Carbohydrate Research, 79 (1980) 155–163

APPLICATION OF *ab initio* MOLECULAR-ORBITAL CALCULATIONS TO THE STRUCTURAL MOIETIES OF CARBOHYDRATES. PART VI

GEORGE A. JEFFREY AND JOHN H. YATES
Department of Crystallography, University of Pittsburgh, Pittsburgh, Pennsylvania 15260 (U.S.A.)
(Received June 4th, 1979; accepted for publication, July 18th, 1979)

ABSTRACT

Ab initio RHF/4-31G molecular-orbital calculations have been conducted on methoxymethyl formate and methoxymethyl acetate as models for examining the *anomeric effect* and stereochemistry of 1-*O*-acetylglycopyranoses. The results indicate that, as with the methyl glycopyranosides, the α-$^4C_1$(D) configurations are more stable than the β-$^4C_1$(D), except that the energy difference is more dependent on the disposition about the glycosidic bond. The lowest-energy conformations occur with glycosidic torsion-angles of $\varphi = 180°$, where the anomeric energy is about 4 kcal/mol. There is a secondary energy-minimum at $\varphi = 90°$, for which the anomeric energy is less, about 2 kcal/mol. This orientation corresponds to the conformation most commonly observed in the crystal structures of peracetylated glycopyranoses. Small differences in the C–O single-bond lengths, which are observed experimentally in both the α and β anomers, are reproduced by the theoretical calculations.

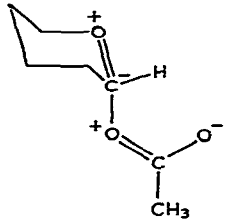
INTRODUCTION

The modeling of the acetal and hemiacetal moieties of carbohydrate molecules with smaller molecules, for which *ab initio* quantum-mechanical calculations are at present possible, has shown that agreement between theory and experiment to the order of 2 pm in C–O bond-lengths and 4° in valence angles may be obtained at the RHF/4-31G level[1]. In this paper, we use methoxymethyl formate (**1**), H_3C-O-CH_2-O-CHO, and methoxymethyl acetate (**2**), H_3C-O-CH_2-O-$COCH_3$, as models for 1-*O*-acetyl glycopyranoses*. The peracetylated aldopyranoses have been examined extensively by X-ray single-crystal structure analysis, and the results from eleven 1-*O*-acetyl derivatives provide a consistent set of experimental structural data for comparison with theory. The replacement of the hydroxyl hydrogen atom of the hemiacetal group in a pyranosid sugar by a carbonyl group introduces significant

---

*Although the compounds of interest are the glycopyranosyl acetates, we used the methoxymethyl formate molecule for most of the theoretical calculations because of the fewer electrons and consequently the considerably less computation required. It was anticipated and shown that substitution of the formate hydrogen atom for a methyl group had only a small effect on the acetal geometry and conformational-energy differences.

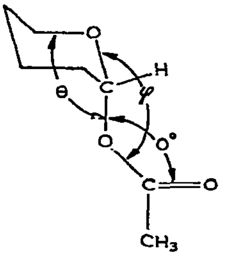
changes in electronic structure, as shown by the changes in the single C–O bond lengths. The glycosidic C–O bond, which is the shortest of the C–O bonds in glycopyranoses and glycopyranosides, becomes one of the longer C–O bonds in the 1-acetates, and the shortest C–O bond in the acetal moiety is that from the anomeric carbon atom to the ring oxygen atom. These structural features would be represented in valence-bond resonance terminology by a significant contribution from the resonance structure 3.



$$H_3C{-}\overset{+}{O}{=\!=}CH_2{\cdots}O{=\!=}C(CH_3){=\!=}\overset{-}{O}$$

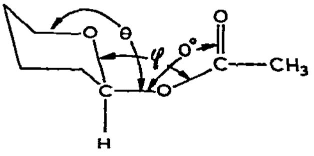
3

Present day, *ab initio* molecular-orbital calculations permit us to calculate the molecular dimensions for the model compounds, 1 and 2, at a level of approximation that is limited by the capacity of the computer used and the cost of computation. For this reason, most of the calculations in this paper refer to 1, rather than 2, although the experimental data are only available for glycopyranosyl acetates.



