Journal of Magnetic Resonance **150,** 1–9 (2001)
doi:10.1006/jmre.2001.2314, available online at http://www.idealibrary.com on IDEAL®

Dependence of ^{13}C NMR Chemical Shifts on Conformations of RNA Nucleosides and Nucleotides

Mohsen Ebrahimi, Paolo Rossi, Christopher Rogers,* and Gerard S. Harbison[1]

*Department of Chemistry and *Department of Mathematics and Statistics, University of Nebraska at Lincoln, Lincoln, Nebraska 68588-0304*

Received December 27, 1999; revised October 26, 2000; published online April 17, 2001

Cross-polarization magic-angle spinning solid-state NMR spectroscopy has been used to investigate the dependence of ^{13}C sugar chemical shifts on specific conformational parameters of a variety of ribonucleotides and ribonucleosides. Solid-state NMR is a valuable tool for nucleoside and nucleotide structural studies since it provides the means to acquire spectra that correspond to single conformations, as opposed to ^{13}C solution NMR methods. The distinct effects of sugar puckering on the C1′, C4′, and C5′ resonances of C2′ *endo* (S type) and C3′ *endo* (N type) furanoid conformations allow us to separate them into two groups. Further analysis of each group reveals an additional dependence of the C1′ and C5′ resonances on the glycosidic and C4′–C5′ exocyclic torsion angles, respectively. However, it is found that the glycosidic conformation cannot independently be determined from sugar ^{13}C chemical shift data. The statistical methods of exploratory data analysis and discriminant analysis are used to construct two canonical coordinates—linear combinations of chemical shifts which give the statistically optimal determination of the conformation from the NMR data.

Key Words: **RNA; sugar pucker; glycosidic angle; exocyclic angle; canonical correlation.**

INTRODUCTION

^{13}C cross-polarization magic-angle spinning (CP-MAS) NMR is a useful way to study the relationship between molecular conformation and carbon chemical shifts. Its primary advantage is that it allows acquisition of spectra of molecules "frozen" in a single conformation, rather than averaged over a Boltzmann distribution of all accessible low-energy conformations, as is the case in the liquid state. In the particular case of nucleosides and nucleotides, where the major conformational flexibility is in and around the sugar ring, a variety of different conformations are available in crystalline materials; these conformations are presumably determined largely by intramolecular and crystal packing forces (*1*). Thus, the observed chemical shifts observed in the solid state correspond to specific sugar ring conformations, rather than to average values over the many conformations which exist in rapid equilibrium in solution (*2*). One approach to determining oligonucleotide and RNA structure in solution has been to measure the steady-state or transient nuclear overhauser enhancement between adjacent protons, either by one-dimensional (*3*) or two-dimensional methods (*4*). Such measurements are still the workhorses of most structural determinations. However, because of the relatively low proton density in nucleic acids, and because spin-diffusion complicates such measurements for larger oligonucleotides, other sources of structural information have been sought. Proton–proton ^{3}J coupling constants obtained via a parameterized Karplus equation (*5, 6*) can be used to determine the proportions of a small number of discrete conformations. However, that approach encounters real difficulties when used in the presence of significant cross-relaxation (*7, 8*).

We have previously shown, via CP-MAS NMR spectroscopy of crystalline deoxynucleosides and deoxynucleotides and of fibrous DNA (*9*), that the ^{13}C chemical shifts of the deoxyribose carbons are profoundly influenced by the sugar pucker, and to a lesser extent by the C4′–C5′ torsion angle (the exocyclic angle). Deoxynucleosides and deoxynucleotides with a ring pucker falling with the S- family, and B-DNA, show ^{13}C chemical shifts at the 3′ and 5′ positions between 5 and 8 ppm higher than those of N-puckered compounds, and A-DNA. The exocyclic angle had a smaller and less-convincingly demonstrated effect, with compounds *gauche–trans* about the C4′–C5′ bond having somewhat lower chemical shifts at the C5′ position than those with the more common *gauche–gauche* conformation. Varani and Tinoco (*10*) used the trends observed in the DNA work to interpret solution ^{13}C spectra of RNA oligonucleotides. The availability of ^{13}C-labeled RNAs, prepared initially to allow the use of three- and four-dimensional heteronuclear NMR techniques to circumvent the problem of spectral overlap in the proton spectrum of RNA (*11*), has made the collection of ^{13}C chemical shift data routine. Empirical interpretation of ^{13}C chemical shifts has been used by many groups (*12–16*), with considerable apparent success, despite the lack of ^{13}C benchmark information specific for RNA.

Recently, Au-Yeung and co-workers (*17, 18*) published ^{13}C NMR data for a series of crystalline oligonucleotides, interpreting them in a similar way to that used in our original paper

[1] To whom correspondence should be addressed. Fax: (402) 472 9402. E-mail: gharbison1@unl.edu.

1090-7807/01 $35.00

and showing similar trends in the chemical shifts. They have also successfully reproduced the NMR data computationally. Previous calculations of ^{13}C chemical shifts had somewhat limited success; the early results of Giessner-Prettre and Pullman (*19*), for example, predicted that the sugar pucker would have its smallest effect at the C3′ position, exactly opposite from experimental findings. Later computational results (*20*) focused on base chemical shifts. With consistent methodological improvement over the past 15 years, such computations are now a valuable adjunct to solid-state and solution NMR studies.

We have previously reported, in a preliminary communication (*21*), a novel approach to the utilization of ^{13}C chemical shift data, based on the analysis not of individual chemical shifts but of linear combinations of shifts, chosen to determine structural parameters in a statistically optimal fashion. In the present paper we present a detailed account of this technique, show how it can be used to determine sugar pucker and exocyclic angle with a high degree of certainty, and discuss the extent to which information about the glycosidic angle can be obtained from ^{13}C chemical shift data.

RESULTS

The nucleoside inosine has three well-characterized polymorphs in which the sugar conformation is distinctly different and therefore is an illustrative example of the effect of conformation on chemical shift. Upon slow evaporation from water at room temperature, inosine crystallizes as a dihydrate in the monoclinic space-group $P2_1$ (*29*), which is isomorphous with guanosine dihydrate and has two chemically distinct molecules in the unit cell. We designate this form polymorph A. Rapid evaporation at room temperature tended to give an orthorhombic polymorph (*30, 31*), also with two chemically distinct inosine molecules in the unit cell but lacking the water of crystallization. This is designated polymorph B. Finally, dehydration of commercial inosine or inosine dihydrate in an oven at 100°C for 24 h gives anhydrous monoclinic inosine (polymorph C), which has a single distinct molecule in the unit cell (*32*). The identity of all three polymorphs was confirmed by X-ray powder diffraction. Polymorph A tends to lose water on storage in the laboratory atmosphere, giving polymorph B; guanosine dihydrate has similarly been observed to dehydrate to anhydrous guanosine (*33*).

