A new Program for the Conformational Analysis by NMR of the Sugar Ring of Nucleosides and Nucleotides in Solution: HETROT. Application to the Sugar Ring of AZT in Solution.

Daniel O. Cicero, Gaetano Barbato and Renzo Bazzo*

Istituto di Ricerche di Biologia Molecolare (IRBM) P. Angeletti,
via Pontina km 30,600 - 00040 - Pomezia, Rome, Italy.

Abstract: A new program for the conformational analysis by NMR of the sugar ring of nucleosides and nucleotides in solution, based on both ^{1}H-^{1}H and long range ^{1}H-^{13}C coupling constants is presented. The experimental conditions for the accurate measurement of long range heteronuclear coupling constants in natural abundance are discussed, together with the implementation of a computational algorithm (HETROT) that calculates the range of values of the five parameters describing the conformation of the furanose ring (P_N, τ_{mN}, P_S, τ_{mS} and x_N) that reproduces all the measured couplings within their estimated errors. HETROT was applied to the conformational analysis of the sugar ring of 3'-azido-3'-deoxythymidine (AZT) in DMSO and D_2O. In both solvents the sugar ring exists in fast equilibrium between roughly equally populated N- and S- type conformers. The degree of definition of the precise geometry of such conformers, and their relative population, that can be obtained from the experimental data is analysed and graphically presented.

After 3'-azido-3'-deoxythimidine (AZT) emerged as the most widely used drug[1] for inhibiting human immunodeficiency virus (HIV) replication, there has been an intense activity to assess its structure and conformation.

A number of groups have investigated AZT in the solid state, showing the presence of two crystallographically independent molecules[2] both in the South conformation (P_S=174°, τ_{mS}=33° for molecule A, and P_S=213°, τ_{mS}=37° for molecule B). Molecular mechanics calculations *in vacuum* were also performed by Herzyk *et al.*[3]. The authors found a global minimum of energy in the North (P_N=25.4°), *anti* (χ=-155°) and *g*+ (γ=54.4°) domain, along with other five preferred local energy minima. The pseudorotational barrier between the N- and S- conformations was estimated to be low (about 3 kcal/mole), thereby suggesting that both sugar puckerings could exist in solution.

This hypothesis was subsequently confirmed by NMR studies[4-7]. The values of ^{1}H-^{1}H coupling constants indicated[4,5] that in DMSO the sugar ring exists as a mixture of N- and S-conformation (P_N=-6°, τ_{mN}=τ_{mS}=33.0, P_S=154°), with a slight preponderance of the latter (x_N=0.43). The first studies of the conformation of AZT in D_2O showed that the sugar was engaged in an equilibrium between N- and S-conformers[6,7], although not all individual J-couplings could be measured owing to the strong coupling between H-2' and H-2". With the synthesis of 2'- and 3'-deuterated analogues of AZT, the complete set of proton J-couplings of the sugar ring in water was obtained[7]. By using the PSEUROT program[8], a more detailed description of the two conformers present in the equilibrium was attempted[7] (P_N=-12°, τ_{mN}=τ_{mS}=33.2, P_S=161°, x_N=0.48)

The PSEUROT program[8] has been one of the more useful tools for the conformational analysis of the sugar ring of nucleosides and nucleotides. It requires the measurement of ^{1}H-^{1}H coupling constants at different temperatures, in order to overcome the problem of insufficient experimental data to fit the five variables that describe the two-state equilibrium of the furanose ring. However, when the values of the coupling constants do not change with temperature, as in the case of AZT[4-7], the number of variables and experimental data for a deoxynucleoside is equal, and the final result is very sensitive to experimental errors and/or theoretical approximations. In the case of a ribose sugar ring the situation is even worse, since the number of observable coupling constants is less than the number of parameters to be fitted.

In order to increase the number of observed parameters, one can make use of $^3J_{H-C}$ coupling constants as they are related to the dihedral angles describing the conformation of the sugar ring. A number of recent works[9,10] demonstrated the possibility of accurate measurement of $^3J_{H-C}$ at natural abundance, thereby providing a practical method for the acquisition of this valuable information. In the present work, both the two-dimensional heteronucleus-coupled ω_1 hetero-half-filtered proton-proton correlation (HETLOC)[9], and the paired satellite selection technique (PASS-TOCSY)[10] in its new modified version (SEL-PASS-TOCSY)[11] were used.

The experimental data were then analysed by means of a new algorithm, HETROT, that combines the information of both ^{1}H-^{1}H and ^{1}H-^{13}C coupling constants to calculate all the possible values for the five variables describing the system (P_N, τ_{mN}, P_S, τ_{mS} and x_N) which can reproduce all the experimental data within their estimated error. The final output is an interval for each variable, which all together represent the feasible space[12] of the variables consistent with the experimental information. This result provides also a direct evaluation of the level of accuracy that can be reached in the assessment of the conformational parameters. This approach was applied to the study of the sugar ring of AZT in DMSO and D_2O solution at 37°C.

METHODS

NMR studies.

AZT was purchased from Aldrich Chem. Co. All spectra were obtained on a Bruker AMX 500 MHz at 37°C. The samples were prepared by dissolving 25 mg of AZT in 0.5 ml of DMSO-*d6*, and 15 mg of AZT in 0.5 ml of D_2O.

2D-HETLOC experiments. The HETLOC experiment[9] (Figure 1) was performed on AZT by acquiring a 256x4096 data point matrix for the sample in DMSO, and a 1024x4096 data point matrix for the sample in D_2O. The BIRD delay[13] and the relaxation delay were set to 400 ms and 500 ms respectively. The ^{1}H pulse for the MLEV-17 mixing period was attenuated by 12 dB (90° pulse duration was 26.3 µs). The duration of trim pulses was set to 2 ms and the total mixing time was 65 ms. Data were processed as a 1024x4096 matrix with Gaussian resolution enhancement.

