

## AN NMR APPROACH TO tRNA TERTIARY STRUCTURE IN SOLUTION

G.T. ROBILLARD, C.E. TARR\*, F. VOSMAN

*Department of Physical Chemistry, University of Groningen, Groningen, The Netherlands*

and

J.L. SUSSMAN

*Department of Structural Chemistry, Weizmann Institute Rehovot, Israel*

Received 1 December 1976

Revised manuscript received 18 January 1977

Atomic coordinates of E. Coli tRNA<sup>Val</sup> have been generated from the X-ray crystal structure of Yeast tRNA<sup>Phe</sup> by base substitution followed by idealization. The NMR spectrum of E. Coli tRNA<sup>Val</sup> was then calculated using these coordinates and ring current calculations. The similarity between the calculated and observed NMR spectra of this tRNA gives credence to the procedure used previously to compare the solution and X-ray crystal structure of Yeast tRNA<sup>Phe</sup> and further suggests a strong similarity between the structures of Yeast tRNA<sup>Phe</sup> and E. Coli tRNA<sup>Val</sup> both in solution and in the crystal. The results presented here indicate that a combination of NMR and atomic coordinates generated from a tRNA model may be a powerful procedure for determining three dimensional structure of a tRNA in solution.

### 1. Introduction

Employing three dimensional X-ray crystal structure information of macromolecules as an aid in understanding the biological functions in solution can be done reliably only when independent evidence confirms that the crystal structure corresponds to the predominant structure in solution. Approaches to obtaining such evidence include measuring reactions in the crystal and in solution [1] as well as applying spectroscopic techniques such as fluorescence to probe the characteristics of particular functional groups in both the crystal and solution [2].

Of the spectroscopic techniques, proton NMR is one of the most sensitive and most general in that it can monitor the immediate environment of each proton in the molecule. In theory, therefore, an NMR spectrum calculated from atomic coordinates of an X-ray crystal structure could be directly compared with a measured NMR spectrum of the molecule in

solution. Similarity between the observed and calculated spectrum would indicate the similarity between the solution and crystal structures.

The calculations of NMR spectra, however, pose significant problems since all parameters which influence the chemical shift of a nucleus are not accurately known and difficult to determine. In nucleic acids this problem is somewhat simplified in that the ring currents arising from the aromatic rings of the individual bases provide the dominant influence in determining the proton chemical shifts. All other shift mechanisms, in diamagnetic systems, with the exception of electric fields generate shifts on the order of 10% of those generated by the ring currents and, in first approximation, may be neglected. Electric fields are important for proteins containing  $\alpha$ -helical regions buried in the interior. However, it is unlikely that such fields develop in tRNA.

With these assumptions we have recently compared the X-ray crystal structure of Yeast tRNA<sup>Phe</sup> with its solution structure by calculating the NMR spectrum from the atomic coordinates of the crystal structure [3] taking into account only ring current shifts [4]. The resulting spectrum agreed remarkably well with

\* Permanent address: Department of Physics, University of Maine, Orono, Maine, USA.

the solution spectrum. The rms (root mean square) error between observed and calculated resonance positions was less than 0.05 ppm (18 Hz). At first sight these results imply that the solution and crystal structure are highly similar. However, the calculations themselves involve several approximations [4]:

- (i) The ring currents, optimized by an iterative procedure to fit experimental NMR spectra and the atomic coordinates of Yeast tRNA<sup>Phe</sup>, were somewhat different than those previously published [5].
- (ii) The low field starting positions for the reversed Hoogsteen and Watson-Crick AU base paired proton resonances were assumed to be different from each other as well as from published values [6].
- (iii) The starting positions for other non Watson-Crick tertiary hydrogen bonds were adjusted to agree with the observed solution spectrum.

One method to confirm these approximations would be to use these optimized ring currents and starting positions and attempt to calculate the solution NMR spectrum of a different tRNA from the three dimensional atomic coordinates of the crystal structure of that molecule. The only problem is that no other tRNA structure has been solved by X-ray crystallographic techniques.

In this paper we present the procedure by which we generated a set of atomic coordinates for E. Coli tRNA<sup>Val</sup> and used these coordinates to calculate the NMR spectrum of this tRNA. This calculation is in very good agreement with the observed NMR spectrum.

