Pure Appl. Chem., Vol. 85, No. 9, pp. 1759–1770, 2013.
http://dx.doi.org/10.1351/PAC-CON-12-10-17

Glycan synthesis, structure, and dynamics: A selection*

Robert Pendrill, K. Hanna M. Jonsson, and Göran Widmalm‡

Department of Organic Chemistry, Arrhenius Laboratory, Stockholm University, SE-106 91 Stockholm, Sweden

Abstract: Glycan structural information is a prerequisite for elucidation of carbohydrate function in biological systems. To this end we employ a tripod approach for investigation of carbohydrate 3D structure and dynamics based on organic synthesis; different experimental spectroscopy techniques, NMR being of prime importance; and molecular simulations using, in particular, molecular dynamics (MD) simulations. The synthesis of oligosaccharides in the form of glucosyl fluorides is described, and their use as substrates for the Lam16A E115S glucosyl synthase is exemplified as well as a conformational analysis of a cyclic β-(1 → 3)-heptaglucan based on molecular simulations. The flexibility of the *N*-acetyl group of aminosugars is by MD simulations indicated to function as a gatekeeper for transitions of glycosidic torsion angles to other regions of conformational space. A novel approach to visualize glycoprotein (GP) structures is presented in which the protein is shown by, for example, ribbons, but instead of stick or space-filling models for the carbohydrate portion it is visualized by the colored geometrical figures known as CFG representation in a 3D way, which we denote 3D-CFG, thereby effectively highlighting the sugar residues of the glycan part of the GP and the position(s) on the protein.

Keywords: carbohydrates; glycosynthase; molecular dynamics; structure; synthesis.

INTRODUCTION

Carbohydrates are involved in structural aspects and many recognition processes where they interact with protein molecules of various kinds [1]. The immune system has to differentiate between self, being the host, altered-self, in the form of tumors, and non-self, as, e.g., microbial invaders [2]. Viral adhesion can exploit specific host factors, and in carbohydrate binding processes there is a complex interplay of the host's defense system and the virus or microbe trying to evade it [3]. These processes are dynamic, and on the molecular level the same molecule may show different 3D structures depending on the interaction partners [4], underscoring the importance of dynamical properties of oligo- and polysaccharides. The structure and dynamics of glycoproteins (GPs) in solution are readily determined by NMR spectroscopy, whereas in the solid state the structural investigation of GPs is often hampered by the dynamic or heterogeneous aspects of the carbohydrate portion [5]. Effective presentation of the structures [6] is important in conveying an understanding of the molecular processes that take place or the functional implications one is describing.

Pure Appl. Chem.* **85, 1759–1900 (2013). A collection of invited papers based on presentations at the 26th International Carbohydrate Symposium (ICS 2012), Madrid, Spain, 22–27 July 2012.
‡Corresponding author

One class of proteins interacting with carbohydrate molecules is glycosidases, also known as glycohydrolases, which catalytically degrade oligo- and polysaccharides by hydrolysis of glycosidic bonds [7]. The reaction route for retaining glycosidases follows a double-displacement mechanism in which a glycosyl-enzyme intermediate is formed [8]. The canonical two-step mechanism involves protonation of the glycosidic oxygen by a carboxyl group of an amino acid with the concomitant or subsequent nucleophilic attack of a carboxylate group from a differently positioned amino acid leading to the cleavage of the glycosidic bond and the formation of a covalent enzyme-substrate intermediate. Subsequently, water, or a sugar residue in the case of a *trans*-glycosylation reaction, assisted by the conjugate base of the first amino acid attacks the anomeric carbon thereby producing a reducing sugar or a glycoside, respectively, together with cleavage of the covalent bond between the substrate and the second amino acid. These retaining glycosidases can be used in chemo-enzymatic synthesis if activated glycosyl donors are used. For example, a *Bacillus licheniformis* β-(1 → 3)-(1 → 4)-glucanase was used for the self-condensation of β-laminaribiosyl fluoride to produce oligomers of different lengths [9]. Note that, since the glycosidase contains an active-site catalytic nucleophile in the form of, for example, a glutamate residue, the substrate used had the β-configuration at the reducing end of the disaccharide. The active nucleophile in the enzyme may be mutated into a non-nucleophilic residue such that this part of the reaction mechanism no longer is available, i.e., a potential glycosynthase [10] has been made and a corresponding activated glycosyl donor should then have the α-anomeric configuration. The use of glycosyl fluorides [11] in chemo-enzymatic synthesis employing glycosynthases [12–14] offers high stereo- and regioselectivity in the reactions.

CHEMO-ENZYMATIC SYNTHESIS AND CONFORMATIONAL ANALYSIS OF A CYCLIC β-(1 → 3)-HEPTAGLUCAN

The oligosaccharides to be used as enzyme substrates, α-laminariheptaosyl fluoride (αL7F) and α-laminaribiosyl fluoride (αL2F), were synthesized essentially as described for xylogluco-oligosaccharides by Fauré et al. [15]. The oligosaccharides (L7 or L2) were peracetylated and subsequently selectively *O*-deacetylated at the reducing end (Fig. 1). The fluoro functionality was introduced with diethylaminosulfur trifluoride (DAST) to give an α/β-anomeric mixture with mainly the β-configuration. Treatment with HF/pyridine resulted in the anomerically pure α-glycosyl fluoride. The oligosaccharides were readily deprotected by treatment under basic conditions and purified by size-exclusion chromatography to give the desired products αL7F and αL2F.

