Help

Glyco3D is a portal of databases covering the three-dimensional features of monosaccharides, disaccharides, oligosaccharides (Conformations and NMR spectra), polysaccharides, glycosyltransferases, lectins, monoclonal antibodies against carbohydrates, and glycosaminoglycan-binding proteins. These databases have been developed with non-proprietary software and they are opened freely to the scientific community. Each individual database stands by itself as it covers a particular field of structural glycosciences.

The condensed IUPAC nomenclature has been adopted for the structural encoding of all structures. Further structural depictions involve chemically relevant representations as well as commonly accepted pictorial representations used in glycobiology and a computing encoding under the GlycoCT convention.

Whereas each of the databases has been set to account for the specific features of the class of molecules covered, the set of databases in Glyco3D share common facilities. An interactive front-end offers the users to extract data from the tables on the relational database on '**Data Query**"' and display the retrieved information in a coherent manner on "**Results**". The "**Data Query"** page comprises primarily two levels of increased complexity to query the database, i.e. **Simple and Advanced** searches. The "**Results**" page details the results which are organized under two tabs namely, **Molecule Information** and **View and Download.**
A search engine is available that scans the full content of all the databases for queries related to sequential information of the carbohydrates or other related descriptors. This is performed under the **search sequence** command.

Monosaccharides. This annotated database contains the 3D structural information of about 150 entries of monosaccharides selected for their biological relevance. For establishing the 3D database, they all have been subjected to geometry optimization throughout molecular mechanics. Some cases exist such as in the case of iduronic acid, idose, and all furanosides where several ring shapes can occur. In these cases, the low energy conformations are available for each entry.

Disaccharides. This annotated database contains the 3D structural information of about 150 entries of disaccharides. These disaccharides constitute molecules in their own rights and they constitute the building blocks of the vast majority of oligosaccharides, complex carbohydrates plants and algal polysaccharides and some bacterial polysaccharides. For each disaccharide, potential energy maps have been calculated and displayed and the atomic coordinates of the low energy conformations can be downloaded.

BiOligo is an annotated database that contains the 3D structural information of more than 250 entries of bioactive oligosaccharides (referred to as "glycan determinants"). For establishing the database, the three-dimensional structures of each constituent was generated using a combination of the available carbohydrate molecular builders. Once constructed, the glycans were subjected to systematic conformational sampling to determine their conformational preferences. In such cases, several low energy conformations (1 to 5) are available for each entry and displayed on the screen whereas the corresponding set of atomic coordinates can be downloaded.

NMR is experimentally based that covers glycans determinants which are a sub-set of the group of bioactive oligosaccharides which constitute the core of the BiOligo database. They have been systematically organized using standard names in the field of glycobiology, into 31 categories and sub-categories. The Glycan Determinants in the NMR Database constitute a sub set of the entries of BiOligo. For each of these glycans, the experimental work encompassed the recording and interpretation of ^{1}H and ^{13}C spectra, along with COSY, TOCSY, HMQC, HMBC correlations spectra. These spectra can be displayed and the corresponding images downloaded.

Polysaccharides. PolySac3DB is an annotated database that contains the 3D structural information of about 160 polysaccharide entries that have been collected from an extensive screening of scientific literature. The molecular representations of the repeat unit, the macromolecular chain, and in some instance, the packing features, can be displayed; the corresponding set of atomic coordinates can be downloaded

Lectins are oligomeric proteins that can specifically recognize carbohydrates, which as per present knowledge act as macromolecular tools to decipher sugar-encoded messages. More than 1,500 lectin three-dimensional structures are available in the database (July 2015). Most of them have been determined by x-ray diffraction, although some neutron diffraction structures are available as well as NMR solution structures or theoretical models. About 70 % of these structures have been determined in complex with a carbohydrate ligand, ranging from monosaccharides to oligosaccharides or glycoproteins. The comment section indicates whether the lectin has been solved in the form of a protein-carbohydrate complex. In that case, the nature of sugar is indicated along with its sequence. The section dealing with molecule information indicates the following items: origin, class, family, species, PDB code, resolution, comment, and reference. Provision is also given to view an image of the source of the protein, along with a still three-dimensional ribbon type representation of the three-dimensional structure, together with access to original 3D information at the Protein Database. Links to NIH sites for references and taxonomy are also provided.

Glycosaminoglycans (abbreviated to as GAGs) comprise a family of complex anionic polysaccharides including heparin, heparan sulfate, chondroitin sulfate, dermatan sulfate, keratan sulfate, and hyaluronic acid. Due to difficulty of obtaining crystals of complexes between protein and glycosaminoglycans, a limited number of structures are available in the Protein Data Bank. Different proteins have been co-crystallized with heparin oligosaccharides. Most of them are of animal origin with the exception of some bacterial enzymes viral proteins. The section dealing with molecule information provides information : protein, classification, GAG type, species, PDB code, resolution, and length of oligosaccharide, comments and reference. Provision is also given to view an image of the source of the protein, along with a graphical representation of the three-dimensional structure. Links to

Antibodies are glycoproteins belonging to the immunoglobulin superfamily. The present database is concerned with the set of high resolution structures of carbohydrate–antibody complexes (about 70). ***Information*** lists the: origin, name of the antibody, nomenclature of the bound carbohydrate, PDB code, resolution, comment, immunoglobulin class, reference to the original article. Provision is given to view a still three-dimensional ribbon type representation of the three-dimensional structure, and to display a three-dimensional representation of the complex which has been constructed from the

reported atomic coordinates. A particular emphasis is given to indicate the conformation of the bound carbohydrate which can be viewed.

Glycosyltransferases (GTs) constitute a ubiquitous group of enzymes that catalyze the synthesis of glycosidic linkages by the transfer of a sugar residue from a donor to an acceptor. At the present time, more than 415 GTs crystal structures are available, that have been grouped in 40 families. The classification of the GTs proteins is made based on their origin: (i) animal; (ii) archea; (iii) bacteria; (iv) plant; (v) virus; (vi) yeast & fungi. Upon selecting one organism, a click opens a new menu that prompts the user to choose among a sub-classification based on the fold, i.e. GT-A or GT-B. For each GT, the information signet provides the following information: Enzyme name, Short name, Origin, Organism, Resulting Linkage, Fold, CAZY family, Mechanism. As regard to the crystal structure, the PDB code, the resolution, the nature of the complex (if any), comments and reference are given. One or more graphical representations are provided, along with the possibility to download the atomic coordinates.