



Review

# Three-Dimensional Structures of Carbohydrates and Where to Find Them

Sofya I. Scherbinina <sup>1,2,\*</sup> and Philip V. Toukach <sup>1,\*</sup>

<sup>1</sup> N.D. Zelinsky Institute of Organic Chemistry, Russian Academy of Science, Leninsky prospect 47, 119991 Moscow, Russia

<sup>2</sup> Higher Chemical College, D. Mendeleev University of Chemical Technology of Russia, Miusskaya Square 9, 125047 Moscow, Russia

\* Correspondence: shcherbininasi@yandex.ru (S.I.S.); netbox@toukach.ru (P.V.T.)

Received: 26 September 2020; Accepted: 16 October 2020; Published: 18 October 2020



**Abstract:** Analysis and systematization of accumulated data on carbohydrate structural diversity is a subject of great interest for structural glycobiology. Despite being a challenging task, development of computational methods for efficient treatment and management of spatial (3D) structural features of carbohydrates breaks new ground in modern glycoscience. This review is dedicated to approaches of chemo- and glyco-informatics towards 3D structural data generation, deposition and processing in regard to carbohydrates and their derivatives. Databases, molecular modeling and experimental data validation services, and structure visualization facilities developed for last five years are reviewed.

**Keywords:** carbohydrate; spatial structure; model build; database; web-tool; glycoinformatics; structure validation; PDB glycans; structure visualization; molecular modeling

## 1. Introduction

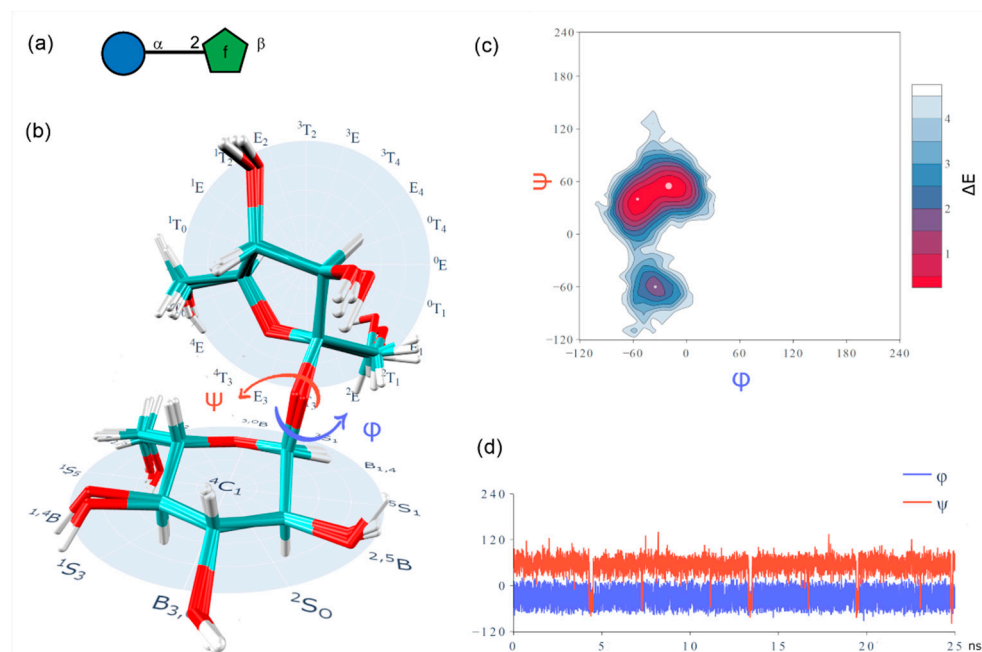
Knowledge of carbohydrate spatial (3D) structure is crucial for investigation of glycoconjugate biological activity [1,2], vaccine development [3,4], estimation of ligand-receptor interaction energy [5–7] studies of conformational mobility of macromolecules [8], drug design [9], studies of cell wall construction aspects [10], glycosylation processes [11], and many other aspects of carbohydrate chemistry and biology. Therefore, providing information support for carbohydrate 3D structure is vital for the development of modern glycomics and glycoproteomics.

As result of growing interest to glycoprofiling, glycan microarrays, carbohydrate active enzymes (CAZy) and glycan-binding proteins (GBP) which are involved in biological processes, several major international projects (e.g., GlySpace [12], GlyCosmos [13], Glycomics@ExPASy [14], GlyGen [15], JCGGDB [16], Glytoucan [17], MIRAGE [18], CFG [19], RINGS [20], GLIC (<https://glic.glycoinfo.org/>), SysGlyco (<https://sysglyco.org/>)) were launched to integrate variety of data produced by glycobiological research. The main goal of existing glycoinformatics projects is to provide versatile resources with user-friendly access helpful for disease diagnostics [21,22], glycobioinformatics studies [23], glycosylation site prediction [24], CAZy activity prognosis [25,26] and other applications.

Appending of structural repositories with 3D structural data opens the way for computational glycobiology and modeling of carbohydrate structures at atomic resolution. Design of novel workflows and techniques to connect carbohydrate spatial structure modes and experimental data with verification, processing, analysis and deposition of associated data has gained increased popularity in glycoscience community [27]. A Carbohydrate Structure Database (CSDB, [28]) module for carbohydrate 3D structure modeling is a demonstrative example of 3D structural data integration facilities (as a database) combined with dedicated interface (as a glycoinformatics project). Further details on CSDB 3D facilities are discussed below.

The typical types of knowledge about a carbohydrate 3D structure include (Figure 1):

- Primary structure (atom connectivity);
- Monosaccharide ring conformation;
- Rotational states of inter-residue and exocyclic linkages and their energies;
- Ring puckering and transitions of glycosidic linkage conformation on a time scale;
- Large-scale spatial arrangement (tertiary structure).

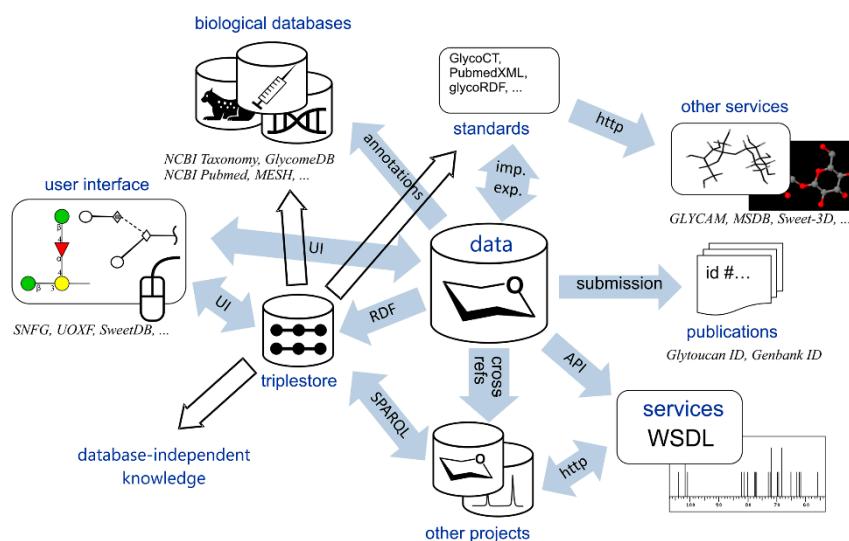


**Figure 1.** Typical components of a carbohydrate 3D structure exemplified on sucrose: (a) primary structure (in Symbol Nomenclature for Glycans (SNFG)); (b) superimposed conformational states and Cremer–Pople diagram; (c) conformational space of a two-torsion glycosidic linkage (Ramachandran plot); (d) transitions of glycosidic dihedrals.

Herein we focus on the important aspects of carbohydrate 3D structure availability to researchers: structural repositories; glycoinformatics tools and workflows to assist structure building, modeling and erroneous molecular geometry data detection and remediation; carbohydrate 3D structure presentation and visualization methods.

## 2. Structural Databases

Structural databases make significant contribution to bringing information technologies to glycoscience [29]. With no focus on spatial structure, glycan databases and online tools have been recently reviewed [30–32]. Depositing huge number of carbohydrates with detailed data for each entry, databases are valuable sources of structural information, biological assignments, references and external links. Structural data are often accompanied by original and sometimes assigned experimental observables: NMR spectra, HPLC and MS profiles, etc. The services built on top of the databases can include 3D structure simulation, validation, and storage. A viewpoint of the authors at the ideal integration of data resources and services in glycoinformatics is summarized in Figure 2. A subject of this review is databases providing theoretical or empirical 3D structures of carbohydrates and related data-mining tools.



**Figure 2.** Networking between glycoinformatics projects and related services that promotes achievement of data integration in glycomics. Reproduced with permission from [29], © 2020 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim.

The majority of existing repositories for carbohydrate 3D structures offer open-access data via web interface. Deposited datasets can be represented by glycoproteins, protein-carbohydrate complexes, poly- and oligosaccharides with 3D structure experimentally resolved or specified by means of NMR, X-ray crystallography, cryoEM, small angle X-ray scattering, etc. [27]. Several databases such as GLYCAM-Web, EK3D, 3DSDSCAR, GlycoMapsDB contain data from molecular dynamics simulations. We have also mentioned databases featuring information on protein structures involving carbohydrate moiety in terms of glycosylation (as post-translational modification, dbPTM), carbohydrate active enzymes (CAZY) and homology modeling (SWISS-MODEL). Table 1 displays currently active structural databases maintaining three-dimensional data on carbohydrates.

For Table 1, we have selected carbohydrate and related databases using the following criteria:

- Database can be freely accessed through web user interface;
- Database must contain experimentally confirmed and/or predicted 3D structures (preprocessed and/or generated on-the-fly from a primary structure input) of glycans, glycoproteins, or protein-carbohydrate complexes;
- Stored 3D structures must be deposited as atomic coordinates in PDB, MOL, or other format, and the structures must contain a saccharide moiety;
- Databases with records linked to other large 3D data collections (e.g., RCSB PDB, PDBe, PDBj, PDBsum, UniProtKB etc.) are included in Table 1 (as long as database entries contain carbohydrate moiety, e.g., as a part of a lectin or an antibody);
- Databases with derived carbohydrate 3D structural data (conformational maps, conformer energy minima, etc.) are included in Table 1 even if they provide no atomic coordinates (e.g., GlycoMapsDB and GFDB).

Despite no fit to the criteria above, assistance of large structure repositories offering only glycan primary structures (e.g., GlyToucan [17] (<https://glytoucan.org/>), UniCarbKB [33] (<http://www.unicarbkb.org/>)) can be useful for cross-referencing of existing carbohydrate resources and serve as supplementation to 3D modeling pipelines.

Some out-of-date projects, such as Complex Carbohydrate Structural Database (CCSD) [34,35], EUROCarbDB [33,36], GlycomeDB [36–38], Glycoconjugate Data Bank [39], GlycoSuite [40,41] are noteworthy as they had shaped the modern vision of structural glycoinformatics.

**Table 1.** Carbohydrate databases with 3D structure support.







Database	Years <sup>a</sup>	Description <sup>b</sup>	Data Coverage	Carbohydrate 3D Structures	References
<i>Structure-centric</i>					
Carbohydrate Structure Database (CSDB) 	2005–present	<ul style="list-style-type: none"> <li>structures from prokaryotes, plants, and fungi</li> <li>taxonomy</li> <li>diseases</li> <li>bibliography</li> <li>curated</li> </ul>	<ul style="list-style-type: none"> <li>24669 structures</li> <li>12521 organisms</li> <li>9353 publications</li> <li>2096 glycosyltransferase activities</li> <li>13378 NMR spectra (<sup>1</sup>H, <sup>13</sup>C)</li> </ul>	<ul style="list-style-type: none"> <li>1327 disaccharide conformational maps</li> <li>3D atomic coordinate generation</li> </ul>	[28,42–44] ( <a href="http://csdb.glycoscience.ru/database">http://csdb.glycoscience.ru/database</a> )
Glycosciences.DE 	1997–present	<ul style="list-style-type: none"> <li>taxonomy</li> <li>bibliography</li> </ul>	<ul style="list-style-type: none"> <li>26559 structures</li> <li>20211 publications</li> <li>3434 NMR spectra (<sup>1</sup>H, <sup>13</sup>C)</li> </ul>	<ul style="list-style-type: none"> <li>13599 3D structure models</li> <li>12098 PDB entries (1880 distinct glycan structures)</li> <li>2585 conformational maps</li> <li>3D atomic coordinate generation</li> </ul>	[45–47] ( <a href="http://www.glycosciences.de/">http://www.glycosciences.de/</a> )
Glyco3D 	2015–present	<ul style="list-style-type: none"> <li>taxonomy</li> <li>bibliography</li> <li>curated</li> </ul>	<ul style="list-style-type: none"> <li>245 monosaccharides</li> <li>125 disaccharides</li> <li>314 bio-oligosaccharides</li> <li>140 polysaccharides</li> <li>415 GT structures</li> <li>88 mAb structures</li> <li>46 GAG structures</li> <li>1662 lectin structures</li> <li>X-ray data *</li> <li>NMR data *</li> <li>molecular modeling data *</li> </ul>	<ul style="list-style-type: none"> <li>3035 3D structures *</li> <li>PDB entries *</li> <li>disaccharide conformational maps *</li> <li>3D atomic coordinate generation</li> </ul>	[48,49] ( <a href="http://glyco3d.cermav.cnrs.fr/home.php">http://glyco3d.cermav.cnrs.fr/home.php</a> )
PolySac3DB 	2012–present	<ul style="list-style-type: none"> <li>polysaccharides</li> <li>taxonomy</li> <li>bibliography</li> <li>curated</li> </ul>	<ul style="list-style-type: none"> <li>157 structures</li> <li>84 publications</li> <li>X-ray data *</li> <li>NMR data *</li> <li>molecular modeling data *</li> </ul>	<ul style="list-style-type: none"> <li>157 3D structures</li> <li>PDB entries *</li> <li>conformational maps *</li> </ul>	[50] ( <a href="http://glyco3d.cermav.cnrs.fr/home.php">http://glyco3d.cermav.cnrs.fr/home.php</a> )
EK3D 	2016–present	<ul style="list-style-type: none"> <li><i>E. coli</i> K antigens</li> <li>bibliography</li> <li>curated</li> </ul>	<ul style="list-style-type: none"> <li>molecular modeling data</li> <li>protein data</li> </ul>	<ul style="list-style-type: none"> <li>72 3D structures</li> <li>3D atomic coordinate generation</li> </ul>	[51] ( <a href="http://www.iith.ac.in/EK3D/">www.iith.ac.in/EK3D/</a> )
3DSDSCAR 	2010–present	<ul style="list-style-type: none"> <li>sialic acid-containing oligosaccharides</li> <li>aqueous molecular dynamics simulations</li> </ul>	<ul style="list-style-type: none"> <li>27 structures</li> </ul>	<ul style="list-style-type: none"> <li>92 3D conformational models</li> </ul>	[52,53] ( <a href="http://aliffishbay.com/Domains/3dsdscar.org/3dsdscar.html">http://aliffishbay.com/Domains/3dsdscar.org/3dsdscar.html</a> )



Table 1. Cont.







Database	Years <sup>a</sup>	Description <sup>b</sup>	Data Coverage	Carbohydrate 3D Structures	References
 MatrixDB	2011–present	<ul style="list-style-type: none"> <li>protein–polysaccharide interactions</li> <li>taxonomy</li> <li>genetic data</li> <li>bibliography</li> <li>curated</li> </ul>	<ul style="list-style-type: none"> <li>58 GAG sequences</li> <li>proteoglycan structures *</li> <li>1507 experiments</li> <li>1058 experimentally supported associations</li> <li>269 publications</li> </ul>	<ul style="list-style-type: none"> <li>3D structures *</li> <li>PDB entries *</li> <li>3D-atomic coordinates generation (GAGs)</li> </ul>	[54–56] ( <a href="http://matrixdb.univ-lyon1.fr/">http://matrixdb.univ-lyon1.fr/</a> )
 EPS-DB	2017–present	<ul style="list-style-type: none"> <li>bacterial exopolysaccharides</li> <li>functional properties</li> <li>genetic data</li> <li>taxonomy</li> <li>bibliography</li> <li>curated</li> </ul>	<ul style="list-style-type: none"> <li>105 structures</li> </ul>	<ul style="list-style-type: none"> <li>85 3D structure models</li> <li>3D-atomic coordinates generation</li> </ul>	[57] ( <a href="http://www.epsdatabase.com">http://www.epsdatabase.com</a> )
 GlycoMDB GLYCAN	2020–present	<ul style="list-style-type: none"> <li>glycan microarrays</li> </ul>	<ul style="list-style-type: none"> <li>5203 glycan microarray samples</li> </ul>	<ul style="list-style-type: none"> <li>1965 3D structures (PDB entries)</li> <li>771 3D structures with glycan ligands (PDB entries)</li> </ul>	[58] ( <a href="http://www.glycanstructure.org/glymdb/">http://www.glycanstructure.org/glymdb/</a> )
 CFG Glycan Structures Database	2006–present	<ul style="list-style-type: none"> <li>mammalian glycan arrays</li> <li>taxonomy</li> <li>biological sources</li> <li>diseases</li> <li>bibliography</li> </ul>	<ul style="list-style-type: none"> <li>N-glycans *</li> <li>O-glycans *</li> </ul>	<ul style="list-style-type: none"> <li>3D-atomic coordinates generation</li> </ul>	[59,60] ( <a href="http://www.functionalglycomics.org/glycomics/molecule/jsp/carbohydrate/carbMoleculeHome.jsp">http://www.functionalglycomics.org/glycomics/molecule/jsp/carbohydrate/carbMoleculeHome.jsp</a> ) ( <a href="http://www.functionalglycomics.org/glycomics/publicdata/selectedScreens.jsp">http://www.functionalglycomics.org/glycomics/publicdata/selectedScreens.jsp</a> )
<i>Glycoproteomic</i>					
 GlycoNAVI TCarp	2020–present	<ul style="list-style-type: none"> <li>diseases</li> <li>genetic data</li> <li>taxonomy</li> <li>bibliography</li> </ul>	<ul style="list-style-type: none"> <li>2723 unique analyzed glycans</li> <li>5814 glycoproteins</li> <li>712 lectins</li> </ul>	<ul style="list-style-type: none"> <li>3D structures *</li> <li>15003 PDB entries</li> <li>3D atomic coordinate generation</li> </ul>	[61] ( <a href="https://glyconavi.org/TCarp/">https://glyconavi.org/TCarp/</a> )
 GlyCosmos	2017–present	<ul style="list-style-type: none"> <li>diseases</li> <li>genetic data</li> <li>taxonomy</li> </ul>	109854 glycans glycolipids * 50113 glycoproteins 1238 lectins 20580 glycoenes	<ul style="list-style-type: none"> <li>3D structures (PDB and UniProtKB entries) *</li> </ul>	[13,62,63] ( <a href="https://glycosmos.org/">https://glycosmos.org/</a> )

Table 1. Cont.