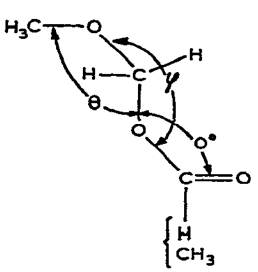
4

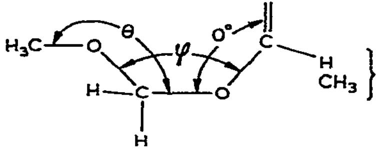
θ = +60°



5

θ = 180°





As in the previous calculations in this series[1], different conformations of **1** and **2** correspond to the different anomeric configurations of the 1-*O*-acetylglycopyranoses, **4** and **5**, as shown next. It is therefore possible, by using methoxymethyl formate and acetate, to investigate the stereochemistry and the anomeric energies for glycopyranosyl acetates and compare these results with those obtained previously for the pyranoses, methyl pyranosides, and pyranosyl fluorides and chlorides by using equivalent approximations and similar models, namely, methoxymethanol, dimethoxymethane, and methoxymethyl fluoride and chloride.

THE MOLECULAR-ORBITAL METHOD

Closed-shell, restricted Hartree–Fock theory was used with a 4-31G basis set for the energy calculations, which are shown in Table I. The first calculations on formate were performed with a DEC KI-10 computer using a version[2] of GAUSSIAN 74, with the standard bond-lengths[3] as shown, R_{CH} = 109 pm, and tetrahedral angles, except at the carbonyl carbon atom, which was assumed to be trigonal planar. The three single C–O bond-lengths were then optimized by using point-by-point energy calculations, keeping the remainder of the geometry fixed. When GAUSSIAN 78 became available, the remaining calculations were carried out with that program on a VAX 11/780 computer. The same bond-lengths were then optimized by using the modified Davidon–Fletcher–Powell scheme[4] included in GAUSSIAN 78. The assumption that replacement of the formate hydrogen atom by a methyl group in the acetate would make only minor differences to the optimized bond-lengths and energy-differences was then checked by conducting limited calculations on methoxymethyl acetate, also using GAUSSIAN 78. These calculations were performed both with standard geometry (with C–CH_3 = 152 pm) and with optimized C–O bond lengths, as with the formate. Total energies were calculated for both **1** and **2** with θ = 60°, namely the α-$^4C_1$(D) model, and with θ = 180°, that is, the β-$^4C_1$(D) model, using selected values of φ, which correspond to different, important orientations about the glycosidic bond. The CH_2-O-C=O torsion-angle was fixed at 0.0°, which is consistent with the short C–O bond (*d*) and the mean value for the orientation of the 1-*O*-acetyl groups (ω) in the experimental data given in Table II, which we used for bond-length comparisons. This *cis* disposition is also observed in electron-diffraction[5] and microwave studies[6] of methyl formate. In the acetate molecule, the methyl group of the acetate was oriented with one C–H bond *cis* to the C=O bond, in accordance with the general trend observed in acetate structures. The other terminal methyl group was oriented with one C–H bond *trans* to the adjacent C–O bond in both the formate and the acetate.

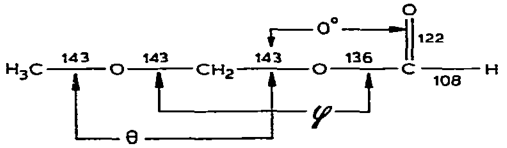


TABLE I

ENERGIES (IN HARTREES[a] AND RELATIVE KCAL/MOL) OF METHOXYMETHYL FORMATE AND METHOXYMETHYL ACETATE, WITH STANDARD GEOMETRY AND OPTIMIZED C–O BOND LENGTHS

