Combination of solid-state NMR, molecular mechanics and DFT calculations for the molecular structure determination of methyl glycoside benzoates

Łukasz Szeleszczuk · Tomasz Gubica · Sebastian Szmeja · Arkadiusz Ciesielski · Michał K. Cyrański · Dariusz M. Pisklak

Received: 4 August 2020 / Accepted: 23 September 2020 / Published online: 12 November 2020
© The Author(s) 2020

Abstract
A reliable method for molecular structure determination, excluding single-crystal X-ray diffraction (SCXRD), has been applied to six methyl glycoside tetrabenzoates. The proposed method is based on a global conformational search using molecular mechanics and subsequent DFT calculations guided by a solid-state NMR experiment. The accuracy of the applied method has been verified on three methyl glycoside benzoates for which the SCXRD analysis has been completed. It appeared that the calculated conformations of unprivileged energy could be found in the solid state. Bulky substituents (benzoates) exerted less energetically favored interactions in crystals in contrast to isolated molecules. Therefore, solid-state NMR was revealed to be an indispensable approach for choosing credible conformations from the calculated conformations.

Keywords Carbohydrates · Molecular structure · Molecular mechanics · DFT · Solid-state NMR · X-ray diffraction

Introduction
It is always worthwhile to compare the molecular structures of similar chemical species. Structural studies are the most basic studies and should precede further physicochemical analysis. The most credible molecular structures are obtained from single-crystal X-ray diffraction (SCXRD). However, this technique requires the formation of appropriate monocrystals, which can be a challenging task even for crystalline substances. Therefore, the accessibility of molecular structures derived from SCXRD is limited.

Abbreviations

| Abbreviation | Description |
|--------------|-------------|
| Bz-Me-α-Glc | Methyl α-D-glucoside tetrabenzoate |
| Bz-Me-β-Glc | Methyl β-D-glucoside tetrabenzoate |
| Bz-Me-α-Gal | Methyl α-D-galactoside tetrabenzoate |
| Bz-Me-β-Gal | Methyl β-D-galactoside tetrabenzoate |
| Bz-Me-α-Man | Methyl α-D-mannoside tetrabenzoate |
| Bz-Me-β-Man | Methyl β-D-mannoside tetrabenzoate |

Supplementary Information The online version of this article (https://doi.org/10.1007/s11224-020-01654-1) contains supplementary material, which is available to authorized users.

Tomasz Gubica
tomasz.gubica@wum.edu.pl
Łukasz Szeleszczuk
lszeleszczuk@wum.edu.pl
Sebastian Szmeja
s.szmeja@gmail.com
Arkadiusz Ciesielski
arek.ciesielski@gmail.com

Michał K. Cyrański
mkc@chem.uw.edu.pl
Dariusz M. Pisklak
dpisklak@wum.edu.pl

1 Department of Physical Chemistry, Faculty of Pharmacy, Medical University of Warsaw, Banacha 1, 02-097 Warsaw, Poland
2 Faculty of Chemistry, University of Warsaw, Pasteura 1, 02-093 Warsaw, Poland
Chemically similar compounds can surprisingly crystallize in various manners (crystals of different sizes and space groups) or even be amorphous. This can often hinder SCXRD measurements. Nevertheless, a comparison of molecular structures among similar chemical species can provide important insights into their structural features. For example, the anomeric effect was thoroughly examined in a series of methyl glycoside acetates because their molecular structures derived from SCXRD were accessible [1].

The cited [1] reported the fortunate coincidence that the molecular structures of six methyl glycoside acetates, the derivatives of both anomers of glucose, galactose, and mannose, were available. Although there were plans to expand the examination of the anomeric effect to a series of six analogous methyl glycoside esters, namely, benzoates (Chart 1), difficulties in SCXRD refinement were encountered.

As mentioned above, in solid-state studies, the SCXRD method provides the most reliable information on the structure and conformation of the studied compounds. However, when it is not possible to obtain an appropriately sized monocrystal, alternative approaches based on dynamically developing methods of molecular modeling are sought. The obtained computational data could be further verified using solid-state NMR spectroscopy.

For compounds in the solid state, intermolecular interactions play a key role in determining solid-phase conformations. Strong intermolecular interactions, when present, stabilize the crystal and mainly influence the conformation of the molecule. In this case, it is necessary to include the full crystal structure in the calculations. However, in the case of molecules in which the intermolecular interactions are limited to the weak van der Waals interactions, single-molecule calculations may allow determination of the conformation of the compound in the crystal. Derivatives of monosaccharides in which hydroxyl groups were substituted with benzoate moieties were selected to investigate this possibility. On the other hand, the introduction of such substituents eliminates the possibility of forming strong intermolecular interactions, such as hydrogen bonds, and the orientations of such bulky aromatic substituents have a primary influence on the $^{13}$C NMR chemical shifts of carbon atoms in a sugar ring.

The aim of this study was to determine to what extent the methods of conformational analysis combined with DFT

![Chart 1](image)

Bz-Me-α-Glc: $R^1=R^3=R^5=H$, $R^2=O\text{Me}$, $R^4=R^6=OBz$
Bz-Me-β-Glc: $R^2=R^3=R^5=H$, $R^1=O\text{Me}$, $R^4=R^6=OBz$
Bz-Me-β-Gal: $R^1=R^3=R^6=H$, $R^2=O\text{Me}$, $R^4=R^5=OBz$
Bz-Me-α-Man: $R^1=R^4=R^5=H$, $R^2=O\text{Me}$, $R^3=R^6=OBz$
Bz-Me-β-Man: $R^2=R^4=R^5=H$, $R^1=O\text{Me}$, $R^3=R^6=OBz$

