Accepted Manuscript

Conformational Preferences in Diglycosyl Disulfides: Nmr and Molecular Modelling Studies

Krisztina Fehér, Richard P. Matthews, Katalin E. Kövér, Kevin J. Naidoo, László Szilágyi

PII: S0008-6215(11)00371-5
DOI: 10.1016/j.carres.2011.07.013
Reference: CAR 5864

To appear in: *Carbohydrate Research*

Received Date: 13 May 2011
Revised Date: 5 July 2011
Accepted Date: 12 July 2011

Please cite this article as: Fehér, K., Matthews, R.P., Kövér, K.E., Naidoo, K.J., Szilágyi, L., Conformational Preferences in Diglycosyl Disulfides: Nmr and Molecular Modelling Studies, *Carbohydrate Research* (2011), doi: 10.1016/j.carres.2011.07.013

CONFORMATIONAL PREFERENCES IN DIGLYCOSYL DISULFIDES: NMR AND MOLECULAR MODELLING STUDIES

Krisztina Fehér[#a,b], Richard P. Matthews[#c,d], Katalin E. Kövér[b], Kevin J. Naidoo*[c,d], László Szilágyi*[a]

Departments of Organic[a] and Inorganic[b] Chemistry, University of Debrecen , H-4010 Debrecen, Pf. 20., Hungary and Scientific Computing Research Unit[c] and Department of Chemistry[d], University of Cape Town, Rondebosch 7701, South Africa

Abstract

The conformations of several 1 1' diglycosyl disulfides were investigated by NMR and computational methods. Experimental data such as NOEs, proton-proton and proton-carbon-13 coupling constants, measured for solutions in DMSO, are in good agreement with values obtained by MD simulations in explicit DMSO. The disulfide torsion angles (C1-S-S-C1') preferentially sample values close to either +90° or –90° (+*g* or –*g*) and appear as the main metric that determines the conformational behavior of these glycomimetics. There is more conformational freedom around the C1-S and C1'-S' bonds (Φ and Ω torsions, respectively) and population cluster analysis allowed to identify up to four allowed conformational regions for each of the +*g* or –*g* forms. Population analysis of the hydroxylic group rotamers, based on proton-proton and proton-carbon-13 couplings as well as on calculated hydrogen bonding statistics, did not reveal any significant intramolecular hydrogen bonds in DMSO solution.

*Authors to whom correspondence should be addressed
e-mail: kevin.naidoo@uct.ac.za
e-mail: lszilagyi@tigris.unideb.hu
[#] These authors contributed equally to computational analyses.

Keywords: diglycosyl disulfides, conformation, MD simulation, hydroxylic rotamer populations, NMR coupling constants, NOEs

1. Introduction

Making and breaking of the glycosidic bond is fundamental to carbohydrate chemistry and biology. Natural glycosides are hydrolyzed by glycosidase enzymes, a reaction often implicated in pathological processes. Resistance to glycoside hydrolases is most efficiently achieved by manipulating the glycosidic linkage; it is therefore of prime importance to investigate the physicochemical properties of this bond.

Natural glycosidic linkages usually comprise *two bonds* (C-X-C) between the sugar ring and an aglycone, or between two sugar rings, with the bridging atom being an *oxygen* (X=O) in the majority of cases. *Three-bond* interglycosidic linkages (3BIGLs, C-X-Y-C) with X=C and

Y=O are not uncommon either but 3BIGLs with *two heteroatoms* (X,Y = O, N, S) were unknown in Nature until the discovery of the *esperamicin-calicheamicin* group of antitumor antibiotics. A striking feature of these structures is the unusual –N-O- glycosidic bond in the oligosaccharide parts of the molecules (for a review, see:[1]) NMR[2] and X-ray[3] data together with force field calculations[4] highlighted the importance of the conformation around the –N-O- linkage to ensure optimal binding to DNA for subsequent cleavage.[5,6]

Introduction of a *disulfide* motif (X,Y = S) to connect two *different* monosaccharide units led to a novel class of disaccharide mimics[7-10] characterized by another two heteroatom-3BIGL (for a review, see Szilágyi et. al[11]). Some neoglycoproteins represent further examples of interesting hybrid structures in which glycosyl units are attached to proteins through S-S linkages.[12-14]

Symmetric diglycosyl disulfides were shown to bind specifically to the plant lectin Concanavalin A[15], furthermore, inhibitory activities against an endogenous lectin were detected in vivo on human tumor cell lines.[16] Based on these results disulfide-linked sugar derivatives were suggested "as new substance platform for lectin-directed drug design".[16] We have recently described specific binding of oligovalent aromatic mannosyl disulfide derivatives to Concanavalin A.[17]

In order to gain insight into biomolecular interactions, structural characterization of the molecules involved is necessary. Experimental and theoretical studies on sugar disulfides have been limited thus far. The crystal and molecular structures of symmetric[18,19] and nonsymmetric diglycosyl disulfide derivatives[20,21] and some alkyl-glycosyl disulfide structures[22,19] have been published. Recently we investigated the chiroptical properties of diglycosyl disulfides and -selenides in solution and in the solid state.[23] Previously the conformations of a 1→4 disulfide disaccharide were estimated in solution on the basis of qualitative NOEs and simple MO considerations,[7] while molecular dynamics simulations without experimental restraints were reported for the symmetric β(1→1)digalactosyl-disulfide.[16]

High resolution NMR spectroscopy holds the most promise for experimental studies of carbohydrate conformations and dynamic properties in solution (for a recent review, see:[24]). Vicinal coupling constants proved valuable to convey information on the conformation at the glycosidic dihedrals[25-27] or in determining the distribution of hydroxymethyl rotamers.[28-30] Nuclear Overhauser effects (NOEs) are useful to estimate internuclear distances but interpretation of the data is complicated by conformational averaging on the NMR timescale and the limited number of observable NOEs between monosaccharide units.[31,32] Supplementing experimental

data with molecular dynamics (MD) simulations provides an efficient approach to handle multiple conformations[33,34,27,35,36,31] and to explore the flexibility of glycosidic linkages.

