Journal of Biomolecular NMR, **21:** 11–29, 2001.
KLUWER/ESCOM

Prediction of proton chemical shifts in RNA

Their use in structure refinement and validation

Jenny A.M.T.C. Cromsigt[a], Cees W. Hilbers[b] & Sybren S. Wijmenga[a,*]
[a]*Department of Medical Biosciences, Medical Biophysics, Umeå University, S90187 Umeå, Sweden;* [b]*NSR Centre for Molecular Structure Design, and Synthesis, Laboratory of Biophysical Chemistry, University of Nijmegen, The Netherlands*

Received 22 February 2001; Accepted 21 May 2001

Key words: chemical shift, RNA, resonance-assignment, structure-refinement, structure-validation

Abstract

An analysis is presented of experimental versus calculated chemical shifts of the non-exchangeable protons for 28 RNA structures deposited in the Protein Data Bank, covering a wide range of structural building blocks. We have used existing models for ring-current and magnetic-anisotropy contributions to calculate the proton chemical shifts from the structures. Two different parameter sets were tried: (i) parameters derived by Ribas-Prado and Giessner-Prettre (GP set) [(1981) *J. Mol. Struct.,* **76**, 81–92.]; (ii) parameters derived by Case [(1995) *J. Biomol. NMR*, **6**, 341–346]. Both sets lead to similar results. The detailed analysis was carried using the GP set. The root-mean-square-deviation between the predicted and observed chemical shifts of the complete database is 0.16 ppm with a Pearson correlation coefficient of 0.79. For protons in the usually well-defined A-helix environment these numbers are, 0.08 ppm and 0.96, respectively. As a result of this good correspondence, a reliable analysis could be made of the structural dependencies of the ^{1}H chemical shifts revealing their physical origin. For example, a down-field shift of either H2′ or H3′ or both indicates a *high-syn/syn* χ-angle. In an A-helix it is essentially the 5′-neighbor that affects the chemical shifts of H5, H6 and H8 protons. The H5, H6 and H8 resonances can therefore be assigned in an A-helix on the basis of their observed chemical shifts. In general, the chemical shifts were found to be quite sensitive to structural changes. We therefore propose that a comparison between calculated and observed ^{1}H chemical shifts is a good tool for validation and refinement of structures derived from NOEs and *J*-couplings.

Abbreviations: cs, cross strand; NMR, Nuclear Magnetic Resonance; NOE, Nuclear Overhauser Effect; pdb, protein data bank; ppm, parts per million; RMDS, root-mean-square-deviation.

Introduction

Chemical shift values carry important structural information, although their application as a tool to derive three-dimensional characteristics of biomolecules has been overtaken by the use of *J*-couplings and NOE effects. However, in the field of protein NMR, chemical shifts have experienced a renewed interest in the past ten years, because of the availability of a large number of detailed 3D structures. This allowed the reliability of chemical shift calculations to be tested and their conformation dependence to be determined (Ösapay and Case, 1991, 1994; Ösapay et al., 1994; Asakura et al., 1992, 1995; Williamson and Asakura, 1993; Case et al., 1994; Williamson et al., 1995). As a result, secondary structure information can now reliably be derived from ^{13}C and ^{1}H shifts (Wishart and Sykes, 1994). It has also been shown that chemical shift derived constraints, used as an extra restraining force in molecular dynamics, improve the accuracy of the derived structure (Ösapay et al., 1994; Celda et al., 1995; Kuszewski et al., 1995a, b)

Earlier, we have analyzed the conformational dependence of ^{1}H chemical shifts of DNA molecules. We demonstrated that these shifts can be predicted with good accuracy (RMSD 0.17 ppm) based on an

*To whom correspondence should be addressed. E-mail: sybren.wijmenga@chem.umu.se

analysis of roughly 20 well-determined solution structures (Wijmenga et al., 1997; Wijmenga and van Buuren, 1998). In a later study, distributions of nucleic acids proton shifts were also analyzed by Dejaegere et al. (1999) using electrostatic and ring-current models (RMSD 0.29 ppm). Another analysis of DNA ^{1}H chemical shifts, which was a strictly empirical determination of intrinsic contributions and additional neighbor contributions to chemical shifts, showed that they can be correlated to a specific triplet sequence in a double helix with high precision (RMSD 0.04 ppm) (Altona et al., 2000).

In addition, we have demonstrated that ^{1}H chemical shifts, when properly implemented, can serve as valuable constraints in structure calculations of DNA and can even partly replace structural restraints derived from NOEs and *J*-couplings (Cromsigt et al., 1998). Furthermore, chemical shifts can be useful as a tool in the validation of NOE and *J*-coupling based structures, as is shown in the study of a DNA three-way junction (van Buuren et al., 2000).

Here, we extend the analysis to the chemical shifts of the non-exchangeable protons in RNAs, using a database of 28 pdb-deposited RNA structures. We address the following four issues:

- The reliability of ^{1}H chemical shift calculations is tested by comparing the calculated and observed ^{1}H chemical shifts of the 28 pdb-deposited RNA structures. For the complete data base a root-mean-square-standard-deviation between calculated and observed chemical shifts is found of 0.16 ppm with a Pearson linear correlation coefficient of 0.79, while for the usually well-defined A-helix environment these numbers are 0.08 ppm and 0.96, respectively.
- As a result of this good correspondence, a reliable analysis could be made of the structural dependencies of the ^{1}H chemical shifts in both a helical and a non-helical environment revealing their physical origin.
- The analysis of the shifts of the base protons in an A-helix environment led to the formulation of a number of rules that describe their main dependence on the nature of the neighboring nucleotide. These rules are of value for their assignment and we discuss possible applications.
- The chemical shifts were generally found to be sensitive to small structural differences but can reliably be calculated. We therefore propose that a comparison between calculated and observed ^{1}H chemical shifts is a good tool for validation and refinement of structures derived from NOEs and *J*-couplings.

Methods

The analysis of the RNA ^{1}H chemical shifts was carried out by comparing calculated to experimental values. They were derived from a database of 28 pdb-deposited RNA structures, containing a variety of different structural building blocks. The chemical shifts were calculated with the program NUCHEMICS (Wijmenga et al., 1997) and subsequently analyzed in EXCEL. The calculation of the ^{1}H chemical shifts is outlined briefly below. For a more detailed description we refer to Wijmenga et al. (1997).

The total calculated chemical shift of ^{1}H nucleus p, δ_p, can be divided into two main categories, a conformation-independent shift, δ_{intrin} and a conformation-dependent shift, δ_{calc},

$$\delta_p = \delta_{intrin} + \delta_{calc}. \quad (1a)$$

The calculated conformational shift, δ_{calc}, is given by,

$$\delta_{calc} = \delta_{rc} + \delta_{ma} + \delta_E. \quad (1b)$$

The term δ_{rc} is the chemical shift induced by ring-currents produced by neighboring aromatic rings and δ_{ma} is the chemical shift due to local magnetic anisotropy effects. Analytical expressions with adjustable parameters (*vide infra*) have been derived for δ_{rc} (Haigh and Mallion, 1980; Johnson and Bovey, 1958) and δ_{ma} (Ribas-Prado and Giessner-Prettre, 1981). For the calculation of δ_{rc} and δ_{ma} we have used the ring-current and local magnetic anisotropy parameters derived by (Ribas-Prado and Giessner-Prettre, 1981) without adjustment. We call this the GP parameter set. In addition, we have used the recently optimized ring-current parameters by Case (Case, 1995). As in the original article (Case, 1995), we assume that these parameters account for the ring-current as well as the magnetic anisotropy contributions of the aromatic rings. We call this the Case parameter set. The term δ_E is the chemical shift term resulting from polarization by an electric field of the electron density along the chemical bond(s) extending from nucleus p (see Giessner-Prettre and Pullman, 1987). The contribution δ_E was found to be negligible (as for DNA, Wijmenga et al., 1997) and was therefore not included in the final calculations. When using the GP parameter set we then have,

$$\delta_{calc} = \delta_{rc} + \delta_{ma} \quad (2a)$$

and using the Case parameter set,

$$\delta_{calc} = \delta_{rc}. \quad (2b)$$

To obtain δ_{rc} and δ_{ma} we sum over all surrounding rings and magnetically anisotropic groups in the molecule. In case the proton concerns a base proton the effect of the base to which it is attached is excluded. Excluded also are possible magnetic anisotropy terms resulting from the sugar ring atoms and from atoms in the phosphate backbone, since their effect has been shown to be small (Pullman and Giessner-Prettre, 1987). To verify whether for a proton the calculated conformational shift (δ_{calc}) is correct, it should be compared with the conformation-dependent part of the observed chemical shift ($\delta_{conf,exp}$),

$$\delta_{exp} = \delta_{ref} + \delta_{conf,exp}, \quad (3)$$

where δ_{exp} is the observed chemical shift and δ_{ref} an experimentally determined reference value. We derive δ_{ref} from the experimental data by using it as an adjustable parameter (Wijmenga et al., 1997),

$$\delta_{ref} = \frac{1}{N}\Sigma(\delta_{exp} - \delta_{calc}) \quad (4)$$

where the sum is over all N protons of a certain type. As a first approximation the same δ_{ref} is taken for H1′, H2′, etc., protons, but depending on the statistics these categories may be subdivided (*vide infra*). The term δ_{ref} is in fact the experimental counterpart of δ_{intrin} and we can take δ_{intrin} equal to δ_{ref}. The total calculated chemical shift δ_p is then given by,

$$\delta_p = \delta_{ref} + \delta_{calc} \quad (5)$$

In addition to comparing the calculated and experimental conformational shifts, one can now compare the total calculated chemical shift (δ_p) with the observed chemical shift (δ_{exp}) to assess the validity of the chemical shift calculations. Henceforth, predicted chemical shifts refer to either the calculated conformational shift (δ_{calc}) or the total calculated chemical shift (δ_p, Equation 5).

We have assessed the quality of the correspondence between experimental and calculated chemical shifts via a number of statistical parameters. First, a Gaussian or normal distribution function (Bevington, 1969) was fitted to the distribution of $\delta_{conf,exp}$-$\delta_{calc,}$ giving a standard deviation $RMSD_G$. When no systematic deviations are present in the calculations the mean of this distribution is expected to be 0. Secondly, the root-mean-square deviation (RMSD) of $\delta_{conf,exp}$-δ_{calc} was calculated. Thirdly, the linear correlation between the calculated conformational shifts (δ_{calc}) and the experimental conformational shifts ($\delta_{exp} - \delta_{ref}$), or alternatively between the total calculated chemical shifts (δ_p, Equation 5) and the experimental chemical shifts (δ_{exp}), was determined via the Pearson correlation coefficient (R) (Bevington, 1969). The term R is a dimensionless index that ranges from -1.0 to 1.0, which reflects the extent two data sets are linearly related. If R has a value of 1, there is a perfect linear relationship, a value of -1 indicates a perfect inverse linear relationship. Ideally, a plot of δ_{exp} versus δ_p or δ_{calc} versus $\delta_{exp} - \delta_{ref}$ should yield a perfect linear relationship, in which case R equals 1. Finally, the linear correlation between δ_{exp} and δ_p (or δ_{calc} and $\delta_{exp} - \delta_{ref}$) was assessed by determining the intercept and slope of a linear regression line through the data points (Bevington, 1969). If the two data arrays are identical, as they ideally should be for δ_{exp} versus δ_p or δ_{calc} versus $\delta_{exp} - \delta_{ref}$, the intercept is 0 and the slope is 1.