The mutation of the 1,3(4)-β-D-glucanase Laminarinase 16A (Lam16A) from *Phanerochaete chrysosporium* [16] to a glycosynthase was made by changing the active site glutamate residue to a serine residue. This form of the enzyme is denoted Lam16A E115S. Employing an activated glycosyl donor, αL7F, we were able to chemo-enzymatically synthesize a cyclic β-(1 → 3)-heptaglucan (Fig. 1) [17]. The intramolecular nucleophilic attack on the C1 atom carrying a fluorine atom by the O3 atom of the terminal glucosyl residue is assisted by the catalytic acid/base Glu120 (Fig. 2). The disaccharide αL2F, which in preliminary studies [18] was shown to self-condensate by the action of Lam16A E115S to larger oligosaccharides, some of which presumably are cyclic, demonstrate that Lam16A E115S accepts different substrates that lead product formation.

Molecular dynamics (MD) with explicit water molecules and Langevin dynamics (LD) simulations were carried out on the cyclic β-(1 → 3)-heptaglucan (cL7), extending the previous MD simulations [17] that employed an MM3 force field, to study its structure and conformational preferences. The molecular simulations of cL7 presented herein utilized a recently developed CHARMM-based force field [19]. Simulations were carried out at 310 K for 150 ns and the glycosidic torsion angles were monitored. Sufficient convergence was obtained for simulations of this duration, and the probability distributions for ϕ_H (H1'-C1'-O3-C3) and ψ_H (C1'-O3-C3-H3) are presented in Fig. 3. Further analysis revealed that two conformational states were present with $\phi_H = 29(15)^\circ$ and $\psi_H = -4(21)^\circ$ for the major

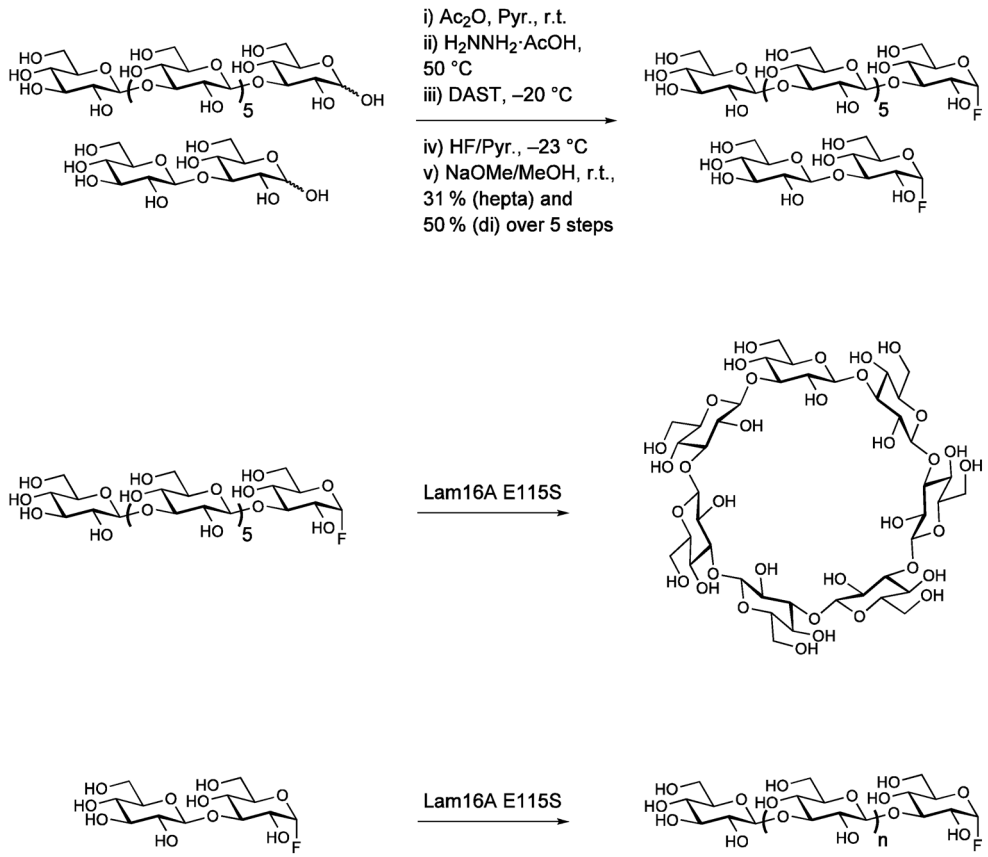

Fig. 1 Synthesis pathway to α-glycosyl fluoride donors (top). Formation of cyclic laminariheptaose (cL7) from αL7F upon Lam16A E115S catalyzed self-condensation (middle). Oligomerization of αL2F under the same conditions (bottom).

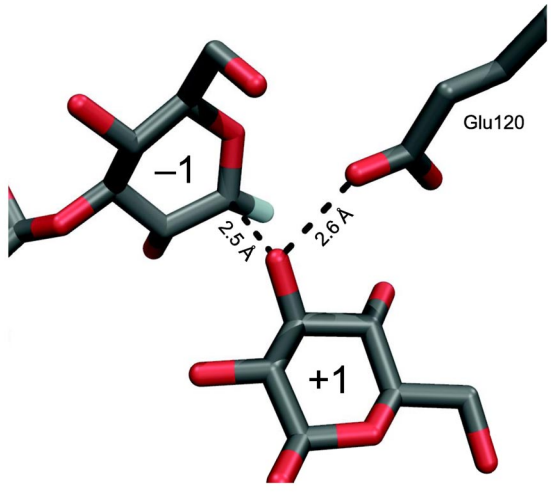

Fig. 2 The active site of Lam16A E115S showing the glucosyl residues in sites +1 and −1 of the bound substrate (αL7F).