The solid-state CP-MAS NMR spectra of these three nucleosides are shown in Fig. 1. Polymorph C, with one molecule per unit cell, and thus five lines in the sugar region (50–100 ppm), has the most straightforwardly interpreted spectrum (Fig. 1a). Because of the absence of any mobile residues in the unit cell, the proton T_1 was long, and a 60-s delay was used between acquisitions, limiting the signal-to-noise that could be obtained within reasonable acquisition times. Nonetheless, the five peaks can be adequately observed, and the slightly increased breadth of the C1′ resonance is clear; this increased linewidth allows it to be assigned. C5′ was assigned by dipolar dephasing (*29*) and

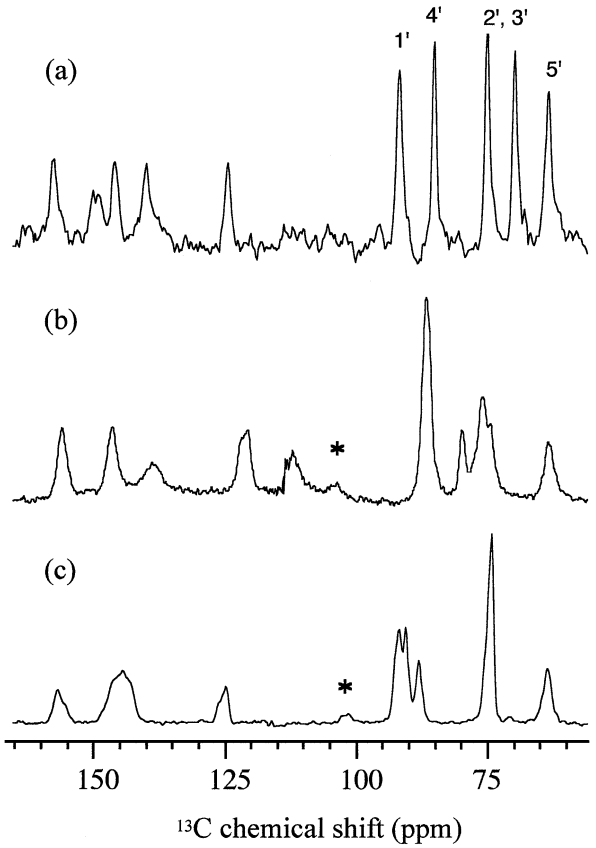

FIG. 1. ^{13}C CP-MAS spectra (a) anhydrous inosine (polymorph C), (b) monoclinic inosine dihydrate (polymorph A), and (c) orthorhombic inosine (polymorph B). MAS sidebands are denoted by asterisks.

C4′ by means of its chemical shift. As is almost always the case in these compounds, C2′ and C3′ cannot be separately assigned with certainty.

The other two forms have two molecules per unit cell, and polymorph A contains water of crystallization, whose twofold hopping motion creates a proton relaxation sink. Its spectra could therefore be acquired with an 8-s delay between acquisitions. Both of these polymorphs illustrate the major difficulties of analyzing the spectra of crystals that have more than one molecule in the unit cell. Where significant degeneracy exists, as is the case for the C5′ signal of both polymorph A and polymorph B, and *a fortiori* for C2′ and C3′ of orthorhombic inosine (polymorph B), there is obviously no difficulty in assigning the spectrum. However, most of the time, the situation is more akin to that of the C2′ and C3′ region of monoclinic inosine dihydrate (polymorph A), where there are four more or less separate resonances. In these circumstances, there is no feasible means of assigning the lines to one or the other molecule in the unit cell or even of correlating them with each other. Therefore, in our analysis we have separated the chemical shifts of species with a single molecule per unit cell, shown in Table 1, and those with more than one molecule per unit cell, shown in Table 2. The latter can

TABLE 1
^{13}C Chemical Shifts and Structural Parameters for Nucleotides and Nucleosides with a Single Conformation in the Asymmetric Unit

Compound	Chemical shifts (ppm)				Conformation			Structural reference
	C1′	C2′, C3′	C4′	C5′	Pucker	γ	χ	
6-Azauridine	92.3	71.8, 74.5	84.2	65.7	N	−ac	gt	*34*
Adenosine	92.9	71.8, 75.5	85.3	63.4	N	−ap	gt	*35*
Adenosine-3′-phosphate dihydrate	93.3	74.0, 75.7	81.4	64.2	N	−ap	gt	*1*
Inosine (polymorph C)	92.0	69.9, 75.3	85.4	63.6	N	−ap	gt	*32*
Adenosine-5′-phosphate monohydrate	89.7	69.1, 76.8	79.3	62.4	N	−ap	gg	*36*
Cytidine	92.4	66.4, 75.5	84.0	60.0	N	−ap	gg	*37*
5-Methyluridine hemihydrate	90.5	67.4, 77.5	84.2	59.2	N	−ap	gg	*38*
2-Thiocytidine dihydrate	95.2	68.0, 77.0	82.9	58.7	N	−ap	gg	*39*
5-Bromouridine	88.4	74.6, 75.3	88.3	63.1	S	−ac	gg	*40*
3-Deazauridine	85.3	75.9, 75.9	85.3	64.4	S	−ac	gg	*41*
Inosine-5′-phosphate monosodium salt octahydrate	88.1	73.6, 78.3	88.1	65.6	S	−ac	gg	*42*
Cytidine-3′-phosphate	86.9	75.4, 78.1	86.9	63.8	S	−ac	gg	*43*
Inosine-5′-phosphate disodium salt heptahydrate	87.4	73.1, 77.5	87.4	64.8	S	−ac	gg	*44*
Adenosine hydrochloride	88.1	73.6, 79.9	88.1	63.8	S	−ac	gg	*45*
Uridine-5′-phosphate disodium salt heptahydrate	86.2	71.0, 76.1	88.3	64.2	S	−ac	gg	*46*
Uridine-3′-phosphate disodium salt tetrahydrate	85.8	73.7, 76.2	88.0	62.6	S	−ac	gg	*47*
5-Hydroxyuridine	87.4	71.1, 76.1	85.0	63.2	S	−ac	gg	*48*
8-Bromoadenosine	92.6	72.1, 74.7	90.4	66.1	S	+sc	gg	*49*
8-Bromoguanosine dihydrate	90.8	72.2, 76.3	88.0	63.6	S	+sc	gg	*50*
Xanthosine dihydrate	88.5	71.9, 73.4	85.3	63.4	S	+sc	gg	*51*

be used to check the consistency of our predictions, but were not used in the primary data analysis.

