Structures of furanosides: Density functional calculations and high-resolution X-ray and neutron diffraction crystal structures

ARTICLE · JANUARY 1999

CITATIONS
10

READS
17

5 AUTHORS, INCLUDING:

Artem G Evdokimov
HarkerBIO LLC
58 PUBLICATIONS **1,587** CITATIONS
SEE PROFILE

Thomas F. Koetzle
Brookhaven National Laboratory, Upton, NY,…
274 PUBLICATIONS **20,356** CITATIONS
SEE PROFILE

Jan M L Martin
Weizmann Institute of Science
352 PUBLICATIONS **14,330** CITATIONS
SEE PROFILE

Structures of Furanosides: Density Functional Calculations and High-Resolution X-ray and Neutron Diffraction Crystal Structures

Artem G. Evdokimov,[†] A. Joseph Kalb (Gilboa),[†] Thomas F. Koetzle,[‡] Wim T. Klooster,[‡] and Jan M. L. Martin*[,§]

Departments of Structural Biology and Organic Chemistry, Weizmann Institute of Science, P. O. Box 26, IL-76100 Rehovot, Israel, and Department of Chemistry, Brookhaven National Laboratory, Upton, New York 11973

Received: September 18, 1998; In Final Form: November 18, 1998

Highly accurate and precise crystal structures of methyl α-D-arabinofuranoside, methyl β-D-ribofuranoside, methyl α-D-lyxofuranoside, and methyl α-D-xylofuranoside have been determined at 100 K by X-ray crystallography. The structures of methyl α-D-arabinofuranoside and methyl β-D-ribofuranoside have also been determined at 15 K by neutron diffraction. Equilibrium (r_e) geometries of the same compounds were computed by means of density functional methods using a variety of exchange-correlation functionals and a sequence of basis sets. The validity of the computed results was assessed by several criteria including agreement between computed and observed bond distances and bond angles, agreement between computed and observed ring conformations, and basis set convergence of the computed geometrical parameters. Particular reference was made to computed internal hydrogen bond parameters, which are especially sensitive to the quality of the theoretical treatment. Because of the intrinsic sensitivity of the conformation of the five-membered ring to bond lengths and bond angles, molecular mechanics and small basis set SCF treatments are wholly inadequate. Local density functional theory also fails because of a tendency to strongly underestimate internal hydrogen bond distances. When the B3LYP exchange-correlation functional is used, bond lengths and bond angles agree with the neutron diffraction values to within their experimental uncertainty and the ring conformation is qualitatively correct, as long as a basis set of at least double-ζ plus polarization quality (such as cc-pVDZ) is used. Further expansion of the basis set leads to more accurate equilibrium bond lengths and bond angles but does not appreciably affect the ring conformation. For methyl α-D-arabinofuranoside, methyl β-D-ribofuranoside, and methyl α-D-xylofuranoside, there is very good correspondence between the best computed and observed ring conformations, even though some intermolecular hydrogen bonds in the crystal give way to internal hydrogen bonds in the predicted gas-phase structures. On the other hand, in the case of methyl α-D-lyxofuranoside, an O2H···O4 internal hydrogen bond between the ring oxygen O4 and the hydroxyl hydrogen of a ring carbon (O2H) in the computed structure leads to a very large change of ring conformation from the northeast corner of the pseudorotation pathway ($\boldsymbol{P} = 28°$, crystal) to the southeast corner ($\boldsymbol{P} = 130°$, computed).

Introduction

There is considerable interest in determining the conformational states of biologically important five-membered ring sugars.[1] For example, changes in the conformation of ribose and deoxyribose rings are associated with major structural alterations in RNA and DNA and can have a decided influence on the biological functions of these macromolecules. The subject of conformation of five-membered rings in general, and that of the furanosides in particular, is especially intriguing and has been studied by various experimental and theoretical approaches.[2−6] In contrast to the six-membered ring, in which bond lengths, bond angles, and dihedral angles can be maintained at their minimum energy values in a well-defined "strain-free" chair conformation, the five-membered ring is strained in all conformations. As a consequence, a five-membered ring such as cyclopentane fluctuates among a large number of conformational states, each of which represents an uneasy compromise among bond-length, bond-angle, and nonbonded strains. The effect of exocyclic substituents on the conformational behavior of the five-membered ring is thus far more complex than the relatively simple and easily predictable substituent effects on six-membered ring conformation.

A systematic conformational study of the simplest furanosides, the methyl pentofuranosides, can help to provide a clearer understanding of the energetics of the five-membered ring and important insights into the effects of exocyclic substituents on ring conformation. The eight methyl D-pentofuranosides comprise a complete set of all eight combinations of substituent configurations on the furanose ring. Until recently, all structural data for the methyl pentofuranosides were limited to the crystal structure of a single member of this group, methyl α-D-lyxofuranoside, determined in 1968 by X-ray crystallography at room temperature.[7] The X-ray crystallographic structure of methyl β-D-ribofuranoside, also at room temperature, was recently reported as well.[8]

We have determined highly accurate and precise structures of four methyl pentofuranosides: methyl α-D-arabinofuranoside,

[†] Department of Structural Biology, Weizmann Institute of Science.
[‡] Brookhaven National Laboratory.
[§] Department of Organic Chemistry, Weizmann Institute of Science.

10.1021/jp9837840 CCC: $18.00
Published on Web 01/26/1999

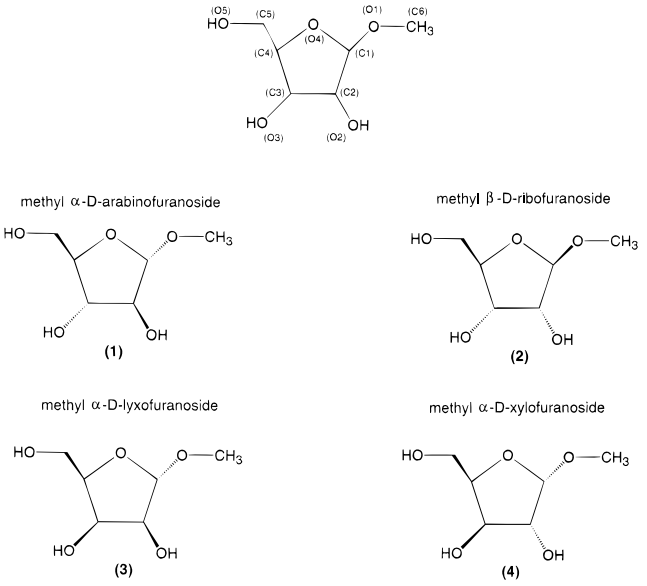

Figure 1. Atom numbering and nomenclature of the four methyl D-pentofuranosides studied.

methyl β-D-ribofuranoside, methyl α-D-lyxofuranoside, and methyl α-D-xylofuranoside (**1**, **2**, **3**, and **4** in Figure 1) by means of cryogenic X-ray and neutron crystallography. The availability of these data and recent advances in density functional theory[9] have prompted us to apply quantum chemical methods to this problem. The highly accurate crystal structures provide reasonable starting coordinates for the calculations. Furthermore, agreement between experimental and calculated values of "hard" structural parameters, such as bond lengths and bond angles, provide an independent, *extrinsic* index of the adequacy of the theory in addition to *intrinsic* verification based on convergence of successively higher levels of theory.

Strictly speaking, nine internal coordinates are necessary to describe the conformation of a five-membered ring. A good approximation can be achieved with a single parameter, the so-called pseudorotational phase angle[1] ***P***. As the five-membered ring accomplishes a complete circuit of its conformational itinerary via all 20 "classical" envelope (E) and twist (T) conformations, ***P*** varies from 0 to 360°. This approximation is especially convenient for simple description and comparison of ring conformations. We have used the definition of ***P*** according to Altona and Sundaralingam.[1]

In this paper, we investigate the suitability of density functional theory (DFT) for predicting ring conformation of the methyl pentofuranosides. It has long been recognized that closure of the five-membered ring imposes extreme sensitivity of the ring dihedral angles to small variations in "hard" parameters such as bond lengths and bond angles. Therefore, only a theoretical treatment that can accurately predict bond lengths and angles can be expected to succeed in predicting correct ring conformation. As we will show, accurate treatment of nonbonded interactions, particularly hydrogen bonding, also appears to be indispensable. It has been amply demonstrated (e.g., ref 10) that modern DFT methods, particularly hybrid methods such as B3LYP (Becke 3-parameter[11]−Lee−Yang−Parr[12]) that include "exact exchange" contributions, are quite successful at predicting "hard" structural parameters such as bond distances and angles. In addition, very recent evidence[13] suggests that these methods may also be quite suitable in the treatment of hydrogen bonding. Here, we have investigated both basis set convergence and the effect of using different forms of the exchange-correlation functional in order to evaluate the degree of sophistication (and the cost in computation time) necessary to deal with the problem of furanoside ring conformation.

Experimental Section

Synthesis and Crystallization of the Methyl Furanosides. Methyl α-D-arabinofuranoside (**1**) was synthesized according to the method of Ness and Fletcher.[14]

Methyl β-D-ribofuranoside (**2**) was purchased from Sigma Chemicals, Inc. and used without further purification.

Methyl α-D-lyxofuranoside (**3**) was synthesized according to the general Pascu thioacetal procedure.[15] D-Lyxose was treated with MeSH in $HCl/ZnCl_2$ to yield D-lyxose methyl thioacetal, which was recrystallized from ethanol. D-lyxose methyl thioacetal was demethylthiolated by the action of $HgO/HgCl_2$ in MeOH, forming **3** in 85% yield (based on thioacetal).