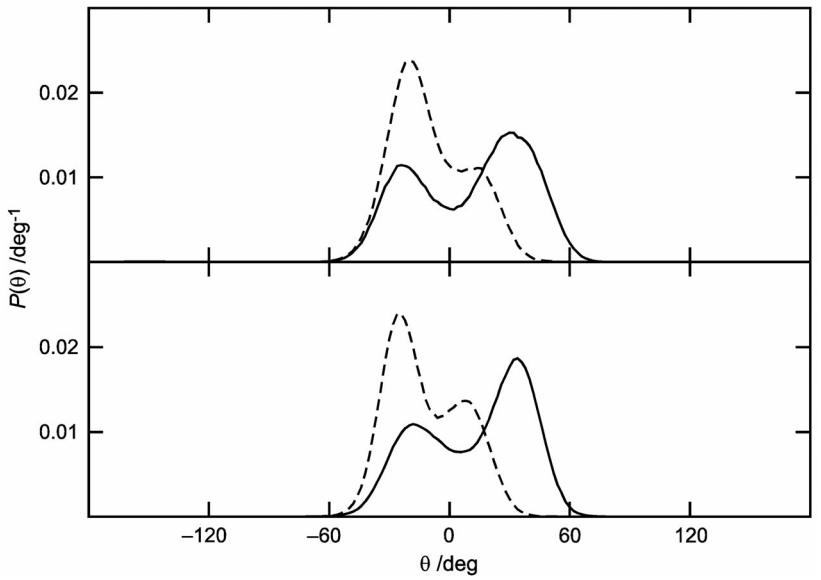

Fig. 3 Probability distribution functions of the glycosidic torsion angles in cL7 from the MD simulation (top) and the LD simulation (bottom); ϕ_H (solid line) and ψ_H (dashed line).

state (RMSD in parenthesis) and $\phi_H = -22(12)^o$ and $\psi_H = -21(10)^o$ for the minor state in the MD simulation; in the LD simulation, the major state is described by $\phi_H = 29(14)^o$ and $\psi_H = -6(20)^o$ and the minor one by $\phi_H = -19(12)^o$ and $\psi_H = -25(9)^o$.

For this geometrically restrained system, the latter simulation methodology, which is several orders of magnitude more rapid [20], gives structural results that are highly similar. These simulations show that two conformational states are populated, which also was the case when the MM3 force field was used [17]. *Trans*-glycosidic heteronuclear coupling constants are sensitive to conformational aspects via Karplus-type relationships [21]. We calculated these from the MD simulation which resulted in $^3J_{H1',C3} = 4.4$ Hz (related to ϕ_H) and $^3J_{C1',H3} = 5.5$ Hz (related to ψ_H) and for the LD simulation, resulting in $^3J_{H1',C3} = 4.6$ Hz and $^3J_{C1',H3} = 5.4$ Hz, respectively. These LD simulations were carried out with a dielectric constant $\varepsilon = 3$; using either a lower value of 1 or higher values of 10 or 80 led to larger deviations for $^3J_{CH}$ values in comparison to those calculated from the MD simulation with explicit water molecules as solvent, consistent with previous investigations that a low value of the dielectric constant should be used if this type of force field is applied without explicit solvent [22]. Compared to those of methyl α-laminaribioside having $^3J_{H1',C3} = 3.9$ Hz and $^3J_{C1',H3} = 4.8$ Hz [23], the $^3J_{C,H}$ values for cL7 do differ somewhat and indicate, together with ^{13}C glycosylation shifts [17], that cL7 conformational preferences are altered compared to methyl α-laminaribioside or laminariheptaose.

THE *N*-ACETYL GROUP AS A GATEKEEPER TO ADDITIONAL CONFORMATIONAL REGIONS

In two previous conformational studies of *N*-acetyl-containing oligosaccharides using MD simulations, we have observed a coupling of the *N*-acetyl torsion with the conformational preference at nearby glycosidic torsion angles. In the trisaccharide β-D-Glc*p*NAc-(1 → 3)-β-D-Gal*p*-(1 → 4)-β-D-Glc*p*NAc-OMe (Fig. 4), a model for the central region of the Le^aLe^x hexasaccharide, we observed a minor conformation appearing to be more easily accessible when the nearby *N*-acetyl group was in a conformation having $\tau \approx 165^o$ [24]. Upon closer investigation, a plausible causal relationship was found involving

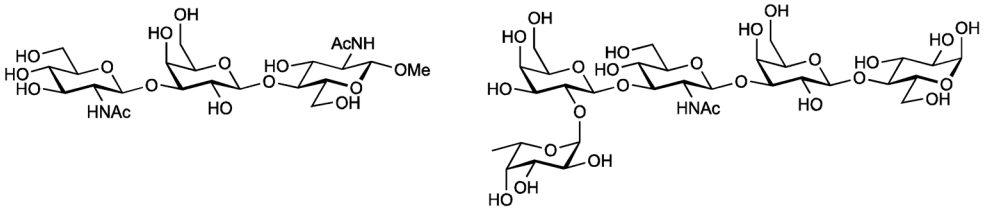

Fig. 4 Schematic structures of the trisaccharide β-D-Glc*p*NAc-(1 → 3)-β-D-Gal*p*-(1 → 4)-β-D-Glc*p*NAc-OMe and the pentasaccharide α-L-Fuc*p*-(1 → 2)-β-D-Gal*p*-(1 → 3)-β-D-Glc*p*NAc-(1 → 3)-β-D-Gal*p*-(1 → 4)-D-Glc*p*.

a sequence of hydrogen bonds, each stabilizing one of several sub-minima occurring during the transition path between the major and minor conformation. This chain of events was initiated by the disruption of the interresidue O5'···HO3 hydrogen bond and the simultaneous formation of a hydrogen bond with the same donor atom but involving the *N*-acetyl carbonyl oxygen instead. Since the latter hydrogen bond is only accessible in a small region of the conformational space for the *N*-acetyl group, it appears that this functional group effectively regulates the conformation at other flexible sites in the molecule [24].

Similarly, the human milk pentasaccharide LNF-1, α-L-Fuc*p*-(1 → 2)-β-D-Gal*p*-(1 → 3)-β-D-Glc*p*NAc-(1 → 3)-β-D-Gal*p*-(1 → 4)-D-Glc*p* (Fig. 4), was shown in an MD simulation to populate a minor conformational state having negative values for the ϕ_H and ψ_H glycosidic torsion angles at the α-(1 → 2)-linkage. The single excursion to this conformational state coincided with the *N*-acetyl torsion being in its minor conformation with $\tau \approx -60°$, and it was thus hypothesized that the *N*-acetyl group acted as a conformational switch [25].

