ELSEVIER

Carbohydrate Research 307 (1998) 199–209

^{13}C-labeled aldopentoses: detection and quantitation of cyclic and acyclic forms by heteronuclear 1D and 2D NMR spectroscopy

Kenneth N. Drew[a], Jaroslav Zajicek[a], Gail Bondo[b], Bidisha Bose[a], Anthony S. Serianni[a,*]

[a] *Department of Chemistry and Biochemistry, University of Notre Dame, Notre Dame, IN 46556, USA*
[b] *Omicron Biochemicals, Inc., 1347 N. Ironwood Drive, South Bend, IN 46615, USA*

Received 19 September 1997; accepted 20 January 1998

Abstract

^{1}H-Decoupled ^{13}C NMR spectra (150 MHz) of the simple aldopentoses (M solutions in $^{2}H_2O$, 28 °C) selectively labeled with ^{13}C at C-1 (D-(1-^{13}C)arabinose **1**, D-(1-^{13}C)lyxose **2**, D-(1-^{13}C)ribose **3**, D-(1-^{13}C)xylose **4**) contain six enriched C-1 signals that were attributed to four cyclic (α- and β-furanoses and pyranoses) and two acyclic (aldehyde, hydrate) forms. Spectral data were collected and processed in a fashion to permit accurate quantitation of the cyclic and acyclic forms. Percentages of forms varied with pentose structure: α-furanose (0.8–7.4%), β-furanose (0.6–13.2%), α-pyranose (20.2–70.8%), β-pyranose (26.9–62.0%), hydrate (0.063–0.095%), aldehyde (0.009–0.042%). Aldehyde was least abundant in solutions of D-xylose and most abundant in solutions of D-ribose, and the hydrate/aldehyde ratio was higher for D-arabinose, D-lyxose, and D-xylose (6.3–7.8) than for D-ribose (2.1). Heteronuclear 2D HMQC–TOCSY and HCCH–TOCSY spectra were also obtained on several selectively and uniformly ^{13}C-labeled model saccharides, respectively, to evaluate the advantages and limitations of these isotope-aided methods to detect ^{1}H signals of specific forms in solution. These methodologies can be extended to studies of suitably ^{13}C-labeled oligosaccharides and oligonucleotides.

Keywords: ^{13}C-labeled aldopentoses; solution composition; quantitative ID ^{13}C NMR; HMQC–TOCSY; HCCH–TOCSY

1. Introduction

The chemical and biological activities of aldoses are intimately related to their solution behaviors. For example, the reduction of aldoses to alditols with reagents such as $NaBH_4$ probably proceeds via the acyclic aldehyde form in solution, and enzymes involved in carbohydrate metabolism often act on one form of a carbohydrate substrate. For example, the glycolytic enzyme, phosphofructokinase (EC 2.7.1.11), is specific for β-D-fructofuranose 6-phosphate (F6P), whereas the enzyme, D-glucose 6-phosphate dehydrogenase (EC 1.1.1.49), acts specifically on β-D-glucopyranose 6-phosphate [1]. Thus, a knowledge of the solution compositions of aldoses, and more importantly, how these compositions depend on

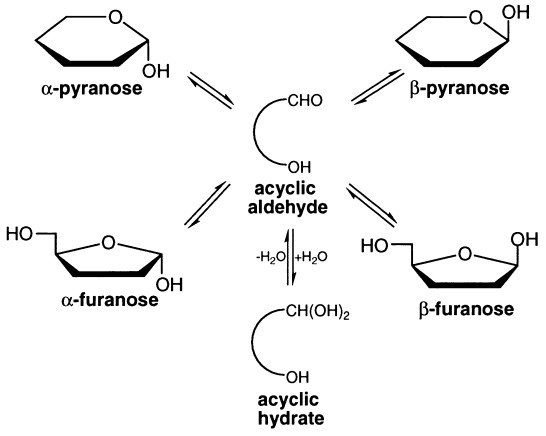
* Corresponding author. Fax: 001 219 631 6924; e-mail: serianni.1@nd.edu

0008-6215/98/$19.00
PII S0008-6215(98)00040-8

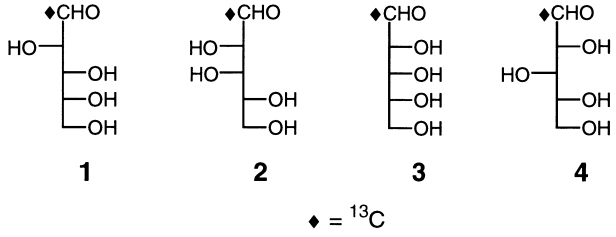
variables such as aldose structure and solvent, is essential to interpreting and predicting carbohydrate reactivity in chemical and biochemical processes.

Simple aldoses exist in aqueous solution in cyclic and/or acyclic forms (Scheme 1), and the relative abundances of these forms depend on structure. For example, aqueous solutions of D-glucose at 31 °C contain primarily α-pyranose (38%) and β-pyranose (62%) [2], whereas those of D-idose at 30 °C contain α-furanose (13.5%), β-furanose (16.5%), α-pyranose (35.9%) and β-pyranose (33.4%) [3a]. These cyclic forms interconvert spontaneously via the acyclic aldehyde [3b] (Scheme 1), which is present at low concentrations for aldoses able to cyclize to furanose and/or pyranose forms (i.e., tetroses, pentoses, hexoses). A second acyclic form, the hydrated aldehyde (hydrate), may also be present [4] (Scheme 1).

While much previous work has focused on the detection and quantitation of the more abundant cyclic forms of aldoses in solution, only a limited number of investigations have reported on the acyclic forms, mainly because their detection is more difficult due to their comparatively low concentrations. To circumvent this problem, selective ^{13}C-labeling of C-1 has proven advantageous to detect and quantify acyclic forms of aldoses in solution [3,5,6], since the resulting ^{13}C NMR spectra are simple and exhibit enhanced signal-to-noise ratios that are required to observe minor forms. In this study, we show that the acyclic forms of the simple aldopentoses (D-(1-^{13}C)arabinose (**1**), D-(1-^{13}C)lyxose (**2**), D-(1-^{13}C)ribose (**3**), D-(1-^{13}C) xylose (**4**)) can be observed and quantified by ^{13}C NMR spectroscopy. In addition, in anticipation of structural studies of ^{13}C-labeled oligosaccharides and oligonucleotides, the advantages and limitations of heteronuclear 2D HMQC–TOCSY [7] and 2D HCCH–TOCSY spectra [8] of selectively and uniformly ^{13}C-labeled model saccharides, respectively, have been examined as a means to detect and assign ^{1}H NMR signals to specific forms in solution.



Scheme 1.

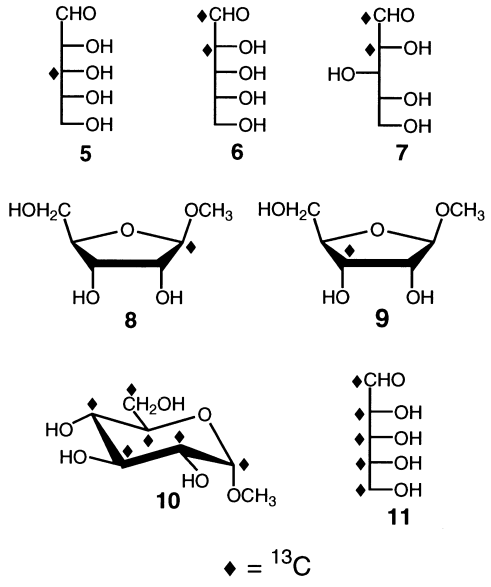


2. Experimental

Preparation of ^{13}C-labeled compounds.—The ^{13}C-labeled compounds used in this investigation were prepared by methods described previously, and thus only a brief account of the synthetic approaches is presented here.