## 2. Materials and methods

E. Coli tRNA<sup>Val</sup> was purified as described earlier [11]. NMR spectra were measured in the continuous wave mode (5 s/kHz) or the correlation mode (1 s/kHz) using a Bruker HX360 MHz spectrometer.

The sample temperature was maintained to  $\pm 1^\circ\text{C}$ .

## 3. Atomic coordinates of E. Coli tRNA<sup>Val</sup>

Fig. 1 shows the cloverleaf structure of Yeast tRNA<sup>Phe</sup> and E. Coli tRNA<sup>Val</sup>. Each molecule contains 76 bases and each corresponding helix and loop contains the same number of bases. Furthermore, there are the same number of base pairs in each helical region.

The encircled bases in the Yeast tRNA<sup>Phe</sup> cloverleaf structure are those which are involved in tertiary structure interactions [7–10]. Comparison of these with the encircled bases in the cloverleaf of E. Coli tRNA<sup>Val</sup> shows that every base in Yeast tRNA<sup>Phe</sup> involved in tertiary interactions corresponds to an identical base in E. Coli tRNA<sup>Val</sup> except that the m<sup>2</sup>G 26–A44 interaction in Yeast tRNA<sup>Phe</sup> is reversed for an A26–G44 in E. Coli tRNA<sup>Val</sup>.

These similarities suggest that the overall structure of E. Coli tRNA<sup>Val</sup> should be similar if not identical to that of Yeast tRNA<sup>Phe</sup> and that the same tertiary interactions should occur. With these assumptions a set of atomic coordinates of E. Coli tRNA<sup>Val</sup> were generated by the following procedure. Holding the atomic coordinates of the phosphate–ribose backbone of Yeast tRNA<sup>Phe</sup> [3] fixed, base substitutions were made at each point where they differed in the two structures. After the base substitutions were completed the new tRNA<sup>Val</sup> structure was idealized by minimizing unfavorable stereochemistry and van der Waals interactions. The resulting atomic coordinates were then considered to represent the three dimensional crystal structure of E. Coli tRNA<sup>Val</sup>.

## 4. Calculation of the E. Coli tRNA<sup>Val</sup> NMR spectrum

The primary objective of this endeavour was to check the validity of the optimized ring currents and low field starting positions for the proton resonances of the various types of hydrogen bonded base pairs which were used to calculate the NMR spectrum of Yeast tRNA<sup>Phe</sup> [4]. Therefore, the same ring currents and starting positions calculated for Yeast tRNA<sup>Phe</sup> were employed along with the newly generated atomic coordinates of the E. Coli tRNA<sup>Val</sup> in the calculation of the NMR spectrum of this tRNA. Furthermore, the same computational procedures were used in this calculation.

The upper portion of fig. 2 presents a 360 MHz proton NMR spectrum of E. Coli tRNA<sup>Val</sup> in Mg<sup>2+</sup> measured at 45°. Previously it had been shown that the region between –15 and –11.3 ppm contained 26  $\pm$  1 proton resonances which are attributable to the 20 hydrogen bonded base pair protons in the helices and five ring NH to ring N tertiary interactions [11,12]. The integrated intensity for each region is given in the

