Values of $\sum r_{ij}^{-6}$ from Table I were plotted in order of increasing value against line widths from the spectral fits of Figures 5 and 6. (A) $tRNA^{Phe}$ (yeast) (line width data from Figure 6); (B) $tRNA_1^{Val}$ (line width data from Figure 5). Many points represent two or more coincidental data values. The *lines* in each plot are the linear least-squares fit to *all* the data except the points in brackets in (A).

for all protons. Table I shows a wide range in the sum of inverse-sixth-power distances for aromatic protons, from 1×10^{-3} Å^{-6} to 38×10^{-3} Å^{-6}. Indeed, region I spectra of $tRNA^{Ph}$,e and $tRNA_1^{Val}$ have peaks whose widths range from 1 to 50 Hz. Additionally, the narrowest peaks are predicted to be singlets, which is what is found in the spectra. Of the 25 resonances resolved into single proton peaks in $tRNA_1^{Val}$, all but one are singlets, and similarly with $tRNA^{Phe}$. In line with the experimental data, 23 singlets of Table I have values of $\sum r_{ij}^{-6}$ less than the smallest doublet value.

If a group of bases (the single-strand regions, for example) had significantly greater motional freedom than others (e.g., helices), we would expect to find several narrow H6 doublets in the spectrum. There is evidence based on T_1 values for $tRNA_1^{Val}$ that the 3′-terminal adenosine and DHU16 have fast motion independent of the rest of the macromolecule (Kastrup, 1976). But most, if not all, other nucleotides appear to reorient approximately with the overall structure. Some limited flexibility around single bonds is not ruled out by these data, however. ^{13}C NMR relaxation time and nuclear Overhauser effect measurements suggest that tRNA ribose rings may have local motion relative to the bases (Bolton & James, 1980).

To further test the correlations times governing relaxation of aromatic protons, we plotted calculated interproton distances from Table I vs. experimental line widths from the computer fits. Since only a few specific assignments are possible at the present time, we simply paired values of $W_{1/2}$ and $\sum r_{ij}^{-6}$ in ascending order. Figure 7 shows the correlation. Many circles on the plot represent two or more coincidental data points. The line in Figure 7A is the least-squares best fit to all the data except those in brackets and has a slope of 0.5×10^{-3} Å^{-6}·s for $tRNA^{Phe}$. According to eq 1, the slope is $[Cf_{ij}(\tau_R)]^{-1}$. With the assumption for the moment that $f_{ij}(\tau_R)$ is the isotropic rotational correlation time of the molecule, τ_R, then $C = 8.0 \times 10^{10}$ s^{-2} Å^6, and $\tau_R = 25 \times 10^{-9}$ s. The NMR rotational correlation time is the same as that which was measured by using fluorescence depolarization of ethidium bromide bound to $tRNA^{Phe}$ (Tao et al., 1970). Such as correspondence suggests that rapid internal motions of large amplitude are not present in $tRNA^{Phe}$ under normal solution conditions. A similar analysis for $tRNA_1^{Val}$ (Figure 7B) has a slope of 0.67×10^{-3} Å^6 Hz^{-1} to yield $\tau_R = 19 \times 10^{-9}$ s.

The four points in brackets in the $tRNA^{Phe}$ plot lie significantly above the best-fit line. They correspond to four predicted broad lines that were not found in the spectrum. Very broad lines ($W_{1/2} > 50$ Hz) could easily have been missed in our computer simulation. Alternately, internal motion of ribose protons relative to base protons would narrow the lines, particularly for very short interproton distances. Finally, the largest values of $\sum r_{ij}^{-6}$ have the greatest uncertainty because of the inverse-sixth-power dependence.

Obviously, the significance of these plots is diminished by the necessity of pairing the variables in order of size, rather than by direct assignment. Nonetheless, the trend of virtually all the data along a single line suggests that line widths are sensitive to a single effective correlation time. Models in which, e.g., $^1/_2$ the peaks are narrowed by a factor of 4 or more due to internal motion would not produce the linear relationship seen. More subtle differences would be hard to detect, however.