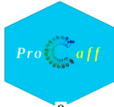


Database	Years <sup>a</sup>	Description <sup>b</sup>	Data Coverage	Carbohydrate 3D Structures	References
 SugarBind	2010–present	<ul style="list-style-type: none"> <li>adherence to pathogens</li> <li>taxonomy</li> <li>diseases</li> <li>bibliography</li> <li>curated</li> </ul>	<ul style="list-style-type: none"> <li>739 lectins</li> <li>204 glycan ligands</li> <li>567 pathogenic agents</li> <li>1266 bindings</li> <li>183 publications</li> </ul>	<ul style="list-style-type: none"> <li>3D lectin structures (PDB entries) *</li> </ul>	[64] ( <a href="https://sugarbind.expasy.org/">https://sugarbind.expasy.org/</a> )
 GlyConnect	2019–present	<ul style="list-style-type: none"> <li>protein glycosylation</li> <li>taxonomy</li> <li>biological sources</li> <li>diseases</li> <li>bibliography</li> <li>curated</li> </ul>	<ul style="list-style-type: none"> <li>2662 glycoproteins</li> <li>3609 glycans</li> <li>246 organisms</li> <li>5675 sites</li> <li>913 publications</li> </ul>	<ul style="list-style-type: none"> <li>3D glycoprotein structures (PDB entries) *</li> </ul>	[65] ( <a href="https://glyconnect.expasy.org/">https://glyconnect.expasy.org/</a> )
 ProGlycProt <small>Protein Glycosylation in Prokaryotes</small>	2012–present	<ul style="list-style-type: none"> <li>prokaryotes</li> <li>taxonomy</li> <li>bibliography</li> <li>homology models *</li> <li>curated</li> </ul>	<ul style="list-style-type: none"> <li>crystal structures</li> <li>61 glycoproteins</li> <li>62 glycosyltransferases</li> <li>38 enzymes/proteins involved in protein glycosylation</li> <li>518 publications</li> </ul>	<ul style="list-style-type: none"> <li>3D structures (PDB entries) *</li> <li>3D homology models (UniProtKB entries) *</li> </ul>	[66,67] ( <a href="http://www.proglycprot.org/">http://www.proglycprot.org/</a> )
 ProCarbDB	2020–present	<ul style="list-style-type: none"> <li>protein-carbohydrate complexes</li> <li>taxonomy</li> <li>bibliography</li> <li>binding affinities</li> <li>curated</li> </ul>	<ul style="list-style-type: none"> <li>5254 complexes</li> <li>867 ligand monomers</li> <li>X-ray data</li> </ul>	<ul style="list-style-type: none"> <li>5254 3D structures (PDB entries)</li> </ul>	[68] ( <a href="http://www.procarbdb.science/procarb/">http://www.procarbdb.science/procarb/</a> )
 Procaff	2019–present	<ul style="list-style-type: none"> <li>protein-carbohydrate complexes</li> <li>taxonomy</li> <li>bibliography</li> </ul>	<ul style="list-style-type: none"> <li>3122 entries</li> <li>228 publications</li> <li>125 organisms</li> <li>354 proteins</li> <li>835 carbohydrates</li> <li>thermodynamic data</li> </ul>	<ul style="list-style-type: none"> <li>335 3D structures (PDB entries)</li> </ul>	[69] ( <a href="https://web.iitm.ac.in/bioinfo2/procaff/index.html">https://web.iitm.ac.in/bioinfo2/procaff/index.html</a> )
 GBSDB GLYCAN	2020–present	<ul style="list-style-type: none"> <li>protein-carbohydrate complexes</li> <li>curated</li> </ul>	<ul style="list-style-type: none"> <li>6402 carbohydrate-containing PDB structures</li> <li>12075 binding sites</li> </ul>	<ul style="list-style-type: none"> <li>6402 3D structures (PDB entries)</li> </ul>	[70] ( <a href="http://www.glycanstructure.org/gbs-db/pdb/">http://www.glycanstructure.org/gbs-db/pdb/</a> )
 PROCARB	2010–present	<ul style="list-style-type: none"> <li>protein-carbohydrate complexes</li> </ul>	<ul style="list-style-type: none"> <li>604 complexes</li> <li>48 modeled glycoproteins</li> <li>100 unique carbohydrate ligands</li> </ul>	<ul style="list-style-type: none"> <li>604 complexes 3D structures (PDB entries)</li> <li>26 N-linked 3D homology models</li> <li>22 O-linked 3D homology models</li> </ul>	[71] ( <a href="http://www.procarb.org/procarb/">http://www.procarb.org/procarb/</a> )

Table 1. Cont.


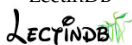





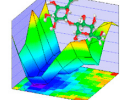
Database	Years <sup>a</sup>	Description <sup>b</sup>	Data Coverage	Carbohydrate 3D Structures	References
UniLectin3D	2019–present	<ul style="list-style-type: none"> <li>lectins</li> <li>taxonomy</li> <li>bibliography</li> <li>curated</li> </ul>	<ul style="list-style-type: none"> <li>2207 structures (1401 interacting with glycan)</li> <li>535 distinct lectin sequences</li> <li>228 distinct glycans</li> <li>896 publications</li> <li>X-ray data</li> </ul>	<ul style="list-style-type: none"> <li>3D structures (PDB entries) *</li> </ul>	[72,73] ( <a href="https://www.unilectin.eu/unilectin3D/">https://www.unilectin.eu/unilectin3D/</a> )
Lectin Frontier 	2015–present	<ul style="list-style-type: none"> <li>lectins</li> <li>taxonomy</li> <li>bibliography</li> </ul>	<ul style="list-style-type: none"> <li>398 structures</li> <li>binding affinities</li> </ul>	<ul style="list-style-type: none"> <li>3D structures (PDB entries) *</li> </ul>	[74] ( <a href="https://acgg.asia/lfdb2/">https://acgg.asia/lfdb2/</a> )
LectinDB 	2006–present	<ul style="list-style-type: none"> <li>lectins</li> <li>taxonomy (all domains, incl. viruses)</li> <li>bibliography</li> <li>curated</li> </ul>	<ul style="list-style-type: none"> <li>789 organisms</li> <li>821 PDB entries</li> </ul>	<ul style="list-style-type: none"> <li>PDB entries *</li> </ul>	[75] ( <a href="http://proline.physics.iisc.ernet.in/lectindb/">http://proline.physics.iisc.ernet.in/lectindb/</a> )
GlycoEpitope 	2006–present	<ul style="list-style-type: none"> <li>epitopes</li> <li>taxonomy</li> <li>diseases</li> <li>functions</li> <li>receptors</li> <li>bibliography</li> <li>curated</li> </ul>	<ul style="list-style-type: none"> <li>178 epitopes</li> <li>624 antibodies</li> </ul>	<ul style="list-style-type: none"> <li>PDB entries (epitopes) *</li> </ul>	[76–78] ( <a href="https://www.glycoepitope.jp/epitopes">https://www.glycoepitope.jp/epitopes</a> )
GlycoCD	2012–present	<ul style="list-style-type: none"> <li>glycan CD antigens</li> <li>bibliography</li> <li>curated</li> </ul>	<ul style="list-style-type: none"> <li>19 glycan CDs</li> <li>44 CRD-CDs</li> </ul>	<ul style="list-style-type: none"> <li>PDB entries *</li> </ul>	[79] ( <a href="http://www.glycosciences.de/glyco-cd/">http://www.glycosciences.de/glyco-cd/</a> )
SACS	2002–present	<ul style="list-style-type: none"> <li>antibodies</li> <li>automatically-updated</li> </ul>	<ul style="list-style-type: none"> <li>3994 entries</li> <li>crystal/EM structure data</li> </ul>	<ul style="list-style-type: none"> <li>PDB entries *</li> </ul>	[80] ( <a href="http://www.bioinf.org.uk/abs/sacs/xslt.cgi?src=antibodies.xml&amp;xsl=summary.xsl">http://www.bioinf.org.uk/abs/sacs/xslt.cgi?src=antibodies.xml&amp;xsl=summary.xsl</a> )
SabDab 	2014–present	<ul style="list-style-type: none"> <li>antibodies</li> <li>automatically-updated taxonomy</li> <li>binding affinities</li> <li>curated</li> </ul>	<ul style="list-style-type: none"> <li>4223 entries</li> <li>111 carbohydrate-containing antigen types</li> <li>experimental data</li> </ul>	<ul style="list-style-type: none"> <li>111 3D structures (PDB entries)</li> </ul>	[81] ( <a href="http://opig.stats.ox.ac.uk/webapps/newsabdab/sabdab/">http://opig.stats.ox.ac.uk/webapps/newsabdab/sabdab/</a> )

Table 1. Cont.

Database	Years <sup>a</sup>	Description <sup>b</sup>	Data Coverage	Carbohydrate 3D Structures	References
 CAZy	1998–present	<ul style="list-style-type: none"> <li>carbohydrate-active enzymes and carbohydrate-binding modules</li> <li>taxonomy</li> <li>genetic data</li> <li>bibliography</li> <li><i>curated</i></li> </ul>	<ul style="list-style-type: none"> <li>CAZy structures *</li> <li>CAZy activities *</li> </ul>	<ul style="list-style-type: none"> <li>7500 <sup>c</sup> 3D structures bearing glycan-containing ligand or a glycan analog revealing enzyme-glycan interactions (PDB entries)</li> </ul>	[82–84] <a href="http://www.cazy.org/">(http://www.cazy.org/)</a>
 dbPTM	2006–present	<ul style="list-style-type: none"> <li>protein post-translational modifications</li> <li>taxonomy</li> <li>diseases</li> <li>genetic data</li> <li>bibliography</li> <li><i>curated</i></li> </ul>	<ul style="list-style-type: none"> <li>32 C-linked glycosylations</li> <li>3289 N-linked glycosylations</li> <li>1860 O-linked glycosylations</li> <li>6 S-linked glycosylations</li> </ul>	<ul style="list-style-type: none"> <li>3D structures (UniProtKB entries) *</li> </ul>	[85–87] ( <a href="http://dbptm.mbc.nctu.edu.tw/">http://dbptm.mbc.nctu.edu.tw/</a> )
SWISS-MODEL Repository 	2004–present	<ul style="list-style-type: none"> <li>3D protein homology models</li> <li>taxonomy</li> <li>regularly updated</li> </ul>	<ul style="list-style-type: none"> <li>glycoprotein structures *</li> <li>1698194 models from SWISS-MODEL for UniProtKB</li> <li>158670 structures from PDB with mapping to UniProtKB</li> </ul>	<ul style="list-style-type: none"> <li>3D structures (PDB and UniProtKB entries) *</li> </ul>	[88–90] <a href="https://swissmodel.expasy.org/repository">https://swissmodel.expasy.org/repository</a>
<i>Specialized</i>					
GlycoMaps DB 	2004–present	<ul style="list-style-type: none"> <li>di- to pentasaccharides</li> </ul>	<ul style="list-style-type: none"> <li><i>in vacuo</i> high-temperature molecular dynamics</li> </ul>	<ul style="list-style-type: none"> <li>conformational maps for 2585 glycosidic linkages</li> </ul>	[91] ( <a href="http://www.glycosciences.de/modeling/glycomapsdb/">http://www.glycosciences.de/modeling/glycomapsdb/</a> )
GFDB <b>GLYCAN</b>	2013–present	<ul style="list-style-type: none"> <li>glycosidic torsion angles</li> <li>clustering analysis</li> </ul>	<ul style="list-style-type: none"> <li>1754 <sup>c</sup> unique glycan sequences in PDB</li> <li>9055 <sup>c</sup> unique fragments with chemical modifications</li> <li>127202 <sup>c</sup> fragment structures</li> </ul>	<ul style="list-style-type: none"> <li>PDB entries *</li> <li>3D-atomic coordinates generation</li> </ul>	[92] ( <a href="http://www.glycanstructure.org/fragment-db">http://www.glycanstructure.org/fragment-db</a> )
GLYCAM-Web <b>GLYCAM</b>	2013–present	<ul style="list-style-type: none"> <li>mammalian glycans</li> </ul>	<ul style="list-style-type: none"> <li>pre-built libraries of predicted 3D structures of common bioglycans</li> </ul>	<ul style="list-style-type: none"> <li>3D structure models *</li> <li>3D-atomic coordinates generation</li> </ul>	<a href="http://glycam.org/Pre-builtLibraries.jsp">http://glycam.org/Pre-builtLibraries.jsp</a>

<sup>a</sup> Where unknown, the year of the first publication is given. <sup>b</sup> Database is marked as *curated* if manual verification of data was reported in the original publication or at the database web site. <sup>c</sup> Published coverage data can be outdated; database interface provides no statistics on current coverage. \* Database provides no search facilities for indicated carbohydrate 3D structural data.

### 3. Carbohydrate 3D Structure Modeling

Methods to probe a 3D structure of carbohydrate-containing biomolecules has been developed for decades. NMR techniques (interatomic distances derived from NOE, and torsion angles derived from coupling constants), X-ray crystallography, and electron cryo-microscopy (the two latter being atomic models built on the basis of electron density map) are among most demanded methods for 3D structural elucidation. These methods have been reviewed [93–96] and are beyond the scope of this review focused in information technologies. For use of instrumental methods for the validation of a simulated structure, please refer to Section 5 “Experimental data validation”.

Structural investigation of large biological systems involving protein-glycan interactions requires leveraging more resources and employing more complex experimental techniques compared to solely oligo- and polysaccharides studies. Advances in NMR methods hold great potential for direct spatial structure determination of carbohydrate-protein complexes in solution based on intermolecular NOEs which affords estimation of atomic contacts between a protein and a carbohydrate ligand [97,98]. Further extraction of NOE-derived distance restraints for a saccharide molecule results in generation of representative conformational ensembles [99–101].

Support of experimental data with computer simulations can significantly improve quality of 3D structures. Quantum mechanics [100,102–106] and molecular dynamics modeling [107–111] are commonly applied to conformation search and NMR signal prediction.

To date, the following theoretical models and methods are applied for *in silico* design of carbohydrate three-dimensional structure [112–116]:

- Molecular mechanics (MM) and molecular dynamics (MD) calculations [117];
- Monte Carlo simulations [118,119];
- Semi-empirical methods [120–123];
- *Ab initio* simulations based on density functional theory (DFT) [124–128];
- Hybrid QM/MM and QM/QM and ONIOM (“our own N-layered integrated molecular orbital and molecular mechanics”) approaches [129–134].

Due to computational limitations, most of publications of the recent decade have reported molecular dynamics approaches in general or dedicated force fields. With increasing computer power, other methods gain interest, however majority of applications of molecular modeling of complex carbohydrates, especially in solution, still use MM/MD methods.

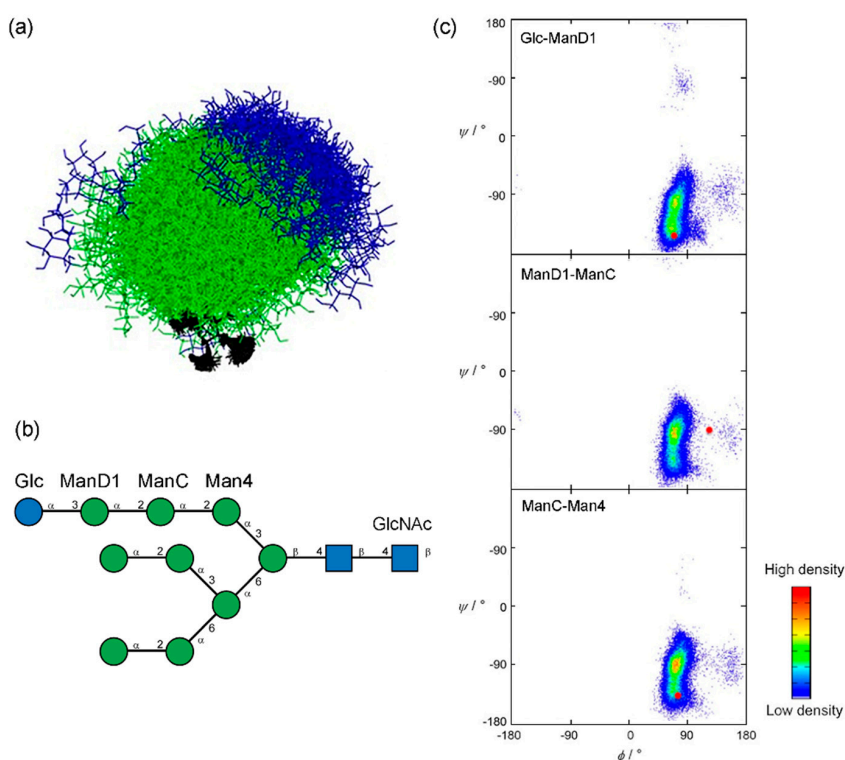
Based on Scopus [135] article count we estimated the application rate for quantum mechanics (10759 publications) and molecular mechanics (14871 publications) methods applied for carbohydrate structure modeling for the recent five years (2015–2020). Search queries included abundant carbohydrate terms, typical glycan moieties, and common modeling approaches (query details are given in Supplementary Table S1). In spite of growing interest to QM approaches in carbohydrate structure simulation, the major contribution to the statistics for such resource-intensive calculations is application of QM to relatively simple model compounds. For complex bioglycans in solution predominance of MM methods is more pronounced [6,8].

#### *Molecular Mechanics and Dynamics*

Molecular dynamics methods have achieved broad scope of application in terms of reasonable computer resource consumption. They fulfill advantageous compromise between calculation accuracy and performance, when applied to glycan molecules and their structural complexity (variety of known monomeric elements, presence of ionogenic groups), high bridge flexibility and stereo-electronic effects [112,113,136,137].

In molecular mechanics simulations, Newtonian mechanics principles are applied to calculate potential energy of a system using parameter set specific for a class of compounds under study (force field). Particular features of carbohydrate moiety, e.g., ring puckering, rotational barriers, hydrogen bonds, must be taken into account to perform precise analysis of molecular behavior *in vacuo* or in solution [138].