Conformation			*Methoxymethyl formate*		*Methoxymethyl acetate*	
	θ(°)	φ(°)	*Standard geometry*	*Optimized C–O*	*Standard geometry*	*Optimized C–O*
α-$^4C_1$	60	−60	−340.74772 246.0		−379.74005 247.0	
		+60	−341.12887 6.7	−341.13014 6.8	−380.12203 7.2	−380.12322 6.8
		+90	−341.13628 2.1			
		+120	−341.13317 4.0			
		180	−341.13959 0.0	−341.14096 0.0	−380.13353 0.0	(−380.13490)[b] (0.0)
β-$^4C_1$	180	±60	−341.12876 6.8	−341.12904 7.5	−380.12206 7.2	−380.12231 7.3
		±90	−341.13288 4.2			
		±120	−341.12718 7.8			
		180	−341.13337 3.9	−341.13392 4.4	−380.12743 3.8	(−380.12798) (4.3)

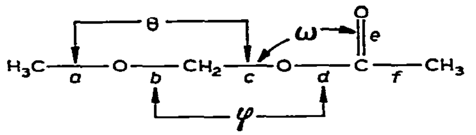
[a] 1 Hartree = 627.544 kcal/mol. [b] Parentheses indicate estimated values.

RESULTS

The results of the energy calculations are shown in Table I. As with the glycopyranoses and methyl glycopyranosides previously examined by this methodology, the α-$^4C_1$(D) type conformations, $\theta = 60°$, have the lowest energy for both model compounds and for all values of φ examined. This is consistent with the *anomeric effect*[7]. An important difference occurs, however, in the energies associated with the *exo-anomeric* angle φ, where the lowest-energy conformation appears at $\varphi = 180°$ for both α and β models, rather than −60°, as in the pyranoses and methyl pyranosides. There are also local minima near $\varphi = 90°$, for both $\theta = 60$ and 180° (see Fig. 1), which were not observed in the pyranose and methyl pyranoside model-calculations. The *anomeric energy*, as defined by the difference between the energies of the lowest-energy conformations for the α and β configurations, $\varphi = 180°$, is 4.4 kcal/mol based on the optimized formate model, or 3.8 kcal/mol based on the unoptimized acetate model. For reasons of economy, the optimized acetate models at $\varphi = 180°$ were not computed. However, comparison of the standard-geometry calculations for the formate and acetate showed that replacing the hydrogen atom by a methyl group makes a difference of <0.5 kcal/mol in the anomeric energies. The optimization

TABLE II

EXPERIMENTAL DATA ON 1-*O*-ACETYLGLYCOPYRANOSES[a]



Molecules[b]	*a*	*b*	*c*	*d*	*e*	*f*	θ	φ	ω	*Ref.*
D *axial*										
α-$^4C_1$, β-$^1C_4$										
1	142.9	139.5	141.6	135.6	119.0	148.3	−55	−69	−6	10
2	144.3	140.8	143.8	136.5	118.9	152.2	64	81	−7	11
3	144.0	140.6	144.8	135.3	119.1	144.9	62	86	9	12
	143.9	140.1	143.7	135.6	121.7	147.2	71	84	−3	12
4	144.0	139.6	143.5	135.9	119.5	147.8	62	82	4	13
5	144.1	139.4	144.2	136.9	118.6	148.8	69	77	6	14
6	144.0	139.7	143.8	135.8	119.9	148.0	62	87	−7	15
7	143.3	138.9	144.1	132.2	119.0	148.0	63	124	1	16
8	144.0	139.0	143.2	137.4	118.5	149.9	65	80	9	17
D *equatorial*										
α-$^1C_4$, β-$^4C_1$										
9	142.2	140.9	140.9	135.5	118.5	149.3	179	88	2	18
10	142.9	141.3	141.6	136.4	117.5	147.5	179	−100	2	19
11	142.3	142.0	140.7	134.5	120.1	145.9	180	−87	5	20

