ARTICLE

ACS chemical biology

Exploring the Structural Diversity of Mammalian Carbohydrates ("Glycospace") by Statistical Databank Analysis

Daniel B. Werz[†,§], René Ranzinger[‡], Stephan Herget[‡], Alexander Adibekian[†], Claus-Wilhelm von der Lieth[‡], and Peter H. Seeberger[†,*]

[†]Laboratory for Organic Chemistry, Swiss Federal Institute of Technology (ETH) Zürich, HCI F 315, Wolfgang-Pauli-Str. 10, CH-8093 Zürich, Switzerland, and [‡]German Cancer Research Center (DKFZ), Im Neuenheimer Feld 280, D-69120 Heidelberg, Germany. [§]Present address: Institute for Organic and Biomolecular Chemistry, Georg-August-Universität Göttingen, Tammannstr. 2, D-37077 Göttingen, Germany

ABSTRACT The diversity of three major classes of mammalian carbohydrates, mainly glycolipids and *O*- and *N*-linked glycans, deposited in the databank GLYCOSCIENCES.de was subjected to statistical analyses. Size, chain length, and branching complexity were accessed and revealed that the average oligosaccharide is composed of about eight monosaccharide units. About a quarter of all oligosaccharides are strictly linear, and the remainder are branched at least once. Glucosamine, galactose, and mannose are dominating and comprise 75% of the monosaccharides within mammalian oligosaccharide frameworks. α-Linked sialic acid, α-linked fucose, and β-linked galactose decorate the majority of reducing termini. Glucose as the most abundant carbohydrate in mammals plays only a very minor role within these structures. Particular emphasis was placed on analyzing the way the monosaccharide units are linked within the oligomeric framework. Just 11 monosaccharide connections account for >75% of all linkages. Thus, the number of structural combinations found in nature, the part of the occupied mammalian glycospace, is much smaller than expected. As a result, a potential set of building blocks for oligosaccharide assembly is presented. This potential building block set was correlated with the accessible 3299 mammalian carbohydrate structures in the GLYCOSCIENCES.de databank. Only 36 building blocks are required to construct 75% of the 3299 mammalian oligosaccharides.

Three major repeating biomacromolecules, polynucleotides, polypeptides, and carbohydrates, are responsible for much of the information transfer in biological systems. Encoding and transmission of information relies on the construction of diverse macromolecules that contain the message. Polynucleotides serve as the blueprint of life in the form of DNA, and polypeptides carry out most of reactions in living cells. Both polymers are strictly linear and derived biosynthetically *via* reliable templated syntheses. DNA is composed of 4 nucleotides, and mammalian proteins have 20 proteinogenic amino acids that determine polymer diversity: 4096 (4^6) hexanucleotides and 64 million (20^6) hexapeptides are possible. Posttranslational modifications such as phosphorylation, glycosylation, and lipidation further increase protein complexity.

The term "carbohydrates" describes a host of different biooligomers composed of monosaccharides. Oligosaccharides are almost always part of glycoconjugates, that is, the combination of a sugar chain with a protein (glycoprotein), a lipid (glycolipid) or both lipid and protein (glycosylphosphatidylinositol (GPI)-anchored proteins) (*1*). Carbohydrate chains can be branched, because each monosaccharide provides different positions around the ring that can be connected. In contrast to amide or phosphate diester linkages, the formation of each glycosidic linkage creates one new stereogenic center. Carbohydrate complexity is increased by the stereocenters that constitute the ring in addition to ring size, linkage position, and branching, as well as further attachments such as sulfation, methylation, and phosphorylation.

*Corresponding author, seeberger@org.chem.ethz.ch.

Received for review August 22, 2007 and accepted September 14, 2007.

Published online October 19, 2007

10.1021/cb700178s CCC: $37.00

TABLE 1. Diversity space of oligonucleotides, peptides, and mammalian oligosaccharides[a]

	Numbers of different oligomers		
Oligomer size	Nucleotides	Peptides	Carbohydrates
1	4	20	20
2	16	400	1360
3	64	8,000	126,080
4	256	160,000	13,495,040
5	1024	3,200,000	1,569,745,920
6	4096	64,000,000	192,780,943,360

[a]The numbers for the mammalian oligosaccharides are based on the 10 mammalian monosaccharides: D-Glc [4], D-Gal [4], D-Man [4], D-Sia [4], D-GlcNAc [3], D-GalNAc [3], L-Fuc [3], D-Xyl [3], D-GlcA [3], and L-IdoA [3]. The number of substitutable OH groups (excluding the anomeric one) is given in square brackets. Commonly, only the pyranose ring forms and not the furanose ring forms of the above-mentioned monosaccharides are found in mammals (*35*, *36*).