Here we use an integrated experimental and computational approach to gain insight into the conformational behavior of a series of diglycosyl-disulfides. Specifically we report our results on the disaccharide mimics D-Glc*p*-*β*S(1→1')S*β*-D-Glc*p*, D-Glc*p*-*β*S(1→1')S*β*-D-Gal*p*, D-Glc*p*-*β*S(1→1')S*β*-D-Man*p,* and D-Glc*p*-*β*S(1→1')S*β*-D-Glc*N*Ac*p* for solutions in DMSO. These are referred to as GSSG, GSSGa and GSSM and GSSGN (Figure 1), respectively to simplify the discussion below.

	R_1	R_2	R_3	R_4
GSSG	**H**	**OH**	**H**	**OH**
GSSGa	**H**	**OH**	**OH**	**H**
GSSM	**OH**	**H**	**H**	**OH**
GSSGN	**H**	**NHAc**	**H**	**OH**

Figure 1. Schematic of the disaccharides studied and definition of the torsion angles about the disulfide interglycosidic linkage.

2. Materials and methods

The compounds investigated in the present study have been synthesized according to published procedures.[10]

2.1 NMR measurements

NMR experiments were run on a Bruker Avance DRX-500 spectrometer. 10-15 mg of the samples were dissolved in DMSO-d_6 and the probe temperature was set to 300K. 2D NOE (NOESY) and rotating frame ROE (ROESY) spectra in phase sensitive mode (TPPI) were obtained using standard Bruker pulse programs; the mixing time being 300 ms in both experiments for each compound. A CW spinlock field of 3.57 kHz strength was used for the ROESY experiments. HOHAHA effects[37,38] and *J*-relayed cross peaks[39,40] were identified by repeating the experiment using a different offset frequency for the spinlock field.[41] The raw datasets typically consisted of 2K 512 complex data points. The cross peak intensities were determined by volume integration from the baseplane corrected spectra. The individually assigned cross peaks were checked for the absence of the artefacts above before converting them into distances using the Isolated Spin Pair

Approximation (ISPA).[42] The distances between 1,3 and 1,5 diaxial protons within the glycopyranosyl rings with $^4C_1$ chair geometry[43,44] were used for calibration of the integrals.

Long-range proton-carbon coupling constants were determined using a sensitivity enhanced, gradient selected ^{13}C-filtered TOCSY pulse sequence called *gs*-HETLOC[45] and HSQMBC experiments.[46,47] Theoretical values for homo- and heteronuclear vicinal couplings involving hydroxyl protons were calculated from MD-simulated torsion angles using the Karplus type equations below,[48,49]

$$^3J_{H,OH} = 10.4\cos^2\varphi - 1.5\cos\varphi + 0.2 \quad (1)$$

$$^3J_{HOCC} = 5.494\cos^2\vartheta - 0.585\cos\vartheta + 0.102 \quad (2)$$

where φ is defined as H_i-C_i-O-H_i and ϑ as C_{i-1}-C_i-O-H_i.

Populations of the hydroxyl rotamers around the C-O bond were estimated from the measured $^3J_{H,OH}$ and $^3J_{HOCC}$ coupling constants[49] and compared with calculated values (see Results and Discussion).

2.2 Computational

2.2.1 Adiabatic maps

The energy landscapes for each disaccharide were explored as a function of the three torsion angles that describe the conformation about the disulfide linkage. They are defined as: $\boldsymbol{\Phi}$ = H1 – C1 – S1 – S1', $\boldsymbol{\Psi}$ = C1– S1 – S1' – C1', and $\boldsymbol{\Omega}$ = S1 – S1' – C1' – H1' (see Fig.1). Adiabatic maps were constructed in three dimensional (, ,) space using the simulated annealing algorithm[50] as described previously for the three-bond glycosidic linkage in isomaltose[36] and panose.[35] The disaccharides were modeled using the CSFF force field[51] with the addition of CSFF consistent parameters for the SS-linkage.[52]

2.2.2 Molecular Dynamics simulations

Based on the three-dimensional adiabatic maps, low lying minima were chosen as starting points for simulations in explicit DMSO. The molecular dynamics program CHARMM33b[53] was used for all the simulations. DMSO MD simulations were conducted using the Strader et al. model for DMSO.[54] The dynamics were run in a cubic box with sides 41.7262 Å applying periodic boundary conditions. The disaccharides were solvated with 612 DMSO molecules. The overlapping DMSO molecules within a heavy atom distance of 3.5 Å from the solute were

subsequently removed. After an initial 1 ns equilibration the MD simulations were extended for a further 10 ns for the conformations that were most consistent with experimental NOE data, while shorter, 5 ns simulations were run for the remaining conformations. The simulations were carried out using the leapfrog Verlet integrator and implementation of the isothermal-isobaric ensemble (NPT) where the pressure and temperature are kept constant (P = 1 bar, T = 300 K) by making use of the Langevin piston method.[55] Data for all respective configurations were stored at intervals of 0.05 ps and 0.5 ps for the 10 ns and 5 ns long simulations, respectively.

2.2.3 Computational Analysis

Population cluster analysis (PCA) was performed using the ART2' adaptive analogue pattern recognition scheme.[56,57] A cluster radius value[58] was used to prevent false clusters from arising. In order to measure the occurrence of intra- and intermolecular hydrogen bonds during the MD simulation hydrogen bonding statistics were calculated. We used a geometric definition of a maximum distance criterion of 2.4 Å between the acceptor oxygen and the hydrogen along with an angle of not less than 100^{o} for (donor oxygen)-hydrogen-(acceptor oxygen) arrangement.

3. Results and discussion

3.1 Conformational analysis of the interglycosidic torsion angles based on adiabatic maps and MD simulations in explicit DMSO

The three dimensional adiabatic energy surface revealed local minima primarily in two main areas characterized by −90^{o} and +90^{o} along the Ψ angle for all disulfide disaccharides. Therefore the conformational space is discussed in terms of these two low energy regions separately. Conformations close to Ψ = −90^{o} and +90^{o} minima are designated as –*g* and +*g* conformations, respectively. The potential energy surfaces shown in Figs. 2 and 3 were obtained by slicing along the Ψ = −90° and Ψ = +90° plane of the three dimensional adiabatic surfaces and contouring at energies of 2 kcal.mol^{-1} from the lowest minimum and up to 12 kcal.mol^{-1}. Within these regions several local minima exist for each of the disaccharides, these are identified by letters A to D in the –*g* conformation (Fig. 2) and E to G in the +*g* conformation (Fig. 3; for a detailed listing of the data, see Supplementary Table S1).