[a] *a, b, c, d, e,* and *f* in pm; θ, φ, and ω in degrees.
[b] Key to the molecules: 1,2,3,4-tetra-*O*-acetyl-β-L-arabinopyranose (**1**), TACARP10; 1,2,3,4-tetra-*O*-acetyl-α-D-lyxopyranose (**2**), TALYXP; 1,2,3,4-tetra-*O*-acetyl-α-D-ribopyranose (**3**), AADRIB; 1,2,3,4,6-penta-*O*-acetyl-α-D-altropyranose (**4**), PACALP; 1,2,6-tri-*O*-acetyl-2-(*N*-acetylacetamido)-2,3-dideoxy-α-D-*threo*-hex-2-enopyranose (**5**), AAXTHP; 1,3,4-tri-*O*-acetyl-2-deoxy-2-fluoro-α-D-xylopyranose (**6**), ACFXYP; 1,2,3,4,6-penta-*O*-acetyl-α-D-gulopyranose (**7**), PACDGP; 1,2,3,4,5-penta-*O*-acetyl-α-D-idopyranose (**8**), PAIDOP; 1,2,3,4-tetra-*O*-acetyl-α-D-arabinopyranose (**9**), ACARAP; methyl 1,2,3,4-tetra-*O*-acetyl-β-D-galactopyranuronate (**10**), MEAGPY; 1,2,3,4-tetra-*O*-acetyl-β-D-xylopyranose (**11**), TAXYLR. The CODE NAMES are the unique reference code in the Cambridge Crystallographic Data Base; they can be used for direct retrieval of the atomic coordinates and crystal data from the Data-Base tape.

of the C–O bonds in the φ = 180° formate model increased the anomeric energy by 0.5 kcal/mol. Estimating the energies of the optimized acetate models for the minimum θ = 60°, φ = 180°, and θ = 180°, φ = 180°, by subtracting the energy gained by optimizing the bond lengths in the formate model, we obtain a value of 4.3 ±0.5 kcal/mol for the estimated anomeric energy for an optimized acetate model. This value is close to that calculated for the aldopyranosyl fluorides (4.5 kcal/mol).

An important difference between these results and those of the glycopyranoses and the methyl glycopyranosides is that the minimum-energy conformations correspond to φ = 180°, rather than —60°, as in the methyl glycopyranosides. There are also secondary minima at φ = 90°, as shown in Fig. 1. These φ = 90° conforma-

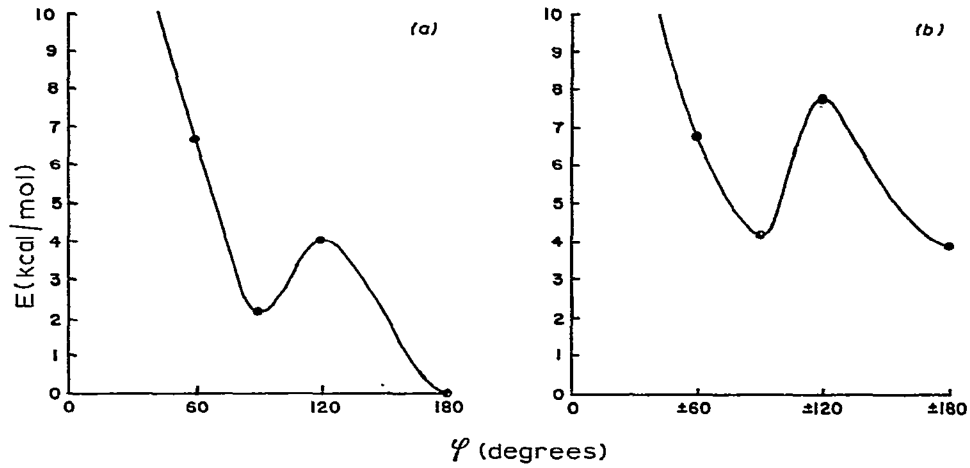


Fig. 1. Variation of energy with glycosidic torsion-angle for methoxymethyl formate. (*a*) α-conformer model. All energy points are relative to 0.0. (*b*) β-conformer model.

tions correspond to local minima 2 kcal/mol above the 180° value for the α model and only 0.3 kcal/mol above the 180° value for the β model.