Not surprisingly, carbohydrate complexity dwarfs that of both DNA and proteins but to date has been assessed on a purely theoretical level (*2*, *3*). We performed calculations regarding the diversity of mammalian carbohydrate structures, based only on the "10 mammalian monosaccharides" [D-glucose (D-Glc), D-galactose (D-Gal), D-mannose (D-Man), D-sialic acid (D-Sia), *N*-acetyl-D-glucosamine (D-GlcNAc), *N*-acetyl-D-galactosamine (D-GalNAc), L-fucose (L-Fuc), D-xylose (D-Xyl), D-glucuronic acid (D-GlcA), and L-iduronic acid (L-IdoA)] and not considering any further attachments. These results are summarized in Table 1. The number of structural combinations encountered in nature, the part of the glycospace that is actually occupied, has not yet been elucidated. A systematic structural analysis of mammalian oligosaccharide structures deposited in glycan databases will aid our understanding of carbohydrate diversity and help to identify a putative set of monosaccharide building blocks for efficient carbohydrate assembly.

Access to pure, structurally defined carbohydrates remains difficult at a time when the automated synthesis of oligonucleotides (*4*, *5*) and oligopeptides (*6*) is common. Although biologically relevant oligosaccharides can be assembled from monosaccharides in a linear fashion on an automated synthesizer (*7*–*10*), no general method for nonspecialists to draw from a set of commercially available building blocks exists yet. The structural complexity of carbohydrates may complicate a comprehensive synthesis approach in the case where too many building blocks are needed.

A better understanding of the structures actually found in nature can guide the selection of building blocks needed for assembly. A database of reliable structures is required to analyze carbohydrate diversity. Relatively limited data sets exist, as carbohydrate isolation and structure elucidation are formidable challenges. The systematic collection of carbohydrates in databases is lagging far behind genomics and proteomics. Currently, no database provides access to all published glycan structures, although several commercial and publicly funded initiatives are working to make glycan structures available in a well-structured and annotated digital representation (*11*). These databases contain mainly information about *O*- and *N*-glycans and glycolipids because their isolation and sequencing is more tractable than that of glycosaminoglycans, for example. Reported here is a detailed statistical analysis of the GLYCOSCIENCES.de databank (*12*, *13*) to elucidate the mammalian glycospace with a focus on the oligosaccharide portions of *O*- and *N*-glycans and glycolipids.

RESULTS AND DISCUSSION

The complexity analysis was based on 3299 oligosaccharides from 38 mammalian species (*14*). Noncarbohydrate portions such as amino or fatty acids at the reducing terminus were not considered. Most oligosaccharides (2128 of 3299, 64%) are of human origin, and the rest are derived from cow, rat, pig, mouse, and

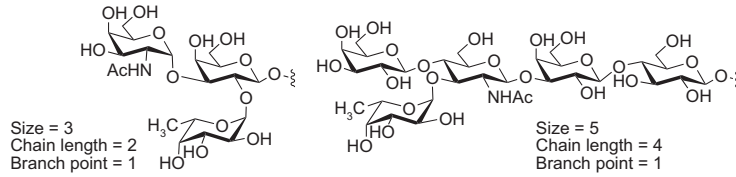

Figure 1. Illustration of size, chain length, and branch points of oligosaccharides.