The main aim is to assess the quality of the model. It is therefore important to consider the potential error sources and how these will affect the error distribution and subsequently the statistics as described above. One can distinguish two types of error: (1) errors related to the experimental data and (2) errors in the model. The first can result from errors in the measurement of the chemical shift, the chemical shift referencing, misassignment of resonances, and most importantly from errors in the used structure(s). These errors will increase the RMSD or can lead to 'outliers' in the database, incidental large deviations. The errors in the model will lead to systematic deviations, also leading to an increased RMSD as well as a reduced linear correlation between predicted and observed chemical shifts. Thus, the RMSD can be viewed as a measure of both the precision and accuracy of the predicted chemical shifts. The Pearson R and the intercept and slope indicate how well the model correlates with the experimental data, i.e., these parameters can be viewed as measures of the accuracy of the predicted chemical shifts.

In addition to determining the statistics for the complete database, we have also determined the statistics for a database where outliers are removed. We call this the reduced database. This was done to investigate the effect of different secondary structures on the chemical shift prediction. A second motivation was that least-squares statistics (including RMSD and correlation coefficients) are quite sensitive to a few outlying data points (see Press et al., 1989; Figure 14.6.1). Robust statistics can remedy this problem, for example, by simply removing the outlying data

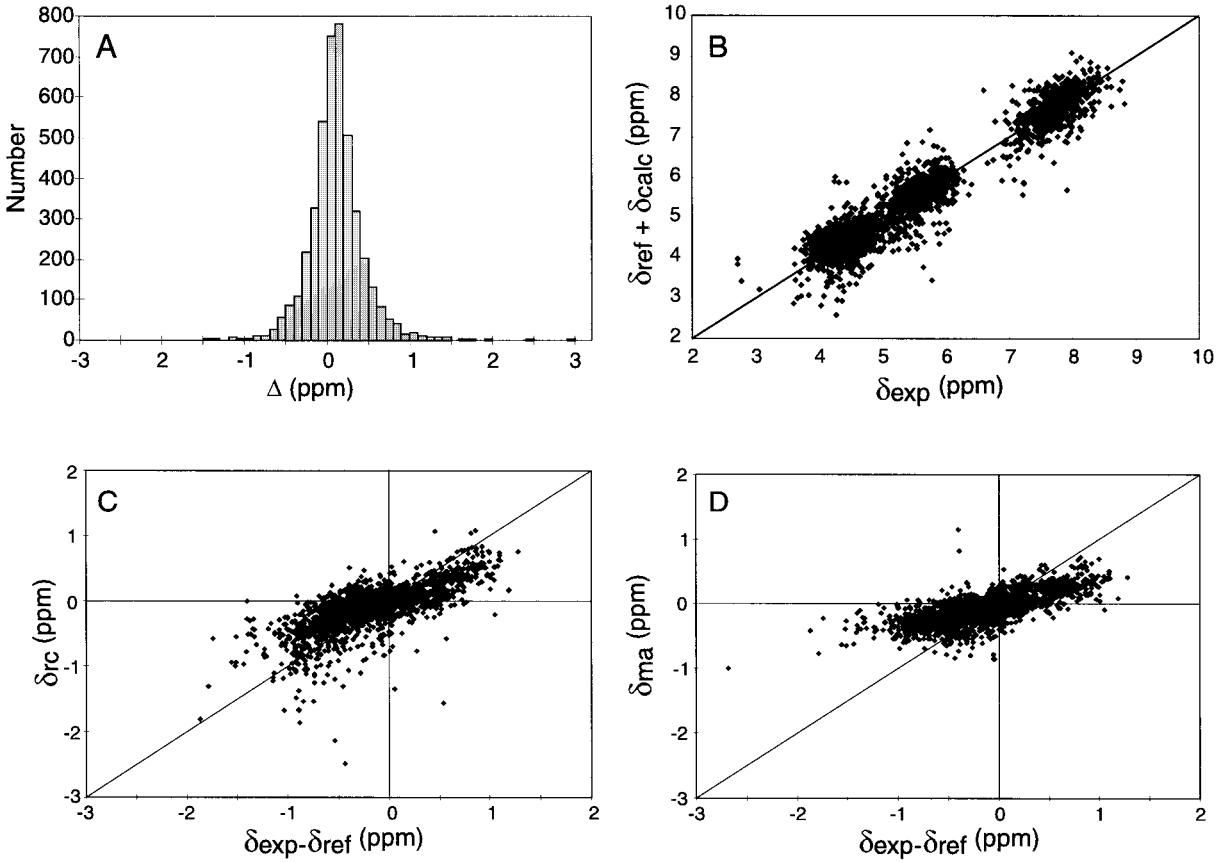

Figure 1. Correlation between predicted shifts and observed shifts for all 4507 non-exchangeable protons. For the shift calculations the GP parameter set was used (see Methods). (A) Distribution of errors between predicted and observed chemical shifts ($\Delta = \delta_{exp} - \delta_{ref} - \delta_{calc}$) The distribution is described by a Gaussian distribution with standard deviation 0.14 ppm and a midpoint of 0.003 ppm (see text). (B) Total calculated shifts ($\delta_{ref} + \delta_{calc}$) versus the observed chemical shifts (δ_{exp}). The trend-line for these data points (solid diagonal line) has the form of $y = 0.08 + 0.97x$ (see text). (C) Ring-current contribution to the calculated conformational shifts. (D) Magnetic-anisotropy contribution to the calculated conformational shifts.

points beyond a certain cutt-off (Press et al., 1989, p. 593 bottom). We have essentially adopted such a procedure. We simply exclude data points for which the difference between predicted and experimental shifts is more than two times the standard deviation and repeat this procedure until no further outliers are removed with a maximum of 10 iterations. For a normal distribution such a procedure does not significantly reduce the number of data points and the recalculated standard deviation does not change significantly. However, when the data set is spoiled by 'outliers' this procedure will remove them and lead to a more correct statistics.

The caveat is that this procedure holds the potential danger that systematic deviations from the applied model are also disregarded in the reduced database. It is therefore important to ascertain that such systematic deviations are not disregarded. A number of extra checks have therefore been invoked. First, we consider the number of data points that are removed. Secondly, we examine the conformational shift range of the complete and reduced data base. Thirdly, we look for general trends, i.e., we check whether the removed data points belong to a category of shifts for which outliers are expected. Finally, we check specific outliers and determine whether they really can be attributed to experimental deviations.

Results and discussion

The nucleic acid molecules used to generate the proton chemical shift database are summarized in Table 1. The set contains RNA molecules for which the ex-

Table 1. Names and PDB entries of the 28 structures used in the analysis

Name	Size[a]	family[b]	rmsd[c]	PDB[d]	Reference
Sarcin/Ricin loop	29	6	1.11 (6)	1SCL	(Szewczak and Moore, 1995)
Rev Responsive element, free	30	5	2.83 (20)	1EBR, 1EBQ	(Peterson and Feigon, 1996)
Rev Responsive element, bound	30	5	2.60 (20)	1EBS	(Peterson and Feigon, 1996)
Pseudoknot VPK	34	1	NR	1RNK	(Shen and Tinoco, 1995)
Pseudoknot APK	32	1	2.0 (4)	1KAJ	(Kang et al., 1996)
Pseudoknot APKA27G	32	1	2.5 (12)	1KPD	(Kang and Tinoco, 1997)
HDV ribozyme loop	19	10	1.02 (50)	1ATO	(Kolk et al., 1997)
24 nucleotide hairpin	24	1	0.6* (7)	1RHT	(Borer et al., 1995)
GU wobble pair (1) duplex	8	30	0.5 (30)	1QES	(McDowell et al., 1997)
GU wobble pair (2) duplex	8	30	0.37 (30)	1QET	(McDowell et al., 1997)
GU wobble pair (3) duplex	8	30	0.65 (30)	1GUC	(McDowell et al., 1997)
GA pair (1) duplex	8	1	0.41 (15)	1MWG	(Wu et al., 1997)
GA pair (2) duplex	8	1	0.52 (14)	1MIS	(Wu and Turner, 1996)
GA pair (3) duplex	8	1	0.7 (11)	1YFV	(Santalucia and Turner, 1993)
FMN-RNA aptamer	35	5	1.24 (4)	1FMN	(Fan et al., 1996)
ATP-binding RNA aptamer	36	10	1.53 (10)	1RAW	(Dieckmann and Feigon, 1997)
HIV-2 TAR argininamide complex	30	20	1.32 (20)	1AJU	(Brodsky and Williamson, 1997)
Helix I from 5S RNA	25	6	1.10 (5)	1ELH	(White et al., 1992)
GC duplex	6	1	0.17 (16)	1PBM	(Popenda et al., 1997)
UGAA tetraloop	12	13	1.10 (13)	1AFX	(Butcher et al., 1997)
GU-quadruplex	6	1	0.71 (8)	1RAU	(Cheong and Moore, 1992)
3′-hairpin of TYMV Pseudoknot	23	10	1.65 (10)	3PHP	(Kolk et al., 1998b)
binding site S8 in 16S rRNA (E. Coli)	23	6	1.24 (10)	1BGZ	(Kalurachchi and Nikonowicz, 1998)
P5 Helix of Group I intron	14	19	0.5 (19)	1C0O	(Colmenarejo and Tinoco, 1999)
Binding site for Phage GA coat protein	21	12	2.0 (21)	17RA	(Smit and Nikonowicz, 1998)
TYMV pseudoknot	44	24	2.1 (10)	1A60	(Kolk et al., 1998; Kolk et al., 1998a)
Lead-dependent Ribozyme	30	26	1.44 (25)	1LDZ, 2LDZ	(Hoogstraten et al., 1998; Legault et al., 1998)
Hairpin from 18S rRNA	19	15	1.05 (15)	1UUU	(Sich et al., 1997)
CUUG tetraloop	12	5	1.3 (13)	1RNG	(Jucker and Pardi, 1995)

[a]Number of nucleotides in the RNA sequence.
[b]Number of NMR structures in PDB file.
[c]Average pairwise RMSD (Å) from all residues from (X) structures.
[d]PDB accession code.
NR: not reported, *only base atoms from stem residues.

perimental chemical shifts are available and one (8 entries) or a set of structures (20 entries) has been deposited in the PDB-database. Also, for each entry in Table 1 the reported RMSD is given. The RNA conformational space is extensively scanned by this set of molecules, because it comprises a variety of different structures: double helices, hairpins, loops, bulges, aptamers, pseudoknots, basepair mismatches, and a quadruplex. We have used the experimental chemical shifts of all the non−exchangeable protons. No corrections were made for differences in measuring conditions. Small errors in referencing may result when using data of experiments conducted at different temperatures and lead to variations in the measured chemical shifts, which translate to an increase of the RMSD of δ_{calc} and $\delta_{exp,conf}$, but would to first order not effect δ_{ref} (Equation 4). A substantial error resulting from the use of reference substance other than the usual TSP or from an error in the referencing would show up in the statistical analysis as a deviating point in the comparison between δ_{calc} and $\delta_{exp,conf}$, or in the derivation of δ_{ref} (Equation 4). The database contains ca. 4500 proton shifts. The shifts were calculated from the known three-dimensional structure, as described in the methods section. When for a molecule more than one structure was available that fulfilled the NMR constraints (e.g., a set of structures), the average of the shifts calculated for a particular type of proton was

used. The reference chemical shifts, δ_{ref}, were obtained using Equation 4 and are presented in Tables 4 and 5.

Overall results of the chemical shift predictions

The overall correspondence between the experimental and predicted ^{1}H shifts, using the GP parameter set for ring-current and magnetic anisotropy (see Methods), is directly visible from the error distribution ($\Delta = \delta_{exp} - \delta_{ref} - \delta_{calc}$) shown in Figure 1A. The error distribution is well described by a Gaussian distribution (correlation coefficient 0.99) with a standard deviation as small as 0.14 ppm and an average value of 0.003 ppm. The good linear correlation between the total calculated chemical shifts ($\delta_{ref} + \delta_{calc}$, Equation 5) and the observed chemical shifts (δ_{exp}) is illustrated in Figure 1B. The regression line through these data points can be described by the equation $y = 0.08 + 0.97x$, indicating that the calculations correctly reproduce the experimental values. Also, the Pearson coefficient, R, for the total calculated (δ_p) and experimental chemical shifts (δ_{exp}) is high, 0.97. Finally, we note that the two major contributions to δ_{calc}, the ring-current effect, δ_{rc}, and the magnetic anisotropy effect, δ_{ma}, amount to ca. 60% and 40%, respectively (Figures 1C and 1D).