(1-^{13}C)Aldopentoses **1**–**4** were prepared from tetroses (D-erythrose and D-threose) and K^{13}CN by the cyanohydrin reduction method [9,10]. D-(3-^{13}C)Ribose (**5**) was prepared from D-(2-^{13}C)erythrose and KCN via cyanohydrin reduction [9,11a]; D-(2-^{13}C)erythrose was prepared from D-(1-^{13}C)threose [11b] by molybdate-catalyzed epimerization [11c]. D-(1,2-$^{13}C_2$)Ribose (**6**) and D-(1,2-$^{13}C_2$)xylose (**7**) were prepared from D-(1-^{13}C) erythrose and D-(1-^{13}C)threose [11b], respectively, and K^{13}CN via cyanohydrin reduction [9,11a]. Methyl β-D-(1-^{13}C)ribofuranoside (**8**) was prepared from D-(1-^{13}C)ribose (**3**) [9,11a]. Fischer glycosidation of **3** gave a mixture of methyl α- and β-(1-^{13}C)ribofuranosides, which were separated by chromatography on Dowex-1×2 (200–400 mesh) ion-exchange resin (Sigma) in the OH^- form [12,13]; methyl β-D-(3-^{13}C)ribofuranoside (**9**) was prepared from **5** in a similar fashion. Methyl α-D-(UL-$^{13}C_6$)glucopyranoside (**10**) was prepared by Fischer glycosidation of D-(UL-$^{13}C_6$)glucose (Cambridge Isotope Laboratories) [14]; the resulting mixture of glycosides was separated by chromatography on Dowex 1×2 (200–400 mesh) ion-exchange resin in the OH^- form [13,14]. D-(UL-$^{13}C_5$) Ribose (**11**) was prepared from D-(UL-$^{13}C_4$)erythrose and K^{13}CN as described previously [9,11a]; D-(UL-$^{13}C_4$)erythrose was prepared by treatment of

D-(UL-$^{13}C_6$)glucose with $Pb(OAc)_4$ as described by Perlin [15], and purification was achieved by chromatography on Dowex-50×8 (200–400 mesh) ion-exchange resin (Sigma) in the Ca^{+2} form [16].



One-dimensional NMR spectroscopy.—^{1}H-Decoupled ^{13}C NMR spectra were obtained on a Varian Unity*Plus* 600-MHz NMR spectrometer operating at 150.861 MHz for ^{13}C using a 3 mm Nalorac ^{1}H/^{13}C dual microprobe. Spectra were obtained at 28 °C on M solutions of the (1-^{13}C)aldose in ^2H_2O; 10,000–20,000 transients were collected for ^{1}H-decoupled ^{13}C NMR spectra, and 20,000–30,000 transients were collected for ^{1}H-coupled ^{13}C NMR spectra with NOE. The latter data were used to extract $^1J_{C1,H1}$ values in order to confirm the assignments of the acyclic C-1 signals. Previous studies [3–6] have shown that C-1 of the acyclic aldehyde resonates at ~205 ppm and that $^1J_{C1,H1} = \sim 180$ Hz in this form, whereas δ_{C1} and $^1J_{C1,H1}$ in the acyclic hydrate are ~91 ppm and ~163 Hz, respectively.

Two-dimensional NMR spectroscopy.—2D HMQC–TOCSY [7] and 2D HCCH–TOCSY spectra [8] were collected in ^2H_2O solvent at 30 °C on a Varian Unity*Plus* 600-MHz NMR spectrometer. Sample concentrations varied from 0.3–1 M. The spin-lock interval and trim pulse preceding the MLEV-16 spin-lock in HMQC–TOCSY experiments were set at 80 and 2 ms, respectively, and carbon decoupling was applied during the mixing period and acquisition using the GARP decoupling scheme. 2D Gradient-enhanced HCCH–TOCSY spectra [8b] were recorded using a DIPSI-3 mixing period of 22.4 ms for the spin-lock and the GARP decoupling scheme for carbon decoupling. All spectra were obtained using the hypercomplex phase-sensitive method [8c]. Depending on the spectral width in the F1 dimension, 128–320 t_1 increments of 160 scans each were sampled in 512 data points for each increment. The relaxation delay between individual scans was 1.4 s. Linear prediction was applied to the 512 complex data points in the F1 dimension. Zero-filling to 2K and Gaussian weighting functions were applied in both dimensions prior to double-Fourier transformation. Average $^1J_{CH}$ and $^1J_{CC}$ values used to set appropriate time intervals in HMQC–TOCSY and HCCH–TOCSY experiments were 160 and 40 Hz, respectively.

3. Results and discussion

General considerations for quantitative ^{13}C NMR studies.—Quantitation of ^{13}C NMR data is subject to errors caused by differential relaxation and/or NOE effects, and insufficient spectral digitization. These errors were evaluated by obtaining sufficiently digitized ^{1}H-decoupled ^{13}C NMR spectra of **3** (26,178 Hz sweep width, 40–210 ppm) with three recycle times ($3\times T_1$, $5\times T_1$, $10\times T_1$) based on C-1 T_1 values for D-ribose measured by inversion-recovery [17a] (aldehyde, 1.52 s; hydrate, 0.75 s; αf, 1.71 s; βf, 1.52 s; αp, 1.66 s; βp, 1.56 s), and the resulting spectra were processed with two exponential weighting functions (3 and 5 Hz). ^{1}H-Decoupled ^{13}C NMR spectra of **3** were also collected in the presence and absence of NOE using gated ^{1}H-decoupling. Signals in all spectra were either integrated manually or by a lineshape fitting routine supplied with the VNMR spectrometer software. We found that data collected with a minimum recycle time of $5\times T_1$ gave reliable and reproducible quantitative results, that data obtained with and without NOE were virtually identical, that data processed with exponential functions of 3 Hz and 5 Hz were indistinguishable, and that signal integration using the lineshape fitting routine gave more reproducible results, presumably due to the reduction or elimination of subjective factors introduced by manual integration. Thus, quantitative ^{13}C NMR spectra (with NOE) of the remaining three aldopentoses were obtained using a $5\times T_1$ recycle time, FIDs were

processed with 3 or 5 Hz exponential line-broadening depending on the S/N, and signals were integrated using the lineshape fitting routine to obtain the percentages of forms in solution.

While peripheral to the present investigation, it is interesting to note that the C-1 T_1 observed for the acyclic hydrate of **3** (0.75 s) is considerably smaller than that observed for the acyclic aldehyde (1.52 s). Similar behavior was observed previously for DL-(1-^{13}C)glyceraldehyde [17b]. The smaller T_1 for the hydrate may be attributed to its higher molecular weight and perhaps to a more extended conformation in solution, both of which would reduce its rotational correlation time and enhance the rate of relaxation. Other factors, such as C-1–H-1 bond length differences, may also contribute to the observed behavior.

Acyclic forms of the D-*aldopentoses*.—The C-1 signals of the acyclic aldehyde and hydrate forms of **1**–**4** were detected in ^{1}H-decoupled ^{13}C NMR spectra of the (1-^{13}C)-labeled compounds obtained at 150 MHz (Fig. 1), yielding δ_{C1} and $^1J_{C1,H1}$ values for the cyclic and acyclic forms (Table 1) and percentages of each form in solution (Table 2). C-1 chemical shifts for hydrate and aldehyde forms are 91.3 ± 0.3 and 206.2 ± 1.2 ppm, respectively (Fig. 1A,B); these assignments were confirmed by $^1J_{C1,H1}$ values (hydrate, 164.3 ± 0.4 Hz; aldehyde, 180.9 ± 2.7 Hz) (Fig. 1C), which are similar to those reported previously on related structures [3,5].