To test the hypothesis that conformational regulation by the *N*-acetyl group occurred in these molecules, additional MD simulations were performed and the results are presented herein. By restraining the *N*-acetyl torsional angle to the different accessible states, the effect on glycosidic torsional degrees of freedom can be observed. Thus, for LNF-1 it was found that when the *N*-acetyl torsion was restrained at $\tau = -60°$, there was a clear increase in the number of excursions to the secondary conformational state at the α-L-Fuc*p*-(1 → 2)-β-D-Gal*p*-linkage compared to the case when the torsion was restrained at $\tau = 90°$, i.e., close to the conformational state determined by experiment [26]. The population having $\phi_H < 0°$ and $\psi_H < 0°$ was found to be 32 % in the simulation having $\tau = -60°$ compared to 4 % in the $\tau = 90°$ simulation. Since the dynamics at the *N*-acetyl torsion in this molecule was found to be relatively slow compared to interconversions at the α-(1 → 2)-linkage, this effect is also clearly visible in the non-restrained trajectory, shown in Fig. 5. The *N*-acetyl torsion makes a transition to the $\tau = -60°$ conformation lasting for approximately 35 ns during the simulation, and throughout this visit it is apparent that the α-L-Fuc*p*-(1 → 2)-β-D-Gal*p*-linkage is more flexible than when the *N*-acetyl group populates its major conformation having $\tau = 90°$.

While it is possible in the case of LNF-1 to test the hypothesis of conformational switching by the *N*-acetyl group without resorting to restraining the torsion angle because of the slow dynamics for the *N*-acetyl torsion, in the trisaccharide the dynamics was found to be much faster with $\tau_N \approx 0.7$ ns as deduced from a number correlation function, making it difficult to assess the impact of the *N*-acetyl conformation on the dynamics at the β-D-Gal*p*-(1 → 4)-β-D-Glc*p*NAc-linkage. However, comparing simulations restraining the *N*-acetyl torsion angle at $\tau = -60°$ and at $\tau = 165°$ respectively, the effect is clear; the population of the minor non-*exo*-anomeric ϕ_H conformation at the β-(1 → 4)-linkage is 9 % in the former and 19 % in the latter simulation. The major *exo*-anomeric ϕ_H conformation is stabilized by the O5'···HO3 hydrogen bond between the two sugar residues of the *N*-acetyllactosamine structural element which must be broken in order to reach the minor non-*exo*-anomeric ϕ_H conformation. With $\tau = 165°$, the *N*-acetyl carbonyl oxygen is suitably placed for accepting hydrogen bonding from HO3 and hence is able to compete with the inter-residual O5'···HO3 hydrogen bond. While the C=O···HO3

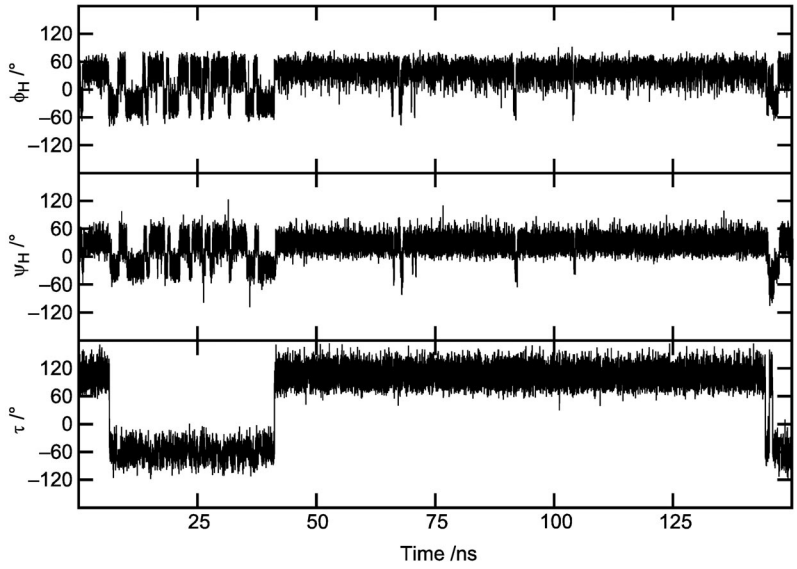

Fig. 5 Time evolution of the glycosidic torsion angles for the α-L-Fuc*p*-(1 → 2)-β-D-Gal*p*-linkage and the *N*-acetyl torsion angle in LNF-1 from an MD simulation without any restraints.

hydrogen bond is not found at all in the $\tau = -60°$ simulation, the population is 13 % in the simulation with $\tau = 165°$. Conversely, for the inter-residual O5'···HO3 hydrogen bond there is a corresponding decline in the population, from 34 % in the former simulation to 25 % in the latter. Thus, with the *N*-acetyl group in the conformation having $\tau = 165°$, the β-(1 → 4)-linkage is made more flexible by reducing the population of the rigidifying O5'···HO3 hydrogen bond.

These results support the idea that there exists a coupling in some cases between the *N*-acetyl conformational state and torsional flexibility at other linkages having a larger impact on the overall molecular shape. It remains to confirm this effect experimentally and to understand the biophysical ramifications of such a conformational coupling.

3D-CFG REPRESENTATION OF GLYCOPROTEIN STRUCTURES

In the area of structural biology, the use of cartoon or ribbon representations of protein secondary and tertiary structure is well established, much due to the superior clarity obtained when atomistic detail is omitted. The glycan part, however, is exclusively represented using sticks or space-filling models, obscuring the identity of the constituting residues. In two dimensions, symbolic representations [27] such as the CFG notation [28] has found widespread use, e.g., in the annotation of mass spectra.