We have also used the recently optimized ring-current parameters (Case, 1995) to calculate δ_{rc}. As in the original article (Case, 1995), we assume that these parameters account for the ring-current as well as the magnetic anisotropy contributions of the aromatic rings. The error distribution is, as for the GP set, well described by a Gaussian distribution (correlation coefficient 0.98) with a standard deviation of 0.13 ppm and an average value of 0.012 ppm. The Pearson correlation between δ_p and δ_{exp} is 0.97, while the slope and intercept are 0.98 and 0.08, respectively.

The Case parameter set yields a slightly lower standard deviation (0.13 ppm) than the GP parameter set (0.14 ppm), based on fitting to a Gaussian distribution. The linear correlation coefficients are quite similar for both sets. The average error, indicating the deviation from zero, shows that the GP set (0.003 ppm) is more accurate than the Case parameter set (0.012 ppm), although the differences are small. Henceforth, the discussion of the predicted chemical shifts will be based on the GP set.

The good correspondence between the predicted and experimental chemical shifts warrants a closer consideration of their behavior in the context of different secondary structures. Helices usually have well-defined structures, which can be reliably derived from NMR data (Allain and Varani, 1997). The database is therefore expected to contain only well-defined and correct helix structures. In contrast, non-helical structures such as loops and bulges frequently encountered in RNA structures (for exact definition we refer to Westhof and Fritsch, 2000) may have a well-defined structure, but can sometimes also be flexible (see for examples the review by Hilbers et al., 1994; Williamson and Boxer, 1989). This means that the database is expected to contain well-defined and correct structures of bulges and loops, but also well-defined bulges and loops for which insufficient NMR data was available to derive their structure correctly. It is well documented that the usual number of NMR constraints poorly defines the conformation of an isolated dinucleotide (see, for example, Wijmenga and van Buuren, 1998; van de Ven and Hilbers, 1988a; Hilbers et al. 1991). Finally, for flexible bulges and loops the NMR derived sets of structures may also not correctly reflect the actual structural ensemble. The last two groups of structures in the database are thus expected to lead to structural 'outliers'. The structural 'outliers' may show up as 'outliers' in the comparison of predicted versus experimental chemical shifts. Based on these considerations one expects to find that the database of helix shifts to contain less 'outliers' than the database of shifts of bulges and loops.

In Table 2, the statistics (RMSD, Pearson R, slope and intercept) are shown for helices, bulges, and loops for the complete database. For the generally well-structured helices the correspondence between predicted and experimental shifts is excellent, RMSD is 0.08 ppm and Pearson R is 0.96. In other words, the applied model leads to chemical shifts which correspond quite well with the experimental data. For bulges and loops the RMSD is substantially higher and the Pearson R lower. We note that the conformational shift range covered by protons in a helical environment is roughly the same as that for bulges and loops (Table 2). Hence, the larger RMSD and reduced R for bulges and loops can not be attributed to a larger conformational shift range and thus to a problem with the model. Instead, it must have some other cause, for example increased structural disorder.

In the bottom part of Table 2, the statistics are given for the reduced database. It is evident that in this case all RMSD values as well as the Pearson R are essentially the same irrespective of their structural origin. Note also that the data base for helix residues is

Table 2. Statistics of chemical shift predictions according to secondary structure[a]

	RMSD[b]	Pearson[b]	Slope[b]	Intercept[b]	$\delta_{conf,exp}$[f]
Complete[c]					
All[d]	0.16	0.79	0.79	0.02	2.53
Helix[d]	0.08	0.96	0.97	−0.003	2.51
Bulge[d]	0.15	0.72	0.59	−0.07	2.74
Loop[d]	0.16	0.67	0.56	−0.07	2.46
Reduced[e]					
All[d]	0.09 (75)	0.95	0.94	−0.005	2.35
Helix[d]	0.08 (98,5)	0.96	0.95	−0.003	2.42
Bulge[d]	0.10 (62)	0.94	0.93	0.007	2.36
Loop[d]	0.09 (73)	0.94	0.91	0.004	2.24

[a]Chemical shifts were predicted using the GP parameter set (see Methods).
[b]RMSD the root-mean-square-difference (see Methods), Pearson correlation values (see Methods) and Slope and Intercept of the linear-regression line (see Methods) of the conformational shifts, i.e. $\delta_{exp} - \delta_{ref}$ versus δ_{calc}.
[c]Complete data set, no outlying data points were discarded (see Methods).
[d]The statistics split according structural element: all structural elements, Helix (24% of residues), Bulge (19% of residues), or Loop (22% of residues); for definition of bulges and loops see text.
[e]Reduced data set; the percentage of data points left after removal of the outliers is indicated between parentheses behind the RMSD values. Note that for helix residues the complete and reduced data set are essentially the same.
[f]The $d_{conf,exp}$ range represent the experimental conformational chemical shift ($\delta_{exp} - \delta_{ref}$). The ranges are derived with a 99% confidence limit, i.e., with 3 times standard deviation. Despite the reduction in size, the conformational shift range (ppm) for all structural elements (All) is essentially the same in the reduced data set as in the complete data set; see Table 3 for a distinction according to proton type.

essentially unaffected (98.5% data points left), showing that the procedure of removing 'outliers' by itself does not overly improve the RMSDs or Pearson R. For bulges and loops about 38% and 27% of the data points were removed, substantial numbers. However, the conformational shift range in the complete and reduced data base is the about same (Table 2). Thus, despite the reduction in size, the model still has to explain essentially the same conformational shift range. In addition, it was checked whether the removed data points could be attributed to structural 'outliers'. One example of a structural 'outlier' concerns the H3′ shift of the G20 residue in the UUCG tetraloop in the FMN-aptamer (Fan et al., 1996), discussed in the Appendix. Another such example is that of the H2′ of the C13 residue located in the flexible loop region of the 24 nucleotide hairpin studied by Borer and coworkers (discussed in the section *Structure Validation and Refinement*). In summary, all these aspects indicate that the removed data points in the shift data base can be attributed to structural 'outliers' and that the improved RMSD and Pearson *R* found for bulges and loops on removing 'outliers' is genuine.

In summary, the most important conclusion is, first, that the chemical shifts of helix segments can be predicted with an RMSD of 0.08 ppm and Pearson *R* of 0.96 as follows from the complete data base. Secondly, the data strongly suggest that the same applies for chemical shifts of non-helical elements. The increased RMSD and lower Pearson *R* seen in the complete data base for some bulges and loops can most likely be attributed either to their higher flexibility or that their structure is less well determined by the NMR data.

In view of these results it is interesting to investigate how the RMSD varies per proton type. This is shown in Table 3 using the complete and reduced database. As can be seen, for all proton types the RMSDs for the reduced data base are quite similar, whether one considers the well-structured helices or all structural elements, confirming for each proton type separately the conclusions drawn for the bulges and loops. It is emphasized again, as indicated in the legend to Table 3, that the conformational shift range ($\delta_{conf,exp}$) is generally not much reduced. There are two notable exceptions. The conformational shift range for H4′ and H5″ is more decreased than for the other protons. This is because there are a few very strongly shifted resonances present in the complete data base that are removed in the reduced data base for both the H4′ and H5″. For the H4′, the reduction in shift range is caused for the most part by the removal of one single data point, the H4′ of G11 of the ATP aptamer. This resonance has an extreme chemical shift of 1.67 ppm ($\delta_{conf,exp}$ −2.7 ppm). Removal of this point alone reduces the conformation shift range already to 1.09 ppm (one point out of 492 H4′ data points). It is back calculated for the set of structures between 0.1 ppm and −1.3 ppm, and is removed, because it lies outside the set cut-off. As discussed in the Appendix, in the context of the extremely shifted H2′ of the same residue (Dieckmann and Feigon, 1997), such a deviation between experimental and calculated shift is structurally relatively minor, because of the 1/r3 dependence of the ring current shifts. The list of H5″ protons contains four H5″ of the C residue in a UUCG tetra-loop present in different RNA molecules. The H5″ of this residue has the unusual chemical shift of

Table 3. Statistics of the chemical shift predictions according to proton type[a]

RMSD[b]	H1′	H2′	H3′	H4′	H5′	H5″	H2	H5	H6	H8
Complete[c]										
All[d]	0.16	0.20	0.13	0.08	0.13	0.08	0.29	0.20	0.17	0.20
Helix[d]	0.09	0.10	0.08	0.06	0.09	0.06	0.12	0.11	0.09	0.09
$\delta_{conf,exp}$[f]	2.13	1.68	1.44	1.32	1.38	1.32	2.34	1.74	1.2	2.28
Reduced[e]										
All[d]	0.09	0.10	0.09	0.08	0.09	0.08	0.11	0.11	0.10	0.09
Helix[d]	0.08	0.09	0.08	0.06	0.09	0.06	0.11	0.11	0.09	0.09
$\delta_{conf,exp}$[f]	1.94	1.50	1.14	0.72	1.08	0.72	2.4	1.44	1.14	2.04

[a]The chemical shifts were predicted with the GP parameter set.
[b]RMSD between $\delta_{exp} - \delta_{ref}$ and δ_{calc} (see Methods).
[c]The complete data base, no outlying data points are removed.
[d]Two sets of RMSDs were derived one for all structural elements and one for helical parts of the molecules.
[e]The reduced data set, i.e. outlying data points are removed (see Methods). For helices the complete and reduced data base are essentially the same (Table 2). Consequently, the RMSDs are hardly or not affected and the conformational ranges are the same.
[f]The $\delta_{conf,exp}$ range represent the experimental conformational chemical shift ($\delta_{exp} - \delta_{ref}$). The ranges are derived with a 99% confidence limit, i.e. with 3 times standard deviation. When considering all structural elements the conformational shift range $\delta_{conf,exp}$ (ppm) is for most proton types only slightly reduced by the removal of outliers.

2.7 ppm, which is caused by the turn in the backbone between C and G, so that ribose resides over the base of the G residue. Similar downfield shifts are observed in DNA tetra loops (Wijmenga et al., 1997). Here, these shifts enlarge the experimental conformational shift range substantially. These unusual resonances of H5″ are removed in the reduced data set, because they are back calculated at ca. 3.85 to 3.99 ppm, which is outside the cut-off. Their removal accounts for most of the reduction in the conformational shift range. Interestingly, in the corrected structure of the UUCG loop by Allain and Varani (1995), which is discussed in the Appendix in the context of the H3′ of the G residue, these H5″ are correctly back calculated for one family of structures (2.9 ppm). Again the removed data points can with good reason be accounted for by structural effects. The experimental conformational shift range of the H5′ is of the same order as that of the H4′ and H5″ protons. For the resonances of the H5′ protons with large $\delta_{conf,exp}$, the back-calculated chemical shifts fall within the error margins and are thus not removed from the data set. For example, the H5′ of G11 of the earlier mentioned ATP aptamer has a chemical shift of 3.2 ppm ($\delta_{conf,exp} - 1.3$ ppm), which is back calculated as 3.1 ppm.

It can also be seen that for the proton types H1′ to H5′/5″ and H6/8, the helix RMSD values vary little (between 0.06 ppm and 0.10 ppm), despite the fact that they reside in such different structural elements as the ribose sugar or the base. The only exceptions are the H2 and H5 protons, which have a slightly higher RMSD of 0.11 ppm to 0.12 ppm. As will be seen later these protons experience considerably larger conformational shifts than the other protons when in a helix environment and their chemical shifts are thus more sensitive to structural changes.