SEL-PASS-TOCSY experiments. SEL-PASS-TOCSY experiments[11] (Figures 2 and 3) were performed on the D_2O sample of AZT to measure the coupling constants involving C-3' and C-4'. Two 1D subspectra were performed with the ^{1}H and ^{13}C pulse carrier positioned at H-3' and C-3' frequencies respectively, and two at H-4' and C-4' frequencies. The selection of each ^{13}C satellite in the corresponding subspectrum was achieved by changing the phase ϕ_9 (45° or 135°). The BIRD delay and the relaxation delay were set to 1.1 s and 6 s respectively. The ^{1}H selective pulse was delivered with an attenuation of 57 dB (90° pulse duration of 3.7 ms),

and the ^{13}C selective pulse was delivered with an attenuation of 69 dB (90° pulse duration of 30.4 ms). The ^1H pulse for MLEV-17 sequence was attenuated by 12 dB (90° pulse duration was 26.3 μs). The duration of trim pulses was set to 2.5 ms and the total mixing time was 73 ms. Data were processed with exponential line broadening.

HETROT computational program.

Calculation of $^3J_{HH}$ and $^3J_{HC}$ coupling constants. A geometrical routine earlier described[14] was applied and translated into a FORTRAN program, which calculates the coordinates of all the atoms as a function of bond angles, bond lengths and the relevant dihedral angles. In a first step, the coordinates of O-5', C-5' and C-4' are defined in a suitable coordinate system in which C-4' is at the origin, C-5' lies in the y-axis and O-5' in the x,y plane. In this system, it is easy to obtain the coordinates of these three atoms as a function of bond lengths and bond angles:

$$O5' = \left(-C5' O5' \sin\left(O5' \hat{C5'} C4'\right), C5' O5' \cos\left(O5' \hat{C5'} C4'\right) - C5' C4', 0\right)$$

$$C5' = (0, -C5' C4', 0) \tag{1}$$

$$C4' = (0,0,0)$$

where C5'O5' and C5'C4' are the corresponding bond lengths and O5' $\hat{C5'}$ C4' the corresponding bond angle.

The coordinates of these three atoms in this system can be converted into the corresponding coordinates in the system in which C-3' is the origin, C-4' lies in the x axis and C-5' lies in the x,y plane by applying the following transformation:

$$X = T(-C3'C4') R(\delta)R(\beta)R(\alpha)X'' \tag{2}$$

where X is the coordinate vector of atom X in the C-3' system, X'' the coordinate vector of the same atom in the C-4' system and T(-C3'C4'), R(δ), R(β), R(α) the transformation matrices:

$$T(-C3'C4') = \begin{bmatrix} 0 \\ -C3'C4' \\ 0 \end{bmatrix} \qquad R(\delta) = \begin{bmatrix} \cos\delta & 0 & -\sin\delta \\ 0 & 1 & 0 \\ \sin\delta & 0 & \cos\delta \end{bmatrix}$$

$$R(\beta) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\beta & \sin\beta \\ 0 & -\sin\beta & \cos\beta \end{bmatrix} \qquad R(\alpha) = \begin{bmatrix} \cos\alpha & \sin\alpha & 0 \\ -\sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{3}$$

The angles α, β and δ are the corresponding Euler angles, and are related to the bond angle C3' $\hat{C4'}$ C5' and the dihedral angle O5'-C5'-C4'-C3' (γ) by the following relations:

$$\sin\alpha = \sin(C3'\hat{C4'}C5')\cos\gamma/(1-\sin^2(C3'\hat{C4'}C5')\sin^2\gamma)^{1/2} \quad 0 \le \alpha \le 2\pi$$

$$\cos\alpha = -\cos(C3'\hat{C4'}C5')/(1-\sin^2(C3'\hat{C4'}C5')\sin^2\gamma)^{1/2}$$

$$\sin\beta = \sin(C3'\hat{C4'}C5')\sin\gamma \qquad\qquad\qquad -\pi/2 \le \beta \le \pi/2$$

$$\cos\beta = (1-\sin^2(C3'\hat{C4'}C5')\sin^2\gamma)^{1/2}$$

$$\sin\delta = -\tan\beta/\tan\gamma \qquad\qquad\qquad\qquad 0 \le \delta \le 2\pi$$

$$\cos\delta = \sin\alpha/\sin(C3'\hat{C4'}C5')$$

(4)

By repeating this conversion process, and incorporating the remaining heavy atoms, it is possible to obtain the coordinates in a common system. For this purpose, it is necessary to have a good estimation of the values of bond lengths, bond angles and torsion angles of the molecule. In the present work we used the empirical correlations between the conformational parameters (P, τ_m, γ and χ) and the internal coordinates based on linear regressions obtained by Altona *et al.*[15] from solid state data. The bond length C3'N4 was taken from the crystal structure of AZT[2], and the bond angle C4'$\hat{C3'}$N4 was corrected with respect to the C4'$\hat{C3'}$O3' reported in the original work of Altona[15], although a similar dependence with P was assumed.

Once the coordinates of the heavy atoms were known, the hydrogen atoms of the molecule were placed by using two empirically derived criteria of symmetry[15]: (i) methylene hydrogen atoms show a tendency to retain local C_{2v} symmetry with a mean H-C-H bond angle of 107.6°; (ii) methine hydrogens on tertiary sp^3 carbon atoms are generally found in positions making equal bond angles to the other three carbon atom substituents.

These regularities allow the calculation of the hydrogen atom coordinates as a function of the heavy atom coordinates. The ^{1}H-^{1}H and ^{1}H-^{13}C dihedral angles were calculated by using standard geometrical relations. The empirical equation of Altona *et al.*[16] and the Karplus relations derived by Cyr and Perlin[17] were used to calculate $^3J_{HH}$ and $^3J_{HC}$ respectively.

