Sugar is not Always Sweet: Naming

Naming oligosaccharides is pretty difficult at a glance. The rules for describing compound names in journals evolve over time, but we should know the basic rules of carbohydrate nomenclature. Here is an example of a compound, which is a complex synthetic intermediate, for naming. (Fig. 1) The name of the compound is phenyl (methyl 5-acetamido-4,7,8,9-tetra-O-acetyltetra-O-deoxy-D-glycero-a-D-galacto-2-nonulosyluronate)-(2→6)-2,3,4-tri-O-benzoyl-1-thio-β-D-galactopyranoside. We now start explaining the naming of the compound.

The compound consists of two monosaccharide units as indicated as “a” and “b”, which are called reducing end and non-reducing end units respectively. The main body of the name is, “galactopyranoside” (unit a) and the other parts are modifiers meaning that the compound is galactoside with several modifications. The compound is a “phenylthio” glycoside of “D”-“galactose” with the “β”-configuration having a “pyran” ring structure, of which the 2-, 3-, and 4-positions are protected as benzoyl esters.

One of the important factors in naming carbohydrates is defining the absolute configuration. The “D” or “L” series is assigned according to the configuration of the highest-numbered center of chirality, which is called “configurational atom” (Fig. 2). This atom is also used to define anomeric configurations of carbohydrates having up to four chiral centers. The stereochemistry of the glycosidic linkage is determined based on a definition; “In the α anomer, the exocyclic oxygen atom at the anomeric center is formally cis, in the Fischer projection, to the oxygen attached to the anomeric reference atom; in the β anomer these oxygen atoms are formally trans.” In this case, the configuration is in the β-configuration (O-5 and O-1 are trans).

Fig. 1. Structure of a compound.

Fig. 2. Stereochemistry of galactose.

Fig. 3. Stereochemistry of sialic acid.
Another important modification is glycosylation at the C-6 position, which is explained below. The main modifier is, “(methyl 5-acetamido-4,7,8,9-tetra-O-acetyl-3,5-dideoxy-α-D-glycero-α-D-galacto-2-nonulo-pyranosylonate)-(2→6)-.” This indicates that the attached sugar (unit \( b \)) is a nine-carbon ketose, a “nonulose” in a pyran ring structure, which is linked to the 6-position of the adjacent sugar unit \( a \) through a glycosidic linkage. It is described as “non-ulos-yl.” The sugar also has a methyl ester of the carboxylic acid, which is indicated as “methyl ...onate.” The description of the stereochemistry in monosaccharides consisting of more than four chiral carbon atoms is described by a combination of stereochemical names derived from monosaccharides consisting of three to six carbons. This particular nona-ose (a sugar composed of nine carbon atoms), sialic acid, has five chiral centers in addition to the anomeric center. The stereochemistry involved here is the “D-glycero” configuration for C-8, which is connected to the C-4–C-7 unit with four continuous chiral centers in the “D-galacto” configuration. Thus, the stereochemistry of the sugar is defined as “D-glycero-D-galacto” (Fig. 3).

There are two remaining important parts in the naming of the nonulose. One is the position of carbonyl group involved in the acetal structure forming the pyran ring, which, in this case, is C-2. The sugar cyclizes via the C-6 hydroxyl group and the C-2 carbonyl group resulting in a six-membered pyranose structure (2-uloxyranose). There is not a need to indicate C-6 as the position is involved in forming a hemiacetal structure; that is inferred from the pyran ring. The last problem is the stereochemistry of the glycosidic linkage. In the case of carbohydrates consisting of more than four chiral centers, the following definition is applied, which is similar to that use for naming of the other sugar: “If multiple configurational prefixes are used, the anomeric reference atom is the highest-numbered atom of the group of chiral centers next to the anomeric center that is involved in the heterocyclic ring and specified by a single configurational prefix.” According to this definition, the anomeric oxygen forming the glycosidic bond is in the same orientation (cis) with the C-7 hydroxyl group, the highest numbered atom in the galacto system. Thus, the glycosidic bond is in the \( \alpha \)-configuration. The rest of the substituents in the molecule are the protecting groups of the other individual functional groups, which are indicated by type and position. Substituents are organized alphabetically.

The nomenclature of the type of carbohydrates is complex but it is logical. A series of related structures and their names are shown in Fig. 4.

Communicated by Osamu Kanie
Editor-in-Chief of Trends in Glycoscience and Glycotechnology

References
http://www.chem.qmul.ac.uk/iupac/2carb/