Comparison of the chemical shift values predicted for DNA with those for RNA shows that the standard deviation between observed and predicted ^{1}H chemical shifts obtained for RNA (0.16 ppm; complete data base Table 2) is similar to that of DNA (0.17 ppm) (Wijmenga et al., 1997). Both DNA and RNA chemical shifts have been calculated with the same expressions for the ring-current and magnetic anisotropy contributions. Both DNA and RNA databases contain a set of diverse structural elements and no corrections were made for differences in measuring conditions nor were parameters optimized to predict the chemical shift. Whether the much better correspondence seen for RNA helices (0.08 ppm) also applies for DNA would require a similar treatment of the DNA database as carried out for RNA.

Analysis of the shifts

The good quality of the results justifies an investigation of the physical origin of the variations in the ^{1}H

chemical shifts of the different types of protons. As will become apparent later, it is convenient to distinguish in δ_{calc} between contributions stemming from their own base, δ_{ib}, and from other fragments, δ_{eb}. In this way the total calculated chemical shift, δ_p, can be delineated as,

$$\delta_p = \delta_{ref} + \delta_{calc} = \delta_{ref} + \delta_{ib} + \delta_{eb}. \quad (6)$$

For base protons δ_{ib} is zero by definition. In Tables 4 and 5 the δ_{ib} and δ_{eb} contributions to the chemical shifts are given for protons in an A-helix environment, together with their experimental chemical shift ranges (δ_{exp}) using the complete database. In a non-A-helix environment the values of δ_{ib}, δ_{eb}, and the experimental ranges δ_{exp} will differ depending on the individual fold. It is therefore not meaningful to tabulate these data as well. Establishing the δ_{ib} and δ_{eb} values for an A-helix environment is important as they can serve in conjunction with Equation 6 as a guide to predict or validate the structural environment of a proton. In the following sections we analyze the experimental and predicted ^{1}H chemical shift of each type of proton in detail.

H1′ resonances

Figure 2A shows a comparison of the calculated conformational chemical shifts (δ_{calc}) of all H1′ protons with their experimental conformational chemical shifts ($\delta_{exp} - \delta_{ref}$, Equation 3). As can be seen, the correspondence between calculated and experimental shifts is good. Note that this requires a different δ_{ref} value for the H1′ of A, G, C, and U (Table 4). The correspondence is significantly worse when just one δ_{ref} value for all H1′ protons is calculated (RMSD 0.16 ppm using different δ_{ref} values; RMSD 0.20 ppm when using one δ_{ref} value). This behavior was also observed in DNA (Wijmenga et al., 1997). Figure 2A also shows that the H1′ protons may exhibit considerable conformational shifts (ranging from −2 ppm to 1.3 ppm). The data points, in Figure 2A, between 0 and ca. 1 ppm are mostly from H1′ protons present in an A-type helix (see also Table 4). Data points outside this region are from protons present in loops and bulges or from protons in non-canonical base pairs.

The data points for H1′ protons in an A-helix environment cluster into four groups depending on the nucleotide (A, G, C or U) in which the H1′ proton resides (Figure 2A). The largest conformational shifts occur in the purines (average δ_{calc}: A: 0.89 ppm; G: 0.63 ppm), while the pyrimidines exhibit smaller conformational shifts (average δ_{calc}: C:0.23 ppm; U:0.12 ppm). The physical reason for this separation can readily be understood when one considers the relative importance of the δ_{ib} and δ_{eb} contributions to δ_{calc}. Figure 2B shows for each nucleotide, the conformational shift (δ_{ib}) of H1′ as a function of the χ-angle. As can be seen the δ_{ib}-curves are similar in shape, but run for each nucleotide at a different level. Because

Table 4. Overview of predicted and experimental chemical shifts for sugar protons in an A-helical environment[a]

		δ_{ref}[b]	δ_{ib}[c]	δ_{eb}[d]	δ_{exp}[d]
H1′	A	5.02	1.17	−0.59–0.03	5.71–6.18
	G	5.20	0.78	−0.48–0.18	5.43–6-13
	C	5.23	0.53	−0.73–0.12	5.12–5.85
	U	5.57	0.31	−0.49–0.11	5.42–5.70
H2′	A	4.66	−0.08	−0.32–0.28	4.31–4.86
	G	4.66	−0.18	−0.32–0.39	4.18–4.82
	C	4.66	−0.20	−0.45–0.43	3.83–4.84
	U	4.66	−0.18	−0.35–0.37	4.18–4.87
H3′	A	4.62	0.02	−0.23–0.30	4.43–4.92
	G	4.62	−0.01	−0.41–0.36	4.21–4.87
	C	4.62	−0.09	−0.39–0.27	4.19–4.73
	U	4.62	−0.14	−0.23–0.31	4.36–4.72
H4′	A	4.35	0.21	−0.31–0.23	4.27–4.67
	G	4.35	0.15	−0.30–0.22	4.29–4.69
	C	4.35	0.07	−0.26–0.19	4.26–4.55
	U	4.35	0.02	−0.35–0.33	4.14–4.60
H5′	A	4.38	0.18	−0.39–0.18	4.25–4.81
	G	4.38	0.12	−0.31–0.18	4.12–4.74
	C	4.38	0.05	−0.23–0.24	4.14–4.75
	U	4.38	0.01	−0.33–0.38	4.04–4.83
H5″	A	4.09	0.13	−0.31–0.32	3.88–4.50
	G	4.09	0.08	−0.28–0.29	3.91–4.48
	C	4.09	0.03	−0.26–0.23	3.82–4.43
	U	4.09	0.07	−0.29–0.25	3.84–4.46

[a]The chemical shifts were calculated using the GP parameter set using all protons in the database (see Methods).
[b]The reference value, δ_{ref}, was calculated by means of Equation 4.
[c]The term δ_{ib} is the conformational shift caused by the proton's own base; it was calculated from a mono-nucleotide using the GP parameter set; the value given is for a glycosidic angle, χ, at 200°, the usual value in an A-helix.
[d]The δ_{eb} ranges represent the calculated conformational shifts caused by neighboring nucleotides ($\delta_{eb} = \delta_{calc} - \delta_{ib}$, see Equations 2a and 6) in an A-helix, for residues in a canonical base-pair flanked on the 5′- and 3′-side by a canonical base-pair. The δ_{exp} ranges represent the experimental chemical shift regions for protons in an A-helix environment. Both the δ_{eb} and δ_{exp} ranges are derived with a 99% confidence limit, i.e., with 3 times standard deviation.

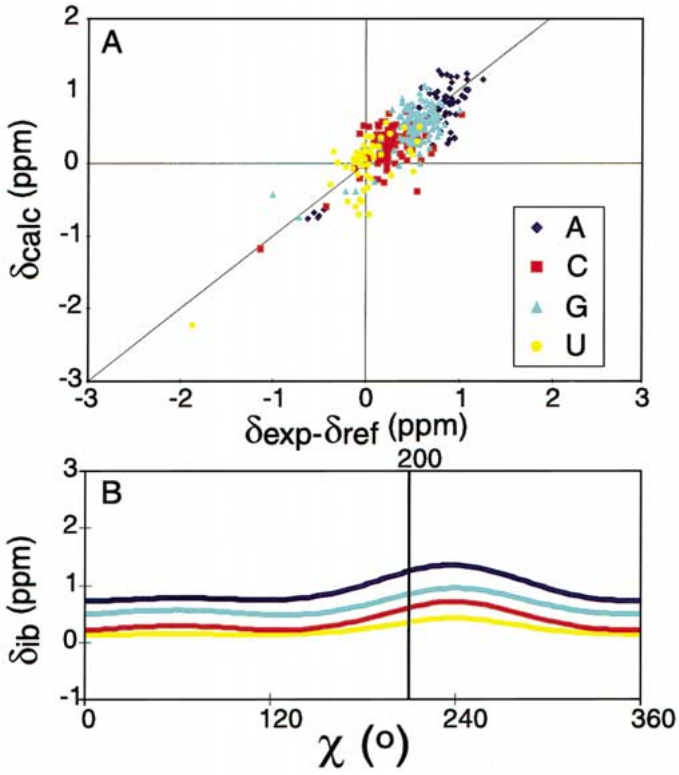

Figure 2. Analysis of H1′ chemical shifts. (A) Correlation between calculated (δ_{calc}) and observed ($\delta_{exp} - \delta_{ref}$) conformational shifts. All H1′ protons in the database are shown and color coded according to residue type; the GP parameter set was used for the calculation of δ_{calc}. (B) Calculated shift (GP parameter set) of H1′ induced by its own base (δ_{ib}) as a function of the glycosidic torsion angle χ for a mono-nucleotide with the sugar in the N-puckered state. The color-coding is the same as in (A).

the H1′ proton points towards the side of the base, the H1′ spin is deshielded by the ring-currents of its own base ($\delta_{ib} > 0$). For example, at $\chi = 200\,^{\circ}$C (normal value in A-helix RNA), δ_{ib} is relatively large for A and G, 1.17 ppm and 0.78 ppm, respectively, while for C and U δ_{ib} is smaller, 0.53 ppm and 0.31 ppm, respectively. In contrast to δ_{ib}, the contribution of neighboring fragments, δ_{eb}, is much smaller (mean value is ~ -0.3 ppm) and has mostly a shielding effect on the chemical shift of H1′ protons ($\delta_{eb} < 0$, Table 4). In conclusion, the major contribution to the conformational shift of H1′ protons comes from their own base and this contribution is mainly responsible for the clustering in the four groups seen in Figure 2A.

That the contribution δ_{eb} of neighboring fragments is relatively small (~ -0.3 ppm) is not surprising, since in an A-type helix the H1′ proton is directed away from both its 5′- and 3′-neigboring base. This may also explain that no correlation of δ_{eb} was observed with the nature of either the 3′- or 5′-neighboring base. In contrast to an A-helix, in a B-DNA helix the H1′ proton is positioned below the base of its 3′-neighbor (when viewed in the 3′ to 5′ direction). Indeed, in a B-DNA helix the conformational shifts, δ_{eb}, are larger and correlate with the nature of the 3′-neighboring base ($\delta_{eb} = \delta_{3'b}$ with values of -0.6 and -0.3 ppm for purines and pyrimidines, respectively) (Wijmenga et al., 1997).

H2′ resonances

The H2′ resonance positions are predicted with good precision (Table 3). In an A-helix environment the resonance positions of the H2′s of pyrimidines and purines strongly overlap (Table 4, δ_{exp}). Analysis of the different contributions to the conformational shifts shows that both δ_{ib} as well as δ_{eb} are essentially the same for all H2′ protons (Table 4). In earlier studies (van de Ven and Hilbers, 1988; Altona et al., 2000) it was found that in DNA the H2′ shifts of the pyrimidines and purines cluster in two separate distributions with midpoints of 2.0 ppm and 2.6 ppm, respectively. This difference between the RNA and DNA results is

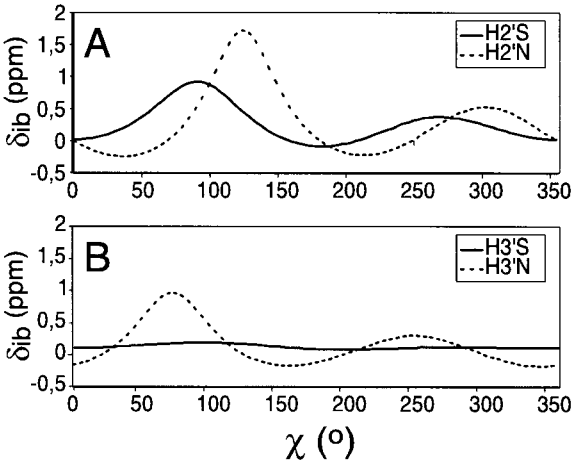

Figure 3. Calculated shifts (δ_{ib}) of H2′ (A) and H3′ (B) in a guanine mono-nucleotide as a function of the glycosidic torsion angle χ and shown for the sugar in an S-puckered state (drawn line) and in an N-puckered state (broken line).

explained by the different puckering of the ribose ring. In DNA, the normal sugar pucker is, C2′-endo, which positions the H2′ proton close to the ring of the base attached to the sugar, leading to a large δ_{ib} for purines and a smaller δ_{ib} for pyrimidines (Wijmenga et al., 1997). In RNA, the sugar pucker is, C3′-endo, and the H2′ proton is directed away from its base. Consequently, its resonance position is only marginally influenced by its own base (Table 4, δ_{ib}) when in the χ-angle is in the normal *anti*-orientation. However, as Figure 3A shows, large down-field shifts are expected for H2′ resonances in north-puckered sugars when the χ-angle is between 90° and 150°, and in south-puckered sugars when this angle is between 60° and 120°. In the Appendix, an example of such a shifted H2′ is discussed.