![Table 1](image)

| Unit cell dimensions | Bz-Me-β-Gal | Bz-Me-α-Man | Bz-Me-β-Man |
|----------------------|-------------|-------------|-------------|
| Exp. *a* (Å)         | 10.0249(3)  | 9.3038(4)   | 12.2653(6)  |
| Exp. *b* (Å)         | 16.2238(5)  | 16.0697(6)  | 9.9444(6)   |
| Exp. *c* (Å)         | 18.3405(6)  | 20.2106(8)  | 12.5476(7)  |
| α (°)                | 90          | 90          | 90          |
| β (°)                | 90          | 90          | 90          |
| γ (°)                | 90          | 90          | 90          |

![Table 2](image)

|                  | Exp. CASTEP-optimized | Gaussian-optimized | Exp. CASTEP-optimized | Gaussian-optimized |
|------------------|-----------------------|--------------------|-----------------------|--------------------|
| Bz-Me-β-Gal      | 0.113                 | 0.702              | 0.702                 | 0.731              |
| Bz-Me-α-Man      | 0.102                 | 0.499              | 0.518                 | 0.499              |
| Bz-Me-β-Man      | 0.114                 | 0.794              | 0.793                 | 0.794              |
calculations and verification of the obtained data based on $^{13}$C CP/MAS NMR spectroscopy will allow reconstruction of the crystal conformation of the analyzed compounds.

Results and discussion

X-ray diffraction

Although the six abovementioned methyl glycoside benzoates differ only in the configurations on separate carbon atoms, major differences among their ability to crystallize were observed. While Bz-Me-β-Gal, Bz-Me-α-Man, and Bz-Me-β-Man were found to form crystals suitable for SCXRD measurements, crystals of the glucosides (Bz-Me-α-Glc and Bz-Me-β-Glc) were not of a sufficient size for this type of crystallographic study, whereas Bz-Me-α-Gal was found to be amorphous.

Bz-Me-β-Gal and Bz-Me-β-Man were found to crystallize in the orthorhombic crystal system and $P2_1 2_1 2_1$ space group, while the crystals of Bz-Me-α-Man belonged to the monoclinic crystal system and $P2_1$ space group (Table S1).

Table 3 Experimental and calculated (CASTEP- and Gaussian-optimized) $^{13}$C NMR chemical shifts (ppm) of methyl glycoside benzoates with the coefficient of determination ($R^2$) between the experimental and calculated values

| Carbon atom | Bz-Me-β-Gal | Bz-Me-α-Man | Bz-Me-β-Man |
|-------------|-------------|-------------|-------------|
|             | Exp.         | CASTEP-optimized | Gaussian-optimized | Exp.         | CASTEP-optimized | Gaussian-optimized | Exp.         | CASTEP-optimized | Gaussian-optimized |
| C1          | 102.1        | 104.53       | 97.8        | 104.30       | 100.58       | 100.16       | 99.1        | 102.25       | 99.00       |
| C2          | 70.3         | 70.16        | 71.2        | 68.84        | 69.07        | 69.74        | 70.7        | 70.76        | 72.16       |
| C3          | 73.7         | 74.83        | 70.2        | 74.54        | 66.0         | 64.75        | 66.05       | 72.12        | 73.32       |
| C4          | 68.3         | 67.39        | 68.50       | 68.84        | 65.1         | 65.76        | 64.1        | 63.84        | 64.73       |
| C5          | 70.3         | 70.83        | 73.62       | 68.9         | 68.64        | 71.15        | 64.1        | 63.84        | 64.73       |
| C6          | 63.6         | 60.64        | 65.28       | 60.2         | 58.37        | 59.40        | 55.7        | 54.43        | 55.56       |
| OMe         | 58.7         | 58.22        | 56.28       | 55.9         | 55.19        | 53.63        | 55.7        | 54.43        | 55.56       |
| $R^2$       | 0.9954       | 0.9885       | 0.9982      | 0.9935       | 0.9961       | 0.9912       |

Fig. 1 Superimposed Bz-Me-α-Man conformations: experimental (blue) and CASTEP-optimized (red)

Fig. 2 Superimposed Bz-Me-α-Man conformations: experimental (blue) and Gaussian-optimized (red)
in those three cases, only one molecule was found to form a unit cell (Z’ = 1). The three glycosides that could not be measured by SCXRD were subjected to powder X-ray diffraction (PXRD) measurements instead (Figs. S25–S28), which proved the crystallinity of Bz-Me-α-Glc and Bz-Me-β-Glc and the amorphousness of Bz-Me-α-Man.

Calculations for Bz-Me-β-Gal, Bz-Me-α-Man, and Bz-Me-β-Man

For half of the studied glycosides, crystal structures were obtained and used to validate the proposed method. The first computational aim of this study was to analyze the differences between the experimental, CASTEP-optimized and Gaussian-optimized conformations of methyl glycoside benzoates. To achieve this, the experimental crystal structures of Bz-Me-β-Gal, Bz-Me-α-Man, and Bz-Me-β-Man were used as the initial geometries for CASTEP geometry optimization, including the optimization of the unit cell dimensions. Simultaneously, single molecules were extracted from the crystal structures and optimized in Gaussian using the same exchange-correlation DFT functional (PBE) and semiempirical dispersion correction (Grimme). The optimized structures were then used for calculations of NMR parameters using GIPAW (for CASTEP calculations) and GIAO (for Gaussian calculations) approaches. Calculations of the NMR shielding constants of the experimental (nonoptimized) crystal structures were not performed since it has been proven multiple times that at least the optimization of hydrogen atoms is mandatory to obtain accurate results, even for structures of very high quality [2, 3].

