February 2001

Calculation of ^{13}C Chemical Shifts in RNA Nucleosides: Structure-^{13}C Chemical Shift Relationships

Paolo Rossi
University of Nebraska - Lincoln

Gerard S. Harbison
University of Nebraska - Lincoln, gharbison1@unl.edu

Published in *Journal of Magnetic Resonance* **151** (2001), 1–8; doi:10.1006/jmre.2001.2350
Published online http://www.idealibrary.com

Submitted December 27, 1999; revised February 19, 2001

Calculation of ^{13}C Chemical Shifts in RNA Nucleosides: Structure-^{13}C Chemical Shift Relationships

Paolo Rossi and Gerard S. Harbison*

Department of Chemistry, University of Nebraska–Lincoln, Lincoln, Nebraska 68588-0304

* Corresponding author — fax (402) 472-9402; email harbison@unlserve.unl.edu

Abstract

Isotropic ^{13}C chemical shifts of the ribose sugar in model RNA nucleosides are calculated using SCF and DFT-GIAO *ab initio* methods for different combinations of ribose sugar pucker, exocyclic torsion angle, and glycosidic torsion angle. Idealized conformations were obtained using structures that were fully optimized by *ab initio* DFT methods starting with averaged parameters from a collection of crystallographic data. Solid-state coordinates of accurate crystal or neutron diffraction structures were also examined directly without optimization. The resulting ^{13}C chemical shifts for the two sets of calculations are then compared. The GIAO-DFT method overestimates the shifts by an average of 5 ppm while the GIAOSCF underestimates the shifts by the same amount. However, in the majority of cases the errors appear to be systematic, as the slope of a plot of calculated vs experimental shifts is very close to unity, with minimal scatter. The values of the ^{13}C NMR shifts of the ribose sugar are therefore sufficiently precise to allow for statistical separation of sugar puckering modes and exocyclic torsion angle conformers, based on the canonical equation model formulated in a previous paper.

Keywords: chemical shift, nucleotide, RNA, *ab initio,* sugar pucker, conformation.

INTRODUCTION

The ^{13}C chemical shift has emerged as a useful probe of the local structure of RNA (*1*) and nucleic acids in general (*2–4*). Carbon chemical shifts are easy to measure, and provide a great deal of qualitative or semi-quantitative structural information, which, combined with more complex multidimensional liquid state NMR techniques, and uniform ^{13}C and ^{15}N labeling (*5*), should be helpful in obtaining accurate structures of RNA oligomers.

Following early semi-empirical studies (*6*), the theoretical study of biomolecular chemical shifts has expanded rapidly. *Ab initio* carbon and proton chemical shifts have been calculated for protein (*7*) and ribonucleoside systems (*4, 8, 9*), with encouraging results. In particular, recent density functional studies of methyl ribofuranosides (*9*) have shown that *ab initio* results can simulate the experimental (*2, 3*) effects of ring pucker on the sugar carbon chemical shifts. Encouraged by the success of this approach, we have undertaken a more general study of the structure-^{13}C chemical shift relationship in RNA nucleosides.

In the accompanying paper (*10*) the four most commonly encountered conformations of crystalline RNA nucleosides and nucleotides were analyzed by CPMAS NMR. The four structures will henceforth be referred to as N-*anti-gg*, N-*anti-gt*, S-*anti-gg*, and S-*syn-gg*. N or S denotes the ribose ring pucker, *syn* or *anti* defines the conformation about the glycosidic bond, whose torsion angle is χ, and *gg* or *gt* is the conformation about the C4′–C5′ exocyclic bond, whose torsion angle is γ. These definitions follow the conventions of Saenger (*11*). Key structural features influence the ribose ^{13}C chemical shifts in a consistent manner. Linear combinations of chemical shifts, or *canonical coordinates*, which give the statistically optimal separation between conformers, were determined. Thus, the S or N sugar puckers are separated along the can1 coordinate

$$\text{can1} = 0.179\delta_{C1'} - 0.225\delta_{C4'} - 0.0585\delta_{C5'} \quad [1]$$

as a function of the ^{13}C chemical shifts of C1′, C4′, and C5′. The *gg* and *gt* orientations about the exocyclic torsion angle γ are separated along the can2 coordinate:

$$\text{can2} = -0.0605(\delta_{C2'} + \delta_{C3'}) - 0.0556\,\delta_{C4'} - 0.0524\,\delta_{C5'} \quad [2]$$

The primary reason for the present work was to examine the accuracy of *ab initio* calculations of ribose carbon chemical shifts. There were other motivations: first, the inability to distinguish between the C2′ and C3′ resonances by simple CP-MAS NMR suggested that perhaps a better canonical model could be found if these two carbons could be assigned by *ab initio* methods. Second, in order to gain a complete picture of structure–chemical shift relationships in RNA, rare conformations need to be examined as well. Unfortunately, there are very few known structures outside the four main families of conformations cited above. This limited sample set would render any empirical study statistically meaningless; however,

should *ab initio* methods prove to be reliable, a theoretical study of the chemical shifts of these more uncommon conformations might well allow them to be identified in RNA oligomers. Finally, the earlier studies of Ghose *et al.* (*8*) and Dejaegere and Case (*9*) were limited in scope and by computational or structural difficulties, and Xu *et al.* (*4*) did not report or utilize the values of C1′ and C2′, which are essential components of our statistical analysis.

