Conformational Analysis of β-D-Ribo-, β-D-Deoxyribo-, β-D-Arabino-, β-D-Xylo-, and β-D-Lyxo-nucleosides from Proton–Proton Coupling Constants

By **Frank A. A. M. de Leeuw** and **Cornelis Altona,*** State University Leiden, Gorlaeus Laboratories, P.O. Box 9502, 2300 RA Leiden, The Netherlands

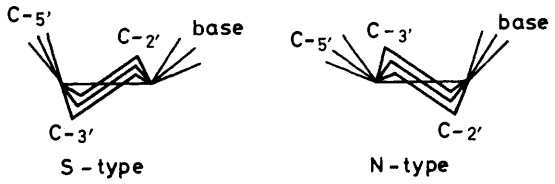
A new n.m.r. coupling constant torsion angle relation is utilized to explore the effect of exocyclic oxygen substituents at C-2′ and C-3′ on $J_{1'2'}$, $J_{2'3'}$, and $J_{3'4'}$ in the furanose ring of β-D-ribo-, β-D-arabino-, β-D-xylo-, and β-D-lyxo-nucleosides. The five vicinal couplings in β-D-deoxyribonucleosides are also investigated. Calculated coupling constants for the full pseudorotational itinerary at different values of the puckering amplitudes are given. It is shown that experimental coupling constants, taken from the literature, can be satisfactorily explained in all cases on the basis of a two-state (N/S)-model of the furanose pseudorotation. Good estimates of both N- and S-geometry and the conformational equilibrium constant in five-membered rings of ribose, deoxyribose, xylose, lyxose, and arabinose derivatives are given. The judicious use of geometrical information obtained by *X*-ray crystallographic studies appeared indispensable for this type of analysis.

A NEW empirical generalization of the classical Karplus equation (1) was introduced in a previous paper.[1] The

$$^3J_{HH} = P_1\cos^2\phi + P_2\cos\phi + P_3 + \Sigma\Delta\chi_i\,[P_4 + P_5\cos^2(\xi_i\phi + P_6|\Delta\chi_i|)] \quad (1)$$

first three terms describe the dependency of the vicinal coupling constant in a given H–C–C–H fragment on the proton–proton torsion angle ϕ. The remaining terms account for the dependency of $^3J_{HH}$ on the electronegativity and relative orientation of substituents S_i. Optimum values of the parameters P_1—P_7 were empirically determined using an iterative least-squares procedure and are based on 315 experimental couplings and corresponding calculated geometries. The analytical expression (1) was essentially derived from MO calculations.

In combination with the concept of pseudorotation equation (1) can be used to deduce the solution conformation of a pentose ring.[2] According to overwhelming n.m.r.[2] and *X*-ray evidence, a pentose ring in nucleosides and nucleotides is generally engaged in a conformational equilibrium described by N- and S-type conformers [3,4] (see Scheme). The N-type conformations are characterized by a positive value of the torsion angle C-1′–C-2′–C-3′–C-4′; in S-type conformations this torsion has a negative value.[3,4] The experimental n.m.r. coupling constants represent time-averaged couplings related to the couplings of the individual conformers and their relative populations in equilibrium as in equation (2)



SCHEME

$$J_{obs} = X_N J_N + (1 - X_N)J_S \quad (2)$$

where J_N and J_S are the coupling constants belonging to the pure N- and S-type conformers and X_N is the mole fraction of the N-type conformer. Using empirical correlations [5] taken from 178 crystal structures of ribo- and deoxyribo-nucleosides the relationship between the pseudorotation parameters (phase angle *P* and puckering amplitude Φ_m) † on the one hand and the proton–proton torsion angles on the other, was established. As the proton–proton torsions in the pure N- and S-type conformers are expressed as a function of P_N, Φ_N and P_S, Φ_S, this gives, in combination with equations (1) and (2), $^3J_{HH}$ as a function of the conformational parameters involved.

$$J_{HH} = f(P_N, \Phi_N, P_S, \Phi_S, X_N) \quad (3)$$

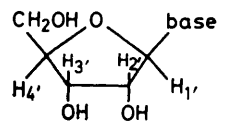
In an earlier paper a method was presented whereby the best fit for the conformational parameters from the observed coupling constants was deduced by means of an iterative least-squares computer program.[2] In this paper a more facile method is described for determination of the conformational parameters in ribo-, deoxyribo-, arabino-, xylo-, and lyxo-nucleosides and -nucleotides. Other approaches to the same problem, advocated by Davies and Danyluk [6] and by Guschlbauer,[7,8] are discussed.

PROCEDURE

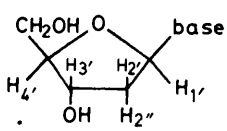
The link between proton–proton torsion angles ϕ_{HH} and the endocyclic torsion angles ϕ_{XX} (X = C, O, N, *etc.*) is usually formed by the implicit or explicit assumption of trigonal (120°) projection symmetry, in other words, of tetrahedral bond angles. This assumption, already of doubtful validity in the case of six-membered rings, is certainly too gross an approximation for five-membered rings. Careful evaluation of neutron diffraction data, in conjunction with *X*-ray results, allowed us to establish semi-empirical correlations between ϕ_{HH} and ϕ_{XX} for ribose and deoxyribose rings.[5] Because the torsion angles ϕ_{XX} are inter-related *via* the pseudorotation equation,[3,4] the ϕ_{HH} values are also inter-related, and it is a simple matter to express the ϕ_{HH} values along the various C–C bonds as

† Throughout this communication Φ_m denotes the puckering amplitude in general. Φ_N and Φ_S are used to denote the puckering amplitude for N- and S-type conformations. P_N and P_S are defined analogously.

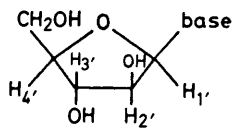
functions of the pseudorotation parameters P and Φ_m.[5] In the present work the established [5] correlations between P, Φ_m, and ϕ_{HH} of ribose (1) and deoxyribose (2) systems are used. For arabino-, lyxo-, and xylo-nucleosides such a direct approach is as yet not feasible because the required large number of solid state data is unavailable. Instead,



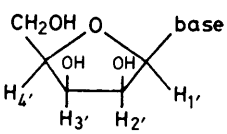
(1)



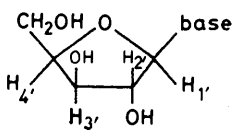
(2)



(3)



(4)



(5)

it is assumed that 2'-H in arabinose (3) and in lyxose (4) are phase-shifted by 120° with respect to 2'-H in the ribose ring and the same approximation is introduced for 3'-H in lyxose and xylose (5). In this way equations (4)—(6) are obtained.

β-D-arabinose

$$\begin{aligned}\phi_{1'2'} &= 3.3° + 1.102\ \Phi_m\cos(P - 144°)\\ \phi_{2'3'} &= 120.2° + 1.090\ \Phi_m\cos P\\ \phi_{3'4'} &= -124.9° + 1.095\ \Phi_m\cos(P + 144°)\end{aligned} \quad (4)$$

β-D-lyxose

$$\begin{aligned}\phi_{1'2'} &= 3.3° + 1.102\ \Phi_m\cos(P - 144°)\\ \phi_{2'3'} &= 0.2° + 1.090\ \Phi_m\cos P\\ \phi_{3'4'} &= -4.9° + 1.095\ \Phi_m\cos(P + 144°)\end{aligned} \quad (5)$$

β-D-xylose

$$\begin{aligned}\phi_{1'2'} &= 123.3° + 1.102\ \Phi_m\cos(P - 144°)\\ \phi_{2'3'} &= -119.8° + 1.090\ \Phi_m\cos P\\ \phi_{3'4'} &= -4.9° + 1.095\ \Phi_m\cos(P + 144°)\end{aligned} \quad (6)$$