The abundances of the acyclic forms of **1**–**4** in aqueous solution depend on aldopentose structure. The percentage of aldehyde form decreases in the order **3** > **1** > **2**≈**4**. These percentages differ slightly from those determined previously [2] by circular dichroism at 20 °C (**1**, 0.03%; **2**, 0.03%; **3**, 0.05%; **4**, 0.02%); temperature and solvent (H_2O versus ^2H_2O) may be responsible for the observed differences. The percentage of hydrate form in solution appears relatively constant for **1**–**3** (0.087–0.095%), but appears significantly lower for **4** (0.063%). However, the hydrate/aldehyde ratio is relatively constant for **1**, **2** and **4** (**1**, 6.3; **2**, 7.8; **4**, 6.9), but significantly lower for **3** (2.1); values of 3.2 and ~5 have been reported previously for D-glucose at 37 °C [6] and idose at 30 °C [3a]. Aqueous solutions of ribose **3** contain the highest percentage of acyclic aldehyde, whereas those of xylose **4** contain the lowest.

Furanose forms were also detected in aqueous (^2H_2O) solutions of **1**–**4**; total percentages are ~9% for **1**, ~2% for **2**, ~21% for **3**, and ~2% for **4** at 28 °C (Table 2). Pyranose forms of **1**–**4** have been detected and quantified previously [2,4].

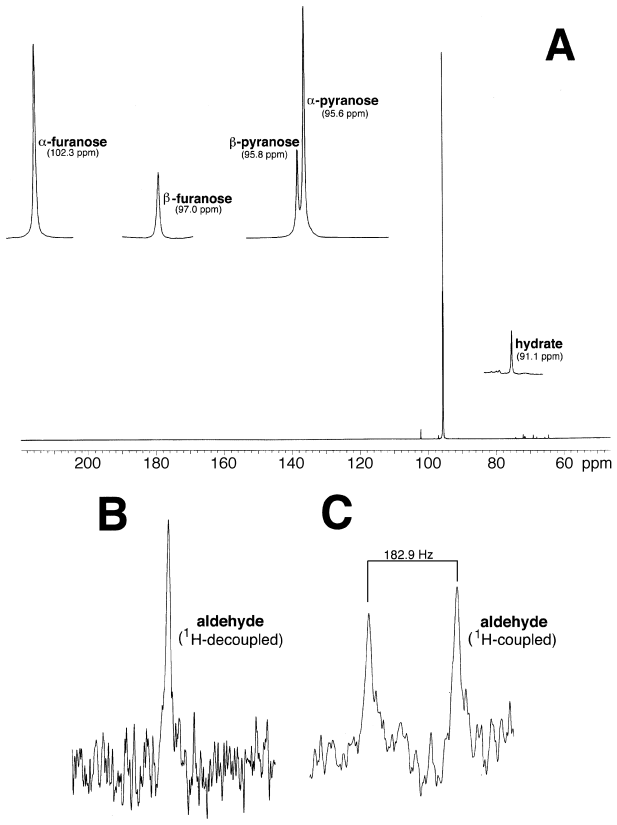


Fig. 1. (A) The ^{1}H-decoupled ^{13}C NMR spectrum (150 MHz) of D-(1-^{13}C)lyxose (**2**) at 28 °C in ^2H_2O showing labeled carbons only. Expanded signals are assigned to C-1 of the α- and β-pyranoses, α- and β-furanoses, and the acyclic hydrate forms (Table 1). (B) The C-1 signal of the acyclic aldehyde form of **2** observed in the ^{1}H-decoupled ^{13}C NMR spectrum. (C) The C-1 signal of the acyclic aldehyde form of **2** observed in the ^{1}H-coupled ^{13}C NMR spectrum showing signal splitting caused by $^1J_{C1,H1}$ (182.9 Hz), which was used to confirm the signal assignment.

Solution composition via isotope-aided NMR methods.—The application of selective ^{13}C-enrichment in the above studies enhances the direct detection of the labeled carbon(s). However, ^{13}C-enrichment may also be exploited to detect *indirectly* the proton signals of the various forms of sugars in solution. We examined this application using several labeled compounds [D-(1-^{13}C)ribose (**3**), D-(3-^{13}C)ribose (**5**), D-(1,2-$^{13}C_2$)ribose (**6**), D-(1,2-$^{13}C_2$)xylose (**7**), methyl β-D-(1-^{13}C)ribofuranoside (**8**), methyl β-D-(3-^{13}C)ribofuranoside (**9**), methyl α-D-(UL-$^{13}C_6$)glucopyranoside (**10**), D-(UL-$^{13}C_5$)ribose (**11**)] and two experimental methods. The first of these, HMQC–TOCSY [7], was applied to selectively ^{13}C-labeled compounds **3** and

Table 1
C-1 Chemical shifts and $^1J_{C1,H1}$ values for the cyclic and acyclic forms of D-aldopentoses **1**–**4**

Compound	NMR parameter	
	δ_{C1} (ppm) [a]	$^1J_{C1,H1}$ (Hz) [b]
D-arabinose (**1**)		
α-furanose	102.6	171.8
β-furanose	96.7	174.2
α-pyranose	98.3	160.6
β-pyranose	94.1	169.0
hydrate	91.6	164.6
aldehyde	207.0	178.2
D-lyxose (**2**)		
α-furanose	102.3	171.4
β-furanose	97.0	obsc [a]
α-pyranose	95.6	167.4
β-pyranose	95.8	161.8
hydrate	91.1	164.2
aldehyde	207.2	182.9
D-ribose (**3**)		
α-furanose	97.8	172.6
β-furanose	102.5	173.0
α-pyranose	95.1	164.6
β-pyranose	95.4	165.4
hydrate	91.0	164.6
aldehyde	204.6	183.4
D-xylose (**4**)		
α-furanose	97.1	obsc [c]
α-furanose	103.2	171.8
α-pyranose	93.8	169.4
β-pyranose	98.2	161.4
hydrate	91.3	163.8
aldehyde	206.0	179.0

[a] ± 0.1 ppm.
[b] ± 0.1 Hz.
[c] Part of signal obscured; *J*-value could not be measured.

Table 2
Percentages of cyclic and acyclic forms of D-aldopentoses **1**–**4** in aqueous solution [a]

Compound	Percent in solution					
	αf	*βf*	*αp*	*βp*	h [c]	a [d]
1 [b]	5.6	3.6	58.5	32.4	0.093	0.016
	5.6	3.6	58.8	32.1	0.090	0.016
	5.6	3.5	58.6	32.3	0.097	0.013
	5.6	3.5	58.8	32.1	0.096	0.013
	(2.5)	(2.0)	(60)	(35.5)		(0.03)
2	1.7	0.59	71.1	26.6	0.091	0.011
	1.7	0.59	70.5	27.2	0.091	0.011
	1.7	0.60	71.1	26.5	0.079	0.011
	1.7	0.60	70.4	27.2	0.084	0.011
	(1.5)	(0.5)	(70)	(28)		(0.03)
3	7.4	13.3	20.3	58.9	0.090	0.040
	7.3	12.9	20.6	59.1	0.087	0.039
	7.6	13.6	19.9	58.9	0.089	0.044
	7.4	13.1	20.0	59.4	0.086	0.043
	(6.5)	(13.5)	(21.5)	(58.5)		(0.05)
4	0.91	0.75	36.8	61.5	0.064	0.011
	0.90	0.74	36.3	62.1	0.069	0.011
	0.81	0.63	36.7	61.8	0.060	0.007
	0.81	0.63	36.1	62.4	0.056	0.007
	(< 1)	(< 1)	(36.5)	(63)		(0.02)

[a] M pentose in ^2H_2O; 28 °C; recycle time, $5\times T_1$.
[b] Two spectra were obtained for each pentose on two separate occasions. Both spectra were processed with a 3 Hz (first entry) and 5 Hz (second entry) line-broadening function. Percentages in parenthesis are those reported previously by Angyal [2], and were determined at 31 and 20 °C for the cyclic and aldehyde forms, respectively.
[c] Hydrate.
[d] Aldehyde.