Figure 2. Representative two-dimensional sections from the (a) GSSG, (b) GSSGa, (c) GSSM and (d) GSSGN adiabatic maps for $\Psi = -90°$. The energy is contoured in increments of 2 kcal.mol^{-1} above the local minimum.

Figure 3. Representative two-dimensional sections from the (a) GSSG, (b) GSSGa, (c) GSSM and (d) GSSGN adiabatic maps for $\Psi = +90°$. The energy is contoured in increments of 2 kcal.mol^{-1} above the local minimum.

The results of the MD simulations in explicit DMSO are displayed in Figs. 4 and 5 by depicting cluster distributions in the Φ/Ω space. Population averaged angles of the clusters of the MD simulation are summarized in Table 1 along with the torsion angles of their starting structures.

Figure 4. Representations of the clusters arising from MD trajectories calculated in DMSO and started in the −*g* conformation. Each symbol identifies a cluster obtained by PCA of the MD run started from the adiabatic minima A-D labeled with symbols ◆,□, ▲ and ○ respectively. Conformational regions are denoted by roman numerals in accordance with Table 1. Data are shown for the –*g* conformations of GSSG (a), GSSGa (b), GSSM (c), and GSSGN (d).

Figure 5. Representations of the clusters arising from MD trajectories calculated in DMSO and started in the +*g* conformation. Each symbol identifies a cluster obtained by PCA of the MD run started from the adiabatic minima A-D labeled with symbols △, ■, and ○ respectively. Conformational regions are denoted by roman numerals in accordance with Table 1. Data are shown for the +*g* conformations of GSSG (a), GSSGa (b), GSSM (c), and GSSGN (d).

For **GSSG** the lowest energy minimum is A in the –*g* conformation followed closely by minimum E with relative potential energy of 0.256 kcal.mol^{-1} in the +*g* conformation as shown in Table S1. All other minima have energies more than 1 kcal.mol^{-1} higher than the global minimum. MD simulations started from each of the adiabatic minima resulted in four conformational regions listed in Table 1 (last column). Inspection of the data shows that for GSSG in the –*g* conformation minima A and D were quasi conserved in the MD simulations yielding cluster groups **II** and **III**, while minimum C shifted along the angle to a region, designated as **I** as shown in Fig. 4a. In contrast, all simulations in the +g conformation resulted in a single region centred close to minimum G, named as region **V** as depicted in Fig. 5a. The population-averaged Ψ values are slightly less than the ideal +/−90^{o} for all clusters except for region **I**, which is characterized by $\Psi = -100^{o}$ (Table 1).

For **GSSGa** the global energy minimum is state A in the -*g* conformation as shown in Table S1, while the second lowest energy minimum is well B with 0.98 kcal.mol^{-1}, also found in the –*g* conformation. The MD simulations started from the different adiabatic minima yielded altogether six regions listed in Table 1. Four regions of the conformational space were sampled in the –*g* conformation, corresponding to the areas around minima A, C, D yielding conformational groups **III, II, I** and an additional new region, conformational area **IV** was found, which emerged from the simulation started from minimum B as shown in Fig. 4b. In the +*g* conformational state the MD simulations sampled conformations in one broad region encompassing all minima E, F and G. By comparison, +*g* conformations of the other derivatives could be clustered along Ω into the two regions designated as V and VI, as seen in Fig. 5b. The population averaged Ψ angles are slightly less than the ideal +/−90^{o} for all clusters, while region **IV** displays a distinctly different Ψ angle (−107^{o}).

For **GSSM** the global energy minimum is in the +*g* conformation E, though it is closely followed by minimum A in the -*g* conformation, which has only 0.2 kcal.mol^{-1} higher energy than E as shown in Table S1. There are yet another two minima which are relatively close to the global minimum: state B with 0.775 kcal.mol^{-1} and F with 0.997 kcal.mol^{-1}. The MD simulations started from the various minima explored only four regions of the conformational space listed in Table 1. In the –*g* conformation only areas around minima D and C were verified by the MD simulation resulting in conformational regions **I** and **II** as shown in Fig. 4c. In the +g conformation a region around minimum E designated as area **VI** and clusters arising from the simulation started from minimum G, named as conformational group **VII**, were identified as shown in Fig. 5c. All resulting regions have similar Ψ values of somewhat smaller than +/−90^{o} with the exception of the cluster starting from G (105^{o}).

For **GSSGN** the global energy minimum is A in the -*g* conformation with all other conformations having significantly higher potential energy as shown in Table S1. The MD simulations for **GSSGN** produced three distinct regions in both of the –*g* (**I, II** and **III**, Fig. 4d) and +*g* (**V, VI** and **VIII**, Fig. 4b) conformational states as listed in Table 1. All six regions displayed very similar Ψ values of somewhat less than the ideal +/−90^{o}.

Table 1
Torsion angles of minimum energy starting structures for MD simulations yielding conformational regions with population averaged torsion angles.

	Adiabatic map minima	Φ	Ψ	Ω	MD population averaged			Conformational regions
					Φ	Ψ	Ω	
GSSG	A	-170	-90	0	-166	-82	38	**III**
	B	60	-90	-40	182	-85	19	**III**
	C	50	-90	50	-2	-100	-21	**I**
	D	30	-90	-180	30	-86	179	**II**
	E	60	90	-50	-5	84	46	**V**
	F	60	90	-50	-22	86	55	**V**
	G	-40	90	40	-22	85	46	**V**
GSSGa	A	180	-90	20	-171	-83	-4	**III**
	B	-140	-90	-160	-50	-107	171	**IV**
	C	30	-90	170	39	-78	164	**II**
	D	30	-90	30	35	-88	38	**I**
	D				-178	-90	35	**III**
	E	-40	90	40	-8	81	33	**V**
	E				21	86	-22	**VI**
	F	80	90	-20	-35	90	44	**V**
	G	40	90	-30	-1	87	-39	**VI**
GSSM	A	-150	-90	170	30	-81	181	**II**
	B	180	-90	-10	21	-91	59	**I**
	C	30	-90	150	-3	-86	194	**II**
	D	50	-90	30	45	-86	43	**I**
	E	60	90	-30	28	82	-56	**VI**
	F	-170	90	20	42	79	-62	**VI**
	G	-20	90	50	174	105	52	**VII**
GSSGN	A	180	-90	30	-164	-85	13	**III**
	B	180	-90	0	-27	-86	-20	**I**
	C	-30	-90	20	13	-90	28	**I**
	D	40	-90	170	50	-89	178	**II**
	E	70	90	-20	37	85	-53	**VI**
	F	-20	90	50	18	84	27	**V**
	G	-40	90	170	12	90	164	**VIII**

