From mannose to small amphiphilic polyol - perfect linearity leads to spontaneous aggregation

Tiina Saloranta,[†] *Anssi Peuronen,*[‡] *Johannes Dieterich,*[§] *Janne Ruokolainen,*[|] *Manu Lahtinen,*[‡] *Reko Leino**[†]

[†] Johan Gadolin Process Chemistry Centre, c/o Laboratory of Organic Chemistry, Åbo Akademi University, Turku, Finland

[‡] Department of Chemistry, University of Jyväskylä, Jyväskylä, Finland

[§] Institut für Physikalische Chemie, Universität Göttingen, Göttingen, Germany

[|] Department of Applied Physics, Nanomicroscopy Center (Aalto-NMC), Aalto University School of Science, Finland

* reko.leino@abo.fi

ABSTRACT

Terminally unsaturated and diastereochemically pure polyol derived from D-mannose shows spontaneous aggregation behavior in water solution. In order to study and clarify this unforeseen phenomenon, a conformational study based on NMR spectroscopy combined with ab initio structure analysis using the COSMO-solvation model was pursued. The results, together with X-

ray diffraction studies, suggest a low energy linear conformation for this particular substrate both in solid states and in solution. For such small-sized acyclic carbohydrate derivatives, the linear conformation appears to be a key prerequisite for the unusual molecular self-assembly reported herein.

INTRODUCTION

Carbohydrate-based amphiphiles are typically composed of a hydrophilic carbohydrate moiety attached to a relatively long, hydrophobic aliphatic carbon chain.[1] In such compounds, the amphiphilicity stems purely from the solubility difference between the two distinct ends of the molecule, while the stereochemistry of the carbohydrate part and the conformational properties play only a minor role. However, as the molecular size becomes smaller, both the conformation and the consequential linearity or non-linearity of the molecule start to have a more significant influence on the amphiphilic behavior.

Acyclic compounds, in general, favor conformations where the steric interactions are minimized. These low-energy conformations are typically characterized by planar zigzag conformations of the carbon backbone minimizing the steric interactions between different substituents. Such reasoning is valid for acyclic carbohydrate derivatives as well, as long as there are no bulky substituents, typically hydroxyl groups, in 1,3-*syn* relationship. Thus, the stereochemistries of, for example, D-mannitol and D-galactitol (see Figure 1) allow these compounds to obtain a planar zigzag conformation, whereas the corresponding D-glucose derivative, D-glucitol, favors a conformation where the C2-C3 bond is rotated 120°. This twist in the carbon chain is due to the *syn*-relationship between the OH-groups at C-2 and C-4 (Figure 1).[2,3]

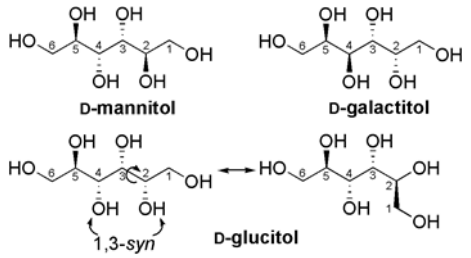
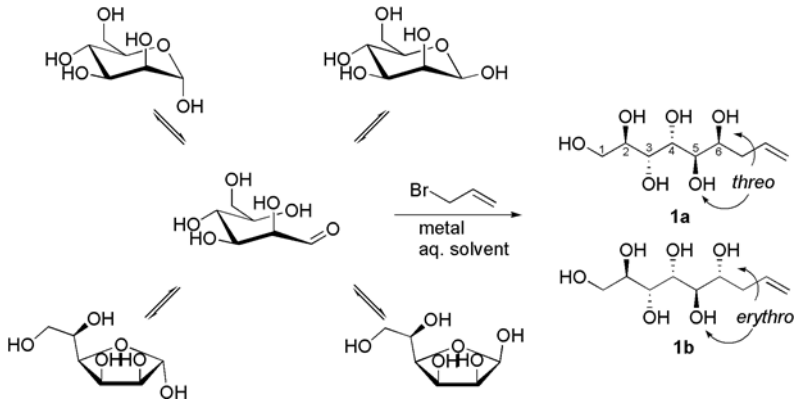


Figure 1. The low-energy conformations of D-mannitol and D-galactitol (top) and D-glucitol (bottom).

The naturally occurring monosaccharides can be utilized as precursors in the synthesis of other functionalized acyclic carbohydrate derivatives. Such approaches generally utilize the tautomeric equilibrium, termed mutarotation, and particularly the presence of the open chain aldehyde form (Scheme 1). To exemplify, metal-mediated allylation of unprotected monosaccharides yields alkene-terminated polyols with multiple chiral carbon atoms with predefined stereochemistry.[4,5] The configurations of C2-C5 stem from the parent monosaccharide and only one stereocenter (C-6) is formed in the allylation reaction. Thereby, the product is formed as a mixture of two diastereoisomers with either *threo* or *erythro* configuration (C-5/C-6). The ratio between the two diastereomers depends on the substrate and reaction conditions. The *threo* form is, however, the generally dominating one (Scheme 1).



Scheme 1. Metal-mediated allylation of unprotected D-mannose yielding alkene-terminated polyol diastereomers.

The diastereoisomers formed can, in general, be separated by acetylation-chromatography-deacetylation manipulations.[4,5] Interestingly, in the case of D-mannose as starting material, the major product diastereomer (**1a**) can be conveniently isolated by precipitation from ethanol.[6] Such alkene-terminated polyols, produced by this protocol, are synthetically attractive products as they are diastereomerically pure and contain multiple functional groups for further derivatizations.

Here, we report that the initially water soluble mannose derived polyol (**1a**) displays an unforeseen and highly interesting solubility behavior, spontaneously aggregating from a stirred water solution at ambient temperature (Figure 2). As investigated by us, the structural analogues derived from other similar monosaccharides do not show such behavior, making the D-mannose derivative **1a** in this sense very unique. This experimental finding suggests some type of highly ordered structure for this particular, potentially amphiphilic, mannose-derived diastereomer in solution that then would explain the spontaneous aggregation process.