In the twelve 1-*O*-acetylglycopyranose molecules, shown in Table II, which form the experimental data for this work, the φ angles vary between 69 and 124° with a mean of 86° for the nine α-pyranosides, and between 87 and 100° with a mean of 90° for the three β-pyranosides. An important factor determining this conformational angle in the sugar molecules is the non-bonding interaction between the substituents at C-2 and the carbonyl group of the acetate, which results in a strong steric repulsion when $\varphi = 180°$. This interaction is, of course, excluded from our model calculations. Were it included, by replacing one of the hydrogen atoms on the central CH_2 by a methyl group or another acetate for example, we can predict that the lowest conformational-energy minima would occur in the region of $\varphi = 90°$, rather than

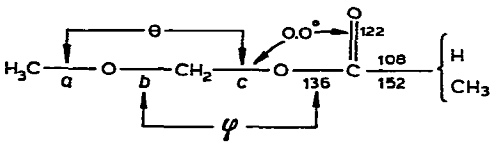
TABLE III

VARIATION OF ANOMERIC ENERGY, $E(\beta\text{-}^4C_1) - E(\alpha\text{-}^4C_1)$, (KCAL/MOL), WITH GLYCOSIDIC TORSION ANGLE (φ)

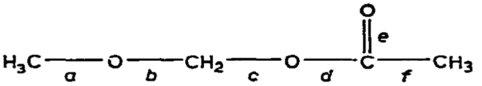
	Methyl glycopyranosides modelled with dimethoxymethane		*1-O-Acetylglycopyranoses modelled with*		
			methoxymethyl formate		*methoxymethyl acetate*
φ	*Standard geometry*	*Optimized geometry*	*Standard geometry*	*Optimized geometry*	*Standard geometry*
+60	2.59	2.38	0.07	0.69	−0.02
+90			2.13		
+120	4.77		3.76		
180	5.09	5.06	3.90	4.42	3.83

TABLE IV

OPTIMIZED C–O BOND-LENGTHS (pm) IN METHOXYMETHYL FORMATE (1), AND METHOXYMETHYL ACETATE (2)



Theory		θ(°)	φ(°)	*a*	*b*	*c*
α-$^4C_1$ models (axial)	1	+60	+60	144.8	140.8	147.1
	2			144.7	140.9	147.2
	1	+60	180	[143.0][a]	139.5	146.5
β-$^4C_1$ models (equatorial)	1	180	+60	143.4	141.3	144.1
	2			143.4	141.3	144.1
	1	180	180	[143.0][a]	140.5	144.2



Experiment	*a*	*b*	*c*	*d*	*e*	*f*
Mean of:						
Nine α-$^4C_1$ or β-$^1C_4$	143.8	139.7	143.6	135.7	119.4	148.3[b]
Three β-$^4C_1$ or α-$^1C_4$	142.5	141.4	141.1	135.5	118.7	147.6[b]

[a] Assumed values. [b] This apparent shortening of the C–CH_3 bond arises, in our view, from the large thermal motion of the terminal methyl groups, for which no corrections were applied in these data.

180°. As shown in Table III, the *anomeric energy* is more sensitive to the value of φ than with the model calculations for methyl glycopyranosides. At $\varphi = 90°$, the anomeric energy is 55% of that at 180°. Therefore, 2 kcal/mol is a reasonable estimate for most 1-*O*-acetylglycopyranoses. A value of 1.3 kcal/mol is estimated from equilibrium experiments[7].

The results of the C–O bond-length optimizations are given in Table IV. In one conformation, the differences between the formate- and acetate-optimized bond-lengths were 0.1 pm, and in the other they were <0.1 pm, justifying the assumption that the distinction between the terminal hydrogen and methyl in the model compounds was relatively small.