To assess how optimization affects the molecular structure, a comparison of the experimental and optimized unit cell dimensions is presented in Table 1. Furthermore, the calculated RMSD values are provided in Table 2. The calculated and experimental 13C NMR chemical shifts are compared in Table 3.

The CASTEP-optimized unit cell dimensions were found to be in very good agreement with the corresponding experimental dimensions. The slight decrease in cell lengths observed after geometry optimization can be explained by differences between the experimental temperature (130 K) and the calculation conditions in which thermal motions have been neglected, as the geometry optimization was performed at 0 K.

The calculated RMSD values proved that the differences between the CASTEP-optimized and experimental conformations in all three studied cases were very small (Fig. 1) and limited mostly to a slight decrease in the C–H bond lengths during the calculations due to differences between the computational (0 K) and experimental (130 K) temperatures, similar to what was observed for the unit cell lengths (Table 1).

Higher RMSD values were obtained between the experimental and Gaussian-optimized conformations. This was

Table 4 Differences between experimentally obtained and calculated (Conformers-calculated Gaussian-optimized) 13C NMR chemical shifts (ppm) for Bz-Me-β-Gal with the coefficient of determination (R²) between experimental and calculated chemical shifts; relative energies (E, kcal/mol) of the optimized conformations calculated using Conformers and Gaussian software

| Conformation | G1  | G2  | G3  | EGO | G4  | G5  | G6  | G7  | G8  | G9  | G10 | G11 |
|--------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| C1           | -2.31 | -1.11 | 0.33 | -2.20 | -1.85 | -1.12 | 1.93 | -0.74 | -3.17 | 3.00 |
| C2           | -0.94 | -0.99 | 2.82 | 1.06 | 1.07 | 0.81 | 0.22 | 0.18 | 1.58 | 0.56 |
| C3           | 0.90 | 0.00 | 0.90 | 0.90 | 0.90 | 0.90 | 0.90 | 0.90 | 0.90 | 0.90 |
| C4           | 0.32 | 0.00 | 2.82 | 1.06 | 1.07 | 0.81 | 0.22 | 0.18 | 1.58 | 0.56 |
| C5           | 0.90 | 0.00 | 2.82 | 1.06 | 1.07 | 0.81 | 0.22 | 0.18 | 1.58 | 0.56 |
| C6           | -0.45 | -0.39 | -1.40 | -1.68 | -2.03 | 1.70 | 2.61 | 2.52 | -4.06 | -3.00 |
| G5e          | 3.31 | 3.86 | 3.97 | 3.42 | 3.59 | 3.98 | 3.00 | 2.45 | 1.00 | 0.94 |
| Etot         | 0.70 | 0.93 | 1.04 | 0.82 | 0.96 | 1.15 | 0.97 | 0.85 | 0.89 | 0.89 |
| Etot         | 0.00 | -0.11 | 1.16 | 0.82 | 0.96 | 1.15 | 0.97 | 0.85 | 0.89 | 0.89 |

Fig. 3 Superimposed Bz-Me-β-Gal conformations: experimental (blue) and Gaussian-optimized G7 (red)
caused mostly by the aperiodicity of the Gaussian calculations, in contrast to the CASTEP calculations, and the resulting inability to include the intermolecular interactions. However, in the case of the Gaussian calculations, the molecules preserved their initial experimental conformations after optimization (Fig. 2).

Subsequently, the optimized structures were used for the NMR parameter calculations, and the obtained results are presented in Table 3. It should be noted here that only those carbon atoms that formed the sugar unit and aglycon (methoxy group) were chosen for comparison. This choice was made due to the overlap of the $^{13}$C CP/MAS NMR signals from aromatic substituents, which made the precise determination of their chemical shift values impossible (Figs. S19–S24).

The calculated and experimental chemical shift values were found to be in very good agreement. Slightly better accuracy was observed for the CASTEP results due to a higher similarity between the experimental and CASTEP-optimized conformations, as proven by the lower values of RMSD (Table 2) and the periodicity of the CASTEP calculations. However, even for the Gaussian calculations, the largest differences between the calculated and experimental NMR chemical shifts did not exceed 3 ppm.

The three refined crystal structures (Bz-Me-β-Gal, Bz-Me-α-Man, and Bz-Me-β-Man) were then used to test the chosen method of conformational search using the Conformers module in Materials Studio 2017, as described in “Conformers calculations.”

**Bz-Me-β-Gal**

In the case of Bz-Me-β-Gal, the experimental solid-state conformation (EGO) was found in the group of ten lowest calculated conformations and ranked as the 4th best with energy higher by approximately 2.5 kcal/mol than the lowest energy conformation (G5) (Table 4). After optimization at the DFT level, only minor differences in the stability order of the conformations were observed. This proves the reliability of the Conformers calculations at the MM level, at least in terms of energy calculations and their application to identify the lowest energy conformation. However, this does not explain why the experimental solid-state conformation (EGO) was found to be merely at the 4th position in terms of energy. However, after an analysis of the NMR calculations results, EGO was found to be the conformation for which the highest coefficient of determination ($R^2$) was obtained. Furthermore, it was the only conformation for which the highest absolute difference between the calculated and experimental NMR chemical shifts did not exceed 3 ppm. A comparison between the experimental and calculated conformations revealed the reasons for both the lowest energies of G5, G6, and G7 compared with EGO as well as some of the observed differences between the calculated and experimental NMR chemical shifts. For example (Fig. 3), in the G7 conformation, a completely different orientation of the benzoate substituent at C6 than in EGO was observed. In the G7 conformation, the aromatic rings of the benzoates at C6 and C4 were found to participate in π–π stacking, which explained why this conformation was the one with the lowest energy. Such interaction was not observed in the experimental conformation. Furthermore, in G7, the H–C1–O–Me moiety was found to be anti-oriented, while in the EGO conformation, this moiety was in the gauche position. This reorientation resulted in a very large difference between the calculated and experimental chemical shifts for the OMe carbon atom.