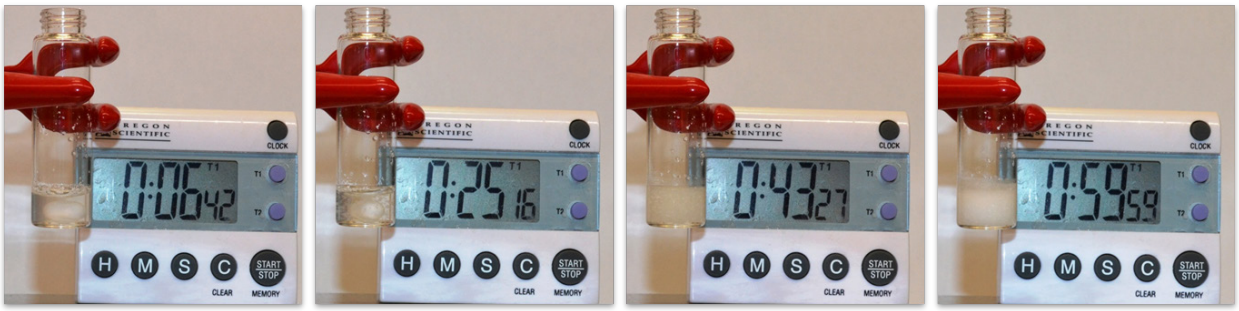


Figure 2. Progress of the spontaneous aggregation of a water solution of **1a** (c = 0.2 M) at room temperature during the first hour after initial dissolution. The elapsed time is shown in the stopwatch display.

EXPERIMENTAL SECTION

General remarks: Bruker Avance 500 MHz and 600 MHz NMR spectrometers were used to record the NMR spectra. PERCH software with spin simulation/iteration techniques was utilized to further analyze the ^{1}H NMR spectra. Optical rotations were recorded with Perkin Elmer 241 polarimeter equipped with a Na-lamp (589 nm). Melting points were measured with a Stuart Scientific apparatus. HRMS were measured using Bruker micrOTOF-Q spectrometer operating in ESI^{+} mode. For complete experimental details, see Supporting Information.

Synthesis: Compounds **1a**, **2a** and **3a** were synthesized according to literature procedures.[5,6] Compound **1a** was isolated by crystallization from EtOH (see Supporting information). For obtaining diastereomerically pure **2a** and **3a**, a peracetylation – chromatography – deacetylation procedure was pursued. For standard procedure, see for example reference 5.

(2*R*,3*R*,4*R*,5*R*,6*S*)-Non-8-ene-1,2,3,4,5,6-hexol (1a)

White solid. mp 186–188 °C. $[\alpha]_D^{20}$ = +22.8 (c = 1.0, H_2O). ^{1}H NMR (600.13 MHz, D_2O, 25 °C): δ = 5.90 (dddd, $J_{8,7b}$ = 6.9 Hz, $J_{8,7a}$ = 7.1 Hz, $J_{8,9cis}$ = 10.2 Hz, $J_{8,9trans}$ = 17.2 Hz, 1 H, H-8), 5.18 (dddd, $J_{9trans,7a}$ = –1.4 Hz, $J_{9trans,7b}$ = –1.5 Hz, $J_{9trans,9cis}$ = –2.1 Hz, $J_{9trans,8}$ = 17.2 Hz, 1 H, H-9_{trans}), 5.13 (dddd, $J_{9cis,7a}$ = –1.1 Hz, $J_{9cis,7b}$ = –1.1 Hz, $J_{9cis,9trans}$ = –2.1 Hz, $J_{9cis,8}$ = 10.2 Hz, 1 H, H-9_{cis}), 3.99 (ddd, $J_{6,5}$ = 1.5 Hz, $J_{6,7b}$ = 5.7 Hz, $J_{6,7a}$ = 8.3 Hz, 1 H, H-6), 3.91 (dd, $J_{4,3}$ = 1.1 Hz, $J_{4,5}$ = 9.4 Hz, 1 H, H-4),3.87 (dd, $J_{1a,2}$ = 3.0 Hz, $J_{1a,1b}$ = –11.9 Hz, 1 H, H-1a), 3.81 (dd, $J_{3,4}$ = 1.1 Hz, $J_{3,2}$ = 8.9 Hz, 1 H, H-3), 3.76 (ddd, $J_{2,1a}$ = 3.0 Hz, $J_{2,1b}$ = 6.5 Hz, $J_{2,3}$ = 8.9 Hz, 1 H, H-2), 3.67 (dd, $J_{1b,2}$ = 6.5 Hz, $J_{1b,1a}$ = –11.9 Hz, 1 H, H-1b), 3.59 (dd, $J_{5,6}$ = 1.5 Hz, $J_{5,4}$ = 9.4 Hz, 1 H, H-5), 2.40 (ddddd, $J_{7a,Hcis}$ = –1.1 Hz, $J_{7a,9trans}$ = –1.4 Hz, $J_{7a,8}$ = 7.1 Hz, $J_{7a,6}$ = 8.3 Hz, $J_{7a,7b}$ = –14.2 Hz, 1 H, H-7a), 2.35 (ddddd, $J_{7b,9cis}$ = –1.1 Hz, $J_{7b,9trans}$ = –1.5 Hz, $J_{7b,6}$ = 5.7 Hz, $J_{7b,8}$ = 6.9 Hz, $J_{7b,7a}$ = –14.2 Hz, 1 H, H-7b) ppm. ^{13}C NMR (150.9 MHz, D_2O, 25 °C): δ = 135.2 (C-8), 117.3 (C-9), 71.0 (C-5), 70.9 (C-2), 69.4 (C-3), 69.3 (C-6), 68.5 (C-4), 63.2 (C-1), 37.6 (C-7) ppm. HRMS calcd for $C_9H_{18}O_6Na$ $[M+Na]^+$ 245.0996, found 245.0984.