Calculation of the feasible space of the variables and degree of matching with the data. These two concepts have been recently illustrated and applied to the study of a 13-residue peptide[12]. From any arbitrary set of values of the five variables describing the two-state equilibrium of the furanose ring (P_N, τ_{mN}, P_S, τ_{mS} and x_N), the averaged expected values of the $^3J_{HH}$ and $^3J_{HC}$ coupling constants can be calculated and compared with the corresponding experimental values. In order to perform this comparison, we assigned to each coupling constant J_i its corresponding error (ΔJ_i) by considering both the experimental source of inaccuracy and the uncertainties inherent in the Karplus equation used to calculate the coupling constants from the dihedral angles. To perform this evaluation the root mean square deviation in a large data set[16] was considered. The estimated total errors turned out to be 0.77 Hz for ^{1}H-^{1}H coupling constants[18], and 1.5 Hz for ^{1}H-^{13}C coupling constants. In this respect, one should appreciate how the relatively large inaccuracy of each single value is counterbalanced by the simultaneous use of several data (indeed as many as possible) to determine a limited set of conformational parameters. However, an unbiased assessment of the reliability of the final conclusions is in our opinion quite desirable, and this can only be obtained by looking at all sources of inaccuracy in a systematic way.

A given set of values of the five variables is considered a *feasible solution* (representing a "possible" system) if condition (5) is fulfilled for all coupling constants:

$$J_i^{exp} - \Delta J_i \leq J_i^{calc} \leq J_i^{exp} + \Delta J_i \tag{5}$$

The ensemble of all feasible sets constitutes what we call the *feasible space* of the variables. Conditions (5) were used as a series of constraints for the variables P_N, τ_{mN}, P_S, τ_{mS} and x_N to calculate analytically the feasible interval for each variable. An additional linear constraint was imposed to τ_{mN} and τ_{mS}, which does not allow these two variables to differ by more than 2.0° (condition (6)):

$$|\tau_{mN} - \tau_{mS}| \leq 2.0° \tag{6}$$

By minimising or maximising the variable under study within the constraints (5)-(6) it is possible to calculate the upper and lower limits of the feasible domain for each conformational parameter. The E04UCF NAG FORTRAN Library Routine was used to perform this calculation, as already described[12].

The feasible space represents the feasible domains of the structural parameters. It includes all possible representations of the system (in terms of types of N and S conformations and their relative populations) that are compatible with the data, within a chosen degree of confidence.

The degree of *matching* with the data (M)[12] can be used as a characterisation of the feasible space:

$$M = \prod_i e^{-\left\{1/2\left[\left(J_i^{exp} - J_i^{calc}\right)/\Delta J_i\right]^2\right\}} \tag{7}$$

The matching M is calculated by an algorithm similar to the one described for the calculation of the feasible space. The variables are constrained within the feasible space by applying the constraints (5)-(6). The matching curves reported in the present work are calculated by maximising M while different values for the variable under study are sequentially assumed and all the other four variables are varied within their feasible domains. An arbitrary matching of 100% was assigned to the solution that exhibits the highest value of M (the "best fit" solution). The matching function M can, with some caution, be taken as an empirical probability function.

RESULTS

^{1}H-^{1}H coupling constants.

Table 1 summarises the ^{1}H-^{1}H coupling constants for AZT both in DMSO and D_2O. The latter were obtained by Gurjar *et al.* from the 2'- and 3'-deuterated analogues[7].

Measurement of ^{1}H-^{13}C coupling constants.

The 2D-HETLOC experiment[9] was used for measuring ^{1}H-^{13}C coupling constants of AZT in DMSO (Figure 1). This kind of experiment provides a very efficient way to obtain long range heteronuclear coupling constants in natural abundance. The long range coupling constants are measured from signal displacement in two F_1 slices through each signal of the cross-peak doublet split by $^1J_{CH}$ in the F_1 dimension (Figure 1b and 1c). The results are listed in Table 2.

Table 1. Experimental ^{1}H-^{1}H Coupling Constants of AZT in DMSO and D_2O at 37°.

Coupling	J^{exp} (Hz)	
	DMSO	D_2O
1'2'	6.7	5.9[a]
1'2"	6.6	6.6[a]
2'3'	7.1	7.6[a]
2'3"	5.0	5.2[a]
3'4'	4.9	5.5

a. From reference 7.
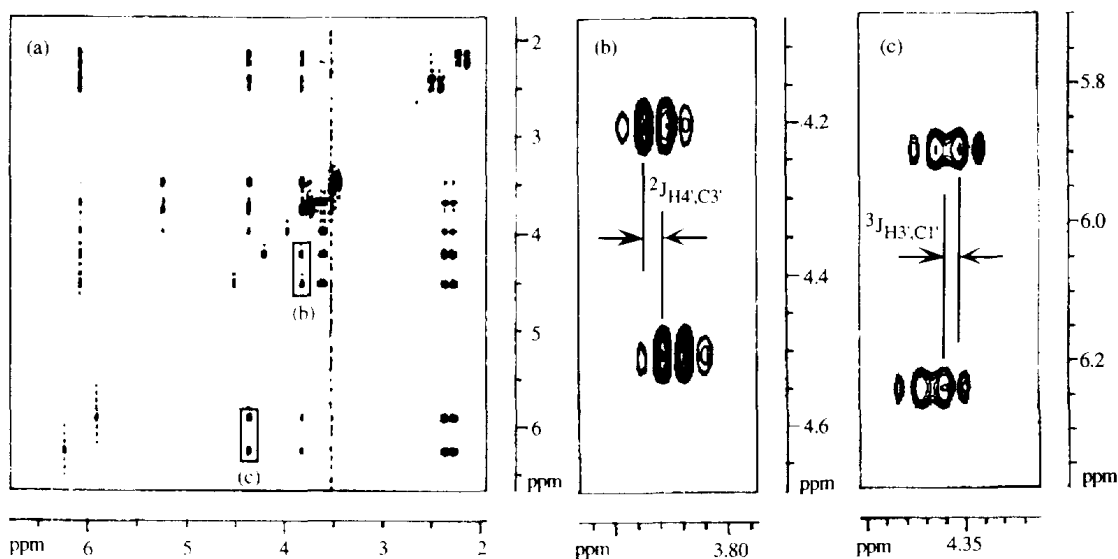
Figure 1. 2D-HETLOC experiment on AZT in DMSO at 37°C. a) selected region showing the spin system of the sugar ring, b) selected region showing cross-peaks between H-4' and H-3', c) selected region showing cross-peaks between H-3' and H-1'.