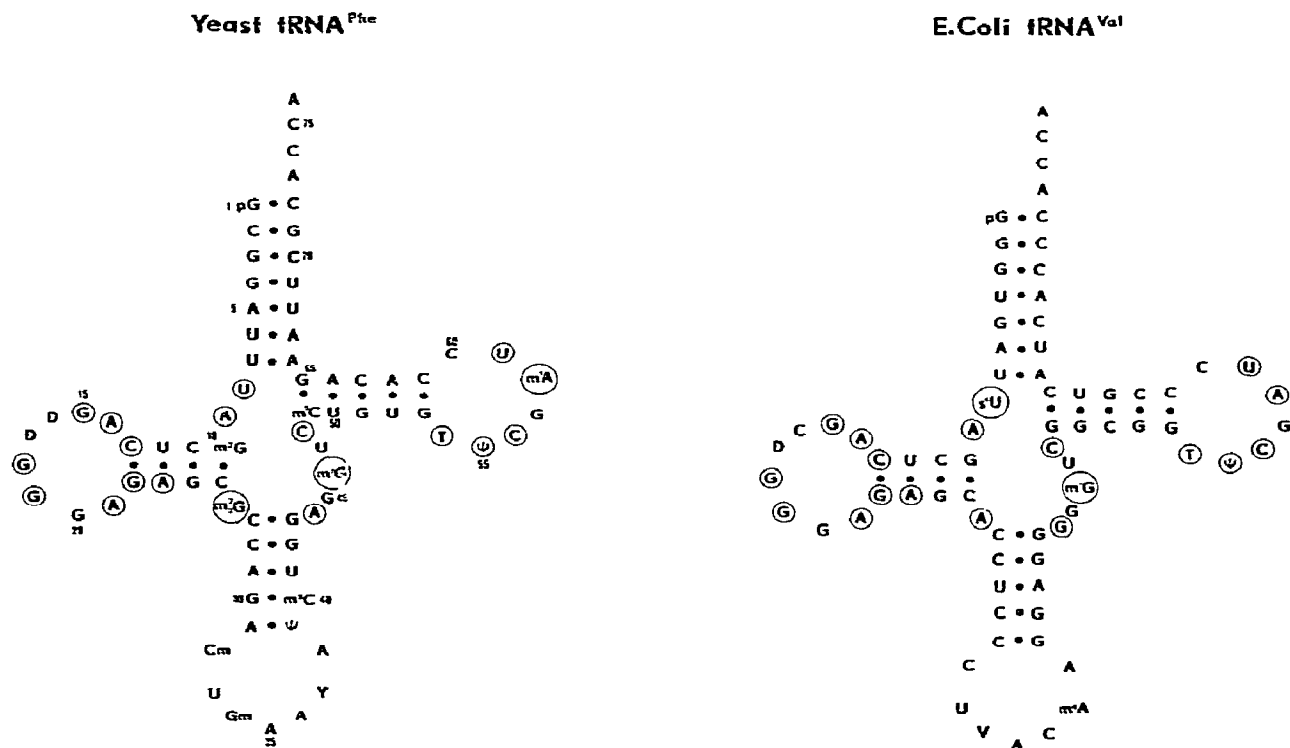


Fig. 1. Cloverleaf structures of Yeast tRNA<sup>Phe</sup> [16–18] and E. Coli tRNA<sup>Val</sup> [19–22]. The encircled bases are those which are involved in tertiary structure interactions [7–10].

table at the side of the figure [11]. The NMR spectrum calculated from the newly generated coordinates and the ring currents discussed above is shown in the lower portion of fig. 2. It is obvious that there is a high degree of similarity between the observed and calculated NMR spectra both with respect to the distribution and position of resonances. The number of resonances calculated for each region of the spectrum is given in the table at the side of the figure. The major differences between the observed and calculated spectra are: (i) in the calculated spectrum, region E, F and G differ by one proton each from the corresponding regions in the observed spectrum; (ii) in region C of the calculated spectrum, 6 resonances are predicted and 6 are found in the observed spectrum but two of the predicted resonances occur at a lower field position than observed.

The positions marked at the bottom of the spec-

trum in fig. 2 are the calculated positions for the resonances specified. The rms error between observed and predicted positions is 0.1 ppm. While the order in which they are listed in a given region may vary depending on the inaccuracies in the atomic coordinates and assumptions used in the calculation of the spectrum it is unlikely that these inaccuracies would be large enough to cause many shifts from one region of the spectrum to another. In the case of Yeast tRNA<sup>Phe</sup> for example when the calculations were repeated on more highly refined coordinates with a crystallographic *R* factor of  $\approx 25$  instead of those with an *R* factor of 39 used initially [3,4] most resonance positions experienced only small changes (unpublished results).

As stated previously, the low field starting positions for the resonances of individual base pair types (see table 1) are the same as those used in the calculation of the NMR spectrum of Yeast tRNA<sup>Phe</sup> [4]. The