Inspection of the relative orientation of the substituents and use of equations (4)—(6) leads to the conclusion that the following approximations are valid:

$$\begin{aligned}J_{1'2'}\ (\text{ribo}) &\simeq J_{1'2'}\ (\text{xylo})\\ J_{1'2'}\ (\text{arabino}) &\simeq J_{1'2'}\ (\text{lyxo})\\ J_{3'4'}\ (\text{ribo}) &\simeq J_{3'4'}\ (\text{arabino})\\ J_{3'4'}\ (\text{xylo}) &\simeq J_{3'4'}\ (\text{lyxo})\end{aligned}$$

In Tables 1—3 the calculated coupling constants for the full pseudorotational circuit at two different values of Φ_m are given. For intermediate values of Φ_m the coupling constant can be accurately determined by interpolation [*e.g.* the interpolated value of $J_{1'2'}$ in ribose for P 162° and Φ_m 38° is 7.81 Hz, the directly calculated value using equation (1) is 7.83 Hz]. A good approximation of the coupling constant values $J_{1'2'}$, $J_{2'3'}$, and $J_{3'4'}$ for 2'-deoxyribose bearing other 2'-substituents such as amino, chlorine, or fluorine can be made from the corresponding ribose and deoxyribose data by means of linear interpolation or

TABLE 1

Calculated coupling constants (Hz) in ribose as a function of P and Φ_m. Intervals of P are 18° except at centres of P found in the solid state

	Φ_m 35°			Φ_m 40°		
P(°)	$J_{1'2'}$	$J_{2'3'}$	$J_{3'4'}$	$J_{1'2'}$	$J_{2'3'}$	$J_{3'4'}$
0	1.02	5.05	8.14	1.02	4.33	8.61
9	1.09	5.11	8.48	1.02	4.40	8.93
18	1.25	5.29	8.71	1.13	4.61	9.13
27	1.53	5.58	8.83	1.39	4.96	9.24
36	1.95	5.97	8.87	1.81	5.42	9.28
54	3.15	6.88	8.71	3.15	6.56	9.13
72	4.63	7.69	8.14	4.85	7.59	8.61
90	6.00	8.03	7.07	6.40	8.03	7.50
108	7.00	7.71	5.53	7.45	7.61	5.78
126	7.55	6.90	3.77	7.99	6.58	3.77
144	7.72	5.99	2.25	8.14	5.45	2.07
153	7.68	5.61	1.69	8.11	4.99	1.48
162	7.55	5.32	1.28	7.99	4.65	1.08
171	7.33	5.14	1.01	7.78	4.44	0.85
180	7.00	5.08	0.85	7.45	4.36	0.77
189	6.56	5.14	0.78	6.99	4.44	0.77
198	6.00	5.32	0.76	6.40	4.65	0.81
216	4.63	5.99	0.76	4.85	5.45	0.87
234	3.15	6.90	0.76	3.15	6.58	0.81
252	1.95	7.71	0.85	1.81	7.61	0.77
270	1.25	8.03	1.28	1.13	8.03	1.08
288	1.02	7.69	2.25	1.02	7.59	2.07
306	1.03	6.88	3.77	1.15	6.56	3.77
324	1.06	5.97	5.53	1.24	5.42	5.78
342	1.03	5.29	7.07	1.15	4.61	7.50
351	1.01	5.11	7.68	1.08	4.40	8.14

extrapolation according to the electronegativity of the 2'-substituent [equation (7) where J_R and J_D are the couplings

$$J = J_R + \Delta\chi(J_D - J_R)/1.3 \quad (7)$$

in ribose and deoxyribose rings, respectively]. In calculations for cases where substituents other than hydrogen or oxygen are present on the C–C fragment along which the $^3J_{HH}$ value is to be determined, it is advisable to adhere to the Huggins' electronegativity scale,[9] on which the parametrization of equation (1) is based. $\Delta\chi$ Values [9] are: H = 0, Si = −0.3, P = 0.05, C = S = 0.4, I = 0.45, Br = 0.75, N = 0.85, Cl = 0.95, O = 1.3, F = 1.7. The azido-group (N_3) behaves as if its $\Delta\chi$-value is *ca.* 1.05.[10] For example, the interpolated values for 2'-deoxy-2'-amino-ribose (P = 162° and Φ_m = 35°) are: $J_{1'2'}$ 8.41, $J_{2'3'}$ 5.57, and $J_{3'4'}$ 1.30 Hz whereas the directly calculated values obtained from equation (1) are: $J_{1'2'}$ 8.42, $J_{2'3'}$ 5.20, and $J_{3'4'}$ 1.32 Hz. The small difference partly stems from the two different parameter sets used in equation (1): for deoxyribose $J_{1'2'}$ and $J_{2'3'}$ are calculated using set D in ref. 1 which was optimized for CH_2-CH fragments whereas these couplings in ribose are calculated using set E in ref. 1, optimized for CH-CH fragments. The coupling constants in 2'-substituted 2'-deoxyarabinose with different substituents can be estimated similarly from the arabino- and deoxy-ribose data presented.

DISCUSSION

β-D-Ribose and β-D-Deoxyribose Sugars.—Equation (1) shows that $^3J_{HH}$ depends on the number of substituents

TABLE 2

Calculated coupling constants (Hz) in deoxyribose as a function of P and Φ_m

	Φ_m 35°					Φ_m 40°				
P(°)	$J_{1'2'}$	$J_{1'2''}$	$J_{2'3'}$	$J_{2''3'}$	$J_{3'4'}$	$J_{1'2'}$	$J_{1'2''}$	$J_{2'3'}$	$J_{2''3'}$	$J_{3'4'}$
0	1.53	7.66	6.95	9.56	7.73	1.34	7.20	6.17	10.27	8.17
9	1.77	8.02	7.01	9.49	8.05	1.53	7.66	6.25	10.20	8.48
18	2.14	8.38	7.21	9.29	8.26	1.87	8.13	6.48	10.00	8.67
27	2.65	8.69	7.51	8.95	8.38	2.39	8.55	6.85	9.66	8.78
36	3.31	8.92	7.90	8.45	8.42	3.10	8.86	7.34	9.14	8.81
54	4.97	8.96	8.79	7.03	8.26	4.97	8.96	8.48	7.58	8.67
72	6.80	8.42	9.49	5.19	7.73	7.06	8.30	9.41	5.48	8.17
90	8.38	7.49	9.60	3.33	6.72	8.81	7.15	9.60	3.33	7.12
108	9.46	6.54	9.01	1.94	5.28	9.93	5.99	8.87	1.79	5.51
126	10.03	5.86	7.94	1.24	3.65	10.47	5.19	7.54	1.15	3.65
144	10.20	5.62	6.83	1.13	2.24	10.61	4.91	6.19	1.22	2.08
153	10.16	5.68	6.38	1.19	1.73	10.58	4.98	5.66	1.38	1.53
162	10.03	5.86	6.04	1.26	1.35	10.47	5.19	5.26	1.53	1.16
171	9.80	6.15	5.83	1.32	1.10	10.26	5.53	5.02	1.64	0.96
180	9.46	6.54	5.76	1.34	0.96	9.93	5.99	4.94	1.68	0.88
189	8.99	6.99	5.83	1.32	0.90	9.45	6.54	5.02	1.64	0.88
198	8.38	7.49	6.04	1.26	0.88	8.81	7.15	5.26	1.53	0.92
216	6.80	8.42	6.83	1.13	0.88	7.06	8.30	6.19	1.22	0.97
234	4.97	8.96	7.94	1.24	0.88	4.97	8.96	7.54	1.15	0.92
252	3.31	8.92	9.01	1.94	0.96	3.10	8.86	8.87	1.79	0.88
270	2.14	8.38	9.60	3.33	1.35	1.87	8.13	9.60	3.33	1.16
288	1.53	7.66	9.49	5.19	2.24	1.34	7.20	9.41	5.48	2.08
306	1.30	7.09	8.79	7.03	3.65	1.22	6.48	8.48	7.58	3.65
324	1.26	6.88	7.90	8.45	5.28	1.22	6.22	7.34	9.14	5.51
342	1.30	7.09	7.21	9.29	6.72	1.22	6.48	6.48	10.00	7.12
351	1.38	7.34	7.01	9.49	7.29	1.25	6.79	6.25	10.20	7.72