Spectra of this sort were obtained for the 20 compounds shown in Table 1 and the 5 compounds in Table 2, all of which have two molecules per unit cell. The multiplicity of the overlapping peaks is denoted by a number in parentheses. Along with the chemical shifts, we give the important conformational parameters derived from the relevant crystal structure. Sugar puckers are simply classified as N or S and exocyclic angles as *gg* or *gt*, while the glycosidic angle χ is specified by the conformational region it lies in: ±sp, ±sc, ±ac, and ±ap (IUPAC-IUB, 1983). The χ values for *syn* and *anti* conformations are $0 \pm 90^\circ$ and $180 \pm 90^\circ$, respectively.

DISCUSSION

Effects of Phosphorylation

In order to compare nucleoside and nucleotide data in a single set, we must first be able to account for the effects of phosphorylation on the ^{13}C chemical shift. Typically, these effects

TABLE 2
^{13}C Chemical Shifts and Structural Parameters for Nucleotides and Nucleosides with More Than One Molecule in the Asymmetric Unit

Compound	Chemical shifts (ppm)				Conformation			Structural reference
	C1′	C2′, C3′	C4′	C5′	Pucker	γ	χ	
8-Bromoinosine	93.1, 94.9	69.6 (2), 73.8 (2)	82.0, 84.0	57.5, 60.0	N (2)	+ac (2)	gg (2)	*52*
Uridine	91.4 (2)	66.5, 68.9, 75.1 (2)	82.3, (2)	61.8, 63.4	N (2)	+ap (2)	gg (2)	*53*
Inosine (orthorhombic; polymorph B)	88.3, 92.2	74.4 (4)	92.4, 90.8	63.6 (2)	S (2)	+sc (2)	gg (2)	*31*
Guanosine dihydrate (monoclinic)	85.7 (2)	72.7, 76.3 (2) 79.2	85.7, 86.9	63.1	S (2)	−ap, −sc	gg (2)	*29*
Inosine dihydrate (monoclinic; polymorph A)	86.8 (2)	74.0, 76.0 (2), 79.9	86.8 (2)	63.6 (2)	S (2)	−ap, −sc	gg (2)	*29*

are smaller than those either of pucker inversion or of rotation about the exocyclic angle. In our previous work (*9*) we estimated the effect of phosphorylation by comparing nucleoside and nucleotide shifts in solution. This, however, is not a completely satisfactory method. Liquid-state chemical shifts of small molecules are averages over a series of conformers, and replacement of a hydroxyl with a bulky phosphate group will not only directly affect the ^{13}C shifts of particular conformers, but also may change the distribution of conformers. However, within the current data set we have sufficient members of the S/*gg*/-ac group with phosphates on the 5′ position to directly estimate the effect of phosphorylation on chemical shifts. The only chemical shift for which we can detect a significant perturbation is that of the directly bonded carbon, C5′, which shows an average 1.4-ppm downfield shift. This, perhaps coincidentally, is the average shift difference at that position in solution between nucleosides and nucleotides in the Sadtler catalog (*54*, *55*). Since we cannot routinely assign the 3′ carbon signal, we cannot directly determine the effect of phosphorylation at that position on the shift of the directly bonded carbon. We therefore assume the effect to be identical to that on the C5′ position. No other clear tendencies are seen in the spectra and, in particular, no influence on more distant carbons can be reliably discerned.

Data Analysis

The raw data examined in this study consisted of the ^{13}C chemical shifts at the C1′, C4′, and C5′ positions and (because the two resonances cannot be unambiguously assigned) the average of the C2′ and C3′ chemical shifts. In the initial analysis, only those compounds which have either one molecule in the unit cell or two molecules with essentially identical conformations were used. Data analysis proceeded in two phases. First, exploratory data analysis, or EDA, was used to ascertain what conformational information resides in the chemical shifts. After this was complete, discriminant analysis was used to determine the linear combination of chemical shifts which gives the best statistical discrimination between different conformations.

The chemical shifts measured for all 20 members of the preliminary group of compounds in which there is only one molecule in the unit cell are given in Table 1, along with their major conformational features and the reference to the relevant crystallographic structure. The EDA phase consists of examining pair plots created by the statistical language S for all pairwise combinations of the assignable chemical shifts. Three- and four-dimensional plots were also examined. Figure 2 shows such "pair plots." Even without identifying individual data points by conformation, the data clearly fall into two clusters, of size 12 and 8, with 3 points lying somewhat further away from the center of the larger group. The C1′ versus C4′ pair plot, in particular, shows clear separation between these two groups. If we now label the data points by conformation, it becomes clear that the smaller group of 8 corresponds to the N-puckered nucleotides, while the smaller group of 12 is the S-puckered group.

Discriminant Coordinates

We chose to adopt the Fisher discriminant function to obtain a statistically optimal separation between the two groups (*56*). The Fisher discriminant function Z, also called the canonical discriminant function, is the linear combination of dependent variables that maximizes the statistical discrimination between two or more populations. Put in other terms, discriminant analysis finds the vector along which the difference in means between the two populations is largest compared with the standard error of the means. In matrix formalism the Fisher discriminant function is expressed as

$$Z = \hat{a}x, \qquad [1]$$

where

$$\hat{a} = [\mathbf{x}_i - \mathbf{x}_j]' S^{-1}_{\text{pooled}}. \qquad [2]$$

The $\mathbf{x}_i$ and $\mathbf{x}_j$ terms are the vectors of the i and j sample distributions. These vectors have as elements the average of the sample variables. For example, in our case $\mathbf{x}_N = \{\delta\mathbf{C}_{1,N}, \delta\mathbf{C}_{4,N}, \delta\mathbf{C}_{5,N}\}$ for the N sample distribution, while S_{pooled} is the weighed average of the covariance matrices of the two groups.