Table 2. Experimental ^{1}H-^{13}C Coupling Constants of AZT in DMSO at 37°C.

	J^{exp} (Hz)						
	H-1'	H-2'	H-2"	H-3'	H-4'	H-5'	H-5"
C-1'	169.0	-4.9	-0.6	3.7	2.5	<1.0	<1.0
C-2'	<1.0	135.0	135.0	-1.2	1.2	<1.0	<1.0
C-3'	1.2	-3.1	-5.3	150.0	-3.7	5.6	2.5
C-4'	1.2	3.8	3.1	-0.9	149.0		
C-5'		<1.0	<1.0	3.7		139.0	139.0

The same experiment was tried for AZT in D_2O. Most of the heteronuclear coupling constants could be measured from this experiment, but it failed to provide coupling constants involving C-3' and C-4' due to overlap between the upfield ^{13}C satellite of H-3' and the downfield ^{13}C satellite of H-4'. To overcome this problem, we have recently proposed a selective version of the 1D PASS-TOCSY experiment[11] (Figure 2). This improved version was necessary, as the frequencies of C-3' and C-5' differ only by 141.9 Hz, and those of C-4' and C-1' by 112.4 Hz. In these conditions, the original sequence of PASS-TOCSY[10] was not able to select the ^{13}C satellites of H-3' without touching the corresponding ^{13}C satellites of H-5', and the same applies to those of H-4' and H-1'. Figure 3 shows the two subspectra of ^{13}C satellites of H-3', with the TOCSY transferred signals to the other protons. The expanded regions show the signal of H-1' in both spectra. The signal displacement corresponds to $^3J_{H1',C3'}$. The whole data set for heteronuclear couplings of AZT in D_2O is listed in Table 3.
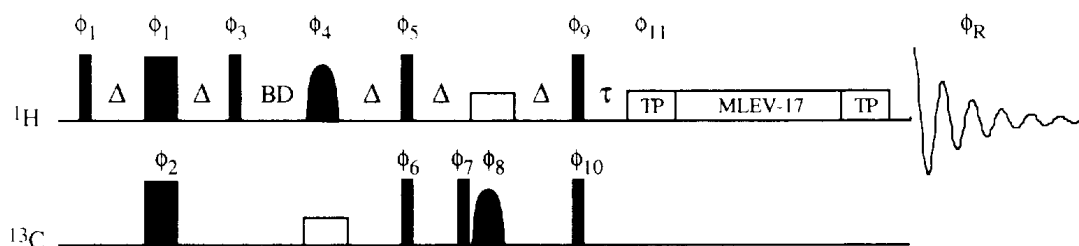

Figure 2. Pulse sequence for SEL-PASS-TOCSY[11]. Narrow and wide bars represent hard 90° and 180° pulses, respectively. Gaussian shaped figures represent selective 90° pulses. Open rectangles represent decoupling, and TP stands for trim pulse. The delay $\Delta = 1/(2 \ ^1J_{CH})$ and $\tau = 1/(4 \ ^1J_{CH})$. BD represents BIRD delay. Phase cycling is as follows: $\phi_1 = x$; $\phi_2 = x$; $\phi_3 = -y$; $\phi_4 = x$; $\phi_5 = y$; $\phi_6 = x, x, -x, -x$; $\phi_7 = y$; $\phi_8 = x, -x$; $\phi_9 = 45°$ or $135°$; $\phi_{10} = y$; $\phi_{11} = x$; $\phi_R = x, -x, -x, x$.

Table 3. Experimental ^1H-^{13}C Coupling Constants of AZT in D_2O at 37°C.

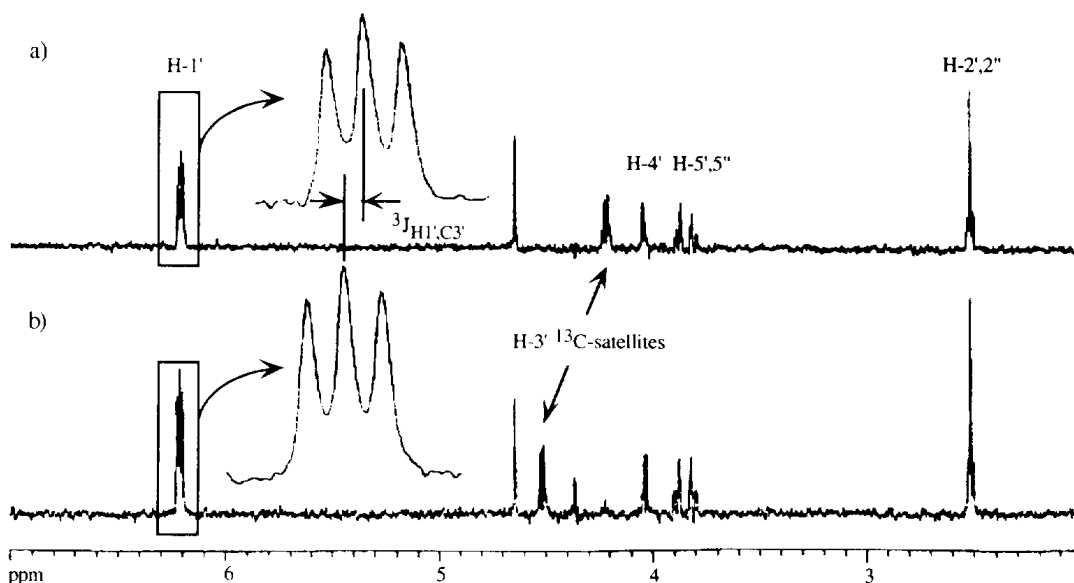|	J^{exp} (Hz)					
	H-1'	H-2',2"	H-3'	H-4'	H-5'	H-5"
C-1'	171.3	-3.7[a]	3.7	1.2		
C-2'	<1.0	136.4[a]	-1.8	1.1	<1.0	<1.0
C-3'	1.3	-4.4[a]	150.6	-2.5	4.9	2.5
C-4'	0.7	3.3[a]	-0.4	150.4	-1.1	-0.7
C-5'	<1.0	<1.0[a]	4.3		143.3	143.8

a. $0.5(J_{H2',C} + J_{H2'',C})$.