Table V gives the Mulliken charges, and overlap-population analysis for the methoxymethyl acetate in the 60°, 60° (α) and 180°, −60° (β) conformations. The only significant difference in the charges is on the anomeric carbon atom, which is slightly (0.05 electron) more positive in the β conformer. The overlap populations for bonds *b* and *c* reflect the lengths of the optimized bonds. Generally, the larger the

TABLE V

MULLIKEN POPULATION-ANALYSIS FOR METHOXYMETHYL ACETATE[a]

Charges	θ(°)	φ(°)	H_3C—	O—	CH_2—	O—	C—	(CH_3)	$=O$
α	60	60	−0.14	−0.66	0.29	−0.73	0.77	−0.50	−0.56
β	180	−60	−0.14	−0.67	0.34	−0.72	0.78	−0.50	−0.57
Overlap population			*a*	*b*	*c*	*d*	*e*	*f*	
α	60	60	0.153	0.174	0.116	0.086	0.476	0.274	
β	180	−60	0.154	0.155	0.148	0.098	0.476	0.276	

[a] For nonhydrogen atoms only. The Mulliken charges on hydrogen atoms ranged from 0.77 to 0.86 and overlap populations from 0.372 to 0.399.

overlap population, the shorter the bond is expected to be. This is shown to be true for the C–O bonds *b* and *c*, and in particular, the longest C–O bond found in the α conformer has the smallest (0.116) overlap population of these two bonds. The C–O bond *d* has an even smaller overlap population, but the bond is known to be short (136 pm in Pople's recommended standard geometry). This shortening appears to result from conjugation of the carbonyl double-bond with this bond.

COMPARISON WITH EXPERIMENTAL DATA

The theoretical-energy calculations refer to the *isolated molecules at rest*. For this reason, comparisons with experimental energy-differences derived from equilibrium experiments, in which the molecules are in solution, are, at best, only qualitative. Comparisons of calculated molecular dimensions with those observed for carbohydrate molecules in the crystalline state have been more rewarding, often providing agreement within 2 pm in bond lengths[8]. This agreement has, in fact, provided the principal justification for continuing with this series of studies, using smaller model compounds for the *ab initio* calculations[1].

The experimental crystal-structure data for the 1-*O*-acetylglycopyranoses are given in Table II and summarized in Table IV. The molecules are rather complex, in that in most instances, all of the hydroxyl groups on the pyranose map are substituted, either by acetyl or other substituents. Although the accuracy of any one structure analysis falls short of the desired standard for this type of comparison, these data provide a self-consistent set of observations relating to the variation in length of the acetal C–O bonds. As shown in Table IV, the agreement between theory and experiment is good. The results for the formate and acetate agree within 0.1 pm, justifying our assumption that the substitution of the hydrogen atom by methyl has little effect on the remainder of the molecule.

For the α *configurations*, the calculations agree with the experimental data, having *b* the shortest C–O bond and *a* and *c* the longest bonds. The very long theoretical value for *c* is an anomalous result, as no C–O bonds are observed experimentally greater than 145 pm*. For the *β configuration*, the relative agreement is $b < c$ and $b_\alpha < b_\beta$, $c_\alpha > c_\beta$. Again, the largest discrepancy is in bond *c*, which is calculated too long, but the difference between the α and β distances is about 3 pm in both theory and experiment.

The conclusions concerning configurational and conformational energy-differences are more difficult to check experimentally, as the calculations refer to isolated molecules *at rest*. Our results indicate that the anomeric energy for glycopyranosyl acetates is very dependent upon the orientation of the glycosidic bond. The minimum in φ near 90° indicates that an exo-anomeric effect is present, but the lower minimum at $\varphi = 180°$ is probably steric in origin and will disappear when the additional ring atom, C-2, on the anomeric carbon atom is added to the model. For this reason, the most reasonable estimate of the anomeric energy for 1-*O*-acetyl groups is that estimated from $\varphi = 90°$ (~2 kcal/mol).