The H2′-position in a RNA double helix is such that its resonance is expected to be influenced by neighboring base-rings. However, investigation of possible sequential patterns did not show a clear correlation of the experimental H2′ shifts with the nature of either the 3′- or 5′-neighboring base. The only distinct pattern observed, concerned the H2′ protons shifts from nucleotides at the 3′- or 5′-end of a helix. The H2′ protons at the 3′-end of a helix have relative low chemical shifts (< 4.16 ppm), whereas the H2′ protons at the 5′-end of a helix have relative high chemical shifts (> 4.82 ppm) compared with the normal H2′ chemical shifts in an A-helix, 3.85–4.85 ppm (Table 4). This effect is also referred to as the phosphate effect, because an OH-group is attached to C3′ of the 3′-terminal ribose, while at the 5′-end the sequence is terminated via a triphosphate attached to C5′ of the terminal ribose. At first sight one would therefore attribute the difference in H2′ shift to the electric field effect (δ_E). As for DNA, we find that the contribution of the electric field is small compared to the effects from ring currents and magnetic anisotropy and it was therefore left out in the final calculations. Interestingly, the correlation between predicted and observed chemical shifts remains good for all of the terminal H2′ protons, i.e. the same δ_{ref} can be used in the calculations even when the electric field effect is disregarded. In other words, conformational variations may account for the so-called phosphate effect. These conformational differences may for instance arise from a higher degree of conformational freedom of the terminal residues, e.g., reduced interactions with neighboring bases or mixed sugar puckers which will affect the H2′ position. We do not have a general definite explanation, but the effect can at least partly be explained from conformational differences.

H3′ resonances

The H3′ resonance positions are predicted with high precision (Table 3). In A-type helices the conformational shifts for H3′ protons are generally quite small (average $\delta_{calc} = -0.1 \pm 0.2$ ppm, Table 4), suggesting that within an A-helix the H3′ chemical shifts are nearly conformation independent. This is, however, not true for other structural elements.

As Figure 3B shows, H3′ resonances may exhibit appreciable conformational shifts. The H3′ resonance position of a particular nucleotide depends, like the H2′ chemical shift, on the orientation of its own base with respect to its sugar, i.e., on the value of the glycosidic torsion angle, χ. This χ-angle dependency of the H3′ proton shifts is not found in DNA (Wijmenga et al., 1997). In DNA, the normally S-puckered sugar ring turns the H3′ away from the base. In contrast, the usually N-puckered sugars in RNA will place the H3′ proton closer to the base-ring, which explains the dependency of the H3′ chemical shifts on the glycosidic torsion angle. Thus, both the H2′- and the H3′ resonances of RNA may strongly depend on the χ-angle, although somewhat differently. The largest H2′ and H3′ shifts occur at different χ-angles, i.e. at 120° and 75°, respectively (Figures 3A and 3B). Down-field shifted H2′ and H3′ resonances are usually indicative of unusual χ-angles. In the Appendix we describe an example of the use of experimental and predicted H2′ and H3′ shifts to trace such an unusual χ-angle.

H4′ resonances

The H4′ resonance positions can also be predicted with high precision (Table 3). Within a nucleotide the contribution of the base (δ_{ib}) to the chemical shift of the H4′ proton is different for different bases, however this contribution is small (0.0–0.2 ppm, Table 4). Also, in an A-helix the conformational contribution to the H4′ chemical shift from other fragments (δ_{eb}) is small, -0.03 ± 0.09 ppm (Table 4), and no clear variation in the distribution can be observed. This is not surprising if one considers the position of H4′ in a double helix. It is located at the outside of the helix, away from neighboring bases. However, deviations from the classical helix structure may lead to large shifts as is illustrated in the examples in the Appendix.

H5′ and H5″ resonances

The H5′ and H5″ protons present a picture similar to the H4′ protons. The chemical shifts are predicted with good precision (Table 3). Again the conformational contributions from neighboring fragments in an A-helix are quite small, i.e., 0.03 ± 0.09 and 0.01 ± 0.09 ppm for H5′ and H5″, respectively (Table 4). Within a nucleotide the contribution from the base varies with different χ-angle values. However, the absolute values are small, e.g., when the glycosidic angle is between 0° and 60° the resonance positions of H5′ and H5″ protons are only slightly affected (maximum 0.3 ppm). When the torsion angle γ is in the trans region somewhat larger effects (up to 0.8 ppm) are expected for the H5″ (data not shown). In conformations clearly deviating from an A-helix considerable shifts may occur. An example is discussed in the Appendix.

Base-protons

The shifts predicted for the base protons, H8, H6, and H5 correlate well with the experimental shifts (Table 3). The calculated conformational shifts of base protons located in A-type helices are generally negative (Table 5). This is caused by the shielding effects generated by the ring-currents of neighboring bases. Also, the shifts of H2 can be predicted quite accurately (Table 3). As for the other base-protons, the conformational shift calculated for H2 protons is generally negative and may be quite large (-2.35 to -0.07 ppm, Table 5). A few outstanding shifts of base protons in a non-helix environment are examined in the Appendix to illustrate the quality of the calculations. The origin of the conformational shifts in an A-helix are discussed in the next section.

Table 5. Overview of predicted and experimental chemical shifts for base protons in an A-helix[a]

		δ_{ref}[b]	δ_{eb}[c]	δ_{exp}[c]
H2	A	8.43	−2.35–−0.07	6.25–8.46
H5	C	6.12	−1.43–−0.02	4.82–5.88
	U	6.04	−1.40–−0.23	4.65–5.82
H6	C	8.16	−1.12–0.09	7.18–8.19
	U	8.25	−0.98–−0.06	7.45–8.12
H8	A	8.64	−1.17–−0.12	7.60–8.37
	G	8.10	−1.14–0.07	7.05–8.08

[a]The chemical shifts were calculated using the GP parameter set using all protons in the database (see Methods).
[b]The reference value, δ_{ref}, was calculated by means of Equation 4.
[c]The δ_{eb} ranges represent the calculated conformational shifts caused by neighboring nucleotides ($\delta_{eb} = \delta_{calc}$ see Equations 2a and 6) in an A-helix environment (for residues in a canonical base-pair flanked on the 5′- and 3′-side by a canonical base-pair). The δ_{exp} ranges represent the experimental chemical shift regions for protons in an A-helix environment. Both the δ_{eb} and δ_{exp} ranges are derived with a 99% confidence limit, i.e., with 3 times standard deviation. Both the δ_{eb} and δ_{exp} ranges were derived using a database with outlying data points removed (see Methods).

Experimental conformational shifts of base protons in an A-helix environment

Table 6 summarizes the experimental conformational shifts ($\delta_{conf,exp} = \delta_{exp} - \delta_{ref}$, Equation 3) of the base protons in an A-helix and how they depend on their neighboring residues. We emphasize that the data were derived purely from the observed patterns in the experimental shifts. Because the correlation between δ_{exp} and δ_p is very good, we use the predicted shifts to interpret the physical origin of the experimental patterns.

H5, H6 and H8 protons

The reference values, δ_{ref}, of the H8 resonances of adenine and guanine units differ significantly (Table 5), showing the different influence of these bases on their ring proton resonance position. In combination with the shifts induced in the double helix environment this leads to distinct chemical shift regions for the respective H8 resonances (Figure 4A). Furthermore, we find that the influence of the double-helix environment on the H8 chemical shifts correlates with the nature of the 5′- and not the 3′-neighbor (Table 6). In a similar manner, the resonances from H6

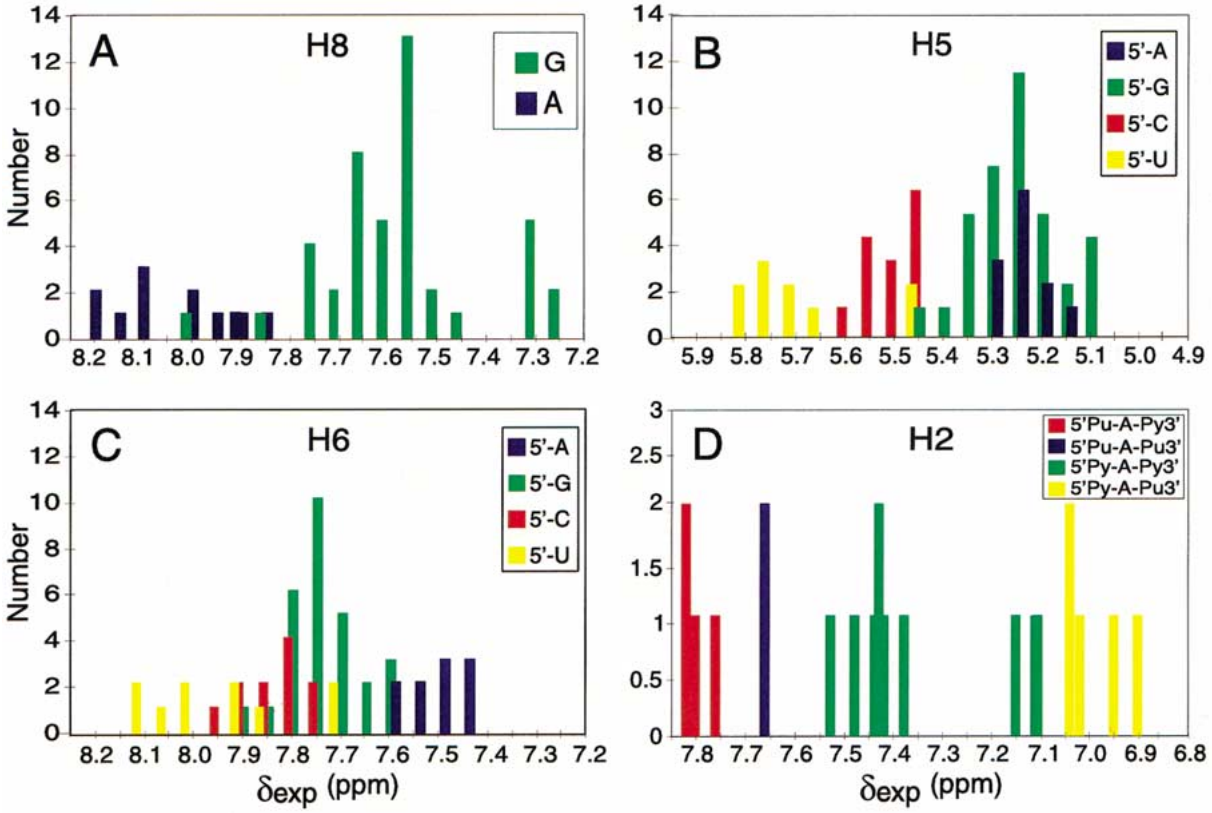

Figure 4. Correlation of experimental shifts of base protons in an A-helix with the type of base and/or neighboring nucleotide (color coding given in inserts). (A) The experimental chemical shift distribution of H8 of A and G. (B) Distribution of the H5 experimental chemical shifts according to 5′-neighbor. (C) Distribution of the H6 experimental chemical shifts according to 5′-neighbor. (D) Distribution of the H2 experimental chemical shifts according to 5′- and 3′-neighbor. The contribution to chemical shift of H2 from the 5′-neighbor actually arises from the base-pairing partner of this 5′-neighboring nucleotide, the 3′-cross-strand residue (see text and legend of Table 6).

and H5 protons show a strong correlation only with the type of 5′-neighbor (Figures 4B and C). We note that no distinction needs to be made between H5 protons belonging to either a C- or U-residue, as their δ_{ref} values are effectively the same (Table 5). The same applies for the H6 protons of a C- and U-residue. The base stacking patterns in an A-helix, shown in Figures 5A-D, explain the dominance of the correlation with the nature of the 5′-neighbor. The H8, H5 and H6 protons are all positioned close to and over the rings of the 5′-nucleotide, and far away from the 3′-neighbor. We find that for each proton type (H5, H6 and H8), protons with 5′-neighboring purines (A and G) resonate up-field from protons with 5′-neighboring pyrimidines (C and U). In other words, purines induce larger conformational shift than pyrimidines; $|\delta_{exp}-\delta_{ref}|$ is larger for protons with purines as 5′-neighbor than with pyrimidines. The observations described above are confirmed by the calculations of the conformational chemical shifts. Ring-current and magnetic anisotropy contributions to chemical shifts of H5, H6 and H8 protons are larger when the 5′-neighbor is a purine than a pyrimidine. This is not surprising as purines have 2 base-rings compared to 1 ring for pyrimidines and thus can create larger ring-currents.

