*Tetrahedron* Vol. 47, No 35, pp 7363-7376, 1991
Printed in Great Britain

0040-4020/91 $3 00+ 00


# Structural Studies of Anti-HIV 3'-α-Fluorothymidine & 3´-α-Azidothymidine By 500 MHz $^1$H-NMR Spectroscopy & Molecular Mechanics (MM2) Calculations

**Janez Plavec, Leo H. Koole, Anders Sandström & Jyoti Chattopadhyaya***

Department of Bioorganic Chemistry, Box 581, Biomedical Center,
University of Uppsala, S-751 23 Uppsala, Sweden

(*Received in UK* 30 *May* 1991)

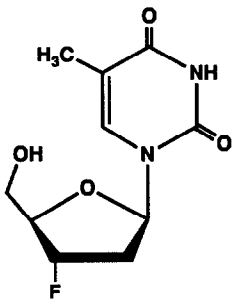
Key Words Anti-HIV, 3'-Fluorothymidine & 3´-Azidothymidine, NMR, Conformation

*Abstract NMR studies of anti-HIV 3'-α-fluorothymidine (FLT) (1) and 3'-α-azidothymidine (AZT) (2) in aqueous solution gave insight into their conformational dynamics [pseudorotation of the furanose ring, rotation around the C4'-C5' bond (γ), and rotation around the glycosidic C1'-N1 bond (χ)] The interpretation of scalar proton-proton coupling data for FLT (1) in solution showed a conformational bias towards a South-type (~90%) puckered furanose conformation Energy minimization by molecular mechanics calculations using the MM2 force field gave molecular structure which is in excellent agreement with the NMR data (P = 151°, $\nu_m$ = 34°) This solution structure is grossly similar to the available averaged crystal structures (P = 171°, $\nu_m$ = 34°) The phase angle P = 151° for FLT (1) in solution indicates that the preferred furanose geometry is an intermediate between the C2'-endo envelope and the C2'-endo/C1'-exo twist conformation while the averaged phase angle in the solid state (P = 171°) indicates an intermediate geometry between the C2'-endo envelope and the C2'-endo/C3'-exo twist structure $^1$H-NMR studies indicated that the North- and South-type pseudorotamer population of the furanose ring in AZT (2) is approximately 1 1 but it failed to provide any definite estimation of P and $\nu_m$ from the analysis of J-coupling constant because of identical chemical shifts of H2' and H2" Molecular mechanics calculations were used to model both the North and South form MM2 calculations in the North region converge virtually to a structure with P = 21 6° (unsymmetrical twist with major C3'-endo and minor C4'-exo twist structure), $\nu_m$ = 40°, χ = -160 8° and γ = 56 9°, while the MM2 calculations in the South region show a predominant puckered geometry represented by P = 160 3° (envelope C2'-endo geometry), $\nu_m$ = 35.5°, χ = -157.8° and γ = 58.3° Note that all six known X-ray structures for AZT belong to the South-type geometry The MM2 calculated South geometry for AZT is grossly similar to one of the two independent molecules found in the asymmetric unit of X-ray crystal structures (P = 175°, $\nu_m$ = 32 3°) Thus discrepancy between solution and solid state structure may indicate the danger of only using crystal structural data for the formulation of structure-activity relationships for a candidate drug*

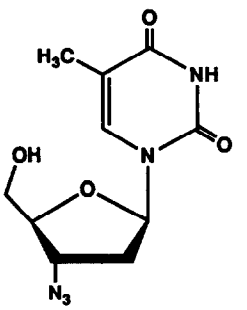
The *in vivo* antiretroviral activity of several 2',3'-dideoxynucleosides has attracted a great deal of interest in their conformational properties To date, the X-ray crystal structures of most active 2',3'-dideoxynucleosides, 3'-α-fluorothymidine and 3'-α-azidothymidine are known[1-8] Interestingly, this has led to the proposal that the inhibitory action may be related to the (preferred) conformation of the modified furanose ring[1] A South-type puckered ring conformation is thought to facilitate anabolic phosphorylation of the 5'-OH group[1] In fact, 2',3'-dideoxynucleosides must be transferred into the 5'-triphosphate species before they become active, either by blocking a site on the reverse transcriptase enzyme, or by terminating the growing DNA chain during reverse transcription[9-13]

In the present work, we have investigated and assessed the conformation of 3'-α-fluorothymidine (FLT) (1) and 3'-α-azidothymidine (AZT) (2) in aqueous solution, using $^1$H-NMR spectroscopy in conjunction with

MM2 molecular mechanics calculations While AZT is currently used in clinical AIDS therapy, the compound FLT is likely to become another drug of clinical importance. Efficacy of FLT as a drug in SIV infected monkeys




1




2

and in HIV-infected human cells has been found to be at least 10 times improved than AZT[12,14] Phase I clinical studies on FLT are now in progress[15] Knowledge of the conformational properties of AZT and FLT in aqueous solution can attribute to our understanding of structure-activity relationships of 2',3'-dideoxynucleosides, and therefore be a guideance in the design of more active substances Evidently, the molecular conformation as observed in the crystal could be influenced by intermolecular crystal packing forces, especially for hydrophilic molecules Thus it has been found that the pentose moiety in adenosine[16] crystallizes in the North form (C3'-*endo*) while the counterpart in adenosine hydrochloride[17] crystallizes in the South form (C2'-*endo*), although the protonation-site in the latter is ~5Å away from the sugar moiety Note that NMR studies of adenosine in solution have clearly shown that the North (N) and South (S) pseudorotamers are present in approximately 2 · 8 ratio, while in acidic solution this North ⇄ South equilibrium dynamically shifts to 4 6 The crystal structure of adenosine represents the structure of the *minor* sugar pseudorotamer, and the crystal of adenosine hydrochloride represents a slightly preponderant sugar pseudorotamer with respect to the population of pseudorotamers that actually exist in neutral or acidic solution, respectively NMR studies, on the other hand, provide insight into the *conformational dynamics* in solution [e g pseudorotation of the furanose ring, rotation around the C4'-C5' bond ($\gamma$), and rotation around the glycosidic C1'-N1 bond ($\chi$)] The information concerning the conformation of the deoxyribose ring in 2',3'-dideoxy-3'-substituted nucleosides is of particular interest since the consequence of pseudorotational population in equilibrium[18] can be studied as a function of solvent, temperature, salt, pH etc by NMR spectroscopy The N-type furanose conformation is characterized by phase angle of pseudorotation $P_N = 0° \pm 90°$, and the S-type by $P_S = 180° \pm 90°$, the amplitude of pucker ($\nu_m$) in both N and S conformations is however around 30 - 40°[18-20] The conformational analysis of the sugar ring in terms of geometry and population of both conformers in solution by $^1$H-NMR spectroscopy depends upon the determination and interpretation of scalar proton-proton coupling constants FLT in solution shows a conformational bias towards a South-type puckered furanose conformation, which deviates considerably from the X-ray crystal structure AZT displays a rapid North ⇄ South conformational equilibrium of the furanose ring in an approximate ratio of 1 1, but it crystallizes in the South form The encountered difference in the solution conformations of FLT and AZT may be rationalized on the basis of two factors (i) the greater electron withdrawing effect of 3'-F versus 3'-$N_3$, which leads to a greater preference for

F-C3'-C4'-O4' *gauche* orientation in **1** in comparison with $N_3$-C3'-C4'-O4' *gauche* orientation in **2** It should be noted that 2'-α-substituted nucleosides display analogous behaviour, i e increase of the electronegativity of the 2'-α-substituent leads to a larger preference for a North-type conformer, in which *gauche* orientation is achieved in the sequence (2'-α-substituent)-C2'-C1'-O4'[28], and (ii) the larger steric impact of 3'-$N_3$ in **2** versus 3'-F in **1** Steric interactions of the 3'-α-$N_3$-substituent in **2** is minimized for its preferential pseudo-equatorial location, which could also explain in part why North is populated to a significantly larger extent in AZT (**2**) than in FLT (**1**) $^1$H-NMR studies and molecular mechanics calculations using the MM2 force field[26] provided models of the North- and South-type furanose rings participating in the conformational equilibrium of AZT These results are also discussed with reference to the South-type conformation of AZT which was encountered in the crystal structure.