In the case of $tRNA_1^{Val}$, the correlation involves calculated interproton distances from the $tRNA^{Phe}$ crystal structure and experimental line widths from the $tRNA_1^{Val}$ spectrum. $tRNA^{Phe}$ and $tRNA_1^{Val}$ belong to the same tertiary structure class (D4V5), so the general trend of line widths should be quite similar. This is borne out in the linear relationship of $\sum r_{ij}^{-6}$ and $W_{1/2}$ for $tRNA_1^{Val}$ (Figure 7B). Quantitative conclusions from this plot are subject to considerable uncertainty, however.

Discussion

Much of the research on NMR of tRNA is now pushing toward understanding how proteins in the protein synthesis cycle interact with this nucleic acid. Protein–nucleic acid complexes have substantially higher molecular weights than that of the tRNA alone so that peaks which are barely resolved in the tRNA may become too broad in the complex. Spectra of the carbon-bound aromatic proton region offer two advantages in this regard. A finite number of peaks are extraordinarily narrow, and these peaks come from bases in single-strand regions as well as in secondary and tertiary structures. Consider, for example, the peak at 8.17 ppm in D_2O spectra of $tRNA^{Phe}$ and $tRNA_1^{Val}$ (Figures 3 and 2). This belongs to H2 of A76, the 3′-terminal base (Kastrup, 1976). It is the narrowest peak in the whole spectrum (less than 1 Hz wide) so that it is easy to resolve out of a complex resonance pattern by use of the Carr–Purcell spin–echo train, and, of course, it is of interest to know what this portion of the tRNA is doing when bound to proteins.

The large range of line widths in the aromatic region is easily rationalized on the basis of the range of interproton distances in tRNA. Four different refinements of crystallographic data on $tRNA^{Phe}$ produced similar predictions of interproton distances. Significant discrepancies in $\sum r_{ij}^{-6}$ occurred in a few cases where distances were short (<2 Å), which is to be expected for a quantity raised to the inverse sixth power. All data sets reinforced the general observation that H2 protons of A are relatively distant from other protons, pyrimidine H6 protons are relatively close to at least one other proton, and selected H8 purine nuclei are distant, particularly when the

base is located at a bend or turn in the regular helical structure. The line widths for aromatic protons predicted by using the simple isotropic reorientation model and interproton distances from the crystal structure are consistent with the spectra, although further work and more assignments will undoubtedly elaborate on this observation.

Given the still limited accuracy of atomic coordinates for the $tRNA^{Phe}$ crystal structure and our use of a simple isotropic model for motion, it would be very surprising to find close correspondence between predicted and observed line widths for all resonances. This is particularly true for protons with very close neighbors, since small errors in interproton distances become large errors in r_{ij}^{-6}. For example, the difference in line widths for a proton with a neighbor 2.0 or 2.5 Å away is a factor of 3.8. An example is the m^7G46 H8 proton near 9.2 ppm in feshly made D_2O solutions. Its width in $tRNA_1^{Val}$ spectra at 28 °C is 16 Hz and about 10 Hz for $tRNA^{Phe}$. According to the average $\sum r_{ij}^{-6}$ from Table I, that resonance should be much broader (about 50 Hz for $tRNA^{Phe}$) because of close proximity to ribose protons. However, one of the structure refinements (Duke I) would predict a width of 20 Hz for $tRNA^{Phe}$ and 15 Hz for $tRNA_1^{Val}$, much closer to the observed values. Given a detailed understanding of tRNA dynamics, carbon-bound proton line widths may become highly accurate gauges of average local structure. At present, they must be regarded as qualitative indicators of conformation.

An important point emerges from the analysis of interproton distances in $tRNA^{Phe}$. Nucleic acids, and in particular tRNA, are distinct from proteins in that, for most aromatic protons, the closest dipole lies on a *different unit* in the polymer. This makes NOE measurements particularly useful for determination of tRNA tertiary structure in solution. Redfield and co-workers have observed NOE's between the imino proton and H2 of secondary AU base pairs, and between methyl groups and purine H8 protons (Sanchez et al., 1980) for tRNA in H_2O. These experiments offer a method for assignment of resonances, and Sanchez et al. (1980 have been able to make tentative assignments of six adenosine H2 peaks based on NOE and ring-current shifts (Arter & Schmidt, 1976). Note that in H_2O UN3H is the proton closest to AH2 of a WC base pair, with a distance of about 2.8 Å. Many other potential NOE's are suggested by the data of Table I, which lists nearest neighbors for all carbon-bound protons. Our own preliminary NOE work confirms that most of the broad, unresolved aromatic peaks are from nuclei close to ribose protons (P. G. Schmidt and E. B. Edelheit, unpublished experiments).