Molecular dynamics simulations consider Newtonian motion equations to observe evolution of a system during a certain timespan. Conformation ensemble generation occurs via calculation of molecular trajectory at given temperature. Accuracy of calculation depends on the employed force field and sufficient conformational sampling. MD simulations are commonly used for interpretation and analysis of the NMR and X-ray observables in the context of carbohydrate 3D structure [139]. Enhanced molecular dynamics sampling technologies, such as replica-exchange MD (REMD) [140,141], Hamiltonian replica-exchange MD (HREX) [142–144], multidimensional swarm-enhanced sampling MD (msesMD) [145,146], Gaussian accelerated MD (GAMD) [147,148] have been reported. Density maps or energy maps built for a set of the glycosidic torsion angles ( $\varphi$ ,  $\psi$ ,  $\omega$ ) are a typical way to report conformational preferences of a glycan provided by population analysis of its MD trajectory. As a representative example, conformational characteristics of highly flexible branched oligosaccharide Glc<sub>1</sub>Man<sub>9</sub>GlcNAc<sub>2</sub> (GM9) were investigated by explicit-water REMD study and validated using paramagnetism-assisted NMR spectroscopy [149] (Figure 3a,b). Due to the structural complexity of GM9, adequate exploration of conformational space requires long-timescale simulations. Regular MD simulations of similar manno-oligosaccharides were reported to fail reproduction of experimental data [150]. Replica-exchange approach implies running periodically swapped parallel replicas of the system at different temperatures. Ensemble of GM9 conformers sampled by this method was consistent with the NMR observables. Populated areas of density maps built for glycosidic linkages of Glc<sub>1</sub>Man<sub>3</sub> branch of GM9 (Figure 3c) were close to crystallographic conformations of a linear Glc<sub>1</sub>Man<sub>3</sub> tetrasaccharide (a GM9 determinant recognized by lectins) from PDB.

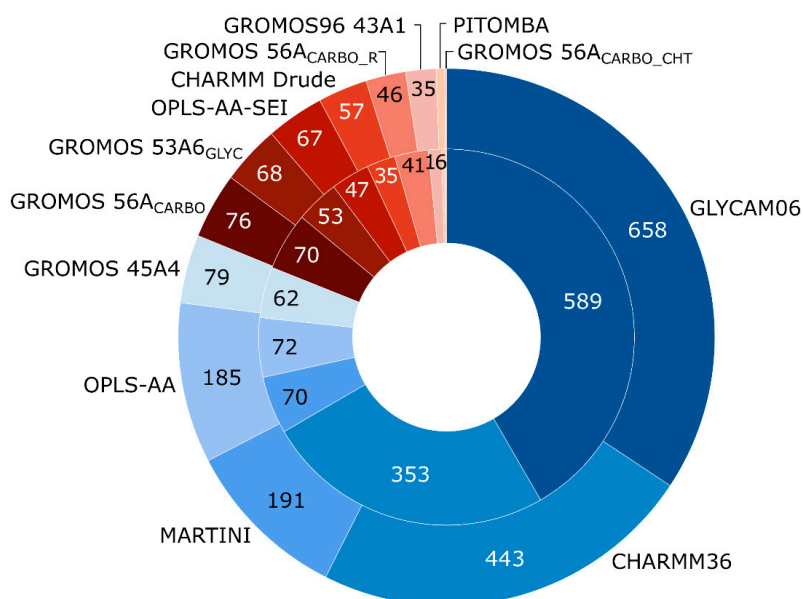


**Figure 3.** NMR-validated conformational analysis of high-mannose oligosaccharide GM9 based on replica-exchange molecular dynamics (REMD) simulation results. (a) Superimposition of 260 GM9 conformers extracted from REMD trajectory (black—GlcNAc, green—Man, blue—Glc). (b) primary structure of the GM9 oligosaccharide (SNFG representation). (c) REMD density maps for  $\varphi$ - $\psi$  torsions of GM9 branch (Glc<sub>1</sub>Man<sub>3</sub>). Red dots locate glycosidic torsion angles derived from crystallographic data of Glc<sub>1</sub>Man<sub>3</sub> tetrasaccharide ligand complexed with the lectin domain of calreticulin (PDB ID: 3O0W). Panels (a) and (c) were reproduced with permission from [149], © 2020 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim.



Force field (or potential energy function) is represented by atomistic parameter set obtained for a considered compound class. Potential energy value can be calculated as a sum of interaction potentials for bonded (covalent bond stretching, angle bending, proper torsions) and non-bonded (electrostatic and van der Waals interactions) terms, and can include other terms (e.g., improper torsions, solvation, hydrogen bonds [151], nonconventional hydrogen bonds [101], for protein-carbohydrate complexes—CH- $\pi$  stacking interactions [152–155], CHI Carbohydrate Intrinsic (CHI) energy contribution [156,157]).

Several force fields developed for general representation of wide range of organic compounds (e.g., Allinger's MM2, MM3, MM4) can be applied to carbohydrate 3D modeling [151,158,159]. Of them, despite being a universal force field, MM3 [160,161] still exhibits good performance on glycans [162–164] (Reviews), [165,166] (exemplary Articles). However, a number of force fields specially tuned for carbohydrates have been developed (Figure 4). In Supplementary Table S2, we provided citation metrics of articles reporting carbohydrate-dedicated and selected general force fields that could be applied to carbohydrate structure modeling. Unfortunately, usage of general force fields could not be adequately estimated via number of citations. Automated full-text analysis and retrieval of data, needed to confirm employment of force fields for carbohydrate molecules, is beyond the scope of this review. Nevertheless, statistical data obtained for general force fields supported in popular MD software packages (e.g., AMBER, CHARMM, GROMACS, Tinker) shows obsolescence of modern force fields above Allinger's ones, and MM3 in particular (see more detailed data, references to original publications and absolute values in Supplementary Table S2).

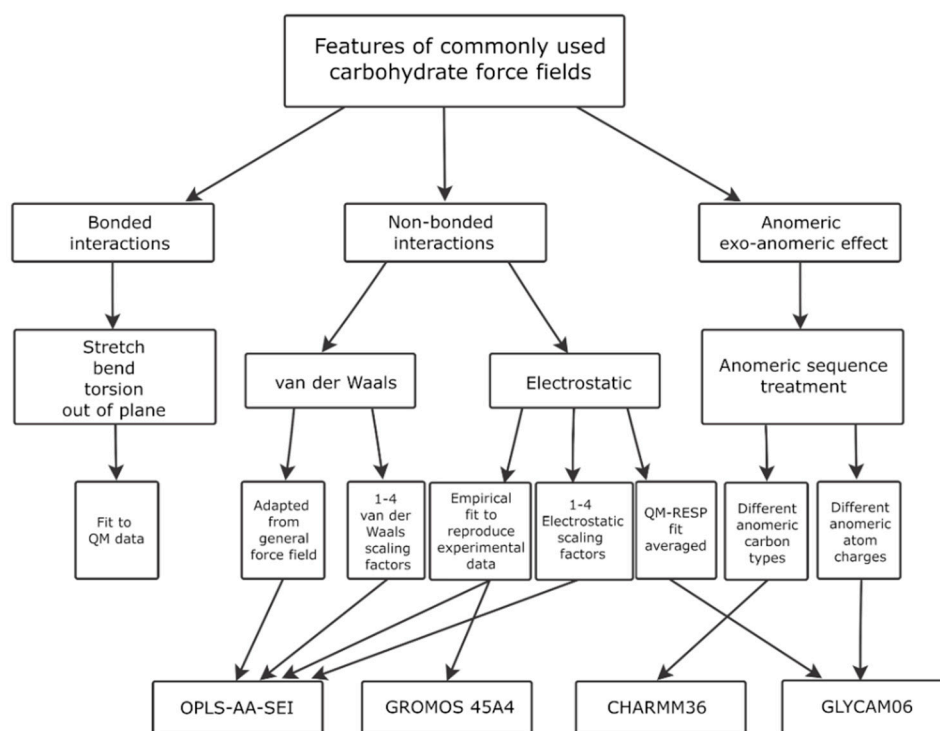


**Figure 4.** Citations of dedicated force fields in carbohydrate studies for the recent five years, according to Scopus. Outer circle shows total citations (number of citing publications) of force fields in 2015–2020. Inner circle shows citations in articles filtered by a carbohydrate topic. See detailed data, references to original publications, absolute values, and carbohydrate filer details in Supplementary Table S2.

Detailed comparisons of all-chemical and dedicated force fields in a context of glycan modeling have been published [114,139,151,167]. CHARMM36, GLYCAM06, GROMOS and OPLS-AA-SEI were reported as commonly used force fields for handling carbohydrate or glycoconjugate molecules. More details are provided in Figure 5.

CHARMM36 force field with modern carbohydrate parameter table (C36 [168]) was derived from CHARMM all-atom biomolecular force field [169,170]. Currently, CHARMM36 parameterization features include monosaccharides in furanose [171] and pyranose [172] forms, glycosidic linkages between monosaccharides [171,173], complex carbohydrates and glycoproteins [174], monosaccharide-linked

sulfate and phosphate groups [175], acyclic carbohydrates and alditols [171], as well as carbohydrate simulations in aqueous solution [176].



**Figure 5.** Digest of the most commonly used carbohydrate force fields with parameterization protocol comparison. Reproduced with permission from [138], © 2020 Elsevier Inc.

GLYCAM06 force field is compatible with carbohydrates of all ring sizes and conformations for both mono- and oligosaccharides built of residues common for mammalian glycans, such as widespread aldoses, N-acetylated amino-sugars, sialic, glucuronic and galacturonic acids [177]. Parameter set was extended to non-carbohydrate moieties such as lipids [178], glycolipids [179,180], lipopolysaccharides [181], proteins and nucleic acids. Parameterization of GLYCAM06 for glycosaminoglycans was reported [182].

GROMOS represents a broad family of carbohydrate force fields. Having been a classic one since 2005, GROMOS 45A4 [183] parameter set is used for explicit-solvent simulation of hexopyranose-based saccharides. In the recent decade, several parameters of 45A4 were optimized in GROMOS 56A<sub>CARBO</sub> [184] including lipopolysaccharides [185]. GROMOS 53A6<sub>GLYC</sub> was improved for explicit-solvent simulations [186] and extended for glycoproteins [187]. GROMOS 56A<sub>CARBO\_R</sub> [188] was designed to improve description of ring conformational equilibria in hexopyranose-based saccharide chains as compared to the previous 56A<sub>CARBO</sub> version. Another modification of 56A<sub>CARBO</sub> named 56A<sub>CARBO\_CHT</sub> [189] was developed for chitosan and its derivatives. Recently, extensions of GROMOS 56A<sub>CARBO/CARBO\_R</sub> parameter set were adapted towards charged, protonated and esterified urines [190] and furanose-based carbohydrates [191]. GROMOS96 43A1 was reported to have good performance on glycan structure simulation in glycoproteins [192,193].

OPLS-AA scaling of electrostatic interactions (SEI) force field [194] consists of improved parameters for conformational changes associated with  $\varphi$ - $\psi$  dihedrals combined with enhanced accuracy of QM relative energy calculation in carbohydrate molecules refined for OPLS-AA biomolecular force field [195,196]. Additionally OPLS force field was improved for explicit-water simulations [197].

Rapidly developing CHARMM Drude polarizable force field for carbohydrates based on classical Drude oscillator has to be mentioned. Parameter sets obtained for hexapyranoses [198] and their aqueous solutions [199], aldopentafuranoses and methyl-aldopentafuranosides [200], carboxylate and

N-acetylamine saccharide derivatives [201], alditols [202] and glycosidic linkages [203] demonstrated significant improvement of QM data reproduction compared to CHARMM additive force field.

MARTINI coarse-grained (CG) force field [204] can be used alternatively to all-atom (AA) level simulations with advantage of modeling large carbohydrate systems (solutions of oligo-, polysaccharides, glycolipids [205–207]) on a long time scale at reasonable computational cost. Blocked ring puckering (only  ${}^4C_1$  conformation is allowed) and restrictions on the anomeric effect and glycosidic bond flexibility cumulatively provide reduction of available degrees of freedom. Another CG model PITOMBA [208] for carbohydrate simulations was developed based on GROMOS 53A6<sub>GLYC</sub> force field.

Docking methods for carbohydrate ligands utilize molecular modeling approaches for protein-carbohydrate complexes for initial geometry generation, conformational sampling, grafting, active site mapping and binding affinity estimation [129,137,209–211]. Accurate reproduction of experimental data requires application of particular scoring function parameterization (empirical, force fields or knowledge-based [212]) and docking protocols, which depend on the interaction types present in a system (CH- $\pi$  interactions, CHI-energy, hydrogen bonding, solvent model, influence of solvent molecules inclusion effects, charged moiety etc.) [8,213–219]. Extension of several docking software packages to handle carbohydrate molecules was reported to improve modeling of biologically relevant systems such as lectin-glycan [220,221], GAG-protein [222–224], or antibody-carbohydrate [225].

#### 4. Model Building and Analysis Tools

Currently available web-based tools along with standalone software packages were developed to facilitate work with carbohydrate 3D structure. Versatile online services for *in silico* molecular modeling allow users to start from a user-friendly structure input, and to automatize further procedures (see Table 2 for references). GLYCAM-Web provides tools for glycan structure prediction, glycosylated protein 3D model generation, grafting and docking. CHARMM-GUI modeler offers options for 3D structure generation and modeling of glycans including N-/O-glycoproteins and glycolipids [226,227]. Biological membranes can be simulated with the assistance of CHARMM-GUI Membrane Builder (by combining features of LPS and glycolipid CHARMM-GUI Modelers) and GNOMM (a tool for building lipopolysaccharide-rich membranes). Noteworthy standalone programming frameworks for structure modeling are Glycosylated (modeling of glycans, glycoproteins and glycosylation) and Rosetta Carbohydrate (loop modeling [228], glycan-to-protein docking, and glycosylation modeling).

To build diverse saccharide 3D models online, one can use such tools as REStLESS and SWEET-II. doGlycans standalone framework can be used for preparation of the atomistic models of glycopolymers, glycolipids and glycoproteins. Complex polysaccharide 3D models can be generated via POLYS and CarbBuilder. Another special class of polysaccharide builders is dedicated to glycosaminoglycans (GAGs) which can be accessed using POLYS GAG-builder and GLYCAM-Web GAG-builder. Recently, another approach for building GAG molecules was reported [229] (exemplary data pipeline only). Unfortunately, application scope of the majority of the existing structure building and modeling services is limited to rigidly defined set of supported sugar residues, and lacks non-carbohydrate moiety support.

Tools for locating and identification of a carbohydrate moiety (e.g., pdb2linucs, GlyFinder, Glycan Reader) are useful for the atomic coordinate analysis and extraction of glycoproteins and protein-carbohydrate complexes deposited in Protein Data Bank (PDB). Automated molecular geometry processing facilities can be accessed via glycoinformatics tools designed for conformational data analysis (CAT, BFMP), nuclear Overhauser effect (NOE) calculation (MD2NOE, Distance Mapping) and 3D structural data analysis related to glycan moieties from PDB (GlyTorsion, GlyVicinity, GS-align).

In Table 2, we summarized freely available tools for generation and processing carbohydrate 3D structural data and divided them into eight categories of application.

**Table 2.** Informatics tools for carbohydrate and glycoprotein modeling, 3D structure prediction and analysis.

Tool	Description	Type <sup>a</sup>	Reference
<i>Structure modeling</i>			
CHARMM-GUI Glycan Modeler	in silico N-/O-glycosylation of proteins; modeling of carbohydrate-only systems	Web-service	[230] ( <a href="http://www.charmm-gui.org/?doc=input/glycan">http://www.charmm-gui.org/?doc=input/glycan</a> )
CHARMM-GUI Glycolipid/LPS Modeler	Glycolipid and lipoglycan structure modeling	Web-service	[230] ( <a href="http://charmm-gui.org/?doc=input/glycolipid">http://charmm-gui.org/?doc=input/glycolipid</a> ) ( <a href="http://charmm-gui.org/?doc=input/lps">http://charmm-gui.org/?doc=input/lps</a> )
Glycosylator	Rapid modeling of glycans and glycoproteins (including glycosylation) based on CHARMM force field	Python framework	[231] ( <a href="https://github.com/tlemmin/glycosylator">https://github.com/tlemmin/glycosylator</a> )
RosettaCarbohydrate	Modeling a wide variety of saccharide and glycoconjugate structures (including loop modeling, glyco-ligand docking and glycosylation)	Python framework	[228,232–234] ( <a href="https://www.rosettacommons.org/docs/latest/application_documentation/carbohydrates/WorkingWithGlycans">https://www.rosettacommons.org/docs/latest/application_documentation/carbohydrates/WorkingWithGlycans</a> )
Azahar	Monte Carlo conformational search and trajectory analysis of glycans	Python framework; PyMol plugin	[235] ( <a href="https://github.com/BIOS-IMASL/Azahar">https://github.com/BIOS-IMASL/Azahar</a> )
Shape	Carbohydrate-dedicated fully automated MM3-based conformation simulation	Standalone software	[236] ( <a href="https://sourceforge.net/projects/shapega/">https://sourceforge.net/projects/shapega/</a> )
Glydict	MM3-based N-glycan structure prediction based on MD simulations	Web-service	[237] ( <a href="http://www.glycosciences.de/modeling/glydict/">http://www.glycosciences.de/modeling/glydict/</a> )
GLYGAL	MM3-based conformational analysis of oligosaccharides	Standalone software	[238]
Fast Sugar Structure Prediction Software (FSPS)	Automatic structure prediction tool for oligo- and polysaccharides in solution	Standalone software	[239–242]
<i>Glycosylation modeling and grafting</i>			
GLYCAM-Web Glycoprotein Builder	Attaching a glycan (user input) to a protein (PDB file)	Web-service	( <a href="http://glycam.org/gp">http://glycam.org/gp</a> )
GlyProt	in silico generation of N-glycosylated 3D models of proteins	Web-service	[243] ( <a href="http://www.glycosciences.de/modeling/glyprot/php/main.php">http://www.glycosciences.de/modeling/glyprot/php/main.php</a> )
Phenix CarboLoad	Loading a carbohydrate structure into protein model and PDB file generation	Python framework	[244] ( <a href="https://www.phenix-online.org/documentation/reference/carbo_load.html">https://www.phenix-online.org/documentation/reference/carbo_load.html</a> )
GLYCAM-Web GlySpec (Grafting)	Prediction of glycan specificity by integrating glycan array screening data and 3D structure	Web-service	[245–249] ( <a href="http://glycam.org/djdev/grafting/">http://glycam.org/djdev/grafting/</a> )

Table 2. Cont.