*Solution conformation of FLT* High-field $^1$H-NMR spectra (500 MHz) of FLT in $D_2O$ were recorded at different sample temperatures (283K, 298K, 303K, 333K, & 353K), the sample concentration was approximately 20 mM Nearly all J-coupling constants were obtained from straightforward first-order analysis of the subspectra For some non-first order patterns we used a routine algorithm to simulate the spectrum to

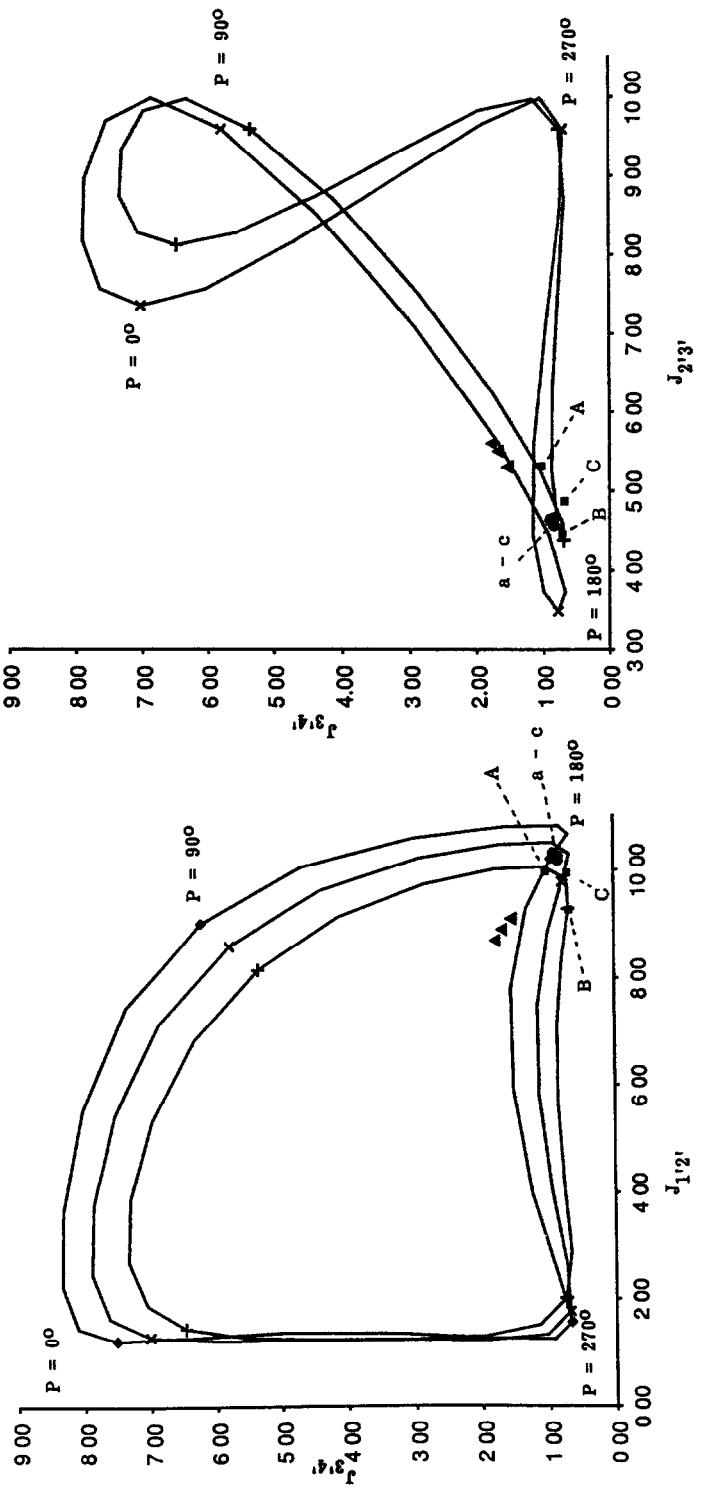
Table 1 J couplings (Hz) of 3'-α-fluorothymidine (FLT) (**1**) recorded at 500 MHz in $D_2O$

| J-couplings | 283K | 298K | 303K | 333K | 353K |
|---|---|---|---|---|---|
| $J_{1'2'}$ | 9 3 | 9 1 | 9 1 | 8 9 | 8 7 |
| $J_{1'2''}$ | 5 6 | 5 7 | 5 7 | 5.8 | 5 9 |
| $J_{2'3'}$ | 4 9 | 5 3 | 5 3 | 5 5 | 5 6 |
| $J_{2''3'}$ | | 1 5 | 1 4 | 1 5 | 1 7 |
| $J_{3'4'}$ | | 1 5 | 1 5 | 1 6 | 1 7 |
| $J_{4'5'}$ | 4 0 | 4 3 | 4 4 | 4 5 | 4 6 |
| $J_{4'5''}$ | 4 0 | 4 3 | 4 4 | 4 5 | 4 6 |
| $J_{2'F}$ | 38 9 | 38 2 | 38 0 | 37 0 | 36 3 |
| $J_{2''F}$ | 21 7 | 22 1 | 22 2 | 22 7 | 22 9 |
| $J_{3'F}$ | 53 3 | 53 4 | 53 4 | 53 6 | 53 8 |
| $J_{4'F}$ | 27 6 | 27 5 | 27 5 | 27 3 | 27 1 |

obtain accurate J-couplings and chemical shifts The results are summarized in Table 1 The conformational equilibrium around the C4'-C5' bond could be characterized following standard methods, i e a relative distribution over the staggered rotamers $\gamma^+$, $\gamma^t$ and $\gamma^-$ can be calculated form the experimental values for $J_{4'5'}$

Table 2 J couplings (Hz) of 3'-α-azidothymidine (AZT) (**2**) recorded at 500 MHz in $D_2O$

| J-couplings | 292K | 318K | 353K |
|---|---|---|---|
| $(J_{1'2'}+J_{1'2''})$ | 12 8 | 13 0 | 13 0 |
| $(J_{2'3'}+J_{2''3'})$ | 13 2 | 13 0 | 13 0 |
| $J_{3'4'}$ | 5 4 | 5 4 | 5 3 |
| $J_{4'5'}$ | 3 5 | 3 7 | 3 9 |
| $J_{4'5''}$ | 4 6 | 4 7 | 4 9 |