5–**9**, and the second, HCCH–TOCSY [8], was applied to uniformly ^{13}C-labeled compounds **10** and **11**.

Application of HMQC–TOCSY [7] to selectively ^{13}C-labeled compounds permits the detection of ^{1}H nuclei via $^1J_{CH}$ and $^2J_{HH}/^3J_{HH}$ coupling pathways. For example, in **8**, polarization transfer occurs from H-1 to C-1 and back to H-1 in order to "label" H-1 magnetization (the HMQC component of the experiment), and the resulting magnetization at H-1 is then transferred via the spin-lock portion of the TOCSY component to other protons within the intact spin-system. As observed for conventional 2D TOCSY, the ability to detect all protons within the spin-system using HMQC–TOCSY depends on the magnitudes of the J_{HH} values between the proton pairs, but an additional key factor is the location of the ^{13}C-label(s) within the spin-system. For methyl β-D-(1-^{13}C)ribofuranoside (**8**), all carbon-bound glycone protons are observed (Fig. 2A) at chemical shifts identical to those reported previously [18], although crosspeak intensity decays significantly between H-1 and H-2, and decays further as the nuclei become more remote from H-1 (i.e., H-5*R*, H-5*S*); the aglycone methyl protons are not detected, as expected. Thus, while $^3J_{H1,H2}$ is small (1.2 Hz) in **8** [12,18], its magnitude is apparently sufficient to allow spin propagation throughout the ring. Placement of the ^{13}C-label at a more "internal" position, such as in methyl β-D-(3-^{13}C)ribofuranoside (**9**), allows the detection of strong crosspeaks for H-2, H-3, H-4, H-5*R* and H-5*S* (note the absence of significant signal loss between H-3 and the other protons), whereas the H-1 signal is weak (Fig. 2B). In this case, the extent of crosspeak truncation from H-2 to H-1 caused by the small $^3J_{H1,H2}$ value is much more apparent than in Fig. 2A.

Extension of the HMQC–TOCSY method to D-(3-^{13}C)ribose (**5**) results in correlations at four ^{13}C

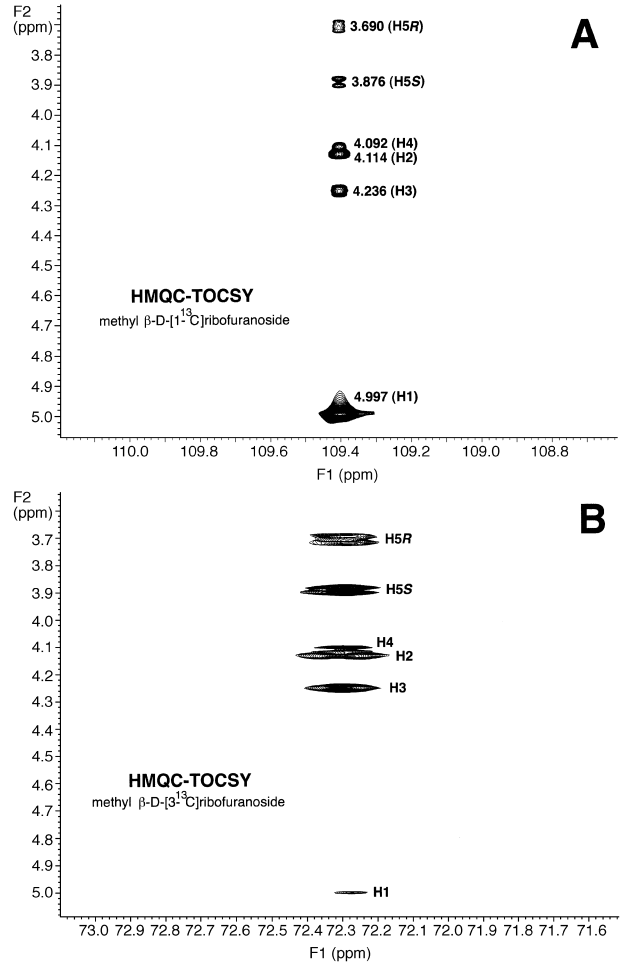


Fig. 2. (A) The 2D HMQC–TOCSY spectrum (600 MHz) of methyl β-D-(1-^{13}C)ribofuranoside **8** (80 ms mixing time) in ^2H_2O showing six crosspeaks at the chemical shift of the labeled carbon (109.4 ppm, F1 dimension). Each ^{1}H signal in **8** (except for the methyl protons) is detected in the F2 dimension; the H-2 and H-4 crosspeaks overlap slightly. (B) The 2D HMQC–TOCSY spectrum (600 MHz) of methyl β-D-(3-^{13}C)ribofuranoside (**9**) (80 ms mixing time) in ^2H_2O showing six crosspeaks at the chemical shift of the labeled carbon (72.28 ppm, F1 dimension). The ^{1}H signals are identical to those observed in (A).

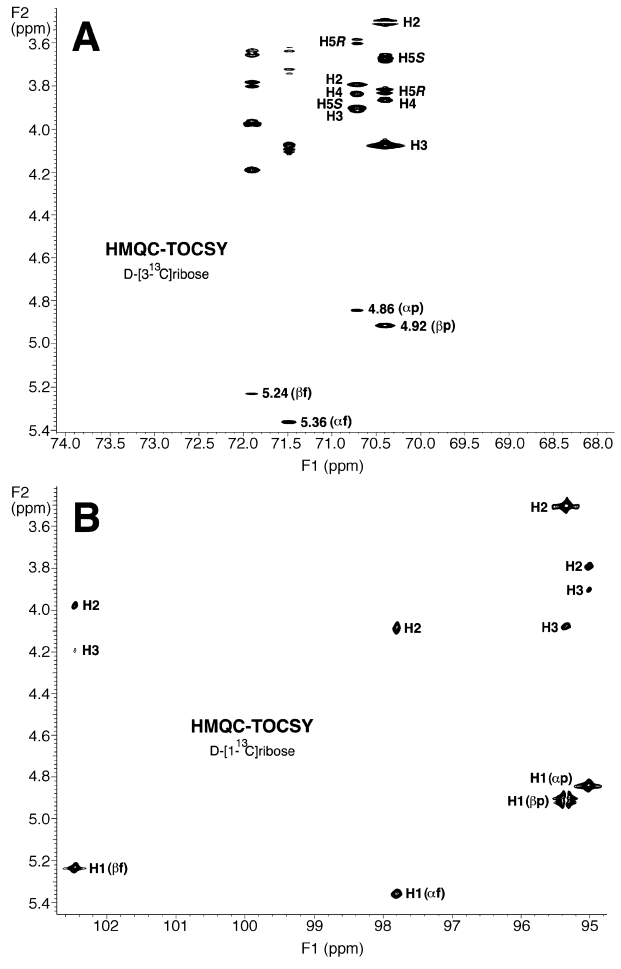


Fig. 3. (A) The 2D HMQC–TOCSY spectrum (600 MHz) of D-(3-^{13}C)ribose (**5**) (80 ms mixing time) in ^2H_2O showing crosspeaks at the C-3 chemical shifts of the four cyclic forms in solution (α-pyranose, 70.7 ppm; β-pyranose, 70.4 ppm; α-furanose, 71.5 ppm; β-furanose, 71.9 ppm). All ^{1}H signals for each form are readily detected (see text for discussion). (B) The 2D HMQC–TOCSY spectrum (600 MHz) of D-(1-^{13}C)ribose (**3**) (80 ms mixing time) in ^2H_2O showing crosspeaks at the C-1 chemical shifts of the four cyclic forms (α-pyranose, 95.0 ppm; β-pyranose, 95.3 ppm; α-furanose, 97.8 ppm; β-furanose, 102.4 ppm). Considerably fewer ^{1}H crosspeaks are detected in comparison to data in (A) (see text for discussion).