Two sets of calculations are presented and compared. First, gas-phase-type calculations are conducted using the four conformations N-*anti-gg*, N-*anti-gt*, S-*anti-gg*, and S-*syn-gg* constructed using average crystallographic bond lengths, angles, and three-bond dihedral angles (*12*), and then fully optimized by *ab initio* DFT methods. A second set of calculations looks at the real solid-state conformation of crystalline nucleotides, using the most accurate crystal structure coordinates available; in the case of X-ray data only the hydrogen positions are minimized while the neutron diffraction structures are used directly without any optimization.

METHOD AND RESULTS

Calculations were performed on either UNIX or onWindows NT workstations running Gaussian 98 (*13*). Model nucleosides representative of the N-*anti-gg*, N-*anti-gt*, S-*anti-gg*, and S-*syngg* conformations were constructed with InsightII (*14*) using average geometric parameters from the literature (*12*). It was found convenient to use a simplified nucleoside model with a pyrimidine base (Fig. 1). This allowed us to maintain the appropriate C1′–N3 bond order and the correct three-dimensional requirements necessary to generate a *syn* or *anti* conformation about the glycosidic bond. Optimization routines were conducted using the B3LYP functional and the 6-31G(d) or the 6-31+G(d) basis sets (*15–19*). Both basis sets gave essentially identical results. Coordinates of the optimized nucleoside structures used are given in the supplementary materials.

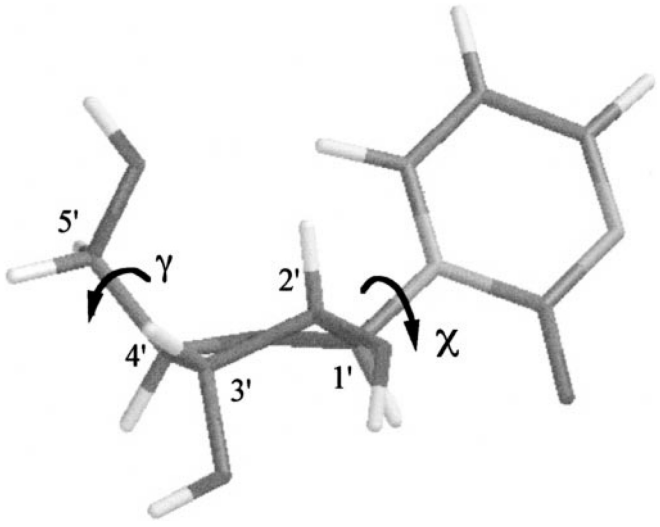

Figure 1. Optimized model nucleoside structure with pyrimidine-type base (S-*anti-gg*).

For calculations direct from crystal structures, the following crystal structures were used: for S-*anti-gg*, 5-hydroxyuridine (*19*); for N-*anti-gg*, 5-methoxyuridine (*20*); for N-*anti-gt*, adenosine 5′-phosphate (*21*); and for *S-syn-gg*, 7-methyl-8-thioxo-7,8-dihydroguanosine monohydrate (*22*). All of the chosen crystal structures have a C–C bond esd between 0.001 and 0.005 Å and an *R* factor of 3% or lower. In this case only the hydrogen positions were optimized; the heavy atom coordinates were left unchanged. The adenosine structure was used directly for the chemical shift calculations without any minimization since the proton positions had been determined with a high degree of accuracy by neutron diffraction.

An initial exploratory survey of GIAO NMR calculations (*23–27*) were conducted with a limited number of SCF and DFT methods. The following basis set and methods were tested: HF/6-311+G(2d,p), B3LYP/6-311+G(2d,p), and B3PW91/6-311+ G(2d,p) (Figure 2). ^{13}C chemical shifts were referenced to TMS, whose T_d symmetric structure was optimized in the 6-31G(d) basis. The chemical shift of TMS was calculated at each level of theory. The usual relationship was used for referencing: $\delta_{iso} = \sigma_{TMS} - \sigma_{iso}$, where σ_{TMS} =182.5994 ppm for B3LYP/6-311+G(2d,p), 185.5643 ppm for B3PW91/6-311+G(2d,p), and 192.5939 ppm for HF/6-311+G(2d,p). For 6-31+G(d) optimizations the reference carbon shift was 182.4075 ppm for B3LYP/6-311+G(2d,p), essentially identical to the 6-31G(d) value.

The slope of the experimental vs calculated chemical shift plot and its standard error values were examined for each level of theory, as in the example shown in Figure 2. These were used to evaluate the accuracy of the calculations. Perfect accuracy would obviously generate a diagonal line. Even though the chemical shift range is limited (50 to 100 ppm), the consistency of the results through the range of conformations studied is encouraging.

Both the GIAO-B3LYP/6-311+G(2d,p) and the GIAO-B3PW91/6-311+G(2d,p) combinations gave very similar results. For the average structure, the slope averaged over all the conformations studied was 1.09 ± 0.04 (mean ± 1 s.d.), respectively, for B3PW91/6-311+G(2d,p) and 1.10 § 0.04 for B3LYP/6-311+G(2d,p). Excessive scatter in the calculated values for the S-*syn-gg* average conformation causes the slope to be dramatically different.

In the calculations using the crystal structure coordinates the situation is considerably improved (slope = 1.02 ± 0.02) since there is no anomalous scatter in the calculated chemical shift values of the S-*syn-gg* conformation. Given the similarity between the two DFT methods, only calculations using B3LYP/6-311+G(2d,p) were pursued further.