---

**Table 6** Differences between the experimentally obtained and calculated (Conformers-calculated Gaussian-optimized) $^{13}$C NMR chemical shifts (ppm) for Bz-Me-α-Man with the coefficient of determination ($R^2$) between experimental and calculated chemical shifts; relative energies ($E$, kcal/mol) of the optimized conformations calculated using Conformers and Gaussian software

| Conformation | G6 | EGO | G7 | G8 | EConformers | R2
|--------------|----|-----|----|----|-------------|----
| C1           | -2.13 | -3.47 | -1.53 | -2.83 | 1.54 | 0.9869 |
| C2           | -1.31 | -0.36 | -2.30 | -3.75 | 1.54 | 0.9802 |
| C3           | -2.30 | -2.04 | -1.59 | -2.75 | 1.54 | 0.80   |
| C4           | -2.53 | -2.23 | -4.83 | -1.74 | 1.54 | 0.9886 |
| C5           | -0.70 | -0.18 | -5.40 | -3.23 | 1.54 | 0.9754 |
| C6           | 0.12 | 0.63 | -4.83 | -1.49 | 1.54 | 0.9754 |
| OMe          | 2.28 | 2.27 | 2.29 | 2.36 | 1.54 | 0.9754 |
| EGO          | -2.31 | -0.81 | -0.80 | -0.80 | 1.54 | 0.9754 |

---

**Fig. 4** Superimposed Bz-Me-α-Man conformations: experimental (blue) and Gaussian-optimized G8 (red)
Calculations for Bz-Me-α-Glc, Bz-Me-β-Glc, and Bz-Me-α-Gal

In the case of Bz-Me-α-Glc, the obtained results were similar to those of Bz-Me-β-Gal. Two low-energy conformations (G2 and G1) that, after the DFT optimization, converged into one minimum were characterized by a very large difference between the calculated and experimental $^{13}$C NMR shifts for C6 (Table 8). Therefore, it is very likely that this conformation is far from the experimental conformation. However, another low-energy conformation (G4) was characterized by a very large $R^2$ value, with the absolute differences between the experimental and calculated $^{13}$C NMR chemical shifts not exceeding 2 ppm. Therefore, it is very likely that the G4 conformation is actually present in the crystal structure.

Almost the same results obtained for Bz-Me-α-Glc were obtained for its anomer, that is, Bz-Me-β-Glc (Table 9). Again, the two lowest energy conformations identified in the Conformers search (G1 and G2) after Gaussian optimization converged to the same minimum, and C3′ (second molecules) was observed, which could not have been formed in the single-molecule calculations performed in Conformers and Gaussian.

### Table 7

| Conformation | EConformers | ENMR | EGAUSSIAN | R2  | RMSD | EGAUSSIAN |
|--------------|-------------|------|-----------|-----|------|-----------|
| C1           | -3.26       | -3.48| -3.67     | 1.53| -3.38| 0.30      |
| C2           | -2.65       | -2.17| -2.11     | 1.47| -2.15| 0.45      |
| C3           | -1.46       | -3.34| -2.88     | 1.58| -1.39| 0.19      |
| C4           | -1.42       | -3.56| -3.50     | 1.42| -3.28| 0.50      |
| C5           | -0.63       | -2.02| -1.41     | 1.32| -1.41| 0.01      |
| C6           | -0.14       | -3.55| -3.46     | 1.43| -1.41| 0.01      |
| C7           | 0.00        | 0.54 | 0.54      | 0.43| 0.19 | 0.01      |

**Bz-Me-α-Man**

In the case of Bz-Me-α-Man, the experimental conformation was ranked as the 2nd best during the conformational search. However, after geometry optimization at the DFT level, conformations G6 and EGO converged into one minimum, which was proven based on the very similar values of the calculated $^{13}$C NMR chemical shifts (Table 3), RMSD (Table 5), and $E_\text{GAUSSIAN}$ (Table 6).

Surprisingly, another very low-energy conformation (G8) was found, for which a very weak correlation between the calculated and experimental $^{13}$C NMR chemical shifts was observed. This was associated with a reorientation of the C6 benzoate residue. While in the experimental conformation, a C–H···π interaction was present between the benzoates at C2 and C6; in G8, this interaction was not observed (Fig. 4). However, in G8, π–π stacking was observed between the benzoates at C4 and C6.

**Bz-Me-β-Man**

In the case of Bz-Me-β-Man, the experimental conformation was ranked as the best during the conformational search. Furthermore, it was also characterized by the lowest differences and highest $R^2$ value between the experimentally obtained and calculated $^{13}$C NMR chemical shifts (Table 3), as well as the lowest $E_\text{GAUSSIAN}$ (Table 7). Again, as in the case of Bz-Me-α-Man, the low-energy conformations G2 and G3, which after Gaussian optimization converged into one minimum, were characterized by the least accurate NMR parameter calculation results. Furthermore, while a C–H···π interaction between the benzoates at C2 and C6 and π–π stacking between the benzoates at C3 and C4 were formed in G2, those interactions were not present in the experimental solid-state conformation (Fig. 5). However, in the refined crystal structure of Bz-Me-β-Man, intermolecular π–π stacking between benzoates at C2 (first molecule) and C3′ (second molecule) was observed, which could not have been formed in the single-molecule calculations performed in Conformers and Gaussian.
and the 3rd best conformation in terms of both $E_{\text{Conformers}}$ and $E_{\text{Gaussian}}$ was characterized by the best fitting of the calculated and experimental NMR results. Although a relatively large absolute difference between the experimental and calculated $^{13}$C NMR shifts for C4 has been observed, it should be noted that some of the differences between those shifts for glycosides with solved crystal structures (Bz-Me-$\alpha$-Gal, Bz-Me-$\alpha$-Man, and Bz-Me-$\beta$-Man) were also in the range of 3 to 4 ppm.