**Figures 1a & 1b:** Values for $J_{1'2'}$, $J_{2'3'}$ and $J_{3'4'}$ were calculated by varying the Phase angle of pseudorotation (P) from 0 - 360° at fixed Puckering amplitude ($\nu_m$) of 35° (+), of 40° (x) and of 45° (◆) using Karplus-type equation as developed by Altona *et al* [18,19]. Data points for compounds **1**, represented by (▲) were obtained through the experimental J-couplings at 500 MHz in $D_2O$ at different temperatures Starting experimental solution geometries for energy optimization, using Allinger's MM2 force field, are shown by (■) and capital letters (A, B and C) Results of these energy optimizations are then exhibited by (●) and small letters (a, b & c), respectively All starting geometries (A-C) actually converge to a set of geometries which have a narrow range of phase angles and puckering amplitudes in the South-region (P = 152 2 - 155 2°, $\nu_m$ = 36 9 - 38 0°)

and $J_{4'5''}$ This showed (Table I) that the $\gamma^+$ rotamer is favoured (49%) over $\gamma^t$ (27%) and $\gamma^-$ (24%) at 298 K[21,22] Elevation of the sample temperature slightly shifts the C4'-C5' conformational equilibrium towards $\gamma^t$ and $\gamma^-$ A graphical method showing the calculated dependence of $J_{1'2'}$ and $J_{3'4'}$ on P and $J_{2'3'}$ and $J_{3'4'}$ on P are shown in Figures 1a and 1b, respectively The three curves in Figure 1a correspond to fixed values of $\nu_m$ at 35°, 40° and 45°, respectively, while the two curves in Figure 1b correspond to fixed values of $\nu_m$ at 35° and 40° The calculations of coupling constants were based on (i) the empirically generalized Karplus equation as developed by Altona *et al*,[18] which relates vicinal proton-proton J-coupling constants to proton-proton torsion angles, and (ii) the relations $\phi$[H1'-C1'-C2'-H2'] = 121 4° +1 03 $\nu_1$, $\phi$[H2'-C2'-C3'-H3'] = 2 4° + 1 06 $\nu_2$ and $\phi$[H3'-C3'-C4'-H4'] = -124 0° + 1 09 $\nu_3$[19] Closed curves are obtained as P varies from 0° (North region) via 180° (South region), to 360° (North region) The experimental data points in Figure 1a [$J_{1'2'}$ = 9 1-8 7 Hz, $J_{3'4'}$ = 1 5 -1 75 Hz (298 to 353 K)] and Figure 1b [$J_{2'3'}$ = 5 3-5 6 Hz, $J_{3'4'}$ = 1 5 -1 75 Hz (298 to 353 K)], which represent time-averaged values of the J-coupling constants in each of the participating conformers, coincide to the South region in both Figures 1a and 1b This shows that the modified ribose ring in FLT **1** is biased towards a South-type conformation Clearly, the position of the experimental data points is determined by the J-values in each of the conformers participating in the conformational equilibrium, and their mole fractions It is possible to construct conodes between the North and South regions, such that the experimental data points fall on the conodes If done so, it follows that the conformation is biased towards the South-type conformation (> ca 90 %) In a subsequent alternative analysis, the conformation of the furanose ring in FLT can be monitored through five vicinal proton-proton J-coupling constants, namely $J_{1'2'}$, $J_{1'2''}$, $J_{2'3'}$, $J_{2''3'}$ and $J_{3'4'}$ These couplings were found to vary considerably upon changing the sample temperature, which greatly facilitated the conformational analysis via the program PSEUROT[23] We used the program PSEUROT[23] for the translation of experimental set of $^3J_{HH}$-coupling constants in terms of a two-state conformational equilibrium of the furanose ring in FLT **1** PSEUROT calculates the best fit of the five conformational parameters needed to characterize a two-state conformational equilibrium to the set of experimental vicinal couplings After convergence, one obtains the best-fit phase angle of pseudorotation (P) and the puckering amplitude ($\nu_m$) of both participating conformers, as well as their relative contributions to the conformational equilibrium For FLT, it was found on the basis of the proton-proton J-couplings data in $D_2O$ (Table 1) that PSEUROT converges towards a South-type conformation with P = 151° and $\nu_m$ = 34°, and a North-type conformation with P = -14° and $\nu_m$ = 40 5° The South conformer was found to be predominant (x(South) = 0 92 at 298 K) It should be noted that the characterisation of the North-type conformation is inherently inaccurate because of its minor contribution to the conformational equilibrium (~10% over the entire temperature traject) It is of interest to compare the apparently preferred furanose conformation in solution to the set of six FLT geometries that were found in the solid state by van Roey et al [1] (structures **I** - **IV**), and Camerman et al [2] (structures **V** and **VI**) The pseudorational parameters are **I** P = 174°, $\nu_m$ = 33°, **II** P = 175°, $\nu_m$ = 34°, **III** P = 171°, $\nu_m$ = 34°, **IV** P = 176°, $\nu_m$ = 32°, **V** P = 164°, $\nu_m$ = 36°, **VI** P = 169°, $\nu_m$ = 32° This means that the average phase angle of pseudorotation of FLT in the solid state is 171°, the average puckering amplitude is 34° It can be concluded that the result of the conformational analysis of the furanose ring in solution (P = 151°, $\nu_m$ = 34°) is grossly similar to the average over the available crystal structures (P = 171°, $\nu_m$ = 34°) The phase angle P = 151° in solution indicates that the preferred furanose geometry is an intermediate between the C2'-*endo* envelope and the C2'-*endo*/C1'-*exo* twist conformation The averaged phase angle in the solid state (P = 171°) indicates an intermediate geometry between the C2'-*endo* envelope and the C3'-*exo*/C2'-*endo* twist structure

Table 3 Comparison of solid state and solution structures of 3'-α–Fluorothymidine (1)

| X-ray structures | | | | Starting Solution geometry (based on PSEUROT)# | | | | MM2 optimized geometry | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Phase angle (P) | Puckering amplitude ($\nu_m$) | γ | χ | Phase angle (P) | Puckering amplitude ($\nu_m$) | γ¶ | χ¥ | Phase angle (P) | Puckering amplitude ($\nu_m$) | γ | χ | Steric Energy (Kcal /mol) |
| | | | | 150.8° | 34 3° | 60° | -154.3° | 152 2° | 38 0° | 58 3° | -160 0° | 18 7 |
| 174°* | 32 5° | 48 3° | -137 6° | - | - | - | - | 154 6° | 37 6° | 55 2° | -142 9° | 18 0 |
| 164°** | 36° | 50 3° | -133 4° | - | - | - | - | 155.2° | 36.9° | 57 1° | -136.0° | 18 1 |

*taken from ref [1]

**taken from ref [2]

#PSEUROT inputs were based on J-couplings at 500 MHz $^1$H-NMR in $D_2O$

¶Population was calculated (49% γ$^+$, 27% γ$^t$ and 24% γ$^-$ using Altona-Hasnoot equation [21,22]

¥The values of long range coupling constants $^3J_{C\text{-}2,H\text{-}1'}$ = 2 4 Hz and $^3J_{C\text{-}6,H\text{-}1'}$ = 3.9 Hz were measured in $D_2O$. Glycosyl torsion (χ) has been calculated using these coupling constants on the basis of modified Karplus type equation by Lemieux[24] and Davies[25] φ[C2-N1-C1'-H1'] = ± 45 3° which corresponds to two glycosidic torsions χ in *anti* range [-64° (high-*anti*) or -154 3° (*anti*)]

Two obvious explanations can be given for this difference. (i) the conformation in the solid state is partly dictated by crystal packing forces, i.e intermolecular interactions such as base-base stacking and hydrogen bonding, and (ii) solvation effects are operative in solution The remaining conformational degree of freedom,

Table 4 Starting and MM2 optimized solution structures of 3'-α–azidothymidine (2) in the North region of pseudorotational cycle