A mixture of anomeric furanosides and pyranosides of D-xylose was synthesized via Fischer methanolysis by allowing 5 g of D-xylose to react with 20 mL of anhydrous methanol containing 20 μL of $SOCl_2$ for 2 h at room temperature. The reaction was stopped by addition of solid $NaHCO_3$, after which the reaction mixture was concentrated in vacuo. The resulting syrup, rich in anomeric furanosides, was taken up in a small volume of water and separated on a column of a basic ion-exchange resin (Dowex 1 × 2) as described in ref 16.

Compounds **1** and **2** were crystallized from ethyl acetate; compounds **3** and **4** were crystallized from acetonitrile.

X-ray Crystallographic Structure Determination. A single crystal protected by a thin film of oil was mounted on a glass fiber and immediately introduced into a stream of liquid nitrogen vapor at a constant temperature of 103 K. In the case of **1**, this was essential since these crystals are unusually hygroscopic.

A Rigaku AFC5R diffractometer, equipped with a rotating molybdenum anode and a graphite monochromator ($\lambda = 0.71073$ Å) was used for data collection. Unit cell parameters were established and refined using the AFC5 software.[17] X-ray data were processed using the Xtal 3.2 package.[18] Crystallographic data are summarized in Table 1.

Structures were solved with SHELXS-86.[19] All non-hydrogen atoms were found directly; hydrogen atoms were introduced according to appropriate peaks in the difference map or, in some cases, geometrically. Structures were refined with SHELXL-92.[20] Structure solution and refinement data are summarized in Table 1.

Neutron Diffraction Structure Determination. All neutron diffraction measurements were carried out in the same general manner. Again, a summary of the crystallographic data is reported in Table 1.

A single crystal was mounted in a halocarbon grease on top of an aluminum pin and immediately sealed in an aluminum container under a helium atmosphere. The container was placed in a DISPLEX model CS-202 closed-cycle refrigerator (APD Cryogenics, Inc.). Following a preliminary examination at 240 K, the crystal was cooled to 15.0 ± 0.5 K, where the temperature was maintained throughout the measurements, and monitored with a Ge resistance thermometer.

Neutron diffraction data were measured on the four-circle diffractometer at beam port H6S of the High Flux Beam Reactor (HFBR) at Brookhaven National Laboratory. The neutron beam, monochromated by germanium (220) planes in transmission geometry, had a wavelength of 1.15942(10) Å as calibrated against KBr; $a_0 = 6.6000$ Å at 298 K.

Unit cell constants were determined from $\sin^2 \theta$ values of 14 and 16 Friedel pairs, respectively, for **1** and **2**. Reflections

TABLE 1: Summary of Essential Crystallographic Data

	X-ray diffraction (103(2) K)				neutron diffraction (15.0(0.5) K)	
compound	**1**	**2**	**3**	**4**	**1**	**2**
space group	$P1$	$P2_12_12_1$	$P2_12_12_1$	$P2_1$	$P1$	$P2_12_12_1$
Z	2	8	4	2	2	8
a (Å)	12.2648(10)	4.7970(10)	4.6770(10)	6.2240(10)	12.248(6)	4.818(1)
b (Å)	6.8312(10)	12.712(3)	10.356(2)	8.142(2)	6.811(2)	12.769(2)
c (Å)	4.601(2)	24.023(5)	15.658(3)	7.4200(10)	4.578(1)	24.199(3)
α (deg)	89.60(3)	90	90	90	89.54(2)	90
β (deg)	100.63(3)	90	90	100.99(3)	100.59(2)	90
γ (deg)	86.67(3)	90	90	90	86.35(2)	90
crystal size (mm^3)	0.004	0.00375	0.002	0.008	0.84	0.243
unique reflections	4019	3418	1690	1953	2014	2244
unique obs reflections[a]	3428	3190	1621	1889	1523	918
θ range	1.69−40.05	1.70−27.66	2.36−27.41	3.09−22.50	4.0−52.5	4.0−52.5
λ	Mo Kα	Mo Kα	Mo Kα	Ag Kα	1.15942(10) Å	1.15942(10) Å
R factor (F_0^2)	0.0630	0.0592	0.0273	0.046	0.116	0.22
R factor[a]	0.0580	0.0487	0.0251	0.047	0.064	0.118
goodness of fit	1.053	1.128	1.094	1.122	1.94[b]	1.51[b]

[a] $F > 2\sigma(F)$ for X-ray, $F_0^2 > 3\sigma(F_0^2)$ for neutron structures. [b] With weights $w^{-1} = \sigma^2 + (0.02F_0^2)^2$.

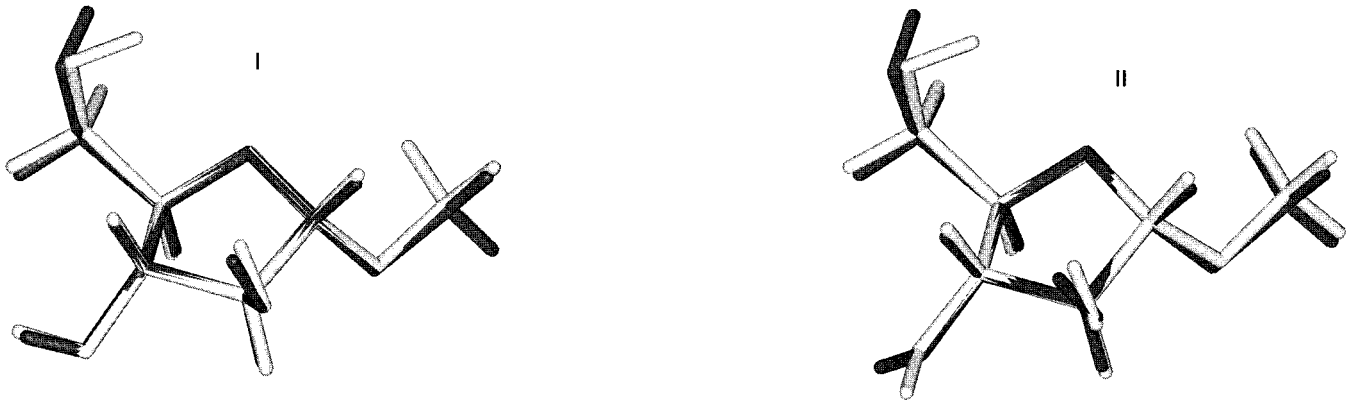

Figure 2. Superimposed (least-squares fit of ring atoms) B3LYP/cc-pVDZ computed (white) and observed X-ray (black) structures of methyl α-D-arabinofuranoside.

were scanned using $\omega-2\theta$ step scans. In the data collection, counts were accumulated at each step for a preset monitor count of the incident beam, amounting to approximately 2.6 and 7.5 s duration respectively for **1** and **2**. The step size was varied to give 65−85 steps per scan in the case of **1** and 60 steps per scan for **2**. Intensities of three reflections were monitored throughout the experiment as a check on experimental stability, which proved to be excellent.

An azimuthal scan of a reflection near $\chi = 90°$ showed maximum intensity variations of 1 and 3% respectively for **1** and **2**, which were within the expected experimental errors.

Integrated intensities, I_0, and variances, $\sigma^2(I_0)$, were derived from the scan profiles as previously described.[21] Lorentz factors were applied, as well as an analytical absorption correction.[22] Averaging over symmetry-related reflections gave internal agreement factors on F^2 of 0.025 and 0.055 respectively for **1** and **2**.

The structural model was refined against all F^2 values using UPALS.[23] The scale factor was varied together with positional and anisotropic displacement parameters for the 46 atoms comprising the two molecules in the asymmetric unit. Attempts to include an extinction correction indicated that this was not significant, and therefore it was omitted. For **2**, a number of atomic displacement tensors (ADT) became nonpositive definite with the introduction of anisotropic displacement parameters. Examination of the individual tensor components revealed that this behavior was caused by the presence of negative values insignificantly different from zero; in any event, the isotropic model should be adequate at 15 K and was therefore applied to **2**. For **1**, anisotropic off-diagonal terms for atom C4 in structure **II** (Figure 2) were set to zero after its ADT became nonpositive definite.

The final difference Fourier map had no residual positive or negative peaks with scattering density exceeding 4% for **1**, and 8% for **2**, of that at the largest carbon atom peak. The relatively high $R(F_0^2)$ and $wR(F_0^2)$ values obtained for **2** result from the high proportion of very weak intensities.

Computational Methods

The Hartree−Fock, MP2 (second-order Møller−Plesset theory[24]), and density functional theory (DFT) calculations were carried out using the GAUSSIAN 94 package[25] running on a DEC Alpha 500/500 workstation and an SGI Origin 2000 minisupercomputer at the Weizmann Institute of Science. The molecular mechanics calculations were carried out using SPARTAN[26] on the DEC Alpha 500/500.