We have for GPs and other glycoconjugates developed a 3D interpretation of the symbol library based on the CFG representation which we denote 3D-CFG. The TbpA-transferrin complex from *Neisseria meningitidis* [29] is shown in Fig. 6 by ribbons for the protein part and a space-filling representation for the glycan part. It is evident that carbohydrates are present and where they are linked to the protein. However, it is not possible to rapidly identify the carbohydrate structure and its constituent monosaccharides. Visualizing the sugars by the 3D-CFG presentation instead immediately shows what they are. The clarification is apparent and this mixed carbohydrate/protein representation efficiently mediates the structural information required.

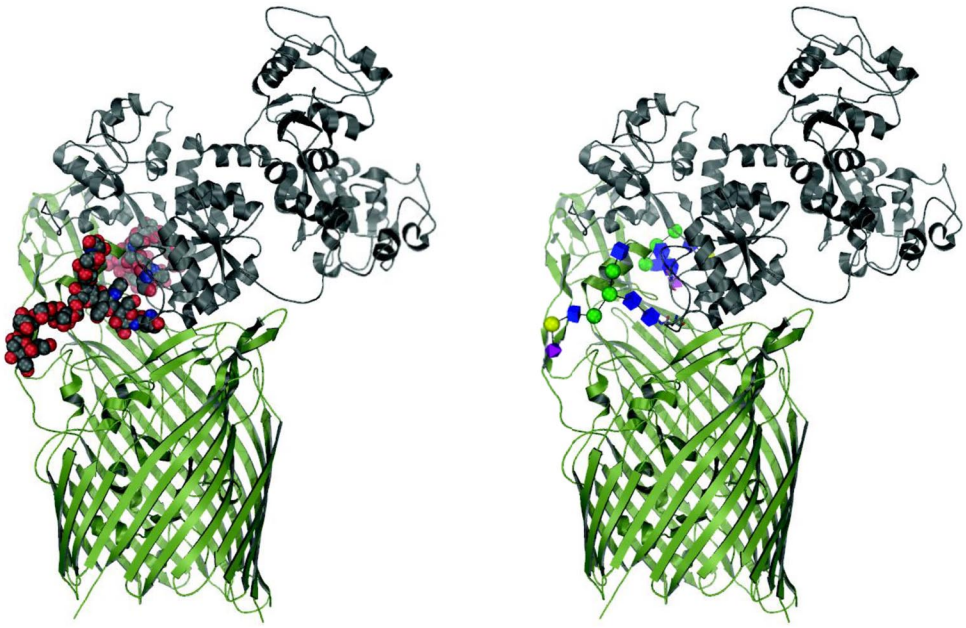

Fig. 6 TbpA-transferrin complex from *N. meningitides*. The glycan part with space-filling representation is shown to the left, and the 3D-CFG representation is presented to the right.

CONCLUSIONS

Chemo-enzymatic synthesis using glycosynthases is an excellent complement to classical organic synthesis since very high stereo- and regioselectivity is achieved in the reactions. Furthermore, the possibility of utilizing different mutations in the active site enables the tailoring of enzymes for various applications, in contrast to just relying on the inherent functionality of the enzyme. We have shown that suitable substrates in the form of glycosyl fluorides can be made by organic synthesis in a few steps from biotechnologically produced oligosaccharides and that different types of oligosaccharides are the result of glycosynthase action, some of which are not easily obtained by other methods.

The methodology of restraining the *N*-acetyl torsion angle is useful when investigating possible conformational coupling where one would otherwise need to wait for a transition to occur into the conformation of interest, or if dynamics are too fast to discern possible remote consequences. In order to study the potential biological importance of such switching using molecular simulations, we envision that the use of such restraints can greatly reduce the amount of computational resources required, as well as allowing the testing of hypotheses in a well-defined manner.

For GPs and other glycoconjugates we have developed the 3D-CFG representation which efficiently and instantly makes it possible to identify the carbohydrate structure and its constituent monosaccharides in these complex molecules.

MATERIALS AND METHODS

General procedures

Laminariheptaose (L7) and laminaribiose (L2) were purchased from Seikagaku Corporation, Tokyo, Japan. Thin layer chromatography (TLC) analysis was performed on precoated Merck 60 F_{254} plates and developed with 8 % H_2SO_4. For purification by column chromatography Davisil silica medium (35–70 micron) was used. NMR spectroscopy was performed on Bruker Avance 400 and 500 MHz

spectrometers at ambient temperature in $CDCl_3$ and at 5 °C in D_2O. ^1H and ^{13}C NMR chemical shifts are referenced to the solvent $CDCl_3$ (δ_H 7.26, residual $CHCl_3$; δ_C 77.23) and for D_2O solutions to external TSP (δ_H 0.00) and external dioxane in D_2O (δ_C 67.40). NMR assignments of the oligosaccharides were performed by 1D ^1H and ^{13}C NMR experiments as well as ^1H,^{13}C-HSQC, ^1H,^{13}C-HETCOR, ^1H,^1H-TOCSY (mixing times of 10–60 ms), phase sensitive ^1H,^1H-DQF-COSY, ^1H,^{13}C-HMBC and ^1H,^{13}C-H2BC experiments. High-resolution mass spectrometry (MS) analysis was performed on a Bruker Daltonics ESI-TOF spectrometer in the positive mode.

α-Laminariheptaosyl fluoride (αL7F)

In a mixture of acetic anhydride and pyridine (1:1, 10 mL), L7 (50 mg, 0.043 mmol) was stirred at r.t. for 17 h. MeOH was added at 0 °C, the mixture allowed to attain r.t. and the solvent was evaporated, co-evaporated with toluene and the solid material was vacuum-dried. The crude product was used without further purification in the next step. Peracetylated L7 was dissolved in dimethyl formamide (DMF) (5 mL) at r.t.; hydrazine acetate (9 mg, 2.5 eq) was added and the mixture heated to 50 °C. After 2 h the starting material was not present by TLC and the reaction mixture was cooled to r.t. and diluted with ethyl acetate. The organic phase was washed with brine and water, dried over Na_2SO_4 and evaporated. The solid material was dissolved in toluene and co-evaporated several times. The crude material was vacuum-dried and used in the next step without further purification. Analysis of the crude product by ^1H NMR showed an upfield shift for the anomeric proton resonances by ~1 ppm; hence the anomeric position had been deprotected.

