

## Introduction

## Description

**Why are glycans so important ?** Glycans, also known as carbohydrates, saccharides or sugars, cover all living cells from humans to virus going through plants or bacteria, among others. These features make glycans the first contact point in any host-guest interaction, spanning from symbiotic to pathogenic relationships. In humans and animals, they are found on cells surfaces constituting the so-called glycocalyx, an enriched coating composed of a dense and complex array of glycans that plays an essential role in host protection against pathogens. Nevertheless, in other species such as prokaryotes, plants or fungi, glycoconjugates comprise the cell wall. Thus, glycans can play diverse roles, including critical functions as cell signalling, molecular recognition, immunity and inflammation. Besides, they are also responsible for energy storage and cellular support.

**Glycans structure** Together with lipids, proteins and nucleic acids, glycans constitute one of the main classes of biomolecules pivotal for living organisms. However, carbohydrates hold structural diversity and information content, potentially several higher than any other biological macromolecule orders. Glycans are indeed complex structures of multiple monosaccharides linked through glycosidic bonds. Monosaccharides are the simplest and most basic carbohydrate unit. The starting building blocks of larger structures are denoted as oligosaccharides if composed of up to 9 or 10 units or polysaccharides if larger. The general chemical formula of monosaccharide is  $(CH_2O)_n$  being the  $n$  value the one that allows first classifications of the monosaccharides into pentoses ( $n=5$ ), hexoses ( $n=6$ ) and heptoses ( $n=7$ ), among others. Another distinctive trait for the classification permits the differentiation between the sugars containing an aldehyde group, aldoses, ketone, ketoses.(1)

Oligo- and poly-saccharides are the result of a condensation polymerisation reaction between these basic monosaccharides. The fact that monosaccharides can be linked together in many different sites and also with different spatial orientations confers glycans a high information content and complexity. It contrasts with the simple ways that building blocks of proteins and nucleic acids—so the amino acids and nucleotides, respectively—are linked together. Protein and nucleic acid biopolymers are linear, and every building block is linked to the next through the same kind of connection.(2) Depending on the forming sugars, polysaccharides can be divided into homopolysaccharides and heteropolysaccharides. As the names indicate, a single monosaccharide type unit comprises homopolysaccharides, while heteropolysaccharides are those constituted by two or more different monomers. (3)

Interestingly, the sugar variety present in eukaryotic glycoproteins or bacteria is completely different. Unlike mammals, where glycans are assembled from a small group of common monosaccharides, the number of bacterial monosaccharides is unknown and potentially endless. In detail, the monosaccharides commonly found in eukaryotic glycans and glycoconjugates include approximately ten sugars : hexoses, amino-hexoses, deoxy-hexoses, pentoses and the well-known family of the sialic acids. However, there are more than one hundred known monosaccharide constituents in bacteria : including peculiar monosaccharides such as hexuronic acids, heptoses, octulosonic acids, branched monosaccharides and monosaccharides with noncarbohydrate substituents **inter alia**.

**Figure 1** depicts the 1D, 2D and 3D descriptors of some of the bacterial monosaccharides.

**d-AAT 2-acetamido-4-amino-2,4,6-trideoxy-d-galactose**



**Acofriose Î±-L**

The PDB logo consists of the letters "PDB" in a white, sans-serif font, centered within a solid blue square.

**Acofriose Î<sup>2</sup>-D**

The PDB logo consists of the letters "PDB" in a white, sans-serif font, centered within a solid blue square.

**Bradyrhizose Î<sup>2</sup>-D**

The PDB logo consists of the letters "PDB" in a white, sans-serif font, centered within a solid blue square.

**Caryose**

The PDB logo consists of the letters "PDB" in a white, sans-serif font, centered within a solid blue square.

**Caryophyllose**

The PDB logo consists of the letters "PDB" in a white, sans-serif font, centered within a solid blue square.

**DAG α-D**

The PDB logo consists of the letters "PDB" in a white, sans-serif font, centered within a solid blue square.

**DAG β<sup>2</sup>-D**

The logo for the Protein Data Bank (PDB), consisting of the letters "PDB" in white, sans-serif font, centered within a solid blue square.

**DHA Î²-D**

The logo for the Protein Data Bank (PDB), consisting of the letters "PDB" in white, sans-serif font, centered within a solid blue square.

**Heptose**

The PDB logo consists of the letters "PDB" in a white, sans-serif font, centered within a solid blue square.

**Heptose 6-deoxy**

The PDB logo consists of the letters "PDB" in a white, sans-serif font, centered within a solid blue square.

**Kdo**

The logo consists of a solid blue square. In the center of the square, the letters "PDB" are written in a white, bold, sans-serif font.

**Ko**

The logo consists of a solid blue square. In the center of the square, the letters "PDB" are written in a white, bold, sans-serif font.

**leg**



The PDB logo consists of the letters "PDB" in a white, sans-serif font, centered within a solid blue square.

**MurnAc**

The PDB logo consists of the letters "PDB" in a white, sans-serif font, centered within a solid blue square.

**Pse**

The logo for the Protein Data Bank (PDB) is a blue square with the letters "PDB" in white, sans-serif font.