The adiabatic maps gave an overall impression of the minimum energy conformations, but some of these regions disappeared or shifted to a slightly different area of the potential energy surface during the MD simulations. The differences between the adiabatic maps and the space sampled by MD may be attributed to solvation effects in the MD simulation. Overall the number of final conformational regions were reduced in comparison with the number of adiabatic minima.

The MD simulations confirmed that the Ψ angle preferences were ca. +/−90^{o}. The distributions of Ψ values in the MD simulations were restricted to a comparatively narrow range

($\leq 24^{o}$); in contrast, the Φ and Ω angles explored significantly wider ranges. The only notable exception was the MD run started from minimum C of GSSM yielding clusters in conformer region **I**, where the Ψ angle assumed values between -75^{o} and -122^{o}.

None of the MD simulations crossed the energy barrier between the –g and +g conformations for any of the derivatives. The conformational energy barriers along Φ or Ω were not overcome either in *–g* conformation during the 10 ns simulation time with the exception of GSSGa where starting from minimum D a transition was observed over the energy barrier at the saddle point between D and A (Fig. 4b). The lifetime of the conformation explored between the two states was 5 ns (Fig. 6). Conformational regions V and VI were both explored in the MD simulation of GSSGa in the *+g* conformation started from minima E (Fig. 5b). On the other hand, three distinctly separate regions were detected for GSSGN with no transitions between them.

Figure 6. The transition around the Φ angle in the *–g* conformer of GSSGa

The preference of conformations where ***Ψ*** is ~90^{o}, rather than 180^{o} or 0^{o}, is supported by crystal structure results for various glycosyl disulfides[18-22] where it was found that in most cases the SS-bond assumes the *–g* conformation. These findings are further supported by simple MO considerations [7] and earlier theoretical calculations[59] of preferred conformations in unrestricted disulfides. Recent chiroptical studies of diglycosyl disulfides and -diselenides indicated conformational equilibrium between the *+g* and *–g* conformers in solution with the predominance of the *–g* forms, however.[23] Direct comparison with CD results may be of little relevance, however, because of significant differences in the experimental conditions; CD spectra were measured for solutions in methanol or water at much lower concentrations.

3.2 Hydrogen bonding statistics based on simulations

We have calculated hydrogen bonding statistics for the four disaccharides from the MD simulations and an overall summary is given in Table 2.

Table 2.
Hydrogen bonding statistics: the unbracketed number refers to the average number of hydrogen bonds per sugar oxygen (as either donor or acceptor), while the numbers in brackets indicate the average lifetime in picoseconds of each such bond.

Disaccharide	Simulation	To DMSO	Internal
GSSG	+g	0.73 (0.50)	0.12 (0.07)
	-g	0.73 (0.51)	0.15 (0.09)
GSSGa	+g	0.91 (0.56)	0.15 (0.09)
	-g	0.87 (0.78)	0.15 (0.12)
GSSM	+g	0.90 (0.53)	0.09 (0.09)
	-g	0.94 (0.51)	0.09 (0.09)
GSSGN	+g	0.75 (0.55)	0.11 (0.09)
	-g	0.72 (0.51)	0.13 (0.11)

The simulation indicated formation of 0.72-0.94 hydrogen bonds, on average, to DMSO solvent molecules per sugar oxygen, with a life span of 0.50-0.78 ps, compared to an average of 0.09 to 0.15 hydrogen bonds to alternate OH acceptors, with a calculated life span of 0.07- 0.09 ps for all the disaccharide mimetics. This provides evidence that DMSO molecules, similar to water molecules,[60,61,36,62] are effectively competing for hydrogen bond formation with the disaccharide and as a result they disrupt internal hydrogen bonds of the sugar molecule.

3.3 NMR parameters from experiment and simulation

The relevance of the computational results can be assessed by comparison with experimental data. NMR parameters such as NOEs and scalar coupling constants were therefore computed from MD trajectories to check against measured values.

3.3.1 Interannular NOEs between protons on the SS-connected glycopyranosyl rings

The observed NOEs can be divided into two categories depending on whether the interacting protons are on the same (intra-ring) or different glycopyranosyl rings (interannular) of the disaccharide. Of the intra-ring interactions the H_i-C_i-O-H_i types could be observed for virtually all OH groups (except when spectral overlap prevented detection) with similar intensities. OH rotamer distributions were therefore estimated on the basis of coupling constants, see below. Interannular NOEs are, on the other hand, relevant with regard to the conformations about the interglycosidic bond. No interannular NOEs are available for GSSG, because of symmetry. Few NOEs were detected for GSSGa, due to very close overlap of the respective resonances for glucose and galactose ring protons. Particularly, the characteristic NOEs between anomeric

protons clearly seen in GSSM and GSSGN, can not be detected in GSSGa because of quasi identical chemical shifts (Supplementary data Table S2). Interannular NOEs observed for GSSM are depicted in Fig. 7 as an example.

Figure 7 ROESY spectrum (τ_{mix} = 300 ms) of GSSM in DMSO. Interannular cross peaks are labeled as in Table 3.

Proton-proton distances derived from experimental NOEs were compared with values obtained from MD simulations in DMSO. It is seen from Table 3 that the MD-averaged values agree with NOE distance constraints within 1 Å in general; the single exception being G-OH2/M1 in GSSM (cf. Fig. 7). These results provide support for the conformational analysis derived from the computer simulations.