Crude material from the previous step was dissolved in dry dichloromethane (DCM, 5 mL), argon was bubbled through the solution which was subsequently cooled to −32 °C. DAST (30 μL, 8 equiv) was added drop-wise with an immediate color change from brown–yellow to pinkish–yellow. The mixture was allowed to attain r.t. and after 2.5 h the reaction was complete according to MS. MeOH was added to quench the reaction, solvent was evaporated and the crude material vacuum-dried. ^1H NMR of crude material showed full conversion to glycosyl fluoride product in an α/β-anomeric mixture.

The solid crude material was dissolved in dry DCM (4 mL) and transferred to a plastic vial. Argon was bubbled through the solution which was subsequently cooled to −22 °C. HF in pyridine (70 %, 0.5 mL) was added in one batch. The mixture was stirred for 3.5 h while the temperature was slowly raised to r.t. The reaction mixture was diluted with DCM and an ice-water mixture. The organic phase was washed with sat. $NaHCO_3$ (aq.) and dried over Na_2SO_4. Crude material was purified on a column of silica (toluene/acetone 3:1) resulting in 40 mg of peracetylated αL7F (44 % over 4 steps) containing a minor amount hydrolysis product. ^1H NMR (selected resonance) δ_H: 5.65, H1, $J_{H1,H2}$ = 2.8 Hz, $J_{F,H1}$ = 53 Hz; ^{13}C NMR (selected resonance) δ_C: 104.8, C1, $J_{F,C1}$ = 210 Hz; ESI-MS: *m/z* $[M+Na]^+$ 2101.5912, expected 2101.5976.

Peracetylated αL7F (48 mg, 0.023 mmol) was dissolved in MeOH (8 mL) and NaOMe/MeOH (1 M, 0.8 mL) was added to the mixture, which was stirred at r.t. for 45 min. Dowex 50WX8 (16-40 mesh) in the H^+-form was added until pH ≈ 7. The solution was filtered and evaporated to dryness. The crude material was dissolved in H_2O and filtered through a short plug of P-2 gel, giving after freeze-drying 29 mg crude product. Purification via size-exclusion chromatography (Bio-Gel P-2 gel) gave 19 mg of αL7F (0.016 mmol, 70 %). ESI-MS: *m/z* $[M+Na]^+$ 1177.3659, expected 1177.3652. ^1H and ^{13}C NMR assignments are compiled in Table 1.

Table 1 ^{1}H and ^{13}C NMR chemical shifts (ppm) of the resonances from αL7F (top) and αL2F (bottom) in D_2O at 5 °C. Sugar residues are denoted by capital letters starting from the reducing end. $^{3}J_{H1,H2}$ values are given in Hz in square brackets, $J_{F,H}$ values in curly brackets and $J_{F,C}$ values in parentheses.

Sugar residue		$^{1}H/^{13}C$					
		1	2	3	4	5	6
→ 3)-α-D-Glc*p*-F	A	5.69	3.82	3.98	3.62	3.87	3.87, 3.83
		[2.7] {53}					
		108.2	71.6	81.3	67.7	74.7	60.8
		(223)	(25)			(2.7)	
→ 3)-β-D-Glc*p*-(1 →	B	4.774	3.56	3.78	3.52	3.50	3.93, 3.75
		[8]					
		103.3	74.1	84.7	68.8	76.4	61.4
→ 3)-β-D-Glc*p*-(1 →	C-F	4.81	3.55	3.79	3.52	3.50	3.93, 3.75
		[8]					
		103.4	74.1/74.2	84.5	68.8	76.4	61.4
β-D-Glc*p*-(1 →	G	4.769	3.35	3.52	3.40	3.48	3.93, 3.75
		[8]					
		103.6	74.2	76.3	70.4	76.8	61.4
→ 3)-α-D-Glc*p*-F	A	5.70	3.83	3.98	3.63	3.87	3.89, 3.84
		[2.8] {53}					
		108.3	71.7	81.7	67.9	74.9	61.0
		(223)	(25)			(3)	
β-D-Glc*p*-(1 →	B	4.75	3.37	3.53	3.42	3.48	3.94, 3.74
		[7.9]					
		103.7	74.3	76.5	70.5	76.9	61.6

α-Laminaribiosyl fluoride (αL2F)

In a mixture of acetic anhydride and pyridine (1:1, 10 mL), L2 (49 mg, 0.14 mmol) was stirred at r.t. for 18 h. MeOH was added at 0 °C, the mixture allowed to attain r.t. and then the solvent was evaporated, co-evaporated with toluene and the solid material vacuum-dried. The crude product was used without further purification in the next step. Peracetylated L2 was dissolved in DMF (5 mL) at r.t.; hydrazine acetate (42 mg, 3.3 eq) was added and the mixture heated to 50 °C. After 1 h the starting material was not present by TLC and the reaction mixture was cooled to r.t. and diluted with ethyl acetate. The organic phase was washed with brine and water, dried over $MgSO_4$, and evaporated. The solid material was dissolved in toluene and co-evaporated several times before it was vacuum-dried.

Crude material from the previous step was dissolved in dry DCM (6 mL), N_2 was bubbled through the solution which was subsequently cooled to −32 °C. DAST (60 µL, 3.2 equiv) was added drop-wise. Temperature was allowed to attain −20 °C and after 30 min the reaction was complete according to TLC. MeOH was added to quench the reaction and the mixture was allowed to attain r.t. The solvent was evaporated and the crude material co-evaporated with toluene and then vacuum-dried.