In the case of Bz-Me-$\alpha$-Gal, it was impossible to compare the calculated and experimental chemical shifts due to the very broad and overlapping signals in its $^{13}$C CP/MAS NMR spectrum (Fig. S21). Therefore, for this glycoside, the calculated chemical shifts, instead of differences between the experimental and calculated shifts, are presented in Table 10. It is worth noting that for this glycoside, the lowest differences between the calculated shifts for the same carbon atoms for different conformations among all the studied compounds were observed. Furthermore, groups of very similar energy yet different conformations have been formed, such as G4, G5, and G7 and G3, G1, and G8. In the first group, it is clear that the optimized conformations are different, which was determined by comparison of the calculated NMR shifts for C4; however, the differences between their $E_{\text{Gaussian}}$ values did not exceed 0.6 kcal/mol. In the second group, the differences between the $^{13}$C NMR shifts for C6 prove their versatility, yet again, the differences between the DFT energies calculated for these conformations did not exceed 0.6 kcal/mol. These observations may stem from the amorphousness of Bz-Me-$\alpha$-Gal, resulting from the multiple conformations of similar energy and the resulting inability to form a stable crystal due to entropic reasons.

### Conclusions

In this paper, the considerable potential of a combined conformational search followed by geometry optimization at the DFT level and subsequent verification using solid-state NMR analysis ($^{13}$C CP/MAS NMR) for molecular structure evaluation has been demonstrated. The method of structure determination applied in this study was first verified on three compounds for which the SCXRD structures were obtained for the first time. Then, the same approach was used to identify conformations for three substituted sugars for which SCXRD analysis was not possible.

While in the case of saccharides with free hydroxyl groups, the solid-state molecular structure is determined by strong intermolecular hydrogen bonds, we have demonstrated that, even for substituted sugars forming weak intermolecular interactions, energetically favored conformations are not always the ones present in the crystal lattice. Fortunately, solid-state NMR chemical shifts of sugar carbon atoms proved to be extremely sensitive to different conformations of bulky and highly anisotropic substituents such as benzoates. Thanks to the screening of the correlations between theoretical and experimental chemical shifts, reliable molecular structures were selected even when SCXRD data were unavailable.

To conclude, in this article, a combined computational-experimental structural analysis and prediction method was presented. It has been shown that such an approach can...
accurately predict the conformation of solid organics, yet there are some limitations. The requirements of this method include the absence of strong intermolecular interactions and the presence of only one molecule in the unit cell ($Z' = 1$). In addition, prior registration of the solid-state $^{13}$C NMR spectra of the studied compounds is mandatory for this method to be applied.

**Experimental**

**General procedures**

Reagents were purchased from POCh (Gliwice, Poland) and used without further purification. Melting points were determined on an Electrothermal IA9200 apparatus with open capillary tubes and are uncorrected. Optical rotations were measured on a Perkin-Elmer 241 polarimeter. $^1$H and $^{13}$C NMR spectra in acetone-$d_6$ solutions were recorded at 300 and 75 MHz, respectively, on a Varian VNMRS-300 spectrometer, and 2D experiments (COSY, HSQC, and HMBC) were performed using standard Varian software.

**Materials**

Bz-Me-α-Glc, Bz-Me-β-Glc, Bz-Me-α-Gal, Bz-Me-β-Gal, and Bz-Me-α-Man were synthesized according to the procedures described in Ref. [4], while Bz-Me-β-Man was obtained according to the synthesis described in Ref. [5]. Basic analytical data for the synthesized compounds are collected in Table 11. The chemical shifts and $^1$H coupling constants for $^1$H and $^{13}$C NMR spectra in acetone-$d_6$ solutions are given in Table 12.

### Solid-state NMR spectroscopy ($^{13}$C CP/MAS NMR)

Cross-polarization (CP) magic angle spinning (MAS) $^{13}$C NMR spectra were recorded at 100.16 MHz on a Bruker DSX-400 spectrometer. Powdered samples were spun at 10 kHz in a 4 mm ZrO$_2$ rotor, with a contact time of 2 ms, a repetition time of 10 s, and a spectral width of 41 kHz for the accumulation of 256 (Bz-Me-α-Glc and Bz-Me-β-Glc), 232 (Bz-Me-α-Gal), or 400 (Bz-Me-β-Gal, Bz-Me-α-Man, and Bz-Me-β-Man) scans. Chemical shifts were calibrated indirectly through the glycine C=O signal recorded at 176.5 ppm relative to TMS.

### Molecular modeling

**CASTEP calculations**

The density functional theory (DFT) geometry optimization and NMR shielding constant calculations were carried out with the CASTEP program [8] implemented in the Materials Studio 2017 software [9] using the plane wave pseudopotential formalism and the Perdew-Burke-Ernzerhof (PBE) exchange-correlation functional [10], defined within the generalized gradient approximation (GGA) with the Grimme method for the dispersion correction (DFT-D) [11]. The calculations were performed with ultrasoft pseudopotentials calculated on the fly.