(2*R*,3*R*,4*R*,5*S*,6*R*)-Non-8-ene-1,2,3,4,5,6-hexol (2a)

White solid. mp 99–101 °C. $[\alpha]_D^{20}$ = –3.5 (c = 1.0, H_2O). ^{1}H NMR (600.13 MHz, D_2O, 25 °C): δ = 5.88 (dddd, $J_{8,7a}$ = 6.6 Hz, $J_{8,7b}$ = 7.5 Hz, $J_{8,9cis}$ = 10.2 Hz, $J_{8,9trans}$ = 17.2 Hz, 1 H, H-8), 5.18 (dddd, $J_{9trans,7b}$ = –1.4 Hz, $J_{9trans,7a}$ = –1.6 Hz, $J_{9trans,9cis}$ = –2.1 Hz, $J_{9trans,8}$ = 17.2 Hz, 1 H, H-9_{trans}), 5.14 (dddd, $J_{9cis,7b}$ = –1.0 Hz, $J_{9cis,7a}$ = –1.2 Hz, $J_{9cis,9trans}$ = –2.1 Hz, $J_{9cis,8}$ = 10.2 Hz, 1 H, H-9_{cis}), 3.96 (dd, $J_{4,3}$ = 2.3 Hz, $J_{4,5}$ = 6.1 Hz, 1 H, H-4), 3.84 (ddd, $J_{6,5}$ = 3.5 Hz, $J_{6,7a}$ = 5.0 Hz, $J_{6,7b}$ = 8.2 Hz, 1 H, H-6), 3.82 (dd, $J_{1a,2}$ = 3.0 Hz, $J_{1a,1b}$ = –11.9 Hz, 1 H, H-1a), 3.78 (ddd, $J_{2,1a}$ = 3.0 Hz, $J_{2,1b}$ = 6.4 Hz, $J_{2,3}$ = 8.2 Hz, 1 H, H-2), 3.70 (dd, $J_{3,4}$ = 2.3 Hz, $J_{3,2}$ = 8.2 Hz, 1 H, H-3), 3.67 (dd, $J_{5,6}$ =

3.5 Hz, $J_{5,4}$ = 6.1 Hz, 1 H, H-5), 3.65 (dd, $J_{1b,2}$ = 6.4 Hz, $J_{1b,1a}$ = –11.9 Hz, 1 H, H-1b), 2.39 (ddddd, $J_{7a,9cis}$ = –1.2 Hz, $J_{7a,9trans}$ = –1.6 Hz, $J_{7a,6}$ = 5.0 Hz, $J_{7a,8}$ = 6.6 Hz, $J_{7a,7b}$ = –14.3 Hz, 1 H, H-7a), 2.34 (ddddd, $J_{7b,9cis}$ = –1.0 Hz, $J_{7b,9trans}$ = –1.4 Hz, $J_{7b,8}$ = 7.5 Hz, $J_{7b,6}$ = 8.2 Hz, $J_{7b,7a}$ = –14.3 Hz, 1 H, H-7b) ppm. ^{13}C NMR (150.9 MHz, D_2O, 25 °C): δ = 134.7 (C-8), 117.6 (C-9), 73.8 (C-5), 71.0 (C-3), 71.0 (C-2), 70.3 (C-6), 69.9 (C-4), 62.8 (C-1), 37.4 (C-7) ppm. HRMS calcd for $C_9H_{18}O_6Na$ $[M+Na]^+$ 245.0996, found 245.1002.

(2*R*,3*S*,4*R*,5*S*,6*R*)-Non-8-ene-1,2,3,4,5,6-hexol (3a)

White solid. mp 115–117 °C. $[\alpha]_D^{20}$ = –1.2 (c = 1.0, H_2O). ^{1}H NMR (600.13 MHz, D_2O, 25 °C): δ = 5.89 (dddd, $J_{8,7a}$ = 6.5 Hz, $J_{8,7b}$ = 7.7 Hz, $J_{8,9cis}$ = 10.2 Hz, $J_{8,9trans}$ = 17.2 Hz, 1 H, H-8), 5.18 (dddd, $J_{9trans,7b}$ = –1.3 Hz, $J_{9trans,7a}$ = –1.6 Hz, $J_{9trans,9cis}$ = –2.1 Hz, $J_{9trans,8}$ = 17.2 Hz, 1 H, H-9$_{trans}$), 5.15 (dddd, $J_{9cis,7b}$ = –1.0 Hz, $J_{9cis,7a}$ = –1.2 Hz, $J_{9cis,9trans}$ = –2.1 Hz, $J_{9cis,8}$ = 10.2 Hz, 1 H, H-9$_{cis}$), 3.95 (ddd, $J_{2,3}$ = 1.6 Hz, $J_{2,1a}$ = 4.8 Hz, $J_{2,1b}$ = 8.0 Hz, 1 H, H-2), 3.86 (ddd, $J_{6,7a}$ = 4.1 Hz, $J_{6,5}$ = 6.6 Hz, $J_{6,7b}$ = 8.3 Hz, 1 H, H-6), 3.79 (dd, $J_{4,5}$ = 1.5 Hz, $J_{4,3}$ = 9.2 Hz, 1 H, H-4), 3.75 (dd, $J_{5,4}$ = 1.5 Hz, $J_{5,6}$ = 6.6 Hz, 1 H, H-5), 3.69 (dd, $J_{3,2}$ = 1.6 Hz, $J_{3,4}$ = 9.2 Hz, 1 H, H-3), 3.69 (dd, $J_{1a,2}$ = 4.8 Hz, $J_{1a,1b}$ = –10.2 Hz, 1 H, H-1a), 3.68 (dd, $J_{1b,2}$ = 8.0 Hz, $J_{1b,1a}$ = –10.2 Hz, 1 H, H-1b), 2.43 (ddddd, $J_{7a,9cis}$ = –1.2 Hz, $J_{7a,9trans}$ = –1.6 Hz, $J_{7a,6}$ = 4.1 Hz, $J_{7a,8}$ = 6.5 Hz, $J_{7a,7b}$ = –14.5 Hz, 1 H, H-7a), 2.25 (ddddd, $J_{7b,9cis}$ = –1.0 Hz, $J_{7b,7trans}$ = –1.3 Hz, $J_{7b,8}$ = 7.7 Hz, $J_{7b,6}$ = 8.3 Hz, $J_{7b,7a}$ = –14.5 Hz, 1 H, H-7b) ppm. ^{13}C NMR (150.9 MHz, D_2O, 25 °C): δ = 134.5 (C-8), 117.8 (C-9), 72.2 (C-6), 71.7 (C-5), 70.0 (C-2), 70.0 (C-4), 69.5 (C-3), 63.2 (C-1), 37.0 (C-7) ppm. HRMS calcd for $C_9H_{18}O_6Na$ $[M+Na]^+$ 245.0996, found 245.1002.