H2 protons

The resonance position of H2 protons in a double helix environment correlates with the nature of both the 5′- and 3′-neighboring nucleotide. Figure 4D shows that H2 protons cluster in separate groups depending on the type of neighboring base. Here we can see that H2 protons with the same 5′-neighbor, that have 3′-neighboring purines resonate up-field from protons with 3′-neighboring pyrimidines. This observation fits with the stacking patterns in an A-helix (Figures 5B and 5C). The H2 base proton of the adenine residues is located in the minor groove in contrast to the H5, H6,

Table 6. Conformational shifts in an A-helix, sequential effects[a]

Atom	$\delta_{conf,exp}$[b] (ppm)		Distribution[c] (ppm)	
H8 A	$\delta_{5'\ Py}$	−0.58	5′-Py-A	7.83–8.29
	$\delta_{5'\ Pu}$	−0.81	5′-Pu-A	7.49–8.17
H8 G	$\delta_{5'\ Py}$	−0.48	5′-Py-G	7.37–7.86
	$\delta_{5'Pu}$	−0.76	5′-Pu-G	6.79–7.88
H5	$\delta_{5'A}$	−0.88	5′-A-C/U	5.10–5.33
	$\delta_{5'G}$	−0.90	5′-G-C/U	4.93–5.46
	$\delta_{5'C}$	−0.59	5′-C-C/U	5.23–5.81
	$\delta_{5'U}$	−0.45	5′-U-C/U	5.31–5.98
H6	$\delta_{5'A}$	−0.71	5′-A-C/U	7.31–7.67
	$\delta_{5'G}$	−0.50	5′-G-C/U	7.46–7.92
	$\delta_{5'C}$	−0.37	5′-C-C/U	7.55–8.10
	$\delta_{5'U}$	−0.30	5′-U-C/U	7.49–8.30
H2	$\delta_{cs3'\ Pu}$	−0.78	3′-csPu // A-3′-Pu	6.85–7.14
	$\delta_{cs3'\ Py}$	−0.33	3′-csPu // A-3′-Py	6.83–8.13
	$\delta_{3'\ Pu}$	−0.64	3′-csPy // A-3′-Pu	6.92–7.82
	$\delta_{3'\ Py}$	−0.30	3′-csPy // A-3′-Py	7.72–7.88

[a]All values are for an A-helix environment, for residues in a canonical base-pair flanked on both the 5′- and 3′-side by a canonical base-pair. Outlying data points have been removed from the data (see Methods).

[b]The term $\delta_{conf,exp}$ (Equation 3) represents experimental conformational shifts caused by neighboring nucleotides (see text). A conformational shift of a certain proton caused by a neighboring nucleotide in the same strand is identified as δ_{pN}, where N stands for the type of nucleotide causing the conformational shift and p indicates the position of N relative to the proton in question. For example, H6 $\delta_{5'A}$ is the conformational shift experienced by a H6 proton from an adenine residue at the 5′-side of the H6 proton. For the conformational shift caused by the base-pairing partner of a neighboring base we use the notation, δ_{cspN} Here, cs stands for cross-strand, p indicates the 3′- or 5′-direction (vide infra), and N stands for the type of nucleotide causing the conformational shift. For these cross-strand shifts the directionality of the cross-strand is used as indicated by p. For example, H2 $\delta_{cs3'pu}$, is the conformational shift experienced by an H2 proton and caused by a purine (A or G) residue at the cross-strand 3′-side of the H2 proton (nomenclature adapted from Wijmenga et al. (1993)).

[c]Chemical shift ranges for protons residing in doublets or triplets in an A-helix. The doublets, for H5, H6 and H8 protons, are two nucleotides in the same strand, e.g. 5′-A-C/U stands for a proton in either a C- or U-residue with an adenine 5′-neighbor; 5′-py-G stands for a proton in a guanine with a pyrimidine as 5′-neighbor. The triplets for H2 protons involve cross-strand stacking interactions. For example, 3′-csPu // A-3′-Py indicates an H2 proton flanked in the same strand on its 3′-side by a pyrimidine and on the cross-strand 3′-side by a purine; // indicates cross-strand stacking interactions. All ranges correspond to a 99% confidence limit, i.e. 3 times standard deviation.

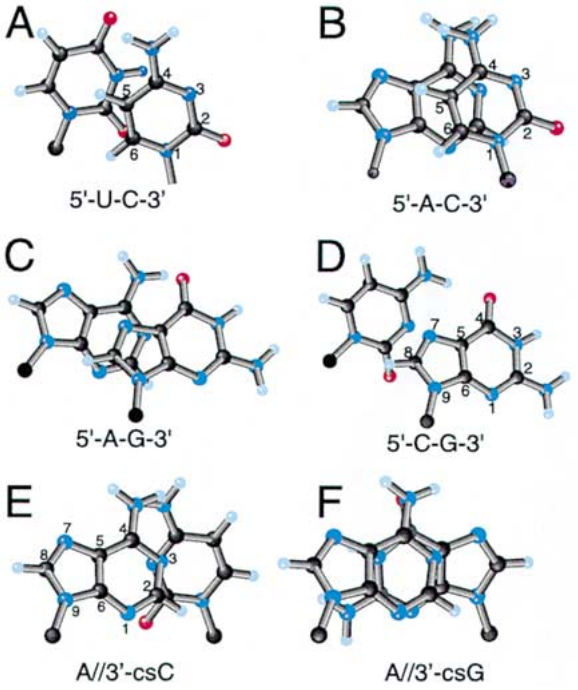

Figure 5. Base stacking patterns in an A-helix. Views are along the helix axis and in the 3′ → 5′ direction along one continuous strand for 5′-U-C-3′ (A), 5′-A-C-3′ (B), 5′-A-G-3′ (C), and 5′-C-G-3′ (D). The stacking of an adenine base with a 3′-cross-stand neighboring C and G is shown in E and F, respectively, using the same directionality as in A to D. The stacking patterns shown in E can be indicated as A // 3′-csC and in G as A // 3′-csG, where // stands for stacking interactions and 3′-cs for the cross-strand base on the 3′-side, using the definitions in the legend of Table 6.

and H8 protons, which are present in the major groove and that is why the H2 chemical shift is affected by its 3′-neighbor. We observe also a correlation with the nature of the 5′-neighbor. In this case, we find that 5′-pyrimidines have a stronger influence on the resonance position than 5′-purines. This is unexpected as purines tend to have larger ring-currents than pyrimidines (see, e.g., the 5′-neighboring influence on the chemical shift of H5, H6 and H8). Stacking patterns in A-helix RNA explain this apparent contradiction. Figures 5E and 5F reveal that the H2 proton is not located close to the base of the 5′-neighor in the same strand, but is in close proximity to the base-pairing partner of this 5′-neighbor in the opposite strand, the cross-strand 3′-neighbor (for nomenclature see the legend of Table 6). This cross-strand stacking nicely resolves this apparent contradiction. It is not the intra-strand 5′-pyrimidine, which causes the extra chemical shift, but the cross-strand 3′-purine. The different influences on

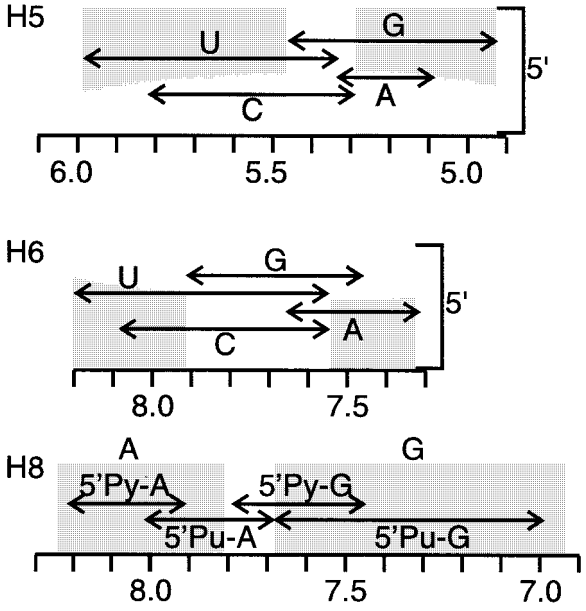

Figure 6. Chemical shift ranges (in ppm) of H5, H6 and H8 in A-helical RNA for residues involved in Watson–Crick base pairs and flanked on both their 5′- and the 3′-side by a canonical base-pair. The ranges (double arrows) correspond to the 95% confidence limit, i.e. 95% chance that an observed shift falls within this range assuming a normal distribution (2 times standard deviation). The top panel shows the shift ranges for H5 of U or C with a 5′-U, C, G, or A neighbor, identified with the capital letter above the double arrow. Similarly, in the middle panel the shift ranges for H6 of C or U are shown with a 5′-U, C, G or A neighbor. In the bottom panel the H8 shift ranges are shown of adenine with either a 5′-pyrimidine neighbor (5′Py-A) or 5′-purine neighbor (5′Pu-A). The shift range of the H8 of guanine with a 5′-pyrimidine neighbor is identified as 5′Py-G and with a 5′-purine neighbor as 5′Pu-G. The grey-shaded boxes indicate the chemical shift ranges where a resonance position can unambiguously be assigned (see text).

the H2 protons as a function of the character of the 5′- and 3′-neighbors are summarized in Table 6.

Assignment of base protons based on chemical shifts

As outlined above the chemical shift of base protons is strongly influenced by neighboring aromatic rings. In an A-helix these influences can be large enough that protons resonate in nearly separate chemical shift ranges depending on the nature of the neighbor, i.e., purine of pyrimidine (see, e.g., Figures 4B and 4D). The separation of the resonance positions can be quite useful in resonance assignments, as we will discuss below.

In Figure 6 the experimental chemical shift distributions have been plotted of H5, H6, and H8 protons in an A-RNA helix. The H5 protons resonate in nearly separate chemical shift ranges depending on the nature of the 5′-neighbor (Figure 6, top panel). This implies that for these protons, on the basis of chemical shift alone, a 5′-neighbor can be assigned. For H5 protons with chemical shifts higher than 5.45 ppm, one can assign a pyrimidine as 5′-neighbor; chemical shifts lower than 5.3 ppm indicate a purine as 5′-neighbor. If a H5 proton resonates down-field from 5.8 ppm, the 5′-neigbor is a uridine; if it resonates up-field from 5.1 ppm, it has a guanine as the 5′-neighbor (see also Table 6).