TABLE 3

Calculated coupling constants (Hz) in arabinose, lyxose, and xylose as a function of P and Φ_m

	Φ_m 35°					Φ_m 40°				
P(°)	$J_{1'2'}$ A/L [a]	$J_{2'3'}$ A	$J_{2'3'}$ L	$J_{2'3'}$ X	$J_{3'4'}$ L/X	$J_{1'2'}$ A/L	$J_{2'3'}$ A	$J_{2'3'}$ L	$J_{2'3'}$ X	$J_{3'4'}$ L/X
0	6.79	7.28	5.05	0.72	3.47	6.50	8.09	4.33	0.93	2.85
9	7.00	7.21	5.11	0.71	3.03	6.80	8.02	4.40	0.90	2.36
18	7.15	6.98	5.29	0.68	2.71	7.05	7.78	4.62	0.83	2.03
27	7.23	6.61	5.59	0.65	2.52	7.21	7.39	4.96	0.74	1.83
36	7.18	6.09	5.97	0.66	2.46	7.21	6.82	5.42	0.67	1.77
54	6.65	4.67	6.88	0.93	2.71	6.65	5.21	6.56	0.79	2.03
72	5.61	2.98	7.70	1.75	3.47	5.43	3.23	7.60	1.59	2.85
90	4.40	1.47	8.04	3.18	4.70	4.00	1.47	8.04	3.18	4.23
108	3.36	0.58	7.72	4.93	6.17	2.83	0.51	7.62	5.19	5.95
126	2.71	0.41	6.92	6.57	7.46	2.13	0.48	6.60	7.04	7.46
144	2.50	0.71	6.01	7.76	8.19	1.91	1.02	5.47	8.31	8.23
153	2.55	0.92	5.63	8.15	8.29	1.96	1.33	5.01	8.70	8.29
162	2.71	1.10	5.34	8.42	8.25	2.13	1.59	4.66	8.96	8.15
171	2.98	1.22	5.16	8.58	8.10	2.42	1.76	4.45	9.10	7.89
180	3.36	1.26	5.10	8.63	7.89	2.83	1.82	4.38	9.14	7.57
189	3.84	1.22	5.16	8.58	7.67	3.37	1.76	4.45	9.10	7.25
198	4.40	1.10	5.34	8.42	7.48	4.00	1.59	4.66	8.96	6.98
216	5.61	0.71	6.01	7.76	7.32	5.43	1.02	5.47	8.31	6.75
234	6.65	0.41	6.92	6.57	7.48	6.65	0.48	6.60	7.04	6.98
252	7.18	0.58	7.72	4.93	7.89	7.21	0.51	7.62	5.19	7.57
270	7.15	1.47	8.04	3.18	8.25	7.05	1.47	8.04	3.18	8.15
288	6.79	2.98	7.70	1.75	8.19	6.50	3.23	7.60	1.59	8.23
306	6.42	4.67	6.88	0.93	7.46	5.97	5.21	6.56	0.79	7.46
324	6.27	6.09	5.97	0.66	6.17	5.76	6.82	5.42	0.67	5.95
342	6.42	6.98	5.29	0.68	4.70	5.97	7.78	4.62	0.83	4.23
351	6.59	7.21	5.11	0.71	4.04	6.21	8.02	4.40	0.90	3.48

[a] A = arabinose; L = lyxose; X = xylose.

and on the electronegativity and orientation of the substituents relative to the coupled protons. This effect is illustrated in Figure 1 where the cisoidal couplings $J_{2'3'}$ in deoxyribose and ribose and the transoidal coupling $J_{2'3'}$ in arabinose and $J_{2''3'}$ in deoxyribose are plotted as function of the phase angle P. The introduction of an oxygen substituent on C-2′ decreases the J value but this change is highly dependent on the actual value of P. For N-type sugars (P *ca.* 9°) the cisoidal coupling decreases by *ca.* 2 Hz, for an S-type sugar (P *ca.* 162°) this amounts to only 0.5 Hz. Calculations of the 1′-2′ coupling constants in α-xylofuranosylamine and α-2′-deoxyribofuranosylamine, based on the formulation of finite perturbation theory in the INDO approximation of SCF molecular orbital theory,[11] reveal the same trend: the introduction of an oxygen substituent in xylose is predicted to decrease the cisoidal coupling with respect to the corresponding coupling in deoxyribose with *ca.*

1.7 Hz for P 9° and with 0.9 Hz for P 162°. For the transoidal coupling the decrease is 2 Hz for the N-type sugar and for the S-type sugar the influence of an extra oxygen substituent vanishes. Obviously, the older methods for electronegativity corrections (a fixed decrease or a decrease proportional to the electronegativity) are hopelessly inaccurate. In the approach

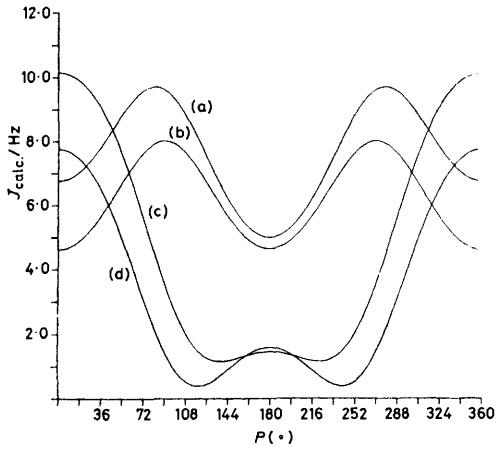


FIGURE 1 $J_{2'3'}$ in deoxyribose (a), ribose (b), and arabinose (d) and $J_{2''3'}$ in deoxyribose (c) as function of the pseudorotation phase angle P (Φ_m 38°)

proposed by Davies and Danyluk [6] the classical Karplus relation $J = A\cos^2\phi + B\cos\phi + C$ with A 11.7, B −0.4, and C 0 Hz is used for a deoxyribose ring. The constants A—C are derived from the observed mean values of $J_{2'3'}$ 6.3 Hz and $J_{1'2'} + J_{3'4'}$ 10.5 Hz under the assumption that a single set of Karplus parameters is valid for the couplings around C-1′–C-2′, C-2′–C-3′, and C-3′–C-4′. This implies that the systematic difference between $J_{3'4'}$ on the one hand and $J_{1'2'}$, $J_{1'2''}$, and $J_{2'3'}$, $J_{2''3'}$ on the other hand that arises from the extra oxygen substituent is averaged over the three couplings. For instance, in the solution conformation of deoxyribonucleoside 3′,5′-cyclic phosphate,[2] described by P 32.1° and Φ_m 43.1°, the torsion angles between H-2″ and H-3′ (161.6°) and between H-3′ and H-4′ (−170.8°) are almost equal. One expects $J_{3'4'}$ to be smaller than $J_{2''3'}$ because of the electronegativity effect of the extra oxygen substituent. The experimental values are $J_{2''3'}$ 10.8 and $J_{3'4'}$ 9.3 Hz. Equation (1) yields 10.0 and 9.0 Hz, respectively. In contrast, from the classical Karplus equation, in conjunction with the Davies–Danyluk parameters, the opposite trend is predicted: $J_{2''3}$ 10.9 and $J_{3'4'}$ 11.8 Hz.