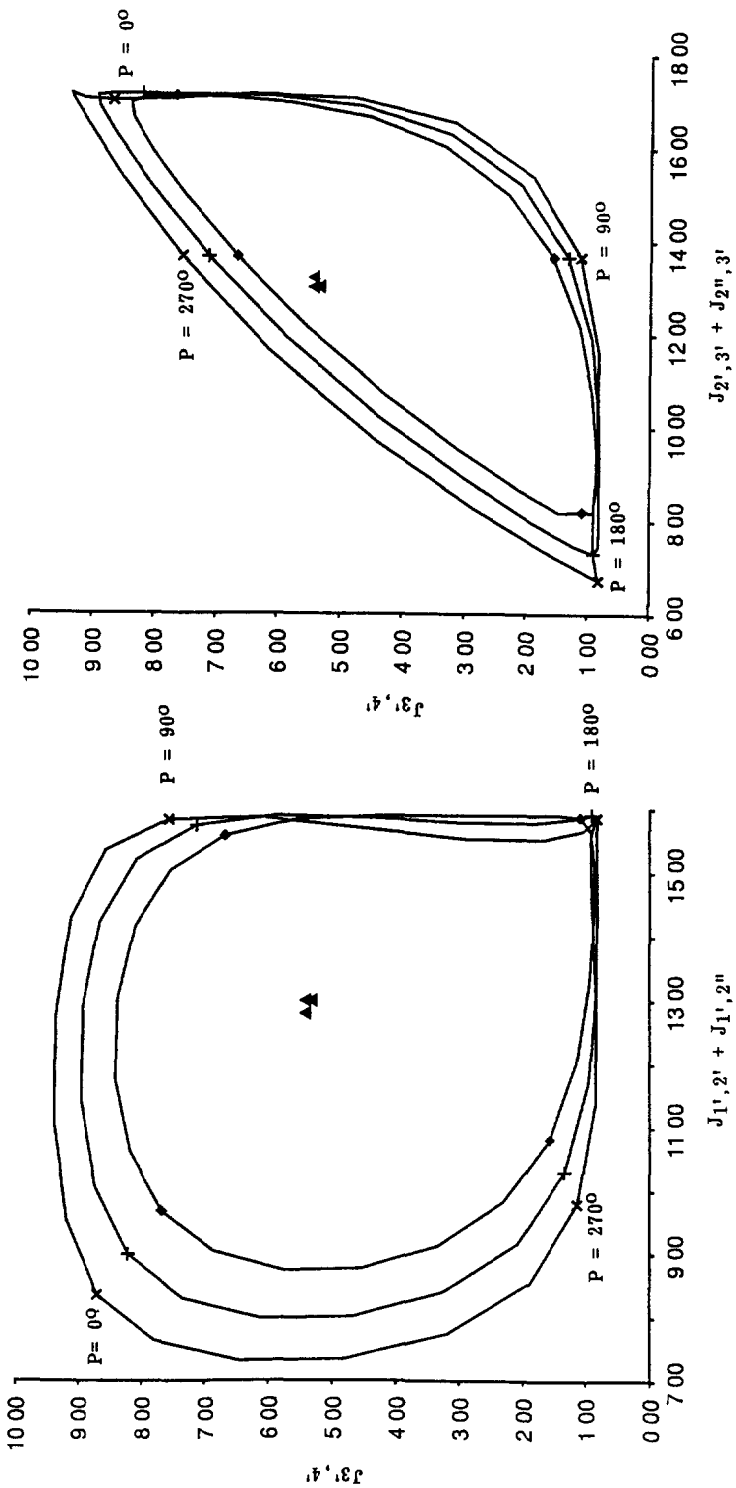
| *Starting geometry* | | | | *MM2 optimized geometry* | | | | |
|---|---|---|---|---|---|---|---|---|
| Phase angle (P) | Puckering amplitude ($\nu_m$) | γ¶ | χ¥ | Phase angle (P) | Puckering amplitude ($\nu_m$) | γ | χ | Steric Energy (Kcal /mol) |
| -41 2° | 28 2° | 60° | -170° | 19 6° | 40 1° | 56 9° | -163 0° | 18 29 |
| -22 6° | 28 7° | 64 2° | -159 4° | 20 6° | 39 5° | 56 8° | -164 5° | 18 36 |
| 0° | 29 5° | 60° | -170 5° | 18 2° | 39 8° | 57 0° | -164 1° | 18 36 |
| 20 3° | 29 3° | 60° | -170 5° | 19 1° | 39 8° | 56 2° | -162 7° | 18 31 |
| 41 4° | 29 0° | 60° | -170 5° | 24 4° | 39 9° | 56 8° | -159 1° | 18 15 |
| -43 2° | 32 5° | 60° | -170 5° | 19 2° | 39 8° | 56 9° | -163 0° | 18 29 |
| -21 0° | 32 9° | 60° | -170 6° | 20 4° | 39 8° | 56 9° | -162 0° | 18 25 |
| 0 0° | 33 4° | 60° | -170 5° | 23 6° | 40 2° | 56 6° | -158 5° | 18 12 |
| 20 8° | 33 2° | 60° | -170 5° | 20 3° | 40 1° | 57 1° | -161 8° | 18 24 |
| 42 5° | 32 9° | 60° | -170 6° | 25 6° | 40 5° | 57.1° | -158 4° | 18 14 |
| -44 4° | 36 1° | 60° | -170 6° | 18 4° | 39 7° | 56 8° | -163 7° | 18 34 |
| -21 6° | 36 4° | 60° | -170 5° | 20 7° | 39 7° | 56 9° | -161 3° | 18 23 |
| 0 1° | 37 0° | 60° | -170 5° | 17 5° | 39 9° | 56 7° | -164 2° | 18 36 |
| 21 4° | 36 7° | 60° | -170 6° | 21 3° | 40 0° | 57 0° | -159 8° | 18 18 |
| 43 7° | 36 6° | 60° | -170 5° | 27 2° | 40 6° | 56 9° | -156 4° | 18 09 |
| -47 1° | 42 7° | 60° | -170 5° | 24 2° | 39 8° | 57 1° | -157 2° | 18 10 |
| -22 9° | 42 6° | 60° | -170 6° | 20 9° | 39 8° | 57 1° | -161 0° | 18 21 |
| 0 0° | 43 3° | 60° | -170 5° | 17 1° | 40 0° | 56 9° | -163 6° | 18 34 |
| 22 7° | 42 9° | 60° | -170 5° | 22 9° | 40.7° | 56 9° | -158 7° | 18 16 |
| 46 3° | 43.2° | 60° | -170 5° | 30 9° | 40 8° | 56 6° | -152 6° | 18 06 |
| | | | | *21 6° | *40 0° | *56 9° | *-160 8° | |

¶ Populations were calculated ( 55% $\gamma^+$, 32% $\gamma^t$ and 13% $\gamma^-$ ) using ref [21,22]

¥ Glycosyl torsion (χ) has been calculated on the basis of $^3J_{C\text{-}2,H\text{-}1'}$ = 1 95 Hz and $^3J_{C\text{-}6,H\text{-}1'}$ = 3 9 Hz measured in $D_2O$ Using modified Karplus equation by Lemieux[24] and Davies[25] we were able to predict φ[C2-N1-C1'-H1'] = ± 50° to which 120° is subtracted to obtain two dihedral angles for χ in *anti* range [-70° (high-*anti* ) or -170° (*anti*)]