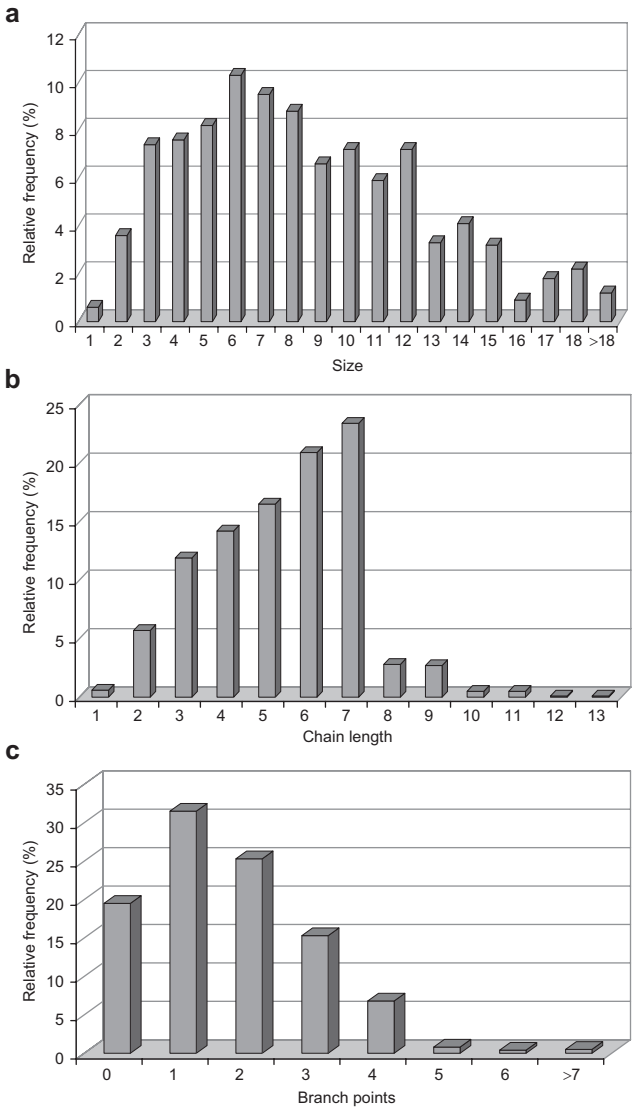

Figure 2. General properties of mammalian carbohydrates. a) Size of mammalian oligosaccharides (for definition see Figure 1). The relative frequency is given in percent. The largest oligosaccharide contains 37 monosaccharides. b) Chain length of mammalian *O*- and *N*-glycans. The longest chain contains 13 monosaccharide moieties. c) Branching complexity of mammalian *O*- and *N*-glycans.

other species. Because the statistical analyses did not reveal any relevant difference between the human and the mammalian set of oligosaccharide structures, we focused our analysis on the mammalian sugars.

Initially, we addressed some basic questions: What is the most common size of an oligosaccharide? What is the typical chain length within a branched oligosaccharide? What portion of oligosaccharides is linear, and what portion is branched? The size of an oligosaccharide is described as the number of monosaccharide units that make up the oligomer. The chain length is the longest path of monosaccharide units from the reducing end to the nonreducing terminus of the chain. The number of terminal residues was calculated as well. It differs by one from the number of branch points. The three definitions are illustrated in Figure 1.

The average oligosaccharide is composed of about 8 monosaccharide units, whereas the sugars in the database vary in size from 1 (*e.g.*, T_N antigen) to 37 monosaccharide units (Figure 2, panel a). Most mammalian structures (95%) have a shorter chain length than eight residues (Figure 2, panel b). The longest mammalian carbohydrate structure in the database has a maximum chain length of 13. About 20% of the oligosaccharide structures in the database are linear. More than 50% of all oligosaccharides are branched once or twice, whereas 22% of the structures are branched three or four times. Few carbohydrates are branched five or more times, with a maximum of nine branch points (Figure 2, panel c).

D-GlcNAc (32%), D-Gal (25%), and D-Man (19%) comprise >75% of all monosaccharide units found in mammalian oligosaccharides (Table 2). Sia and L-Fuc are found less frequently (8% each). The most abundant monosaccharide in nature, D-Glc, which makes up cellulose, starch, and glycogen, astonishingly plays only a minor role within mammalian *O*- and *N*-glycans. Three different monosaccharides dominate the nonreducing terminus, the site most often recognized by carbohydrate-binding proteins: D-Gal, Sia, and L-Fuc each cap 25% of oligosaccharides. D-Man and D-GlcNAc each terminate 8% of mammalian oligosaccharides (Table 2).