Using variables C1′, C4′, and C5′ we determined a first canonical coordinate:

$$\text{can1} = 0.179\delta_{\text{C1}'} - 0.225\delta_{\text{C4}'} - 0.0585\delta_{\text{C5}'}. \qquad [3]$$

Note that 1.4 ppm must been added to the $\delta_{\text{C5}'}$ chemical shift of the nucleosides to account for the effect of phosphorylation on the directly bonded carbon.

We can obtain can1 values for each of the 20 chemical shifts in our data set; along this coordinate, the two clusters are cleanly separated by a large gap. For the N-puckered group, can1 $= -5.94 \pm 0.53$ (mean ± 1 standard deviation), while for the S pucker, can1 $= -7.73 \pm 0.33$. The difference between these two means is clearly more than twice the sum of the standard deviations.

Another way to test the statistical significance of this separation is to order the points by can1 value and then perform a one-sided Wilcoxon rank-sum test. Letting $N = \text{C}(20, 12) = 125, 970$, the probability that the data would be ranked perfectly along can1 by pure chance is $1/N$. This coordinate therefore can be used to determine the ring pucker with a high degree of certainty.

An alternative method for analyzing the data might use the pseudorotation angle P in a linear or nonlinear regression model to predict the group. In fact, a model predicting P from can1 fits as well as any and is parsimonious. However, there are two disadvantages of this model. First, the data don't really look linear. There seem to be two widely separated clusters. There are no intermediate points between clusters along the can1

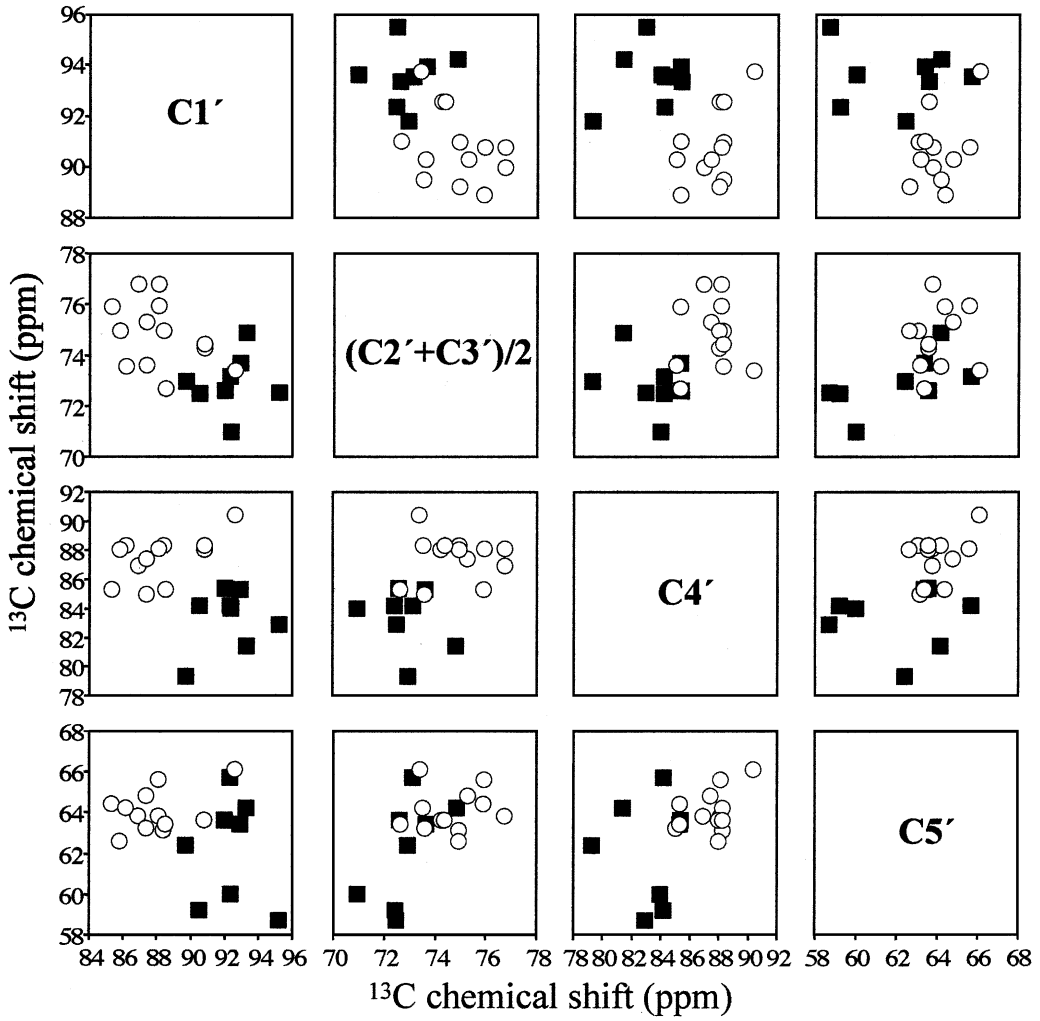

FIG. 2. ^{13}C chemical shift pair plots for the (○) S-puckered (■) N-puckered nucleosides and nucleotides used in this study.

direction, so any model will be forced to interpolate into this region. This last consideration also makes it difficult to compare nonlinear models. Since the groups are separated, there is no way to accurately model the boundary region. It therefore seems wise at this time to treat the sugar pucker as a two-state problem.

Exocyclic Torsion Angle

The exocyclic torsion angle γ appears to have quite distinct and characteristic effects on the chemical shift. Specifically, a plot of C1′ versus C5′ chemical shift separates the N group into two clusters, each of size 4. These clusters are identified with a difference in the torsion angle γ, corresponding to the *gg* and *gt* configurations. In Fig. 3, we have denoted *gt* configuration with filled squares, while the open circles indicate a *gg* configuration.

While the comparison of *gg* with *gt* conformers is based on fewer data points than the comparison of ring puckers, it agrees with previous observations of the effect of the exocyclic angle on DNA chemical shifts (*9*). We have constructed a second canonical coordinate, can2, which maximally separates the two clusters.

$$\text{can2} = -0.0605(\delta_{C2'} + \delta_{C3'}) - 0.0556\delta_{C4'} - 0.0524\delta_{C5} \quad [4]$$

This coordinate, it should be noted, is considerably different from that which we gave in our preliminary report (*21*). While the largest average difference in chemical shifts is at the C5′ position, the C5′ chemical shift also has considerable scatter within each group, and the average (C2′ + C3′) shift is actually a better discriminant than that of C5′. This is reflected in the larger weighting of this shift in the canonical coordinate.