Figure 3. SEL-PASS-TOCSY spectra of AZT in D_2O (sequence in Figure 2). Proton and carbon carrier frequencies were set to those of H-3' and C-3' respectively. Phase ϕ_9 was set to 45° to select the downfield ^{13}C satellite of H-3' (a), or to 135° to select the corresponding upfield satellite (b). The expanded regions show the H-1' signal. $^3J_{H1',C3'}$ could be measured from the relative displacement of these signals.

Conformational analysis

As a test for the reliability of the structures of AZT calculated by HETROT, the coordinates of the heavy atoms of the sugar ring and its substituents for the two molecules of the crystal structure were calculated, using as input the reported values[2] of P_S and τ_{mS} ($P_S=174°$, $\tau_{mS}=33°$ for molecule A, and $P_S=213°$, $\tau_{mS}=37°$ for molecule B). Figure 4 shows the superposition of the crystal structure[2] and the same molecules calculated by the HETROT program. The rms deviation in the coordinates was of 0.020Å and 0.024Å for molecules A and B respectively.
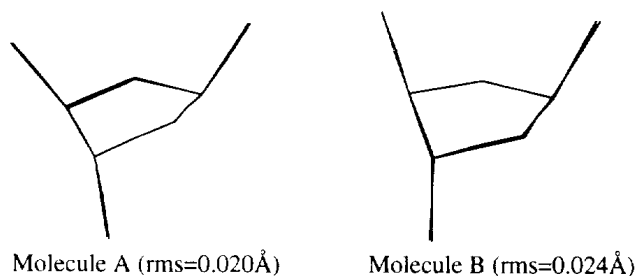

Molecule A (rms=0.020Å)   Molecule B (rms=0.024Å)

Figure 4. Superposition of crystal structure of molecules A and B and the corresponding calculated molecule with the same values of P and τ_m by the program HETROT.

HETROT results for AZT in DMSO. Two independent analyses using two sets of coupling constants were performed (Table 4). In the first one, the feasible space for each of the five conformational parameters (P_N, τ_{mN}, P_S, τ_{mS} and x_N) was calculated, using only ^1H-^1H spin couplings of Table 1. The second analysis involved the same calculation, but including all the available coupling constants.

In a second step, the matching curves[12] for each variable were calculated. The results are plotted in Figure 5.

HETROT results for AZT in D₂O. Four different sets of coupling constants were used to perform the calculation of the feasible space of the variables describing the sugar ring of AZT in D₂O. The first one included the ^1H-^1H coupling constant information originally measured for AZT in D_2O^6, that is ($J_{H1',H2'}+J_{H1',H2''}$), ($J_{H2',H3'}+J_{H2'',H3'}$), $J_{H3',H4'}$. The second set includes the former plus the heteronuclear coupling constants of Table 3. The third set comprises all the homonuclear coupling constants determined by Gurjar *et al.*[7], and the fourth one the whole set of ^1H-^1H and ^1H-^{13}C coupling constants of Tables 1 and 3. Each set of coupling constants was used by HETROT to calculate the feasible space of each variable, and the results are shown in Table 5. The application of our methodology in this case allows to compare different sets of experimental parameters as to their relative informative content. As shown in Table 5, a somewhat different conformational scenario (more or less precise) can ensue from the application of different measurements.

Table 4. Feasible space for the conformational variables describing AZT in DMSO at 37° C.

Coupling Set	P_N (°)	τ_{mN} (°)	P_S (°)	τ_{mS} (°)	x_N
Homonuclear	-44.1 - 55.9	25.9 - 48.1	121.8 - 190.4	26.3 - 47.9	0.33 - 0.55
All couplings	4.8 - 55.7	32.5 - 48.0	142.1 - 166.4	32.5 - 47.8	0.40 - 0.50

Table 5. Feasible space for the conformational variables describing AZT in D₂O at 37° C.

Coupling Set	P_N (°)	τ_{mN} (°)	P_S (°)	τ_{mS} (°)	x_N
Reduced Homo[a]	-90.0 - 83.1	17.2 - 50.0	90.0 - 270.0	17.3 - 50.0	0.15 - 0.66
Red. Homo + Het[b]	11.8 - 52.5	20.6 - 50.0	120.4 - 223.5	20.2 - 50.0	0.28 - 0.67
Homonuclear[c]	-42.4 - 51.6	22.5 - 43.3	110.1 - 199.3	22.5 - 43.2	0.36 - 0.64
All couplings[d]	14.3 - 51.7	25.7 - 42.7	126.6 - 169.1	25.6 - 42.8	0.42 - 0.63