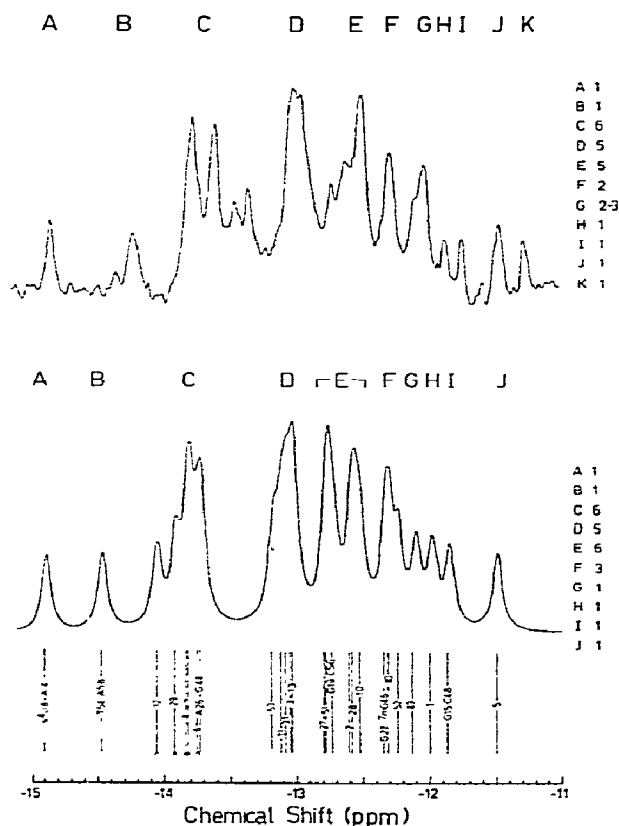


Fig. 2. Upper: 360 MHz proton NMR spectrum of E. Coli tRNA<sup>Val</sup> in H<sub>2</sub>O containing 15mM MgCl<sub>2</sub>, 10mM cacodylate pH 7, 1mM EDTA and 0.1M NaCl. The sample temperature was 45°. The spectrum is an accumulation of 500 sweeps using correlation spectroscopy with a sweep rate of 1 s/kHz. The chemical shift is given in PPM relative to DSS. Lower: Ring current calculated spectrum using coordinates for E. Coli tRNA<sup>Val</sup> and computational procedures described in the text. Secondary structure AU resonances are marked (A) at the bottom of the line.

U8—A14 reversed Hoogstein tertiary base pair in Yeast tRNA<sup>Phe</sup>, however, is replaced by an s<sup>4</sup>U8—A14 base pair in E. Coli tRNA<sup>Val</sup>. The experiments of Reid et al. [11] have demonstrated that conversion of s<sup>4</sup>U to U by removal of the sulfur with cyanogen bromide treatment caused the -14.9 ppm resonance to shift upfield to -14.3 ppm. This suggests that the s<sup>4</sup>U—A reversed Hoogstein hydrogen bonded proton resonance is offset to lower field than a standard AU reversed Hoogstein resonance. Thus using an offset of -15.45 ppm,

Table 1

Base pair type	Low field offset (ppm)
Watson-Crick AU	-14.35
Watson-Crick GC	-13.54
reversed Hoogstein AU	-14.90
reversed Hoogstein s <sup>4</sup> U-A	-15.45
G N <sub>1</sub> H--N <sub>1</sub> A	-14.2
G N <sub>1</sub> H--N <sub>7</sub> G	-14.2
G N <sub>1</sub> H--O <sub>2</sub> C	-12.1

the s<sup>4</sup>U8—A14 resonance is placed at -14.9 ppm.

### 5. Ring current chemical shift tables

Shulman et al. [13] have employed the ring current isoshielding contours of Giessner-Prettre and Pullman [14] together with the predicted structure of an ideal A'-RNA helix [15] to generate a table which allowed one to calculate the ring current shift experienced by a particular hydrogen-bonded base pair proton due to the nearest neighbour base pair above and below. Since the optimized ring current values which we have employed [4] are substantially lower than those used by Shulman et al. [13], we have generated a new set of shifts which are presented in table 2. These shifts were computed with the same program and ring currents used to calculate the shifts experienced by the hydrogen bonded protons of the tRNA itself. The coordinates employed in the calculation, however, were generated from the optimized A-RNA coordinates of Arnott et al. [23] from which a double helix containing all possible base pair combinations was constructed. Since the adenine and guanine rings occupying next nearest neighbour positions also contribute to the shift experienced by a given proton, the shifts arising from these bases in the next nearest neighbour positions are included in parentheses. These new shift values together with the new starting positions for the various types of hydrogen bonded base pair proton resonances (see table 1) allows a more accurate calculation of expected ring current shifts and spectral positions. It should be emphasized, however, that calculations based on these values are expected to be reasonably accurate only for helical portions of the molecule.