The following exchange-correlation functionals were considered: (a) LDA (local density approximation), which is a combination of the Slater exchange functional[27] and the Vosko−Wilk−Nusair[28] correlation functional; (b) BLYP (Becke−Lee−Yang−Parr), which is a combination of the Becke (1988)[29] gradient-corrected exchange functional with the Lee−Yang−Parr (LYP)[12] correlation functional; (c) BP86 (Becke−Perdew 1986), in which the aforementioned nonlocal exchange functional is combined with Perdew's 1986 correlation functional;[30] (d) the popular B3LYP (Becke 3-parameter−Lee−Yang−Parr)

TABLE 2: Computed and Observed Structural Parameters (Å, deg) for Methyl α-D-Arabinofuranoside Structure I[a]

	MM3	HF/4-21G	HF/cc-pVDZ	B3LYP/4-21G	B3LYP/cc-pVDZ	exptl X-ray[b]	exptl neutron[c]
r(C1−C2)	1.549	1.527	1.538	1.555	1.548	1.544	1.547
r(C2−C3)	1.533	1.538	1.524	1.543	1.536	1.526	1.541
r(C3−C4)	1.525	1.549	1.525	1.539	1.537	1.531	1.535
r(C4−O4)	1.427	1.472	1.414	1.492	1.440	1.431	1.440
r(O4−C1)	1.432	1.436	1.403	1.475	1.431	1.432	1.430
r(C1−O1)	1.420	1.422	1.378	1.432	1.400	1.401	1.398
α(C1,C2,C3)	103.8	100.8	103.5	103.9	103.8	104.0	103.4
α(C2,C3,C4)	100.9	103.6	102.1	104.2	103.0	102.5	102.5
α(C3,C4,O4)	103.0	105.1	102.7	102.1	103.2	103.0	103.4
α(C4,O4,C1)	106.5	109.3	108.8	105.7	107.4	107.2	106.5
α(O4,C1,C2)	106.0	104.6	106.5	106.9	107.0	106.3	107.3
r(O5H···O4)	2.442	2.298	2.398	2.094	2.298	2.927	2.963
τ(C1,C2,C3,C4)	26.0	−35.7	24.4	22.2	20.7	22.4	21.3
τ(C2,C3,C4,O4)	−42.5	19.1	−38.7	−38.8	−36.8	−38.6	−38.1
τ(C3,C4,O4,C1)	43.6	6.5	39.6	41.0	39.8	41.0	41.2
τ(C4,O4,C1,C2)	−26.1	−30.0	−23.6	−27.3	−26.5	−26.4	−27.3
τ(O4,C1,C2,C3)	−1.5	40.5	−2.2	2.5	2.2	1.0	2.4
P	54.11	152.20	52.99	58.00	58.80	57.17	59.03

[a] In this and subsequent tables, *r* represents a bond distance in Å, α a bond angle in deg, and τ a torsion angle in deg. ***P*** is the pseudorotation angle (in deg) as defined in ref 1. [b] Averaged estimated standard deviation (AESD): 0.0037 Å on bond distances, 0.2° on bond angles. [c] AESD: 0.009 Å on bond distances, 0.5° on bond angles.

functional, in which Becke's hybrid local/nonlocal/Hartree−Fock exchange functional[11] is combined with the LYP correlation functional; (e) B3PW91 (Becke 3-parameter−Perdew−Wang 1991), in which the newer Perdew−Wang (1991) exchange functional[31] is substituted for LYP.

The excellent performance of B3LYP for such varied properties as geometries, harmonic frequencies,[10,32] ionization potentials and electron affinities,[33] molecular charge distributions,[34] electrostatic potentials,[35] and reactivity indices[35] such as the Fukui function has previously been demonstrated. For our purposes, it is particularly noteworthy that B3LYP yields qualitatively correct geometries even in problematic cases. For instance, structures for carbon clusters obtained using elaborate conventional ab initio methods[36] were correctly reproduced by B3LYP[37] but not by BP86.[38] (In addition, similar conclusions appear to hold for proton-bound dimers.[39]) Since there is some evidence[35] that the B3PW91 method is somewhat more suitable for the calculation of charge distributions than B3LYP, particularly in strongly polar systems, we consider this functional here since the systems under study contain strongly polar bonds.

Comparison between B3LYP and BLYP on one hand and BLYP (or BP86) and LDA on the other hand, permits assessment of the importance of "exact exchange" effects and of nonlocal corrections, respectively, while a comparison between B3LYP and standard Hartree−Fock calculations permits assessment of the importance of electron correlation.

The following three basis sets were considered: (a) the cc-pVDZ (correlation consistent polarized valence double-ζ) basis set of Dunning,[40] which is a [3s2p1d/2s1p] contraction of a (9s4p1d/4s1p) primitive set; (b) the 4-21G(*) basis set obtained by augmenting the standard 4-21G basis set,[41] which is a [3s2p/2s] contraction of a (7s3p/3s) primitive set with d-type polarization functions on the oxygen atoms, using the same d-exponents as in the cc-pVDZ basis set; (c) the Dunning cc-pVTZ (correlation consistent valence triple-ζ) basis set[40] with the d functions on hydrogen and the f functions on first-row atoms deleted. This basis set, which is a [4s3p2d/3s2p] contraction of a (10s5p2d/5s3p) primitive set, will be denoted TZ2P (for triple-ζ plus two polarization) in this work.

In addition, we evaluated structures obtained using the MM3[42] molecular mechanics force field.

Optimizations were started off from the experimental geometries and carried out using the Schlegel algorithm[43] in redundant internal coordinates.[44] In those cases where a lower-level method yielded a conformation significantly different from that of a higher-level method, a higher-level optimization was performed from that geometry to rule out convergence to different local minima or starting point bias. Harmonic frequency calculations were also carried out at some of the lower levels of theory in order to confirm that the optimized structures obtained were true local minima.

Results and Discussion

1. Methyl α-D-Arabinofuranoside. The crystallographic unit cell contains two distinct structures (labeled **I** and **II**) which differ mainly by the rotation angle around the C3−O3 bond but are otherwise quite similar (Figure 2).

All relevant geometry data for structures **I** and **II** can be found in Tables 2 and 3, respectively.

The X-ray and neutron data for structure **I** agree to within their combined standard uncertainties and show pseudorotation angles of 57.2 and 59.0°, respectively. All computed structures considered give values close to these values, except for HF/4-21G, which at 152.2° represents a completely different structure. The MM3 bond distances are very close to the observed values, as are, to a lesser extent, the bond angles. While the agreement for the bond angles is also fairly good for the B3LYP/4-21G level, bond distances, particularly those involving oxygen, are too long by 0.04−0.06 Å. This problem is remedied by enlarging the basis set to cc-pVDZ, at which level the computed geometry (particularly for the bond distances) agrees surprisingly well with the neutron values; in fact, they are within the standard deviation of the experimental values for most parameters. Considering that the neutron values were measured at 15 K and should be larger than true r_e distances by about 0.01 Å (in fact, they should be close to r_z distances[45]), some error compensation undoubtedly exists between neglect of the $r_z - r_e$ difference (proper account for which would require at least third derivatives of the energy) and the effect of basis set incompleteness, which is expected to make the B3LYP/cc-pVDZ bond distances too long by a similar amount.[10] All the computed structures exhibit an O5H···O4 internal hydrogen bond between the ring oxygen (O4) and the hydroxyl hydrogen (O5H) of the exocyclic carbon C5, which is absent in the experimental structures. It is noteworthy that this interaction does not affect the ring

TABLE 3: Computed and Observed Structural Parameters (Å, deg) for Methyl α-D-Arabinofuranoside Structure II

	MM3	HF/4-21G	HF/cc-pVDZ	B3LYP/4-21G	B3LYP/cc-pVDZ	exptl X-ray[a]	exptl neutron[b]
r(C1−C2)	1.548	1.527	1.536	1.553	1.546	1.542	1.534
r(C2−C3)	1.533	1.538	1.526	1.541	1.538	1.535	1.533
r(C3−C4)	1.524	1.549	1.519	1.528	1.527	1.526	1.534
r(C4−O4)	1.428	1.472	1.416	1.494	1.442	1.428	1.427
r(O4−C1)	1.432	1.436	1.401	1.476	1.429	1.429	1.433
r(C1−O1)	1.420	1.422	1.378	1.430	1.400	1.397	1.396
α(C1,C2,C3)	104.1	100.8	103.6	103.9	103.8	104.1	103.9
α(C2,C3,C4)	101.3	103.6	102.3	103.9	103.1	102.2	102.1
α(C3,C4,O4)	103.2	105.1	102.9	102.1	103.2	104.1	104.1
α(C4,O4,C1)	106.6	109.3	109.6	107.3	108.4	106.5	105.9
α(O4,C1,C2)	105.9	104.6	106.4	106.6	106.8	106.7	107.4
r(O5H···O4)	2.351	2.298	2.375	2.082	2.285	3.005	3.020
τ(C1,C2,C3,C4)	24.1	−35.7	25.1	27.8	22.9	19.5	19.9
τ(C2,C3,C4,O4)	−41.0	19.1	−37.7	−39.6	−36.8	−36.9	−37.5
τ(C3,C4,O4,C1)	43.2	6.6	37.5	37.0	37.7	41.2	41.2
τ(C4,O4,C1,C2)	−27.0	−30.0	−21.1	−19.7	−22.9	−28.2	−28.2
τ(O4,C1,C2,C3)	0.3	40.5	−4.1	−5.4	−1.3	4.0	3.7
P	56.39	152.19	50.01	46.79	53.68	61.49	61.03

[a] Averaged estimated standard deviation (AESD): 0.0037 Å on bond distances, 0.2° on bond angles. [b] AESD: 0.010 Å on bond distances, 0.6° on bond angles.