The Hartree–Fock (SCF) method was quickly tested. A slope of 0.80 was found and this value was deemed unacceptable. Therefore SCF calculations were abandoned and will be briefly mentioned in the Discussion section. Attempts to reduce

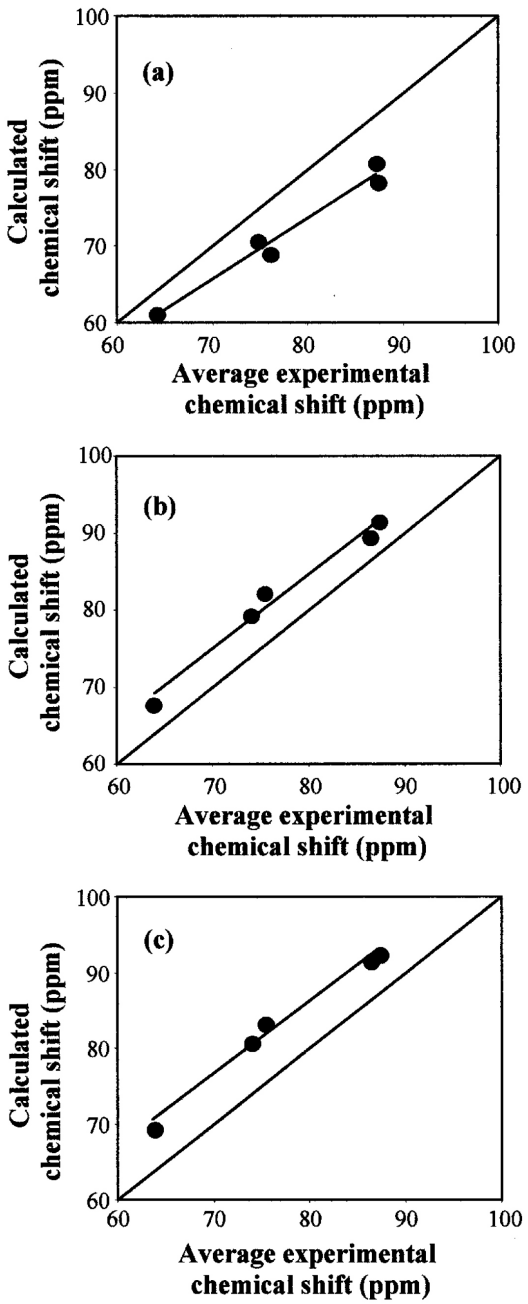

Figure 2. Plot of calculated vs experimental ^{13}C NMR chemical shift of the average equilibrium S-*anti-gg* coordinates calculated by different methods: (a) HF/6-311+CG(2d,p) (slope = 0.80 ± 0.08), (b) B3LYP/6-311CG(2d,p) (slope = 0.98 ± 0.07), (c) B3PW91/6-311+CG(2d,p) (slope = 0.96 ± 0.08).

the basis set size were also made and quickly dismissed as the slope deviates from unity rapidly with a reduction in the basis set.

Structural parameters for the fully optimized average structures, χ, γ, and P, are listed in Table 1. A full comparison between the bond lengths, two-bond angles, and three-bond dihedral of the crystal structures, and the average optimized structure is also attached in the supplementary materials. The calculated chemical shifts for the *S-anti-gg, S-syn-gg, N-anti-gg,* and *N-anti-gt* structures are listed in Table 2 for both the average and the crystalline coordinates. The difference in parts per million between the calculated shift and the experimental average shift of the same structure type is given in parentheses for each chemical shift. On average, computations yielded ^{13}C chemical shift values that were systematically different from experiment by 4.23, 6.17, and 4.88 ppm for the crystal structure calculation using B3LYP/6-311+G(2d,p), the average structure using B3LYP/6-311+G(2d,p), and the average structure with the B3PW91/6-311+G(2d,p) method, respectively.

In order to visualize the trends in chemical shift as a function of structure, the differences listed above were used to correct the calculated chemical shifts, and bring them closer to the experimental values for more meaningful comparison. Examples of these corrected chemical shifts are shown in Figure 3. All data, including the anomalous S-*syn-gg* average conformation, were included in computing the systematic differences.

TABLE 1 Parameters of Average, Fully Optimized RNA Nucleoside Structures[a]

Type	χ	γ	δ	Pseudorotation phase angle
S-*anti-gg*	–130.0 (–ac)[b]	50.96	148.68	161.9 (^{2}E)[c]
S-*syn-gg*	63.3 (+sc)	52.51	146.99	166.5 (^{2}E)
N-*anti-gg*	–163.6 (–ap)	51.96	85.08	12.20 (^{3}E)
N-*anti-gt*	–171.2 (–ap)	–174.64	84.22	14.9 (^{3}E)

Note. Angles are given in degrees.
[a] Type refers to structure type as described in the text.
[b] Glycosidic angle range.
[c] Ring puckering type as defined by Saenger (*9*).

TABLE 2 Calculated ^{13}C Chemical Shift of RNA Sugar Carbons[a]