Geometry optimization was carried out using the Broyden-Fletcher-Goldfarb-Shanno (BFGS) [12] optimization scheme and the smart method for finite basis set correction. The quality of calculations was set to ultrafine as implemented in the CASTEP standards. The convergence criteria were set at $5 \times 10^{-6}$ eV/atom for the energy, $1 \times 10^{-2}$ eV/Å for the interatomic forces, $2 \times 10^{-3}$ GPa for the stresses, and $5 \times 10^{-4}$ Å for the

| Parameters | d-Glucoside | d-Galactoside | d-Mannoside |
|------------|-------------|---------------|-------------|
| mp (°C) | α | β | α | β | α | β |
| (EtOH) | 103–104.5 | 160.5–161.5 | 73–75 | 131.5–132.5 | 135–136 | 147–148 |
| (EtOH + hexane) | | | (EtOH) | | (MeOH) | |
| (MeOH) | 101.5–102.5 | | | | | |
| (solvent) | | | (EtOH) | | (MeOH) | |
| lit. mp (°C) | 103–105 | 158–160 | 63–64 | 51–52 | 136–137 | 148–149 |
| (EtOH) [6] | | (EtOH) [4] | (EtOH) [4] | (EtOH) [4] | (MeOH) [4] | (MeOH) [5] |
| 73–74 | | | | | | |
| (MeOH) [4] | | | | | | |
| [α]$_D$ | α | + 81 | + 143 | + 103 | − 66 | − 147 |
| (c 1.0, CHCl$_3$) | (c 1.0, CHCl$_3$) | (c 1.0, CHCl$_3$) | (c 1.0, CHCl$_3$) | (c 1.0, CHCl$_3$) | (c 1.0, CHCl$_3$) | (c 1.0, CHCl$_3$) |
| lit. [α]$_D$ | 88.4 | 27.6 | 140.2 | 97.7 | − 66.9 | − 153 |
| (c 0.5, CHCl$_3$) [4] | (c 1.0, CHCl$_3$) [7] | (c 1.0, CHCl$_3$) [4] | (c 1.0, CHCl$_3$) [4] | (c 1.0, CHCl$_3$) [4] | (c 2.5, CHCl$_3$) [4] | (c 1.1, CHCl$_3$) [5] |
displacements. The fixed basis set quality method for the cell optimization calculations and the $5 \times 10^{-7}$ eV/atom tolerance for SCF were used. The kinetic energy cutoff for the plane waves was set to 630 eV. Brillouin zone integration was performed using discrete $3 \times 2 \times 1$ (Bz-Me-β-Man) and $2 \times 3 \times 2$ (Bz-Me-α-Man) Monkhorst-Pack k-point sampling for a primitive cell.

The optimized structures were then used for the NMR parameter calculations using the gauge including projector augmented wave (GIPAW) method of Pickard and Mauri [13]. To facilitate comparison of the theoretical and experimental data, the calculated shielding constants ($\sigma_{\text{iso}}$) were converted to chemical shifts ($\delta$) using the following equation: $\delta = \sigma_{\text{TMS}} - \sigma_{\text{iso}}$, where $\sigma_{\text{TMS}}$ stands for the shielding constant of the tetramethylsilane (TMS) carbon atom calculated at the same level of theory.

### Conformers calculations

A conformational search of the studied molecules was performed using the Conformers module of Materials Studio 2017 by applying the Metropolis Monte Carlo algorithm-based Boltzmann jump search method. A total of $10^3$ conformers per molecule were generated by applying 500 perturbations per jump within the torsion angle window adjusted to accept 50% of the generated conformers. Geometry optimization was performed on each conformer at the molecular mechanics level by the provided Smart algorithm (COMPASS force field, ultrafine quality and maximum $5 \times 10^4$ iterations). Then, the ten lowest energy conformations for each molecule were kept for the Gaussian DFT calculations.

### Gaussian calculations

The DFT geometry optimization and NMR shielding constant calculations were carried out with the Gaussian 16 program [14]. The molecules either from the asymmetric units of the X-ray structures or resulting from the Conformers search (as described in the “Conformers calculations” section) were taken as the starting geometries to optimize all atom positions at the PBEPBE/6-311G(d,p) with Grimme dispersion correction level of theory. The resulting geometries were used for calculation of the NMR shielding constants using the GIAO approach at the same level of theory. To compare the theoretical and experimental data, the calculated chemical shielding constants ($\sigma_{\text{iso}}$) were converted to chemical shifts ($\delta$) using the following equation: $\delta = \sigma_{\text{TMS}} - \sigma_{\text{iso}}$, where $\sigma_{\text{TMS}}$ stands for the shielding constant of the tetramethylsilane (TMS) carbon atom calculated at the same level of theory.

### Table 12

Chemical shifts (ppm) and $^1H$ coupling constants (Hz) for $^1H$ and $^{13}C$ NMR spectra in acetone-$d_6$ solutions for methyl glycoside tetrabenzoates$^a$