In planning a comprehensive, general, and linear synthetic approach, a set of building blocks containing the

TABLE 2. Abundance of monosaccharides and terminal monosaccharide moieties in mammalian carbohydrates

Monosaccharide	Abundance (%)
Abundance of monosaccharides	
D-GlcNAc	31.8
D-Gal	24.8
D-Man	18.9
D-Sia[a]	8.3
L-Fuc	7.2
D-GalNAc	4.8
D-Glc	2.5
D-GlcA	0.3
D-Xyl	0.1
L-IdoA	0.1
Others	1.2
Abundance of terminal monosaccharide moieties	
α-D-Sia[a]	26.1
α-L-Fuc	23.8
β-D-Gal	23.0
α-D-Man	8.2
β-D-GlcNAc	8.0
α-D-Gal	2.3
α-D-GalNAc	2.3
β-D-GalNAc	2.2
α-D-Glc	0.8
Other	3.3

[a] *N*-Acetyl and *N*-glycolyl.

proper protective groups to install all possible connectivities and stereogenic centers is mandatory. Different protective groups mark hydroxyl groups that serve as nucleophiles during chain extension and those that remain latent. The protective groups used should also control the stereochemical outcome of glycosylation reactions. To further complicate matters, the sterics, electronics, and conformation of the monosaccharides are fundamentally influenced by the choice of protective groups (*15–17*). One aim of this study is to derive a minimal set of putative monosaccharide building blocks required to assemble the majority of mammalian oligosaccharides in a strictly linear fashion. Procurement of monosaccharide building blocks is a formidable challenge (*18*), and a defined number of reliable standard components for oligomer construction would help in synthetic planning and for practical reasons. These building blocks would be utilized in the buildup of linear and branched molecules by solution- and solid-phase methods.

Each of the 10 mammalian monosaccharide units can in principle be connected to its neighbors in a variety of different ways, including different anomeric configurations and one or more branchings. To construct all theoretically possible mammalian oligosaccharides by linear chemical synthesis, 224 different building blocks would be required (for further details, see Supporting Information). Because of this large number, special care was taken to elucidate the stereochemistry at the anomeric position (α or β) and the position of linkages within mammalian carbohydrates (Table 3). The results illustrated in Figure 3 are stunning: 80% of the monosaccharide linkages within the oligomers can be constructed using only 13 building blocks. The most frequently occurring connections are 4-linked β-GlcNAc, capping α-Sia, capping α-Fuc, capping β-Gal, 2-linked α-Man, and 3-linked β-Gal.

On the basis of this analysis, key building blocks needed for the construction of mammalian oligosaccharides can be designed. In Figure 4, we suggest putative building blocks **1–20** to obtain the most abundant linkages. For this purpose, benzyl groups (Bn) were selected for the permanent protection of hydroxyls. Where participating groups are needed to ensure anomeric specificity, pivaloyl (Piv) (*19*), acetyl (Ac), or benzoyl (Bz) groups are placed for permanent protection. 9-Fluorenylmethyl carbonate (Fmoc) was selected as a temporary protecting group and also serves as a participating group in the C2 position for temporary protection (*20*). To install branched carbohydrate structures, two or more temporary protecting groups that can be

TABLE 3. The 20 most abundant monosaccharide units with their linkage modes found in mammalian oligosaccharides

Monosaccharide unit	Abundance (%)
(4→1)β-D-GlcNAc	21.7
α-D-Sia[a]	8.8
α-L-Fuc	8.0
β-D-Gal	7.8
(2→1)α-D-Man	7.1
(3→1)β-D-Gal	5.2
(3→1)(6→1)β-D-Man	4.7
(3→2)β-D-Gal	4.3
α-D-Man	2.8
β-D-GlcNAc	2.7
(6→2)β-D-Gal	2.4
(3→1)(4→1)β-D-GlcNAc	2.3
(2→1)(4→1)α-D-Man	2.1
(4→1)β-D-Glc	1.6
(2→1)β-D-Gal	1.5
(2→1)(6→1)α-D-Man	1.5
(3→1)β-D-GlcNAc	1.2
(2→1)(3→1)β-D-Gal	1.0
(3→1)(6→1)β-D-Gal	0.9
α-D-Gal	0.8

[a] *N*-Acetyl and *N*-glycolyl.