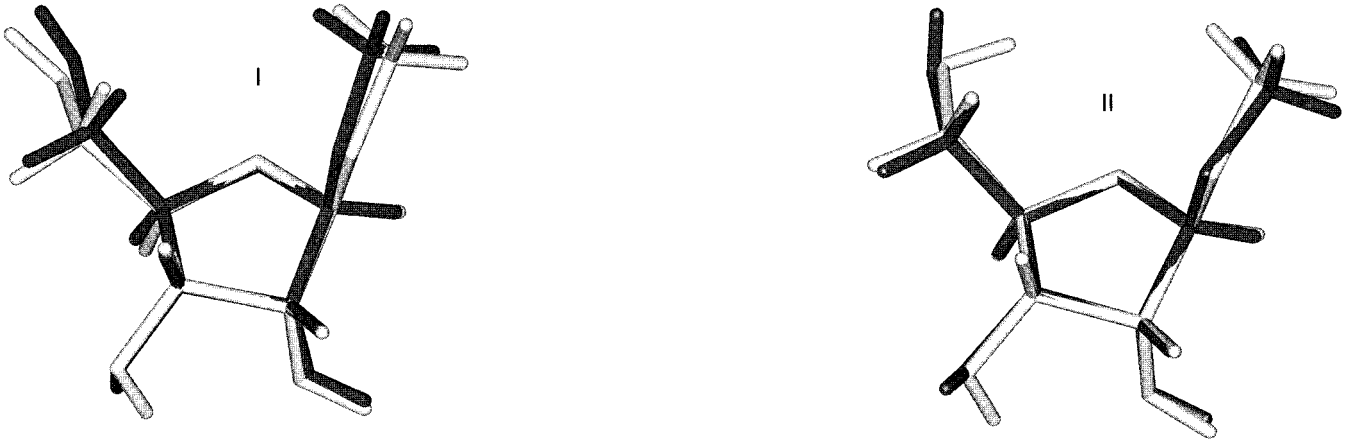

Figure 3. Superimposed (least-squares fit of ring atoms) B3LYP/cc-pVDZ computed (white) and observed X-ray (black) structures of methyl β-D-ribofuranoside.

conformation at all, contrary to the case for methyl-β-D-ribofuranoside, where a similar interaction has a small but significant effect on the ring (see below).

Interestingly, while a basis of at least DZP quality appears to be required at the SCF level for a correct structure, the 4-21G basis set appears to be sufficient at the B3LYP level.

At the B3LYP/cc-pVDZ level, structure **II** is more stable than structure **I** by 1.21 kcal/mol, having a O3H···O2 interaction (3.08 Å) that is absent in structure **I**.

The computed ring conformation displays somewhat greater sensitivity to the level of theory. Otherwise, most of the trends seen for structure **I** are observed here; agreement between the B3LYP/cc-pVDZ and the neutron structure is equally satisfying.

2. Methyl β-D-Ribofuranoside. The crystallographic asymmetric unit contains two distinct structures, labeled **I** and **II** in Figure 3. The relevant computed and experimental data are presented in Tables 4 and 5, respectively.

With the notable exception of *r*(C1−O1), the X-ray and neutron data for structure **I** agree to within the combined uncertainties. In structure **II**, we find additional discrepancies for **r**(C2−C3) and **r**(C3−C4).

Upon optimization at the B3LYP/cc-pVDZ level starting from the two X-ray structures, we obtain the two structures depicted in Figure 3.

Both calculated structures exhibit an internal hydrogen bond O3H···O2; the crucial difference between **I** and **II** is that the latter has an additional O5H···O4 hydrogen bond, which directly involves one of the ring atoms. As we will see below, this has profound consequences for the sensitivity of the computed structure to the level of theory.

The most conspicuous feature about the computed geometries for **II** is the wild variation of the pseudorotation angle with the level of theory. For instance, within the cc-pVDZ basis set, ***P*** varies from −51.2°, using the local density approximation, via −43.6°, for the BP86 functional, to −16.4°, for BLYP, −14.5°, for B3LYP, and −15.5°, for B3PW91. Within the B3LYP functional, it varies from −58.8°, with the 4-21G basis set, via −77.2°, with the 4-21G(*) basis set, to −14.5°, with the cc-pVDZ basis set. Since it appeared by no means certain that the latter value was converged with respect to the basis set, we reoptimized structures **I** and **II** in a still larger (TZ2P) basis set. These calculations took a solid 2 weeks each on the SGI Origin 2000 of the faculty of chemistry. Eventually, this led to a geometry for structure **II** that *quantitatively* differs from its B3LYP/cc-pVDZ counterpart in a manner consistent with experience (particularly, the shortening of bond distances as the basis set is extended) but that *qualitatively* yields the same conformation, $\boldsymbol{P} = -11.8°$.

These variations are very strongly correlated with quite drastic changes in the internal hydrogen bond distances. In particular, the O5H···O4 distance changes from 1.92 Å, at the B3LYP/4-21G level, over 2.35 Å, at the B3LYP/cc-pVDZ level, to 2.44

TABLE 4: Computed and Observed Structural Parameters (Å, deg) for Methyl β-D-Ribofuranoside Structure I

	MM3	HF/ 4-21G	HF/ cc-pVDZ	MP2/ cc-pVDZ	B3LYP/ 4-21G	B3LYP/ cc-pVDZ	B3PW91/ cc-pVDZ	B3LYP/ TZ2P	exptl X-ray[a]	exptl neutron[b]
r(C1−C2)	1.532	1.521	1.521	1.525	1.529	1.530	1.526	1.528	1.506	1.518
r(C2−C3)	1.519	1.529	1.525	1.532	1.544	1.537	1.532	1.534	1.523	1.540
r(C3−C4)	1.529	1.541	1.534	1.545	1.553	1.545	1.540	1.541	1.515	1.530
r(C4−O4)	1.431	1.460	1.418	1.442	1.489	1.442	1.434	1.442	1.444	1.443
r(O4−C1)	1.431	1.427	1.398	1.411	1.453	1.415	1.409	1.416	1.407	1.405
r(C1−O1)	1.420	1.417	1.385	1.409	1.448	1.409	1.404	1.409	1.405	1.433
α(C1,C2,C3)	101.5	101.5	100.9	100.2	101.6	101.4	101.2	101.5	100.8	101.0
α(C2,C3,C4)	101.7	101.7	101.7	101.4	101.6	101.7	101.6	102.1	103.9	103.4
α(C3,C4,O4)	105.1	104.6	105.5	106.9	105.9	106.1	106.2	105.8	106.3	106.2
α(C4,O4,C1)	109.9	110.4	111.5	108.6	108.4	110.4	110.3	110.6	109.1	110.2
α(O4,C1,C2)	106.5	105.6	105.9	105.8	106.4	106.4	106.4	106.4	106.4	106.4
r(O5H···O4)	3.740	3.767	3.731	3.737	3.819	3.776	3.769	3.788	3.582	3.749
r(O3H···O2)	2.035	2.094	2.195	2.110	1.979	2.108	2.090	2.160	2.575	2.534
τ(C1,C2,C3,C4)	37.7	39.3	36.5	37.0	38.9	36.1	36.3	35.7	32.5	32.2
τ(C2,C3,C4,O4)	−35.0	−30.4	−27.3	−22.9	−28.8	−27.4	−27.1	−27.3	−18.4	−19.3
τ(C3,C4,O4,C1)	17.7	9.1	6.1	−2.2	6.8	6.8	6.0	7.2	−4.5	−2.8
τ(C4,O4,C1,C2)	6.8	16.4	18.1	27.0	18.6	17.1	18.2	16.3	26.3	24.4
τ(O4,C1,C2,C3)	−28.4	−35.0	−34.4	−40.5	−36.3	−33.7	−34.5	−32.8	−36.5	−35.1
P	8.56	−5.60	−9.67	−22.34	−9.13	−8.49	−9.91	−7.55	−25.99	−23.45

[a] Averaged estimated standard deviation (AESD): 0.003 Å on bond distances, 0.2° on bond angles. [b] AESD: 0.012 Å on bond distances, 0.7° on bond angles.

TABLE 5: Computed and Observed Structural Parameters (Å, deg) for Methyl β-D-Ribofuranoside Structure II