Table 3.
Interannular ^{1}H-^{1}H distances from NMR and simulations in DMSO.

Molecule	NOE assignment	Interproton distance	
		NMR	MD (average)
GSSGa	G1 / Ga1	n.a[#]	3.5
	G1 / Ga2	n.a[#]	5.0
	G1 / Ga3	4.8	4.7
	G1 / Ga5	5.0	5.2
	G1 / Ga-OH2	n.*d*[&]	5.9
	G1 / Ga-OH6	n.*d*[&]	6.8
	G2 / Ga2	n.a[#]	6.4
	G2 / Ga4	n.a[#]	7.2
GSSM	G1 / M1	3.7	3.3
	G2 / M1	4.5	4.4
	G-OH2 / M1	3.8	5.4
	G-OH6 / M1	4.2	5.2
	G2 / M2	4.3	5.3
	G4 / M2	4.8	5.6
GSSGN	G1 / N1	3.4	3.4
	G2 / N1	3.7	4.7
	G-OH2 / N1	4.2	5.1
	G2 / N-Me	4.5 +1[$]	5.8
	G-OH6 / N-Me	5.0 +1[$]	6.3

n.a[#] Not available because of spectral overlap.
n.d[&] Not detected
[$] Pseudoatom correction for Me groups

As a general trend, the shortest interannular distances are observed between anomeric protons both experimentally and calculated.

3.3.2 ^{1}H-^{1}H and ^{1}H-^{13}C *J*-couplings involving OH protons.

Three-bond coupling constants were calculated from the MD-simulated torsion angles using Karplus equations (1) and (2) (see Methods section). The results and comparison with experimental data are summarized in Tables 4 & 5.

Table 4.

^{1}H-^{1}H coupling values in Hz obtained from experimental NMR studies and MD simulations

Molecule	Coupling	NMR*	MD average**
GSSG	Glu-OH2	6.1	5.5
	Glu-OH3	4.5	4.5
	Glu-OH4	5.4	6.2
	Glu-OH6	5.5	5.1
	Glu'-OH2	6.1[#]	4.9
	Glu'-OH3	4.5[#]	4.8
	Glu'-OH4	5.4[#]	6.0
	Glu'-OH6	5.5[#]	5.1
GSSGa	Glu-OH2	5.7	4.7
	Glu-OH3	4.8	4.8
	Glu-OH4	5.3	4.8
	Glu-OH6	5.6	5.0
	Gal-OH2	6.1	5.1
	Gal-OH3	5.8	5.4
	Gal-OH4	5.0	4.3
	Gal-OH6	5.6	4.4
GSSM	Glu-OH2	5.9	5.1
	Glu-OH3	4.8	4.5
	Glu-OH4	5.4	4.8
	Glu-OH6	5.7	5.3
	Man-OH2	5.4	4.2
	Man-OH3	5.4	4.5
	Man-OH4	5.0	4.4
	Man-OH6	5.7	5.2
GSSGN	Glu-OH2	5.9	5.2
	Glu-OH3	n.a[&]	4.7
	Glu-OH4	5.0	4.8
	Glu-OH6	5.1	4.8
	GlcNAc-NH	9.3	8.1
	GlcNAc-OH3	5.4	5.0
	GlcNAc-OH4	n.a[&]	4.6
	GlcNAc-OH6	5.6	5.0

* The ^{1}H-^{1}H couplings obtained from NMR experiments have a standard deviation of 0.3 Hz.
**Averaged over the +g and –g conformations, standard deviation 0.4-0.8 Hz.
[&] Not available because of spectral overlap.
[#] Only one set of values available by NMR because of symmetry

Table 5.

^{1}H-^{13}C coupling values in Hz obtained from experimental NMR studies and MD simulations

Molecule	Coupling	NMR*	MD average**
GSSM	Glu-OH2/C1	3.6	3.3
	Glu-OH2/C3	2.2	2.0
	Glu-OH3/C2	2.5	2.8
	Glu-OH3/C4	2.9	2.7
	Glu-OH4/C3	2.5	1.4
	Glu-OH4/C5	3.4	5.7
	Glu-OH6/C5	2.3	4.8
	Man-OH2/C1	n.a	3.4
	Man-OH2/C3	1.8	2.5
	Man-OH3/C2	2.8	3.2
	Man-OH3/C4	2.3	2.9
	Man-OH4/C5	3.4	3.4
	Man-OH6/C5	2.3	2.7
	Man-OH4/C3	2.3	1.7
GSSGa	Glu-OH2/C1	3.0	3.2
	Glu-OH2/C3	2.0	3.6
	Glu-OH3/C2	1.6	2.2
	Glu-OH3/C4	2.0	2.2
	Glu-OH4/C3	2.3	2.3
	Glu-OH4/C5	3.0	3.6
	Glu-OH6/C5	2.6	3.1
	Gal-OH2/C1	3.3	3.2
	Gal-OH2/C3	2.6	2.4
	Gal-OH3/C2	2.0	2.3
	Gal-OH3/C4	3.0	3.5
	Gal-OH4/C3	1.6	2.1
GSSGN	Glu-OH2/C1	2.9	3.2
	Glu-OH2/C3	1.0	1.6
	Glu-OH3/C2	1.6	3.0
	Glu-OH3/C4	n.a*	2.7
	Glu-OH4/C3	3.4	2.7
	Glu-OH6/C5	2.3	4.0
	GlcN-OH3/C4	1.1	2.8
	GlcNAc-OH3/C2	3.4	2.6
	GlcNAc-OH4/C5	1.8	5.5
	GlcNAc-NH/C1	n.a[#]	5.6

* The ^{1}H-^{13}C couplings obtained from NMR experiments have a standard deviation of 0.4 Hz.
** Averaged over the +g and –g conformations, standard deviation 0.4-0.8 Hz.
[#] Not available because of spectral overlap.

Good correlation was observed between experimental and calculated values for all ^{1}H-^{1}H coupling constants. The majority of the calculated $^{3}J_{HOCC}$ couplings also show reasonable agreement with measured values. Differences exceeding +/- 1Hz may be due to partial exchange of OH-protons on the experimental side and/or substituent effects not properly taken into account in eq. (2) in the calculations. $^{3}J_{HC}$ coupling constants are known to be notoriously sensitive to substituent effects (see[63] and references cited therein). A graphical representation (using GSSGa as an illustration) of the correlation between experimental and calculated values is shown in Figure 8.