Also, for H6 resonances a separation depending on the 5′-neighbor can be observed although not as outspoken as for H5 resonances. Still for the shaded regions in the middle panel of Figure 6, a reliable distinction between the presence of a purine or pyrimidine 5′-neighbor can be made. H6 protons with chemical shifts lower than 7.55 ppm have a purine as 5′-neighbor; chemical shifts lower than 7.45 ppm indicate a adenine as 5′-neighbor. H6 protons with chemical shifts higher than 7.9 ppm have pyrimidines on their 5′-side; chemical shifts higher than 8.2 point indicate the presence of an uridine.

The H8 proton chemical shifts nicely separate into two distinct resonance regions, one for the adenine H8 protons the other for the guanine H8 protons (Figure 6, bottom panel, shaded regions). H8 protons resonating up-field from 7.7 ppm belong to a guanine residue, H8 protons resonating down-field from 7.8 ppm are located in an adenine. In addition one can observe a distribution in these shifts depending on the 5′-neighbor. One can assign a pyrimidine as 5′-neigbor in case the adenine H8 is resonating above 8.0 ppm and a purine in case the guanine H8 is resonating below 7.4 ppm.

Unfortunately, on the basis of the proton shifts alone it is not straightforward to identify a resonance as arising from a H5, H6 or a H8 proton. For example, the H5 protons resonate in the same spectral region as the H1′ protons and their resonances may thus be confused with those of the latter. To resolve this ambiguity one may correlate the H5 with H6 resonances via a (^{1}H,^{1}H) TOCSY or COSY-type experiments, thereby discriminating between H5 and H1′ resonances but also between H6 and H2 as well as H8 resonances. Alternatively, using hetero-correlation-type experiments one may correlate, on one hand, the H5-proton and the C5-carbon shifts and on the other hand the H1′- and the C1′-shifts. A discrimination of the H1′- and H5-proton shifts then directly follows from the C1′- and C5-carbon shifts, because C1′ and C5 resonate in

separate spectral regions (C1′: 87–98 ppm, C5: 93–109 ppm). Moreover, because the chemical shift distributions of the C5 resonances of U and C residues are separated to a large extent (C5(U) 99–109 ppm, C5(C) 93–101 ppm) these resonances may be attributed to either U or C. Thus, the combined C5 and H5 resonance positions of a C5-H5 moiety in a pyrimidine may be used to decide whether the signals belong to a di-nucleotide of the type, 5′-Py-U, 5′-Pu-U, 5′-Py-C or 5′-Pu-C, where Py stands for pyrimidine (C and U) and Pu stands for purine (A and G).

The H6, H8 and H2 protons all resonate between 7.1 ppm and 8.2 ppm and can therefore not be distinguished on the basis their chemical shifts alone. However, as mentioned already the H6 resonances can be identified via their correlation to the H5 resonances in a (^{1}H,^{1}H) TOCSY or COSY-type spectrum and thereby distinguished from the H8 and H2 resonances. Furthermore, establishing the position of H6 resonances confirms or refines the assignment of the H5 resonances (*vide supra*). This leaves the analysis of the overlapping H8 and H2 resonances. At this point one may again take recourse to the hetero-nuclear correlation experiments. By correlating the H2-proton and the C2-carbon chemical shifts and the H8- proton and C8-carbon shifts one can discriminate between the H2 and H8 resonances because the C2 and C8 carbons resonate in separate regions, i.e., C2 carbons resonate between 149 and 158 ppm and C8 carbons between 132 and 146 ppm. Subsequently, the H8 resonances can be assigned to di-nucleotide sequences in the double helix in the way discussed above. The possibility to attribute the H5 and H6/8 resonances to certain types of di-nucleotide sequences in a double helix is particularly useful because it limits the number of assignments to a specific residue during the sequential assignment procedure.

Structure validation and refinement

As shown, the proton chemical shifts in RNAs can be predicted with good accuracy and precision from their three-dimensional structure. This suggests that, vice versa, a comparison between predicted and observed ^{1}H chemical shifts may serve as an independent method to validate structures derived on the basis of NOEs and J-couplings.

The RMSD values in Table 2 may serve as yardsticks in qualifying the shift comparison. Helices usually have well-defined structures and are characterized by an overall RMSD value of ca. 0.08 ppm. Therefore, RMSD values for double-helix regions of ca. 0.10 ppm or lower indicate good structures. As can be seen in Table 2, overall bulges and loops give rise to considerably higher RMSD values (0.15 and 0.16 ppm). On the other hand, when ‘outliers’ are removed the RMSD of loop and bulge structures is also around 0.10 ppm. As discussed earlier the higher RMSD values can be attributed to flexibility or lower definition, due to insufficient NMR constraints, of these structural elements. Therefore, we suggest that for loops and bulges an RMSD value of 0.10 ppm or lower can be used to indicate good and well-defined structures, while higher values indicate either flexibility or lack of definition due to insufficient NMR constraints.

The predicted shifts are generally very sensitive to structural changes. This means that a detailed comparison of the predicted versus the experimental shifts can also detect local discrepancies between the derived and the actual structure. The analysis of the shift of the H2′-proton in cytidine 13 (C13) of the 24 nucleotide hairpin studied by Borer and collaborators (Borer et al., 1995; Nooren et al., 1998) provides an example of how the chemical shift can be used in structure validation and subsequent structure refinement.

The C13 residue is part of the hairpin loop, which closes the stem. The loop region is less well defined than the stem region and has not yet been fully refined (Borer et al., 1995). The H2′ resonance of C13 is found in the spectral region characteristic of a regular A-RNA helix (4.20 ppm). The predicted resonance position, however, is shifted significantly up-field (1.34 ppm). Usually such large negative, conformational shifts result from ring-current and magnetic anisotropy effects induced by neighboring aromatic rings. Examination of the position of the residue in the structure shows indeed that, the sugar moiety of C13 is in close proximity of the base of residue 14. In fact, the H2′ is pointing towards the adenine rings, which accounts for the predicted, large up-field shift. Moving the C13 sugar ring away from the base of A14 should bring the predicted shift in the range of the experimental shift. We note in passing that conformational shifts strongly depend on distance ($\delta_{\mathrm{rc/ma}} \propto 1/r^3$) and relative small structural adjustments may change the values of the shifts dramatically. The situation described above is a typical example, in which the loop structure is less well defined because of a paucity of restraints. The use of chemical shifts in the structure improvement and validation may, especially in these cases, be very helpful.

Conclusion

This study shows that ^{1}H chemical shifts of RNA can be predicted with high reliability. As a result, the physical origin of the structural dependencies of the ^{1}H chemical shifts could be established in helical and non-helical environments. We find that the experimental chemical shifts of H2′ and H3′ protons can be used to determine the χ-angle, i.e., down-field shifts indicate a *high-syn*/*syn*-orientation. In an A-helix, the chemical shifts of H5, H6 and H8 protons are mainly affected by their 5′-neighbor. For H2 proton shifts large contributions are seen from both the intra-strand and cross-strand 3′-neighbor (definition in legend of Table 6). As a result for an A-helix H5 and H6 and H8 resonances can be assigned solely on the basis of their observed chemical shifts. We also investigated a number of examples of predicted and experimental chemical shifts which deviated from the expected values and found that predicted shifts are good indicators of errors in structures derived from NOEs and *J*-couplings. Chemical shifts can therefore be used in structure validation and refinement. The computer program that performs the chemical shift predictions (NUCHEMICS) is available on request from the authors.

Acknowledgements

This work was supported by grants from the Swedish National Research Council, the Bioteknik Medel Umeå University and Stiftelsen för Strategisk Forskning/Structural Biology Network (S.W.) and Kempe Foundation (J.C.).

Appendix. Examples of predictions of conformational shifts in a non-helix environment

H2′

To illustrate that the quality of the correlation between observed and predicted shifts is also good in situations where unusual chemical shifts are observed we discuss the shift of the H2′ resonance of residue G30 in the ATP-aptamer (Dieckmann and Feigon, 1997). This resonance exhibits the largest down-field shift of all H2′ protons in the database used in this study. It is found at 5.64 ppm while the normal range covered by the H2′ resonances is 3.9–4.9 ppm (Table 4). The ATP-aptamer structure is composed of an internal loop region flanked by two helical stems. Opposite of the loop a single bulged guanine residue, G30, is found. This guanine forms a base pair with the 3′-terminal residue of the loop, G17. To make this happen, the sugar ring of G30 is south-puckered and the glycosidic torsion angle is in the *syn*-domain (78°). Figure 3A shows that for a nucleotide large down-field shifted H2′ resonances are expected for north-puckered sugars when the χ-angle is between 90°and 150°, and for south-puckered sugars when this angle is between 60°and 120°. In other words, given that G30 is south-puckered and the χ-angle is ca. 78°, one indeed predicts the experimentally observed large down-field shift for the H2′ resonance.

H3′

This concerns the χ-angle of the G20 residue of the UUCG tetraloop in the FMN-aptamer (Fan et al., 1996). Its H3′ resonance is shifted from its reference value (δ_{ref}, Table 4) down-field by $\sim$ 1 ppm to 5.6 ppm. Based on the deposited structure ($\chi = 14°$) we predict a down-field conformational shift of only 0.3 ppm (Figure 3B). The very large conformational shift observed for this guanine residue ($\sim$ 1 ppm) suggests therefore that in the structure the χ-angle has to be adjusted, namely to $\sim 70°$ (cf. Figure 3B). At this value of the χ-angle, the H2′ resonance is expected to shift only moderately ($\sim$ 0.2 ppm, Figure 3A). Indeed, the experimental shift of the H2′ proton corresponds to only a small down-field shift ($\sim$ 0.2 ppm). Hence, both the H2′ and H3′ experimental shifts point to a value of $\sim 70°$ for the glycosidic angle. This value was indeed found in two other studies of UUCG loops (Allain and Varani, 1995; Green and Varcarcel, 1996).

H4′

The G11 H4′ resonance of the ATP aptamer (Dieckmann and Feigon, 1997) is shifted substantially ($\sim$ 3 ppm) up-field to 1.67 ppm. The resonance under consideration was assigned with certainty to the H4′ proton of G11, amongst others, by specific base substitution (Dieckmann and Feigon, 1997). The calculations show also very large ring-currents for this proton leading to a large negative conformational shift (-5.5 ppm). In this case the ring-current contribution is overestimated. However, ring-current shifts strongly depend on distance ($\delta_{rc} \propto 1/r^3$) and relative small structural adjustments may change the values of the shifts dramatically. The up-field shift indicates that the

proton is located directly above or below an aromatic ring. Indeed, G11 is part of an 11-nucleotide bulge, which becomes structured upon binding of AMP. This residue forms a base-pair with G7, the first nucleotide of the bulge. The base pairing of G11 with G7 creates a sharp turn in the backbone and brings the aromatic rings of A12 close to the H4′ of G11, which in turn accounts for the observed large up-field shift.

H5′ and H5″

In conformations clearly deviating from an A-helix considerable shifts may occur for these protons. For example, both the G11 H5′ and H5″ resonances of the ATP aptamer (Dieckmann and Feigon, 1997) at 3.60 ppm and 3.06 ppm, are shifted up-field with respect to their usual position in the A-helix (Table 4). The G11 residue in the ATP aptamer is, as discussed in the preceding section, located in a sharp turn of a large loop, placing the ribose moiety of this residue close to the base of its neighbor, A12, thus explaining the up-field shifts. The calculations confirm the up-field shift for these resonances.