	MM3	HF/ 4-21G	HF/ cc-pVDZ	MP2/ cc-pVDZ	B3LYP/ 4-21G	B3LYP/ 4-21G(*)	B3LYP/ cc-pVDZ	B3LYP/ TZ2P	exptl X-ray[a]	exptl neutron[b]	LDA/ cc-pVDZ	BLYP/ cc-pVDZ	BP86/ cc-pVDZ	B3PW91/ cc-pVDZ
r(C1−C2)	1.532	1.520	1.521	1.525	1.527	1.538	1.530	1.528	1.527	1.520	1.505	1.542	1.528	1.525
r(C2−C3)	1.519	1.533	1.527	1.536	1.575	1.587	1.540	1.536	1.511	1.539	1.542	1.555	1.557	1.535
r(C3−C4)	1.528	1.538	1.533	1.544	1.556	1.554	1.544	1.539	1.516	1.539	1.538	1.557	1.559	1.539
r(C4−O4)	1.432	1.470	1.425	1.452	1.503	1.475	1.451	1.450	1.435	1.453	1.453	1.473	1.479	1.443
r(O4−C1)	1.432	1.428	1.394	1.414	1.459	1.439	1.418	1.419	1.432	1.441	1.422	1.437	1.440	1.412
r(C1−O1)	1.420	1.415	1.383	1.407	1.439	1.421	1.407	1.406	1.384	1.413	1.378	1.426	1.411	1.401
α(C1,C2,C3)	101.5	101.9	101.2	100.7	104.0	104.2	101.7	101.8	101.6	102.1	102.2	102.0	102.2	101.5
α(C2,C3,C4)	101.5	101.9	101.7	101.7	103.7	103.5	102.0	102.3	102.4	101.4	101.7	102.2	102.2	101.8
α(C3,C4,O4)	105.5	104.9	105.5	107.0	106.6	106.5	106.3	106.0	106.4	106.0	106.5	106.5	106.6	106.4
α(C4,O4,C1)	109.6	110.3	111.4	108.4	106.3	106.8	110.3	110.5	109.8	110.1	106.1	109.8	107.4	110.1
α(O4,C1,C2)	106.6	105.2	105.7	105.2	104.4	105.7	106.0	106.0	104.6	104.6	101.6	106.0	103.0	105.9
r(O5H···O4)	2.403	2.273	2.420	2.256	1.917	1.967	2.351	2.439	3.288	3.389	1.931	2.351	2.203	2.330
r(O3H···O2)	2.037	2.074	2.190	2.085	1.808	1.868	2.091	2.153	3.262	3.398	1.870	2.079	2.024	2.074
τ(C1,C2,C3,C4)	37.6	38.6	36.5	35.9	20.2	8.3	35.2	35.3	37.5	38.0	29.2	34.5	31.1	35.5
τ(C2,C3,C4,O4)	−34.8	−27.9	−26.9	−19.9	2.3	13.5	−24.4	−25.4	−26.6	−26.5	−3.6	−23.1	−9.1	−24.3
τ(C3,C4,O4,C1)	17.4	5.9	5.5	−5.8	−25.4	−32.3	2.9	4.5	4.1	4.0	−25.4	1.6	−18.2	2.3
τ(C4,O4,C1,C2)	7.3	19.1	18.7	29.7	38.7	37.9	20.3	18.7	20.1	20.7	44.4	21.1	38.6	21.2
τ(O4,C1,C2,C3)	−28.7	−36.1	−34.7	−41.3	−36.4	−28.0	−34.9	−33.9	−35.9	−36.7	−45.6	−35.0	−43.3	−35.7
P	7.96	−10.25	−10.55	−27.24	−58.84	−77.16	−14.46	−11.82	−12.40	−12.96	−51.16	−16.45	−43.56	−15.45

[a] Averaged estimated standard deviation (AESD): 0.003 Å on bond distances, 0.2° on bond angles. [b] AESD: 0.012 Å on bond distances, 0.7° on bond angles.

Å, at the B3LYP/TZ2P level, a change of more than one-half of an angström! As a matter of fact, the B3LYP/4-21G value is so "uncomfortably" short that it forces the whole structure to distort (Figure 4), leading to the dramatic change observed in ***P***.

The changes in the O3H···O2 hydrogen bond distance, while not as spectacular, are still quite large, from 1.81 Å, at the B3LYP/4-21G level, over 2.09 Å, at the B3LYP/cc-pVDZ level, to 2.15 Å, at the B3LYP/TZ2P level. Since these changes do not directly involve a ring atom, however, their impact on the overall shape of the molecule is not as dramatic. This conclusion is borne out by comparison with structure **I**, which has only an O3H···O2 hydrogen bond but not its O5H···O4 counterpart. While the variation in the O3H···O2 bond distance is almost as strong as that for structure **II** (from 1.98 Å at the B3LYP/4-21G level to 2.16 Å at the B3LYP/TZ2P level), the pseudorotation angle ***P*** varies by less than 2°.

2.1. A Brief Digression: The Water Dimer as a Model System for Hydrogen Bonding. Since the quality of the treatment of hydrogen bonds appears to be such a crucial factor in the predicted geometry for **II**, we have investigated the performance of the levels of theory used in the present study for the water dimer, perhaps the most extensively investigated prototype hydrogen-bonded system in the literature. Theoretical studies of this system have been reviewed by Scheiner;[46] very recently, Handy and co-workers[13] computed a complete potential surface for the water dimer using an adaptation of the B3PW91 functional in conjunction with the cc-pVTZ basis set. Some recent ab initio calibration studies of this system can be found in refs 47 and 48.

Because of the well-known importance of basis set superposition error (BSSE) for weak molecular interactions,[49] the interaction energy was computed both with and without the counterpoise (CP) correction.[50] The results are presented in Table 6.

The most accurate estimates for the water dimer interaction energy are 5.4 ± 0.7 kcal/mol experimentally[51] and 5.0 ± 0.1 kcal/mol from an exhaustive ab initio convergence study by Halkier et al.[48] using coupled cluster methods and basis sets with as many as 574 basis functions. With the largest basis set

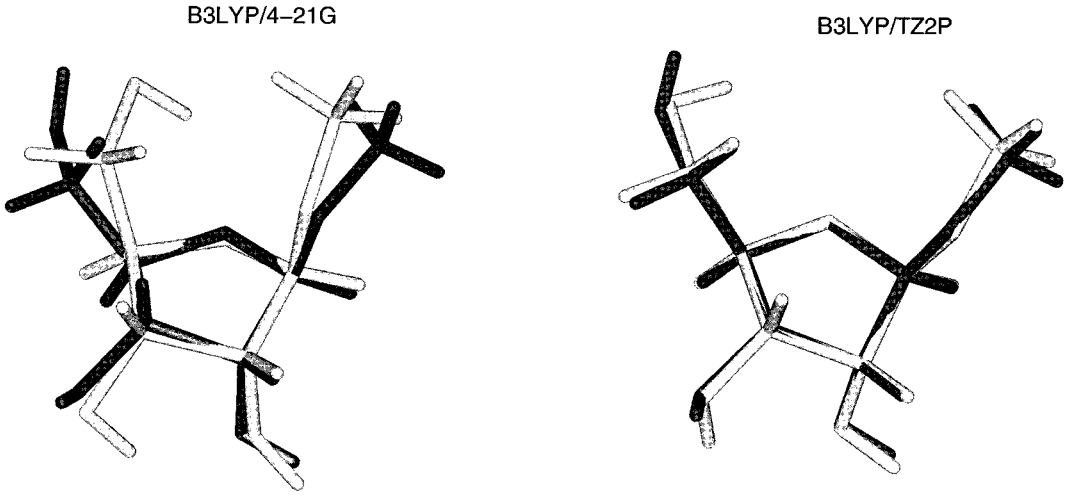

Figure 4. Comparison of X-ray (black) and computed (white) structures of methyl β-D-ribofuranoside II at the B3LYP/4-21G and B3LYP/TZ2P levels of theory.

TABLE 6: Performance of the Methods Used in This Work for the Geometry (Å) and Binding Energy (kcal/mol) of the Water Dimer

	ΔE uncorrected	ΔE with counterpoise correction	r(O···H)	r(O···O)
HF/4-21G	10.6	6.0	1.846	2.818
HF/cc-pVDZ	5.76	3.88	2.036	2.986
MP2/cc-pVDZ	7.47	3.70	1.942	2.909
B3LYP/4-21G	14.0	6.2	1.768	2.765
B3LYP/cc-pVDZ	8.27	4.31	1.909	2.883
LDA/cc-pVDZ	13.8	7.0	1.822	2.543
BLYP/cc-pVDZ	8.67	3.70	1.922	2.899
BP86/cc-pVDZ	8.14	3.92	1.893	2.871
B3PW91/cc-pVDZ	7.29	3.94	1.912	2.883
B3LYP/TZ2P	6.14	4.56	1.945	2.909
B3LYP/cc-pVTZ	6.08	4.49	1.948	2.911
B3LYP/AVDZ	4.68	4.53	1.947	2.910
B3LYP/AVTZ	4.57	4.56	1.954	2.917
MP2/AVQZ[48]	5.04	4.80		2.895
CCSD(T)/AVTZ[48]	5.11	4.68	1.943	2.895
CCSD(T)/AVQZ[48]	5.06	4.86		
MP2/AV5Z (a)[48]	4.98	4.85		
CCSD(T)/V5Z (a)48	5.08	4.83		
exptl[51,52]	5.4 ± 0.7	5.4 ± 0.7		2.946
ref[48] [a]	5.0 ± 0.1	5.0 ± 0.1		2.90

[a] Exhaustive ab initio calibration study.

used for the furanoside (TZ2P), the uncorrected and CP-corrected interaction energies, 6.14 and 4.56 kcal/mol, respectively, fairly tightly bracket the experimental value. The O···H and O···O distances at that level are 1.95 and 2.91 Å, respectively. The latter is about 0.04 Å shorter than the experimental value[52] of 2.946 Å. However, Halkier et al.[48] persuasively argue that this latter value should be revised downward to 2.90 Å, since a discrepancy of 0.05 Å with the true structure is almost an order of magnitude larger than could reasonably be expected for their[48] best calculations.

At the B3LYP/cc-pVDZ level, the uncorrected and CP-corrected interaction energies are 8.27 and 4.31 kcal/mol, respectively. The increase in BSSE with this basis set somewhat shortens the O···H and O···O distances to 1.91 and 2.88 Å, respectively. However, with the 4-21G basis set, dramatic shortenings to 1.77 and 2.77 Å are observed, together with an unrealistically large uncorrected interaction energy of 14.0 kcal/mol. (The CP-corrected value, 6.2 kcal/mol, is still too high but at least in the right range.) This exaggerated hydrogen bond strength is consistent with the fact that the energy difference between structures **I** and **II** of methyl β-D-ribofuranoside is calculated as 6.86 kcal/mol at the B3LYP/4-21G level, compared to 2.49 kcal/mol at the B3LYP/TZ2P level.