Fig. 8. Correlation between experimental and calculated coupling constants for GSSGa (cf. Tables 4 & 5). Color coding and symbols: *, blue: $^{3}J_{HOCH}$ measured; o, red: $^{3}J_{HOCH}$ calculated; •, green: $^{3}J_{HOCC}$ measured; o, black: $^{3}J_{HOCC}$ calculated. Error bars are added in matching colours.

According to the Karplus equations (1) & (2) coupling constants of the order of ~ 5.4 Hz and ~ 2.8 Hz, for $^{3}J_{HOCH}$ and $^{3}J_{HOCC}$, respectively, represent rotation-averaged values. The data in Tables 4 and 5 are close to these values and therefore indicate quasi free rotation around the C-O bond for all hydroxyl groups investigated. When both couplings are available for a particular OH further information can be deduced, namely, populations of each of the OH rotameric states can be estimated in terms of *antiperiplanar* (*ap*) and *synclinal* (*sc*) conformations around the C-O bond.[49] Population analyses of the conformational preferences of the same hydroxyl groups were performed for both the +*g* and –*g* simulations by binning the sampled rotamers in increments of 5^{o}. Rotamers within 5^{o} of any transitions, i.e. +*sc* to -*sc*, etc. were excluded from the final summation. OH rotameric populations that were estimated from NMR coupling constants show good correlation with values obtained from MD simulations (Table 6).

Rotameric stabilization may be promoted by stereo-electronic effects and/or hydrogen bonding. Calculated H-bonding statistics presented above are, however, ruling out the occurrence of persistent hydrogen bonds in DMSO solution. This is (also) in line with experimental and theoretical results demonstrating the disruption of internal hydrogen bonds in solvents with increased polarity such as water or DMSO.[60-62] The data in Table 6 indicate no preferred rotameric state for the majority of the OH groups in GSSGa, GSSM and GSSGN. Low percentage of the *ap* rotamers for Glu-OH3 and Gal-OH4 in GSSGa and one of the *sc* rotamers for GlcNAc-OH3 in GSSGN were found, however, by experiment and calculation alike (Table

6). According to the considerations above this restriction of the rotameric freedom may be due stereo-electronic effects rather than engagement into internal hydrogen bonds of these particular OH groups.

Table 6.

NMR and MD rotamer population distributions of OH groups

Compound	Group	NMR measured rotamers (%)*			MD calculated rotamers (%)T		
		P(-*sc*)	P(*ap*)	P(+*sc*)	P(-*sc*)	P(*ap*)	P(+*sc*)
GSSGa	Glu-OH2	36	37	17	32	40	21
	Glu-OH3	17	9	28	26	4	49
	Glu-OH4	33	23	37	35	10	45
	Glu-OH6	36	29	n.a.	45	40	13
	Gal-OH2	40	43	28	40	40	18
	Gal-OH3	37	17	37	36	19	40
	Gal-OH4	60	9	29	60	18	20
GSSM	Glu-OH2	38	48	20	37	42	20
	Glu-OH3	35	27	28	28	6	47
	Glu-OH4	34	27	44	24	22	50
	Glu-OH6	37	23	n.a.	46	42	8
	Man-OH2	13	n.a.	34	11	53	34
	Man-OH3	23	33	34	11	25	44
	Man-OH4	30	23	45	18	14	50
	Man-OH6	37	23	n.a.	45	38	15
GSSGN	Glu-OH2	38	35	n.a.	40	40	17
	Glu-OH3	n.a.	n.a.	n.a.	55	42	1
	Glu-OH4	29	30	41	30	15	50
	Glu-OH6	31	23	n.a.	44	41	5
	GlcNAc-OH3	0	45	33	3	37	57

*As determined from three-bond proton-proton- ($^3J_{H,OH}$) and proton-carbon ($^3J_{HOCC}$) coupling constants (see text). The *+sc, -sc* and *ap* notations refer to the *synclinal* or *antiperiplanar* orientation of proton OH(*i*) with respect to C(*i-1*). Populations P(*n*) are approximate (+/−10%) values.

TMD rotamer populations calculated for the *+g/-g average* conformation.

4. Conclusion

In summary, experimental NMR data such as NOEs, proton-proton and proton-carbon-13 coupling constants supplemented with MD calculations in explicit DMSO have clearly indicated that the main metric to determine the conformations in 1 1' diglycosyl disulfides is the disulfide

torsion angle (C1-S-S-C1') which preferentially samples values close to either +90° or –90° (+*g* or –*g*). Significantly more conformational freedom was observed around the C1-S and C1'-S' bonds (Φ and Ω torsions, respectively) and population cluster analysis (PCA) allowed to identify up to four allowed conformational regions for each of the +*g* or –*g* forms.

Regarding conformational similarities vs. differences between the members of the current panel of diglycosyl disulfide structures it can be stated in general that substituent effects (steric or electronic) of groups close to the disulfide bridge have larger impact on the conformational distributions than changes at remote positions. Visual comparison of adiabatic maps and MD simulations both show the highest conformational similarity of GSSG to GSSGa, less similarity to GSSGN and even less to GSSM. In particular, in the –g form conformational groups I, II and III are sampled by GSSG, GSSGa and GSSGN whereas only groups I and II are present for GSSM. In the +g form conformational group V, which is dominant for GSSG, is shared only by GSSGa and GSSGN, but is not present for GSSM. This is in line with the general trend stated above. The change of the 4-OH group from equatorial (GSSG) to axial (GSSGa) position is far away from the disulfide bridge and therefore does not result in large conformational differences. The 2-NAc group in GSSGN is closer to disulfide bridge, however, the substitution does not significantly affect the stereoelectronic charaterisitics of GSSGN compared to GSSG. The GSSM derivative features the largest difference compared to GSSG because the equatorial 2-OH group, next to the disulfide linkage, in the latter is changed for an axial one in the mannosyl unit of the former.