Because of the smaller number of data points, it is difficult to make a useful estimate of the significance level of the difference in can2 between *gg* and *gt* conformers. The standard deviation of the overall data set, with respect to the respective mean can2 values of 16.62 (*gg*) and 17.08 (*gt*), is 0.16. This takes into

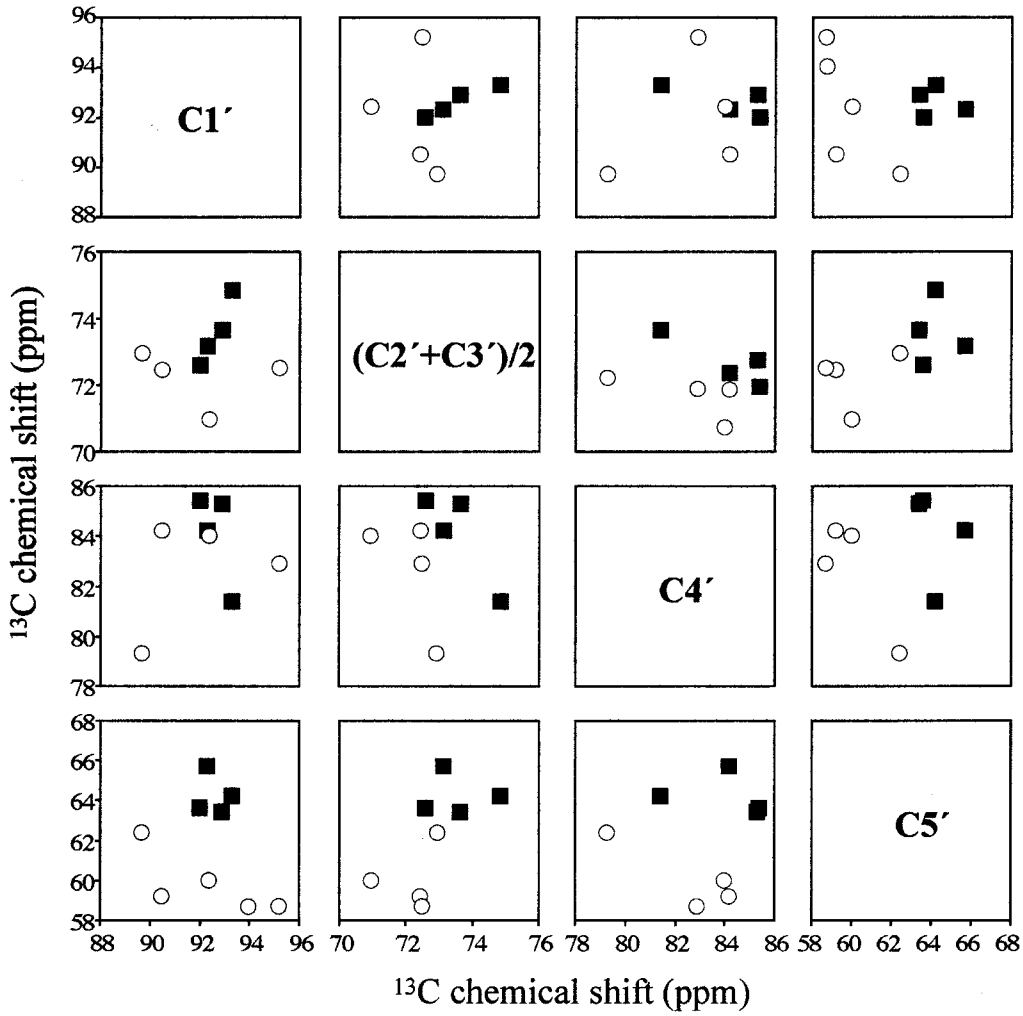

FIG. 3. ^{13}C chemical shift pair plots for the (○) N/*gg*- and (■) N/*gt* nucleosides and nucleotides used in this study.

account that there were five parameters extracted from the data points (the three coefficients and the two means). The standard deviation is, as before, considerably smaller than the difference between the means and suggests that can2 is a statistically useful measure of the exocyclic conformation. Clearly, though, the analysis would benefit from a larger set of experimental data.

Glycosidic Torsion Angle

Figure 4 shows the can1 and can2 values for the 20 nucleotides. It is clear that N are well separated from S along can1, and *gt* from *gg* along can2, and that the two coordinates are linearly independent of each other. It is also clear from this plot and from EDA that there is no independent coordinate similar to can1 and can 2 separating species with *syn* glycosidic angles from those with an *anti* configuration about the glycosidic bond. In fact, the five *syn* compounds are those which form the "handle" protruding outward from the S cluster in the direction of the N cluster in the EDA analysis. At best, the *syn* glycosidic linkages attached to S-puckered sugar rings can be distinguished from *anti* linkages by their relatively high values of can1, but in the presence of significant dynamics it is unlikely that this is a reliable measure. We are forced to conclude that this aspect of the configuration cannot be determined from the carbon chemical shift.

Limitations of the Present Work

One problem with the present data analysis arises from the difficult of definitively assigning C2′ and C3′, necessitating the use of the sum of these two shifts in the analysis. It is possible that a further improvement in the discrimination coordinates could be obtained if such assignment were possible. We can investigate this possibility by examining the relationship between the magnitude of the difference between C2′ and C3′ chemical shifts and conformation. This is done in Fig. 5, where we plot the absolute value of the difference in these two shifts against their sum. In this representation, the two individual chemical shifts fall along the two diagonals; which diagonal is which