Type	C1′	C2′	C3′	C2′ + C3′	C4′	C5′
Average equilibrium coordinates B3PW91/6-311+G(2d,p)						
S-anti-gg	91.66(4.47)	82.35	79.42	161.77(12.8)	89.68(3.48)	67.92(4.29)
S-syn-gg	103.36(14.7)	70.17	79.00	149.17(1.61)	95.81(8.69)	66.60(2.97)
N-anti-gg	96.42(3.97)	79.48	72.66	52.14(8.60)	85.78(3.15)	63.61(3.93)
N-anti-gt	97.97(5.24)	78.13	77.25	155.39(7.99)	86.03(2.00)	69.08(5.35)
Average equilibrium coordinates B3LYP/6-311+G(2d,p)						
S-*anti-gg*	92.81(5.62)	83.68	80.85	164.53(15.5)	91.57(5.37)	69.29(5.66)
S-*syn-gg*	104.94(16.3)	71.21	80.48	151.68(0.90)	97.91(10.8)	67.91(4.28)
N-*anti-gg*	97.70(5.25)	80.88	73.61	154.48(10.9)	87.36(4.73)	64.81(5.13)
N-*anti-gt*	99.21(6.48)	79.44	78.49	157.93(10.5)	87.61(3.60)	70.49(6.80)
Crystalline coordinates B3LYP/6-311+G(2d,p)						
S-*anti-gg*	91.22(4.02)	81.41	77.69	159.1(10.1)	89.20(3.00)	67.15(3.52)
S-*syn-gg*	93.42(4.80)	74.47	78.44	152.9(2.12)	92.33(5.20)	68.30(4.66)
N-*anti-gg*	94.20(1.75)	82.14	71.47	153.6(10.1)	89.20(4.71)	67.15(0.84)
N-*anti-gt* 9	8.13(5.40)	81.24	77.03	158.3(10.9)	90.69(6.67)	70.49(6.76)

[a] The chemical shifts are in ppm from the absolute shift of TMS. Type refers to structure type as described in the text. Each value in parentheses is the magnitude of the difference in ppm between the calculated shift and the population average experimental shift value for the corresponding type. The B3LYP/6-311+G(2d,p) shifts are not scaled.

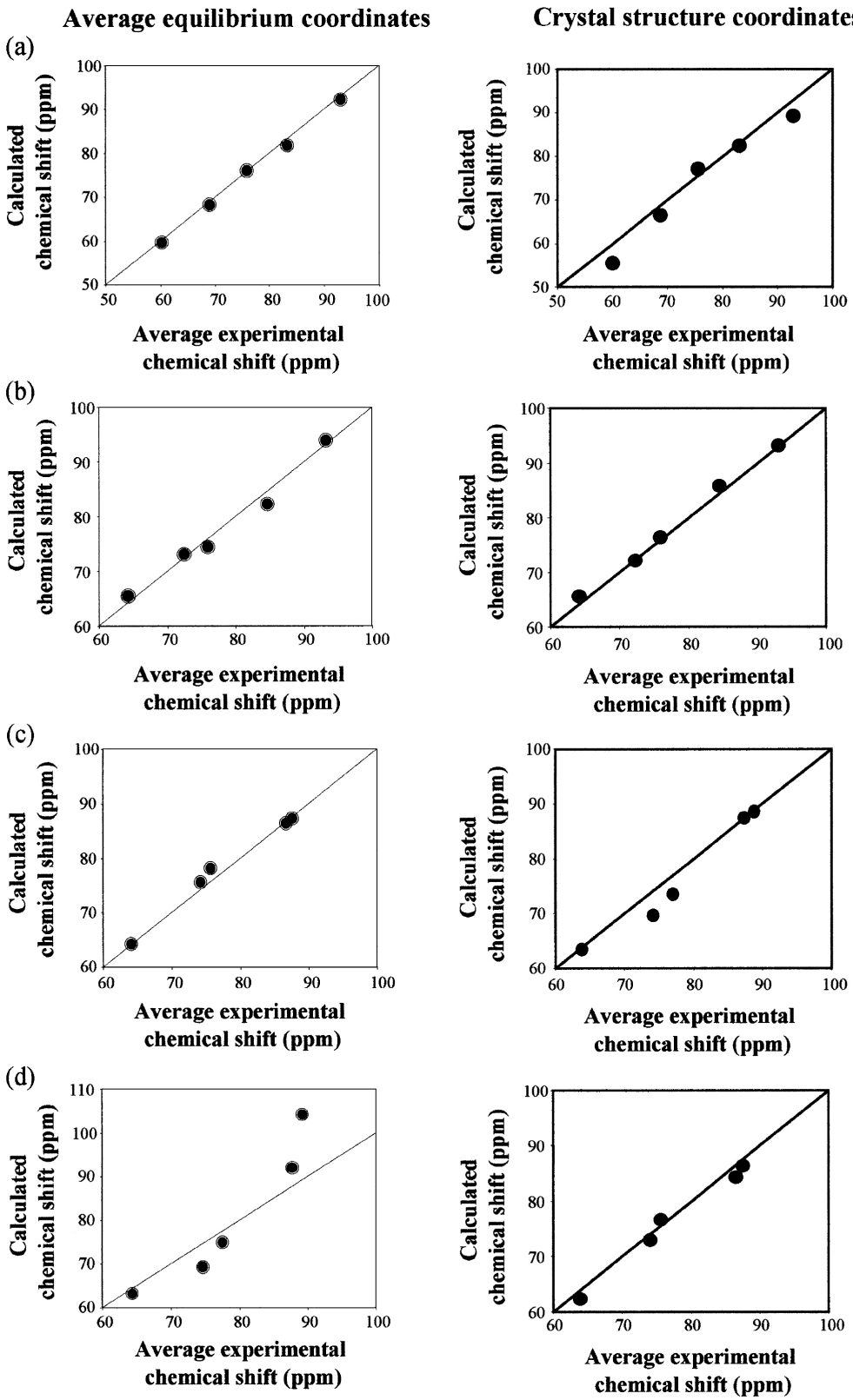

Figure 3. Plot of calculated (B3LYP/6-311CG(2d,p)) vs experimental ^{13}C NMR chemical shifts after correction. The structures used were (a) N-*anti-gg*, (b) N-*anti-gt*, (c) S-*anti-gg*, and (d) S-*syn-gg*.