Figure 2 shows the calculated coupling constants in the ribofuranose ring as a function of the pseudorotational phase angle P for different values of the puckering amplitude Φ_m. Note that the dependency on Φ_m differs for cisoidal and transoidal couplings: both in the N-type and in the S-type sugars $J_{2'3'}$ decreases with increasing Φ_m, whereas $J_{1'2'}$ in S-type and $J_{3'4'}$ in N-type increase with increasing Φ_m. The $J_{1'2'}$ value in the N-conformer and $J_{3'4'}$ in the S-conformer are practically independent of Φ_m, because the proton–proton torsion angles in these cases are *ca.* 90° (or 270°). This gives a (flattened) minimum in the Karplus-curve when plotting J *versus*

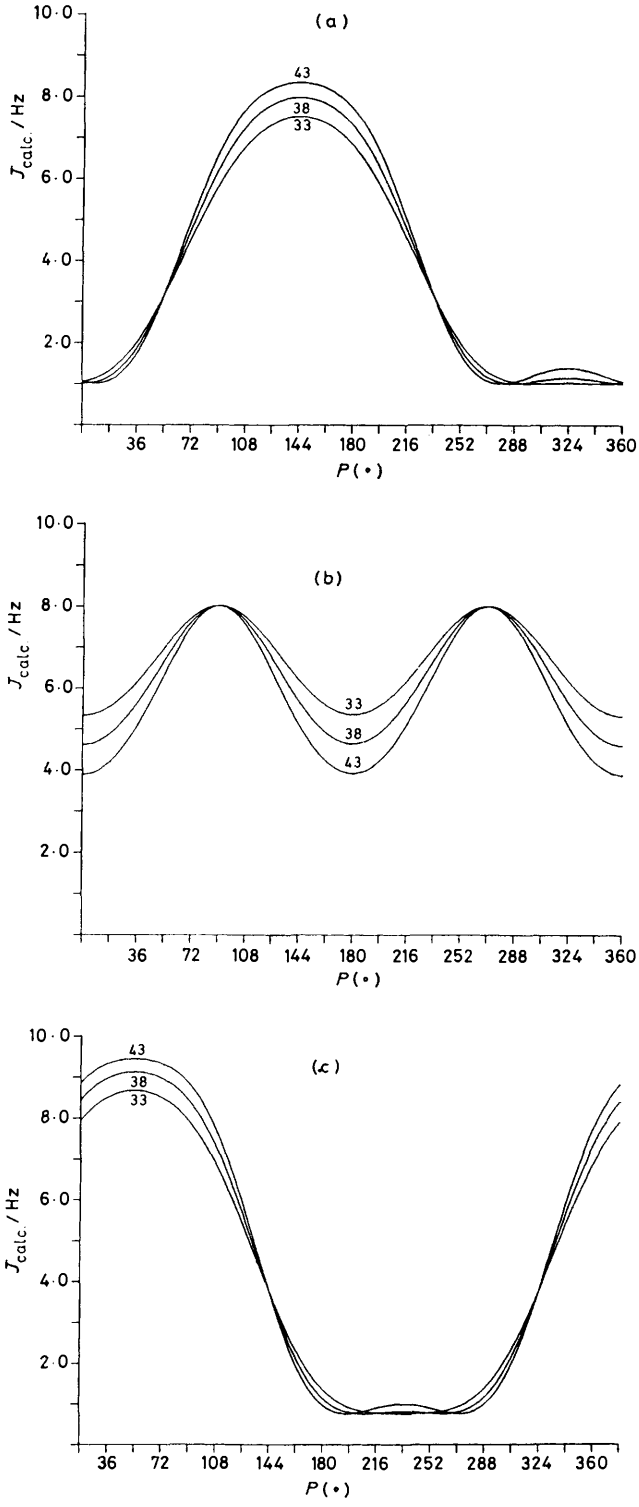


FIGURE 2 Calculated coupling constants $^3J_{HH}$/Hz in ribofuranose as function of the pseudorotation phase angle P. The calculations were carried out for three values of the puckering amplitude, Φ_m 33, 38, and 43°: (a) $J_{1'2'}$, (b) $J_{2'3'}$, (c) $J_{3'4'}$

the proton–proton torsion angle ϕ_{HH}, according to equation (1).

The curves for a deoxyribofuranose (Figure 3) have a similar shape. Comparison of Figures 2c and 3c reveals a systematic difference in the $J_{3'4'}$ coupling for a N-type sugar: in the deoxy-series this coupling is calculated to be smaller by *ca.* 0.5 Hz. This is caused by the β-effect: [1] in a ribofuranose the electronegativity of C-2′ is moderated by the β-oxygen substituent O-2′ and the electronegativity correction term [equation (1)] is smaller than in the case of deoxyribose. In S-type sugar conformations the correction term for the C-2′ substituent vanishes or becomes slightly positive: here the deoxyribose H-3′–H-4′ coupling is calculated to be greater than the ribose coupling. Experimentally a difference in coupling constants for N-type sugars is found: the average $J_{3'4'}$ in 3′,5′-cyclic ribonucleoside (9.8 Hz) appears significantly higher than the average $J_{3'4'}$ in 3′,5′-cyclic deoxyribonucleosides (9.3 Hz), in agreement with the present calculations.

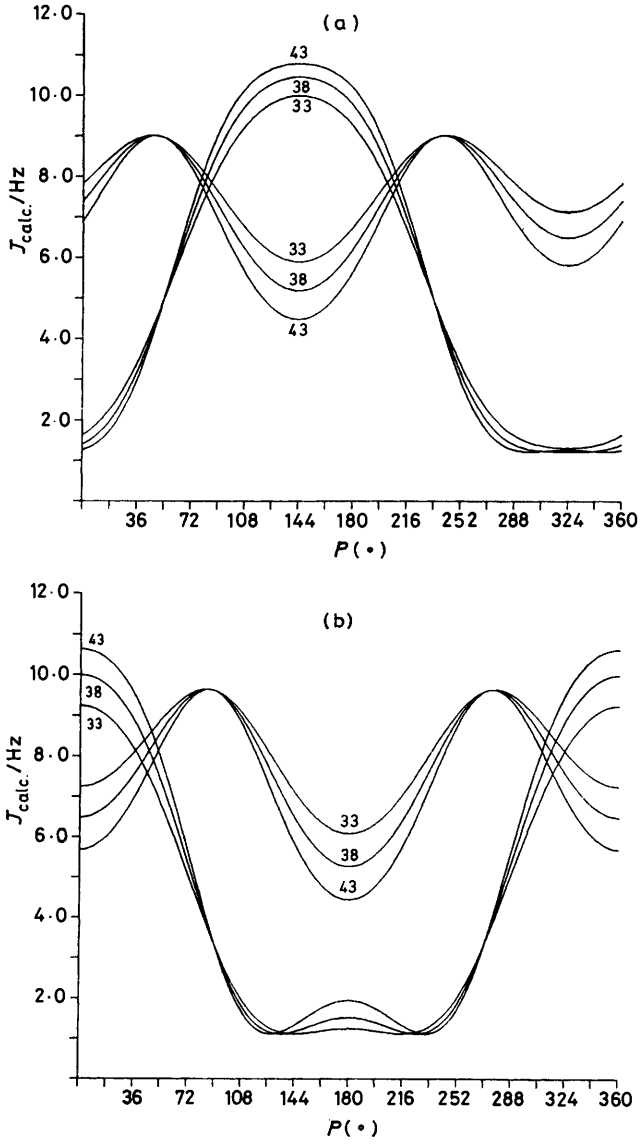


FIGURE 3 Calculated coupling constants $^3J_{HH}$/Hz in deoxyribofuranose as function of the pseudorotation phase angle P. The calculations were carried out for three values of the puckering amplitude, Φ_m 33, 38, and 43°: (a) couplings along C-1′–C-2′, (b) couplings along C-2′–C-3′, (c) $J_{3'4'}$