chemical shifts assignable to the C-3 carbons of the four cyclic forms of D-ribose in aqueous solution: α-furanose **12** (71.5 ppm), β-furanose **13** (71.9 ppm), α-pyranose **14** (70.7 ppm) and β-pyranose **15** (70.4 ppm) [19] (Fig. 3A). The crosspeak pattern observed for β-D-ribofuranose (**13**) is similar to that observed for **9** (Fig. 2B); similar crosspeak intensities are observed for H-2, H-3, H-4, H-5*R* and H-5*S*, whereas that for H-1 is reduced due to the small $^3J_{H1,H2}$ value in **13** (1.9 Hz). Five crosspeaks are observed for α-D-ribofuranose **12**; the H-2, H-3 and H-4 signals overlap at ~4.1 ppm, and those attributed to H-5*R* (~3.64 ppm) and H-5*S* (~3.74 ppm) are weak. Signal loss from H-2 to H-1 is less evident than in **13** due to the larger $^3J_{H1,H2}$ value in **12** (4.2 Hz). The weak H-5*R*/H-5*S* signals suggest that $^3J_{H3,H4}$ and/or $^3J_{H4,H5R/S}$ values in **12** are smaller than corresponding values in **13**; indeed, $^3J_{H3,H4}$ in methyl α-D-ribofuranoside (3.3 Hz) is smaller than $^3J_{H3,H4}$ in methyl β-D-ribofuranoside (6.9 Hz) [18]. Six relatively intense crosspeaks are observed for β-D-ribopyranose **15** at 4.92, 4.08, 3.87, 3.83, 3.67 and 3.51 ppm, and can be assigned to H-1, H-3, H-4, H-5*R*, H-5*S* and H-2, respectively [20,21]. Six crosspeaks are also

observed for α-D-ribopyranose **14** at 4.86, 3.92, 3.90, 3.84, 3.80, and 3.60 ppm, and can be assigned to H-1, H-3, H-5*S*, H-4, H-2, and H-5*R*, respectively [20,21]. Thus, the HMQC–TOCSY spectrum of D-(3-^{13}C)ribose contains crosspeaks for all carbon-bound protons in the four cyclic forms. In contrast, the HMQC–TOCSY spectrum of D-(1-^{13}C)ribose **3** contains considerably fewer correlations (Fig. 3B); two or three correlations corresponding to H-1, H-2, and H-3 are observed in each of the four cyclic forms. These results further demonstrate the advantages of "internal" labeling for maximal signal detection.

The problem of crosspeak truncation observed in HMQC–TOCSY data due to the presence of small $^3J_{HH}$ values can be circumvented by use of the HCCH–TOCSY method [8]. Here, crosspeak detection is mediated by the large $^1J_{CH}$ and $^1J_{CC}$ values, which range from 140–170 Hz and 35–45 Hz in carbohydrates, respectively [14,19,22]. In order to apply this method, however, the sample must be uniformly ^{13}C-labeled. Consequently, HCCH–TOCSY spectra can be complex, since a full set of proton correlations will be observed for each labeled carbon. For example, the HCCH–TOCSY spectrum of methyl α-D-(UL-$^{13}C_6$)glucopyranoside (**10**) (Fig. 4A) contains signals along the F1 axis for the six labeled carbons. Each carbon shows correlations to each of the ring protons; for example, C-6 (signal at 62.3 ppm) shows correlations to seven protons: H-1 (4.90 ppm), H-2 (3.66 ppm), H-3 (3.76 ppm), H-4 (3.50 ppm), H-5 (3.74 ppm), H-6 (3.97 ppm), H-6′ (3.85 ppm) [14]. The same crosspeaks are observed at C-1–C-5 (although crosspeak intensities vary), thus providing redundant information with respect to ^{1}H signal assignments.

The potential complexity of 2D HCCH–TOCSY spectra is made apparent in studies of D-(UL-$^{13}C_5$)ribose (**11**) (Fig. 4B). The carbon signals at 84.0 (C-4 of **13**) and 84.3 ppm (C-4 of **12**) reveal the same ^{1}H correlations observed using C-3 as the probe nucleus (Fig. 3A), but spectral resolution is reduced. The carbon signal at 76.3 ppm (C-2 of **13**) shows the same correlations as that at 84.0 ppm (C-4 of **13**), as expected. In contrast to these ribofuranose data, the interpretation of the ribopyranose crosspeaks is more challenging. The C-1 signals provide the best resolution (95.4 ppm, **15**; 95.1 ppm, **14**), and show ^{1}H correlations identical to those observed in Fig. 3A. For the ribopyranoses, data redundancy arising from crosspeaks at various

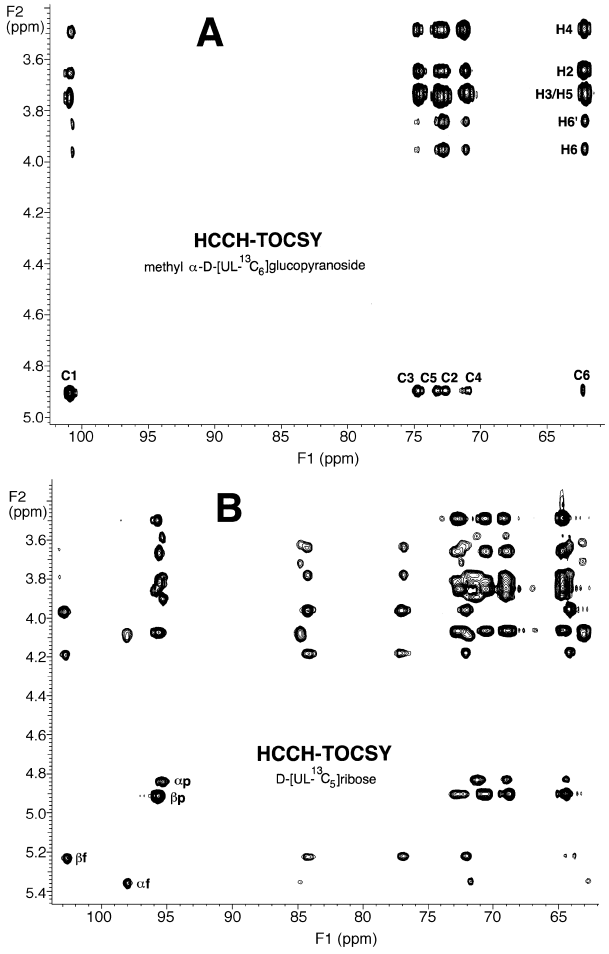


Fig. 4. (A) The 2D HCCH–TOCSY spectrum (600 MHz) of methyl α-D-(UL-$^{13}C_6$)glucopyranoside (**10**) in ^2H_2O. The six labeled carbons (F1 dimension) show identical correlations to seven glycone protons (the H-3 and H-5 signals overlap), although crosspeaks vary in intensity at each carbon. (B) The 2D HCCH–TOCSY spectrum (600 MHz) of D-(UL-$^{13}C_5$)ribose (**11**) in ^2H_2O showing proton correlations with all labeled carbons in the four cyclic forms. Significant crosspeak overlap is observed in the 63–73 ppm region of the spectrum, which complicates the assignment of proton signals to individual forms.

carbons of each form is unavailable due to the significant crosspeak overlap in the 63–73 ppm region of the ^{13}C spectrum. Thus, while the 2D HCCH–TOCSY method is, in principle, more likely to provide a full set of ^{1}H correlations than 2D HMQC–TOCSY (since it is not affected by $^3J_{HH}$ magnitude), the increased data complexity may hinder spectral interpretation, necessitating the implementation of the 3D version of the method to sort out all of the correlations.