Acknowledgments

We thank Dr. Andrew A.-H. Wang for a copy of the MIT group $tRNA^{Phe}$ coordinates and Drs. Al Redfield, Brian Reid, and Ralph Hurd for data in advance of publication. Special thanks go to Mark Weber for his invaluable help with the computer programing.

References

Arter, D. B., & Schmidt, P. G. (1976) *Nucleic Acids Res. 3*, 1437–1447.

Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. F., Jr., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T., & Tasumi, M. (1977) *J. Mol. Biol. 112*, 535–542.

Bolton, P. H., & James, T. L. (1980) *Biochemistry 19*, 1388–1392.

Borer, P. N., Kan, L. S., & Ts'o, P. O. P. (1975) *Biochemistry 14*, 4847–4863.

Campbell, I. D., Dobson, C. M., Williams, R. J. P., & Xavier, A. V. (1973) *J. Magn. Reson. 11*, 172–181.

Campbell, I. D., Dobson, C. M., Williams, R. J. P., & Wright, P. E. (1975) *FEBS Lett. 57*, 96–99.

Davanloo, P., Sprinzl, M., & Cramer, F. (1979) *Biochemistry 18*, 3189–3199.

Gamble, R. C., Schoemaker, H. J. P., Jekowsky, E., & Schimmel, P. R. (1976) *Biochemistry 15*, 2791–2799.

Hurd, R. E., & Reid, B. R. (1979) *Biochemistry 18*, 4017–4024.

Johnston, P. D., Figueroa, N., & Redfield, A. G. (1979) *Proc. Natl. Acad. Sci. U.S.A. 76*, 3130–3134.

Kan, L. S., Ts'o, P. O. P., Sprinzl, M., Haar, F. v. d., & Cramer, F. (1977) *Biochemistry 16*, 3143–3154.

Kastrup, R. V. (1976) Ph.D. Thesis, University of Illinois, Urbana.

Kastrup, R. V., & Schmidt, P. G. (1978) *Nucleic Acids Res. 5*, 257–269.

Kearns, D. R., Patel, D., & Shulman, R. G. (1971) *Nature (London) 229*, 338–339.

Ladner, J. E., Jack, A., Robertus, J. D., Brown, R. S., Rhodes, D., Clark, B. F. C., & Klug, A. (1975) *Nucleic Acids Res. 2*, 1629–1637.

Quigley, G. J., Seeman, N. C., Wang, A. H.-J., Suddath, F. L., & Rich, A. (1975) *Nucleic Acids Res. 2*, 2329–2342.

Reid, B. R., Ribeiro, N. S., McCollum, L., Abbate, J., & Hurd, R. E. (1977) *Biochemistry 16*, 2086–2094.

Reid, B. R., McCollum, L., Ribeiro, N. S., Abbate, J., & Hurd, R. E. (1979) *Biochemistry 18*, 3996–4005.

Robillard, G. T., Tarr, C. E., Vosman, F., & Reid, B. R. (1977) *Biochemistry 16*, 5261–5273.

Sanchez, V., Redfield, A. G., Johnston, P. D., & Tropp, J. (1980) *Proc. Natl. Acad. Sci. U.S.A. 77*, 5659–5662.

Schmidt, P. G., & Kastrup, R. V. (1978) *Biomolecular Structure and Function* (Agris, P. F., Ed.) pp 517–525, Academic Press, New York.

Sussman, J. L., & Kim, S.-H. (1976) *Biochem. Biophys. Res. Commun. 68*, 89–96.

Sussman, J. L., Holbrook, S. R., Warrant, R. W., Church, G. M., & Kim, S.-H. (1978) *J. Mol. Biol. 123*, 607–630.

Tao, T., Nelson, J. H., Jr., & Cantor, C. R. (1970) *Biochemistry 9*, 3514–3524.