* Average of MM2 optimized geometries of the North-type furanose of AZT

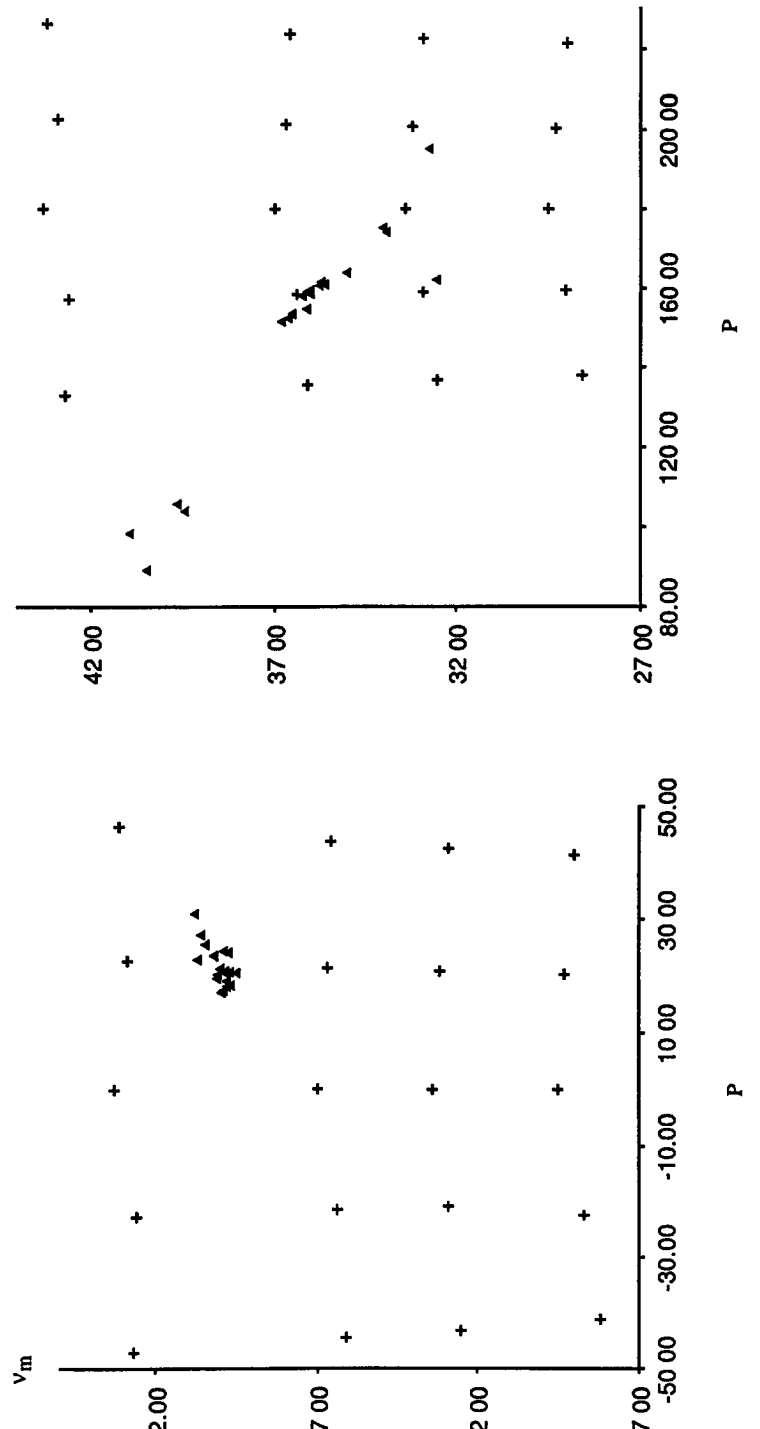
i e rotation around the C1'-N1 bond determining *syn* or *anti* orientation of the base was probed via the vicinal $^{13}C$ - $^1H$ J-coupling constants $^3J_{H1'\text{-}C2}$ and $^3J_{H1'\text{-}C6}$, using the Karplus type equation as proposed by Lemieux[24] and Davies[25] [$^3J_{H1'\text{-}C2} = 6\,7\cos^2\phi - 1\,3\cos\phi$] Based on the experimental observations that $^3J_{H1'\text{-}}$




**Figures 2a & 2b:** Values for ($J_{12'} + J_{1'2''}$), ($J_{23'} + J_{2''3'}$) and $J_{34'}$ were calculated by varying the Phase angle of pseudorotation (P) from 0 - 360° at fixed Puckering amplitude ($\nu_m$) of 30° (♦), of 35° (+) and of 40° (×) using Karplus-type equation as developed by Altona *et al* 18,19 Data points for compound **2**, represented by (▲) were obtained through the experimental J-couplings at 500 MHz in $D_2O$ at different temperatures

$_{C2}$ = 2 4 Hz, and $^{3}J_{H1'-C6}$ = 3 9 Hz, it can be safely concluded that the thymine base is in the *anti*-domain, which is normal for pyrimidine bases. The next stage of our conformational analysis of FLT comprised a set of MM2 molecular mechanics calculations[26] A starting geometry was constructed on the basis of the NMR-PSEUROT derived structural model, i e the C4'-C5' conformation was placed in $\gamma^{+}$, the furanose ring was put in a South conformation (P = 151°, and $\nu_m$ = 34°), and the base was placed in the *anti* domain By employing Lemieux[24] and Davies[25] equation for determination of glycosyl torsion, we obtained two dihedral angles for $\chi$ in *anti* range [-63 7 (high-*anti* ) or -154 3 (*anti*)] because of even functional nature of "cos" function We performed MM2 calculation with both dihedral angles for $\chi$, and the one in *anti* range i e $\chi$ = -154 3° was preferred by 2 Kcal/mol over the high-*anti* After MM2 optimization, we obtained a final geometry with the following characteristics (see also Table 3) The torsion angle $\gamma$[O5'-C5'-C4'-C3'] was calculated to be 58 3°, which favourably agrees with the $\gamma$ values for crystal stuctures **I - VI**, i e 48 3°, 51 9°, 45 8°, 50 5°, 50 3°, and 53 4°, respectively (average 50 0°) The pseudorotational parameters in the MM2 optimized geometry of FLT are P = 152° and $\nu_m$ = 38° Thus, the phase angle agrees perfectly with the PSEUROT result (vide supra), while the puckering amplitude is predicted 4° larger by MM2 *Thus, MM2 predicts the same mode of puckering as found by NMR J-coupling analysis, but at the same time slightly overestimates the extent of puckering* The glycosidic torsion angle $\chi$[C2-N1-C1'-O4'] is found to be -160° (Table 3) in the MM2 optimized model (Figure 4b) This number is also in good agreement with the X-ray data, which showed for the structures I - VI -137 6°, -153 2°, -129 5°, -149 4°, -133 4°, and -153 5°, respectively [average -144°, (Figure 4a)] These results confirm that MM2 molecular mechanics calculations constitute a very valuable complement to NMR studies on e g modified, flexible nucleosides In the present case it is found that the torsion angles $\gamma$ and $\chi$, as well as the phase angle of pseudorotation are very closely predicted, while the puckering amplitude of the furanose ring is overestimated by 4°