Recently Arter and Schmidt [24] published a series

Table 2  
Ring current shifts experienced by ring NH...N hydrogen bonded protons in A-RNA helices <sup>a)</sup>

5' U 0.0				A 0.54 (0.11)	3'
C 0.0				G 0.26 (0.08)	
G 0.1				C 0.04	
A 0.25				U 0.02	
	U	A			
	Ω	Υ			
U 0.08				A 0.0	
C 0.15				G 0.0	
G 0.39 (0.03)				C 0.0	
3' A 0.50 (0.05)				U 0.0	5'
5' U 0.0				A 0.73 (0.1)	3'
C 0.0				G 0.42 (0.07)	
G 0.03				C 0.10	
A 0.09				U 0.05	
	C	G			
	Ω	Ω			
U 0.07				A 0.03	
C 0.13				G 0.0	
G 0.43 (0.05)				C 0.0	
3' A 0.65 (0.08)				U 0.0	

a) All values given in ppm

of tables of ring current shifts experienced by the protons of the four bases when they are in the helical conformation of A-RNA, A'-RNA and B-DNA. Their calculated shifts for the ring NH...ring N hydrogen bonded protons in the A-RNA helix differ by approximately 10% with those presented here in table 2. The differences seem to arise either from our use of optimized ring currents which are approximately 10% different from the ring currents of Giessner-Prettre and Pullman [14] used by Arter and Schmidt [24] or from the different computational procedures employed or a combination of both.

## 6. Expected accuracy of the NMR spectrum calculation

The assumption that the structure of E. Coli tRNA<sub>1</sub><sup>Val</sup> is essentially the same as that of Yeast tRNA<sup>Phe</sup> forms the foundation of the experiments presented above. While this is a reasonable assumption it should be realized that any errors either in the X-ray crystal structure of Yeast tRNA<sup>Phe</sup> or ring currents or offset parameters will probably be carried over to the calculated spectrum for the E. Coli tRNA<sub>1</sub><sup>Val</sup> structure. Therefore, no more

accuracy should be expected from the calculated compared with observed spectra of E. Coli tRNA<sub>1</sub><sup>Val</sup> than was initially present in the Yeast tRNA<sup>Phe</sup> NMR spectra [4]. This is reflected in the rms error of 0.1 ppm between the calculated and observed resonance position in the spectra of E. Coli tRNA<sub>1</sub><sup>Val</sup> as compared to 0.05 ppm between those of Yeast tRNA<sup>Phe</sup>. In this regard it is worth noting that in the calculated spectrum of Yeast tRNA<sup>Phe</sup> the resonance which deviated most from its position in the observed spectrum was that of AU 12 (see resonance at -14.05 ppm, fig. 3a, spectrum b). Likewise in the present calculated spectrum of E. Coli tRNA<sub>1</sub><sup>Val</sup> the resonance for AU 12 is predicted at too low a field position (see fig. 2). Furthermore we pointed out that a small adjustment in the DHU arm could well alter the positions of GC 11, AU 12 and GC 13 bringing the calculated spectrum of Yeast tRNA<sup>Phe</sup> into much closer alignment with the observed spectrum [4]. A similar adjustment in the case of the E. Coli tRNA<sub>1</sub><sup>Val</sup> coordinates might have an analogous effect on the calculated spectrum presented above. Since a number of bases in the DHU arm and loop region are involved in tertiary structure interactions it is entirely possible that: (i) the solution structure in this region is slightly different from that seen in the crystal or (ii) some minor adjustment in the X-ray coordinates might be necessary which would still be within the limits of the electron density contours. Experiments are in progress to discriminate between these two options. A second reason for an increase in the rms error for the calculated versus observed resonance positions in the spectra presented above would be that there are small but real differences between the structures of Yeast tRNA<sup>Phe</sup> and E. Coli tRNA<sub>1</sub><sup>Val</sup>. Since it is difficult to predict what these differences might be we have constrained the structure of E. Coli tRNA<sub>1</sub><sup>Val</sup> to resemble, as closely as possible, that of Yeast tRNA<sup>Phe</sup>. This again could generate a higher RMS error between observed and calculated spectral positions. In spite of this RMS error of 0.1 ppm, it is worth noting that the agreement between the observed and calculated tRNA<sub>1</sub><sup>Val</sup> spectrum, fig. 2, is better than the agreement between the observed and calculated tRNA<sup>Phe</sup> spectra for three of the four structures calculated in fig. 3.

