

wwPDB.org

Improving Carbohydrates in the PDB for 2020

~10% of PDB structures contain carbohydrates.

To improve the ability to search for these structures and their complex chemistry (e.g., stereo-isomers, anomeric configurations, branched chains), wwPDB is embarking on a new remediation effort that will:

- Standardize Chemical Component Dictionary definitions
- Adopt carbohydrate nomenclature following IUPAC-IUBMB recommendations
- Provide a uniform representation for oligosaccharides
- Adopt glycoscience community naming conventions
- Incorporate glycoscience community tools to properly identify, validate, and biocurate glycosylation sites in PDB structures

New sequence descriptions will be included in PDB structure data files:

Condensed IUPAC:

Implementation

LFucpa1-6[DManpa1-3DManpb1-4DGlcpNAcb1-4][LFucpa1-3]DGlcpNAcb1-ASN

LINUCS:

PDB ID 1b5f

 $[[ASN]{[(4+1)][b-D-GlcpNAc]{[(3+1)][a-L-Fucp]}](4+1)][b-D-GlcpNAc]{[(4+1)][b-D-Manp]{[(3+1)][a-D-Manp]}](6+1)][a-L-Fucp]{}](6+1)[[a-L-Fucp]{}](6+1)][a-L-Fucp]{}(6+1)[[a-L-Fucp]{}](6+1)][a-L-Fucp]{}(6+1)[[a-L-Fucp]{}](6+1$

These descriptions can be translated into **Symbolic** representations used by glycoscientists:



Example based on PDB Entry 6cmg

This project is described at wwpdb.org/documentation/carbohydrate-remediation and examples are available at github, github.com/pdbxmmcifwg

Information about this project will be updated at wwPDB.org.