The corrected chemical shifts were used to compute the canonical coordinates can1 and can2 (Figure 4). Such treatment of the data is useful for visualization purposes and the offset data itself still statistically meaningful. The offset in the chemical shift clearly produces an offset in the positions on the plot. Nonetheless, the differences between the values of can1 and can2 that separate N and S ring pucker and the *gg*-*gt* exocyclic angle are maintained. A visual idea of group classification may be gained by examining the midpoint of the group means that appear as solid and dashed lines in the can1–can2 plot (Fig. 4). However, a statistically more accurate classification can be done by conducting an *F*-distribution test (see the *Statistical Analysis* section).

In order to interpret the *F*-distribution analysis the following guideline is useful: a lower percent confidence indicates closer proximity to the distribution midpoint, and lack of overlap between high percent confidence levels of two groups is a measure of statistical separation. At this stage it should also be pointed out that the small number of experimental observations of the *gt* and *gg* types makes their distribution very wide and therefore, in a rigorous statistical approach the *gg*–*gt* classification along can2 would be considered problematic. Nonetheless, in all the cases the N-*anti*-*gg* and N-*anti*-*gt* types fall within their 75 and 90% *gg* and *gt* confidence contours, respectively. Even though, visually, the N-*anti*-*gt* point crosses over the midpoint of cluster means, the classification is still correct. Continuing on, the N-*anti*-*gg* falls within the same range as the experimental value. The S-*anti*-*gg* belongs to the S group within the 95% confidence and both N-*anti*-*gg* and N-*anti*-*gt* clearly belong to the N group within the 90% confidence interval. The anomalous chemical shift values found in the case of the average equilibrium S-*syn*-*gg* are such that the structure cannot be classified as either S nor N. However, calculations on the crystalline coordinates clearly classify it as within the S grouping. The classification on the basis of confidence ellipsoids is identical for both the B3LYP and the B3PW91 calculations.

Statistical Analysis

The goal of this analysis is to quantitatively assign the position of any set of chemical shifts with respect to the experimentally determined can1 and can2 clusters (*10*). Two basic techniques can be used: (1) midpoint of cluster means and (2) *F*-distribution test (*28*). The first method can be used to obtain a rough classification based on visual inspection of the canonical plot in Fig. 4. The dashed line is the midpoint of cluster means of the N-*gg* and N-*gt* groups. The solid line is the midpoint of cluster means of the N and S groups. For example, if the can2 score obtained from a set of C2′ + C3′, C4′, and C5′ chemical shifts is below about –16.8 ppm, then we can qualitatively say that the structure is N-*gt*, and so on. For the *gg* and *gt* clusters the midpoint of cluster means is obtained using the equation

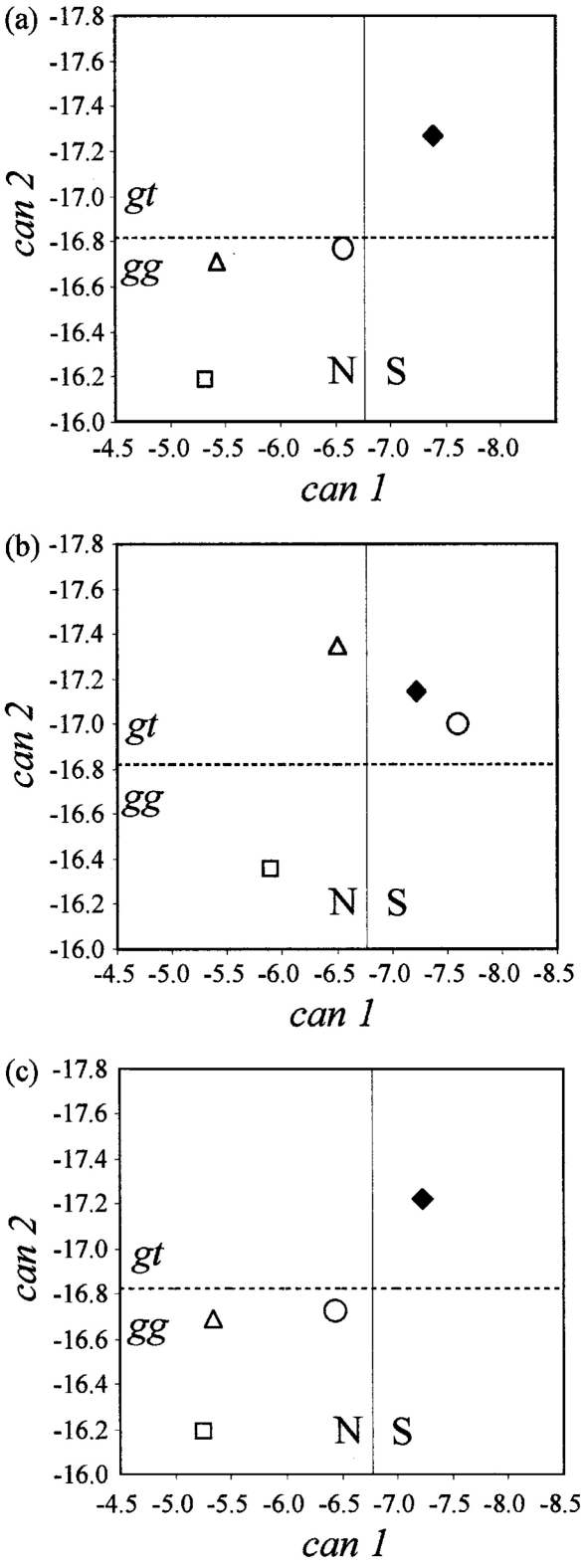

Figure 4. Plot of canonical variables, can2 vs can1. The solid line is the midpoint of the cluster means for the N and S groups at can1D–6.77, the dashed line is the midpoint of the *gg* and *gt* cluster means can2D–16.82. The points on the plot are the coordinates from the canonical scores of the calculated positions. □, N-*anti*-*gg*; ○, S-*syn*-*gg*; Δ, N-*anti*-*gt*; ♦, S-*anti*-*gg*.