An attractive alternative approach giving maximum assignment information yet relatively low complexity is embodied in 2D HMQC–TOCSY

spectra of appropriately multiply ^{13}C-labeled compounds. In this case, several independent and complementary sets of ^{1}H correlations are available for data redundancy, and the likelihood of lost correlations due to the presence of small $^{3}J_{HH}$ values can be reduced by judicious placement of the ^{13}C labels. This approach is illustrated in the HMQC–TOCSY spectrum of D-(1,2-$^{13}C_2$)ribose (**6**) (Fig. 5A), which shows two sets of ^{1}H correlations for the four cyclic forms. For the β-furanose, a complete set of ^{1}H correlations is not detected at the C-1 chemical shift (Fig. 5B); only crosspeaks to H-1, H-2, and H-3 are observed, and the latter two are weak. In contrast, ^{1}H correlations detected at C-2 are complete (Fig. 5C), since the label at C-2 bypasses the $^{3}J_{H1,H2}$ bottleneck. Similar observations are made for the α-furanose, α-pyranose and β-pyranose (Fig. 5A–C). Thus, thoughtful application of multiple ^{13}C-labeling, based on prior knowledge of approximate $^{3}J_{HH}$ values and thus the efficiency of spin-propagation via TOCSY in a given ring configuration, can yield relatively simple 2D HMQC–TOCSY spectra that still contain the desired ^{1}H chemical shift information.

The value of HMQC–TOCSY as applied to multiply ^{13}C-labeled compounds was further examined in a study of the solution composition of D-(1,2-$^{13}C_2$)xylose (**7**). Complete ^{1}H correlations were observed for β-xylopyranose using C-1 as the probe nucleus, whereas a weaker and incomplete set of crosspeaks was observed for α-xylopyranose (Fig. 6A); the latter behavior is attributed to the smaller value of $^{3}J_{H1,H2}$ observed in this anomer (αp, $^{3}J_{H1,H2} = 3.6$ Hz; βp, $^{3}J_{H1,H2} = 7.9$ Hz). In contrast, a complete set of crosspeaks was observed for the α- and β-xylopyranoses using C-2 as the probe nucleus (Fig. 6B), and chemical shift assignments are consistent with those reported by Collins and Ferrier [21] and Wu *et al.* [20]. More importantly, however, ^{1}H correlations for H-1–H-4, and H-1–H-3 were also observed for the α- and β-xylofuranoses, respectively, using C-1 as the probe nucleus, and complete correlations were observed for each furanose at C-2 (Fig. 6C). This result is noteworthy, since these forms are present in low abundance (~1%, Table 2), thus providing a measure of the sensitivity of the method.

In addition to correlations for minor furanose forms, several ^{1}H correlations (H-1–H-3) could be detected in HMQC–TOCSY data at the C-1 chemical shift of the hydrate form of D-(1,2-$^{13}C_2$) xylose (**7**) (91.3 ppm; Table 1, Fig. 6C), thereby

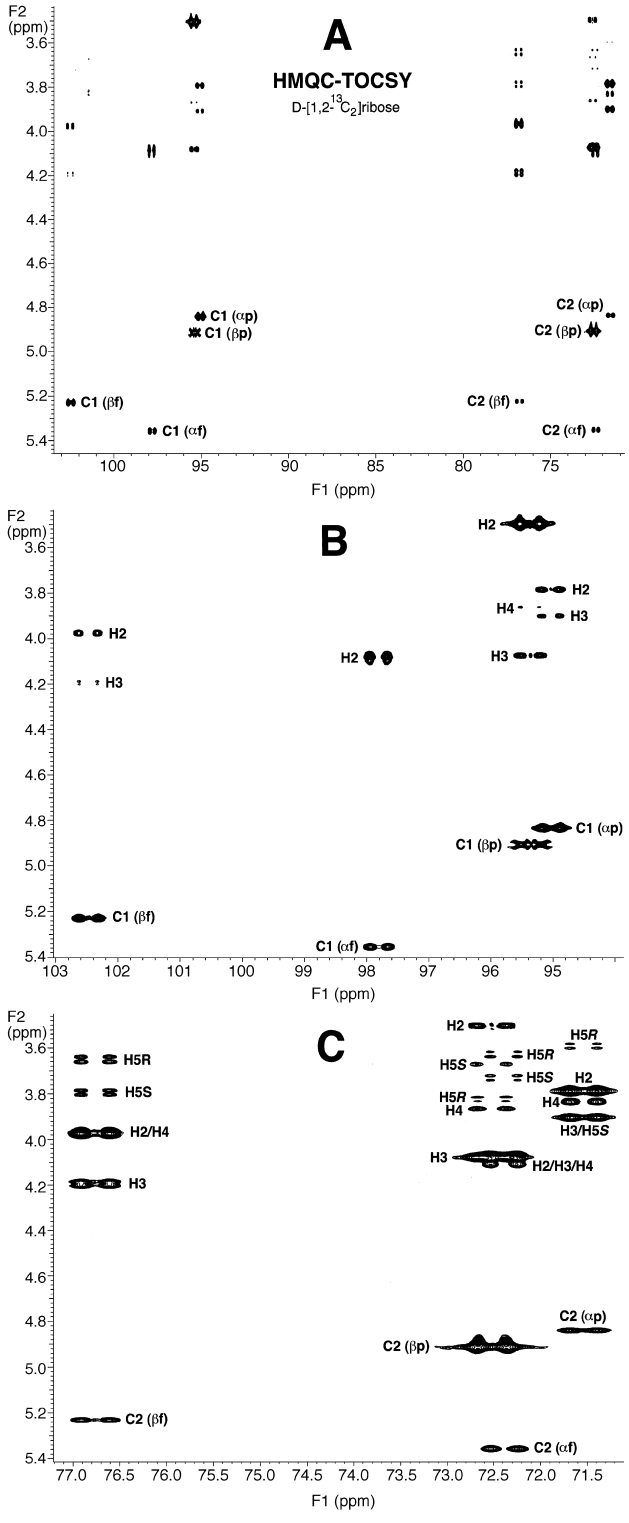


Fig. 5. (A) The 2D HMQC–TOCSY spectrum (600 MHz) of D-(1,2-$^{13}C_2$)ribose (**6**) in $^{2}H_2O$ showing two sets of crosspeaks for each of the four cyclic forms. Crosspeaks appear as doublets in the ^{13}C dimension (F1) due to the presence of $^{1}J_{C1,C2}$ in each form (42.6–47.0 Hz) [19]. (B) Expansion of the C-1 region of (A) showing crosspeak assignments to some of the non-exchangeable protons in the four cyclic forms. (C) Expansion of the C-2 region of (A) showing crosspeak assignments to all of the non-exchangeable protons in the four cyclic forms.