Statistical analysis of a large number of solid-state data [5] on mono-nucleosides and -nucleotides reveals average values of P_N 9° and P_S 162°. If these values are indicative for the conformational behaviour in solution it can be predicted that the cisoidal coupling $J_{2'3'}$ is only slightly dependent on the position of the N ⇌ S conformational equilibrium, whereas the transoidal couplings $J_{1'2'}$ and $J_{3'4'}$ vary strongly with the equilibrium constant, but in an opposite manner. Altona and Sundaralingam [4] predicted that the sum $J_{1'2'} + J_{3'4'}$ should be practically independent of the magnitude of the equilibrium constant for ribose (and deoxyribose) rings engaged in a conformational equilibrium between N and S geometries, provided these geometries are close to the centres of the normal pseudorotational ranges in terms of P. The effect of flattening or puckering of the sugar ring (Φ_m) on $\Sigma(1'2' + 3'4')$ was also explicitly considered.[4] It is of interest to re-evaluate these predictions in the light of the generalized electronegativity correction embodied in equation (1). Now the full psuedorotational range available to the pentose ring can be explored with confidence. Figure 4 shows the calculated value of $\Sigma(1'2' + 3'4')$ as a function of the phase angle P for different values of the puckering amplitude Φ_m in ribo- and deoxyribo-furanose rings. In accord with the earlier work [4] the Φ_m value is of relatively minor influence on this sum, at least in the normal ranges of P. On the basis of the average solid-state geometry one calculates for the pure N and S ribose conformers a $\Sigma(1'2' + 3'4')$ value of 9.8 ± 0.4 and 9.0 ± 0.5 Hz, respectively. This indicates that, provided the solution geometry is similar to the solid-state geometry, one

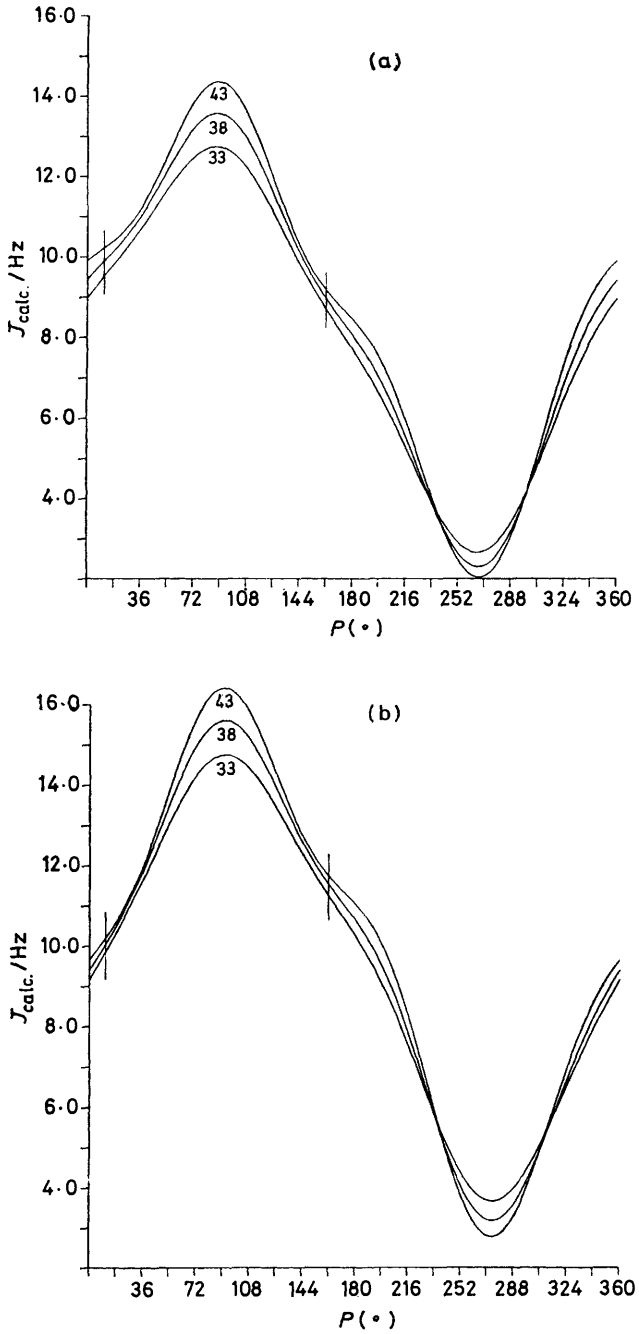


FIGURE 4 Calculated sum $J_{1'2'} + J_{3'4'}$ as function of the pseudorotation phase angle P. The calculations were carried out for three values of the puckering amplitude, Φ_m 33, 38, and 43°: (a) $J_{1'2'} + J_{3'4'}$ in ribofuranose, (b) $J_{1'2'} + J_{3'4'}$ in deoxyribofuranose. Vertical bars indicate centres of pseudorotational ranges found in the solid-state (ref. 5)

again predicts only a slight dependence of this sum on the equilibrium composition. An experimental plot [12,13] of $J_{1'2'}$ *versus* $J_{3'4'}$ in a series of riboses with different N : S ratios, gives a straight line with a slope of −0.86 and an intercept of 9.0 Hz. The calculated slope from the individual coupling constants (N-conformer P 9°, Φ_m 38°, S-conformer P 162°, Φ_m 38°) is −0.89 which agrees well with the experimental value. The calculated intercept is 8.9 Hz. In contrast to the conclusions drawn by Rabczenko *et al.*[12] it now appears that both slope and intercept of the $J_{1'2'}$–$J_{3'4'}$ plot are determined by electronegativity effects on these couplings as well as by the geometric factor (see below).

Figure 4(b) indicates that the predicted range for $\Sigma(1'2' + 3'4')$ in β-D-deoxyribofuranoses is somewhat larger, 9.9 (N-type)—11.5 Hz (S-type). The overall experimental plot of $J_{1'2'}$ *versus* $J_{3'4'}$ in deoxyriboses (Figure 5) gives a slope of −0.95 (0.09) * and an intercept of 10.5 (0.3) Hz, for data derived from monomers, dimers, and trimers.[6,14-21] However, a closer look reveals that

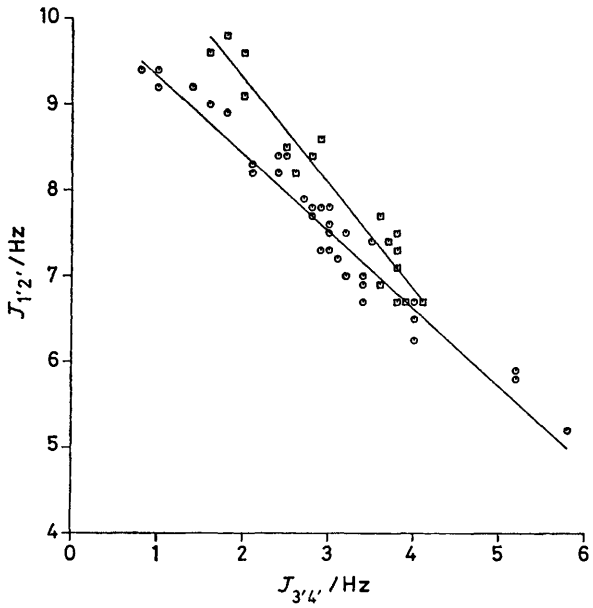


FIGURE 5 Dependence of the coupling constants $J_{1'2'}$ on $J_{3'4'}$ for various deoxyribofuranosylnucleosides. The compounds are subdivided into two classes: ○ monomers and 5'-OH terminal residues, □ 3'-OH terminal residues and central residues

there is a small but consistent difference between monomer and 5'-OH terminal residue data on the one hand and data from central and 3'-OH terminal residues on the other hand. The former gives an intercept of 10.2 (0.2) Hz and a slope of −0.90 (0.06), the latter an intercept of 11.8 (0.5) Hz and a slope of −1.23 (0.15). In a full pseudorotation analysis of dApdA and dApdA-pdA [14] by means of the program PSEUROT (which computes the pseudorotation parameters and mole fraction, given a set of experimental coupling constants) it was concluded that the phase angle P in the S-type conformer of monomers and 5'-OH terminal residues is significantly higher (P *ca.* 166°) than the corresponding value in 3'-OH terminal and central residues (P *ca.* 153°). The calculated values are: slope −1.19, intercept 11.5 Hz for P_N 9°, P_S 166.5°, and Φ_m 38° and slope −1.31, intercept 12.5 Hz for P_N 9°, P_S 154°, and Φ_m 38°, which reproduces the same trend as experiment, although the calculated values are *ca.* 10% too high. Several factors