RESULTS AND DISCUSSION

In order to elucidate the possible underlying structural details for the observed aggregation and for validating the hypothesis on amphiphilicity, a comprehensive conformational study of the D-mannose derived alkene-terminated polyol **1a** by NMR spectroscopy was performed. The D-glucose and D-galactose derived analogues (**2a** and **3a**), not showing any aggregation behavior in water solution under similar conditions, were examined as reference compounds. Furthermore, the most stable conformations were simulated using ab initio wave function methods. Finally, the aggregation of **1a** was studied by X-ray diffraction techniques revealing the very high level of structural order as a plausible explanation for the spontaneous aggregation.

NMR spectroscopic study. Earlier, Lewis and coworkers have performed extensive conformational studies on a series of carbohydrate based polyols up to heptitols.[3,7-9] The reported results were predominantly derived from Karplus equation with the corresponding proton–proton vicinal coupling constants ($^3J_{H,H}$) as input data.[10] More recently, Murata formulated a more universal *J*-coupling based conformational model for acyclic structures.[11,12] Relevant for the present study are the structures with dioxygenated fragments. When two protons are in gauche relationship in such structures, the $^3J_{H,H}$ should be less than 3 Hz, whereas for the corresponding anti-orientation, the $^3J_{H,H}$ varies between 7 and 10 Hz (Figure 3). Thus, for linear carbohydrate derivatives in planar zigzag conformation, $^3J_{H,H}$ should be either small or large as the corresponding dihedral angles are 60° (gauche) or 180° (anti), respectively. In contrast, medium sized coupling constants ($^3J_{H,H}$ = 3–7 Hz) are characteristic for nonplanar conformations.

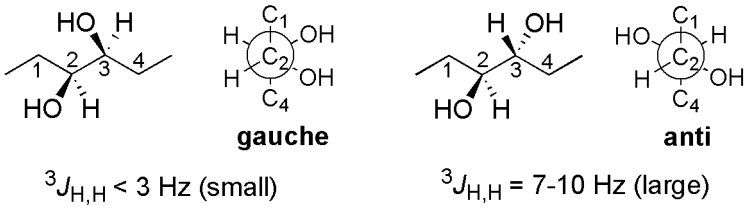
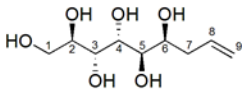
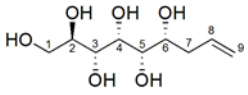
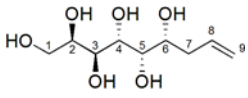


Figure 3. $^3J_{H,H}$ coupling constants in deoxygenated fragments.

Herein, the ^{1}H NMR spectra of the alkene-terminated derivatives of D-mannose (**1a**), D-glucose (**2a**) and D-galactose (**3a**) were studied in detail. Typically, the signals in ^{1}H NMR spectra of such structures are overlapping and computational tools are required for accurate interpretation of the NMR data. For this purpose, PERCH software with spin simulation/iteration techniques was utilized.[13] The $^3J_{H,H}$ coupling constants relevant for **1a**, **2a** and **3a** are given in Table 1.

Table 1. $^3J_{H,H}$ values for compounds **1a, 2a and 3a** (given in Hz).

Entry	Structure	$^3J_{H2,H3}$	$^3J_{H3,H4}$	$^3J_{H4,H5}$	$^3J_{H5,H6}$
1	 **1a**	8.9 (erythro)	1.1 (*threo*)	9.4 (erythro)	1.5 (*threo*)
2	 **2a**	8.2 (erythro)	2.3 (*threo*)	6.1 (*threo*)	3.5 (*threo*)
3	 **3a**	1.6 (*threo*)	9.2 (erythro)	1.5 (*threo*)	6.6 (*threo*)

Characteristically, two adjacent protons in *threo* relationship occur in gauche-orientation while anti-orientation is expected for protons in *erythro* relationship when the carbohydrate backbone is

in the linear conformation. The coupling constant pattern for the D-mannose derivative **1a** (Table 1, Entry 1) follows these rules as the configuration is ideal for linear conformation in the absence of 1,3-*syn* OH-groups. Owing to the configuration of **1a**, the structure specific $^{3}J_{H,H}$ values for anti- and gauche-orientations can thus be deduced, i.e., ~9 Hz for anti-orientation and 1–1.5 Hz for gauche-orientation.

In contrast to **1a**, the D-galactose derived analogue **3a** has two OH-groups (O-4 and O-6) in 1,3-*syn* relationship leading to a twist in the carbon backbone. This can be observed from the intermediate sized coupling constant (6.6 Hz) between the H-5 and H-6 protons (Table 1, Entry 3). The configuration of the corresponding D-glucose derived analogue **2a** disfavors the linear conformation even more clearly due to *syn*-relationship between both O-4/O-6 and O-3/O-5. For this structure, the $^{3}J_{H,H}$ values for protons in *threo* relationship vary between 2.3 Hz and 6.1 Hz indicating that the conformation is heavily distorted from the linear one. Based on the $^{3}J_{H,H}$ coupling constant patterns presented herein, it can be concluded that the D-glucose and D-galactose derivatives **2a** and **3a**, disfavor the linear zigzag conformation while the configuration of the D-mannose derivative **1a** is evidently ideal for the perfectly linear conformation.

Further support for this conclusion is gained from NOESY experiments (see Supporting Information). An NOE correlation between the CH_2-protons at C7 and H-4 is observed for both **2a** and **3a**. This is possible since the nonlinear conformation of the carbon chain allows the carbohydrate backbone to come closer to the hydrophobic end. For the corresponding D-mannose derived structure (**1a**), this NOE correlation is not observed as a result of the linear zigzag conformation of this structure.