Halkier et al.[48] found that the use of (diffuse-function) "augmented" basis sets such as aug-cc-pVTZ[53] cuts BSSE by as much as an order of magnitude at the SCF level. Indeed, we arrive at the same conclusion at the B3LYP level; with the aug-cc-pVDZ basis set (which is the same size as TZ2P), the uncorrected and corrected interaction energies are 4.68 and 4.53 kcal/mol, respectively, a difference that narrows to 4.57 vs 4.56 kcal/mol with the aug-cc-pVTZ basis set. The geometrical parameters at the B3LYP/aug-cc-pVDZ level, r(O···H) = 1.95 Å and r(O···O) = 2.91 Å, agree to better than two decimal places with the B3LYP/TZ2P geometry.

Finally, Halkier et al. found that the MP2 basis set limit for the dimer interaction parameters (binding energy, intermonomer geometric parameters) was quite close to the CCSD(T) basis set limit, even though the MP2 basis set limit for the monomer geometry and total atomization energies is significantly different from the exact result.

2.2. Methyl β-D-Ribofuranoside (Continued). From this digression, we conclude that the lowest adequate level of theory for predicting the ring conformation (and related structural features) of the furanosides will be B3LYP/cc-pVDZ. The formidable extra computational cost of B3LYP/TZ2P would be required only if quantitative accuracy in the computed r_e bond distances (particularly for the internal hydrogen bonds) were of the utmost importance.

The calculated bond distances at both the B3LYP/cc-pVDZ and the B3LYP/TZ2P levels agree well with the neutron results, with two exceptions: r(C1−O1) in structure **I** and r(O4−C1) in structure **II**, for which the discrepancies are somewhat outside the experimental uncertainty. For such comparisons, then, the smaller basis set is quite adequate and certainly more economical.

As previously seen by Handy and co-workers[13] for the water dimer, the LDA exchange-correlation functional is wholly inadequate for systems with hydrogen bonds. Interestingly, the computed hydrogen bond distances in structure **II** are quite different between the BLYP and BP86 functional, and also differ mildly between B3LYP and B3PW91, suggesting a fairly strong dependence on nonlocal corrections to the correlation part of the functional.

As we have mentioned above, structure **I** is adequately predicted even at the B3LYP/4-21G level of theory. We

TABLE 7: Computed and Observed Structural Parameters (Å, deg) for Methyl α-D-Lyxofuranoside

	MM3	HF/4-21G	HF/cc-pVDZ	MP2/cc-pVDZ	B3LYP/4-21G	B3LYP/cc-pVDZ	exptl X-ray[a]	B3PW91/cc-pVDZ
r(C1−C2)	1.540	1.513	1.531	1.519	1.521	1.523	1.540	1.519
r(C2−C3)	1.530	1.533	1.556	1.545	1.557	1.550	1.541	1.545
r(C3−C4)	1.526	1.549	1.535	1.564	1.572	1.567	1.527	1.562
r(C4−O4)	1.428	1.457	1.413	1.452	1.508	1.452	1.463	1.445
r(O4−C1)	1.429	1.429	1.389	1.421	1.470	1.420	1.454	1.413
r(C1−O1)	1.420	1.419	1.384	1.400	1.432	1.402	1.411	1.397
α(C1,C2,C3)	104.6	101.4	103.2	100.9	102.6	101.8	103.3	101.7
α(C2,C3,C4)	102.1	103.0	103.3	101.8	102.4	102.4	99.7	102.2
α(C3,C4,O4)	103.4	104.3	104.6	106.4	105.7	106.1	102.7	106.2
α(C4,O4,C1)	106.7	110.9	107.4	105.7	104.8	107.6	109.0	107.2
α(O4,C1,C2)	105.7	105.5	105.5	102.6	101.6	103.3	105.2	103.1
r(O2H···O3)	2.027	2.109	2.316	3.008	2.979	3.014	3.451	2.990
r(O2H···O4)	3.162	3.498	2.949	2.272	2.217	2.419	4.137	2.373
r(O3H···O2)	2.551	3.341	2.433	2.023	1.892	2.028	2.789	2.003
τ(C1,C2,C3,C4)	19.4	−36.9	0.3	−30.2	−30.0	−28.1	39.0	−28.0
τ(C2,C3,C4,O4)	−37.5	25.4	−22.9	4.9	3.4	5.1	−43.8	4.4
τ(C3,C4,O4,C1)	42.7	−3.2	39.9	24.6	25.7	22.3	33.0	23.5
τ(C4,O4,C1,C2)	−29.7	−20.9	−39.9	−44.9	−45.1	−41.2	−7.7	−42.3
τ(O4,C1,C2,C3)	4.9	35.8	23.0	46.7	46.7	42.8	−20.4	43.4
P	62.53	166.08	89.62	129.82	128.95	130.58	28.13	129.45

[a] Averaged estimated standard deviation (AESD): 0.0015 Å on bond distances, 0.1° on bond angles.

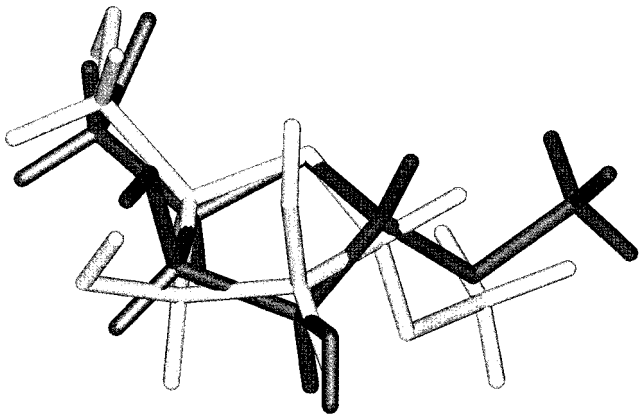

Figure 5. Superimposed (least-squares fit of ring atoms) B3LYP/cc-pVDZ computed (white) and observed X-ray (black) structures of methyl α-D-lyxofuranoside.

attributed this to the absence of an internal hydrogen bond involving the ring oxygen O4, thereby obviating the necessity of accurate treatment of hydrogen bonds. It is noteworthy that the predicted ring conformation for structures **I** and **II** is nearly the same (***P*** differs by only approximately 4°), although **II** has an internal O5H···O4 bond that **I** lacks. In this case, it is only inaccurate overestimates of the H-bond contribution that cause a false distortion of the ring in low-level predictions of structure **II**.

Given the excellent performance of the MP2 method for the water dimer, we have also considered MP2/cc-pVDZ structures for the two ribofuranoside structures. Despite an approximately 4-fold increase in CPU time over the B3LYP/cc-pVDZ optimizations, the MP2/cc-pVDZ results on the whole do not suggest a great improvement over the former.

3. Methyl α-D-Lyxofuranoside. All relevant data have been given in Table 7. In this case, the crystallographic asymmetric unit only contains one unique structure (Figure 5).

With the B3LYP and B3PW91 exchange-correlation functionals, the computed structure exhibits two internal hydrogen bonds: one, O3H···O2, involving exocyclic groups only and the other, O2H···O4, directly involving the ring oxygen. To bring O2H sufficiently close to O4 to permit this hydrogen bond (even at the elongated distance of 2.42 Å), the O2H group must be in a quasiaxial position, while in the crystal, it is in a quasiequatorial position. (In the crystal, O4 is involved in an external hydrogen bond with HO3 of the adjacent molecule.) This forces the entire ring to adopt a completely different conformation, which explains the startling difference of 103° between the pseudorotation angles ***P*** and the rather large discrepancies between predicted and observed X-ray bond distances, e.g., the discrepancy in *r*(C3−C4) being 0.04 Å.

The B3PW91/cc-pVDZ structure is quite close to its B3LYP counterpart. Compared to B3LYP/cc-pVDZ, B3LYP/4-21G shortens O3H···O2 and O2H···O4 by 0.14 and 0.20 Å, respectively; despite this, ***P*** only changes by −1.6°. At the HF/4-21G level, however, the change in ***P*** by +35.5° (relative to B3LYP/cc-pVDZ) represents a qualitative change in the geometry; in this case, it consists of the complete absence of the O2H···O4 and O3H···O2 hydrogen bonds, with there instead being a somewhat unrealistically short O2H···O3 bond. To rule out the existence of two distinct local minima, this geometry was reoptimized at the B3LYP/cc-pVDZ level, which led back to the structure with the O2H···O4 bond. This serves as an illustration of the pitfalls of using a low-level "ab initio" method like HF/4-21G for structures of flexible biomolecules.

Using a larger cc-pVDZ basis set at the SCF level affects ***P*** dramatically, but the computed value still differs quite substantially from the B3LYP/cc-pVDZ and B3PW91/cc-pVDZ results. The latter two are essentially the same as the MP2/cc-pVDZ result, obtained at substantially greater cost in CPU time.

Finally, the MM3 structure does predict the existence of O3H···O2 but together with a short OH2···O3 bond rather than the OH2···O4 and ends up being quite different from both the B3LYP/cc-pVDZ and the X-ray structures.

4. Methyl α-D-Xylofuranoside. Relevant results are given in Table 8, while computed and observed structure are compared in Figure 6.

Since no internal hydrogen bonds are present in either the computed gas-phase or the observed crystal structures, it comes as no surprise that the conformation predicted at the B3LYP/cc-pVDZ level is almost identical to the X-ray result. Differences between computed and observed bond distances and bond angles display similar trends as noted for the other compounds.

While MM3 does yield a rather good geometry, its pseudorotation angle ***P*** is still off by about 15°.