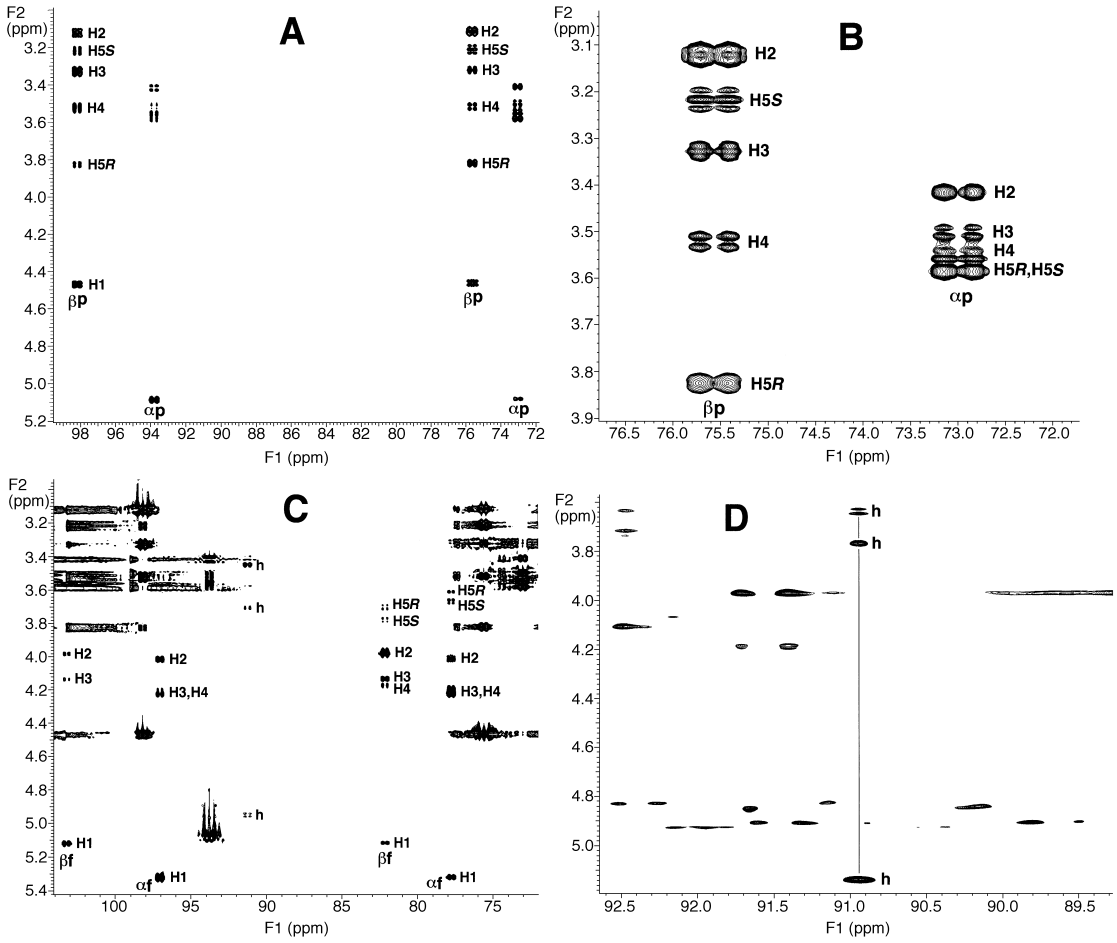


Fig. 6. (A) The 2D HMQC–TOCSY spectrum (600 MHz) of D-(1,2-$^{13}C_2$)xylose (**7**) in ^2H_2O showing two sets of crosspeaks for the predominant α- and β-pyranoses. Crosspeaks appear as doublets in the ^{13}C dimension due to the presence of $^1J_{C1,C2}$ in each form. (B) Expansion of the C-2 region of (A) showing crosspeak assignments to some of the non-exchangeable protons in the α- and β-pyranoses. The H-1 crosspeaks are not plotted. (C) The 2D HMQC–TOCSY spectrum of **7** showing two groups of crosspeaks for the minor α- and β-furanoses (97.1 and 103.2 ppm, respectively), and crosspeaks due to the hydrate form (h, δ_{C1} 91.3). Crosspeaks appear as doublets in the ^{13}C dimension due to the presence of $^1J_{C1,C2}$ in each form. Some spectral noise was purged from the data to facilitate the observation of these very weak crosspeaks. (D) A portion of the 2D HMQC–TOCSY spectrum (600 MHz) of D-(1-^{13}C)ribose **3** in ^2H_2O showing crosspeaks attributed to the hydrate form (h, δ_{C1} 91.0).

providing new ^1H chemical shift information for this very minor form (< 0.1%, Table 2). Likewise, the HMQC–TOCSY spectrum of **3** contains at least four crosspeaks at the C-1 chemical shift of the hydrate form (91.0 ppm, Table 1) (Fig. 6D). In addition to the H-1 signal at ~5.14 ppm, two closely spaced crosspeaks at ~3.63 and ~3.65 ppm, and one crosspeak at ~3.78 ppm, are detected. A very weak crosspeak was also detected at ~3.90 (not shown in Fig. 6D), which is probably associated with the hydrate form. Thus, the presence of a single ^{13}C-label at C-1 in **3** allows at least four hydrate proton signals to be detected, presumably corresponding to H-1–H-4, although only the H-1 signal can be assigned from the available data.

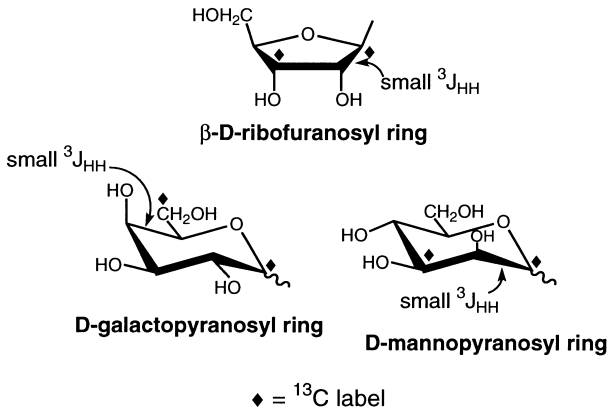


Fig. 7. Double ^{13}C-labeling strategy to circumvent the loss of HMQC–TOCSY crosspeaks caused by the presence of small $^3J_{HH}$ values in some furanosyl and pyranosyl rings (see text for discussion).

4. Summary

This investigation describes the first detection and quantitation of the hydrate forms of the simple D-aldopentoses **1**–**4** in aqueous solution. In addition, while the detection and crude quantitation of the aldehyde form of **3** by ^{13}C NMR spectroscopy have been reported previously [19], the detection and quantitation of the aldehyde forms of **1**, **2** and **4** using this experimental approach have not been documented previously. Since the percentages of these acyclic forms are low (< 0.1%), quantitation by ^{13}C NMR would not be possible without ^{13}C-labeling at C-1. It should be appreciated that Maple and Allerhand [6] previously detected and quantified the acyclic hydrate and aldehyde forms of D-(1-^{13}C)glucose using "ultra-high resolution ^{13}C NMR spectroscopy", and the lowest concentration detected by these investigators was 0.0024% (% aldehyde, 37 °C). In the present study, forms present at levels as low as ~0.009% were detected and quantified reliably using conventional NMR hardware.

In addition to the "direct" detection of ^{13}C nuclei for the assignment of forms in solution, we have shown that ^{13}C-labeled site(s) can also be exploited to assist in the assignment of proton signals in these forms. The effects of ^{13}C-enrichment on COSY [23], INADEQUATE [24], TOCSY [25], and HMBC [26] spectra of carbohydrates and their derivatives, and the wealth of information and editing features that may be derived from the use of such data, have been previously described [12,27–30]. In this report, the advantages and limitations of heteronuclear 2D HMQC–TOCSY and 2D HCCH–TOCSY methods as applied to selectively and uniformly ^{13}C-labeled compounds, respectively, have been examined. It is well known that the ability to detect all protons in furanose and pyranose rings via TOCSY depends on the magnitudes of $^3J_{HH}$ values within these rings, with ring configurations containing small $^3J_{HH}$ values (e.g., $^3J_{H1,H2}$ in β-ribofuranosyl rings; $^3J_{H1,H2}$ in α- and β-mannopyranosyl rings; $^3J_{H4,H5}$ in galactopyranosyl rings) (Fig. 7) being problematic. This limitation, however, can be minimized or eliminated in 2D HMQC–TOCSY spectra by incorporating two or more ^{13}C labels in the molecule, with at least one label on each side of the coupling fragment producing the small $^3J_{HH}$ value (Fig. 7). This approach provides two independent spin-propagation pathways in the molecule with some redundant correlations to permit a full assignment of ^1H signals without introducing unnecessary spectral complexity. The selection of the sites for ^{13}C-labeling will also depend on other considerations, such as ease of synthesis and the desire to extract other structural parameters made available by labeling particular carbons (e.g., J_{CH} and J_{CC} values). Alternatively, uniformly ^{13}C-labeled compounds can be used in conjunction with HCCH–TOCSY, although the 3D version of this method will likely be required, especially in studies of more complex systems such as oligosaccharides [31a] and oligonucleotides [31b], in order to resolve overlapping signals. The choice of experimental approach rests in large part on the relative ease of preparing uniformly versus selectively labeled samples, the anticipated level of complexity of the spectral data to be collected, and the type of information that needs to be extracted from the data. It should appreciated that, while the synthesis of uniformly labeled compounds may appear easier to achieve, the NMR data derived from these samples will be inherently more complex. On the other hand, efficient synthetic methods exist that allow ready access to a wide range of singly and multiply ^{13}C-labeled precursors [10,11,32], and an investment of effort targeting the isotopic labeling may result in a more straightforward and accurate measurement of spectral parameters.