Thermal analysis. Furthermore, the linearity/nonlinearity seems to have a considerable effect on the melting point. The tentatively linear D-mannose derivative **1a** melts at 186–188 °C (non-

corrected, initially measured melting point). In contrast, the non-linear structures **2a** and **3a** melt at significantly lower temperatures: 99–101 °C and 115–117 °C. The similar effect has previously been observed with sugar alcohols, i.e., D-mannitol, D-galactitol and D-glucitol as well (for structures, see Figure 1). Mannitol and galactitol with linear conformations melt at 166–168 °C and 188–189 °C, respectively, whereas the melting point of the non-linear glucitol is significantly lower (110–112 °C).[14]

To further investigate the thermal behavior of the D-mannose derivative **1a**, differential scanning calorimetry (DSC) analyses of both bulk sample precipitated from EtOH and aggregated samples were carried out (for DSC scans and data see Supporting Information). Under heating, the bulk product of **1a** shows a melting endotherm (onset temperature) at 181.9 °C preceded by a solid-solid transformation at 142.9 °C. The respective exothermic events of crystallization and solid-solid transformation with well-matching enthalpies to the endotherms are observed when cooling the sample. These thermal events are well reversible and show very little hysteresis as evidenced by a second heating/cooling cycle. Unlike the bulk, the aggregation product of **1a** shows no solid-solid transformation in the first heating cycle. Also, the melting point onset temperature is increased by ca. 8 °C compared to the bulk sample of **1a**. Cooling of the melt results in close reproduction of the thermal behavior of the bulk sample, indicating that the aggregate presents a metastable polymorph which, after melting, crystallizes to the polymorph represented by the bulk. This is confirmed by a second heating/cooling cycle which shows a good match of the temperatures and enthalpies to the bulk sample.

Computational structure analysis. In order to gain quantitative insight into the relevant structural parameters, each monomeric structure (**1a**, **2a** and **3a**) was optimized computationally. The computations were performed both under gas phase conditions and under an implicit water

solvation through the COSMO solvation model using the ab initio DF-LMP2/aug'-cc-pVTZ level of theory for the optimizations.[15,16] This level of theory provides an accuracy comparable to canonical MP2 at a fraction of the computational cost.[17,18] The optimized gas-phase geometries for **1a**, **2a** and **3a** together with the corresponding COSMO-solvated geometries are depicted in Figures 4 and 5. The calculations (see supporting information for details) evidently support the observed planarity difference between the diastereochemically different structures. While at first sight it appears that the optimized geometries are not fully consistent with the observed coupling constants presented in Table 1, this is, however, not surprising as especially the structures **2a** and **3a** presumably occur in multiple Boltzmann distributed conformational states in the experimental setup. For the structure **1a**, in turn, the linear conformation clearly dominates and the optimized geometry is fully in line with the observed coupling constant pattern and the corresponding conformation discussed above.

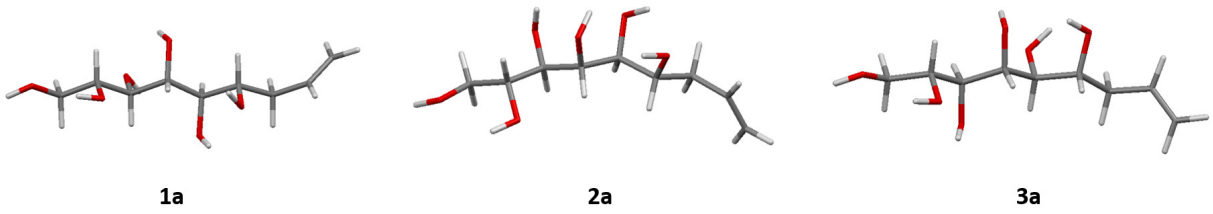


Figure 4. Optimized gas-phase geometries for **1a**, **2a** and **3a**.

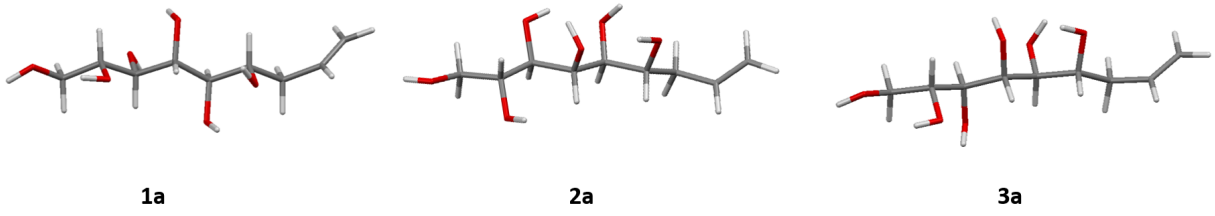


Figure 5. Optimized COSMO-solvated geometries for **1a**, **2a** and **3a**.

In these systems, the planarity is proportional to the angle between *C2−C4−C8* and this angle can thus be utilized as a relevant measure of planarity. For perfectly planar structures, this angle should be 180° (for numeric values of angle of planarity for **1a**, **2a** and **3a**, see Table 2). It can be observed that the D-mannose derived structure **1a** favors an almost perfectly planar form in both gas-phase and in implicit water solvation model. In contrast, the D-glucose and D-galactose derived structures **2a** and **3a** seem to be nonplanar under both conditions studied. However, the implicit solvation reduces the bend of all structures (see Table 2 for angles and Supporting Information for coordinates). This provides quantitative support for the hypothesis derived from the NOESY experiments, i.e., that the planarity of structure **1a** may be relevant for the observed aggregation behavior. Also, the differing energy penalties associated with forcing the structures into a linear conformation in a crystal can explain the experimentally observed differences in the melting points between **1a** and **2a**/**3a**. The reduced bend for the structure **2a** in solution is due to a favored internal hydrogen bond network straightening the backbone. The concurrency between the formation of an intermolecular and intramolecular hydrogen bond network could be another cause for the difference in aggregation. Microsolvation studies coupled with a global optimization of the optimal bonding pattern and an analysis of the underlying energy landscapes can be used to prove this hypothesis.[19] For further studies, dynamic simulations could be helpful in correctly assigning the driving force of the aggregation **1a**.