Tool	Description	Type <sup>a</sup>	Reference
<i>Biological membranes and micelles</i>			
CHARMM-GUI Membrane Builder	Building complex glycolipid-/LPS-/LOS-containing biological membrane systems	Web-service	[230,250–253] ( <a href="http://www.charmm-gui.org/?doc=input/membrane.bilayer">http://www.charmm-gui.org/?doc=input/membrane.bilayer</a> )
GNOMM (gram-negative outer membrane modeler)	Automated building of lipopolysaccharide-rich bacterial outer membranes (3D model preparation for MD simulations in GROMACS)	Standalone software	[254] ( <a href="http://thalis.biol.uoa.gr/GNOMM/">http://thalis.biol.uoa.gr/GNOMM/</a> )
Micelle Maker	Micelle building based on broad range of starting lipids and glycolipids (3D model preparation using AMBER software package and GLYCAM library)	Web-service	[255] ( <a href="http://micelle.icm.uu.se/">http://micelle.icm.uu.se/</a> )
<i>Carbohydrate moiety identification</i>			
Cheminformatics Tool for Probabilistic Identification of Carbohydrates (CTPIC)	Identification of small saccharides and their derivatives (input in SDF or MOL format)	Web-service	[256] ( <a href="http://ctpic.nmrfa.wisc.edu/">http://ctpic.nmrfa.wisc.edu/</a> ) ( <a href="https://github.com/htdashhti/ctpic">https://github.com/htdashhti/ctpic</a> )
Sails	Automated identification of linked sugars	Python framework	( <a href="https://github.com/glycojones/sails">https://github.com/glycojones/sails</a> )
GlyFinder	Locating relevant carbohydrate-containing structures in Protein Data Bank	Part of web-service pipeline	[257,258] ( <a href="https://dev.glycam.org/portal/gf_home/">https://dev.glycam.org/portal/gf_home/</a> )
pdb2linucs	Extraction of carbohydrate data from a PDB file	Web-tool	[259] ( <a href="http://www.glycosciences.de/tools/pdb2linucs/">http://www.glycosciences.de/tools/pdb2linucs/</a> )
GLYCAM-Web PDB-preprocessor	Processing of PDB files with (glyco-)proteins for AMBER-style output	Web-service	( <a href="http://glycam.org/pdb">http://glycam.org/pdb</a> )
Sugar identification program	Identifying the residue names of carbohydrates in a PDB file	Standalone software	( <a href="http://glycam.org/docs/othertoolsservice/downloads/downloads-software/">http://glycam.org/docs/othertoolsservice/downloads/downloads-software/</a> )
Glycan Reader	Automated sugar identification and simulation preparation for carbohydrates and glycoproteins in PDB files	Web-service	[260,261] ( <a href="http://glycanstructure.org/glycanreader/">http://glycanstructure.org/glycanreader/</a> ) ( <a href="http://www.charmm-gui.org/?doc=input/glycan">http://www.charmm-gui.org/?doc=input/glycan</a> )
<i>Structure building and model preparation</i>			
doGlycans	Preparing carbohydrate structures (including polysaccharides, glycolipids and glycoproteins) for GROMACS atomistic simulations	Python framework	[262] ( <a href="https://bitbucket.org/biophys-uh/doglycans/src/master/">https://bitbucket.org/biophys-uh/doglycans/src/master/</a> )
GLYCAM-Web Carbohydrate builder	3D structure prediction of carbohydrates and related macromolecules using GLYCAM06 force field and MD in AMBER (successor of GLYCAM Biomolecule Builder ( <a href="http://glycam.org/old/biombuilder/biomb_index.jsp">http://glycam.org/old/biombuilder/biomb_index.jsp</a> ))	Web-service	[177] ( <a href="http://glycam.org/">http://glycam.org/</a> )
SWEET-II	Rapid 3D model construction of oligo- and polysaccharides with MM3 optimization	Web-service	[263,264] ( <a href="http://www.glycosciences.de/modeling/sweet2/">http://www.glycosciences.de/modeling/sweet2/</a> )

Table 2. Cont.

Tool	Description	Type <sup>a</sup>	Reference
REStLESS API	3D structure generation of carbohydrates and derivatives from CSDB Linear notation with MMFF94 optimization (including aglycone moiety)	Web-service	[265] ( <a href="http://csdb.glycoscience.ru/database/core/translate.html#from">http://csdb.glycoscience.ru/database/core/translate.html#from</a> )
<i>Polysaccharide builders</i>			
POLYS	3D structure generation of poly- and complex oligosaccharides from MM2-precalculated glycosidic linkage torsions and energy minimization	Web-service	[266,267] ( <a href="https://bitbucket.org/polys/polys/src/default/">https://bitbucket.org/polys/polys/src/default/</a> ) ( <a href="http://glycan-builder.cermav.cnrs.fr/">http://glycan-builder.cermav.cnrs.fr/</a> )
CarbBuilder	Building of 3D structures of polysaccharides in CHARMM force field from pre-calculated glycosidic linkage torsions	Standalone software	[268,269] ( <a href="https://people.cs.uct.ac.za/~mkuttel/Downloads.html">https://people.cs.uct.ac.za/~mkuttel/Downloads.html</a> )
GAG-builder	Translating of GAG sequences into 3D models based on POLYS glycan builder	Web-service	[270] ( <a href="http://glycan-builder.cermav.cnrs.fr/gag/">http://glycan-builder.cermav.cnrs.fr/gag/</a> ) ( <a href="http://matrixdb.univ-lyon1.fr/">http://matrixdb.univ-lyon1.fr/</a> )
GLYCAM-Web GAG Builder	Modeling of GAG 3D structure in GLYCAM06 force field using AMBER MD package	Web-service	[271] ( <a href="http://glycam.org/gag">http://glycam.org/gag</a> )
<i>Docking</i>			
BALLDock/SLICK	Protein-carbohydrate complex docking software	Standalone software, a module in docking software	[272,273] ( <a href="https://ball-project.org/download/">https://ball-project.org/download/</a> )
HADDOCK	Modeling of biomolecular complexes with support of glycosylated proteins	Web-service	[274] ( <a href="https://wenmr.science.uu.nl/haddock2.4/library">https://wenmr.science.uu.nl/haddock2.4/library</a> )
Vina-Carb	CHI-energy functions implemented in AutoDock Vina software	Standalone software	[156,157] ( <a href="http://glycam.org/docs/othertoolsservice/download-docs/publication-materials/vina-carb/">http://glycam.org/docs/othertoolsservice/download-docs/publication-materials/vina-carb/</a> )
GLYCAM-Web Antibody docking	Docking of an antibody (from a PDB file) to a glycan antigen (from a library or user input)	Web-service	( <a href="http://glycam.org/ad">http://glycam.org/ad</a> )
Cluspro	Sulfated GAG docking (as one of options)	Web-service	[275,276] ( <a href="https://cluspro.bu.edu/login.php">https://cluspro.bu.edu/login.php</a> )
GAGDock (DarwinDock)	Modification of DarwinDock method for sulfated glycosaminoglycans	Algorithm	[277]
GlycoTorch Vina	Docking of sulfated glycosaminoglycans based on Vina-Carb	Standalone software	[278] ( <a href="http://ericboittier.pythonanywhere.com/">http://ericboittier.pythonanywhere.com/</a> )



Table 2. Cont.

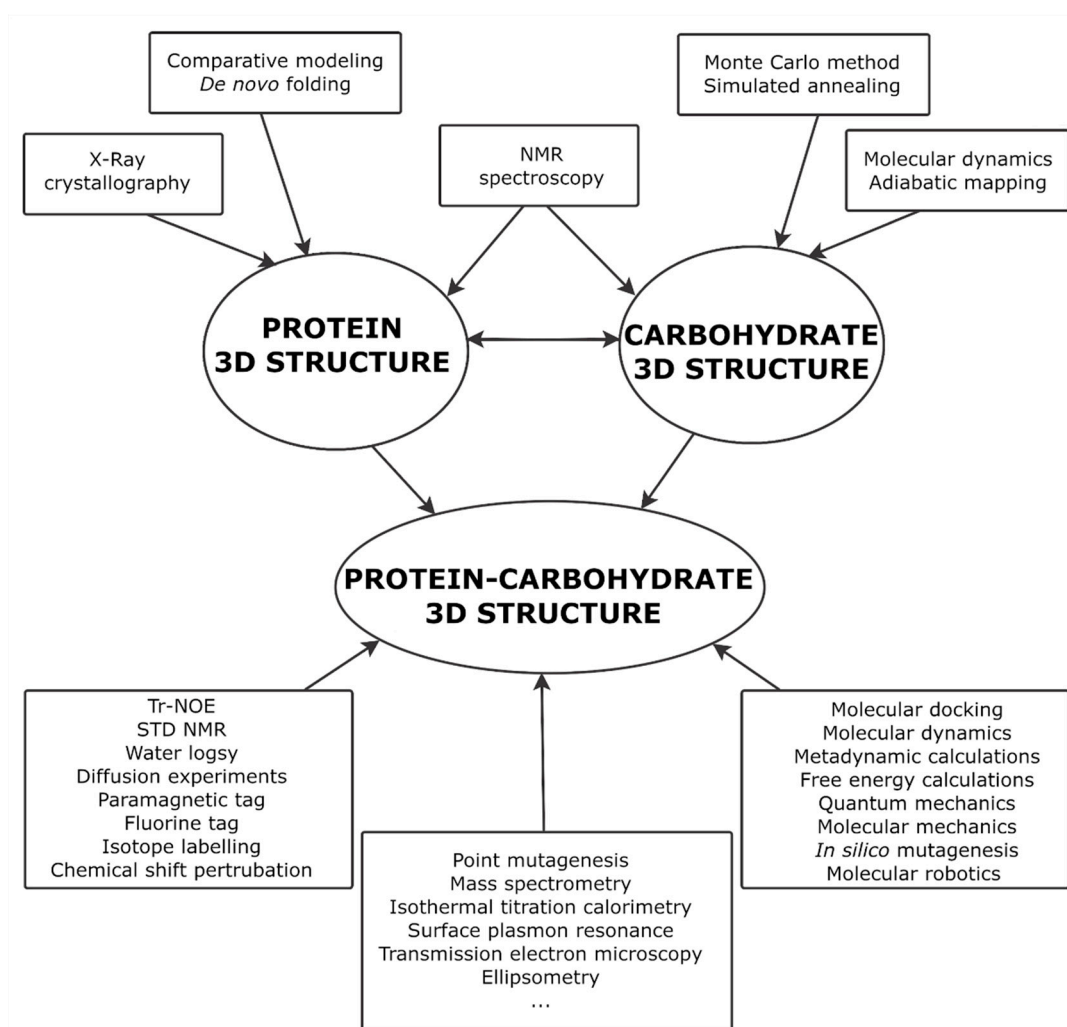
Tool	Description	Type <sup>a</sup>	Reference
<i>Structural data analysis</i>			
Conformational Analysis Tool (CAT)	Analysis of carbohydrate molecular trajectory data derived from MD simulations	Standalone software	[279] ( <a href="http://www.md-simulations.de/CAT/">http://www.md-simulations.de/CAT/</a> )
Best-fit, Four-Membered Plane (BFMP)	Analysis of conformational data from crystal structures and MD simulations of carbohydrates	Standalone software	[280] ( <a href="http://glycam.org/docs/othertoolsservice/download-docs/publication-materials/bfmp/">http://glycam.org/docs/othertoolsservice/download-docs/publication-materials/bfmp/</a> )
Distance Mapping	Estimation of nuclear Overhauser effects in disaccharides	Web-tool	( <a href="http://www.glycosciences.de/modeling/distmap/">http://www.glycosciences.de/modeling/distmap/</a> )
MD2NOE	Calculation of Nuclear Overhauser effect build-up curves from long MD trajectories	Standalone software	[281] ( <a href="http://glycam.org/docs/othertoolsservice/download-docs/publication-materials/md2noe/">http://glycam.org/docs/othertoolsservice/download-docs/publication-materials/md2noe/</a> )
GS-align	Glycan structure alignment and similarity calculation	Standalone software	[282] ( <a href="http://www.glycanstructure.org/gsalign">http://www.glycanstructure.org/gsalign</a> )
GlyTorsion	Analysis of torsion angles in carbohydrates from Protein Data Bank	Web-tool	[283] ( <a href="http://www.glycosciences.de/tools/glytorsion/">http://www.glycosciences.de/tools/glytorsion/</a> )
GlyVicinity	Analysis of amino acids in the vicinity of carbohydrate residues derived from Protein Data Bank	Web-tool	[284] ( <a href="http://www.glycosciences.de/tools/glyvicinity/">http://www.glycosciences.de/tools/glyvicinity/</a> )

<sup>a</sup> Web-service implies an automated pipeline for running a specific software (e.g., molecular modeling, structure building, carbohydrate coordinate extraction, format conversion). It results in 3D structural data output starting from primary structure input or atomic coordinate file upload. Web-tool is employed for 3D structural data processing and analysis without 3D structural data output; it is a simpler application designed primarily for statistics and visualization. Other types are self-explanatory.

## 5. Experimental Data Validation

Vast variety of methods provide information about 3D structure of individual glycans and glycan moieties of glycoproteins and protein-carbohydrate complexes (Figure 6) [285,286]. The following approaches are most utilized for 3D structural data validation [287–289]:

- Combination of carbohydrate simulated geometry data with X-ray crystallographic data analysis [225,290];
- Analysis of inter-glycosidic NMR spin couplings, which depend on glycosidic bond torsions [114,291,292];
- Deriving nuclear Overhauser effects (NOEs) from relative populations of the interatomic distances, with subsequent comparison to the experimental NOEs in solution [99,293,294];
- Purely informatic detection of errors, such as incompatible atomic coordinates originating from incorrect processing or simulation [295–298];
- Simulation by other computational methods at higher levels of theory [102,103,105,108].

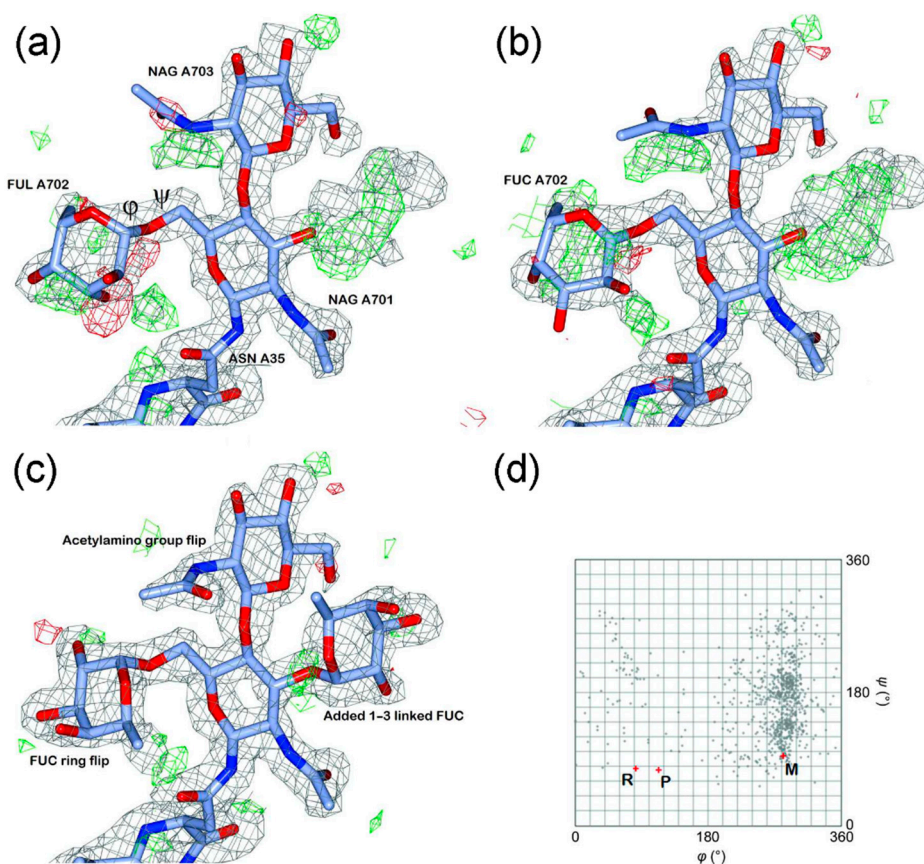


**Figure 6.** Interplay of the instrumental and computational methods in the 3D structure determination of carbohydrates, proteins, and protein-glycoconjugate complexes. Reproduced from [285] © 2020 The authors. Published by Wiley-VCH Verlag GmbH & Co. KGaA.

Unfortunately, most of the data obtained on the basis of crystallographic experiments can dramatically differ from glycan conformations in solution or have poor resolution which needs further adjustment [299,300]. Moreover, not all of the objects of interest can be obtained as a single crystal.

Electron cryo-microscopy gains popularity for carbohydrate 3D structural research [301], however, this method requires additional refinement procedures due to resolution restrictions of the obtained density maps [302–304]. Recently, cryo-EM data were used for the refinement of SARS-CoV-2 spike glycoprotein structure using Privateer (see Table 3 for references) software [305,306].

Van Beusekom et al., illustrated [295] quality improvement of the PDB glycan structure model with incorrect (1–6)-linked fucose annotation, poor fit to the electron density, and missing (1–3)-linked fucose (Figure 7a) with the help of PDB-REDO (Figure 7b) and CARP (Figure 7d) tools (see Table 3 for references). Structure model obtained by PDB-REDO treatment was further manually inspected (Figure 7c): corrections were made for acetylamino group geometry, distorted (1–6)-linked fucose ring conformation, and (1–3)-linked fucose residue was added. Despite successful automated resolution of residue annotation problem and poor electron density refinement, complete revision could not be achieved without manual intervention.

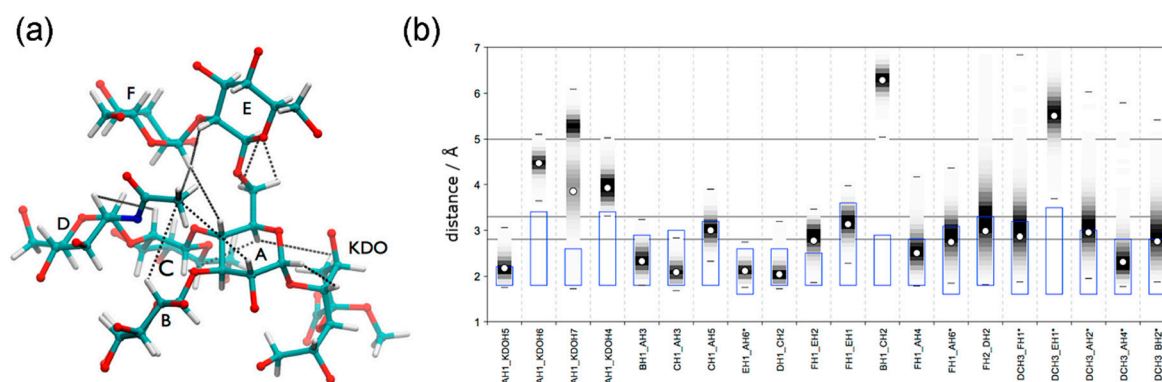


**Figure 7.** X-ray diffraction data refinement of N-glycan moiety from PDB ID 2Z62.  $2mF_o - DF_c$  electron density map contoured at  $1\sigma$  is displayed in grey; positive and negative  $mF_o - DF_c$  difference electron density maps contoured at  $3\sigma$  are displayed in green and red, respectively. (a) Original glycan structure model from the PDB entry. (b) PDB-REDO model with properly renamed fucose residue and improved fit to the electron density. (c) Manually rebuilt model based on PDB-REDO results. (d) CARP distribution plot for glycosidic  $\phi$ - $\psi$  torsions of FUC(1-6)NAG (from panel (a)) in PDB. Characteristic points: R, model refined with PDB-REDO; P, original PDB model; M, manually rebuilt model. Reproduced from [295], © 2020 The authors. Published by John Wiley & Sons, Inc.