* 95% Confidence limit in parentheses.

may contribute to this small discrepancy. In our experience, it is often found that $J_{1'2'}$ in the S-type conformer is slightly overestimated. For example, the extrapolated coupling constant $J_{1'2'}$ for 100% stacking in the dAp part of dApdApdA [14] is 9.6 Hz, the calculated value is 9.85 Hz. Moreover, there are indications that the experimental value of $J_{3'4'}$ in the S-type conformer might be smaller than 1 Hz,[14] the minimum value calculated by the use of equation is *ca.* 1.0 Hz. This effect, combined with the overestimation of $J_{1'2'}$ will increase the absolute value of both the calculated slope and intercept.

Elucidation of Solution Geometry.—β-D-*Ribose and* β-D-*deoxyribose.* In the ribofuranose ring only three experimental couplings are available whereas the solution conformation is fully described by five parameters (P_N, Φ_N, P_S, Φ_S, and the mole fraction of N-conformer X_N). This means that, for elucidation of the solution conformation, at least two degrees of freedom must be eliminated. In the graphical method presented by Guschlbauer [7,8] this problem was ' solved ' by (i) the assumption of equal puckering in the N- and S-conformer ($\Phi_N = \Phi_S$) and (ii) the postulate that the phase angles are inter-related by symmetry ($P_S = 180° - P_N$).

In view of the average puckering found in the *X*-ray structures, the first assumption ($\Phi_N = \Phi_S$) appears reasonably justified. Nevertheless, nucleoside 5'-phosphates exhibit a relatively low puckering amplitude in S-type conformers [5'-rNMPs centre around $\langle\Phi_S\rangle = 35 \pm 3°$ (eight *X*-ray structures) and 5'-dNMPs around $\langle\Phi_S\rangle = 34°$ (three structures)] whereas the N-type conformers are more puckered [5'-rNMPs centre around $\langle\Phi_N\rangle = 38 \pm 3°$ (18 structures) and 5'-dNMPs around $\langle\Phi_N\rangle = 38°$ (two structures)].[5] If this assumption is made, the results must be handled with caution.

The second assumption made by Guschlbauer [7,8] is physically untenable. This assumption is only valid for molecules having C_{2v} symmetry, *i.e.* when the N- and S-conformer are exact mirror images, and this is not the case for the ribose ring. For monomeric nucleosides and nucleotides in the solid state the statistical centres of the respective P_N and P_S ranges approximately obey (±8°) the postulated symmetry about *P* 90°, but this is not necessarily true for individual molecules. Indeed, full pseudorotational analyses of some selected nucleotides in solution (see below and ref. 2) indicate that the equality $P_S = 180° - P_N$ is not obeyed. In the case of proline rings the deviations are even more severe.[22]

It must be concluded that it is impossible to elucidate a completely unambiguous description of the conformation equilibrium with a 5 : 3 ratio of parameters and observables. However, a good estimate can be made according to the following procedure. Given that the sum of $J_{1'2'} + J_{3'4'}$ equals 9.4 ± 0.5 Hz, as is observed in the large majority of cases, one may confidently assume that the solution geometry is similar to the solid state geometry. The mole fraction of S-conformer can be reliably estimated then with the aid of Figure 6. In Figure 6 the ranges of $J_{1'2'}$ and $J_{3'4'}$, calculated according to the solid state geometry, as function of the mole fraction of S-conformer are plotted. The puckering amplitude is estimated from the value of $J_{2'3'}$. As can be seen from Table 1 the dependency on Φ_m is expressed by $\delta J/\delta\Phi_m = -0.14$ Hz/degree. A ' high ' value of $J_{2'3'}$ (*e.g.* 5.2 Hz) indicates a low puckering amplitude (Φ_m 35°). If the sum $\Sigma(1'2' + 3'4')$ is significantly smaller than 9.0 Hz this probably indicates a shift of the phase angle *P* to a lower value for the N-conformer and/or to a higher value for the S-conformer rather than a decreased puckering amplitude.

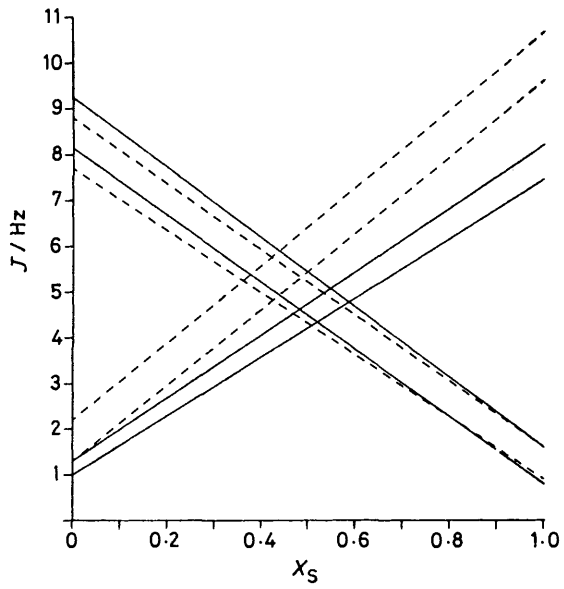


FIGURE 6 Calculated values of $J_{1'2'}$ and $J_{3'4'}$ as function of the mole fraction of S-conformer (X_S) in ribofuranose (solid line) and deoxyribofuranose (dashed line)

In the deoxy-series an analogous reasoning can be employed. The necessary data are given in Figure 6 and in Table 2.

In a recent publication Olsen [23] proposed a three-state model of the furanose pseudorotation. In addition to the principal N and S puckering domains, she included the unusual E (O-4'-*endo*) as a stable intermediate in the analysis. The E-conformer was claimed to be of major importance in various nucleosides and nucleotides in which bulky substituents force the base into unusual *syn*-glycosyl arrangements. Examples are 6-methyl-2'-deoxyuridine and its 3'- and 5'-monophosphate derivatives, for which Olsen [23] calculates an E-population of 31—37% on the basis of observed vicinal couplings and a classical Karplus-type equation. Some of these couplings indeed appear to deviate from the normal pattern of values, which fact prompted George *et al.*[24] to postulate a flattening of the sugar ring (within the framework of a two-state model). In view of the possible biological importance of significant E-populations, discussed by Olsen,[23] we decided to reinvestigate this problem. The coupling constants of the three methylated uridine compounds mentioned above,

determined at eight different temperatures,[24] were subjected to a two-state analysis. The best-fit values obtained for the phase angles are P_N 51.1° and P_S 207.8°, when the puckering amplitude is constrained ($\Phi_N = \Phi_S = 36°$). The overall r.m.s. deviation amounts to only 0.39 Hz which is less than the r.m.s. deviation obtained by Olsen in a three-state analysis (0.52 Hz).[23] Both phase angles are shifted to larger values, but P_S falls within the region found in the solid-state.[5] P_N appears outside the solid-state region. However, only five crystal structures with an N-type deoxyribose ring are known and on statistical grounds there is no reason to reject a P_N value of 51.1°. This phase angle can be rationalized by assuming a pseudorotation shift to a C-4′-*exo* conformer to avoid the steric interaction between the 2-keto oxygen of uracil and C-5′. In our opinion the coupling constants cannot serve as evidence in favour of a stable O-4′-*endo* intermediate.