While the present studies were limited to simple monosaccharides, it should be apparent that these methods can be extended to more complex carbohydrates (e.g., oligosaccharides) to permit ^1H signal assignments, extract coupling constants, and/or assess solution conformation; indeed, the spectral complexity exhibited by D-ribose can be likened to that expected from a tetrasaccharide. In the latter regard, these and related heteronuclear 2D methods (e.g., HMQC–NOESY), and their 3D variants, can be used to measure heteronuclear spin-couplings (e.g., J_{CH} values) by taking advantage of crosspeak displacements to obtain both magnitude and sign information [33–35]. It is also conceivable that pattern recognition methods based on the heteronuclear approaches described herein might lead to more efficient automated assignments of oligosaccharide structure. Thus, the integrated use of these methods with appropriate ^{13}C-labeling is expected to promote NMR investigations of the solution properties of carbohydrates and carbohydrate-containing molecules, either free in solution or bound to chemical or biological receptors.

Acknowledgement

This work was supported by a grant from Omicron Biochemicals, Inc., South Bend, IN.

References

[1] S.J. Benkovic and K.J. Schray, *Adv. Enzymol.*, 44 (1976) 139–164.
[2] S.J. Angyal, *Adv. Carbohydr. Chem. Biochem.*, 42 (1984) 15–68.
[3] (a) J.R. Snyder and A.S. Serianni, *J. Org. Chem.*, 51 (1986) 2694–2702. (b) A.S. Serianni, J. Pierce, S.-G. Huang, and R. Barker, *J. Am. Chem. Soc.*, 104 (1982) 4037–4044.
[4] S.J. Angyal, *Adv. Carbohydr. Chem. Biochem.*, 49 (1991) 19–35.
[5] J.R. Snyder and A.S. Serianni, *J. Am. Chem. Soc.*, 111 (1989) 2681–2687.
[6] S.R. Maple and A. Allerhand, *J. Am. Chem. Soc.*, 109 (1987) 3168–3169.
[7] S.S. Wijmenga, K. Hallenga, and C.W. Hilbers, *J. Magn. Reson.*, 84 (1989) 634–642.
[8] (a) S.W. Fesik, H.L. Eaton, E.T. Olejniczak, E. Zuiderweg, L.P. McIntosh, and F.W. Dahlquist, *J. Am. Chem. Soc.*, 112 (1990) 886–888. (b) L.E. Kay, G.Y. Xu, A.U. Singer, D.R. Muhandiram, and J.D. Forman-Kay, *J. Mag. Reson. Ser. B*, 101 (1993) 333–337. (c) D.J. States, R.A. Haberkorn, and D.J. Ruben, *J. Mag. Reson.*, 48 (1982) 286–293.
[9] A.S. Serianni, H.A. Nunez, and R. Barker, *Carbohydr. Res.*, 72 (1979) 71–78.
[10] A.S. Serianni, T. Vuorinen, and P.B. Bondo, *J. Carbohydr. Chem.*, 9 (1990) 513–541.
[11] (a) A.S. Serianni and P.B. Bondo, *J. Biomol. Struct. Dyn.*, 11 (1994) 1133–1148. (b) A.S. Serianni, E.L. Clark, and R. Barker, *Carbohydr. Res.*, 72 (1979) 79–91. (c) M.L. Hayes, N.J. Pennings, A.S. Serianni, and R. Barker, *J. Am. Chem. Soc.*, 104 (1982) 6764–6769.
[12] P.C. Kline and A.S. Serianni, *J. Am. Chem. Soc.*, 112 (1990) 7373–7381.
[13] P.W. Austin, F.E. Hardy, J.G. Buchanan, and J. Baddiley, *J. Chem. Soc.*, (1963) 5350.
[14] C.A. Podlasek, J. Wu, W.A. Stripe, P.B. Bondo, and A.S. Serianni, *J. Am. Chem. Soc.*, 117 (1995) 8635–8644.
[15] A.S. Perlin, *Methods Carbohydr. Chem.*, 1 (1962) 64–66.
[16] S.J. Angyal, G.S. Bethell, and R. Beveridge, *Carbohydr. Res.*, 73 (1979) 9–18.
[17] (a) R.L. Vold, J.S. Waugh, M.P. Klein, and D.E. Phelps, *J. Chem. Phys.*, 48 (1968) 3831–3832. (b) A.S. Serianni, and R. Barker, *J. Magn. Reson.*, 49 (1982) 335–340.
[18] A.S. Serianni and R. Barker, *J. Org. Chem.*, 49 (1984) 3292–3300.
[19] M.J. King-Morris and A.S. Serianni, *J. Am. Chem. Soc.*, 109 (1987) 3501–3508.
[20] G.D. Wu, A.S. Serianni, and R. Barker, *J. Org. Chem.*, 48 (1983) 1750–1757.
[21] P.M. Collins and R.J. Ferrier, *Monosaccharides*, J. Wiley and Sons, New York, 1995, p 534.
[22] J. Wu, P.B. Bondo, T. Vuorinen, and A.S. Serianni, *J. Am. Chem. Soc.*, 114 (1992) 3499–3505.
[23] A. Bax and R. Freeman, *J. Magn. Reson.*, 42 (1981) 164–168.
[24] A. Bax, R. Freeman, and S.P Kempsell, *J. Am. Chem. Soc.*, 102 (1980) 4849–4851.
[25] (a) A. Bax and D.G. Davis, *J. Magn. Reson.*, 65 (1985) 355–360. (b) D.G. Davis and A. Bax, *J. Am. Chem. Soc.*, 107 (1985) 2820–2821.
[26] M.F. Summers, L.G. Marzilli, and A. Bax, *J. Am. Chem. Soc.*, 108 (1986) 4285–4294.
[27] J. Wu and A.S. Serianni, *Carbohydr. Res.*, 226 (1992) 209–219.
[28] J.M. Duker and A.S. Serianni, *Carbohydr. Res.*, 249 (1993) 281–303.
[29] J. Wu and A.S. Serianni, *Biopolymers*, 34 (1994) 1175–1186.
[30] A.S. Serianni and C.A. Podlasek, *Carbohydr. Res.*, 259 (1994) 277–282.
[31] (a) C.T. Weller, J. Lustbader, K. Seshadri, J.M. Brown, C.A. Chadwick, C.E. Kolthoff, S. Ramnarain, S. Pollak, R. Canfield, and S. Homans, *Biochemistry*, 35 (1996) 8815–8823. (b) A. Pardi and E.P. Nikonowicz, *J. Am. Chem. Soc.*, 114 (1992) 9301–9302.
[32] J. Wu, A.S. Serianni, and P.B. Bondo, *Carbohydr. Res.*, 226 (1992) 261–269.
[33] C. Biamonti, C.B. Rios, B.A. Lyons, and G.T. Montelione, *Adv. Biophys. Chem.*, 4 (1994) 51–120.
[34] J.V. Hines, S.M. Landry, G. Varani, and I. Tinoco, Jr., *J. Am. Chem. Soc.*, 116 (1994) 5823–5831.
[35] J.P Marino, H. Schwalbe, S.J. Glaser, and C. Griesinger, *J. Am. Chem. Soc.*, 118 (1996) 4388–4395.