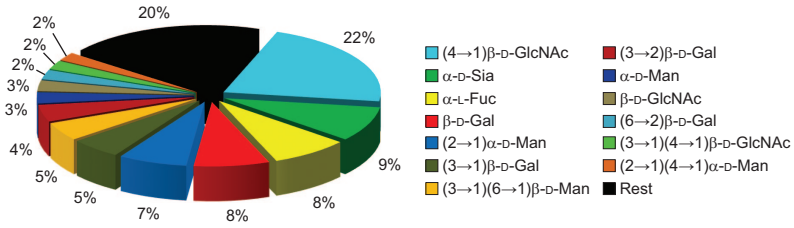

Figure 3. The 13 most abundant monosaccharide units (with linkage mode and position) found in mammalian oligosaccharides.

cleaved chemoselectively are necessary. Levulinoyl ester (Lev) and *p*-methoxybenzyl (PMB) were selected as other temporary protecting groups (*21*). Similar protecting group schemes can be selected alternatively. However, the large majority of the building blocks presented in Figure 4 have been tested successfully for their utility in solution- and solid-phase oligosaccharide syntheses (*22–30*). The Sia and β-Man building blocks **2** and **7** represent a special challenge concerning selective glycosidic bond construction. So far, α-Sia and β-Man bonds have been constructed using disaccharide modules in place of monosaccharide units (*31*). These challenges are currently being addressed.

With a set of putative building blocks in hand (Figure 4), the construction of mammalian carbohydrates contained in the GLYCOSCIENCES.de databank was simulated. The results of these calculations are impressive: 60% of the 3299 mammalian oligosaccharides are accessible with only 25 building blocks (see Figure 5, panel a)! Just 11 more building blocks are needed to construct 75% of oligosaccharides. To produce 90% of all structures, a set of 65 building blocks is required. The number of building blocks to access the last 10% of mammalian oligosaccharides increases tremendously (Figure 5, panel a). The occurrence of rare monosaccharide units not commonly found in mammals, such as D-Fuc, L-arabinose, L-rhamnose, and D-galacturonic acid, as

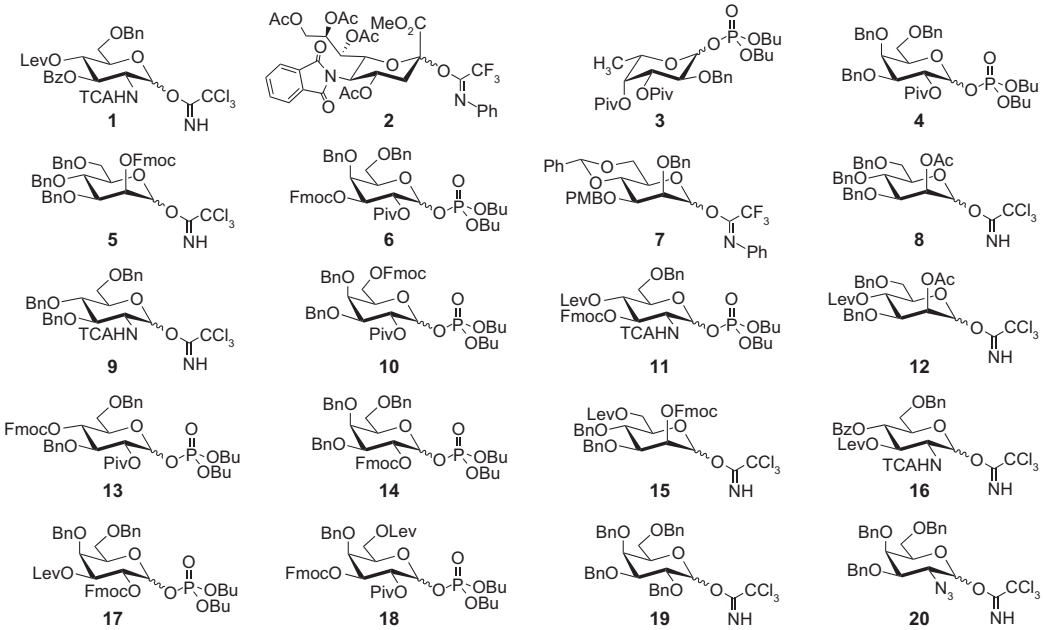

Figure 4. Putative monosaccharide building blocks 1–20 sorted by their relative abundance in mammalian oligosaccharides. Fmoc, Lev, and PMB serve as temporary protecting groups, whereas Bn, Ac, Piv, Bz, and TCA serve as permanent protecting groups. The numbering differs slightly from that of Table 3 because (3→1)β-D-Gal and (3→2)β-D-Gal units require the same building block 6.