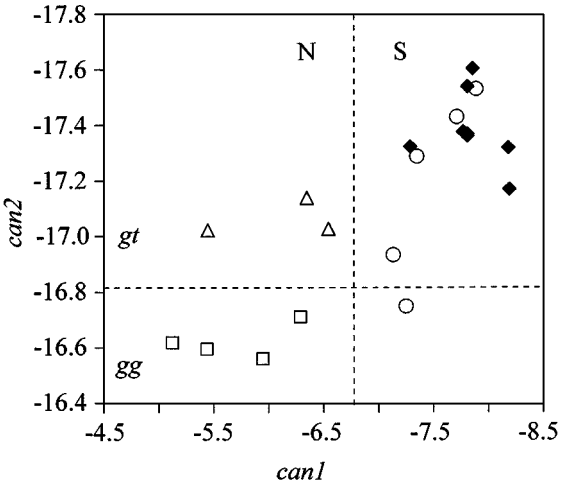

FIG. 4. Canonical coordinates can1 and can2, calculated as described in the text, for the 20 nucleosides and nucleotides used in this study, identified by their ring pucker, exocyclic angle, and glycosidic angle, as follows: (□): N/*gg*/*anti*, (△): N/*gt*/*anti*, (○): S/*gg*/*syn*, (◆): S/*gg*/*anti*.

depends on the sign of the difference. It is clear that the difference between C2′ and C3′ discriminates primarily between *gg* and *gt* conformers. In fact, the line separating the *gg* and *gt* clusters is approximately parallel to one of the diagonals, implying that the dependence of can2 on the average C2′ and C3′ chemical shift is actually concentrated in one of the two components. This component is likely C3′, if we make the assumption that C2′ is generally of higher frequency than C3′. On the other hand, the difference between these two chemical shifts is largely independent of the sugar pucker and the glycosidic angle, indicating that an ability to separately assign C2′ and C3′ will not improve the determination of these conformational parameters.

This data set also is relatively poor in nucleosides and nucleotides in rarer sugar conformations. While there are a few other crystalline nucleotides and nucleosides that do not fall into the four classes discussed in this work, their rarity probably

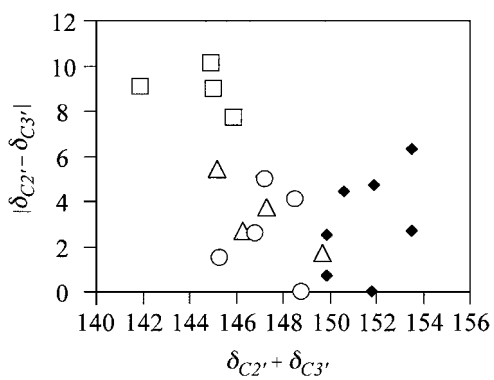

FIG. 5. Sums and absolute differences of the chemical shifts of the unassignable sugar resonances C2′ and C3′ for the 20 nucleosides and nucleotides used in this study, identified by the same notation used in the previous figure.

means that the analysis of their chemical shifts will primarily depend on computational methods (*57*).

Comparison with Published Solution NMR Structures

Calculation of can1 and can2 values for liquid-state structures (*12–16*) shows broad agreement with the conclusions of this solid-state NMR study, with a couple of significant exceptions. The can1 values for S- and *syn* puckered bases in solution seem to be uniformly less than those observed in the solid state, so that range of can1 values in solution appears to be somewhat lower. It seems reasonable to attribute this discrepancy to dynamics; *syn* nucleosides almost invariably are found in RNA loop regions, and these regions are likely highly mobile. If this conjecture is indeed true, it implies that most loop regions in RNA either undergo significant rotation about the glycosidic bond or alternatively significant pucker interconversion. In either case such dynamics would average the chemical shifts in solution and lead to lower can1 values. With this proviso, and with one exception (*10*) that has since been corrected in the literature, deductions from can1 values are in complete agreement with published solution NMR conformations. In contrast, there are several cases where can2 values from published ^{13}C chemical shifts are markedly discrepant with the published conformations. These discrepancies lead to exocyclic torsion angles being determined as *gt* where chemical shift data would suggest *gg*, or vice versa. Such discrepancies seem to be concentrated in regions near the end of a helix, where some "fraying" of the nucleotide may be evident and possibly where the solution structure determination may have been complicated by internal dynamics.

CONCLUSIONS

The canonical coordinates derived in this work appear to be the optimal way of extracting the sugar pucker and exocyclic angle from RNA chemical shift data. They are mathematically straightforward and can be used to determine conformations with a meaningful statistical likelihood. This should be readily compatible with existing methods of structural determination.

MATERIALS AND METHODS

Sample Preparation

Nucleosides and nucleotides used were purchased from Sigma Chemical Co. (St. Louis, MO) and ICN Pharmaceuticals, Inc. (Cleveland, OH). Crystals of adenosine-3′-phosphate dihydrate, adenosine-5′-phosphate monohydrate, monoclinic inosine dihydrate, monosodium inosine-5′-phosphate octahydrate, and 2-thiocytidine dihydrate were obtained by slow evaporation of aqueous solutions. Guanosine dihydrate and 8-bromoguanosine dihydrate were crystallized by slowly cooling hot saturated aqueous solutions of these substances. The crystals of cytidine and 5-methyluridine hemihydrate were grown from solutions

of 30% aqueous ethanol. Disodium uridine-5′-phosphate heptahydrate, disodium uridine-3′-phosphate tetrahydrate, and disodium inosine-5′-phosphate heptahydrate were crystallized from a 25% acetone/water mixture by the diffusion technique. A solution of adenosine was titrated with concentrated HCl to pH 2.41. The protonated crystals then were obtained from aqueous ethanol solution by evaporation at 4°C. Anhydrous inosine was prepared by oven-heating at 100°C for 24 h. Crystals of adenosine, uridine, deazauridine, 8-bromoadenosine, 8-bromoinosine, 5-hydroxyuridine, 5-bromouridine, xanthosine dihydrate, cytidine-3′-monophosphate, 6-azauridine, 6-thiopurine riboside, and 5-iodouridine were already in the desired forms as purchased. Crystals of orthorhombic inosine were obtained by fast evaporation at 20°C.

^{13}C NMR Data Acquisition

All NMR measurements were performed using a homebuilt solid-state spectrometer operating at an external field strength of 7.1 T. Sample quantities ranged from 25 to 100 mg; spectra were run in cylindrical sapphire rotors in a homebuilt probe incorporating Doty Scientific stator design. All nucleotide and nucleoside spectra were obtained using CP-MAS, with matched rotating frame B_1 field strengths equivalent to 62.5 kHz. Spinning rates between 2.8 and 3.7 kHz were used. Spectra were obtained using a proton 90° pulse of 4.0 μs and spectral width of $\pm$10 kHz, using high-power proton decoupling during acquisition. The contact time was 1.00 ms for all experiments. Transients were averaged with a recycle delay ranging from 4 to 16 s for hydrated and 30 to 60 s for anhydrous crystalline samples. Typically, a total of 218 to 1024 free induction decay signals were collected for each sample and processed by zero-filling before Fourier transformation to make the estimation of peak positions possible. Spinning sidebands in some of the spectra were eliminated by employing the SELTICS pulse sequence (*22*). All shifts are referenced to external TMS; however, the downfield resonance of adamantane, at 38.56 ppm from TMS, was used as a primary reference. No corrections are made for susceptibility effects, which we expect to be well under 1 ppm.