NMR techniques are a powerful approach to investigate conformational and dynamic behavior of carbohydrate moieties in biomolecules [307–310]. However, the nature of NOE enhancement factor has been hampering obtaining the sufficient number of distance restraints [99]. In the case of saccharides with their multiple rotatable bonds, the stable 3D structure was difficult to define, making molecular modeling essential for this class of compounds. Adjustment of experimental conditions helped to

overcome the mentioned limitation and to reproduce crystal structures of oligosaccharides by modeling with NOE-derived distance restraints [100,101].

Since there is no direct way to derive detailed three-dimensional representation from the observed NOE intensities, additional molecular modeling protocols are required to establish comprehensive view of conformational space at the atomic level [311–313]. Frank et al., demonstrated conformation filtering based on the observed NOE obtained by molecular dynamics in explicit solvent [314]. As a representative example, Figure 8 depicts  $^1\text{H}$ - $^1\text{H}$  spatial contacts and conformation selection criteria illustrated by *Moraxella catarrhalis* lgt2 $\Delta$  bacterium heptasaccharide, which adopts an unusual conformation.



**Figure 8.** *M. catarrhalis* lgt2 $\Delta$  structure validation based on NOE data analysis. (a) Characteristic proton-proton contacts; (b) NOE-filtered (blue boxes) sampling of proton-proton distances from MD simulation (grey shades). Reproduced from [314], © 2020 The authors. Licensee MDPI, Basel, Switzerland.

## 6. Protein Data Bank and Its Validation

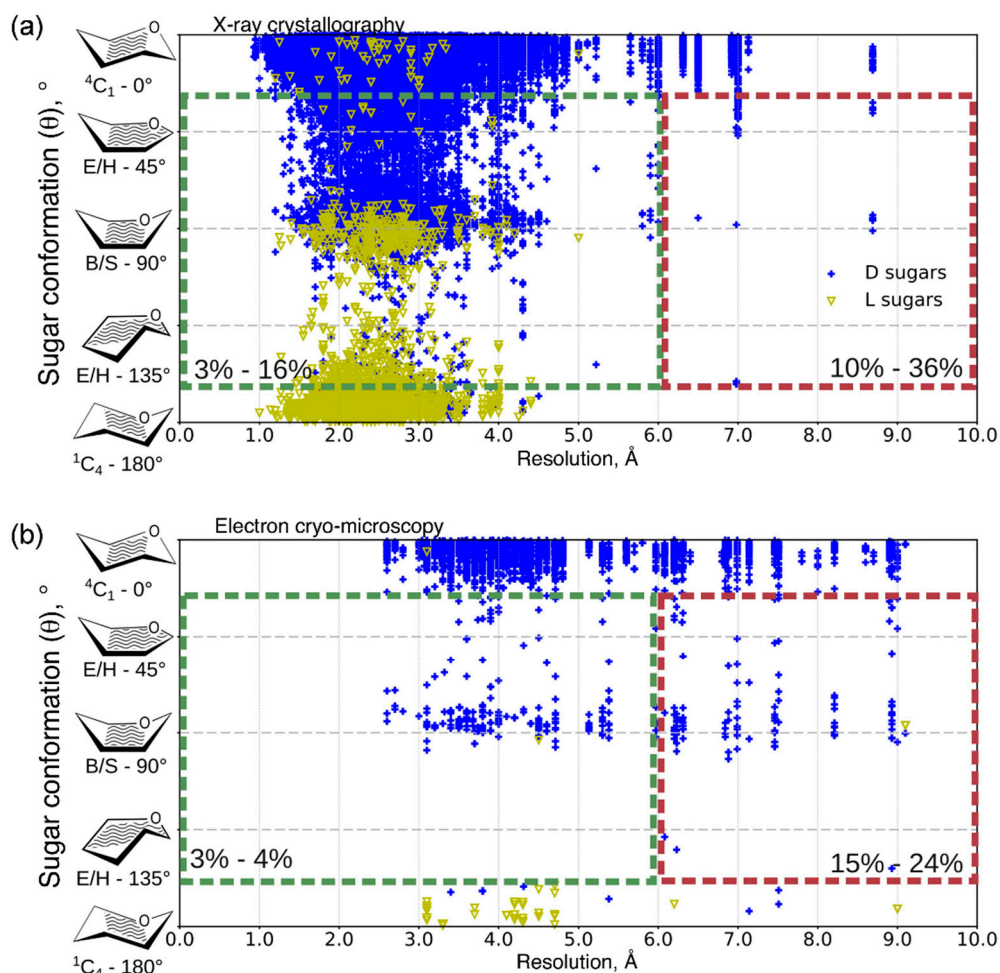
Protein Data Bank (PDB) [315] and Cambridge Structural Database (CSD) [316] are historically considered the main repositories of experimentally determined carbohydrate three-dimensional structures. CSD is reported to deposit over 4000 crystal structures of oligosaccharides [93]. Unlike Cambridge Structural Database, Protein Data Bank provides open access to the entire structural archive. Carbohydrate moieties deposited in PDB are usually represented as covalently bound to protein or imply non-covalently bound protein-carbohydrate complex formation [302]. According to recent reports, as at November 18, 2019 Protein Data Bank contained ~13500 carbohydrate structures representing ~9.4% of total database records [317].

Despite being a valuable source of 3D structural data for glycoscientists, PDB lacks convenient search facilities for glycan structures. Some projects have developed data-mining tools capable of retrieving bioglycan molecular geometry data from PDB: Glycan Reader ([GlycanStructure.org](http://www.glycanstructure.org)) [260,261] (<http://www.glycanstructure.org/>), pdb2linucs (GLYCOSCIENCES.de) [47,259,318] ([http://www.glycosciences.de/database/start.php?action=form\\_pdb\\_data](http://www.glycosciences.de/database/start.php?action=form_pdb_data)), GlycoNAVI TCarp [61] (<https://glyconavi.org/TCarp/>) (<https://gitlab.com/glyconavi/pdb2glycan>) and GlyFinder (GLYCAM-Web) [257,258] ([https://dev.glycam.org/portal/gf\\_home/](https://dev.glycam.org/portal/gf_home/)).

Another issue of concern related to Protein Data Bank is large proportion of errors in deposited coordinates, leading to requirement for a thorough checkup and development of data remediation services [319]. Commonly occurring problems associated with nomenclature, poor glycan geometry, linkage errors, missing or surplus atoms can seriously decline the quality of the obtained 3D structures [300,320,321]. Using Privateer software, it was discovered [299],[301] that PDB deposits significant number of erroneous N-glycosylated structures with pyranose ring distortions, considering preferred adoption of  $^4\text{C}_1$  conformation for D-sugars and  $^1\text{C}_4$  conformation for L-sugars (Figure 9). In most cases, poor electron density of carbohydrate moiety results in anomalous high-energy pyranose ring conformations (envelopes, half-chairs, boats, skew boats, etc.). To obtain a reasonable structure



model, experimental data refinement programs should be applied to derive geometric restraints for sugar monomers. Notably, despite a cryo-EM method has a resolution limit disadvantage, observed results indicate larger content of atypical conformations solved by X-ray crystallography, as compared to cryo-EM data.



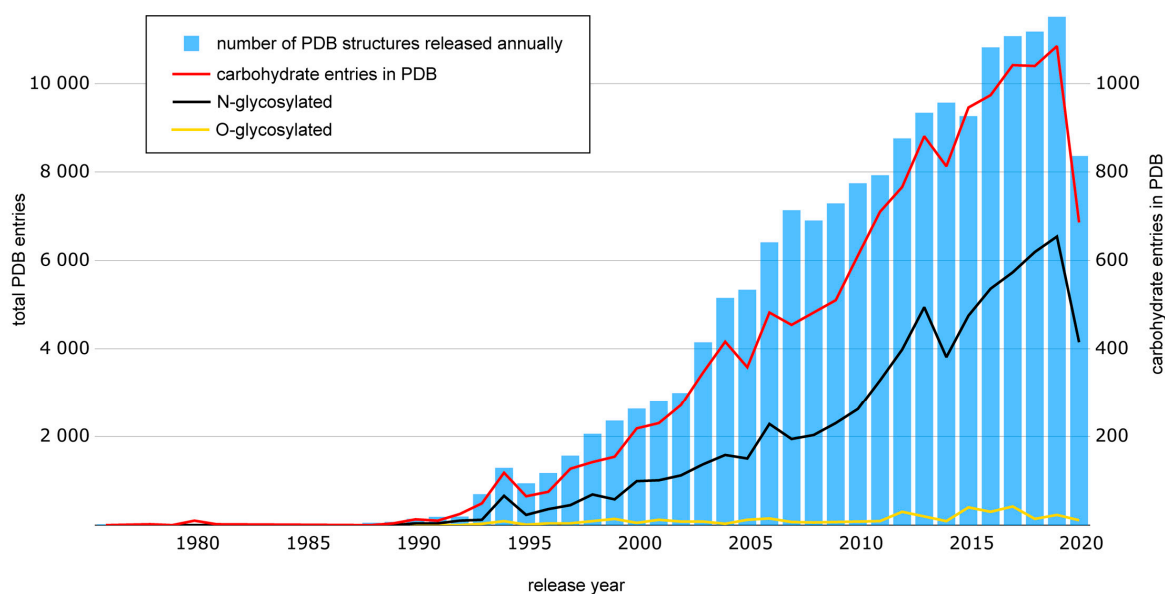
**Figure 9.** Distribution of D- (shown in blue) and L-pyranoside (shown in yellow) ring conformations as function of resolution for all sugar moieties in N-glycosylated proteins in PDB (on April 2019) solved with (a) X-ray crystallography and (b) electron cryo-microscopy. Non-chair conformations are bordered by dotted line boxes for 0.0-6.0 Å (green) and 6.0-10.0 Å (red) resolution ranges; the percentage of structures is given in the boxes. Reproduced with permission from [301], © 2020 Elsevier Ltd.

Exceptions for the relevancy of high-energy conformations were found in complexes involving carbohydrate-active enzymes, which force pyranose ring distortion enabling catalytic transformation of a carbohydrate substrate via transition states (e.g., glycosidic bond hydrolysis) [322]. Fushinobu has performed glycosidic torsion analysis for a set of PDB entries of crystal structure complexes bound to ligands bearing lacto-N-biose I (LNB, both  $\alpha$ - and  $\beta$ -anomers) disaccharide unit presented in type-1 antigens. The study was supported by GlycoMaps DB (see Table 1 for references) [323]. Obtained  $\varphi$ - $\psi$  data for LNBs bound to various proteins was plotted against corresponding free energy maps. Distortion of the energetically favored ring conformation strongly depended on substrate catalytic and recognition mechanisms.

To date, existing tools for carbohydrate structural error detection and correction in PDB files (Table 3) cannot be used directly as an integral part of Protein Data Bank. Nevertheless, initiative aimed at improvement of quality at wwPDB was carried out via collaboration with glycoscience community in July 2020 [324] (<https://www.wwpdb.org/documentation/carbohydrate-remediation>). It includes data annotation and validation of carbohydrate-containing records.

Proportion of carbohydrate-containing structures in PDB has been recently reported in [302]. Figure 10 presents our analysis of data published in the framework of Protein Data Bank carbohydrate remediation project. 14117 PDB entries from carbohydrate remediation list ([https://cdn.rcsb.org/wwpdb/docs/documentation/carbohydrateRemediation/PDB\\_carbohydrate\\_list.list](https://cdn.rcsb.org/wwpdb/docs/documentation/carbohydrateRemediation/PDB_carbohydrate_list.list)) were sorted by release year and plotted against the growth of PDB structures released annually (<https://www.rcsb.org/stats/growth/growth-released-structures>) (as on August 10, 2020; 167,327 PDB entries were available). Obtained results indicated that ~8.4% of PDB records contained a carbohydrate moiety. Additionally, each PDBx/mmCIF file corresponding to PDB ID from carbohydrate remediation list was parsed to reveal the presence of N- or O-glycosylation site annotations, which resulted in ~4.2% (7076 N-glycosylated entries) and 0.2% (362 O-glycosylated entries) of total database records. A few S- and C-glycans (24 entries in total) were neglected.

Statistics on glycans in Protein Data Bank was reported [259,302,317,325], as well as tools that could facilitate collection of statistical data (Glycan Reader [70,260,261], GlyFinder [258], pdb2linucs and pdb-care [326]).



**Figure 10.** Deposition statistics of carbohydrate-containing structures in Protein Data Bank based on carbohydrate remediated list data. Data for 2020 cover seven of twelve months. See detailed data in Supplementary Tables S3–S4.



**Table 3.** Tools for structural validation of carbohydrates.

Tool	Description	Type <sup>a</sup>	Reference
CNS	Macromolecular structure determination and refinement (including carbohydrates and glycoproteins) based on X-ray and NMR data	Standalone software	[327–330] ( <a href="http://cns-online.org/v1.3/">http://cns-online.org/v1.3/</a> )
pdb-care	Identification and assigning carbohydrate structures using atom types and coordinates from PDB files	Web-tool	[326] ( <a href="http://www.glycosciences.de/tools/pdb-care/">http://www.glycosciences.de/tools/pdb-care/</a> )
CARP	Glycoprotein 3D quality evaluation based on the analysis of glycosidic torsion angles from PDB	Web-tool	[283] ( <a href="http://www.glycosciences.de/tools/carp/">http://www.glycosciences.de/tools/carp/</a> )
GlyProbity	Accuracy and internal consistency check of carbohydrate 3D structures	Part of web-service pipeline	[257] ( <a href="https://dev.glycam.org/portal/gf_home/">https://dev.glycam.org/portal/gf_home/</a> )
PDB2Glycan	3D structure analysis and validation of glycoprotein PDB entries	Part of web-service pipeline	[61] ( <a href="https://glyconavi.org/TCarp/">https://glyconavi.org/TCarp/</a> ) ( <a href="https://gitlab.com/glyconavi/pdb2glycan">https://gitlab.com/glyconavi/pdb2glycan</a> )
PDB-REDO	Glycoprotein structure model improvement and validation	Web-service; standalone software	[295,325] ( <a href="https://pdb-redo.eu/">https://pdb-redo.eu/</a> )
Coot	Refinement and validation of glycoprotein 3D structure from cryoEM and X-ray crystallography data	Standalone software	[298,331] ( <a href="https://www2.mrc-lmb.cam.ac.uk/personal/pemsley/coot/">https://www2.mrc-lmb.cam.ac.uk/personal/pemsley/coot/</a> )
Rosetta Carbohydrate	Refinement of glycoprotein 3D structure from cryoEM and X-ray crystallography data, based on correction of conformational and configurational errors in carbohydrates	Python framework	[296] ( <a href="https://www.rosettacommons.org/docs/latest/application_documentation/carbohydrates/WorkingWithGlycans">https://www.rosettacommons.org/docs/latest/application_documentation/carbohydrates/WorkingWithGlycans</a> )
Privateer	Automated validation of carbohydrate conformation data based on 3D structure analysis	Standalone software	[297,332] ( <a href="https://smb.slac.stanford.edu/facilities/software/ccp4/html/privateer.html">https://smb.slac.stanford.edu/facilities/software/ccp4/html/privateer.html</a> )
Phenix	Determination, refinement and validation of macromolecular structure (including carbohydrates and glycoproteins) from cryoEM, X-ray diffraction and neutron diffraction crystallography data	Standalone software	[244] ( <a href="http://phenix-online.org/">http://phenix-online.org/</a> )
Motive Validator	Automatic custom residue validation in biomolecules, including carbohydrates	Web-service	[333] ( <a href="https://webchem.ncbr.muni.cz/Platform/MotiveValidator">https://webchem.ncbr.muni.cz/Platform/MotiveValidator</a> )
ValidatorDB	Pre-computed validation results of ligands and non-standard residues in PDB (including carbohydrates)	Web-service	[334] ( <a href="http://webchem.ncbr.muni.cz/Platform/ValidatorDb">http://webchem.ncbr.muni.cz/Platform/ValidatorDb</a> )

<sup>a</sup> See footnote <sup>a</sup> to Table 2.

## 7. 3D Structure Input and Visualization

Carbohydrate structure visualization in publications and computer interfaces is extremely important in terms of perception universality, unambiguity, and machine-readability. Hence, carbohydrate input [335–337] and visualization [338,339] tools are actively developing. Feature comparison of glycan sketchers, builders and viewers (occasionally including 3D ones) was reported in a recently published review [340]. In our review, we gave more emphasis to 3D visualization approaches.

Being informative to represent glycan primary structure, most of graphical input tools such as GlycanBuilder [341], DrawRINGS [342], SugarSketcher [343], DrawGlycan-SNFG [344,345] and GlycoGlyph [337] are inappropriate for obtaining 3D structural models and their visualization due to lack of underlying modeling and insufficient data conversion functionality.

At present, glycan 3D molecular models can be built in user-friendly software allowing constructing glycans from individual saccharide components. Free web-tools, such as GLYCAM-Web, CHARMM-GUI, POLYS glycan builder, GAG-builder, SWEET-II should be noted (more references are listed in Table 2). A few commercial molecular modeling software is equipped with special plugins for glycan 3D structure building based on a list of predefined monosaccharide templates, e.g., Sugar Builder tool in HyperChem (<http://www.hyper.com/?tabid=360>) software [346] or Azahar [235] plugin in PyMol package (Schrödinger software) (<https://pymol.org/2/>)[347].