a. The sums of homonuclear couplings between H-2',2" and H-1' or H-3' were used. b. The same as *a* plus the heteronuclear set. c. All homonuclear couplings of Table 1 were used. d. The same as *c* plus the heteronuclear set.
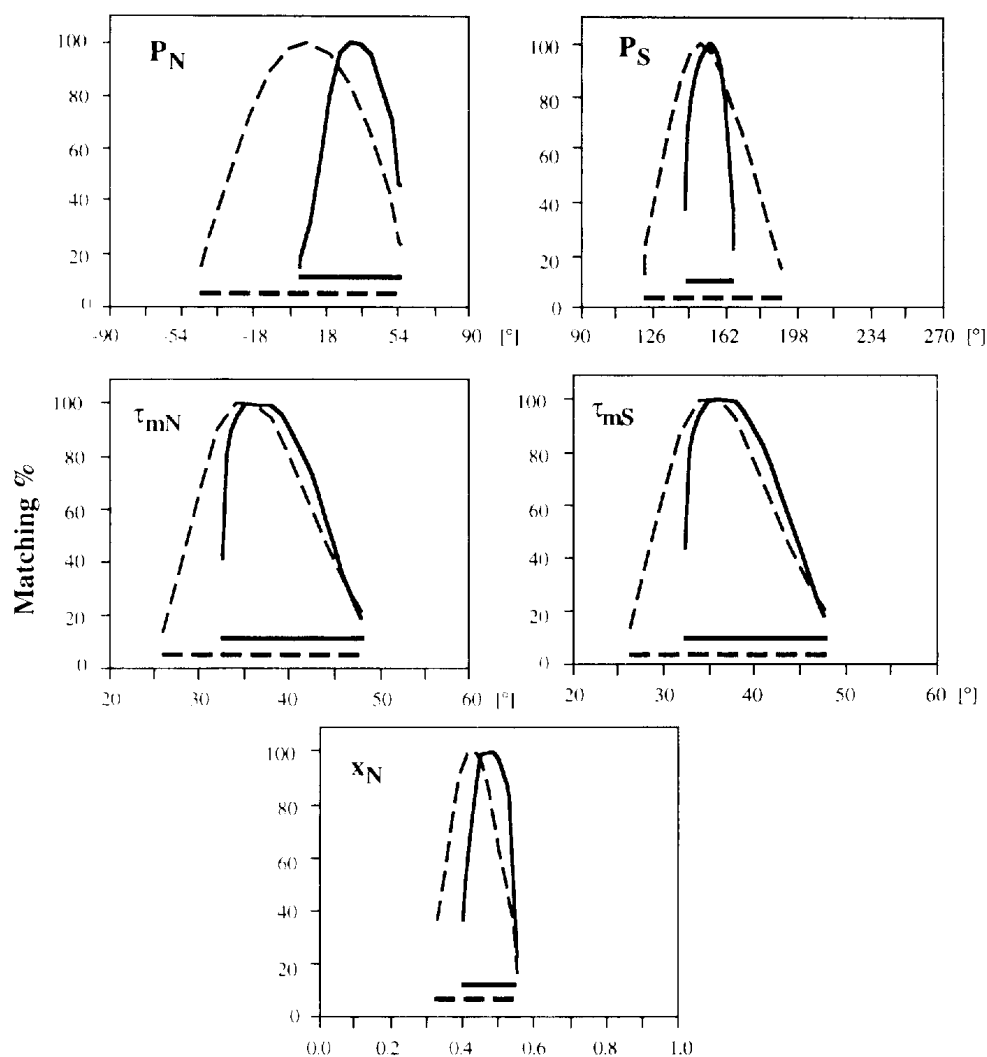
Figure 5. Matching curves for the five variables describing the two-state equilibrium of the ring sugar of AZT in DMSO at 37°C. Dashed lines and continuous line correspond to matching curves using only ^{1}H-^{1}H spin and the complete set of coupling constants, respectively. Horizontal lines at the bottom represent the feasible space of each variable (see Table 4).

DISCUSSION

According to the pseudorotational model developed by Altona *et al.*[19], the conformation of a furanose ring can be described by two parameters: the phase angle of pseudorotation P, and the maximum puckering amplitude τ_m. In the case of nucleosides and nucleotides, a generally accepted model involves a two-state equilibrium between a N- and S-conformer. The complete description of the conformation of the sugar ring can be achieved, within the limits of this approximation, by five parameters: P_N, τ_{mN}, P_S, τ_{mS} and a molar fraction, x_N. In the early '80, Altona proposed a generalised Karplus equation[16] to correlate the coupling constant and the ^{1}H-^{1}H dihedral angle, and an empirical correlation between the ^{1}H-^{1}H dihedral angles and the pseudorotational parameters developed also by his group[15]. These two pieces of information allowed the

implementation of a computational program, PSEUROT[8], which calculates the "best values" of the five parameters that can reproduce a set of ^{1}H-^{1}H coupling constants. In order to increase the ratio observables/unknowns, the authors suggested to measure the coupling constants at different temperatures.

Our analysis allows a critical revision of the PSEUROT method and generates a more general methodology which is practical and also exerts an adequate control on the significance and reliability of the final conclusions.

The new methodology is incorporated in an algorithm and translated in the so called HETROT program, which performs a series of operations. First, the parametrizations of the internal coordinates of the molecule with respect to the conformational parameters P and τ_m are used[15]. By using a geometrical routine[14], it is possible to derive the coordinates of all the heavy atoms. The protons are then placed using empirically derived criteria of symmetry[15]. As a consequence, it is no longer necessary to use the relations between the proton dihedral angles and P and τ_m, which constitute a second step of approximation. Secondly, the program calculates the whole interval for each variable that can reproduce the experimental data set within the estimated error. As we already pointed out[12,18], this result allows a direct estimation of the degree of accuracy that can be achieved in the description of each variable. As a contrary, the "best fit" solution, as calculated by PSEUROT, constitutes only one solution within the complete set of all feasible solutions[12]. Thirdly, the inclusion of heteronuclear coupling constants allows an increase in the ratio observables/variables, without a need for the variation of the ^{1}H-^{1}H coupling constants with the temperature. This condition imposes a limitation to the type of system that can be studied by PSEUROT: systems having roughly equal population of N- and S-conformers, or presenting only one of the two with a prominent difference in energy, will hardly show any variation of the coupling constants with temperature.