^{13}C NMR Spectral Interpretation

Assignments of the carbon resonances were made initially by comparison with spectral data of analogous nucleosides and nucleotides in solution (*23, 24*). The resonance observed at the highest field was attributed to C5′. This was easily confirmed by noting its faster dipole evolution in the delayed decoupling experiment (*25*). Assignment of the C1′ resonance was readily established by observation of its greater linewidth and occasionally resolved doublet splitting lineshape due to ^{14}N quadrupolar effects (*26–28*). The C4′ signal is in general at considerably higher chemical shift than the 2′ and 3′ signals and can be assigned on that basis; however, without isotopic labeling, the latter two resonances cannot be separately assigned with any reliability using current methods.

ACKNOWLEDGMENT

This work was supported by NSF under Grant MCB-9604521.

REFERENCES

1. M. Sundaralingam, Stereochemistry of nucleic acids constituents. III. Crystal and molecular structure of adenosine 3′-phosphate dihydrate (adenylic acid *b*), *Acta Crystallogr.* **21,** 495–506 (1966).
2. D. B. Davis, Conformations of nucleosides and nucleotides, *Prog. NMR Spectrosc.* **12,** 135–225 (1978).
3. J. Tropp and A. G. Redfield, Environment of ribothymidine in transfer ribonucleic acid studied by means of nuclear Overhauser effect, *Biochemistry* **20,** 2133–2040 (1981).
4. D. J. Patel, L. Shapiro, and D. Hare, DNA and RNA: NMR studies of conformations and dynamics in solution, *Q. Rev. Biophys*. **20,** 35–112 (1987).
5. F. A. A. M. DeLeeuw and C. Altona, Conformational analysis of β-D-ribo-, β-D-deoxyribo-, β-D-arabino-, β-D-xylo-, β-D-lyxonucleosides from proton–proton coupling constants, *J. Chem. Soc. Perkin Trans.* **2,** 375–379 (1982).
6. A. Bax and L. Lerner, Measurement of ^1H–^1H coupling constants in DNA fragments by 2D NMR, *J. Magn. Reson*. **79,** 429–438 (1988).
7. G. S. Harbison, Interference between J-couplings and cross-relaxation in solution NMR spectroscopy: Consequences for macromolecular structure determination, *J. Am. Chem. Soc.* **115,** 6023–6024 (1993).
8. L. Zhu, B. R. Reid, M. Kennedy, and G. P. Drobny, Modulation of *J* couplings by cross relaxation in DNA sugars, *J. Magn. Reson*. **111,** 195–202 (1994).
9. R. Santos, P. Tang, and G. S. Harbison, Determination of the DNA sugar pucker using ^{13}C NMR spectroscopy, *Biochemistry* **28,** 9372–9378 (1989).
10. G. Varani and I. Tinoco Jr., Carbon assignments and heteronuclear coupling constants for an RNA oligonucleotide from natural abundance ^{13}C–^1H correlated experiments, *J. Am. Chem. Soc*. **113,** 9349–9354 (1991).
11. E. P. Nikonowicz and A. Pardi, Application of four-dimensional heteronuclear NMR to the structure determination of a uniformly ^{13}C labeled RNA, *J. Am. Chem. Soc*. **114,** 1082–1083 (1992).
12. F. M. Jucker and A. Pardi, Solution structure of the CUUG hairpin loop: A novel RNA tetraloop motif, *Biochemistry* **34,** 14416–14427 (1995).
13. N. L. Greenbaum, I. Radhakrishnan, D. Hirsh, and D. J. Patel, Determination of the folding topology of the SL1 RNA from *Caenorhabditis elegans* by multidimensional heteronuclear NMR, *J. Mol. Biol*. **252,** 314–327 (1995).
14. G. Varani and I. Tinoco Jr., RNA structure and NMR spectroscopy, *Q. Rev. Biophys*. **24,** 479–532 (1991).
15. E. P. Nikonowicz and A. Pardi, An efficient procedure for assignment of the proton, carbon and nitrogen resonances in ^{13}C/^{15}N labeled nucleic acids, *J. Mol. Biol*. **232,** 1141–1156 (1993).
16. S. E. Butcher, T. Dieckmann, and J. Feigon, Solution structure of the conserved 16 S-like ribosomal RNA UGAA tetraloop, *J. Mol. Biol*. **268,** 348–358 (1997).
17. 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).
18. X.-P. Xu, Ph.D. thesis, Chinese University of Hong Kong, Hong Kong, China, 1999.
19. C. Giessner-Prettre and B. Pullman, Quantum mechanical calculations of NMR chemical shifts in nucleic acids, *Q. Rev. Biophys*. **20,** 113–172 (1987).
20. 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).

21. 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.

22. J. Hong and G. S. Harbison, Magic-angle spinning sideband elimination by temporary interruption of the chemical shift, *J. Magn. Reson*. **105,** 128–136 (1993).

23. A. J. Jones, D. M. Grant, M. W. Winkely, and R. K. Robins, Carbon-13 magnetic resonance. XVII. Pyrimidine and purine nucleosides, *J. Am. Chem. Soc*. **92,** 4079–4087 (1970).

24. M. C. Thorpe, W. Coburn, and J. A. Montgomery, Carbon-13 nuclear magnetic resonance spectra of some 2-, 6-, and 2,6-substituted purines, *J. Magn. Reson*. **15,** 98–112 (1974).

25. S. J. Opella and M. H. Frey, Selection of non protonated carbon resonances in solid-state nuclear magnetic resonance, *J. Am. Chem. Soc*. **101,** 5854–5856 (1979).

26. A. Naito, S. Ganapathy, and C. A. McDowell, High resolution solid state ^{13}C NMR spectra of carbons bonded to nitrogen in a sample spinning at the magic angle, *J. Chem. Phys*. **74,** 5393–5397 (1981).