| Spectral data | D-Glucoside | D-Galactoside | D-Mannoside |
|--------------|-------------|---------------|-------------|
| $\alpha$     | $\beta$     | $\alpha$      | $\beta$     |
| H1 ($\delta_{\text{H1-H2}}$) | 5.35 d (3.6) | 5.13 d (7.8) | 5.43 d (3.6) | 5.13 d (7.5) | 5.13 d (1.5) | 5.25 d (0.9) |
| H2 ($\delta_{\text{H2-H3}}$) | 5.46 dd (10.2) | 5.54 dd (9.6) | 5.79 dd (10.5) | 5.83–5.80 m | 5.77 dd (3.3) | 5.99 dd (3.3) |
| H3 ($\delta_{\text{H3-H4}}$) | 6.22 t (9.9) | 6.04 t (9.6) | 6.06 dd (3.3) | 5.83–5.80 m (2.4) | 5.90 dd (10.2) | 5.83 dd (10.2) |
| H4 ($\delta_{\text{H4-H1}}$) | 5.84 t (9.5) | 5.79 t (9.6) | 6.16 dd (0.9) | 6.09 d (2.4) | 6.24–6.18 m | 6.14 t (9.9) |
| H5 ($\delta_{\text{H5-H6a-H6b}}$) | 4.72–4.56 m | 4.53 ddd (3.0, 4.5) | 4.84 pt (6.9, 6.0) | 4.77–4.68 m | 4.64–4.58 m | 4.49 ddd (2.7, 3.9) |
| H6a ($\delta_{\text{H6a-H6b}}$) | 4.72–4.56 m | 4.72 dd (12.0) | 4.65 dd (11.4) | 4.77–4.68 m | 4.82–4.76 m | 4.83 dd (12.3) |
| H6b | 4.72–4.56 m | 4.61 dd | 4.53 dd | 4.59–4.50 m | 4.64–4.58 m | 4.65 dd |
| OCH$_3$ | 3.58 s | 3.56 s | 3.59 s | 3.60 s | 3.63 s | 3.57 s |
| COC$_6$H$_5$ | 8.11–7.36 | 8.10–7.36 | 8.16–7.30 | 8.13–7.31 | 8.18–7.33 | 8.19–7.32 |
| C1 | 97.05 | 101.57 | 98.58 | 101.81 | 98.67 | 100.70 |
| C2 | 71.84 | 72.05 | 70.27 | 72.03 | 70.18 | 70.97 |
| C3 | 70.71 | 73.43 | 69.50 | 70.07 | 70.58 | 73.32 |
| C4 | 69.54 | 69.80 | 70.61 | 68.81 | 66.58 | 67.96 |
| C5 | 67.66 | 71.81 | 67.77 | 70.96 | 68.67 | 73.03 |
| C6 | 62.93 | 62.88 | 63.47 | 62.05 | 62.55 | 63.54 |
| OCH$_3$ | 54.98 | 56.19 | 55.99 | 56.12 | 54.87 | 57.42 |
| COC$_6$H$_5$ | 165.56, 165.48, 165.54, 165.29, 165.43, 165.37, 165.45, 165.13, 165.03, 164.91, 166.10, 165.88 | 166.45, 166.15, 166.45, 166.15, 166.03, 165.45, 166.15, 166.45, 166.15, 166.03, 165.45, 166.15 |
| COC$_6$H$_5$ | 133.55–128.46 | 133.50–128.45 | 134.61–129.36 | 133.66–128.41 | 133.43–128.41 | 134.92–129.41 |

$^a$ Spectral data originally assigned (in the case of Bz-Me-β-Man for the first time), coherent (except for Bz-Me-β-Gal) with appropriate data from Ref. [4].
Presentation of calculated data

In Tables 1, 2, and 3, the description “CASTEP-optimized” refers to the results obtained for the experimental crystal structures (Bz-Me-β-Gal, Bz-Me-α-Man, and Bz-Me-β-Man) subjected to the DFT calculations in CASTEP, as described in “CASTEP calculations.”

In Tables 2, 3, 4, 5, 6, 7, 8, 9, and 10, the description “Gaussian-optimized” refers to the results obtained for the single molecules either extracted from the experimental crystal structures (Bz-Me-β-Gal, Bz-Me-α-Man, and Bz-Me-β-Man) or found in the Conformers search (Bz-Me-α-Glc, Bz-Me-β-Glc, Bz-Me-α-Gal, Bz-Me-β-Gal, Bz-Me-α-Man, and Bz-Me-β-Man) as described in “Conformers calculations” and, in both cases, subjected to DFT calculations in Gaussian, as described in “Gaussian calculations.”

In Tables 4, 5, 6, 7, 8, 9, and 10, the abbreviations G1–G10 refer to the ten lowest energy conformations found during the Conformers search as described in “Conformers calculations,” while EGO refers to the experimental conformation for which the geometry optimization was performed in the same way as for G1–G10, first at the molecular mechanics level (as described in “Conformers calculations”) and then at the DFT level in Gaussian, as described in “Gaussian calculations.” Furthermore, to facilitate reception of the results, the relative energy values are presented in Tables 4, 5, 6, 7, 8, 9, and 10, with the zero energy assigned to the lowest value in each case.

In Tables 4 and 6, 7, 8, 9, and 10, the abbreviation “E Conformers” refers to the energy calculated for the chosen conformer in the Conformers module at the molecular mechanics level (force field COMPASS), as described in “Conformers calculations,” while “E Gaussian” refers to the energy calculated for the chosen conformer in Gaussian 16 at the DFT level, as described in “Gaussian calculations.”

For comparison of the conformations of structures under investigation, the direct root-mean-square deviation (RMSD) of atomic positions for single molecules was calculated (Tables 2 and 5) according to the equation:

$$ RMSD = \sqrt{\frac{\sum d^2}{n}}. $$

where $d$ is the distance between each of the $n$ pairs of equivalent atoms in two optimally superposed structures.

To facilitate the analysis of the data in Tables 4 and 6, 7, 8, 9, and 10, a 3-color scale was applied. In this scale, the 50th percentile (midpoint) was calculated, and the cell that holds this value was colored yellow. The cell that holds the minimum value was colored blue, and the cell that holds the maximum value was colored red. All other cells were colored proportionally. For this coloring, the data within Tables 4 and 6, 7, 8, and 9 were grouped within the same calculated parameter ($E_{\text{Conformers}}$; $E_{\text{Gaussian}}$; $R^2$; NMR chemical shift differences) without splitting for individual atoms in the case of NMR data.