TABLE 8: Computed and Observed Structural Parameters (Å, deg) for Methyl α-D-Xylofuranoside

	MM3	B3LYP/cc-pVDZ	X-ray[a]
r(C1−C2)	1.523	1.533	1.526
r(C2−C3)	1.528	1.547	1.526
r(C3−C4)	1.547	1.557	1.547
r(C4−O4)	1.437	1.446	1.451
r(O4−C1)	1.424	1.409	1.421
α(C1,C2,C3)	102.0	101.4	100.9
α(C2,C3,C4)	102.7	102.6	102.4
α(C3,C4,O4)	106.9	106.5	105.3
α(C4,O4,C1)	108.8	110.2	110.1
α(O4,C1,C2)	104.0	106.3	104.0
τ(C1,C2,C3,C4)	−31.1	−31.7	−37.2
τ(C2,C3,C4,O4)	12.5	18.4	21.6
τ(C3,C4,O4,C1)	12.8	4.0	4.2
τ(C4,O4,C1,C2)	−33.2	−25.4	−28.4
τ(O4,C1,C2,C3)	39.9	35.7	40.8
P	142.50	154.39	155.68

[a] Averaged estimated standard deviation (AESD): 0.0015 Å on bond distances, 0.1° on bond angles.

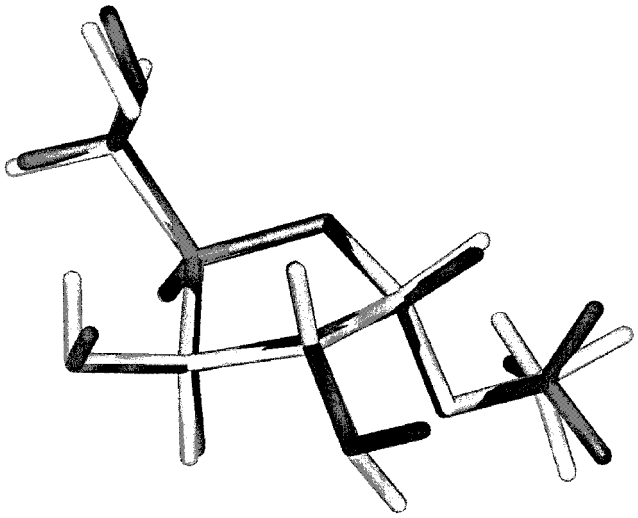

Figure 6. Superimposed (least-squares fit of ring atoms) B3LYP/cc-pVDZ computed (white) and observed X-ray (black) structures of methyl α-D-xylofuranoside.

Conclusions

In this paper, the structures of methyl α-D-arabinofuranoside, methyl β-D-ribofuranoside, methyl α-D-lyxofuranoside, and methyl-α-D-xylofuranoside were studied experimentally by X-ray crystallography at 100 K and, for the first two, by neutron diffraction at 15 K. Gas-phase structures of all four molecules were computed using density functional theory with a variety of basis sets and exchange-correlation functionals; for comparison, Hartree−Fock and second-order many-body perturbation theory (MP2) calculations were also carried out.

Because of the intrinsic sensitivity of the conformation of the five-membered ring to bond lengths and bond angles, molecular mechanics and small basis set SCF treatments are wholly inadequate. In cases where internal hydrogen bonds are prominent, the local density approximation fails because of a tendency to strongly underestimate hydrogen bond distances and overestimate hydrogen bond strengths. Even with more sophisticated exchange-correlation functionals, however, unrealistically short and strong internal hydrogen bonds are likewise obtained with basis sets of less than double-ζ plus polarization quality due to strong basis set superposition error. B3LYP/cc-pVDZ is the lowest level of theory that consistently yields qualitatively correct geometries for all structures investigated here. In addition, B3LYP/cc-pVDZ computed structural parameters by and large agree with low-temperature neutron diffraction data to within their experimental uncertainty. Further expansion of the basis set will lead to more accurate equilibrium bond lengths and bond angles, but such improvement is neither significant compared to the experimental uncertainties in neutron diffraction structures nor is it important for accurate prediction of ring conformation of five-membered ring systems. MP2/cc-pVDZ yields results comparable to those from B3LYP/cc-pVDZ at much greater computational cost. In short, B3LYP/cc-pVDZ appears to represent the best compromise between accuracy and computational cost.

It is worth noting that the B3LYP/cc-pVDZ optimizations reported took, on average, about 2 days CPU time each on a DEC Alpha 500/500 workstation. For comparison, the neutron diffraction data collection for a fairly large crystal of methyl α-D-arabinofuranoside took approximately 2.5 weeks at HFBR. For equivalent accuracy, neutron diffraction data collection times would scale approximately linearly with the number of atoms *N*, while in our experience, B3LYP/cc-pVDZ computation times in this size range scale approximately as $N^{2.66}$. Recent advances[54] in ab initio methods for large systems however open up the prospect of essentially linear scaling with *N*. In addition, the ongoing improvements in CPU performance of workstation-class computers will further reduce the premultiplier of *N*. On the other hand, recent advances in neutron Laue methods[55] may considerably shorten data collection times.

For methyl α-D-arabinofuranoside and methyl β-D-ribofuranoside, there is very good correspondence between the best computed and observed ring conformations, even though some intermolecular hydrogen bonds in the crystal give way to internal hydrogen bonds in the predicted gas-phase structures. On the other hand, in the case of methyl α-D-lyxofuranoside, an O2H···O4 internal hydrogen bond that exists in the predicted gas-phase structure but not in the crystal structure (where instead the donor and acceptor are involved in external hydrogen bonds) leads to a drastic change in ring conformation compared to the crystal.

This study demonstrates the intrinsic difficulty in arriving at convenient generalizations concerning substituent effects on ring conformation in the five-membered ring systems. Not only does the intuitive approach, as embodied in the elaborate molecular mechanics force field, fail to predict the correct structure but even low-level quantum mechanical methods (small basis set, Hartree−Fock or local density approximation) fail. If it were not for the outstanding case of methyl α-D-lyxofuranoside, however, one might have arrived at erroneously optimistic conclusions, since the results of the analyses for methyl α-D-arabinofuranoside, methyl β-D-ribofuranoside, and methyl α-D-xylofuranoside, taken by themselves, suggest that less extensive (and certainly less expensive!) treatments might be adequate. The importance of extending the experimental and theoretical treatment to the remaining four members of the family of methyl furanosides is clear.

Acknowledgment. J.M. is a Yigal Allon Fellow, the Incumbent of the Helen and Milton A. Kimmelman Career Development Chair, and an Honorary Research Associate ("Onderzoeksleider in Eremandaat") of the National Science Foundation of Belgium. He acknowledges support from the Minerva Foundation, Munich, Germany. A.E. acknowledges a Doctoral Fellowship of the Feinberg Graduate School, Weizmann Institute of Science (WIS), and thanks Dr. Felix Frolow (Department of Chemical Services, WIS) for his helpful instruction in the performance of X-ray diffraction experiments

and C. Koehler, III for technical assistance during the neutron diffraction experiments at Brookhaven National Laboratory (BNL). The authors thank Dr. Kim Baldridge (San Diego Supercomputer Center, La Jolla, CA) for critical reading of the manuscript prior to submission. Work at BNL was carried out under Contract DE-AC02-98CH10886 with the U.S. Department of Energy, Office of Basic Energy Sciences.

Supporting Information Available: Tables of atomic coordinates and displacement parameters from the neutron and X-ray diffraction experiments and geometries in Cartesian coordinates for the calculated structures. This material is available free of charge via the Internet at http://pubs.acs.org. The neutron structures for **1** (CCDC112946) and **2** (CCDC112948), as well as the x-ray structures for **1** (CCDC112947), **2** (CCDC112949), **3** (CCDC112951), and **4** (CCDC112950) have also been deposited with the Cambridge Crystallographic Data Centre (CCDC) under the identifiers given in parentheses.