Table 2. Relevant angle of planarity (*C*2-*C*4-*C*8) for the optimized structures in gas-phase and in solution (all values in degrees).

Structure	gas-phase	solution (ε=80.4)
1a	176.6	177.6
2a	156.1	168.3
3a	169.2	171.0

X-ray diffraction studies and cryo-TEM imaging. Additionally, structural analysis of the solid aggregation product may offer insights to the relationship between the intramolecular structure and molecular packing which in turn can aid in deducing factors inducing the spontaneous aggregation of **1a**. To investigate its solid state structure, single crystals of **1a** were grown from an aqueous solution and were subjected to single crystal X-ray diffraction analysis (full details in Supporting Information).[20] The analysis reveals that the carbohydrate backbone of D-mannose derivative **1a** adopts a linear conformation in the solid state (Figure 6, Tables S3-S6). This is exemplified by the angle of planarity [∠(C2-C4-C8) = 177.73(7)°] which corresponds almost perfectly to the theoretical value obtained using the solvation model (177.6°). Due to the orientation of the hydroxyl group at the achiral C1 carbon, four of the OH-groups (O1, O2, O5 and O6) of **1a** are located below the plane generated by the carbohydrate backbone whereas two OH-groups (O3 and O4) reside above the plane. This ensures effective intermolecular hydrogen bonding (HB) scheme where **1a** is bonded to five distinct adjacent molecules. It is noteworthy that these intermolecular interactions occur via either two or three OH-groups with molecules that reside either above and below the plane, or parallel to the plane, respectively, guaranteeing high rigidity throughout the crystal lattice. Interestingly, the crystal structure of D-mannitol[21] shows

very similar HB-connectivity parallel to the carbohydrate plane, compared to **1a**, resulting in similar packing of these two compounds along the *c*-axes of the respective unit cells. However, a noticeable difference arises from the HB-connectivities of O1 and O2 of D-mannitol which engage in hydrogen bonding with altogether four adjacent molecules (HB-pattern with repeating single graph set[22] $R_3^3(9)$) instead of three as observed for **1a** (HB-pattern with alternating graph sets $R_2^2(10)$ and $R_4^4(8)$).

A close examination of the packing of **1a** also allows us to speculate on the amphiphilic character of this specific D-mannose derivative. The effects of incorporating a large hydrophobic substituent into a polyol backbone are generally observed in the solid state as a formation of layered structures due to segregation of hydrophilic and hydrophobic parts of the molecules whereas the effects of small substituents, such as the allyl group in **1a**, are not as profound. The crystal structure of **1a**, viewed along the crystallographic *c*-axis (Figure 6), shows arrays of molecules in which the hydrophobic allyl groups are parallel to each other and point alternately up and down. The hydrogen bonding pattern extends in all three dimensions, and thus the hydrophobic effect of the allyl group is not significant enough to induce a layered packing of the molecules (*cf.* benzyl group in an aldonamide derivative of D-glycero-D-gulo-heptono-1,4-lactone[23]). However, it should be noted that these observations only concern the solid state structure of **1a** and do not necessarily reflect its amphiphilic behavior in solution.

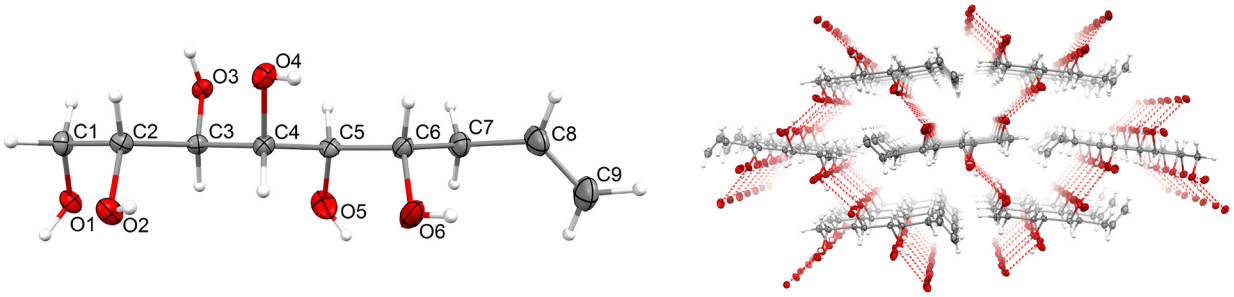


Figure 6. Left: asymmetric unit of crystal structure of **1a**. Right: intermolecular hydrogen bonding pattern (dotted lines) observed in crystal structure of **1a** and viewed along crystallographic *c*-axis. Thermal ellipsoids are presented at the 50% probability level. Disordered H-atoms of O-5 and O-6 are omitted from the figure.

The conclusions drawn from the structural analysis of single crystals of **1a** are valid for the aggregation product of **1a** only if it presents a structural match to the measured single crystals. Therefore, powder X-ray diffraction (PXRD) analyses of **1a** bulk powder crystallized from ethanol, and the precipitate, obtained from spontaneous aggregation of **1a** from an aqueous solution, were conducted and the results were compared to a simulated PXRD pattern obtained from the single crystal data of **1a**. A side-by-side comparison (Figure 7) reveals that the simulated pattern agrees well with the PXRD pattern of the aggregation product implying that both the spontaneous aggregation and slow crystallization of **1a** from an aqueous solution yield the same structure form. This can be further established by carrying out a Pawley analysis, (Figure S2) in which least-squares fit of the diffraction data is performed using the established unit cell parameters, space group setting and the peak profile parameters. The refined unit cell parameters (comparison of unit cells in Table S7) show a good fit to the single crystal unit cell with a somewhat anisotropic cell expansion (ca. 0.5 % elongation of *a* and *c* cell axes whereas *b* axis shows a 1.7 % lengthening). The overall increase in cell volume is typical considering the different measurement temperatures. Compared to the aggregation product and single crystals, the bulk precipitate clearly presents another polymorph of **1a** illustrating the significance of crystallization conditions to induce the crystallization of a specific structural form. For **1a**, such behavior can be expected on the basis of rich polymorphism of the parent non-functionalized D-mannitol.[24-26]