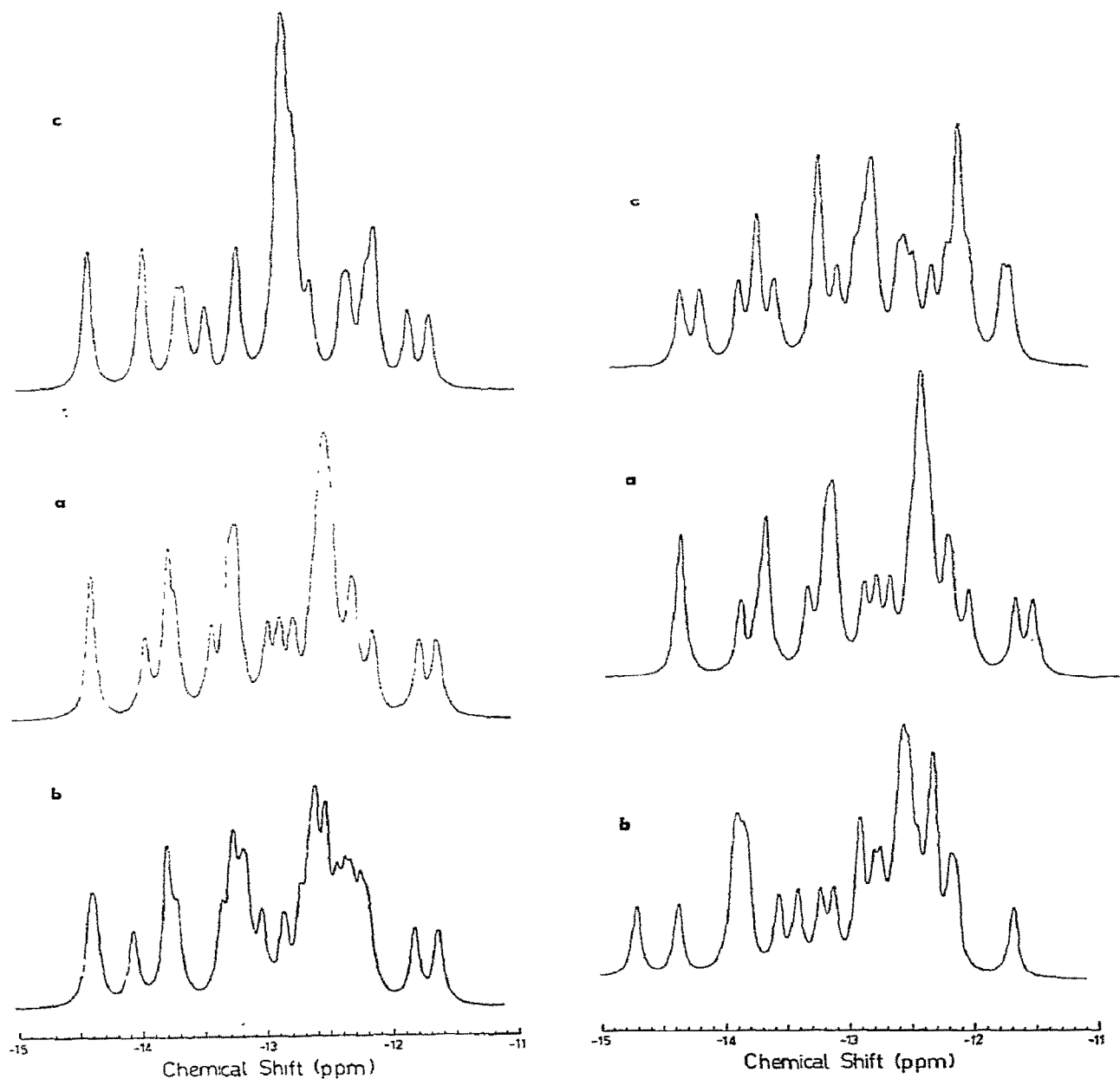


Fig. 3. (a) Spectrum a – computer simulation of the observed Yeast tRNA<sup>Phe</sup> NMR spectrum. Spectrum b – NMR spectrum calculated from the coordinates of Sussman and Kim [3,4]. Spectrum c – NMR spectrum calculated from the coordinates of Quigley et al. [27]. (b) Spectrum a – computer simulation of the observed Yeast tRNA<sup>Phe</sup> NMR spectrum. Spectrum b – NMR spectrum calculated from the coordinates of Stout et al. [28]. Spectrum c – NMR spectrum calculated from the coordinates of the MRC group (personal communication). All calculations were done using the same ring currents and offset parameters [4].