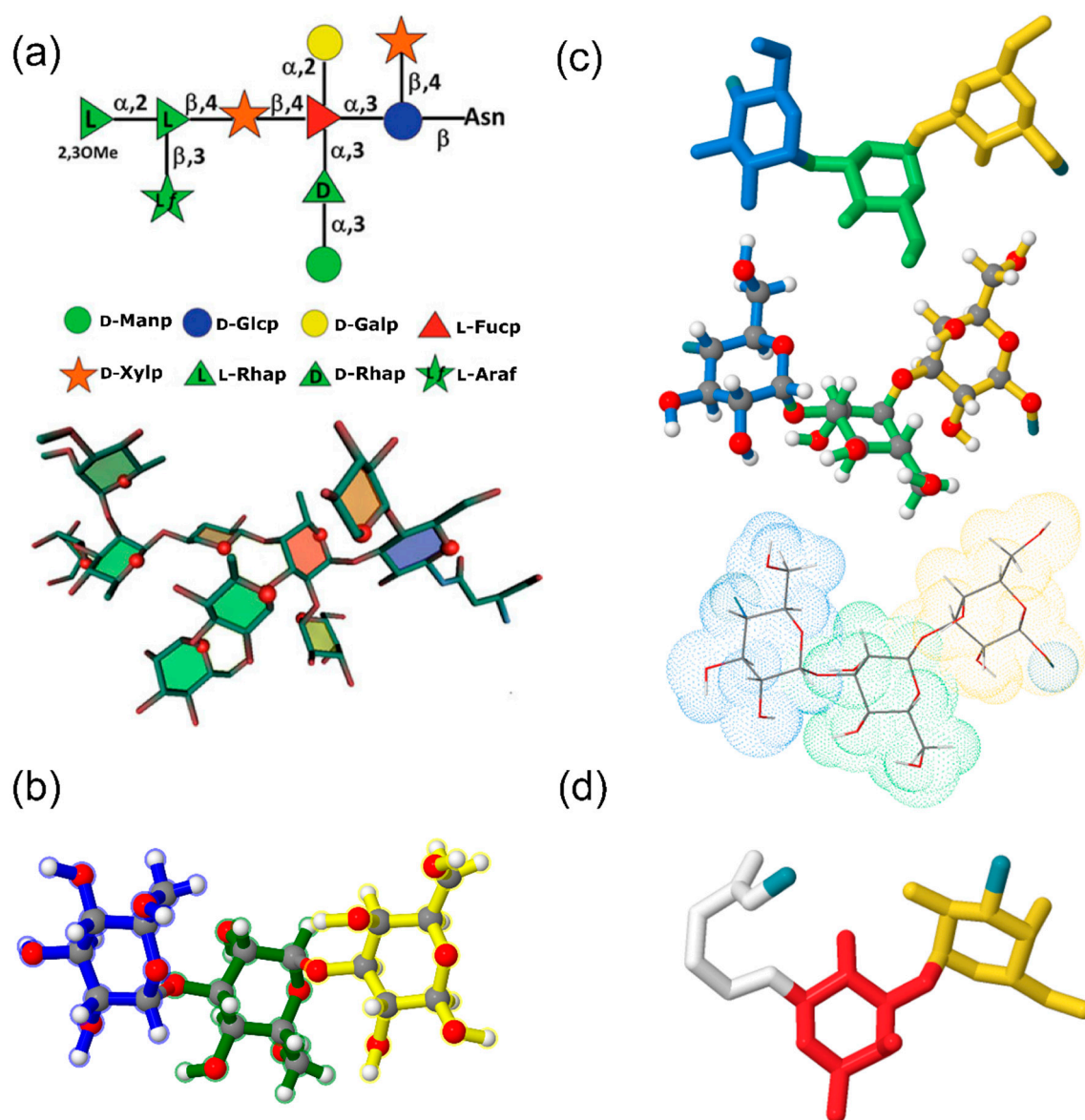
To render 3D glycan structure and its conformational features, it should be recorded using a notation which includes atomic coordinates, such as MOL [348] or PDB [349]. All-atom visualization based on atomic coordinates is supported by the majority of existing molecular modeling software. Several carbohydrate structure databases utilize interactive 3D visualization using open-source software engines. As one of the pioneers, GLYCOSCIENCES.de portal developed PDB2MultiGIF [350] (<http://www.glycosciences.de/modeling/pdb2mgif/>) visualization pipeline which generates an animated image of 3D model from a PDB file using RasMol [351] (<http://www.openrasmol.org/>). RasMol visualization was included in W3-SWEET [263] (ancestor of SWEET-II) pipeline developed by same project. Nowadays, more advanced interactive visualization applications have been developed for carbohydrate 3D molecule presentation. Jmol/JSmol [352] (<http://www.openrasmol.org/>) visualization applet is useful to display 3D models of carbohydrate molecules applied in numerous projects, such as CSDB, GLYCOSCIENCES.de, GLYCAM-Web and EK3D (see references in Table 1). NGL [353,354] (<http://nglviewer.org/>), LiteMol [355] (<https://www.litemol.org/>) and Mol\* [356] (<https://www.rcsb.org/news?year=2020&article=5efe0f606378d876901146f8>) (<https://molstar.org/>) 3D viewers are handy for processing macromolecular PDB data stored in glycoproteomics databases (UniLectin3D, Glycan Binding Site DB, ProCarbDB, GlycoNAVI, ProCaff, etc.; see references in Table 1) and general proteomics repositories such as PDB [315] (<http://www.wwpdb.org/>), UniProtKB [357] (<https://www.uniprot.org/>) or SWISS-MODEL [90] (<https://swissmodel.expasy.org/repository>).

NGL viewer was developed mainly for convenient protein macromolecule structure processing. It allows only ball-stick representation for small molecules or non-peptide fragments, such as saccharide residues. LiteMol (and its successor, Mol\*) viewer could be applied for the visualization of an arbitrary glycan with facility of highlighting carbohydrate fragments or displaying specific interactions in protein-carbohydrate complex structure. Due to these features, it was implemented in multiple carbohydrate structure databases (e.g., CSDB, Glyco3D, MatrixDB, and EPS-DB).

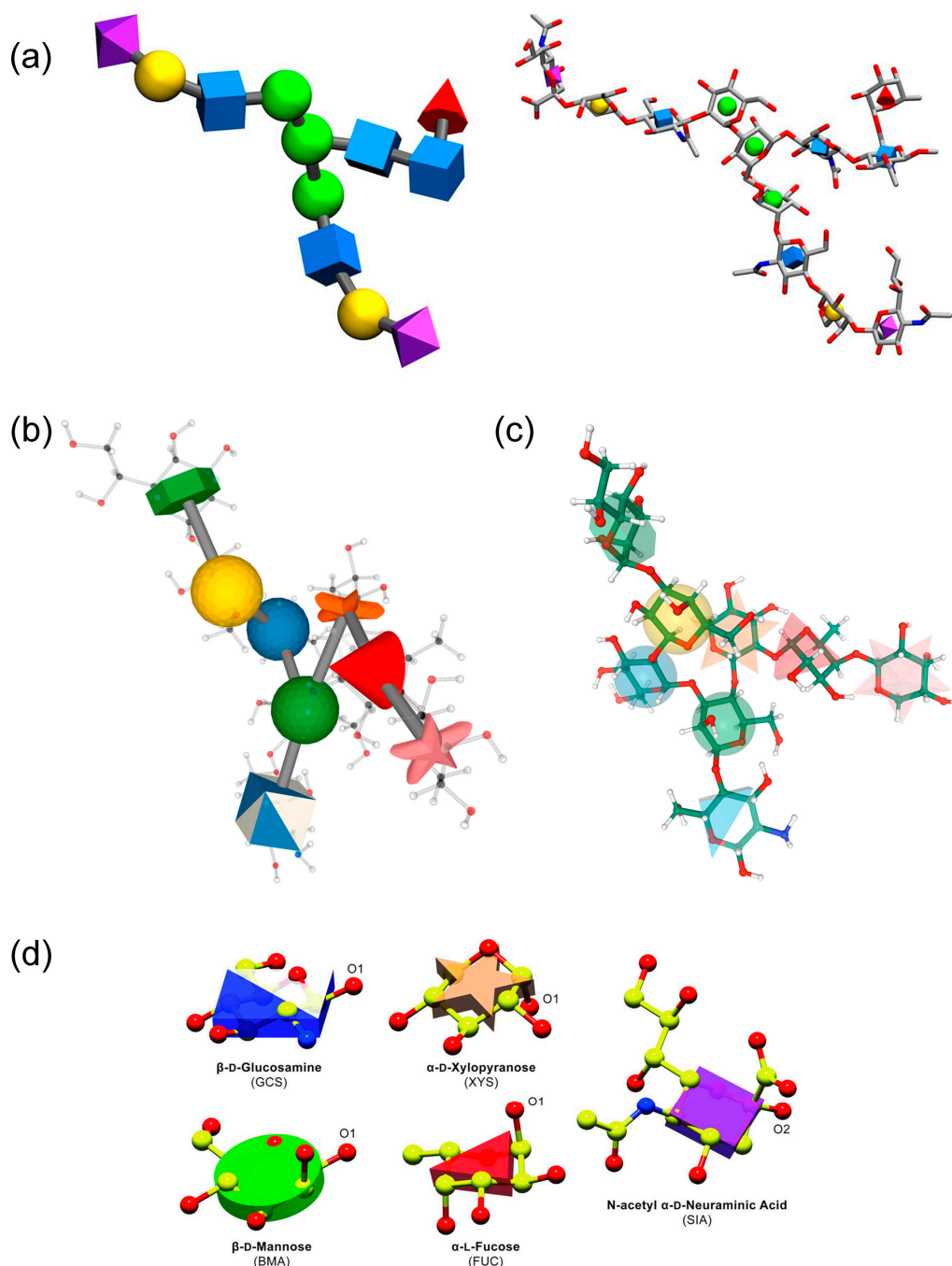
Despite the absence of the experimental 3D structural data, a number of carbohydrate databases have opportunity to simulate 3D atomic coordinates for deposited or inputted compounds from primary structure owing to tools developed by glycoinformatics community. CSDB (REStLESS API [265]), GLYCOSCIENCES.de (SWEET-II [264,350]) and GLYCAM-Web (<http://glycam.org/>) portals make it possible to generate 3D atomic coordinates recorded in PDB (all) and MOL (CSDB) file formats. POLYS developed by Glyco3D project enables the construction of polysaccharides in PDB format; it was introduced in MatrixDB and EPS-DB databases. More details are provided in Table 2.

Atomic coordinates and all-atom molecular models have not been popular in publications due to a lack of human readability. First attempts [358,359] of prof. Kuttel et al., to visualize carbohydrate molecules in an efficient and simple way were made by developing PaperChain and Twister graphic algorithms as a part of CarboHydra [360] and Visual Molecular Dynamics [361] software packages. Later, group of prof. Pérez suggested to restrict visualized molecule to skeletal atoms via conditional cycle plane coloring in accordance with the color code adopted in SNFG [338] visualization scheme (SweetUnityMol software [362], Figure 11a). Another UnityMol visualization approach called Umbrella Visualization [363,364] was tailored for N-glycan structures. Azahar plugin for PyMol [235] affords cartoon models with polygons and rods. Several solutions for convenient visualization came up with the development of SNFG notation [339]. Thus, group of prof. Woods proposed to combine molecular structure elements with 3D SNFG icons (Figure 12a). Such convenient visualization technique was integrated in LiteMol (Figure 12b) [365] and Mol\* (Figure 12c) [324,356]. 3D SNFG visualization plugins are available via Visual Molecular Dynamics platform [366] (<http://glycam.org/docs/othertoolsservice/2016/06/03/3d-symbol-nomenclature-for-glycans-3d-snfg/>) and UCSF Chimera [367] visualization software Tangram plugin ([https://github.com/insilichem/tangram\\_snfg](https://github.com/insilichem/tangram_snfg)). Designed as part of CCP4mg [368] molecular-graphics software, Glycoblcks [369] representation of monosaccharides uses shapes and colors, identical to those in SNFG (Figure 12d). Available as PyMol plugin developed by Widmalm group (<http://www.organ.su.se/gw/doku.php?id=3dcfg>), 3D-CFG representation [370] based on CFG notation [371] (often referred to as a predecessor of SNFG) should also be noted as earlier approach to interpretation of carbohydrate 3D structures based on a symbol library.

Considering efficiency and usability of 3D representation based on SNFG concept, which grows popular among glycoscientists, the development of alternative solutions in carbohydrate 3D structure representations has a potential for application in glycoinformatics projects. Support of colored residues in 3D structures implemented via JSmol on GLYCOSCIENCES.de portal was reported [47] (Figure 11b). Similarly, CSDB project has developed a 3D viewer ([http://csdb.glycoscience.ru/database/core/show\\_3d.php?csdb=-3\)aDManp\(1-3\)\[Ac\(1-2\)?DGlcpN\(1-6\)\]bDGal?\(1-](http://csdb.glycoscience.ru/database/core/show_3d.php?csdb=-3)aDManp(1-3)[Ac(1-2)?DGlcpN(1-6)]bDGal?(1-)) with carbohydrate residue coloring according to the SNFG notation in the framework of a modeling module based on RESTLESS API. In this tool, user can visualize input structure with help of sticks, balls and sticks, or van der Waals spheres (Figure 11c). Options for aglycone moiety (white) and pseudo-atoms (polymeric repeats, blue caps) are supported (Figure 11d).



**Figure 11.** Glycan structure colored according to SNFG, or superimposed with 3D SNFG, as implemented in SweetUnityMol (a), GLYCOSCIENCES.de (via JSmol) (b), and CSDB (via JSmol) (c,d), see text. Panel (a) was reproduced with permission from [372], © Springer Japan 2017.



**Figure 12.** Glycan structure colored according to SNFG, or superimposed with 3D SNFG, as implemented in 3D-SNFG (a), LiteMol (b), Mol\* (c); monosaccharide presentation in Glycoblocks (d). Panel (a) was reproduced with permission from [366], © 2020, Oxford University Press. Panel (d) was reproduced from [369], © 2020 The authors. Published by John Wiley & Sons, Inc.

## 8. Conclusions

Development of glycoinformatics resources makes great impact on treating enormous masses of data sets produced by glyco-related research. Tools for carbohydrate 3D structural information retrieval provide a framework for experimental and computational data quality validation. Data sources based on conformational ensemble generation and analysis assist structure–function and structure–activity relationship prediction of biologically relevant bioglycans and glycoconjugates. In this review, we have summarized existing facilities on working with glycan spatial features that can provide harmonious

network of structural databases, web-services, tools and standalone software for modeling and processing structural data. Further advances in this field will help building better understanding of glycan participation in biological processes and supply glycoscience community with user-friendly access to voluminous data collections.

**Supplementary Materials:** Supplementary Materials can be found at <http://www.mdpi.com/1422-0067/21/20/7702/s1>. Supplementary Tables: S1. Application rate for quantum mechanics and molecular mechanics methods in 2015–2020. S2. Citing of force fields in 2015–2020. S3. Protein Data Bank statistics. S4. PDB carbohydrate remediation list data.

**Author Contributions:** S.I.S. and P.V.T. contributed equally. All authors have read and agreed to the published version of the manuscript.

**Funding:** The work with carbohydrate molecular modeling and PDB data was funded by Russian Foundation for Basic Research grant 18-04-00094. The work with structural databases, glycoinformatic tools and visualization was funded by Russian Science Foundation grant 18-14-00098.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Abbreviations

3D	Three-dimensional
AA	All-atom
CAZy	Carbohydrate-Active Enzyme
CD	Cluster of Differentiation
CFG	Consortium for Functional Glycomics
CG	Coarse-grained
CHI	Carbohydrate Intrinsic
CRD	Carbohydrate Recognition Site
Cryo-EM	Electron cryo-microscopy
CSD	Cambridge Structural Database
DFT	Density Functional Theory
FUC	$\alpha$ -L-fucopyranose
GAG	Glycosaminoglycan
GAMD	Gaussian Accelerated MD
GBP	Glycan-Binding Protein
GM9	$\text{Glc}_1\text{Man}_9\text{GlcNAc}_2$
HPLC	High Performance Liquid Chromatography
HREX	Hamiltonian Replica-Exchange MD
INIOM	Our own N-layered integrated molecular orbital and molecular mechanics
LNB	Lacto-N-biose I
MD	Molecular Dynamics
MM	Molecular Mechanics
MS	Mass-spectrometry
mseMD	Multidimensional swarm-enhanced sampling MD
NAG	2-acetamido-2-deoxy- $\beta$ -D-glucopyranose
NMR	Nuclear Magnetic Resonance
NOE	Nuclear Overhauser Effect
PDB	Protein Data Bank
PDBe	Protein Data Bank Europe
PDBj	Protein Data Bank Japan
PDBsum	Database of Structural Summaries of PDB Entries
QM	Quantum Mechanics
RCSB PDB	Research Collaboratory for Structural Bioinformatics Protein Data Bank
REMD	Replica-exchange MD
SNFG	Symbol Nomenclature for Glycans
UniProtKB	UniProt Knowledgebase
wwPDB	Worldwide Protein Data Bank