*Solution conformation of AZT* In our attempts to elucidate the conformation of AZT (**2**), we have recorded its 500MHz $^{1}$H-NMR spectra at the sample temperatures 292K, 318K and 353K, a sample concentration of approximately 20 mM was used It was possible to extract some of the J-coupling constants except most important $J_{1'2'}$ , $J_{1'2"}$, $J_{2',3'}$ and $J_{2"3'}$ because of identical chemical shifts for H2' and H2" protons (Table 2) Note that it was not possible to perform pseudorotational calculations using PSEUROT program[23] because we were unable to extract individual $J_{1'2'}$, $J_{1'2"}$, $J_{2',3'}$ and $J_{2"3'}$ couplings[27] We have been however able to estimate ($J_{1'2'}$ + $J_{1'2"}$ ) and ($J_{2'3'}$ + $J_{2"3'}$) by analysis of $^{1}$H-NMR spectra Further studies of the conformational analysis of the furanose ring of AZT was based on these data (Table 2) The data show that almost no changes in the J-couplings are found upon varying the sample temperature This has further complicated the conformational analysis of the furanose moiety (vide infra) Conformational analysis of the C4'-C5' bond could be easily performed As in the case of FLT, it is found that the $\gamma^{+}$ rotamer is predominantly populated (55%) at 292 K while populations of $\gamma^{t}$ and $\gamma^{-}$ are 32 and 13% respectively The conformation around the C1'-N1 bond ($\chi$) was again based on the vicinal $^{1}$H, $^{13}$C J-couplings, $^{3}J_{H1'-C2}$ and $^{3}J_{H1'-C6}$, using the equation of Lemieux[24] and Davies[25] From the experimental data $^{3}J_{H1'-C2}$ = 2 0 Hz and $^{3}J_{H1'-C6}$ = 3 9 Hz, it can be concluded that the thymine base is in the *anti* domain [$\chi$ = -70° (high-*anti* ) or -170° (*anti*)] , as was also found for FLT We however performed MM2 calculation with glycosidic torsion $\chi$ = -170° (see Table 4) Conformational analysis of the furanose ring in AZT using the PSEUROT procedure was impossible since H2' and H2" protons showed identical chemical shifts and therefore we were left with only option which was to use temperature invariant




**Figures 3a & 3b:** (3a) Starting geometries (+) for MM2 minimization for North-type furanose of AZT (2) were chosen with Phase angle of pseudorotation (P = -41.2 to 46 3°) and the puckering amplitudes ($\nu_m$ = 28.2 - 43 2°). Note that the resulting geometries after energy optimization (▲) converge to a narrow range of P = 17.1 to 30 9° and $\nu_m$ = 39 5 - 40.8°; (3b) Starting geometries (+) for South-type furanose were chosen with Phase angle of pseudorotation (P = 132 9° to 226.3°) and the puckering amplitudes ($\nu_m$ = 28 6 - 42 7°) Most of the resulting geometries after energy optimization (▲) converge to a narrow range of P = 151 4 to 175.2° and $\nu_m$ = 32 5 - 36 8°

($J_{1'2'} + J_{1'2''}$) and ($J_{2'3'} + J_{2''3'}$) and $J_{3'4'}$ coupling constants[27] We turned to a graphical method showing the calculated dependencies of ($J_{1'2'} + J_{1'2''}$) and $J_{3'4'}$ on P (Fig. 2a) and ($J_{2'3'} + J_{2''3'}$) and $J_{3'4'}$ on P (Fig 2b) The

Table 5 Starting and MM2 optimized solution structures of 3'-α–azidothymidine (2) in the South region of pseudorotational cycle

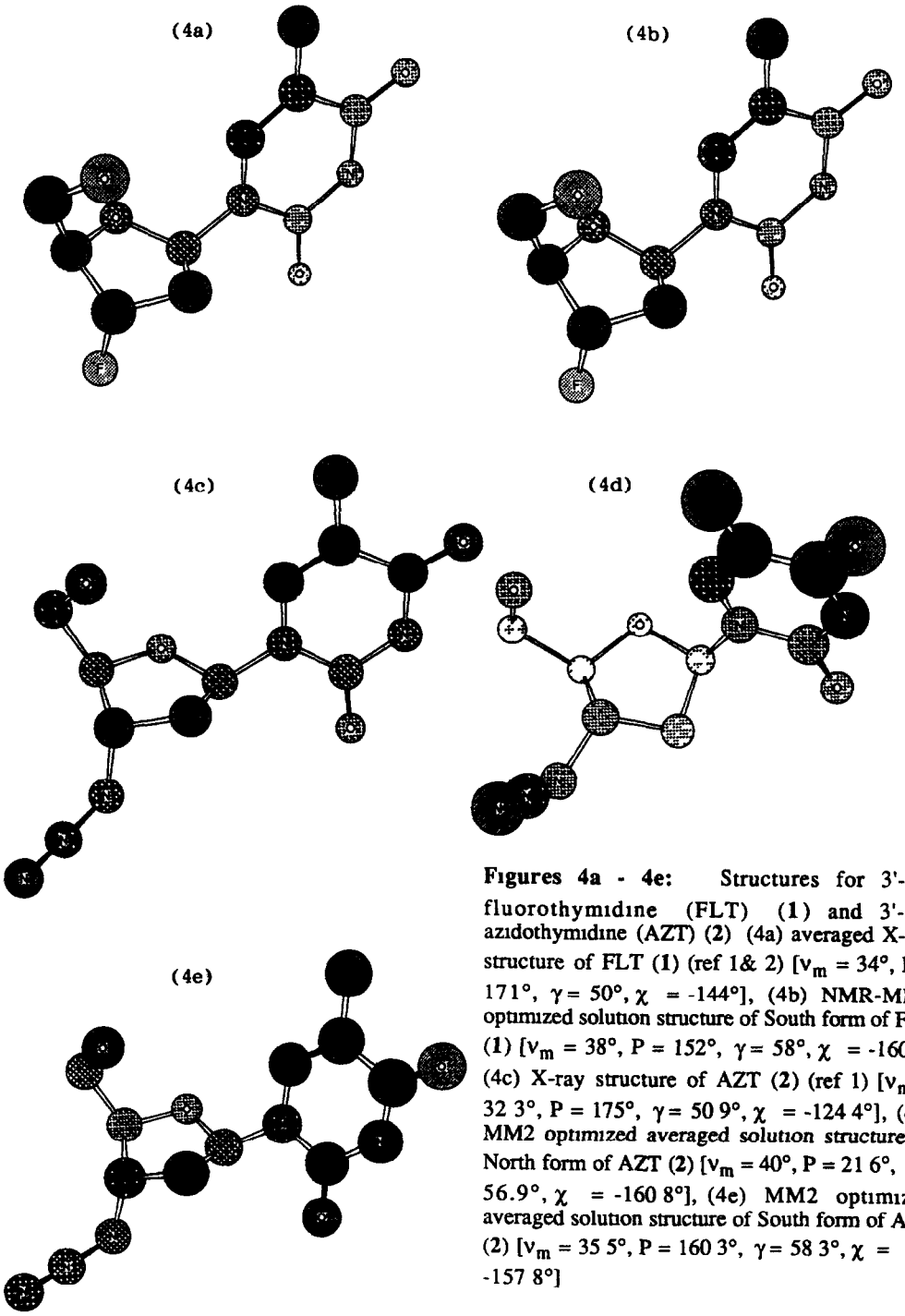
| *Starting geometry* | | | | *MM2 optimized geometry* | | | | |
|---|---|---|---|---|---|---|---|---|
| Phase angle (P) | Puckering amplitude ($\nu_m$) | γ¶ | χ¥ | Phase angle (P) | Puckering amplitude ($\nu_m$) | γ | χ | Steric Energy (Kcal /mol) |
| 137 9° | 28 6° | 60° | -170 5° | #89 2° | #40 4° | #61 4° | #-160 3° | 18 80 |
| 159 5° | 29 0° | 60° | -170 5° | 151 4° | 36 8° | 58 3° | -149 1° | 18 90 |
| 180 0° | 29 5° | 60° | -170 5° | 160 9° | 35 6° | 58 1° | -156 5° | 19 07 |
| 200 3° | 29 3° | 60° | -170 5° | 158 0° | 36 2° | 58 4° | -159 7° | 19 05 |
| 221 4° | 29 0° | 60° | -170 6° | 160 7° | 35 8° | 58 4° | -161 7° | 19 14 |
| 136 8° | 32 5° | 60° | -170 6° | #98 4° | #40 9° | #61 4° | #-157 4° | 18 87 |
| 159 0° | 32.9° | 60° | -170 5° | 152 3° | 36 6° | 58 1° | -149 5° | 18 91 |
| 180 0° | 33 4° | 60° | -170 6° | 163 9° | 35 0° | 58.2° | -157 6° | 19 15 |
| 200 8° | 33 2° | 60° | -170 6° | 161 5° | 35 7° | 58 4° | -161 4° | 19 15 |
| 222 5° | 32 9° | 60° | -170 5° | #195 1° | #32 7° | #55 9° | #-168 8° | 20 21 |
| 135 6° | 36 1° | 60° | -170 6° | #104 0° | #39 4° | #61 8° | #-154 8° | 18 94 |
| 158 4° | 36 4° | 60° | -170 5° | 153 4° | 36 5° | 58 2° | -149 9° | 18 93 |
| 180 0° | 37 0° | 60° | -170 5° | 159 0° | 36 1° | 58 4° | -158 1° | 19 05 |
| 201 4° | 36 7° | 60° | -170 5° | 162 1° | 32 5° | 58 0° | -162 2° | 19 18 |
| 223 7° | 36 6° | 60° | -170 5° | 175 2° | 34 0° | 59 0° | -165 6° | 19 59 |
| 132 9° | 42 7° | 60° | -170 5° | #105 8° | #39 6° | #61 7° | #-153 3° | 18 94 |
| 157 1° | 42 6° | 60° | -170 6° | 154 6° | 36 1° | 58 0° | -149 9° | 18 97 |
| 180 0° | 43 3° | 60° | -170 6° | 159 4° | 36 0° | 58 5° | -159 0° | 19 07 |
| 202 7° | 42 9° | 60° | -170 5° | 174 1° | 33.9° | 58 5° | -164 7° | 19 54 |
| 226 3° | 43 2° | 60° | -170 6° | 158 5° | 36 0° | 58 5° | -161 7° | 19 09 |
| | | | | *160 3° | *35 5° | *58 3° | *-157 8° | |