## 7. Sensitivity of spectral calculations to structural differences

Although we have suggested that these calculations are very sensitive to structural changes and thus may be employed to search out correct structures, it is difficult to present a convincing measure of the sensitivity. In figs. 3a and 3b we attempt to provide such a measure by comparing spectra calculated from the coordinates of four partially refined X-ray crystallographic structures of Yeast tRNA<sup>Phe</sup>. The middle spectrum (a) in each figure is a computer simulation of the measured NMR spectrum while spectra b and c are the spectra calculated from coordinates from the orthorhombic (fig. 3a) and monoclinic (fig. 3b) crystal forms. There are obviously significant differences between the spectra calculated from the four sets of coordinates. These differences reflect small variations in bond lengths and angles as well as non-bonded contacts which arise principally from the methods of data refinement [26] and, thus, will certainly decrease as refinement continues. The point which must be emphasized here is that these small differences are readily detectable due to the sensitivity of the calculations.

This comparison clearly illustrates that the approach presented above is sensitive enough to detect small differences in conformations and find the one which corresponds to the correct solution structure. The accuracy of these calculations will increase further as the number of resonance positions calculated also increases. It is, therefore, essential to expand the calculations to include non exchangeable proton resonances, particularly those from methyl and methylene groups which often occur in loop regions of the tRNA. Due to the careful work of Kan et al. [29,30] the assignment of these resonances is relatively straightforward by comparison with hydrogen bonded proton resonances. In addition they can provide more accurate information concerning the conformation of the loop regions as well as the tertiary structure interactions which occur between loop regions (Robillard et al. manuscript in preparation).

## 8. Conclusions

In addition to supporting the optimized ring current values and the low field starting positions presented earlier [4], these results provide rather clear evidence

for a high degree of similarity between the structure of Yeast tRNA<sup>Phe</sup> and E. Coli tRNA<sup>Val</sup> both in the crystal and in solution.

The process of crystallizing biological macromolecules and determining the three dimensional crystal structures from X-ray diffraction techniques is time consuming and not always successful. An alternative approach which may be applicable, at least in the case of tRNA, is the combination of NMR spectroscopy and atomic coordinates generated from a proposed model. The procedures and results presented in this paper and earlier [4] suggest that it may be possible, relying on what has already been learned about the folding of a tRNA molecule from X-ray crystallographic studies, to generate three dimensional models of tRNA whose structures are not known, and by comparing their calculated versus observed NMR spectra, to refine these structures to the correct solution structure. Further experiments are now in progress to refine the present coordinates of Yeast tRNA<sup>Phe</sup> and E. Coli tRNA<sup>Val</sup> by this approach in order to test the feasibility of such a procedure.

## Acknowledgement

The authors wish to thank Professor B. Reid for the samples of E. Coli tRNA<sup>Val</sup> used in these measurements. We also wish to acknowledge the ZWO (Netherlands Foundation for Pure Research) for support of the 360 MHz NMR facility at Groningen. Part of this work has been supported by the grants from US National Institute of Health (CA 15802) and National Science Foundation (GB 40814). JLS is a fellow of the Arthritis Foundation. The efforts of the M.I.T. and the M.R.C. groups to supply us with tapes of coordinates are gratefully acknowledged.

## Note added in proof

Bases G10 and G45 in fig. 1 should be encircled.

## References

- [1] L.A. Davis and G.P. Hess, *J. Mol. Biol.* 82 (1974) 27.
- [2] G.L. Rossi and S.A. Bernhard, *J. Mol. Biol.* 49 (1970) 85.