Estimation of the hydroxylic group rotamer populations, based on calculated hydrogen bonding statistics and ^{1}H-^{1}H / ^{1}H-^{13}C *J*-coupling analysis, did not reveal any significant intramolecular hydrogen bonds in DMSO solution. Knowledge of the conformational preferences in this novel class of glycomimetics will certainly contribute to the design of further derivatives and provide insight into biologically relevant interactions such as binding to lectins[15,17] or tumor cells.[16]

Acknowledgements

This work is based upon research supported by the South African Research Chairs Initiative of the Department of Science and Technology and National Research Foundation to KJN and by the South African Hungarian collaboration programs (TéT DAK-1/02, ZA-26/2006 and ZA-20/2008) and by OTKA (NK-68578), TÁMOP-4.2.2-08/1/2008-0019 and TÁMOP-4.2.1./B-09/1/KONV-2010-0007 grants in Hungary.

References

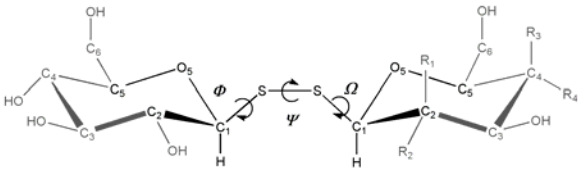
1. Smith, A. L.; Nicolaou, K. C. *J. Med. Chem.* **1996,** *39*, 2103-2117.
2. Walker, S.; Valentine, K. G.; Kahne, D. *J. Am. Chem. Soc.* **1990,** *112*, 6428-6429.
3. Lee, M. D.; Dunne, T. S.; Chang, C. C.; Ellestad, G. A.; Siegel, M. M.; Morton, G. O.; McGahren, W. J.; Borders, D. B. *J. Am. Chem. Soc.* **1987,** *109*, 3466-3468.
4. Walker, S.; Gange, D.; Gupta, V.; Kahne, D. *J. Am. Chem. Soc.* **1994,** *116*, 3197-3206.
5. Biswas, K.; Pal, S.; Carbeck, J. D.; Kahne, D. *J. Am. Chem. Soc.* **2000,** *122*, 8413-8420.
6. Sissi, C.; Aiyar, J.; Boyer, S.; Depew, K.; Danishefsky, S.; Crothers, D. M. *Proc. Nat. Acad. Sci.USA* **1999,** *96*, 10643-10648.
7. Chakka, N.; Johnston, B. D.; Pinto, B. M. *Can. J. Chem.* **2005,** *83*, 929-936.
8. Morais, G. R.; Falconer, R. A. *Tetrahedron Lett.* **2007,** *48*, 7637-7641.
9. Stellenboom, N.; Hunter, R.; Caira, M. R.; Szilágyi, L. *Tetrahedron Lett.* **2010,** *51*, 5309-5312.
10. Szilágyi, L.; Illyés, T. Z.; Herczegh, P. *Tetrahedron Lett.* **2001,** *42*, 3901-3903.
11. Szilágyi, L.; Varela, O. *Curr. Org. Chem.* **2006,** *10*, 1745-1770.
12. Davis, B. G.; Lloyd, R. C.; Jones, J. B. *J. Org. Chem.* **1998,** *63*, 9614-9615.
13. Watt, G. M.; Boons, G. J. *Carbohydr. Res.* **2004,** *339*, 181-193.
14. Bernardes, G. J. L.; Gamblin, D. P.; Davis, B. G. *Angew. Chem.-Int. Ed.* **2006,** *45*, 4007-4011.
15. Pei, Z. C.; Aastrup, T.; Anderson, H.; Ramström, O. *Bioorg. Med. Chem. Lett.* **2005,** *15*, 2707-2710.
16. André, S.; Pei, Z. C.; Siebert, H. C.; Ramström, O.; Gabius, H. J. *Bioorg. Med. Chem.* **2006,** *14*, 6314-6326.
17. Murthy, B. N.; Sinha, S.; Surolia, A.; Jayaraman, N.; Szilágyi, L.; Szabó, I.; Kövér, K. E. *Carbohydr. Res.* **2009,** *344*, 1758-1763.
18. Potrzebowski, M. J.; Michalska, M.; Blaszczyk, J.; Wieczorek, M. W.; Ciesielski, W.; Kazmierski, S.; Pluskowski, J. *J. Org. Chem.* **1995,** *60*, 3139-3148.
19. Hürzeler, M.; Bernet, B.; Vasella, A. *Helv. Chim. Acta* **1992,** *75*, 557-588.
20. Brito, I.; López-Rodriguez, M.; Bényei, A.; Szilágyi, L. *Carbohydr. Res.* **2006,** *341*, 2967-2972.
21. Brito, I.; Szilágyi, L.; López-Rodríguez, M. *Acta Cryst.* **2008,** *E64*, 2472-2473.
22. Grayson, E. J.; Ward, S. J.; Hall, A. L.; Rendle, P. M.; Gamblin, D. P.; Batsanov, A. S.; Davis, B. G. *J. Org. Chem.* **2005,** *70*, 9740-9754.
23. Kurtán, T.; Pescitelli, G.; Salvadori, P.; Kenéz, A.; Antus, S.; Szilágyi, I.; Illyés, T. Z.; Szabó, I. *Chirality* **2008,** *20*, 379-385.
24. Kövér, K. E.; Szilágyi, L.; Batta, G.; Uhrín, D.; Jiménez-Barbero, J. In *Comprehensive Natural Products II Chemistry and Biology;* Mander, L.; Lui, H.-W. Eds.; Elsevier: Oxford, 2010; pp. 197-246.
25. Mulloy, B.; Frenkiel, T. A.; Davies, D. B. *Carbohydr. Res.* **1988,** *184*, 39-46.
26. Tvaroska, I.; Hricovini, M.; Petrakova, E. *Carbohydr. Res.* **1989,** *189*, 359-362.
27. Best, R. B.; Jackson, G. E.; Naidoo, K. J. *Spectroscopy Lett.* **2002,** *35*, 625-632.
28. Bock, K.; Duus, J. O. *J. Carbohydr. Chem.* **1994,** *13*, 513-543.
29. Rockwell, G. D.; Grindley, T. B. *J. Am. Chem. Soc.* **1998,** *120*, 10953-10963.
30. Barnett, C. B.; Naidoo, K. J. *J. Phys. Chem. B* **2008,** *112*, 15450-15459.
31. Asensio, J. L.; Jimenezbarbero, J. *Biopolymers* **1995,** *35*, 55-73.
32. Cumming, D. A.; Carver, J. P. *Biochemistry* **1987,** *26*, 6664-6676.
33. Thaning, J.; Stevensson, B.; Ostervall, J.; Naidoo, K. J.; Widmalm, G.; Maliniak, A. *J. Phys. Chem. B* **2008,** *112*, 8434-8436.