¶ Populations were calculated ( 55% $\gamma^+$, 32% $\gamma^t$ and 13% $\gamma^-$ ) using ref [21,22]

¥ Glycosyl torsion (χ) has been calculated on the basis of $^3J_{C\text{-}2,H\text{-}1'} = 1\ 95$ Hz and $^3J_{C\text{-}6,H\text{-}1'} = 3\ 9$ Hz measured in $D_2O$ Using modified Karplus equation by Lemieux[24] and Davies[25] we were able to predict ϕ[C2-N1-C1'-H1'] = ± 50° to which 120° is subtracted to obtain two dihedral angles for χ in *anti* range [-70° (high-*anti*) or -170° (*anti*)]

* Average of MM2 optimized geometries (excluding the geometries marked by #) of the South-type furanose of AZT

three curves in Figures 2a and 2b correspond to fixed values of $\nu_m$ at 30°, 35° and 40° The calculations were performed identically as described for FLT, but using different electronegativity for 3'-$N_3$ = 0 85 Closed curves are obtained as P varies from 0° (North region) via 180° (South region), to 360° (North region) The experimental data points in Figure 2a [($J_{1'2'} + J_{1'2''}$) = 12 8-13.0 Hz, $J_{3'4'}$ = 5 4-5 3 Hz (292 to 353K)] and




**Figures 4a - 4e:** Structures for 3'-α-fluorothymidine (FLT) (1) and 3'-α-azidothymidine (AZT) (2) (4a) averaged X-ray structure of FLT (1) (ref 1& 2) [$\nu_m$ = 34°, P = 171°, $\gamma$ = 50°, $\chi$ = -144°], (4b) NMR-MM2 optimized solution structure of South form of FLT (1) [$\nu_m$ = 38°, P = 152°, $\gamma$ = 58°, $\chi$ = -160°], (4c) X-ray structure of AZT (2) (ref 1) [$\nu_m$ = 32 3°, P = 175°, $\gamma$ = 50 9°, $\chi$ = -124 4°], (4d) MM2 optimized averaged solution structure of North form of AZT (2) [$\nu_m$ = 40°, P = 21 6°, $\gamma$ = 56.9°, $\chi$ = -160 8°], (4e) MM2 optimized averaged solution structure of South form of AZT (2) [$\nu_m$ = 35 5°, P = 160 3°, $\gamma$ = 58 3°, $\chi$ = -157 8°]

Figure 2b [($J_{2'3'} + J_{2''3'}$) = 13 2-13 0 Hz, $J_{3'4'}$ = 5 4-5 3 Hz (292 to 353K)] represent time-averaged values of the J-coupling constants in each of the participating conformers. The experimental data points in both Figures 2a and 2b nearly coincide at the spot which is half-way between the North (P = 0°) and South (P = 180°) region in the pseudorotational cycle in all three graphs at $\nu_m$ = 30, 35 and 40° It is of interest to note that the experimental data points can be extrapolated to any two points in North and South regions for $\nu_m$ = 30, 35 or 40° The crude estimation of North and South population from graphical representations in Figure 2a or 2b does not allow us to define their exact geometrical nature The graphical representation in Figure 2b however showed us quite expectedly[18] that it is possible to arbitrarily restrict[19] the regions for phase angle of pseudorotation (P) and puckering amplitudes ($\nu_m$) in both South (130 to 230°, $\nu_m$ = 30-45°) and North (-40 to +45°, $\nu_m$ = 30-45°) conformers At this point of our search for conformational space for North- and South-type conformers for AZT, we initiated molecular mechanics calculations (MM2) as an independent method of arriving at an energy minimized structure Clearly, MM2 calculations[26] constitute an important part of the construction of a structural model for AZT in solution We decided to perform two sets of MM2 calculations The first starts from P-values between -40° and + 45° (North region), while the second set starts from P-values between 130° and 230° (South region) In 20 MM2 calculations in the North region, the puckering amplitude and the phase angle were varied between 30-45° and -40 to 45°, respectively while the torsions $\gamma$ and $\chi$ were set at 60°, and −170° in the starting geometries without any constraint The results of these calculations are summarized in Table 4 and Fig 3a All 20 MM2 calculations in the North region converge to virtually a similar structure with P = 17 1 to 30.9° (average P = 21 6°), $\nu_m$ = 39.5 to 40 8° (average $\nu_m$ = 40°), $\chi$ = -152 6 to -164 5° (averge $\chi$ = -160.8°) and $\gamma$ = 56 2 to 57 1° (averge $\gamma$ = 56 9°) (see Table 4, Fig 3a) The molecular model representing the average MM2 minimized North structure is shown in Figure 4d. In 20 MM2 calculations in the South region, the puckering amplitude and the phase angles were varied between 30-45° and 130 to 230°, respectively while the torsions $\gamma$ and $\chi$ were set at 60°, and −170° in the starting geometries without any constraint The results of these calculations are summarized in Table 5 and Fig 3b Fifteen resulting structures out of 20 results in the South region converge to virtually a similar structure with P = 151 4 to 175 2° (average P = 160 3°), $\nu_m$ = 32 5 to 36 8° (average $\nu_m$ = 35 5°), $\chi$ = -149 1 to -165 6° (averge $\chi$ = -157 8°) and $\gamma$ = 58 0 to 59 0° (averge $\gamma$ = 58 3°) (see Table 5, Fig 3b) The molecular model representing the average MM2 minimized South structure is shown in Figure 4e The North form has P = 22° which is close to unsymmetrical twist with major C3'-*endo* and minor C4'-*exo* twist structure, while the South form with P = 160° is close to the envelope C2'-*endo* geometry Note that all six X-ray crystal structure known for AZT[3-8] belong to the South-type geometry, and they distinctly fall into two groups (i) P = 175°, $\nu_m$ = 32° and (ii) P = 215°, $\nu_m$ = 36° *The MM2 calculated South geometry for AZT is grossly similar to one of the available X-ray crystal structures (P = 175°, $\nu_m$ = 32.3°)[1] (Fig 4c) which implies an unsymmetrical twist with major C2'-endo and minor C3'-exo pucker*