- [3] J.L. Sussman and S.H. Kim, *Biochim. Biophys. Res. Commun.*, **68** (1975) 89.
- [4] G.T. Robillard, C.E. Tarr, F. Vosman and H.J.C. Berendsen, *Nature* **262** (1976) 363.
- [5] C. Giessner-Prettre and B. Pullman, *Compt. Rend. Acad. Sci. (Paris)* **261** (1965) 2521.
- [6] D.R. Lightfoot, K.L. Wong, D.R. Kearns, B.R. Reid and R.G. Shulman, *J. Mol. Biol.* **78** (1973) 71.
- [7] S.J. Kim, J.L. Sussman, F.L. Suddath, G.J. Quigley, A. McPherson, A.H.J. Wang, N.C. and A. Rich, *Proc. Natl. Acad. Sci. US* **71** (1974) 4970-4974.
- [8] J.D. Robertus, J.E. Ladner, J.T. Finch, D. Rhodes, R.S. Brown, B.F.C. Clark and A. Klug, *Nature* **250** (1974) 546-551.
- [9] G.J. Quigley, A.H.J. Wang, N.C. Seeman, F.L. Suddath, A. Rich, J.L. Sussman and S.H. Kim, *Proc. Natl. Acad. Sci. US* **72** (1975) 4866-4870.
- [10] J.E. Ladner, A. Jack, J.D. Robertus, R.S. Brown, D. Rhodes, B.F.C. Clark and A. Klug, *Proc. Natl. Acad. Sci. US* **72** (1975) 4414-4418.
- [11] B.R. Reid, N.S. Ribeiro, G. Gould, G. Robillard, C.W. Hilbers and R.G. Shulman, *Proc. Natl. Acad. Sci. US* **72** (1975) 2049-2053.
- [12] B.R. Reid and G.T. Robillard, *Nature* **257** (1975) 287.
- [13] R.G. Shulman, C.W. Hilbers, D.R. Kearns, B.R. Reid and Y.P. Wong, *J. Mol. Biol.* **78** (1973) 57-68.
- [14] C. Giessner-Prettre and B. Pullman, *J. Theoret. Biol.* **27** (1970) 87-95.
- [15] S. Arnott, *Progr. Biophys. Mol. Biol.* **22** (1971) 181-213.
- [16] U.L. RajBhandary, S.H. Chang, A. Stuart, R.D. Faulkner, R.M. Hoskinson and H.G. Khorana, *Proc. Natl. Acad. Sci. US* **57** (1967) 751-758.
- [17] U.L. RajBhandary et al., *J. Biol. Chem.* **243** (1968) 598-608.
- [18] F. Nakanishi, N. Furutachi, M. Funamizu, D. Grunberger, and I.B. Weinstein, *J. Am. Chem. Soc.* **92** (1970) 7617-7619.
- [19] M. Yaniv and B.G. Barrell, *Nature* **222** (1969) 278-279.
- [20] F. Harada, F. Kimura and S. Nishimura, *Biochim. Biophys. Acta* **195** (1969) 590-592.
- [21] F. Harada, F. Kimura and S. Nishimura, *Biochemistry* **10** (1971) 3269-3276.
- [22] F. Kimura, F. Harada and S. Nishimura, *Biochemistry* **10** (1971) 3277-3283.
- [23] S. Arnott, D.W.L. Hukins, S.D. Dover, *Biochem. Biophys. Res. Commun.* **48** (1972) 1392-1399.
- [24] D.B. Arter and P.G. Schmidt, *Nucleic Acid Res.* **3** (1976) 1437-1447.
- [25] C.W. Haigh and R.B. Mallion, *Mol. Phys.* **22** (1971) 955-970.
- [26] J.L. Sussman and S.H. Kim, *Science* **192** (1976) 853-858.
- [27] G.T. Quigley, N.C. Seeman, A.H.T. Wang, F.L. Suddath and A. Rich, *Nucleic Acid Res.* **2** (1975) 2329-2341.
- [28] C.D. Stout, H. Mizuno, J. Rubin, T. Brennan, S.T. Rao and M. Sundaralingam, *Nucleic Acid Res.* **3** (1976) 1111-1123.
- [29] L.S. Kan, P.O.P. Ts'o, F. van der Haar, M. Sprinzl and F. Cramer, *Biochem. Biophys. Res. Commun.* **59** (1974) 22-26.
- [30] L.S. Kan, P.O.P. Ts'o, F. van der Haar, M. Sprinzl and F. Cramer, *Biochemistry* **14** (1975) 3278-3291.