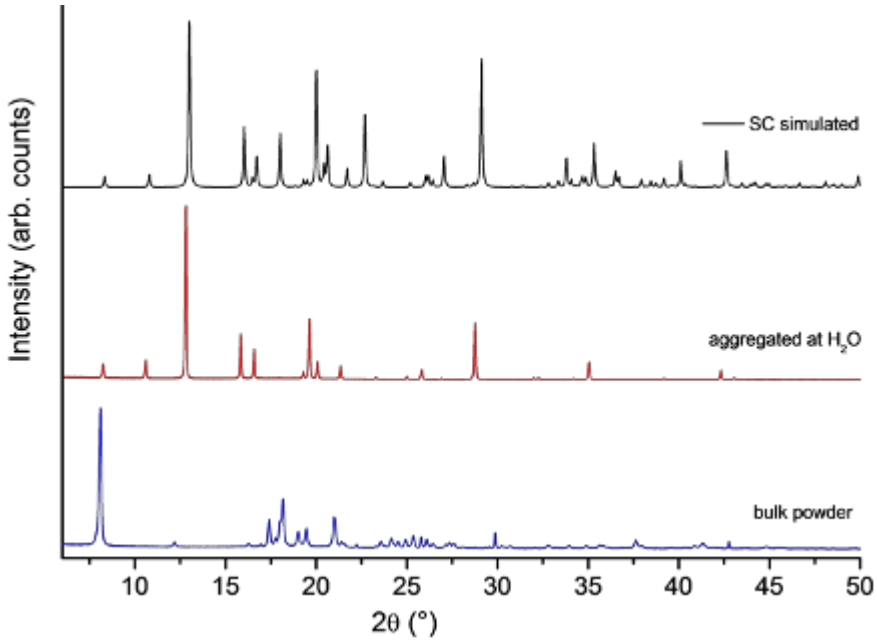


Figure 7. Simulated (single crystal, SC) and experimental powder X-ray diffraction patterns of compound **1a**.

Furthermore, cryo-TEM imaging was performed in order to directly observe the aggregation behavior in water solutions. As the water vitrification process is very rapid, the solution microstructure in the frozen cryo-TEM specimen is expected to be preserved same as prior to vitrification. Figures S3 and S4 in Supporting Information show cryo-TEM images taken from the sample which was vitrified 37 min after the initial dissolution. The image clearly shows small spherical aggregates with diameter approximately 7-8 nm, which further form larger aggregation networks.

CONCLUSIONS

To conclude, the NMR spectroscopic data, theoretical analysis, thermal analysis and crystallographic data support the perfectly linear conformation of the alkene-terminated D-

mannose derived polyol (**1a**). This type of high level of structural order that is due to favorable relative stereochemistry of this structure is suggested to play a crucial role in the observed aggregation behavior of the water solution of this compound. In turn, the corresponding nonlinear analogues derived from D-glucose and D-galactose (**2a** and **3a**) do not show similar aggregation behavior. The perfectly regular three-dimensional structure of **1a** encourages to search further applications for this molecule. For example, the terminal alkene functionality could possibly be utilized in various coupling reactions, thus opening possibilities to synthesize novel hydrophilic functional materials. An intriguing thought is also the possible role of similar enantiomerically or diastereomerically pure small molecule amphiphiles as templates for mirror symmetry breaking and the origin of biomolecular homochirality.

ACKNOWLEDGMENT

This work is part of the activities at the Johan Gadolin Process Chemistry Centre, a Centre of Excellence financed by Åbo Akademi University. JMD wishes to thank Ricardo A. Mata for fruitful discussions. Dr Jari Sinkkonen is likewise acknowledged for fruitful comments. TS gratefully acknowledges a post-doctoral researcher position from the Department of Chemical Engineering, Åbo Akademi University 2013-15.

ASSOCIATED CONTENT

Supporting Information. Experimental details, ^{1}H, ^{13}C NMR and NOESY spectra, details for DSC measurements, xyz-coordinates for optimized geometries, details for X-ray diffraction and cryo-TEM imaging. CCDC 1406085 contains the supplementary crystallographic data for this

paper. This data can be obtained free of charge from the Cambridge Crystallographic Data Centre via www.ccdc.cam.ac.uk/data_request/cif.

AUTHOR INFORMATION

Corresponding Author

* reko.leino@abo.fi

Present Addresses

† Johannes Dieterich: Scientific Computing & Modelling, De Boelelaan 1083, 1081 HV Amsterdam, The Netherlands.

Notes

The initial findings on aggregation of **1a** together with NMR spectroscopic data and computational results have been discussed in part earlier in the DrS Thesis of the first author; See: Saloranta, T. Development of Simple and Efficient Synthetic Strategies for Production of Fine Chemicals of Pharmaceutical relevance – Metal-Mediated Allylation Combined with Applied Catalysis. DrS Thesis, Åbo Akademi University, Turku, Finland, 2012. This thesis is also available via www.doria.fi.

REFERENCES

(1) Queneau, Y.; Chamber, S.; Besset, C.; Cheaib, R. *Carbohydr. Res.* **2008,** *343,* 1999−2009.

(2) Collins, P. M.; Ferrier, R. J. *Monosaccharides: Their Chemistry And Their Roles In Natural Products*: John Wiley & Sons, Chichester, 1995, pp. 35–36.

(3) Hawkes, G. E.; Lewis, D. *J. Chem. Soc. Perkin Trans. 2,* **1984,** 2073−2078.

(4) Schmid, W.; Whitesides, G. M. *J Am. Chem. Soc.* **1991,** *113, 6674 – 6675.*

(5) Kim, E.; Gordon, D. M.; Schmid, W.; Whitesides, G. M. *J. Org. Chem.,* **1993,** *58,* 5500−5507.

(6) Saloranta, T.; Müller, C.; Vogt, D.; Leino, R. *Chem. Eur. J.,* **2008,** *14,* 10539−10542.

(7) Lewis, D.; *J. Chem. Soc. Perkin Trans., 2,* **1986,** 467−470.