To date there is no report of any AZT crystal structure representing the North-type geometry although the North and South-type furanose are equally populated in solution Clearly it is the overall solution structure that has a direct implication in our understanding of the structure-activity relationship of an antiviral substance in the biological system Note that X-ray structures of AZT[3-8] only partly represent the overall structural view found in solution This discrepancy between solution and solid state structures may indicate the danger of only using crystal structural data for the formulation of structure-activity relationships for a candidate drug Molecular

structures of small hydrophilic molecules in the solid state can be more or less determined by crystal packing forces, and no solvation effects or inherent dynamics can be examined We suggest that it is therefore better to rely also on data that were obtained in aqueous solution. J-coupling data as measured by high-field NMR reflect the *dynamic structural properties* of modified nucleosides, and molecular mechanics calculations can be helpful in order to translate this information into structural models

**Acknowledgements:** We thank Swedish Board for Technical Development and Swedish Natural Science Research Council and Medivir AB, Lunastigen 7, S-141 44 Huddinge, Sweden for generous financial supports and Wallenbergstiftelsen, Forskningsrådsnämnden (FRN) and University of Uppsala for funds for the purchase of a 500 MHz Bruker AMX NMR spectrometer. Financial Support from the European Molecular Biology Organization (EMBO) through two-year EMBO fellowship to LHK is gratefully acknowledged

## REFERENCES

1 P van Roey and R F Schinazi, *Antiviral Chem & Chemotherapy*, **1**, 93 (1990)
2 N Camerman, D Mastropaolo and A Camerman, *Proc Natl Acad Sci*, **87**, 3534 (1990)
3 G I Birnbaum, J Giziewicz, E J Gabe, T-S Lin and W H Prusoff, *Can J Chem*, **65**, 2135 (1987)
4 A Camerman, D Mastropaolo and N. Camerman, *Proc Natl Acad Sci*, **84**, 8239 (1987).
5 P. van Roey, J M Salerno, W L Duax, C K Chu, M K Ahn and R F Schinazi, *J Am Chem Soc*, **110**, 2277 (1988)
6 R Parthasarathy and H Kim, *Biochem & Biophys Res Comm*, **152**, 351 (1988).
7 I Dyer, J.N. Low, P. Tollin, R H wilson and R.W Howie, *Acta Cryst*, **C44**, 767 (1988)
8 G V Gurskaya, E N. Tsapkina, N V Skaptasova, A A Lindeman, S V and Yu. T Struchkov, *Dokl Akad Nauk USSR*, **291**, 854 (1986).
9 R Pauwels, M Baba, J Balzarini, P Herdewijn, J Desmyter, M J Robins, R Zou, D Madej and E de Clercq, *Biochem Pharm*, **37**, 1317, (1988).
10 L Vrang, H Bazin, G Ramaud, J Chattopadhyaya and B Öberg, *Antiviral Res*, **7**, 139, (1987)
11 B Eriksson, L Vrang, H Bazin, J Chattopadhyaya and B Öberg, *Antimicrobial Agents and Chemotherapy*, **31**, 600, (1987)
12 H Bazin, J Chattopadhyaya, R Datema, A-C Ericson, G Gilljam, N G Johansson, J Hansen, R Koshida, K Moelling, B. Öberg, G Ramaud, G Stening, L Vrang, B Wahren and J-C Wu, *Biochem Pharmacol*, **38**, 109-119 (1989)
13 F Barré-Sinoussi, J.C. Chermann,, F Rey, M T Bygetre, S Chamaret, J Gruest, C Dauguet, C Axler-Blin, F Vézinet-Brun, C Rouzioux, W Rozenbaum and L Montagnier *Science* (Washington, D C ) **220**, 868 (1983)
14 B Lundgren, D Bottiger, E. Ljungdahl-Ståhle, E Norby, L Ståhle, B Wahren and B Öberg, *J Acquired & Immune Deficiency Syndromes*, (in press, 1991)
15 Phase I clinical trial is presently conducted by American Cyanamide in USA
16 T F Lai and R E Marsh, *Acta Cryst*, **B28**, 1982, (1972)
17 K Shikata, T Ueki and T. Mitsui, *Acta Cryst*, **B29**, 31, (1973)
18 C Altona and M Sundaralingam *J Am Chem Soc* **94**, 8205 (1972) and *ibid* **95**, 2333 (1973)
19 H P M de Leeuw, C A G Haasnoot, and C Altona *Isr J Chem* **20**, 108 (1980)
20 W Saenger In "Principles of Nucleic Acids Structure" Springer Verlag, New York, 1984
21 Limiting values for $J_{4'5'}$ and $J_{4'5''}$ in the staggered C4'-C5' rotamers are as follows $\gamma+$ $J_{4'5'}$ = 2 4 Hz, $J_{4'5''}$ = 1 3 Hz Rotamer $\gamma^t$ $J_{4'5'}$ = 2 6 Hz, $J_{4'5''}$ = 10 5 Hz Rotamer $\gamma-$ $J_{4'5'}$ = 10.6 Hz, $J_{4'5''}$ = 3 8 Hz See also ref 25
22 C A G Haasnoot, F A A M de Leeuw, H P M de Leeuw and C Altona *Recl Trav Chim Pays-Bas* **98**, 576 (1979)
23 F A A M de Leeuw and C Altona *J Comp Chem* **4**, 438 (1983) and PSEUROT QCPE Program No 463
24 R U Lemieux, T L Nagabhushan, B Paul *Can J Chem* **50**, 773 (1972)
25 D B Davies, Progress in NMR Spectroscopy, Pergamon Press, **12**, 135 (1978)
26 Prof Allinger's MM2 force field as implented by J W Ponder, which is marketted as Chem3D plus (version 3.0) by Cambridge Scientific Computing, Cambridge, Massachusetts, USA
27 L J Rinkel and C Altona, *J Biom Struct Dynam* **4**, 621 (1987)
28 W K Olson *J Am Chem Soc* **104**, 278 (1982), C A.G. Haasnoot, F A A M de Leeuw, H P M de Leeuw and C Altona, *Org Mag Res* **15**, 43 (1981)