The α/β-anomeric mixture was dissolved in dry DCM (5 mL) and transferred to a plastic vial. N_2 was bubbled through the solution which was subsequently cooled to −23 °C. HF in pyridine (70 %, 2 mL) was added in one batch. The mixture was stirred for 3.5 h while the temperature was allowed to attain r.t. The reaction mixture was diluted with DCM and an ice-water mixture. The organic phase was washed with sat. $NaHCO_3$ (aq.) and dried over $MgSO_4$. Crude material was purified on a column of silica (toluene/EtOAc, gradient 4:1 → 2:1) resulting in 70 mg of peracetylated αL2F (78 % over 4 steps). ^{1}H NMR (selected resonances) δ_H: 5.65, H1, $J_{H1,H2}$ = 2.7 Hz, $J_{F,H1}$ = 53 Hz; 4.65, H1', $J_{H1',H2'}$ =

8.3 Hz; ^{13}C NMR (selected resonances) δ_C: 104.1, C1, $J_{F,C1}$ = 228 Hz; 100.9, C1'. ESI-MS: *m/z* [M + Na]$^+$ 661.1784, expected 661.1750.

Peracetylated αL2F (70 mg, 0.11 mmol) was dissolved in NaOMe/MeOH (0.05 M, 6 mL) and stirred at r.t. for 1.5 h. Dowex 50WX8 (16-40 mesh) in the H$^+$-form was added until pH ≈ 7. The solution was filtered and evaporated to dryness. The crude material was dissolved in H_2O and eluted with H_2O through a short SepPak C-18 column. Purification via size-exclusion chromatography (Superdex 30, prep. grade) gave 25 mg of αL2F (0.07 mmol, 64 %). ESI-MS: *m/z* [M + Na]$^+$ 367.1010, expected 367.1011. ^{1}H and ^{13}C NMR assignments are compiled in Table 1.

Molecular simulations

The ϕ_H and ψ_H glycosidic torsion angles are defined by the pertinent consecutive C-O-C-H sequence of atoms and the τ torsion angle by the C1-C2-N-C atoms.

MD simulations were performed using CHARMM [30] (parallel version, 35b4). The CHARMM2011 [31] force field was used for the simulation of cL7, while for LNF-1 and the trisaccharide the PARM22/SU01 [32] force field was used.

The oligosaccharides were built from internal coordinates using the CHARMM2011 force field and subsequently the potential energy was minimized. For cL7, a harmonic restraint of 1 kcal·mol^{-1}·Å^{-1} was used to maintain dihedral angles at the values specified in the internal coordinates table while applying the steepest descent method, 1000 steps, followed by adopted-basis Newton-Raphson, 5000 steps or until the RMS gradient was less than 0.01 kcal·mol^{-1}·Å^{-1}. The minimization was subsequently repeated without the dihedral restraints. For the two linear oligosaccharides, minimization was performed using steepest descent, 100 steps, followed by adopted-basis Newton-Raphson, 1000 steps or until the RMS gradient was less than 0.01 kcal·mol^{-1}·Å^{-1}.

The oligosaccharides were then placed in pre-equilibrated cubic boxes with 29.972 Å sides containing 900 TIP3P water molecules [33,34]. After removing water residues within 2.6 Å of the solute, 842, 854, and 867 water residues remained in the cL7, LNF-1 and the trisaccharide boxes, respectively. Subsequently, energy minimization was performed consisting of steepest descent, 100 steps, and adopted-basis Newton-Raphson, 5000 steps or until the RMS gradient was less than 0.01 kcal·mol^{-1}·Å^{-1}. Velocities were assigned at 100 K and reassigned in 5 K steps to 310 K during 40 ps. Production runs were performed for 150 ns with a time step of 2 fs and constant temperature and pressure (1 atm). The first ns was considered as the equilibration period, and hence it was not included in the analysis. Non-bonded forces were switched to zero between 10 and 12 Å and electrostatics were handled using the particle mesh Ewald approach. Bonds involving hydrogen atoms were restrained using the SHAKE procedure and overall rotation and translation was removed every 1000 steps. Coordinates were saved every 2 ps for analysis.

For the simulations having the *N*-acetyl torsion angle restrained, a harmonic restraint of 10 kcal·mol^{-1}·deg^{-1} was applied to the dihedral defined by the C1-C2-N-C atoms.

Starting from the energy minimized structure of cL7 prior to solvation for the MD simulation, 150 ns of LD were simulated using a time step of 1 fs. The collision frequency was 50 ps^{-1} for non-hydrogen atoms, and the heat bath temperature was 310 K.

3D-CFG representation

Octave, version 3 (http://www.gnu.org) [35] was used to read the PDB file for the X-ray crystallographic structure and to create a compiled graphics objects (CGO) file describing the 3D-CFG representation the glycan. Subsequently, these files were used together with the original PDB file to render images using PyMol (The PyMOL Molecular Graphics System, Version 1.5.0.1 Schrödinger, LLC).

ACKNOWLEDGMENTS

This work was supported by grants from the Swedish Research Council and The Knut and Alice Wallenberg Foundation. Computing resources were kindly provided by the Center for Parallel Computers (PDC), Stockholm, Sweden.