β-D-*Lyxose*.—In the case of lyxonucleosides a full pseudorotational analysis of published coupling constant data[25] was undertaken. For this analysis it was assumed that all eleven compounds have similar N- and S-type geometries, but that the molar ratio N : S differs from compound to compound. For the sake of simplicity, differences in electronegativity of OH and OCH_3 substituents due to the β-effect were neglected. A pilot calculation, in which all the parameters involved (P_N, Φ_N, P_S, Φ_S and 11 mole fractions) were refined, indicates that the N ⇄ S equilibrium is biased toward the N-conformer. In a second calculation the ill-defined S-conformer was constrained to assume the geometry of lyxofuranosyluracil as found in its crystal structure[25] (P_S 183.5°, Φ_S 37.6°). This resulted, within the limits of error, in the same N-conformer geometry and mole fractions as was found for the unconstrained calculation; the residual r.m.s. deviation was only 0.17 Hz (33 couplings). The results are given in Table 4 and depicted in Figure 7. It is noted that in the O-2′-methylated compounds the equilibrium is further shifted towards the N-conformer than in the O-3′-methylated or non-methylated compounds. This conformational behaviour can be rationalized by assuming an intramolecular hydrogen bond O-3′–H · · · O-5′–H which tends to stabilize the minor S-conformer. From Figure 7 an estimate of the mole fraction in lyxonucleosides can be made.

β-D-*Arabinoses*.—The published data for arabinoses[26] were handled in a similar manner (21 molecules, 63 couplings). No distinct preference for either N- or S-conformer is calculated in this case. The best-fit parameters are P_N 10.9°, Φ_N 39.8°, P_S 158.8°, and Φ_S 30.4° with an overall r.m.s. deviation of 0.17 Hz. (Table 5, set a). The surprisingly low puckering amplitude of the S-conformer is probably an artifact, because, judging from the solid state data,[5] no difference

TABLE 4

Observed [25] and calculated coupling constants (Hz) for lyxonucleosides. Pseudorotation parameters for the N-conformer and the mole fraction were optimized, the pseudorotation parameters for the minor S-conformer were fixed (see text)

	$J_{1'2'}$		$J_{2'3'}$		$J_{3'4'}$		
Compound	exp	calc	exp	calc	exp	calc	X_N
3′-m-lyxo-C,DCl †	5.6	5.6	4.8	4.8	4.6	4.6	0.73
2′-m-lyxo-C,DCl	6.8	6.5	4.5	4.8	2.9	3.5	1.00
2′-m-lyxo-C + 0.08 mmol NaOD	6.7	6.5	4.6	4.8	3.3	3.5	1.00
2′-m-lyxo-C + 0.3 mmol NaOD	6.5	6.5	4.6	4.8	3.6	3.6	0.98
2′-m-lyxo-C + 0.75 mmol NaOD	6.4	6.4	4.7	4.8	3.6	3.6	0.97
3′,5′-m_2-lyxo-C,DCl	5.9	5.9	4.9	4.8	4.3	4.3	0.81
3′,5′-m_2-lyxo-C + 0.3 mmol NaOD	6.4	6.1	4.9	4.8	4.2	4.0	0.88
2′,5′-m_2-lyxo-C	6.7	6.5	4.8	4.8	3.5	3.5	1.00
lyxo-U	6.3	6.3	4.8	4.8	3.8	3.8	0.93
lyxo-C,DCl	6.2	6.3	4.8	4.8	3.7	3.8	0.93
lyxo-C	5.9	6.0	5.0	4.8	4.2	4.2	0.82
	P_N −5.1°				P_S 183.5° *		
	Φ_N 37.2°				Φ_S 37.6° *		

* Constrained value. † m = methyl, C = cytidine, U = uridine.

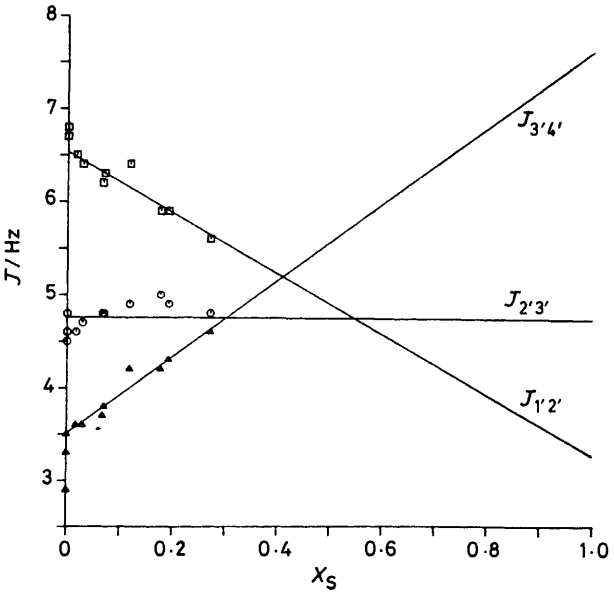


FIGURE 7 Experimental coupling constants as function of the calculated mole fraction of S-conformer in lyxofuranoses. The limiting coupling constants were calculated using P_N −5.1°, Φ_N 37.2° and P_S 183.5°, Φ_S 37.6° and connected by straight lines

in puckering amplitude between N- and S-type is expected. In order to gain insight into the reliability of the calculated geometries further computations were carried out. It was found that the output parameters Φ_N and Φ_S are strongly correlated. For this reason the puckering amplitudes Φ_N and/or Φ_S were constrained to assume a constant value of 36°. The results are shown in Table 5, sets b and c. The changes in P_N, P_S, and in

TABLE 5

Calculated mole fractions of N-conformer in arabinoses. Coupling constant data were taken from ref. 24

Compound	X_N [a]	X_N [b]	X_N [c]
ara C	0.47	0.52	0.50
2′-m-ara C, $(CD_3)_2SO$	0.52	0.57	0.55
3′,5′-m_2-ara C, $(CD_3)_2SO$	0.16	0.19	0.19
2′,3′,5′-m_3-ara C, $(CD_3)_2SO$	0.33	0.37	0.36
ara U	0.53	0.61	0.59
5-amino-ara U, protonated	0.47	0.52	0.50
2′-m-ara A, protonated	0.73	0.79	0.77
2′-m-ara A, neutral, 55 °C	0.69	0.76	0.73
3′-m-ara A, protonated, 50 °C	0.50	0.55	0.54
3′-m-ara A, neutral, 75 °C	0.42	0.47	0.45
3′-m-ara A, pD = 14, 50 °C	0.11	0.13	0.12
3′-m-ara A, $(CD_3)_2SO$	0.36	0.40	0.39
5′-m-ara A, protonated	0.69	0.75	0.73
5′-m-ara A, neutral	0.58	0.63	0.61
5′-m-ara A, neutral, 50 °C	0.58	0.63	0.62
5′-m-ara A, pD = 14	0.62	0.67	0.65
2′,3′-$m_2$ara A, protonated	0.54	0.60	0.58
2′,3′-m_2-ara A, pD = 14	0.48	0.53	0.52
3′,5′-m_2 ara A, neutral, 55 °C	0.42	0.47	0.45
2′,5′-m_2-ara A, neutral, 55 °C	0.65	0.71	0.69
2′,3′,5′-m_3-ara A, neutral, 55 °C	0.46	0.51	0.49

[a] P_N 10.9°, P_S 158.8°, Φ_N 39.8°, Φ_S 30.4°, r.m.s. 0.17 Hz. [b] P_N 10.2°, P_S 156.8°, Φ_N 34.2°, Φ_S 36.0° (constrained value), r.m.s. 0.22 Hz. [c] P_N 16.6°, P_S 159.6°, $\Phi_N = \Phi_S = 36.0°$ (constrained value), r.m.s. 0.22 Hz.

the mole fractions are consistent with the unconstrained approach within a range of 5%. The constrained iterations are almost equally satisfactory with an r.m.s. deviation of only 0.22 Hz. On the basis of the available solid-state information the results of the constrained calculations presumably are the most reliable ones. In Figure 8 the experimental coupling constants are plotted against the calculated mole fraction of the S-conformer. From this Figure a fair estimate of the mole fractions in arabino-nucleosides can be made. A few deviations of individual points from the regression lines are noted. Aside from possible experimental errors these deviations may arise from individual small changes in the detailed geometry of particular molecules from the overall N- and/or S-type geometry deduced for the arabino-sugars. It may well be possible to trace these changes according to the principles laid down above, but for the moment we prefer to abstain from going into further detailed analysis.