27. N. Zumbulyadis, P. M. Henrichs, and R. H. Young, Quadrupole effects in the magic-angle-spinning spectra of spin-1/2 nuclei, *J. Chem. Phys*. **75,** 1603–1611 (1981).

28. A. C. Olivieri, L. Frydman, M. Grasselli, and L. E. Diaz, Analysis of ^{13}C, ^{14}N residual dipolar coupling in the ^{13}C CP/MAS NMR spectra of ribonucleosides, *Magn. Reson. Chem*. **26,** 281–286 (1988).

29. U. Thewalt, C. E. Bugg, and R. E. Marsh, Crystal structure of guanosine dihydrate and inosine dihydrate, *Acta Crystallogr*. **B26,** 1089–1101 (1970).

30. E. Subramanian and J. J. Madden, Syn conformation for inosine, the wobble nucleoside in some tRNA, *Biochem. Biophys. Res. Commun*. **50,** 691–696 (1973).

31. E. Subramanian, Inosine (hypoxanthine riboside), $C_{10}H_{12}N_4O_5$, *Cryst. Struct. Commun*. **8,** 777–785 (1979).

32. A. R. I. Munns and P. Tollin, Crystal and molecular structure of inosine, *Acta Crystalloger* **B26,** 1101–1113 (1970).

33. Y. Sugawara, Y. Iimura, H. Iwasaki, H. Urabe, and H. Saito, Reversible crystal transition of guanosine between the dihydrate and anhydrous states coupled with adsorption–desorption process, *J. Biomol. Struct. Dynam*. **11,** 721–729 (1994).

34. C. H. Schwalbe and W. Saenger, 6-Azauridine, a nucleoside with unusual ribose conformation, *J. Mol. Biol*. **75,** 129–143 (1973).

35. T. F. Lai and R. E. Marsh, Crystal structure of adenosine, *Acta Crystallogr.* **B28,** 1982–1989 (1972).

36. J. Kraut and L. H. Jensen, Refinement of the crystal structure of adenosine 5′-phosphate, *Acta Crystallogr.* **16,** 79–88 (1963).

37. S. Furberg, S. C. Petersen, and C. Roemming, Refinement of the crystal structure of cytidine, *Acta Crystallogr.* **18,** 313–320 (1965).

38. D. J. Hunt and E. Subramanian, Crystal structure of 5-methyluridine, *Proc. R. Soc. A* **295,** 320–333 (1969).

39. G. H. Y. Lin, M. Sundaralingam, and S. K. Arora, Stereochemistry of nucleic acids and their constituents. XV. Crystal and molecular structure of 2-thiocytidine dihydrate, a minor constituent of transfer ribonucleic acid, *J. Am. Chem. Soc*. **93,** 1235–1241 (1971).

40. J. Iball, C. H. Morgan, and H. R. Wilson, Fibers of guanine nucleosides and nucleosides, *Proc. R. Soc. A* **295,** 320–333 (1966).

41. C. H. Schwalbe and W. Saenger, 3-Deazauridine: Crystal structure, conformational parameters, and molecular orbital calculations, *Acta Crystallogr.* **B29,** 61–69 (1973).

42. S. T. Rao and M. Sundaralingam, Stereochemistry of nucleic acids and their constituents. V. The crystal and molecular structure of a hydrated monosodium inosine 5′-phosphate. A commonly occurring unusual nucleotide in the anticodons of tRNA, *J. Am. Chem. Soc.* **91,** 1210–1217 (1969).

43. C. E. Bugg and R. E. Marsh, Structure of the monoclinic form of cytidylic acid b, *J. Mol. Biol.* **25,** 67–82 (1967).

44. N. Nagashima and Y. Iitake, The structure of barium inosine-5′-phosphate and disodium inosine-5′-phosphate, *Acta Crystallogr.* **B24,** 1136–1138 (1968).

45. K. Shikata, T. Ueki, and T. Mitsui, The crystal and molecular structure of adenosine hydrochloride, *Acta Crystallogr.* **B29,** 31–38 (1973).

46. T. P. Seshadri, M. A. Viswamitra, and G. Kartha, Structure of disodium uridine 5′-phosphate heptahydrate, *Acta Crystallogr.* **B36,** 925–927 (1980).

47. M. A. Viswamitra, B. S. Reddy, M. N. G. James, and G. J. B. Williams, The molecular and crystal structure of disodium uridine-3′-phosphate tetrahydrate, *Acta Crystallogr.* **B28,** 1108–1116 (1972).

48. U. Thewalt and C. E. Bugg, The crystal structure of 5-hydroxyuridine, *Acta Crystallogr.* **B29,** 1393–1398 (1973).

49. S. S. Tavale and H. M. Sobell, Crystal and molecular structure of 8-bromoguanosine and 8-bromoadenosine, two purine nucleosides in the syn conformation, *J. Mol. Biol*. **48,** 109–123 (1970).

50. C. E. Bugg and U. Thewalt, Effects of halogen substituents on base stacking in nucleic acid components: The crystal structure of 8-bromoguanosine, *Biochem. Biophys. Res. Commun.* **37,** 623–629 (1969).

51. G. Koyama, H. Nakamura, H. Umezawa, and Y. Iitaka, Xanthosine dihydrate, *Acta Crystallogr.* **B32,** 969–972 (1976).

52. H. Sternglanz, J. M. Thomas, and C. E. Bugg, Stacking patterns of halogenated purines: Crystal structure of 8-bromoinosine, *Acta Crystallogr.* **B33,** 2097–2102 (1977).

53. E. A. Green, R. D. Rosenstein, R. Shiono, D. J. Abraham, B. L. Trus, and R. E. Marsh, The crystal structure of uridine, *Acta Crystallogr.* **B31,** 102–107 (1975).

54. Sadtler Research Laboratories, Inc., "The Sadtler Collection of ^{13}C NMR Spectra," Sadtler Research, Philadelphia, PA, 1989.

55. Sadtler Research Laboratories, Inc., "The Sadtler Collection of ^{13}C NMR Spectra," Sadtler Research , Philadelphia, PA, 1993.

56. R. A. Johnson and D. A. Wichern, "Applied Multivariate Statistical Analysis," 3rd ed., Prentice Hall, Englewood Cliffs, NJ, 1992.

57. Paolo Rossi and Gerard S. Harbison, Calculation of ^{13}C Chemical Shifts in RNA Nucleosides: Structure- ^{13}C Chemical Shift Relationships, submitted for publication.