$$\hat{m} = 1/2(\bar{y}_{gt} + \bar{y}_{gg}) \quad [3]$$

with $\bar{y}_{gt} = \hat{\alpha}_{gt}$ and $\bar{y}_{gg} = \hat{\alpha}_{gg}$, where

$$\bar{x}_{gt} = \{\, \bar{\delta}''_{(2+3)/2,Ngt}, \bar{\delta}''_{4,Ngt}, \bar{\delta}''_{5,Ngt} \,\} \quad [4]$$

and $\hat{a}$ is the can2 vector. The same analysis is made for the N and S clusters.

The *F*-distribution makes it possible to quantitatively position any set of ribose chemical shifts with respect to the confidence ellipsoid of each experimentally determined cluster of can1 and can2 scores. In principle, the *F*-distribution is the multidimensional version of the univariate student-*t* test. The *p*-dimensional normal distribution of *n* observations is given by

$$p(n-1)/(n-p)F_{p,n-p}(\alpha), \quad [5]$$

where α is the $100(1-\alpha)\%$ confidence region. This number is then compared to

$$n(\bar{x}-\bar{\mu})'S^{-1}(\bar{x}-\bar{\mu}), \quad [6]$$

where $\bar{\mu}$ is the vector of chemical shifts to be tested, S is the covariance matrix for the experimental cluster in exam, and $\bar{x}$ has the same meaning as above. If the condition

$$n(\bar{x}-\bar{\mu})'S^{-1}(\bar{x}-\bar{\mu}) < p(n-1)/(n-p)F_{p,n-p}(\alpha) \quad [7]$$

is met, then the $\bar{\mu}$ data set is within the confidence ellipsoid. The methods that were just outlined here were used to test the accuracy of the calculated chemical shifts with respect to experiment (*10*). When appropriate, cross-testing was performed in order to rule out the possibility that a structure may belong to more than one group.

DISCUSSION

Two sets of ^{13}C chemical shifts of ribose in nucleosides were calculated for fully optimized average coordinates and for crystal structure coordinates. Comparison to the data of Xu and coworkers (*4*) is really not possible since those workers did not report their values of C1′ and C2′ and because their choice of a chemical shift reference is not specified. Assuming that the same (correct) value was used as in the present work, our results are not in agreement. This is likely because our results indicate that previous workers omitted prior optimization of the structures, which our results demonstrate can be very important. Even though the optimization process appears to modify the starting bonds and dihedral angles only slightly (with respect to the average crystalline coordinates) the effects on the calculated chemical shifts can be dramatic, as in the case of S-*syn*-*gg*. Although the reasons for the large discrepancy between the experimental and calculated values C1′ and C4′are to be traced back to the optimization process, the precise culprit has not been pinpointed. The optimized structures are found to have the pseudorotation angle in the same range as the starting coordinates. Bond lengths are in most cases within the X-ray structure tolerances, with a few bonds longer by about 0.02 Å. Angles around the ribose ring are for the most part identical within 3°. Some large discrepancies (8–15°) in the v_0 and v_1 dihedral angles were noted in both S pucker types of the minimized conformations when compared to the X-ray coordinates. This did not appear to affect the calculated values of C1′ and C4′ in the S-*anti*-*gg* conformation.

Operating within the narrow constraints of Gaussian 98, a few basis sets and methods were investigated. The quality of the calculated ^{13}C NMR shifts was established by comparison with experiment as shown in Figures 2 and 3. Chemical shifts trends are clearly visible in Figure 3. The N-type ribose sugar presents a set of chemical shifts that are widely spaced in the range 50 to 100 ppm. Conversely, the S-type sugar has a very close pair of resonances at the C1′ and C4′ and another pair formed by C2′ and C3′. Distinction between the N-*gg* and and N-*gt* is less obvious. In nucleoside systems the inclusion of correlation is necessary to obtain NMR shifts of carbons C1′ and C4′in line with the remaining parts of the molecule, as a result of the electron lone pairs on O4′. Beyond that, either B3LYP or B3PW91 gives about the same results when used in combination with the 6-311+G(2d,p) basis set. In order to properly compare calculated and experimental results a scaling factor of some type needs to be applied. Even though vibrational averaging, bulk susceptibility, and solvation effects as the underlying reasons for this discrepancy have been studied and discussed at length (*29–31*), a consensus set of rules for correction has not been clearly established. In order to correct the chemical shift offset the average systematic deviation between the calculated structures and the experimental was subtracted from the calculated value. This deviation was found to be between 6.17 ppm for B3LYP/6-311+G(2d,p) with the average equilibrium structures and 4.22 ppm with B3LYP/6-311+G(2d,p) using X-ray coordinates. The 4.22-ppm deshielding determined by B3LYP/6- 311+G(2d,p) with respect to experimental data is slightly lower than that previously reported for nucleosides (*9*). It is also important to note that for the crystal structure calculations there is not really a "proper" way to reference the heavy atoms to TMS as their coordinates have not been minimized. However, the values were found to be very close to other DFT values of the optimized structures. Even though a separation along can1 and can2 is clearly apparent irrespective of scaling, a more complete analysis requires scaling of the calculated shift.