## References

1. Hricovini, M. Structural Aspects of Carbohydrates and the Relation with their Biological Properties. *Curr. Med. Chem.* **2004**, *11*, 2565–2583. [[CrossRef](#)] [[PubMed](#)]
2. Varki, A. Biological roles of glycans. *Glycobiology* **2017**, *27*, 3–49. [[CrossRef](#)]
3. Buddhadeb, M.; Dimitrios, M. Applications of Molecular Dynamics Simulations in Immunology: A Useful Computational Method in Aiding Vaccine Design. *Curr. Proteom.* **2006**, *3*, 259–270.
4. Kuttel, M.M.; Ravenscroft, N. The Role of Molecular Modeling in Predicting Carbohydrate Antigen Conformation and Understanding Vaccine Immunogenicity. In *Carbohydrate-Based Vaccines: From Concept to Clinic*; American Chemical Society: Washington, DC, USA, 2018; Volume 1290, pp. 139–173.
5. Mishra, S.K.; Calabró, G.; Loeffler, H.H.; Michel, J.; Koča, J. Evaluation of Selected Classical Force Fields for Alchemical Binding Free Energy Calculations of Protein-Carbohydrate Complexes. *J. Chem. Theory Comput.* **2015**, *11*, 3333–3345. [[CrossRef](#)] [[PubMed](#)]
6. Woods, R.J.; Tessier, M.B. Computational glycoscience: Characterizing the spatial and temporal properties of glycans and glycan–protein complexes. *Curr. Opin. Struct. Biol.* **2010**, *20*, 575–583. [[CrossRef](#)]
7. Hadden, J.A.; Tessier, M.B.; Fadda, E.; Woods, R.J. Calculating Binding Free Energies for Protein–Carbohydrate Complexes. In *Glycoinformatics*; Lütteke, T., Frank, M., Eds.; Humana Press: New York, NY, USA, 2015; pp. 431–465.
8. Woods, R.J. Predicting the Structures of Glycans, Glycoproteins, and Their Complexes. *Chem. Rev.* **2018**, *118*, 8005–8024. [[CrossRef](#)]
9. Frank, M. Computational Docking as a Tool for the Rational Design of Carbohydrate-Based Drugs. In *Carbohydrates as Drugs*; Seeberger, P.H., Rademacher, C., Eds.; Springer: Cham, Switzerland, 2014; pp. 53–72.
10. Purcell, S.C.; Godula, K. Synthetic glycoscapes: Addressing the structural and functional complexity of the glycocalyx. *Interface Focus* **2019**, *9*, 20180080. [[CrossRef](#)]
11. Nagae, M.; Yamaguchi, Y. Function and 3D Structure of the N-Glycans on Glycoproteins. *Int. J. Mol. Sci.* **2012**, *13*, 8398–8429. [[CrossRef](#)]
12. Aoki-Kinoshita, K.F.; Lisacek, F.; Mazumder, R.; York, W.S.; Packer, N.H. The GlySpace Alliance: Toward a collaborative global glycoinformatics community. *Glycobiology* **2020**, *30*, 70–71. [[CrossRef](#)]
13. Yamada, I.; Shiota, M.; Shinmachi, D.; Ono, T.; Tsuchiya, S.; Hosoda, M.; Fujita, A.; Aoki, N.P.; Watanabe, Y.; Fujita, N.; et al. The GlyCosmos Portal: A unified and comprehensive web resource for the glycosciences. *Nat. Methods* **2020**, *17*, 649–650. [[CrossRef](#)]
14. Mariethoz, J.; Alocci, D.; Gastaldello, A.; Horlacher, O.; Gasteiger, E.; Rojas-Macias, M.; Karlsson, N.G.; Packer, N.H.; Lisacek, F. Glycomics@ExpASY: Bridging the Gap. *Mol. Cell. Proteom.* **2018**, *17*, 2164. [[CrossRef](#)] [[PubMed](#)]
15. Khasay, R.; Vora, J.; Navelkar, R.; Mousavi, R.; Fochtman, B.; Holmes, X.; Pattabiraman, N.; Ranzinger, R.; Mahadik, R.; Williamson, T.; et al. GlyGen data model and processing workflow. *Bioinformatics* **2020**, *36*, 3941–3943. [[CrossRef](#)]
16. Maeda, M.; Fujita, N.; Suzuki, Y.; Sawaki, H.; Shikanai, T.; Narimatsu, H. JCGGDB: Japan Consortium for Glycobiology and Glycotechnology Database. In *Glycoinformatics*; Lütteke, T., Frank, M., Eds.; Humana Press: New York, NY, USA, 2015; pp. 161–179.
17. Tiemeyer, M.; Aoki, K.; Paulson, J.; Cummings, R.D.; York, W.S.; Karlsson, N.G.; Lisacek, F.; Packer, N.H.; Campbell, M.P.; Aoki, N.P.; et al. GlyTouCan: An accessible glycan structure repository. *Glycobiology* **2017**, *27*, 915–919. [[CrossRef](#)]
18. York, W.S.; Agravat, S.; Aoki-Kinoshita, K.F.; McBride, R.; Campbell, M.P.; Costello, C.E.; Dell, A.; Feizi, T.; Haslam, S.M.; Karlsson, N.; et al. MIRAGE: The minimum information required for a glycomics experiment. *Glycobiology* **2014**, *24*, 402–406. [[CrossRef](#)] [[PubMed](#)]
19. Comelli, E.M.; Head, S.R.; Gilmartin, T.; Whisenant, T.; Haslam, S.M.; North, S.J.; Wong, N.-K.; Kudo, T.; Narimatsu, H.; Esko, J.D.; et al. A focused microarray approach to functional glycomics: Transcriptional regulation of the glycome. *Glycobiology* **2006**, *16*, 117–131. [[CrossRef](#)]
20. Aoki-Kinoshita, K.F. RINGS: A Web Resource of Tools for Analyzing Glycomics Data. In *A Practical Guide to Using Glycomics Databases*; Aoki-Kinoshita, K.F., Ed.; Springer: Tokyo, Japan, 2017; pp. 299–334.

21. Gouridine, J.-P.F.; Brush, M.H.; Vasilevsky, N.A.; Shefchek, K.; Köhler, S.; Matentzoglou, N.; Munoz-Torres, M.C.; McMurry, J.A.; Zhang, X.A.; Robinson, P.N.; et al. Representing glycophenotypes: Semantic unification of glycobiology resources for disease discovery. *Database* **2019**, 2019. [[CrossRef](#)]
22. Alocci, D.; Ghraichy, M.; Barletta, E.; Gastaldello, A.; Mariethoz, J.; Lisacek, F. Understanding the glycome: An interactive view of glycosylation from glyco compositions to glycoepitopes. *Glycobiology* **2018**, *28*, 349–362. [[CrossRef](#)] [[PubMed](#)]
23. Barnett, C.B.; Aoki-Kinoshita, K.F.; Naidoo, K.J. The Glycome Analytics Platform: An integrative framework for glycoinformatics. *Bioinformatics* **2016**, *32*, 3005–3011. [[CrossRef](#)]
24. Reily, C.; Stewart, T.J.; Renfrow, M.B.; Novak, J. Glycosylation in health and disease. *Nat. Rev. Nephrol.* **2019**, *15*, 346–366. [[CrossRef](#)] [[PubMed](#)]
25. Katoh, T.; Yamamoto, K. Glycoenzymes in Glycan Analysis and Synthesis. In *Glycoscience: Biology and Medicine*; Taniguchi, N., Endo, T., Hart, G.W., Seeberger, P.H., Wong, C.-H., Eds.; Springer: Tokyo, Japan, 2015; pp. 379–389.
26. Montgomery, A.P.; Xiao, K.; Wang, X.; Skropeta, D.; Yu, H. Chapter Two—Computational Glycobiology: Mechanistic Studies of Carbohydrate-Active Enzymes and Implication for Inhibitor Design. In *Advances in Protein Chemistry and Structural Biology*; Karabencheva-Christova, T., Ed.; Academic Press: Cambridge, MA, USA, 2017; Volume 109, pp. 25–76.
27. Copoiu, L.; Malhotra, S. The current structural glycome landscape and emerging technologies. *Curr. Opin. Struct. Biol.* **2020**, *62*, 132–139. [[CrossRef](#)]
28. Toukach, P.V.; Egorova, K.S. Carbohydrate structure database merged from bacterial, archaeal, plant and fungal parts. *Nucleic Acids Res.* **2016**, *44*, D1229–D1236. [[CrossRef](#)]
29. Egorova, K.S.; Toukach, P.V. Glycoinformatics: Bridging isolated islands in the sea of data. *Angew. Chem. Int. Ed.* **2018**, *57*, 14986–14990. [[CrossRef](#)] [[PubMed](#)]
30. Lisacek, F.; Mariethoz, J.; Alocci, D.; Rudd, P.M.; Abrahams, J.L.; Campbell, M.P.; Packer, N.H.; Stähle, J.; Widmalm, G.; Mullen, E.; et al. Databases and Associated Tools for Glycomics and Glycoproteomics. In *High-Throughput Glycomics and Glycoproteomics: Methods and Protocols*; Lauc, G., Wührer, M., Eds.; Humana Press: New York, NY, USA, 2017; pp. 235–264.
31. Abrahams, J.L.; Taherzadeh, G.; Jarvas, G.; Guttman, A.; Zhou, Y.; Campbell, M.P. Recent advances in glycoinformatic platforms for glycomics and glycoproteomics. *Curr. Opin. Struct. Biol.* **2020**, *62*, 56–69. [[CrossRef](#)] [[PubMed](#)]
32. Li, X.; Xu, Z.; Hong, X.; Zhang, Y.; Zou, X. Databases and Bioinformatic Tools for Glycobiology and Glycoproteomics. *Int. J. Mol. Sci.* **2020**, *21*, 6727. [[CrossRef](#)] [[PubMed](#)]
33. Campbell, M.P.; Peterson, R.; Mariethoz, J.; Gasteiger, E.; Akune, Y.; Aoki-Kinoshita, K.F.; Lisacek, F.; Packer, N.H. UniCarbKB: Building a knowledge platform for glycoproteomics. *Nucleic Acids Res.* **2014**, *42*, D215–D221. [[CrossRef](#)]
34. Doubet, S.; Bock, K.; Smith, D.; Darvill, A.; Albersheim, P. The Complex Carbohydrate Structure Database. *Trends Biochem. Sci.* **1989**, *14*, 475–477. [[CrossRef](#)]
35. Doubet, S.; Albersheim, P. CarbBank. *Glycobiology* **1992**, *2*, 505–507. [[CrossRef](#)]
36. von der Lieth, C.W.; Freire, A.A.; Blank, D.; Campbell, M.P.; Ceroni, A.; Damerell, D.R.; Dell, A.; Dwek, R.A.; Ernst, B.; Fogh, R.; et al. EUROCarbDB: An open-access platform for glycoinformatics. *Glycobiology* **2011**, *21*, 493–502. [[CrossRef](#)]
37. Ranzinger, R.; Herget, S.; Wetter, T.; von der Lieth, C.W. GlycomeDB—Integration of open-access carbohydrate structure databases. *BMC Bioinf.* **2008**, *9*, 384. [[CrossRef](#)]
38. Ranzinger, R.; Frank, M.; von der Lieth, C.W.; Herget, S. Glycome-DB.org: A portal for querying across the digital world of carbohydrate sequences. *Glycobiology* **2009**, *19*, 1563–1567. [[CrossRef](#)]
39. Nakahara, T.; Hashimoto, R.; Nakagawa, H.; Monde, K.; Miura, N.; Nishimura, S. Glycoconjugate Data Bank: Structures—An annotated glycan structure database and N-glycan primary structure verification service. *Nucleic Acids Res.* **2008**, *36*, D368–D371. [[CrossRef](#)] [[PubMed](#)]
40. Cooper, C.A.; Harrison, M.J.; Wilkins, M.R.; Packer, N.H. GlycoSuiteDB: A new curated relational database of glycoprotein glycan structures and their biological sources. *Nucleic Acids Res.* **2001**, *29*, 332–335. [[CrossRef](#)] [[PubMed](#)]

41. Cooper, C.A.; Joshi, H.J.; Harrison, M.J.; Wilkins, M.R.; Packer, N.H. GlycoSuiteDB: A curated relational database of glycoprotein glycan structures and their biological sources. 2003 update. *Nucleic Acids Res.* **2003**, *31*, 511–513. [[CrossRef](#)]
42. Toukach, P.V. Bacterial carbohydrate structure database 3: Principles and realization. *J. Chem. Inf. Model.* **2011**, *51*, 159–170. [[CrossRef](#)] [[PubMed](#)]
43. Egorova, K.S.; Toukach, P.V. CSDB\_GT: A new curated database on glycosyltransferases. *Glycobiology* **2017**, *27*, 285–290. [[CrossRef](#)]
44. Toukach, P.V.; Knirel, Y.A. New database of bacterial carbohydrate structures. *Glycoconj. J.* **2005**, *22*, 216–217.
45. Lütteke, T.; Bohne-Lang, A.; Loss, A.; Goetz, T.; Frank, M.; Von Der Lieth, C.W. GLYCOSCIENCES.de: An Internet portal to support glycomics and glycobiology research. *Glycobiology* **2006**, *16*, 71R–81R. [[CrossRef](#)]
46. Lütteke, T. Glycan data retrieval and analysis using GLYCOSCIENCES.de applications. In *A Practical Guide to Using Glycomics Databases*, 1st ed.; Aoki-Kinoshita, K.F., Ed.; Springer: Tokyo, Japan, 2017; pp. 335–350.
47. Bohm, M.; Bohne-Lang, A.; Frank, M.; Loss, A.; Rojas-Macias, M.A.; Lutteke, T. Glycosciences.DB: An annotated data collection linking glycomics and proteomics data (2018 update). *Nucleic Acids Res.* **2019**, *47*, D1195–D1201. [[CrossRef](#)]
48. Pérez, S.; Sarkar, A.; Rivet, A.; Breton, C.; Imberty, A. Glyco3D: A Portal for Structural Glycosciences. In *Glycoinformatics*; Lütteke, T., Frank, M., Eds.; Humana Press: New York, NY, USA, 2015; pp. 241–258.
49. Pérez, S.; Sarkar, A.; Rivet, A.; Drouillard, S.; Breton, C.; Imberty, A. Glyco3D: A Suite of Interlinked Databases of 3D Structures of Complex Carbohydrates, Lectins, Antibodies, and Glycosyltransferases. In *A Practical Guide to Using Glycomics Databases*; Aoki-Kinoshita, K.F., Ed.; Springer: Tokyo, Japan, 2017; pp. 133–161.
50. Sarkar, A.; Pérez, S. PolySac3DB: An annotated data base of 3 dimensional structures of polysaccharides. *BMC Bioinf.* **2012**, *13*, 302. [[CrossRef](#)]
51. Kunduru, B.R.; Nair, S.A.; Rathinavelan, T. EK3D: An *E. coli* K antigen 3-dimensional structure database. *Nucleic Acids Res.* **2016**, *44*, D675–D681. [[CrossRef](#)]
52. Veluraja, K.; Selvin, J.F.A.; Venkateshwari, S.; Priyadarzini, T.R.K. 3DSDSCAR—A three dimensional structural database for sialic acid-containing carbohydrates through molecular dynamics simulation. *Carbohydr. Res.* **2010**, *345*, 2030–2037. [[CrossRef](#)] [[PubMed](#)]
53. Veluraja, K.; Fermin Angelo Selvin, J.; Jasmine, A.; Hema Thanka Christlet, T. Three Dimensional Structures of Carbohydrates and Glycoinformatics: An Overview. In *Current Trends in Bioinformatics: An Insight*; Wadhwa, G., Shanmughavel, P., Singh, A.K., Bellare, J.R., Eds.; Springer: Singapore, 2018; pp. 55–87.
54. Chautard, E.; Fatoux-Ardore, M.; Ballut, L.; Thierry-Mieg, N.; Ricard-Blum, S. MatrixDB, the extracellular matrix interaction database. *Nucleic Acids Res.* **2011**, *39*, D235–D240. [[CrossRef](#)] [[PubMed](#)]
55. Launay, G.; Salza, R.; Multedo, D.; Thierry-Mieg, N.; Ricard-Blum, S. MatrixDB, the extracellular matrix interaction database: Updated content, a new navigator and expanded functionalities. *Nucleic Acids Res.* **2015**, *43*, D321–D327. [[CrossRef](#)]
56. Clerc, O.; Deniaud, M.; Vallet, S.D.; Naba, A.; Rivet, A.; Perez, S.; Thierry-Mieg, N.; Ricard-Blum, S. MatrixDB: Integration of new data with a focus on glycosaminoglycan interactions. *Nucleic Acids Res.* **2019**, *47*, D376–D381. [[CrossRef](#)]
57. Birch, J.; Van Calsteren, M.R.; Perez, S.; Svensson, B. The exopolysaccharide properties and structures database: EPS-DB. Application to bacterial exopolysaccharides. *Carbohydr. Polym.* **2019**, *205*, 565–570. [[CrossRef](#)]
58. Cao, Y.; Park, S.-J.; Mehta, A.Y.; Cummings, R.D.; Im, W. GlyMDB: Glycan Microarray Database and analysis toolset. *Bioinformatics* **2020**, *36*, 2438–2442. [[CrossRef](#)]
59. Raman, R.; Venkataraman, M.; Ramakrishnan, S.; Lang, W.; Raguram, S.; Sasisekharan, R. Advancing glycomics: Implementation strategies at the consortium for functional glycomics. *Glycobiology* **2006**, *16*, 82R–90R. [[CrossRef](#)] [[PubMed](#)]
60. Venkataraman, M.; Sasisekharan, R.; Raman, R. Glycan Array Data Management at Consortium for Functional Glycomics. In *Glycoinformatics*, 1st ed.; Lütteke, T., Frank, M., Eds.; Humana Press: New York, NY, USA, 2015; pp. 181–190.
61. Yamada, I.; Angata, K.; Watanabe, Y.; Ono, T. Databases for glycoconjugates (GlyCosmos Glycoproteins and Glycolipids, GlycoProtDB, GlycoNAVI:TCarp, GlycoPOST). *Glycoforum* **2020**, *23*, A2.
62. Yamada, I.; Aoki-Kinoshita, K.F. Integration of Glycoscience Data in GlyCosmos Using Semantic Web Technologies. In *Program and Abstracts for 2017 Annual Meeting of the Society for Glycobiology, Portland, OR, USA, 5–8 November 2017*; Oxford University Press Inc.: Cary, NC, USA, 2017; p. 61.