Single-crystal X-ray diffraction

The crystal structure determinations of Bz-Me-β-Gal, Bz-Me-α-Man, and Bz-Me-β-Man were carried out by the X-ray diffraction method on a single crystal. In all those cases, measurements were performed at 130 K using MoKα radiation on a Bruker D8 Venture diffractometer with Photon II CPAD. The subsequent data treatment and reduction were performed according to procedures described in Ref. [15]. All structures were solved and refined using the SHELX package [16]. The refinement was based on $F^2$ for all reflections except for those with very negative $F^2$. The weighted $R$ factor, $wR$, and all goodness-of-fit $S$ values were based on $F^2$. The nonhydrogen atoms were refined anisotropically. The hydrogen atoms were located from a difference map and were refined isotropically. Scattering factors, in their analytical form, were taken from the International Tables for Crystallography [17]. The crystal data and structure refinement parameters for the obtained crystals are collected in Table S1 (Supplementary material).

Acknowledgments The authors thank Prof. A. Temeriusz for proofreading the manuscript and discussions. The authors also thank Ms. V. Kowalska for recording NMR spectra in solution.

Compliance with ethical standards

Conflict of interest The authors declare that there are no conflicts of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.
References

1. Gubica T, Zimniak A, Szeleszczuk Ł, Dąbrowska K, Cyrański MK, Kańska M (2019) Influence of acetylation on anomeric effect in methyl glycosides. Molecular Physics 117(3):349–358. https://doi.org/10.1080/00268976.2018.1512723

2. Szeleszczuk Ł, Pisklak DM, Zieleńska-Pisklak M (2018) Does the choice of the crystal structure influence the results of the periodic DFT calculations? A case of glycine alpha polymorph GIPAW NMR parameters computations. J Comput Chem 39(14):853–861. https://doi.org/10.1002/jcc.25161

3. Szeleszczuk Ł, Pisklak DM, Zielińska-Pisklak M (2016) Effects of structural differences on the NMR chemical shifts in cinnamic acid derivatives: comparison of GIAO and GIPAW calculations. Chemical Physics Letters 653:35–41. https://doi.org/10.1016/j.cplett.2016.04.075

4. Esmurziev A, Simic N, Sundby E, Hoff BH (2009) H-1 and C-13 NMR data of methyl tetra-O-benzoyl-D-pyranosides in acetone-d(6). Magn Reson Chem 47(5):449–452. https://doi.org/10.1002/mrc.2400

5. Ness RK, Fletcher HG, Hudson CS (1950) The reaction of 2,3,4,6-tetrabenzoyl-alpha-D-glucopyranosyl bromide and 2,3,4,6-tetrabenzoyl-alpha-D-mannopyranosyl bromide with methanol - certain benzoylated derivatives of D-glucose and D-mannose. Journal of the American Chemical Society 72(5):2200–2205. https://doi.org/10.1021/ja01161a091

6. Willard JJ, Pacsu E (1960) New method of removing xanthate groups from carbohydrates - chemical structure of methyl alpha-D-glucopyranoside monoxanthate. Journal of the American Chemical Society 82(16):4347–4350. https://doi.org/10.1021/ja01501a056

7. Willard JJ, Brimacombe JS, Brueton RP (1964) Synthesis OF 3-O-((benzythio)carbonyl)-beta-D-glucopyranose + methyl2,4,6-tri-O-benzoyl-alpha + beta-D-glucopyranosides. Can J Chem-Rev Can Chim 42(11):2560–&

8. Clark SJ, Segall MD, Pickard CJ, Hasnip PJ, Probert MJ, Reifson K, Payne MC (2005) First principles methods using CASTEP. Z Kristall 220(5-6):567–570. https://doi.org/10.1524/zkri.220.5.567.65075

9. BIOVIA Materials Studio. http://accelrys.com/products/collaborative-science/biovia-materials-studio/.

10. Perdew JP, Burke K, Ernzerhof M (1996) Generalized gradient approximation made simple. Phys Rev Lett 77(18):3865–3868. https://doi.org/10.1103/PhysRevLett.77.3865

11. Grimme S (2006) Semiempirical GGA-type density functional constructed with a long-range dispersion correction. J Comput Chem 27(15):1787–1799. https://doi.org/10.1002/jcc.20495

12. Pfrommer BG, Cote M, Louie SG, Cohen ML (1997) Relaxation of crystals with the quasi-Newton method. J Comput Phys 131(1):233–240. https://doi.org/10.1006/jcph.1996.5612

13. Pickard CJ, Mauri F (2001) All-electron magnetic response with pseudopotentials: NMR chemical shifts. Phys Rev B 63(24):13. https://doi.org/10.1103/PhysRevB.63.245101

14. Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Scalmani G, Barone V, Petersson GA, Nakatsuji H, Li X, Caricato M, Marenich AV, Bloino J, Janesko BG, Gomperts R, Mennucci B, Hratchian HP, Ortiz JV, Izmaylov AF, Sonnenberg JL, Williams, Ding F, Lipparrini F, Egid F, Goings J, Peng B, Petrone A, Henderson T, Ranasinghe D, Zakrzewski VG, Gao J, Rega N, Zheng G, Liang W, Hada M, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao A, Nakai H, Vreven T, Throssell K, Montgomery JR, Peralta JE, Ogliaro F, Bearpark MJ, Heyd JJ, Brothers EN, Kudin KN, Staroverov VN, Keith TA, Kobayashi R, Normand J, Raghavachari K, Rendell AP, Burant JC, Iyengar SS, Tomasi J, Cossi M, Millam JM, Klene M, Adamo C, Cammi R, Ochterski JW, Martin RL, Morokuma K, Farkas O, Foresman JB, Fox DJ (2016) Gaussian 16 Rev. C.01. Wallingford, CT

15. APEX3 (2012) Bruker AXS. Madison, WI, USA

16. Sheldrick GM (2008) A short history of SHELX. Acta Crystallogr Sect A 64:112–122. https://doi.org/10.1107/s0108767307043930

17. Wilson AJC (1992) In: International tables for crystallography, vol C. Kluwer, Dordrecht

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.