The next issue to be examined in this work was the assignment of measured values for the C2′and C3′ resonances and their possible impact on the canonical model. The calculations clearly indicate a consistent upfield shift in the *endo* carbon in the case of N-*anti*-*gg* (about 7 ppm for C3′-*endo*) and S-*anti*-*gg* (about 9 ppm for C2′-*endo*). The difference in the measured values for C2′ and C3′ values for N-*anti*-*gg* is about 9 ppm, and the calculated values of the ^{13}C chemical shift are also different by this amount, so ordering the resonances and reexamining the data is very tempting. However, the mea-

sured C2′–C3′ chemical shift difference in the case of the remaining three structures in the range 3–3.5 ppm and the inversion of the values of C2′ and C3′ in going from S to N is not very apparent. Even so, the C2′and C3′ were ordered on the basis of magnitude and the experimental data reanalyzed. Unfortunately, these results confirm the hypothesis in the accompanying paper (*10*) that the ability to assign C2′and C3′ does not appear to give any additional structural information.

The final step was the assignment of the set of calculated NMR shifts to a particular conformation based on the *F*-distribution test. In the average minimized structure, the midpoint of cluster means is crossed in the case of N-*anti*-*gt* and S-*syn*-*gg*. The N-*anti*-*gt* appears to fall under the N-*anti*-*gg* after crossing the *gt*–*gg* group borderline and the S-*syn*-*gg* crosses the S–N border into the N-*anti*-*gg* sector. However, under the conditions of the quantitative *F*-distribution criterion, three out of four calculated conformations were found to belong to the right group at the 95% confidence or better. The classification results are identical for both the GIAO-B3PW91 and the GIAO-B3LYP chemical shifts. Much to our satisfaction, the crystal structures yielded widely separated and well-defined groups. Clearly, the apparent discrepancy between analysis based on the midpoint of cluster means, and the results of the *F*-distribution test, raises serious doubts about the applicability of sophisticated statistical treatments when low-accuracy scaling methods are employed. It is likely that the scaling issue will have to be addressed using a full vibrational average of the TMS chemical shift, since the barrier to rotation of methyl groups in the reference compound is low (on the order of a few hundred centimeters^{-1}), while the chemical shift of the fully eclipsed, fourth-order saddle point obtained by rotating all methyls by 60° is calculated to be 5 ppm from the minimum energy conformation. Significantly, this value brings the chemical shifts into almost perfect agreement with experimental data. A full vibrational and torsional average of the chemical shift of TMS would probably be helpful in clarifying the offset between theory and experiment.

CONCLUSIONS

A protocol for accurate calculation of isotropic carbon chemical shifts of the ribose sugar in RNA nucleosides was presented. Unconstrained minimization of average crystal structure coordinates with the standard B3LYP/6-31+G(d) method gives, generally, reliable starting structures. However, in the small sample tested, direct chemical shift calculation from crystalline coordinates, optimizing only hydrogen positions, appears to give consistently good results through the entire range of conformations. DFT-GIAO with either B3LYP/6-311+G(2d,p) or B3PW91/ 6-311+G(2d,p) appears to be ideally suited for inexpensive calculation of NMR properties of the entire nucleoside system, while the lack of electron correlation limits the applicability range of even simpler SCF methods.

The C2′ and C3′ resonances were ordered based on the magnitude of the calculated shifts and the measured data were reexamined. Unfortunately, this new information does not give any new insight into the structure–chemical shift relationship, nor a better separation of the *gg* and *gt* groups along can2. The calculations we conducted show excellent agreement between theory and experimental findings. We are currently elucidating vibrational and torsional effects on the ^{13}C NMR tensor in nucleoside structures and examining uncommon nucleoside motifs. These studies should give, we hope, a more complete view of structure–chemical shift relationship in RNA.

ACKNOWLEDGEMENT

This research was supported by NSF Grant MCB-9604521.

REFERENCES

1. M. Ebrahimi, C. Rogers, and G. S. Harbison, ^{13}C chemical shifts and the determination of DNA and RNA structure, in *Spectroscopy of Biological Molecules: Modern Trends* (P. Caramona, R. Navarro, and A. Hernanz, Eds.), Kluwer Academic, Dordrecht, The Netherlands (1997).

2. R. A. Santos, P. Tang, and G. S. Harbison, Determination of the DNA sugar pucker using ^{13}C NMR spectroscopy, *Biochemistry* **28,** 9372–9378 (1989).

3. P. Tang, R. A. Santos, and G. S. Harbison, Two-dimensional solid-state NMR studies of the conformation of oriented A-DNA, *Adv. Magn. Reson.* **13,** 225–255 (1989).

4. X. Xu, W. A. K. Chiu, and S. C. F. Au-Yeung, Chemical shift and structure relationship in nucleic acids: Correlation of backbone torsion angles γ and *α* with ^{13}C chemical shift, *J. Am. Chem. Soc.* **120,** 4230–4231 (1994).

5. E. P. Nikonowicz, A. Sirr, P. Legault, F. Juker, L. M. Baer, and A. Pardi, Preparation of ^{13}C and ^{15}N labelled RNAs for heteronuclear multidimensional NMR studies, *Nucl. Acids Res.* **20,** 4507–4513 (1992).

6. C. Giessner-Prettre and B. Pullman, Quantum mechanical calculations of NMR chemical shifts in nucleic acids, *Q. Rev. Biophys.* **20,** 113–172 (1987).