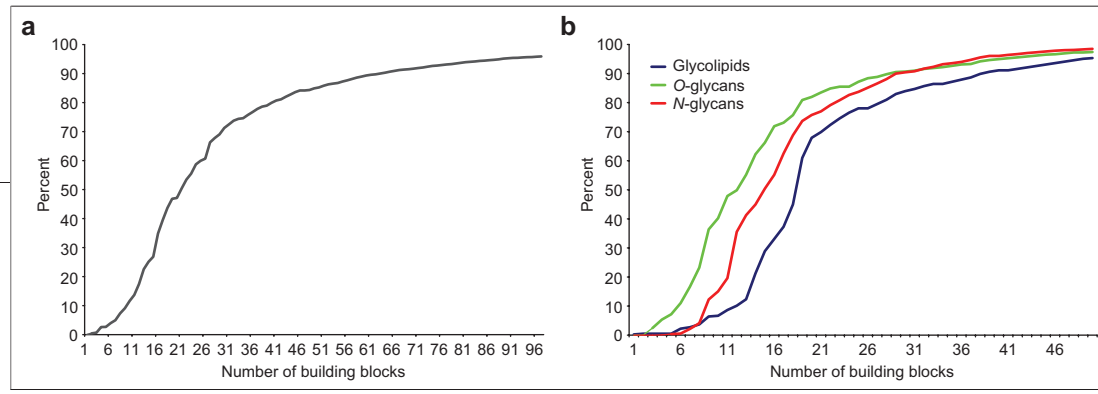

Figure 5. Number of building blocks required for synthetic access to mammalian carbohydrates. a) Percentage of accessible mammalian carbohydrates correlated to the number of building blocks. b) Percentage of accessible mammalian carbohydrates split into different classes (glycolipids and *N*- and *O*-linked glycans) and correlated to the number of building blocks.

well as unusual linkages of L-Fuc and Sia, are likely the result of erroneous assigned databank entries. Microorganisms that live in mammals express a much broader variety of carbohydrate moieties, and linkages and may be the source of the additional sugars.

Evaluating the accessibility of the different classes reveals that for each class even fewer building blocks are required to reach a certain number of structures. Glycolipids in comparison show a greater variety of different linkages than *N*- and *O*-linked glycans (Figure 5, panel b).

A rather small number of building blocks is sufficient to access the majority of the mammalian glycospace. In many cases, the reducing terminal units can be introduced by building blocks containing a temporary protecting group to further reduce this number. However, in the case of branched structures, such an approach may be problematic. Therefore, our synthetic strategy is based on general principles with special capping building blocks.

Conclusion. Herein, we report the first attempt to describe mammalian oligosaccharide diversity ("glycospace") based on a glycan databank. Carbohydrate sizes, chain lengths, and branching complexity have been examined. Analysis of monosaccharide connectivities within the oligomeric structures guided us to identify a set of putative monosaccharide building blocks suitable for the linear solution- and solid-phase assembly of mammalian oligosaccharides. This potential building block set was correlated with the accessible 3299 mammalian carbohydrate structures in the GLYCOSCIENCES.de databank. Only 36 building blocks are needed to construct 75% of the 3299 mammalian oligosaccharides.

METHODS

Preparation of Data Sets. The GLYCOSCIENCES.de (version 02/2006) data set including taxonomical annotation data from CarbBank (*32, 33*) was used to generate a subset of all mammalian oligosaccharides, resulting in a total number of 5339 distinct structures. With the help of a recently developed dictionary database that contains a controlled vocabulary for monosaccharides in a novel machine-readable form, noncarbohydrate moieties were removed, thus resulting in 3299 distinct oligosaccharide structures. This set was stored in a MySQL database in the LINUCS notation for oligosaccarides (*34*) and used for subsequent analyses.

Analysis. An analysis tool (JAVA) was developed that implements a parsing section for the LINUCS notation, analytical functions, and an output generator. The output generator stores the results in temporary database tables that are subsequently used to prepare the final result tables.

Analytical Functions. The analysis tool implements basic functions to extract monosaccharide basetypes and substituents with the help of the dictionary database from a given set of oligosaccharides. It computes size and topology-related parameters such as size distribution, chain length, and number of terminal residues. The topology can be examined with different options regarding naming conventions (CarbBank, normalized IUPAC) and definitions of the detected building block fragments (including or excluding connected residues, substituents, or positional information). This procedure allows for different views of oligosaccharide residue connectivities. In addition, an option was implemented to calculate the percentage of structures that can be synthesized on the basis of a given list of building blocks.