References and Notes

(1) Altona, C.; Sundaralingam, M. *J. Am. Chem. Soc.* **1972**, *94*, 8205.
(2) Dzakula, Z.; De Rider, M. L.; Markley, J. L. *J. Am. Chem. Soc.* **1996**, *118*, 12796 and references therein.
(3) Tomimoto, M.; Go, N. *J. Phys. Chem.* **1995**, *99*, 563 and references therein.
(4) Rudnicki, W. R.; Leysyng, B.; Harvey, S. C. *Biopolymers* **1994**, *34*, 383 and references therein.
(5) Church, T. J.; Carmichael, I.; Serianni, A. S. *J. Am. Chem. Soc.* **1997**, *119*, 8946 and references therein.
(6) Cros, S.; Hervé du Penhoat, C.; Pérez, S.; Imberty, A. *Carbohydr. Res.* **1993**, *248*, 81.
(7) Groth, P.; Hammer, H. *Acta Chem. Scand.* **1968**, *22*, 2059. It should be pointed out that the coordinates for methyl α-D-lyxofuranoside listed there and in the associated entry MLYXRA in the Cambridge Structural Database are, in fact, those of the mirror compound methyl-α-L-lyxofuranoside, although the accompanying figures are those of the correct compound.
(8) Podlasek, C. A.; Stripe, W. A.; Carmichael, I.; Shang, M.; Basu, B.; Serianni, A. S. *J. Am. Chem. Soc.* **1996**, *118*, 1413.
(9) For general reviews, see: (a) Parr, R. G.; Yang, W. *Density functional theory of atoms and molecules*; Oxford University Press: Oxford, 1989. (b) Handy, N. C. In *Lecture Notes in Quantum Chemistry II*; Roos, B. O., Ed.; Lecture Notes in Chemistry 64; Springer: Berlin, 1994. (c) Parr, R. G.; Yang, W. *Annu. Rev. Phys. Chem.* **1995**, *46*, 701.
(10) Martin, J. M. L.; El-Yazal, J.; François, J. P. *Mol. Phys.* **1995**, *86*, 1437.
(11) (a) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648. See also for the slightly modified parameters commonly used, the following: (b) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. *J. Phys. Chem.* **1994**, *98*, 11623.
(12) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.
(13) Mok, D. K. W.; Handy, N. C.; Amos, R. D. *Mol. Phys.* **1997**, *92*, 667.
(14) Ness, R. K.; Fletcher, H. G. *J. Org. Chem.* **1957**, *22*, 2007.
(15) Green, J. W.; Pascu, E. *J. Am. Chem. Soc.* **1937**, *59*, 1206.
(16) Wu, J.; Serianni, A. S. *Carbohydr. Res.* **1991**, *210*, 51−71.
(17) *MSC/AFC Diffractometer Control Software*; Molecular Structure Corporation: The Woodlands, TX, 1991.
(18) *Xtal3.2 User's Manual*; Hall, S. R., Flack, H. D., Stewart, J. M., Eds.; University of Western Australia: Nedlands (Perth) Australia; University of Geneva: Switzerland; University of Maryland: Washington, DC, 1992.
(19) Sheldrick, G. M. *Acta Crystallogr. A* **1990**, *46*, 467.
(20) Sheldrick, G. M. *SHELXL93, Program for the refinement of crystal structures*; University of Göttingen: Göttingen, 1993.
(21) McMullan, R. K.; Epstein, J.; Ruble, J. R.; Craven, B. M. *Acta Crystallogr. B* **1979**, *35*, 688.
(22) de Meulenaar, J.; Tompa, H. *Acta Crystallogr.* **1965**, *19*, 1014. Templeton, L. K.; Templeton, D. H. *Abstr. Am. Crystallogr. Assoc. Meet.*; Storrs, CT, 1973; p 143.
(23) Lundgren, J.-O. *Crystallographic Computer Programs*; Report UUIC-B13-4-05; Institute of Chemistry, University of Uppsala: Uppsala, Sweden.
(24) Møller, C.; Plesset, M. S. *Phys. Rev.* **1934**, *46*, 618.
(25) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Gill, P. M. W.; Johnson, B. G.; Robb, M. A.; Cheeseman, J. R.; Keith, T.; Petersson, G. A.; Montgomery, J. A.; Raghavachari, K.; Al-Laham, M. A.; Zakrzewski, V. G.; Ortiz, J. V.; Foresman, J. B.; Cioslowski, J.; Stefanov, B. B.; Nanayakkara, A.; Challacombe, M.; Peng, C. Y.; Ayala, P. Y.; Chen, W.; Wong, M. W.; Andres, J. L.; Replogle, E. S.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Binkley J. S.; DeFrees, D. J.; Baker, J.; Stewart, J. P.; Head-Gordon, M.; Gonzalez, C.; Pople, J. A. *GAUSSIAN 94*, revision D.4; Gaussian, Inc.: Pittsburgh, PA, 1995.
(26) SPARTAN 4.1; Wavefunction, Inc.: Irvine, CA, 1995.
(27) Kohn, W.; Sham, L. *J. Phys. Rev.* **1965**, *140*, 1133.
(28) Vosko, S. H.; Wilk, L.; Nusair, M. *Can. J. Phys.* **1980**, *58*, 1200.
(29) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098.
(30) Perdew, J. P. *Phys. Rev. B* **1986**, *33*, 8822.
(31) Perdew, J. P.; Wang, Y. *Phys. Rev. B* **1992**, *45*, 13244.
(32) Devlin, F. J.; Finley, J. W.; Stephens, P. J.; Frisch, M. J. *J. Phys. Chem.* **1995**, *99*, 16883. Rauhut, G.; Pulay, P. *J. Phys. Chem.* **1995**, *99*, 3093. Wong, M. W. *Chem. Phys. Lett.* **1996**, *256*, 391.
(33) De Proft, F.; Geerlings, P. *J. Chem. Phys.* **1997**, *106*, 3270.
(34) De Proft, F.; Martin, J. M. L.; Geerlings, P. *Chem. Phys. Lett.* **1996**, *250*, 393.
(35) (a) De Proft, F.; Martin, J. M. L.; Geerlings, P. *Chem. Phys. Lett.* **1996**, *256*, 400. (b) Geerlings, P.; De Proft, F.; Martin, J. M. L. In *Theoretical and Computational Chemistry: Recent developments and applications of modern density functional theory*; Seminario, J., Ed.; Elsevier: New York, 1996; Vol. 4, pp 773−809.
(36) Martin, J. M. L.; Taylor, P. R. *J. Phys. Chem.* **1996**, *100*, 6047 and references therein.
(37) Martin, J. M. L.; El-Yazal, J.; François, J. P. *Chem. Phys. Lett.* **1995**, *242*, 570.
(38) Hutter, J.; Lüthi, H. P.; Diederich, F. *J. Am. Chem. Soc.* **1994**, *116*, 750.
(39) Martin, J. M. L.; Aviyente, V.; Lifshitz, C. *J. Phys. Chem. A* **1997**, *101*, 2597.
(40) Dunning, T. H., Jr. *J. Chem. Phys.* **1989**, *90*, 1007.
(41) Pulay, P.; Fogarasi, G.; Pang, F.; Boggs, J. E. *J. Am. Chem. Soc.* **1979**, *101*, 2550.
(42) (a) Allinger, N. L.; Yuh, Y. H.; Lii, J. H. *J. Am. Chem. Soc.* **1989**, *111*, 8551. (b) Lii, J. H.; Allinger, N. L. *J. Am. Chem. Soc.* **1989**, *111*, 8566, 8576.
(43) Schlegel, H. B. In *New Theoretical Concepts for Understanding Organic Reactions*; Bertran, J., Ed.; Kluwer: Dordrecht, The Netherlands, 1989 and references therein.
(44) Peng, C. Y.; Ayala, P. Y.; Schlegel, H. B.; Frisch, M. J. *J. Comput. Chem.* **1996,** *17*, 49.
(45) For an review of different types of observed bond distances, see: Kuchitsu, K.; In *Accurate Molecular Structures*; Domenicano, A., Hargittai, I., Eds.; International Union of Crystallography and Oxford University Press: Oxford, 1992.
(46) Scheiner, S. *Annu. Rev. Phys. Chem.* **1994**, *45*, 23.
(47) Xantheas, S.; Dunning, T. H., Jr. *J. Chem. Phys.* **1993**, 99, 9774. Mas, E. M.; Szalewicz, K. *J. Chem. Phys.* **1996**, *104*, 7606. Feyereisen, M. W.; Feller, D.; Dixon, D. A. *J. Phys. Chem.* **1996**, *100*, 2993.
(48) Halkier, A.; Koch, H.; Jørgensen, P.; Christiansen, O.; Beck Nielsen, I. M.; Helgaker, T. *Theor. Chem. Acc.* **1997**, *97*, 150 (journal formerly known as *Theor. Chim. Acta*).
(49) For example, see: Hobza, P.; Zahradnik, R. *Chem. Rev.* **1988**, *88*, 871.
(50) Boys, S. F.; Bernardi, F. *Mol. Phys.* **1970**, *19*, 553.
(51) Curtiss, L. A.; Frurip, D. J.; Blander, M. *J. Chem. Phys.* **1979**, *71*, 2703. Reimers, J.; Watts, R.; Klein, M. *Chem. Phys.* **1982**, *64*, 95.
(52) Odutola, J. A.; Dyke, J. M. *J. Chem. Phys.* **1980**, *72*, 5062.
(53) Kendall, R. A.; Dunning, T. H., Jr.; Harrison, R. J. *J. Chem. Phys.* **1992**, *96*, 6796.
(54) White, C. A.; Johnson, B. F. G.; Gill, P. M. W.; Head-Gordon, M. *Chem. Phys. Lett.* **1996**, *253*, 268. Stratmann, R. E.; Scuseria, G. E.; Frisch, M. J. *Chem. Phys. Lett.* **1996**, *257*, 213. Schwegler, E.; Challacombe, M.; Head-Gordon, M. *J. Chem. Phys.* **1997**, *106*, 9708. Challacombe, M.; Schwegler, E. *J. Chem. Phys.* **1997**, *106*, 5526. Strain, M. C.; Scuseria, G. E.; Frisch, M. J. *Science* **1996**, *271*, 51. Millam, J. M.; Scuseria, G. E. *J. Chem. Phys.* **1997**, *106*, 5569. Challacombe, M.; Schwegler, E. *J. Chem. Phys.* **1996**, *105*, 2726. Burant, J. C.; Scuseria, G. E.; Frisch, M. J. *J. Chem. Phys.* **1996**, *105*, 8969.
(55) Niimwa, N.; Minezaki, Y.; Nonaka, T.; Castagna, J.-C.; Cipriani, F.; Høghøj, P.; Lehmann, M. S.; Wilkinson, C. *Nat. Struct. Biol.* **1997**, *4*, 909. Helliwell, J. R. *Nat. Struct. Biol.* **1997**, *4*, 874. Habash, J.; Raftery, J.; Weisgerber, S.; Cassetta, A.; Lehmann, M. S.; Høghøj, P.; Wilkinson, C.; Campbell, J. W.; Helliwell, J. R. *J. Chem. Soc., Faraday Trans.* **1997**, *93*, 4313.