34. Wormald, M. R.; Petrescu, A. J.; Pao, Y. L.; Glithero, A.; Elliott, T.; Dwek, R. A. *Chem. Rev.* **2002,** *102*, 371-386.
35. Best, R. B.; Jackson, G. E.; Naidoo, K. J. *J. Phys. Chem. B* **2002,** *106*, 5091-5098.
36. Best, R. B.; Jackson, G. E.; Naidoo, K. J. *J. Phys. Chem. B* **2001,** *105*, 4742-4751.
37. Bax, A.; Davis, D. G. *J. Magn. Reson.* **1985,** *63*, 207-213.
38. Bax, A. *J. Magn. Reson.* **1988,** *77*, 134-147.
39. Farmer, B. T.; Macura, S.; Brown, L. R. *J. Magn. Reson.* **1987,** *72*, 347-352.
40. Szilagyi, L.; Pusztahelyi, Z. S. *Magn. Reson. Chem.* **1992,** *30*, 107-117.
41. Neuhaus, D.; Keeler, J. *J. Magn. Reson.* **1986,** *68*, 568-574.
42. Neuhaus, D.; Williamson, M. P. *The Nuclear Overhauser Effect in Structural and Conformational Analysis.*; VCH Publishers, 1989.
43. Mo, F.; Jensen, L. H. *Acta Cryst.B-Struct. Sci.* **1978,** *34*, 1562-1569.
44. Szilágyi, L.; Forgó, P. *Carbohydr. Res.* **1993,** *247*, 129-144.
45. Uhrin, D.; Batta, G.; Hruby, V. J.; Barlow, P. N.; Kover, K. E. *J. Magn. Reson.* **1998,** *130*, 155-161.
46. Williamson, R. T.; Marquez, B. L.; Gerwick, W. H.; Kövér, K. E. *Magn. Reson. Chem.* **2000,** *38*, 265-273.
47. Kövér, K. E.; Batta, G.; Fehér, K. *J. Magn. Reson.* **2006,** *181*, 89-97.
48. Fraser, R. R.; Kaufman, M.; Morand, P.; Govil, G. *Can. J. Chem.* **1969,** *47*, 403.
49. Kövér, K. E.; Beke, T.; Lipták, A.; Perczel, A. *J. Comput. Chem.* **2009,** *30*, 540-550.
50. Naidoo, K. J.; Brady, J. W. *Chem. Phys.* **1997,** *224*, 263-273.
51. Kuttel, M.; Brady, J. W.; Naidoo, K. J. *J. Comput. Chem.* **2002,** *23*, 1236-1243.
52. Venter, G. A.; Matthews, R. P.; Naidoo, K. J. *Mol. Simul.* **2008,** *34*, 391-401.
53. Brooks, B. R.; Brooks, C. L.; Mackerell, A. D.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caflisch, A.; Caves, L.; Cui, Q.; Dinner, A. R.; Feig, M.; Fischer, S.; Gao, J.; Hodoscek, M.; Im, W.; Kuczera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R. W.; Post, C. B.; Pu, J. Z.; Schaefer, M.; Tidor, B.; Venable, R. M.; Woodcock, H. L.; Wu, X.; Yang, W.; York, D. M.; Karplus, M. *J. Comput. Chem.* **2009,** *30*, 1545-1614.
54. Strader, M. L.; Feller, S. E. *J. Phys. Chem. A* **2002,** *106*, 1074-1080.
55. Feller, S. E.; Zhang, Y. H.; Pastor, R. W.; Brooks, B. R. *J. Chem. Phys.* **1995,** *103*, 4613-4621.
56. Carpenter, G. A.; Grossberg, S. *Appl. Optics* **1987,** *26*, 4919-4930.
57. Pao, Y. H. *Adaptive Pattern Recognition and Neural Networks.*; Addison-Wesley: New York, 1989.
58. Karpen, M. E.; Tobias, D. J.; Brooks, C. L. *Biochemistry* **1993,** *32*, 412-420.
59. Burns, J. A.; Whitesides, G. M. *J. Am. Chem. Soc.* **1990,** *112*, 6296-6303.
60. Woodcock, H. L.; Brooks, B. R.; Pastor, R. W. *J. Am. Chem. Soc.* **2008,** *130*, 6345-+.
61. Gonzalez-Outeirino, J.; Kirschner, K. N.; Thobhani, S.; Woods, R. J. *Can. J. Chem.* **2006,** *84*, 569-579.
62. Kirschner, K. N.; Woods, R. J. *Proc. Nat. Acad. Sci.USA* **2001,** *98*, 10541-10545.
63. Májer, G.; Borbás, A.; Illyés, T.-Z.; Szilágyi, L.; Bényei, A. C.; Lipták, A. *Carbohydr. Res.* **2007,** *342*, 1393-1404.

Conformational preferences in diglycosyl disulfides: NMR and molecular modelling studies

Krisztina Fehér, Richard P. Matthews, Katalin E. Kövér, Kevin J. Naidoo, László Szilágyi*

The disulfide torsion angles (Ψ) sample values close to $+90^{o}$ or -90^{o} ($+g$ or $-g$) and up to four allowed conformational regions around Φ and Ω were identified for each of the Ψ values.



HIGHLIGHTS

Conformations of diglycosyl disulfides were studied by NMR and computations

The disulfide torsion angles (Ψ) are close to $+90^{o}$ or -90^{o} ($+g$ or $-g$)

Up to four conformational regions around Φ and Ω are allowed for each Ψ values

No significant intramolecular hydrogen bonds were revealed in DMSO solution.