REFERENCES

1 Previous papers in this series where this methodology is used are: (*a*) G. A. JEFFREY, J. A. POPLE, AND L. RADOM, *Carbohydr. Res.*, 25 (1972) 117–131; (*b*) G. A. JEFFREY, J. A. POPLE, AND L. RADOM, *ibid.*, 38 (1974) 81–95; (*c*) G. A. JEFFREY, J. A. POPLE, J. S. BINKLEY, AND S. VISHVESHWARA, *J. Am. Chem. Soc.*, 100 (1978) 373–379; (*d*) G. A. JEFFREY AND J. H. YATES, *ibid.*, 101 (1979) 820–825; (*e*) M. D. NEWTON, G. A. JEFFREY, AND S. TAKAGI, *ibid.*, 101 (1979) 1997–2002.

2 *Quantum Chemistry Program Exchange*, No. 236. Chemistry Department, Indiana University, Bloomington, Indiana.

3 J. A. POPLE AND M. GORDON, *J. Am. Chem. Soc.*, 89 (1967) 4253–4261.

4 J. B. COLLINS, P. V. R. SCHLEYER, J. S. BINKLEY, AND J. A. POPLE, *Chem. Phys.*, 64 (1976) 5142–5151, and references therein.

5 J. M. O'GORMAN, W. SHAND, JR., AND V. SCHOMAKER, *J. Am. Chem. Soc.*, 72 (1950) 4222–4228.

6 R. F. CURL, JR., *J. Chem. Phys.*, 30 (1958) 1529–1536.

7 For a discussion of the anomeric and exo-anomeric effects, and primary references thereto, see J. F. STODDART, *Stereochemistry of Carbohydrates*, Wiley–Interscience, New York, 1971, pp. 72–92.

8 J. A. POPLE, in H. F. SCHAEFER (Ed.), *Stereochemistry of Carbohydrates*, Vol. 4, Plenum Press, New York, 1977, pp. 1–24.

9 S. SAMDEL AND H. H. SEIP, *J. Mol. Struct.*, 28 (1975) 193–203.

10 V. J. JAMES AND J. D. STEVENS, *Cryst. Struct. Commun.*, 3 (1974) 19–22.

11 P. HERPIN, R. FAMERY, J. AUGÉ, AND S. DAVID, *Acta Crystallogr., Sect. B*, 32 (1976) 215–220.

12 V. J. JAMES AND J. D. STEVENS, *Cryst. Struct. Commun.*, 6 (1977) 241–246.

13 J. OLLIS, V. J. JAMES, AND J. D. STEVENS, *Cryst. Struct. Commun.*, 4 (1975) 215–220.

14 B. KOJIĆ-PRODIĆ, V. ROGIĆ, AND Ž. RUŽIĆ-TOROŠ, *Acta Crystallogr., Sect. B*, 32 (1976) 1833–1838.

15 G. KOTHE, P. LUGER, AND H. PAULSEN, *Acta Crystallogr., Sect. B*, 32 (1976) 2710–2712.

16 V. J. JAMES AND J. D. STEVENS, *Cryst. Struct. Commun.*, 6 (1977) 119–122.

17 P. LUGER AND H. PAULSEN, *Carbohydr. Res.*, 51 (1976) 169–178.

18 V. J. JAMES AND J. D. STEVENS, *Cryst. Struct. Commun.*, 3 (1974) 187–190.

19 K. NIMGIRAWATH, V. J. JAMES, AND J. D. STEVENS, *Cryst. Struct. Commun.*, 4 (1975) 617–622.

20 V. J. JAMES, K. NIMGIRAWATH, AND J. D. STEVENS, *Cryst. Struct. Commun.*, 5 (1976) 851–856.

---

*A microwave analysis of methyl formate[6] gave a value of 143.7 ±1.0 pm for this bond, although some earlier, electron-diffraction studies of methyl formate and acetate[5] had suggested longer values of 147 and 146 pm, with uncertainties of 4 pm. In methyl vinyl ether, the C–O bond lengths are[9] 136.0(3) and 142.8(3) pm.