The possibility of using the information of heteronuclear coupling constants was earlier suggested by Serianni[20]. However, the need for ^{13}C-labelled nucleosides impaired the practical use of this information. In the present work we demonstrate the possibility of measuring long range heteronuclear coupling constants for a nucleoside with a good level of accuracy, in condition of natural abundance. For this purpose, the HETLOC[9] and the SEL-PASS-TOCSY[11] experiments can be profitably used. The main advantage of these experiments resides in the fact that the heteronuclear coupling constant is not measured from a splitting, but from a difference in chemical shift between two F_1 slices or between two subspectra. In this way, it is possible to measure accurately coupling constants that are in the order of 1 Hz.

The HETROT program was applied to the study of the conformation of the sugar ring of AZT in DMSO using both the homonuclear coupling set and the complete set reported in Tables 1 and 2. AZT constitutes a clear example of system with roughly equal population of N- and S-conformers, and therefore the difference in energy between the two forms is too low to allow a significant variation of relative population with temperature. As a consequence, only five ^{1}H-^{1}H coupling constants can be used to derive five variables. Although this situation is far from ideal, a "best fit" solution was attempted by means of PSEUROT[4]. We calculated the whole feasible space for each variable with the HETROT program, using only homonuclear coupling constants (Table 4, first row). Each interval represents the domain of the variable that can reproduce the experimental data set within the estimated error. The "best fit" solution calculated by PSEUROT represents only a point within this interval. As an example, let's consider the case of P_S. The "best fit" solution[4] is 154°, whereas, the complete feasible space ranges from 122° to 191°.

A significant narrowing of the intervals occurs when the heteronuclear couplings listed in Table 2 are included in the calculation (Table 4, second row). Clearly, all the variables are better defined, and in the case of P_S, the total allowed domain drops from 69° to 24° ($142°\leq P_S<166°$). As a matter of fact, the "best" solution

derived from homonuclear coupling constants for PN (-6°) is not allowed when the whole set of couplings is used (5°≤PN≤56°).

A further characterisation of this feasible space was attempted, by calculating the matching curves[12]. The graphics depicted in Figure 5 show the degree of reproduction of the experimental data set when the variable under study is fixed at a certain point of the feasible space. The "best fit" solution is represented by the five maxima of the curves, and a 100% matching was arbitrarily assigned to this solution. These curves indicate that all the points within the feasible interval reproduce the data with a good matching (notice that M is define as a multiplicative function (equation 7), and therefore one can consider a bad matching to be at least one order of magnitude lower than the best one). This result demonstrates that the only "best" solution in terms of the rms deviation between calculated and experimental couplings is not a good characterisation of the complete feasible space.

From a conformational point of view, AZT in DMSO presents a sugar ring in equilibrium between "normal" types of N- and S-conformers, with a slight preponderance of the latter ($0.40 \leq x_N \leq 0.50$). This slight but significant preponderance of the S-conformer can be also corroborated by the values of $^2J_{C1',H2'}$ (-4.6 Hz) and $^2J_{C1',H2''}$ (-0.6 Hz), since it has been previously reported[20b] that the magnitude of $|^2J_{C1',H2'}|-|^2J_{C1',H2''}|$ appears to be a useful parameter to assess N/S distribution in 2'-deoxyribonucleosides. In fact, $|^2J_{C1',H2'}|$ is higher than $|^2J_{C1',H2''}|$ for S-conformers, and similar in N-conformers.

For the study of AZT in D_2O we performed four calculations of the feasible space of the variables. In the first run (Table 5, first row) only the sums of homonuclear couplings between H-2',2" and H-1' or H-3' were used. In this case, three ^1H-^1H coupling constants were used to define five variables. Although it is possible to calculate a "best fit" solution also in this case, the feasible space clearly indicates the complete lack of information.

A significant improvement can be achieved by including the heteronuclear set (Table 5, second row). All the variables, except for the puckering amplitudes, present a much better definition. The latter are mainly defined by coupling constants between *cis*-nuclei[20a] and for 2'-deoxyribonucleosides, this information is mainly encoded in the coupling constants $J_{H2',H3'}$, $J_{H2'',H1'}$ and $J_{H3',C5'}$. Clearly, in the present calculation only $J_{H3',C5'}$ contributes to the definition of the puckering amplitude and this diminished informational content is responsible for a poor definition of τ_m.

The third calculation involves the whole set of homonuclear coupling constants (Table 5, third row), and is used as a control to clarify whether the effort of synthesising the 2'- and 3'-deuterated analogues gave a better information than the one already encoded in the set of heteronuclear coupling constants. By comparing the second and third row we can safely conclude that this is not the case. P_N is already better defined by the inclusion of the heteronuclear coupling constants to the reduced homonuclear set (40° versus 94° of allowed domain), and a similar definition could be achieved for P_S and x_N. A slightly better definition for the puckering amplitudes can be obtained by using the complete set of homonuclear constants, but these parameters are far for being well defined by the experimental data.

Finally, a fourth calculation was performed using the complete set of both homonuclear and heteronuclear coupling constants (Table 5, fourth row). The introduction of the latter, as compared with the third row, clearly defines better all the variables, specially in the case of P_S ($127° \leq P_S \leq 169°$). The measurement of heteronuclear couplings for the deuterated analogues of AZT synthesised by Gurjar *et al*[7] could certainly constitute a further improvement of this result.

From a conformational point of view, the sugar ring of AZT in D_2O is again engaged in an equilibrium between "normal" types of N- and S-conformers. As a difference with the DMSO situation, no preponderance

of one of these two conformers can be established $(0.42 \leq x_N \leq 0.63)$, and maybe this change in the relative population contributes to the slight but experimentally significant differences found between coupling constants in DMSO and D_2O.