7. A. C. de Dios, J. G. Pearson, and E. Oldfield, Secondary and tertiary structural effects on protein NMR chemical shifts: an ab initio approach, *Science* **256,** 234–237 (1993).

8. R. Ghose, J. P. Marino, K. B. Wiberg, and J. H. Prestegard, Dependence of ^{13}C chemical shifts on glycosidic torsional angles in ribonucleic acids, *J. Am. Chem. Soc.* **116,** 8827–8828 (1994).

9. A. P. Dejaegere and D. A. Case, Density functional study of ribose and deoxyribose chemical shifts, *J. Phys. Chem.* **102,** 5280–5289 (1998).

10. M. Ebrahimi, P. Rossi, C. Rogers, and G. S. Harbison, Dependence of ^{13}C NMR chemical shifts on conformations of RNA nucleosides and nucleotides, *J. Magn. Reson.* **150,** 1–9 (2001).

11. W. Saenger, "Principles of Nucleic Acid Structure," Springer-Verlag, New York (1984).

12. A. Gelbin, B. Schneider, L. Clowney, S. Hsieh, W. K. Olson, and H. M. Berman, Geometric parameters in nucleic acids: sugar and phosphate constituents, *J. Am. Chem. Soc.* **118,** 519–529 (1996).

13. M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, V. G. Zakrzewski, J. A. Montgomery, Jr., R. E. Stratmann, J. C. Burant, S. Dapprich, J. M. Millam, A. D. Daniels, K. N. Kudin, M. C. Strain, O. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, C. Adamo, S. Clifford, J. Ochterski, G. A. Petersson, P. Y. Ayala, Q. Cui, K. Morokuma, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. Cioslowski, J. V. Ortiz, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. Gomperts, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, C. Gonzalez, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, J. L. Andres, C. Gonzalez, M. Head-Gordon, E. S. Replogle, and J. A. Pople, Gaussian, Inc., Pittsburgh, PA (1998).

14. InsightII Version 98.0 from MSI San Diego, CA.

15. A. D. Becke, Density-functional thermochemistry. III. The role of exact exchange, *J. Chem. Phys.* **98,** 5648–5652 (1993).

16. C. Lee, W. Yang, and R. G. Parr, Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density, *Phys. Rev. B* **37,** 785–789 (1988).

17. B. Miehlich, A. Savin, H. Stoll, and H. Preuss, Results obtained with the correlation energy density functionals of Becke and Lee, Yang and Parr, *Chem. Phys. Lett.* **157,** 200–206 (1989).

18. J. A. Pople, G. W. Trucks, K. Raghavachari, L. A. Curtiss, and C. Jones, Gaussian-1 theory of molecular energies for second-row compounds, *J. Chem. Phys.* **93,** 2537–2545 (1990).

19. U. Thewalt and C. E. Bugg, Crystal structure of 5-hydroxyuridine, *Acta Crystallogr. Sect. B* **29,** 1393–1398 (1973).

20. G. I. Birnbaum, W. J. P. Blonski, and F. E. Hruska, Structure and conformation of the anticodon nucleoside 5-methoxyuridine in the solid state and in solution, *Can J. Chem.* **61,** 2299–2304 (1983).

21. T. F. Lai and R. E. Marsh, Refinement of the crystal structure of adenosine 50-phosphate, *Acta Crystallogr. Sect. B* **28,** 1982–1979 (1972).

22. S. B. Larson, E. M. Henry, G. D. Kini, and R. K. Robins, Structure of 7-methyl-8-thioxo-7,8-Dihydroguanosine monohydrate, *Acta Crystallogr. Sect. C* **46,** 506–508 (1990).

23. R. Ditchfield, Self-consistent perturbation theory of diamagnetism. I. A gauge-invariant LCAO method for N.M.R. chemical shifts, *Mol. Phys.* **27,** 789–807 (1974).

24. K. Wolinski, J. F. Hinton, and P. Pulay, Efficient implementation of the gauge-independent atomic orbital method for NMR chemical shift calculations, *J. Am. Chem. Soc.* **112,** 8251–8260 (1990).

25. V. G. Malkin, O. L. Malkina, M. K. Casida, and D. R. Salahub, Nuclear magnetic resonance shielding tensor calculated with sum-over-states density functional perturbation theory, *J. Am. Chem. Soc.* **116,** 5898–5908 (1994).

26. G. Schreckenbach and T. Ziegler, Calculation of NMR shielding tensor using gauge-including atomic orbitals and modern density functional theory, *J. Phys. Chem.* **99,** 606–611 (1995).

27. J. P. Perdew, K. Burke, and Y. Wang, Generalized gradient approximation for the exchange-correlation hole of a many-electron system, *Phys. Rev. B* **54,** 16,533–16,539 (1996).

28. R. A. Johnson and D. A. Wichern, *Applied Mulivariate Statistical Analysis,* 3rd ed., Prentice–Hall, Englewood Cliffs, NJ (1992).

29. D. B. Chesnut, The ab initio computation of nuclear magnetic resonance chemical shielding, in *Reviews in Computational Chemistry,* Vol. 8, (K. B. Lipkowitz and D. B. Boyd, Eds.), VCH, New York (1996).

30. C. J. Jameson, Gas-phase NMR spectroscopy, *Chem. Rev.* **91,** 1375–1395 (1991).

31. A. C. de Dios and C. J. Jameson, The NMR chemical shift: insight into structure and environment, in *Annual Reports on NMR Spectroscopy,* (G. A. Webb, Ed.), Vol. 29, pp. 1–69, Academic Press, London (1994).