Base protons

The H8 proton of residue A26 of the FMN-RNA aptamer complex (Fan et al., 1996) resonates at 6.75 ppm. This represents an up-field shift of 1.28 ppm compared to its position in an A-helix (8.03 $\pm$ 0.11 ppm). The A26 residue is part of the large internal loop of the molecule. The internal loop segment zippers up to generate a continuous helix with the flanking stem segments. This creates an intercalation site for the aromatic flavin mononucleotide (FMN), which is positioned opposite A26 in the helix. A26 is stacked between the G10-U12-A25 triple and the G9-G27 pair, with the base triple on its 5′-side and the base pair on its 3′-side. The H8 of A26 is especially close to G27 ($\sim$ 4 Å). In an A-helix a H8 proton would be too far away from its 3′-neighbor ($\sim$ 6.3 Å) to experience much of a ring-current (vide infra). But in the present case, the H8 experiences ring-currents from both A25 and G27, thus explaining its up-field shifted resonance. Despite the unusual position of the H8 proton, its chemical shift is predicted correctly, i.e. within 2 standard deviations from the experimental value, although the ring-current contribution is slightly overestimated (the predicted resonance position is 6.45 ppm).

An example of a down-field shifted resonance is that of the H5 proton of C12 in a hairpin loop present in eukaryotic 18S rRNA (Sich et al., 1997). The H5 has an experimental chemical shift of 6.13 ppm (compared to 5.36 $\pm$ 0.17 ppm in an A-helix). The C12 residue is part of a loop 5′-$G_8U_9U_{10}U_{11}C_{12}$-3′ in which it is looped out. In a normal A-helix, the chemical shift of H5 is influenced by its 5′-neighboring residue causing the resonance to shift to lower values (*vide infra*). Because of the looped out position of the C12 residue the resonance position of its H5 proton is not influenced by its 5′-neighbor and should be very close to the reference value. Indeed, the experimental shift of 6.13 ppm is virtually the same as the reference value, δ_{ref}, which is 6.12 ppm (Table 5). The unusual chemical shift is predicted correctly, 6.07 ppm.

The H2 proton of residue A11 of the FMN-RNA aptamer complex (Fan et al., 1996) has an experimental chemical shift of 8.43 ppm (compared to 7.28 $\pm$ 0.30 ppm in an A-type helix). The former resonance position has the exact same value as the reference value, δ_{ref}, of H2 protons (Table 5). Indeed, the A11 residue is looped out from the internal loop and no residues seem close enough to contribute to the chemical shift. The predicted chemical shift represents the experimental shift quite well (8.37 ppm).

References

Allain, F.H.T. and Varani, G. (1995) *J. Mol. Biol.*, **250**, 333–353.

Allain, F.H.T. and Varani, G. (1997) *J. Mol. Biol.*, **267**, 338–351.

Altona, C., Faber, D.H. and Westra Hoekzema, A.J.A. (2000) *Magn. Reson. Chem.*, **38**, 95–107.

Asakura, T., Niizawa, Y. and Williamson, M.P. (1992) *J. Magn. Reson.*, **98**, 646–653.

Asakura, T., Taoka, K., Demura, M. and Williamson, M.P. (1995) *J. Biomol. NMR*, **6**, 227–236.

Bevington, P.R. (1969) *Data Reduction and Error Analysis for the Physical Sciences*, McGraw-Hill, New York, NY.

Borer, P.N., Lin, Y., Wang, S., Roggenbuck, M.W., Gott, J.M., Uhlenbeck, O.C. and Pelczer, I. (1995) *Biochemistry*, **34**, 6488–6503.

Brodsky, A.S. and Williamson, J.R. (1997) *J. Mol. Biol.*, **267**, 624–639.

Butcher, S.E., Dieckmann, T. and Feigon, J. (1997) *J. Mol. Biol.*, **268**, 348–358.

Case, D.A. (1995) *J. Biomol. NMR*, **6**, 341–346.

Case, D.A., Dyson, H.J. and Wright, P.E. (1994) *Methods Enzymol.*, **239**, 392–416.

Celda, B., Biamonti, C., Arnau, M.J., Tejero, R. and Montelione, G.T. (1995) *J. Biomol. NMR*, **5**, 161–172.

Cheong, C. and Moore, P.B. (1992) *Biochemistry*, **31**, 8406–8414.

Colmenarejo, G. and Tinoco, Jr., I. (1999) *J. Mol. Biol.*, **290**, 119–135.

Cromsigt, J.A.M.T.C., van Buuren, B.N.M., Zdunek, J., Schleucher, J., Hilbers, C.W. and Wijmenga, S.S. (1998). In *Magnetic Resonance and Related Phenomena*, Ziessow, D., Lubitz, W.

and Lendzian, F. (Eds.), Technische Universität Berlin, Berlin, pp. 132–133.
Dejaegere, A., Bryce R.A. and Case, D.A. (1999) In *Modeling NMR Chemical Shifts. Gaining Insight into Structure and Environment*, Facelli, J.C. and de Dios, A.C. (Eds.), Americal Chemical Society, Washington, DC, pp. 194–206.
Dieckmann, T. and Feigon, J. (1997) *J. Biomol. NMR*, **9**, 259–272.
Fan, P., Suri, A.K., Fiala, R., Live, D. and Patel, D.J. (1996) *J. Mol. Biol.*, **258**, 480–500.
Giessner-Prettre, C. and Pullman, B. (1987) *Q. Rev. Biophys.*, **20**, 113–172.
Green, M.R. and Valcarcel, J. (1996) *TIBS*, **21**, 296–301.
Haigh, C.W. and Mallion, R.B. (1980) *Prog. NMR. Spect.*, **13**, 303–344.
Hilbers, C.W., Blommers, M.J.J., van de Ven, F.J.M., van Boom, J.H. and van der Marel, G.A. (1991), *Nucleosides Nucleotides*, **10**, 61–80.
Hilbers, C.W., Heus, H.A., van Dongen, M.J.P. and Wijmenga, S.S. (1994) In *Nucleic Acids and Molecular Biology*, Eckstein, F. and Lilley, D.M.J. (Eds.), Springer-Verlag, Berlin, pp. 56–104.
Hoogstraten, C.G., Legault, P. and Pardi, A. (1998) *J. Mol. Biol.*, **284**, 337–350.
Johnson, C.E. and Bovey, F.A. (1958) *J. Chem. Phys.*, **29**, 1012–1014.
Jucker, F.M. and Pardi, A. (1995) *Biochemistry*, **34**, 14416–14427.
Kalurachchi, K. and Nikonowicz, E.P. (1998) *J. Mol. Biol.*, **280**, 639–654.
Kang, H.S. and Tinoco, Jr., I. (1997) *Nucleic Acids Res.*, **25**, 1943–1949.
Kang, H.S., Hines, J.V. and Tinoco, Jr., I. (1996) *J. Mol. Biol.*, **259**, 135–147.
Kolk, M.H., Heus, H.A. and Hilbers, C.W. (1997) *EMBO J.*, **16**, 3685–3692.
Kolk, M.H., van der Graaf, M., Fransen, C.T.M., Wijmenga, S.S., Pleij, C.W.A., Heus, H.A. and Hilbers, C.W. (1998a) *EMBO J.*, **17**, 7498–7504.
Kolk, M.H., van der Graaf, M., Wijmenga, S.S., Pleij, C.W.A., Heus, H.A. and Hilbers, C.W. (1998b) *Science*, **280**, 434–438.
Kolk, M.H., Wijmenga, S.S., Heus, H.A. and Hilbers, C.W. (1998c) *J. Biomol. NMR*, **12**, 423–433.
Kuszewski, J., Gronenborn, A.M. and Clore, M. (1995a) *J. Magn. Reson. B.*, **106**, 92–96.
Kuszewski, J., Gronenborn, A.M. and Clore, M. (1995b) *J. Magn. Reson. B.*, **107**, 293–297.
Legault, P., Hoogstraten, C.G., Metlitzky, E. and Pardi, A. (1998) *J. Mol. Biol.*, **284**, 325–335.
McDowell, J.A., He, L.Y., Chen, X.Y. and Turner, D.H. (1997) *Biochemistry*, **36**, 8030–8038.
Nooren, I.M.A., Wang, K.Y., Borer, P.N. and Pelczer, I. (1998) *J. Biomol. NMR*, **11**, 319–328.
Ösapay, K. and Case, D.A. (1991) *J. Am. Chem. Soc.*, **113**, 9436–9444.
Ösapay, K. and Case, D.A. (1994) *J. Biomol. NMR*, **4**, 215–230.
Ösapay, K., Theriault, Y., Wright, P.E. and Case, D.A. (1994) *J. Mol. Biol.*, **244**, 183–197.
Peterson, R.D. and Feigon, J. (1996) *J. Mol. Biol.*, **264**, 863–877.
Popenda, M., Biala, E., Milecki, J. and Adamiak, R.W. (1997) *Nucleic Acids Res.*, **25**, 4589–4598.
Press, W.H., Flannery, B.P., Teukolsky, S.A. and Vetterling, W.T. (1989) *Numerical Recipes in Pascal: The Art of Scientific Computing*, Cambridge University Press, Cambridge.
Ribas-Prado, R. and Giessner-Prettre, C. (1981) *J. Mol. Struct.*, **76**, 81–92.
Santalucia, J. and Turner, D.H. (1993) *Biochemistry*, **32**, 12612–12623.
Shen, L.X. and Tinoco, Jr., I. (1995) *J. Mol. Biol.*, **247**, 963–978.
Sich, C., Ohlenschläger, O., Ramachandran, R., Görlach, M. and Brown, L.R. (1997) *Biochemistry*, **36**, 13989–14002.
Smit, J.S. and Nikonowicz, E.P. (1998) *Biochemistry*, **37**, 13486–13498.
Szewczak, A.A. and Moore, P.B. (1995) *J. Mol. Biol.*, **247**, 81–98.
van Buuren, B.N.M., Overmars, F.J.J., Ippel, J.H., Altona, C. and Wijmenga, S.S. (2000) *J. Mol. Biol.*,**304**, 371–383.
van de Ven, F.J.M. and Hilbers, C.W. (1988a) *Eur. J. Biochem.*, **178**, 1–38.
van de Ven, F.J.M. and Hilbers, C.W. (1988b) *Nucleic Acids Res.*, **16**, 5713–5726.
Varani, G., Cheong, C. and Tinoco, Jr., I. (1991) *Biochemistry*, **30**, 3280–3289.
Westhof, E. and Fritsch, V. (2000) *Struct. Folding Design*, **8**, R55-R65.
White, S.W., Nilges, M., Huang, A., Brunger, A.T. and Moore, P.B. (1992) *Biochemistry*, **31**, 1610–1621.
Wijmenga, S.S. and van Buuren, B.N.M. (1998) *Prog. NMR. Spec.*, **32**, 287–387.
Wijmenga, S.S., Mooren, M.M. and Hilbers, C.W. (1993) In *NMR of Macromolecules, A Practical Approach*, Roberts, G.C.K. (Ed.), Oxford University Press, New York, NY, pp. 217–288.
Wijmenga, S.S., Kruithof, M. and Hilbers, C.W. (1997) *J. Biomol. NMR*, **10**, 337–350.
Williamson, M.P. and Asakura, T. (1993) *J. Magn. Reson. B.*, **101**, 63–71.
Williamson, J.R. and Boxer, S.G. (1989) *Biochemistry*, **28**, 2819–2831.
Williamson, M.P., Kikuchi, J. and Asakura, T. (1995) *J. Mol. Biol.*, **247**, 541–546.
Wishart, D.S. and Sykes, B.D. (1994) *Methods Enzymol.*, **239**, 363–392.
Wu, M. and Turner, D.H. (1996) *Biochemistry*, **35**, 9677–9689.
Wu, M., Santalucia, J. and Turner, D.H. (1997) *Biochemistry*, **36**, 4449–4460.