CONCLUSIONS

The introduction of heteronuclear coupling constants in the conformational analysis of the sugar ring of nucleosides and nucleotides greatly improves the level of accuracy that can be obtained for the five variables describing this system. By using recently proposed 2D- and 1D-NMR experiments[9-11], the accurate measurement of long range ^{13}C-^{1}H coupling constants in natural abundance is possible. The coupling constant values can be correlated with the furanose conformation using the parametrization of Altona[15], avoiding the application of a second step of approximation[21]. The use of HETROT allows a systematic and efficient description of the whole range of values of the variables, that are in agreement with the experimental information. In this way, a more accurate description of the conformational properties of the sugar ring of AZT both in DMSO and D_2O was possible. The complete description is graphycally given by the so called matching curves that show the degree of matching with the data allowed by each value of the conformational parameters compatible with the data (feasible). In summary, the sugar ring of AZT exists in fast equilibrium between roughly equally populated N- and S - type conformers. The analysis of the degree of definition of the precise geometry of such conformers, and their relative populations, that can be obtained by the data is a novel feature of the methodology.

For those systems that do not change conformation with temperature, HETROT constitutes the only alternative to the accurate study of the conformational properties of the furanose moiety. For other systems in which the coupling constants change with temperature, the inclusion of heteronuclear coupling constants at different temperatures would greatly improve the degree of definition attainable for the variables.

REFERENCES

1.   De Clerq, E. *J. Med. Chem.* **1986**, *29*, 1561-1569.

2.   a) Gurskaia, G. V.; Tsapkina, E. N.; Skaptsova, N. V.; Kraevskii, A. A.; Lindeman, S. V.; Struchkov, I. T. *Dokl. Akad. Nauk. SSSR* **1986**, *291*, 859-862. b) Brinbaum, G. J.; Giziewicz, J.; Gabe, E. J.; Lin, T-S.; Prusoff, W. H. *Can. J. Chem.* **1987**, *65*, 2135-2139. c) Camerman, A.; Mastropaolo, D.; Camerman, N. *Proc. Natl. Acad. Sci. USA* **1987**, *84*, 8239-8242. d) Van Roey, P.; Salerno, J. M.; Duax, W. L.; Chu, C. K.; Ahn, M. K.; Schinazi, R. F. *J. Am. Chem. Soc.* **1988**, *110*, 2277-2282. e) Parthasarathy, R.; Kim, H. *Biochem. Biophys. Res. Commun.* **1988**, *152*, 351-358. f) Van Roey, P.; Salerno, J. M.; Duax, W. L.; Chu, C. K.; Ahn, M. K.; Schinazi, R. F. *Proc. Natl. Acad. Sci. USA* **1989**, *86*, 3929-3933.

3.   Herzyk, P.; Beveridge, A.; Neidle, S. *Biochem. Biophys. Res. Commun.* **1987**, *145*, 1356-1361.

4.   Swapna, G. V. T.; Jagannadh, B.; Gurjar, M. K.; Kunwar, A. C.; *Biochim. Biophys. Res. Commun.* **1989**, *164*, 1986-1092.

5.   Jagannadh, B.; Reddy, D. V.; Kunwar, A. C. *Biochim. Biophys. Res. Commun.* **1991**, *179*, 386-391.

6.   Plavec, J.; Koole, L. H.; Sandström, A.; Chattopadhyaya, J. *Tetrahedron* **1991**, *47*, 7363-7376.

7.   Gurjar, M. K.; Kunwar, A. C.; Reddy, D. V.; Islam, A.; Lalitha, S. V. S.; Jagannadh, B.; Rama Rao A. V. *Tetrahedron* **1993**, *49*, 4373-4382.

8. de Leeuw, F. A. A. M.; Altona, C. *J. Comp. Chem.* **1983**, *4*, 428-437.

9. a) Kurtz, M.; Schmieder, P.; Kessler, H. *Angew. Chem. Int. Ed. Engl* **1991**, *30*, 1329-1331. b) Wollborn, U.; Leibfritz, D. *J. Magn. Reson.* **1992**, *98*, 142-146.

10. Fukushi, E.; Kawabata, J. *J. Magn. Reson. (A)* **1994**, *108*, 103-105.

11. Bazzo, R.; Barbato, G.; Cicero, D. O. *J. Magn. Reson.* in press.

12. Cicero, D. O.; Barbato, G.; Bazzo, R. *J. Am. Chem. Soc.* **1995**, *117*, 1027-1033.

13. Garbow, J. R.; Weitekamp, D. P.; Pines, A. *Chem. Phys. Lett.* **1982**, *93*, 504-511.

14. Wrenn Wooten, E.; Edge, C. J.; Bazzo, R.; Dwek, R. A.; Rademacher, T. W. *Carbohydr. Res.* **1990**, *203*, 13-17.

15. de Leeuw, H. P. M.; Haastnoot, C. A. G.; Altona, C. *Isr. J. Chem.* **1980**, *20*, 108-126.

16. Haasnoot, C. A. G.; de Leeuw, F. A. A. M.; Altona, C. *Tetrahedron* **1980**, *36*, 2783-2792.

17. Hamer, G. K.; Balza, F.; Cyr, N.; Perlin, A. S. *Can. J. Chem.* **1978**, *56*, 3109-3126.

18. Cicero, D. O.; Iribarren, A. M.; Bazzo, R. *Appl. Magn. Reson.* **1994**, *7*, 95-105.

19. Altona, C.; Sundaralingam, M. *J. Am. Chem. Soc.* **1972**, *94*, 8205-8212.

20. a) Serianni, A. S.; Barker, R. *J. Org. Chem.* **1984**, *49*, 3292-3300. b) Bandyopadhyay, T.; Wu, J.; Serianni, S. *J. Org. Chem.* **1993**, *58*, 5513-5517.

21. Haasnoot, C. A. G.; de Leeuw, F. A. A. M.; de Leeuw, H. P. M.; Altona, C. *Org. Magn. Reson.* **1981**, *15*, 43-57.