Acknowledgment: We thank the ETH Zürich and the Swiss National Science Foundation (Grant 200021-101593/1) for generous support. D.B.W. acknowledges the Alexander von Humboldt Foundation (AvH) for a Feodor Lynen Research Fellowship and the Deutsche Forschungsgemeinschaft (DFG) for an Emmy Noether Fellowship. A.A. is a Kekulé scholar of the Fonds der Chemischen Industrie.

Supporting Information Available: This material is available free of charge *via* the Internet.

REFERENCES

1. Seeberger, P. H, and Werz, D. B. (2005) Automated synthesis of oligosaccharides as a basis for drug discovery, *Nat. Rev Drug Discovery 4*, 751–763.
2. Schmidt, R. R. (1986) New methods of glycoside and oligosaccharide syntheses–are there alternatives to the Koenigs–Knorr method? *Angew. Chem., Int. Ed. Engl. 25*, 212–235.
3. Laine, R. A. (1994) A calculation of all possible oligosaccharide isomers both branched and linear yields 1.05×10^{12} structures for a reducing hexasaccharide: the isomer barrier to development of single-method saccharide sequencing or synthesis systems, *Glycobiology 4*, 759–767.
4. Caruthers, M. H. (1985) Gene synthesis machines: DNA chemistry and its uses, *Science 230*, 281–285.