Note that no indications whatsoever are found for the three-state conformational model for arabino-nucleosides proposed by Ekiel *et al.*[25] on the basis of a $J_{2'3'}$ *versus* $J_{3'4'}$ plot. Such plots can be quite misleading, however, because many parameters are involved (electronegativity effects, pseudorotation parameters, mole fractions) which should be analysed separately.

Conclusions.—By the combined use of a generalized Karplus equation and the concept of pseudorotation the coupling constants for the full pseudorotation itinerary can be calculated. From these values a good estimate of

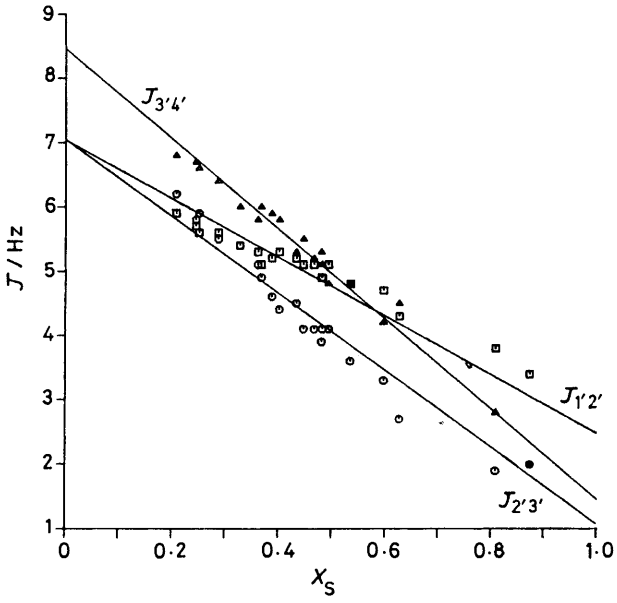


FIGURE 8 Experimental coupling constants as function of the calculated mole fractions (Table 5, case b) of S-conformer in arabinose. The limiting coupling constants were calculated using P_N 10.2°, Φ_N 34.2° and P_S 156.8°, Φ_S 36.0° and connected by straight lines (results from set b, Table 5)

both N- and S-geometry and the equilibrium constant in ribose, deoxyribose, xylose, lyxose, and arabinose derivatives can be made. However, one should remember that a complete description requires five parameters, whereas only three experimental values are available in ribose, lyxose, xylose, and arabinose systems. The judicious use of geometrical information obtained by *X*-ray crystallographic studies therefore appears necessary. In cases where no information from the solid-state is available, the observable coupling constants still allow us to define narrow ranges within conformation space which are compatible with experiment.

For deoxyribose systems a maximum of five vicinal coupling constants is obtained. In this case one has a one-to-one relationship between observables (*i.e.* the proton–proton couplings) and the parameters pertaining to the N $\rightleftharpoons$ S equilibrium. In our experience it is found that the best-fit parameters are quite sensitive towards experimental errors in the coupling constants and it is advisable to constrain one or more parameters to values which are in agreement with information obtained from other sources, for instance from the solid-state.

We thank Drs. C. A. G. Haasnoot and C. S. M. Olsthoorn for allowing us to use their unpublished data.

[1/1318 *Received*, 17*th August*, 1981]

REFERENCES

[1] C. A. G. Haasnoot, F. A. A. M. de Leeuw, and C. Altona, *Tetrahedron*, 1980, **36**, 2783.
[2] C. A. G. Haasnoot, F. A. A. M. de Leeuw, H. P. M. de Leeuw, and C. Altona, *Org. Magn. Reson.*, 1981, **15**, 43.
[3] C. Altona and M. Sundaralingam, *J. Am. Chem. Soc.*, 1972, **94**, 8205.
[4] C. Altona and M. Sundaralingam, *J. Am. Chem. Soc.*, 1973, **95**, 2333.
[5] H. P. M. de Leeuw, C. A. G. Haasnoot, and C. Altona, *Isr. J. Chem.*, 1980, **20**, 108.
[6] D. B. Davies and S. S. Danyluk, *Biochemistry*, 1974, **13**, 4417.
[7] W. Guschlbauer and T.-D. Son, *Nucl. Acids Res. Spec. Publ.*, 1975, **1**, s85.
[8] W. Guschlbauer, *Biochim. Biophys. Acta*, 1980, **610**, 47.
[9] M. L. Huggins, *J. Am. Chem. Soc.*, 1953, **75**, 4123.
[10] C. Altona and C. A. G. Haasnoot, *Org. Magn. Reson.*, 1980, **13**, 417.
[11] A. Jaworski and I. Ekiel, *Int. J. Quantum Chem.*, 1979, **16**, 615.
[12] A. Rabczenko, K. Jankowski, and D. B. Davies, *Studia Biophysica*, 1978, **67**, 33.
[13] D. B. Davies, 'Progress in Nuclear Magnetic Resonance Spectroscopy,' eds. J. W. Emsley, J. Feeney and L. H. Sutcliffe, Pergamon Press, Oxford, 1978, vol. 12, p. 135.
[14] C. S. M. Olsthoorn, L. J. Bostelaar, J. H. van Boom, and C. Altona, *Eur. J. Biochem.*, 1980, **112**, 95.
[15] E. Westhof, H. Plach, I. Cuno, and H.-D. Lüdemann, *Nucl. Acids Res.*, 1977, **4**, 939.
[16] K. N. Slessor and A. S. Tracey, *Carbohydr. Res.*, 1973, **27**, 407.
[17] W. Guschlbauer and K. Jankowski, *Nucl. Acids Res.*, 1980, **8**, 1421.
[18] J. Cadet, R. Ducolomb, and C. Taieb, *Tetrahedron Lett.*, 1975, 3455.
[19] D. J. Wood, K. K. Ogilvie, and F. E. Hruska, *Can. J. Chem.*, 1975, **53**, 2781.
[20] T. Schleich, B. J. Blackburn, R. D. Lapper, and I. C. P. Smith, *Biochemistry*, 1972, **11**, 137.
[21] C. Altona, J. H. van Boom, and C. A. G. Haasnoot, *Eur. J. Biochem.*, 1976, **71**, 557.
[22] C. A. G. Haasnoot, F. A. A. M. de Leeuw, H. P. M. de Leeuw, and C. Altona, *Biopolymers*, 1981, **20**, 1211.
[23] W. K. Olsen, *Nucl. Acids Res.*, 1981, **8**, 1251.
[24] A. L. George, F. E. Hruska, K. K. Ogilvie, and A. Holy, *Can. J. Chem.*, 1978, **56**, 1170.
[25] I. Ekiel, E. Darzynkiewicz, G. I. Birnbaum, and D. Shugar, *J. Am. Chem. Soc.*, 1979, **101**, 4724.
[26] I. Ekiel, M. Remin, E. Darzynkiewicz, and D. Shugar, *Biochim. Biophys. Acta*, 1979, **562**, 177.