REFERENCES

1. H.-J. Gabius, S. André, J. Jiménez-Barbero, A. Romero, D. Solís. *Trends Biochem. Sci.* **36**, 298 (2011).
2. G. A. Rabinovich, Y. van Kooyk, B. A. Cobb. *Ann. N.Y. Acad. Sci.* **1253**, 1 (2012).
3. S. Sato, M. Quellet, C. St-Pierre, M. J. Trembley. *Ann. N.Y. Acad. Sci.* **1253**, 133 (2012).
4. H.-C. Siebert, S. André, S.-Y. Lu, M. Frank, H. Kaltner, J. A. van Kuik, E. Y. Korchagina, N. Bovin, E. Tajkhorshid, R. Kaptein, J. F. G. Vliegenthart, C.-W. von der Lieth, J. Jiménez-Barbero, J. Kopitz, H.-J. Gabius. *Biochemistry* **42**, 14762 (2003).
5. M. Nagae, Y. Yamaguchi. *Int. J. Mol. Sci.* **13**, 8398 (2012).
6. M. Kuttel, Y. Mao, G. Widmalm, M. Lundborg. *Proc. 7th IEEE Int. Conf. eScience*, 395 (2011). http://dx.doi.org/10.1109/eScience.2011.61.
7. B. O. Petersen, M. Krah, J. Ø. Duus, K. K. Thomsen. *Eur. J. Biochem.* **267**, 361 (2000).
8. M. Faijes, A. Planas. *Carbohydr. Res.* **342**, 1581 (2007).
9. J.-L. Viladot, V. Moreau, A. Planas, H. Driguez. *J. Chem. Soc., Perkin Trans. 1* 2383 (1997).
10. B. Cobucci-Ponzano, M. Moracci. *Nat. Prod. Rep.* **29**, 697 (2012).
11. M. Yokoyama. *Carbohydr. Res.* **327**, 5 (2000).
12. L. F. Mackenzie, Q. Wang, R. A. J. Warren, S. G. Withers. *J. Am. Chem. Soc.* **120**, 5583 (1998).
13. C. Malet, A. Planas. *FEBS Lett.* **440**, 208 (1998).
14. V. Boyer, S. Fort, T. P. Frandsen, M. Schülein, S. Cottaz, H. Driguez. *Chem.—Eur. J.* **8**, 1389 (2002).
15. R. Fauré, M. Saura-Valls, H. Brumer, A. Planas, S. Cottaz, H. Driguez. *J. Org. Chem.* **71**, 5151 (2006).
16. J. Vasur, R. Kawai, E. Andersson, K. Igarashi, M. Sundgren, M. Samejima, J. Ståhlberg. *FEBS J.* **276**, 3858 (2009).
17. J. Vasur, R. Kawai, K. H. M. Jonsson, G. Widmalm, Å. Engström, M. Frank, E. Andersson, H. Hansson, Z. Forsberg, K. Igarashi, M. Samejima, M. Sandgren, J. Ståhlberg. *J. Am. Chem. Soc.* **132**, 1724 (2010).
18. J. Ståhlberg, J. Vasur, R. Kawai, K. H. M. Jonsson, G. Widmalm. Unpublished results.
19. O. Guvench, E. Hatcher, R. M. Venable, R. W. Pastor, A. D. MacKerell Jr. *J. Chem. Theory Comput.* **5**, 2353 (2009).
20. G. Widmalm, R. W. Pastor. *J. Chem. Soc., Faraday Trans.* **88**, 1747 (1992).
21. E. Säwén, T. Massad, C. Landersjö, P. Damberg, G. Widmalm. *Org. Biomol. Chem.* **8**, 3684 (2010).
22. B. J. Hardy, W. Egan, G. Widmalm. *Int. J. Biol. Macromol.* **17**, 149 (1995).
23. U. Olsson, A. S. Serianni, R. Stenutz. *J. Phys. Chem. B* **112**, 4447 (2008).
24. M. Zaccheus, R. Pendrill, T. A. Jackson, A. Wang, F.-I. Auzanneau, G. Widmalm. *Eur. J. Org. Chem.* 4705 (2012).
25. E. Säwén, B. Stevensson, J. Östervall, A. Maliniak, G. Widmalm. *J. Phys. Chem. B* **115**, 7109 (2011).
26. P. Cagas, K. Kaluarachchi, C. A. Bush. *J. Am. Chem. Soc.* **113**, 6815 (1991).
27. S. Kornfeld, E. Li, I. Tabas. *J. Biol. Chem.* **253**, 7771 (1978).
28. A. Varki, R. D. Cummings, J. D. Esko, H. H. Freeze, P. Stanley, J. D. Marth, C. R. Bertozzi, G. W. Hart, M. E. Etzler. *Proteomics* **9**, 5398 (2009).

29. N. Noinaj, N. C. Easley, M. Oke, N. Mizuno, J. Gumbart, E. Boura, A. S. Steere, O. Zak, P. Aisen, E. Tajkhorshid, R. W. Evans, A. R. Gorringe, A. B. Mason, A. C. Steven, S. K. Buchanan. *Nature* **483**, 53 (2012).
30. B. R. Brooks, C. L. Brooks III, A. D. MacKerell Jr., L. Nilsson, R. J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, A. Caflisch, L. Caves, Q. Cui, A. R. Dinner, M. Feig, S. Fischer, J. Gao, M. Hodoscek, W. Im, K. Kuczera, T. Lazaridis, J. Ma, V. Ovchinnikov, E. Paci, R. W. Pastor, C. B. Post, J. Z. Pu, M. Schaefer, B. Tidor, R. M. Venable, H. L. Woodcock, X. Wu, W. Yang, D. M. York, M. Karplus. *J. Comput. Chem.* **30**, 1545 (2009).
31. O. Guvench, S. S. Mallajosyula, E. P. Raman, E. Hatcher, K. Vanommeslaeghe, T. J. Foster, F. W. Jamison II, A. D. Mackerell Jr. *J. Chem. Theory Comput.* **7**, 3162 (2011).
32. R. Eklund, G. Widmalm. *Carbohydr. Res.* **338**, 393 (2003).
33. W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein. *J. Chem. Phys.* **79**, 926 (1983).
34. S. R. Durell, B. R. Brooks, A. Ben-Naim. *J. Phys. Chem.* **98**, 2198 (1994).
35. J. W. Eaton, D. Bateman, S. Hauberg. *GNU Octave Manual Version 3*, Network Theory Ltd.