5. Caruthers, M. H. (1991) Chemical synthesis of DNA and DNA analogs, *Acc. Chem. Res. 24*, 278–284.
6. Atherton, E., and Sheppard, R. C. (1989) *Solid-Phase Peptide Synthesis: A Practical Approach*, Oxford Univ. Press, Oxford.
7. Sears, P., and Wong, C.-H. (2001) Toward automated synthesis of oligosaccharides and glycoproteins, *Science 291*, 2344–2350.
8. Plante, O. J., Palmacci, E. R., and Seeberger, P. H. (2001) Automated solid-phase synthesis of oligosaccharides, *Science 291*, 1523–1527.
9. Love, K. R., and Seeberger, P. H. (2004) Automated solid-phase synthesis of protected tumor-associated antigen and blood group determinant oligosaccharides, *Angew. Chem., Int. Ed. 43*, 602–605.
10. Werz, D. B., Castagner, B., and Seeberger, P. H. (2007) Automated synthesis of tumor-associated carbohydrate antigens Gb-3 and Globo-H: incorporation of α-galactosidic linkages, *J. Am. Chem. Soc. 129*, 2770–2771.
11. von der Lieth, C.-W. (2004) An endorsement to create open access databases for analytical data of complex carbohydrates, *J. Carbohydr. Chem. 23*, 277–297.
12. Lutteke, T., Bohne-Lang, A., Loss, A., Goetz, T., Frank, M., and von der Lieth, C.-W. (2006) GLYCOSCIENCES.de: an Internet portal to support glycomics and glycobiology research, *Glycobiology 16*, 71R–81R.
13. Less complex oligosaccharides may dominate as a result of experimental difficulties in sequencing larger structures. This problem is common to all the glycan databanks.
14. About 47% of the 3299 structures are derived from *N*-linked glycoproteins, 19% of them from *O*-linked glycoproteins and 17% from glycolipids. The rest are not assigned.
15. Mootoo, D. R., Konradsson, P., Udodong, U., and Fraser-Reid, B. (1988) "Armed" and "disarmed" *n*-pentenyl glycosides in saccharide couplings leading to oligosaccharides, *J. Am. Chem. Soc. 110*, 5583–5584.
16. Ye, X. S., and Wong, C.-H. (2000) Anomeric reactivity-based one-pot oligosaccharide synthesis: a rapid route to oligosaccharide libraries, *J. Org. Chem. 65*, 2410–2431.
17. Orgueira, H. A., Bartolozzi, A., Schell, P., and Seeberger, P. H. (2002) Conformational locking of the glycosyl acceptor for stereocontrol in the key step in the synthesis of heparin, *Angew. Chem., Int. Ed. 41*, 2128–2131.
18. Boons, G. J., Ed. (1998) *Carbohydrate Chemistry*, Blackie, London, U.K.
19. Kunz, H., and Harreus, A. (1982) Glycosidsynthese mit 2,3,4,6-tetra-*O*-pivaloyl-α-D-glucopyranosylbromid, *Liebigs Ann.* 41–48.
20. Roussel, F., Knerr, L., Grathwohl, M., and Schmidt, R. R. (2000) *O*-Glycosyl trichloroacetimidates bearing Fmoc as temporary hydroxy protecting group: a new access to solid-phase oligosaccharide synthesis, *Org. Lett. 2*, 3043–3046.
21. Koeners, H. J., Verhoeven, J., and van Boom, J. H. (1980) Synthesis of oligosaccharides by using levulinic ester as an hydroxyl protecting group, *Tetrahedron Lett. 21*, 381–382.
22. Building blocks **1** and **16**:Love, K. R., and Seeberger, P. H. (2005) Solution syntheses of protected type II Lewis blood group oligosaccharides: study for automated synthesis, *J. Org. Chem. 70*, 3168–3177.
23. Building block **2**:Tanaka, K., Goi, T., and Fukase, K. (2005) Highly efficient sialylation towards alpha(2–3)- and α(2–6)-Neu5Ac-Gal synthesis: significant 'fixed dipole effect' of *N*-phthalyl group on alpha-selectivity, *Synlett 19*, 2958–2962.
24. Building blocks **3**, **6**, **11**, **13**, and **14**:Love, K. R., and Seeberger, P. H. (2004) Automated solid-phase synthesis of protected tumor-associated antigen and blood group determinant oligosaccharides, *Angew. Chem., Int. Ed. 43*, 602–605.
25. Building blocks **4**, **8**, and **12**:Hewitt, M. C., and Seeberger, P. H. (2001) Automated solid-phase synthesis of a branched Leishmania cap tetrasaccharide, *Org. Lett. 3*, 3699–3702.
26. Building block **5**:Wu, X., Grathwohl, M., and Schmidt, R. R. (2002) Efficient solid-phase synthesis of a complex, branched *N*-glycan hexasaccharide: Use of a novel linker and temporary-protecting-group pattern, *Angew. Chem., Int. Ed. 41*, 4489–4493.
27. Building blocks **7**, **10**, **15**, **17**, and **18**: Kröck, L., Oberli, M., Werz, D. B., Seeberger, P. H. Unpublished results.
28. Building blocks **9**:Blatter, G., Beau, J. M., and Jacquinet, J. C. (1994) The use of 2-deoxy-2-trichloroacetamido-D-glucopyranose derivatives in syntheses of oligosaccharides, *Carbohydr. Res. 260*, 189–202.
29. Building block **19**:Wang, C. C., Lee, J. C., Luo, S. Y., Fan, H. F., Pai, C. L., Yang, W. C., Lu, L. D., and Hung, S. C. (2002) Synthesis of biologically potent α1→2-linked disaccharide derivatives *via* regioselective one-pot protection-glycosylation, *Angew. Chem., Int. Ed. 41*, 2360–2362.
30. Building block **20**:Belén Cid, M., Bonilla, J. B., Alfonso, F., and Martín-Lomas, M. (2003) Synthesis of new hexosaminyl D- and L-chiro-inositols related to putative insulin mediators, *Eur. J. Org. Chem.* 3505–3515.
31. Ratner, D. M., Swanson, E. R., and Seeberger, P. H. (2003) Automated synthesis of a protected *N*-linked glycoprotein core pentasaccharide, *Org. Lett. 5*, 4717–4720.
32. Doubet, S., and Albersheim, P. (1992) Carbbank, *Glycobiology 2*, 505.
33. Doubet, S., Bock, K., Smith, D., Darvill, A., and Albersheim, P. (1989) The complex carbohydrate structure database, *Trends Biochem. Sci. 14*, 475–477.
34. Bohne-Lang, A., Lang, E., Forster, T., and von der Lieth, C.-W. (2001) LINUCS: linear notation for unique description of carbohydrate sequences, *Carbohydr. Res. 336*, 1–11.
35. de Lederkremer, R. M., and Colli, W. (1995) Galactofuranose-containing glycoconjugates in trypanosomatids, *Glycobiology 5*, 547–552.
36. Marlow, A. L., and Kiessling, L. L. (2001) Improved chemical synthesis of UDP-galactofuranose, *Org. Lett. 3*, 2517–2519.