(8) Lewis, D.; Angyal, S. *J. Chem. Soc. Perkin Trans., 2,* **1989,** 1763−1765.

(9) Gillies, D. G.; Lewis, D. *J. Chem. Soc. Perkin Trans., 2,* **1985,** 1155−1159.

(10) Haasnoot, C. A. G. F.; de Leeuw, A. A. M.; Altona, C. *Tetrahedron*, **1980,** *36,* 2783−2792.

(11) Matsumori, N.; Kaneno, D.; Murata, M.; Nakamura, H.; Tachibana, K. *J. Org. Chem.,* **1999,** *64,* 866–876.

(12) See also: Bifulco, G.; Dambruoso, P.; Gomez-Paloma, L.; Riccio, L. R. *Chem. Rev.,* **2007,** *107,* 3744−3779.

(13) Laatikainen, R.; Niemitz, M.; Weber, U.; Sundelin, J.; Hassinen, T.; Vepsäläinen, J. *J. Magn. Res. Ser. A,* **1996,** *120,* 1−10.

(14) *The Merck Index: An Encyclopedia of Chemicals, Drugs and Biologicals,* 13th ed. O'Neil, M. J.; Smith, A.; Heckelman, P. E.; Budavari, S.; Eds.; Merck & Co., Inc., New Jersey, 2001; entries 5769, 4353 and 8797.

(15) aug-cc-pVTZ basis sets on all non-hydrogen atoms, cc-pVTZ on hydrogen atoms

(16) Dieterich, J. M.; Oliveira, J. C. A.; Mata, R. A.; *J. Chem. Theo. Comp.,* **2012,** *8,* 3053−3060.

(17) MOLPRO is a package of ab initio programs written by Werner, H.-J.; Knowles, P. J.; Knizia, G.; Manby, F. R.; Schütz, M.; Celani, P.; Korona, T.; Lindh, R.; Mitrushenkov, A.; Rauhut, G.; Shamasundar, K. R.; Adler, T. B.; Amos, R. D.; Bernhardsson, A.; Berning, A.; Cooper, D. L.; Deegan, M. J. O.; Dobbyn, A. J.; Eckert, F.; Goll, E.; Hampel, C.; Hesselmann, A.; Hetzer, G.; Hrenar, T.; Jansen, G.; Köppl, C.; Liu, Y.; Lloyd, A. W.; Mata, R. A.; May, A. J.; McNicholas, S. J.; Meyer, W.; Mura, M. E.; Nicklaß, A.; O'Neill, D. P.; Palmieri, P.; Peng, D.; Pflüger, K.; Pitzer, R.; Reiher, M.; Shiozaki, T.; Stoll, H.; Stone, A. J.; Tarroni, R.; Thorsteinsson, T.; Wang M.

(18) Werner, H.-J.; Manby, F. R.; Knowles, P. *J. Chem. Phys.,* **2003,** *118,* 8149−8160.

(19) Wales, D. J. *Energy Landscapes:* Cambridge University Press, Cambridge, 2003.

(20) **Crystal Data** for $C_9H_{18}O_6$ (M =222.23 g/mol): orthorhombic, space group $P2_12_12$ (no. 18), a = 21.1909(5) Å, b = 8.8667(2) Å, c = 5.55223(12) Å, V = 1043.22(4) Å^3, Z = 4, T = 123.01(10) K, μ(CuKα) = 1.014 mm^{-1}, *Dcalc* = 1.415 g/cm^3, 10568 reflections measured (8.346° ≤ 2Θ ≤ 153.47°), 2160 unique (R_{int} = 0.0291, R_{sigma} = 0.0198) which were used in all calculations. The final R_1 was 0.0307 (I > 2σ(I)) and wR_2 was 0.0758 (all data).

(21) Crystallographic data of D-mannitol (α, β and κ polymorphs) were obtained from Cambridge Structural Database, using CCDC ConQuest v. 1.16 (Build RC5) 2013, CSD version 5.35 (entries up to February 2014); Allen, F. H. *Acta Crystallogr.* **2002**, *B58*, 380–388.

(22) Etter, M. C.; MacDonald, J. C.; Bernstein, J. *Acta Crystallogr.* **1990**, *B46*, 256–262.

(23) Metta-Magaña, A. J.; Reyes-Martínez, R.; Tlahuext, H. *Carbohydr. Res.* **2007**, *342*, 243–253.

(24) Berman, H. M.; Jeffrey, G. A.; Rosenstein, R. D. *Acta Cryst.* **1968**, *B24*, 442–449.

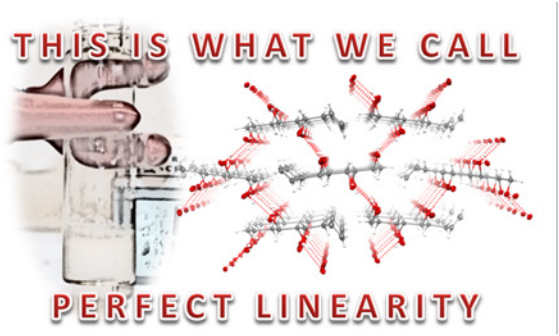
(25) Kim, H. S.; Jeffrey, G. A.; Rosenstein, R. D. *Acta Cryst.* **1968**, *B24*, 1449–1455.

(26) Fronczek, F. R.; Kamel, H. N.; Slattery, M. *Acta Cryst.* **2003**, *C59*, o567–o570 and references therein.

"For Table of Contents Use Only"

From mannose to small amphiphilic polyol - perfect linearity leads to spontaneous aggregation

Tiina Saloranta, Anssi Peuronen, Johannes Dieterich, Janne Ruokolainen, Manu Lahtinen, Reko Leino



Terminally unsaturated D-mannose derived polyol exhibits spontaneous aggregation from aqueous solution owing to its rigid backbone and semi-hydrophobic nature. This unconventional phenomenon is here shown by combinations of NMR spectroscopy, ab initio structure analysis, X-ray diffraction and thermal analysis to be due to the low energy linear conformation maintained by this particular substrate both in solid state and in solution.