63. Shiota, M.; Tsuchiya, S.; Ono, T.; Kuoka, T.; Miura, N.; Hiraki, A.; Yamada, I.; Shinmachi, D.; Aoki, N.P.; Kim, J.-D. *The GlyCosmos Web Portal: Glycan structures, glycozymes, glycoproteins, pathways, diseases and more!* In Program and Abstracts for 2018 Annual Meeting of the Society for Glycobiology, New Orleans, LA, USA, 5–8 November 2018; Oxford University Press Inc.: Cary, NC, USA, 2018; pp. 1070–1071.
64. Mariethoz, J.; Khatib, K.; Alocci, D.; Campbell, M.P.; Karlsson, N.G.; Packer, N.H.; Mullen, E.H.; Lisacek, F. SugarBindDB, a resource of glycan-mediated host–pathogen interactions. *Nucleic Acids Res.* **2016**, *44*, D1243–D1250. [[CrossRef](#)]
65. Alocci, D.; Mariethoz, J.; Gastaldello, A.; Gasteiger, E.; Karlsson, N.G.; Kolarich, D.; Packer, N.H.; Lisacek, F. GlyConnect: Glycoproteomics Goes Visual, Interactive, and Analytical. *J. Proteome Res.* **2019**, *18*, 664–677. [[CrossRef](#)]
66. Choudhary, P.; Nagar, R.; Singh, V.; Bhat, A.H.; Sharma, Y.; Rao, A. ProGlycProt V2.0, a repository of experimentally validated glycoproteins and protein glycosyltransferases of prokaryotes. *Glycobiology* **2019**, *29*, 461–468. [[CrossRef](#)]
67. Bhat, A.H.; Mondal, H.; Chauhan, J.S.; Raghava, G.P.S.; Methi, A.; Rao, A. ProGlycProt: A repository of experimentally characterized prokaryotic glycoproteins. *Nucleic Acids Res.* **2012**, *40*, D388–D393. [[CrossRef](#)] [[PubMed](#)]
68. Copoiu, L.; Torres, P.H.M.; Ascher, D.B.; Blundell, T.L.; Malhotra, S. ProCarbDB: A database of carbohydrate-binding proteins. *Nucleic Acids Res.* **2020**, *48*, D368–D375. [[CrossRef](#)] [[PubMed](#)]
69. Siva Shanmugam, N.R.; Jino Blessy, J.; Veluraja, K.; Michael Gromiha, M. ProCaff: Protein–carbohydrate complex binding affinity database. *Bioinformatics* **2020**, *36*, 3615–3617. [[CrossRef](#)] [[PubMed](#)]
70. Cao, Y.; Park, S.-J.; Im, W. A Systematic Analysis of Protein–Carbohydrate Interactions in the PDB. *Glycobiology* **2020**. [[CrossRef](#)]
71. Malik, A.; Firoz, A.; Jha, V.; Ahmad, S. PROCARB: A Database of Known and Modelled Carbohydrate-Binding Protein Structures with Sequence-Based Prediction Tools. *Adv. Bioinf.* **2010**, *2010*, 1–9. [[CrossRef](#)]
72. Bonnardel, F.; Mariethoz, J.; Salentin, S.; Robin, X.; Schroeder, M.; Perez, S.; Lisacek, F.; Imberty, A. UniLectin3D, a database of carbohydrate binding proteins with curated information on 3D structures and interacting ligands. *Nucleic Acids Res.* **2019**, *47*, D1236–D1244. [[CrossRef](#)] [[PubMed](#)]
73. Bonnardel, F.; Perez, S.; Lisacek, F.; Imberty, A. Structural Database for Lectins and the UniLectin Web Platform. In *Lectin Purification and Analysis: Methods and Protocols*; Hirabayashi, J., Ed.; Humana: New York, NY, USA, 2020; pp. 1–14.
74. Hirabayashi, J.; Tateno, H.; Shikanai, T.; Aoki-Kinoshita, K.F.; Narimatsu, H. The Lectin Frontier Database (LfDB), and data generation based on frontal affinity chromatography. *Molecules* **2015**, *20*, 951–973. [[CrossRef](#)]
75. Chandra, N.R.; Kumar, N.; Jeyakani, J.; Singh, D.D.; Gowda, S.B.; Prathima, M.N. Lectindb: A plant lectin database. *Glycobiology* **2006**, *16*, 938–946. [[CrossRef](#)]
76. Kawasaki, T.; Nakao, H.; Takahashi, E.; Tominaga, T. GlycoEpitope: The integrated database of carbohydrate antigens and antibodies. *Trends Glycosci. Glycotechnol.* **2006**, *18*, 267–272. [[CrossRef](#)]
77. Kawasaki, T.; Nakao, H.; Tominaga, T. GlycoEpitope: A database of carbohydrate epitopes and antibodies. In *Experimental Glycoscience*; Taniguchi, N., Suzuki, A., Ito, Y., Narimatsu, H., Kawasaki, T., Hase, S., Eds.; Springer: Tokyo, Japan, 2008; pp. 429–431.
78. Aoki-Kinoshita, K.F. Using Databases and Web Resources for Glycomics Research. *Mol. Cell. Proteom.* **2013**, *12*, 1036. [[CrossRef](#)] [[PubMed](#)]
79. Kumar, S.; Lütteke, T.; Schwartz-Albiez, R. GlycoCD: A repository for carbohydrate-related CD antigens. *Bioinformatics* **2012**, *28*, 2553–2555. [[CrossRef](#)]
80. Allcorn, L.C.; Martin, A.C.R. SACS—Self-maintaining database of antibody crystal structure information. *Bioinformatics* **2002**, *18*, 175–181. [[CrossRef](#)] [[PubMed](#)]
81. Dunbar, J.; Krawczyk, K.; Leem, J.; Baker, T.; Fuchs, A.; Georges, G.; Shi, J.; Deane, C.M. SABDab: The structural antibody database. *Nucleic Acids Res.* **2014**, *42*, D1140–D1146. [[CrossRef](#)]
82. Lombard, V.; Golaconda Ramulu, H.; Drula, E.; Coutinho, P.M.; Henrissat, B. The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res.* **2014**, *42*, D490–D495. [[CrossRef](#)]
83. Cantarel, B.L.; Coutinho, P.M.; Rancurel, C.; Bernard, T.; Lombard, V.; Henrissat, B. The Carbohydrate-Active EnZymes database (CAZy): An expert resource for Glycogenomics. *Nucleic Acids Res.* **2009**, *37*, D233–D238. [[CrossRef](#)]



84. Terrapon, N.; Lombard, V.; Drula, E.; Coutinho, P.M.; Henrissat, B. The CAZy Database/the Carbohydrate-Active Enzyme (CAZy) Database: Principles and Usage Guidelines. In *A Practical Guide to Using Glycomics Databases*; Aoki-Kinoshita, K.F., Ed.; Springer: Tokyo, Japan, 2017; pp. 117–131.
85. Lee, T.-Y.; Huang, H.-D.; Hung, J.-H.; Huang, H.-Y.; Yang, Y.-S.; Wang, T.-H. dbPTM: An information repository of protein post-translational modification. *Nucleic Acids Res.* **2006**, *34*, D622–D627. [[CrossRef](#)]
86. Huang, K.-Y.; Su, M.-G.; Kao, H.-J.; Hsieh, Y.-C.; Jhong, J.-H.; Cheng, K.-H.; Huang, H.-D.; Lee, T.-Y. dbPTM 2016: 10-year anniversary of a resource for post-translational modification of proteins. *Nucleic Acids Res.* **2016**, *44*, D435–D446. [[CrossRef](#)] [[PubMed](#)]
87. Huang, K.-Y.; Lee, T.-Y.; Kao, H.-J.; Ma, C.-T.; Lee, C.-C.; Lin, T.-H.; Chang, W.-C.; Huang, H.-D. dbPTM in 2019: Exploring disease association and cross-talk of post-translational modifications. *Nucleic Acids Res.* **2019**, *47*, D298–D308. [[CrossRef](#)] [[PubMed](#)]
88. Kopp, J.; Schwede, T. The SWISS-MODEL Repository of annotated three-dimensional protein structure homology models. *Nucleic Acids Res.* **2004**, *32*, D230–D234. [[CrossRef](#)]
89. Kiefer, F.; Arnold, K.; Künzli, M.; Bordoli, L.; Schwede, T. The SWISS-MODEL Repository and associated resources. *Nucleic Acids Res.* **2009**, *37*, D387–D392. [[CrossRef](#)]
90. Bienert, S.; Waterhouse, A.; de Beer, T.A.P.; Tauriello, G.; Studer, G.; Bordoli, L.; Schwede, T. The SWISS-MODEL Repository—New features and functionality. *Nucleic Acids Res.* **2017**, *45*, D313–D319. [[CrossRef](#)]
91. Frank, M.; Lutteke, T.; Von Der Lieth, C.W. GlycoMapsDB: A database of the accessible conformational space of glycosidic linkages. *Nucleic Acids Res.* **2007**, *35*, 287–290. [[CrossRef](#)] [[PubMed](#)]
92. Jo, S.; Im, W. Glycan fragment database: A database of PDB-based glycan 3D structures. *Nucleic Acids Res.* **2013**, *41*, D470–D474. [[CrossRef](#)]
93. Perez, S. X-Ray Diffraction and Crystallography of Oligosaccharides and Polysaccharides. In *Encyclopedia of Biophysics*; Roberts, G.C.K., Ed.; Springer: Berlin/Heidelberg, Germany, 2013; pp. 2767–2777.
94. Gimeno, A.; Valverde, P.; Ardá, A.; Jiménez-Barbero, J. Glycan structures and their interactions with proteins. A NMR view. *Curr. Opin. Struct. Biol.* **2020**, *62*, 22–30. [[CrossRef](#)] [[PubMed](#)]
95. Blaum, B.S.; Neu, U.; Peters, T.; Stehle, T. Spin ballet for sweet encounters: Saturation-transfer difference NMR and X-ray crystallography complement each other in the elucidation of protein-glycan interactions. *Acta Crystallogr. Sect. F Struct. Biol. Commun.* **2018**, *74*, 451–462. [[CrossRef](#)] [[PubMed](#)]
96. Battistel, M.D.; Azurmendi, H.F.; Yu, B.; Freedberg, D.I. NMR of glycans: Shedding new light on old problems. *Prog. Nucl. Magn. Reson. Spectrosc.* **2014**, *79*, 48–68. [[CrossRef](#)] [[PubMed](#)]
97. Schubert, M. Insights into Carbohydrate Recognition by 3D Structure Determination of Protein–Carbohydrate Complexes Using NMR. In *NMR in Glycoscience and Glycotechnology*; Kato, K., Peters, T., Eds.; The Royal Society of Chemistry: London, UK, 2017; pp. 101–122.
98. Cléry, A.; Schubert, M.; Allain, F.H.T. NMR Spectroscopy: An Excellent Tool to Understand RNA and Carbohydrate Recognition by Proteins. *Chimia* **2012**, *66*, 741–746. [[CrossRef](#)]
99. Valverde, P.; Delgado, S.; Martínez, J.D.; Vendeville, J.-B.; Malassis, J.; Linclau, B.; Reichardt, N.-C.; Cañada, F.J.; Jiménez-Barbero, J.; Ardá, A. Molecular Insights into DC-SIGN Binding to Self-Antigens: The Interaction with the Blood Group A/B Antigens. *ACS Chem. Biol.* **2019**, *14*, 1660–1671. [[CrossRef](#)]
100. Aeschbacher, T.; Zierke, M.; Smieško, M.; Collot, M.; Mallet, J.-M.; Ernst, B.; Allain, F.H.-T.; Schubert, M. A Secondary Structural Element in a Wide Range of Fucosylated Glycoepitopes. *Chem. Eur. J.* **2017**, *23*, 11598–11610. [[CrossRef](#)]
101. Zierke, M.; Smieško, M.; Rabbani, S.; Aeschbacher, T.; Cutting, B.; Allain, F.H.T.; Schubert, M.; Ernst, B. Stabilization of Branched Oligosaccharides: Lewisx Benefits from a Nonconventional C–H…O Hydrogen Bond. *J. Am. Chem. Soc.* **2013**, *135*, 13464–13472. [[CrossRef](#)] [[PubMed](#)]
102. Sattelle, B.M.; Almond, A. Is N-acetyl-d-glucosamine a rigid 4C1 chair? *Glycobiology* **2011**, *21*, 1651–1662. [[CrossRef](#)] [[PubMed](#)]
103. França, B.A.; da Silva, C.O. Specific rotation of monosaccharides: A global property bringing local information. *Phys. Chem. Chem. Phys.* **2014**, *16*, 13096–13102. [[CrossRef](#)] [[PubMed](#)]
104. Ling, Z.; Edwards, J.V.; Nam, S.; Xu, F.; French, A.D. Conformational analysis of xylobiose by DFT quantum mechanics. *Cellulose* **2020**, *27*, 1207–1224. [[CrossRef](#)]
105. McMahon, C.M.; Isabella, C.R.; Windsor, I.W.; Kosma, P.; Raines, R.T.; Kiessling, L.L. Stereoelectronic Effects Impact Glycan Recognition. *J. Am. Chem. Soc.* **2020**, *142*, 2386–2395. [[CrossRef](#)]

106. Zhang, W.; Meredith, R.; Pan, Q.; Wang, X.; Woods, R.J.; Carmichael, I.; Serianni, A.S. Use of Circular Statistics To Model  $\alpha$ Man-(1 $\rightarrow$ 2)- $\alpha$ Man and  $\alpha$ Man-(1 $\rightarrow$ 3)- $\alpha$ / $\beta$ Man O-Glycosidic Linkage Conformation in <sup>13</sup>C-Labeled Disaccharides and High-Mannose Oligosaccharides. *Biochemistry* **2019**, *58*, 546–560. [[CrossRef](#)]
107. Saraboji, K.; Håkansson, M.; Genheden, S.; Diehl, C.; Qvist, J.; Weininger, U.; Nilsson, U.J.; Leffler, H.; Ryde, U.; Akke, M.; et al. The Carbohydrate-Binding Site in Galectin-3 Is Preorganized To Recognize a Sugarlike Framework of Oxygens: Ultra-High-Resolution Structures and Water Dynamics. *Biochemistry* **2012**, *51*, 296–306. [[CrossRef](#)]
108. Turupcu, A.; Bowen, A.M.; Di Paolo, A.; Matagne, A.; Oostenbrink, C.; Redfield, C.; Smith, L.J. An NMR and MD study of complexes of bacteriophage lambda lysozyme with tetra- and hexa-N-acetylchitohexaose. *Proteins Struct. Funct. Bioinf.* **2020**, *88*, 82–93. [[CrossRef](#)]
109. Turupcu, A.; Oostenbrink, C. Modeling of Oligosaccharides within Glycoproteins from Free-Energy Landscapes. *J. Chem. Inf. Model.* **2017**, *57*, 2222–2236. [[CrossRef](#)]
110. Zhang, W.; Turney, T.; Meredith, R.; Pan, Q.; Sernau, L.; Wang, X.; Hu, X.; Woods, R.J.; Carmichael, I.; Serianni, A.S. Conformational Populations of  $\beta$ -(1 $\rightarrow$ 4) O-Glycosidic Linkages Using Redundant NMR J-Couplings and Circular Statistics. *J. Phys. Chem. B* **2017**, *121*, 3042–3058. [[CrossRef](#)]
111. Sattelle, B.M.; Almond, A. Shaping up for structural glycomics: A predictive protocol for oligosaccharide conformational analysis applied to N-linked glycans. *Carbohydr. Res.* **2014**, *383*, 34–42. [[CrossRef](#)] [[PubMed](#)]
112. Frank, M. Conformational Analysis of Carbohydrates—A Historical Overview. In *Bioinformatics for Glycobiology and Glycomics: An Introduction*; Von Der Lieth, C.W., Lütteke, T., Frank, M., Eds.; Wiley: Hoboken, NJ, USA, 2009; pp. 335–357.
113. Frank, M. Predicting Carbohydrate 3D Structures Using Theoretical Methods. In *Bioinformatics for Glycobiology and Glycomics: An Introduction*; Von Der Lieth, C.W., Lütteke, T., Frank, M., Eds.; Wiley: Hoboken, NJ, USA, 2009; pp. 359–388.
114. Toukach, F.V.; Ananikov, V.P. Recent advances in computational predictions of NMR parameters for the structure elucidation of carbohydrates: Methods and limitations. *Chem. Soc. Rev.* **2013**, *42*, 8376–8415. [[CrossRef](#)]
115. French, A.D. Computerized Models of Carbohydrates. In *Polysaccharides: Bioactivity and Biotechnology*; Ramawat, K.G., Mérillon, J.-M., Eds.; Springer: Cham, Switzerland, 2015; pp. 1397–1440.
116. French, A.D.; Johnson, G.P. Computerized Molecular Modeling of Carbohydrates. In *The Plant Cell Wall: Methods and Protocols*; Popper, Z.A., Ed.; Humana: New York, NY, USA, 2020; pp. 513–539.
117. Feng, T.; Li, M.; Zhou, J.; Zhuang, H.; Chen, F.; Ye, R.; Campanella, O.; Fang, Z. Application of molecular dynamics simulation in food carbohydrate research—A review. *Innov. Food Sci. Emerg. Technol.* **2015**, *31*, 1–13. [[CrossRef](#)]
118. Dowd, M.K.; Kiely, D.E.; Zhang, J. Monte Carlo-based searching as a tool to study carbohydrate structure. *Carbohydr. Res.* **2011**, *346*, 1140–1148. [[CrossRef](#)]
119. Zhang, W.; Howell, S.C.; Wright, D.W.; Heindel, A.; Qiu, X.; Chen, J.; Curtis, J.E. Combined Monte Carlo/torsion-angle molecular dynamics for ensemble modeling of proteins, nucleic acids and carbohydrates. *J. Mol. Graph. Modell.* **2017**, *73*, 179–190. [[CrossRef](#)]
120. Rahal-Sekkal, M.; Sekkal, N.; Kleb, D.C.; Bleckmann, P. Structures and energies of D-galactose and galabiose conformers as calculated by *ab initio* and semiempirical methods. *J. Comput. Chem.* **2003**, *24*, 806–818. [[CrossRef](#)] [[PubMed](#)]
121. Barnett, C.B.; Naidoo, K.J. Ring Puckering: A Metric for Evaluating the Accuracy of AM1, PM3, PM3CARB-1, and SCC-DFTB Carbohydrate QM/MM Simulations. *J. Phys. Chem. B* **2010**, *114*, 17142–17154. [[CrossRef](#)]
122. Govender, K.; Gao, J.; Naidoo, K.J. AM1/d-CB1: A Semiempirical Model for QM/MM Simulations of Chemical Glycobiology Systems. *J. Chem. Theory Comput.* **2014**, *10*, 4694–4707. [[CrossRef](#)]
123. Govender, K.K.; Naidoo, K.J. Evaluating AM1/d-CB1 for Chemical Glycobiology QM/MM Simulations. *J. Chem. Theory Comput.* **2014**, *10*, 4708–4717. [[CrossRef](#)]
124. Gould, I.R.; Bettley, H.A.-A.; Bryce, R.A. Correlated *ab initio* quantum chemical calculations of di- and trisaccharide conformations. *J. Comput. Chem.* **2007**, *28*, 1965–1973. [[CrossRef](#)]
125. French, A.D.; Johnson, G.P.; Cramer, C.J.; Csonka, G.I. Conformational analysis of cellobiose by electronic structure theories. *Carbohydr. Res.* **2012**, *350*, 68–76. [[CrossRef](#)]
126. Schnupf, U.; Momany, F.A. DFT Energy Optimization of a Large Carbohydrate: Cyclomaltohexacosaoose (CA-26). *J. Phys. Chem. B* **2012**, *116*, 6618–6627. [[CrossRef](#)] [[PubMed](#)]



127. Devarajan, A.; Markutsya, S.; Lamm, M.H.; Cheng, X.; Smith, J.C.; Baluyut, J.Y.; Kholod, Y.; Gordon, M.S.; Windus, T.L. *Ab Initio* Study of Molecular Interactions in Cellulose I $\alpha$ . *J. Phys. Chem. B* **2013**, *117*, 10430–10443. [[CrossRef](#)]
128. Chan, B. Aqueous-Phase Conformations of Lactose, Maltose, and Sucrose and the Assessment of Low-Cost DFT Methods with the DSCONF Set of Conformers for the Three Disaccharides. *J. Phys. Chem. A* **2020**, *124*, 582–590. [[CrossRef](#)]
129. Ishida, T. Computational analysis of carbohydrate recognition based on hybrid QM/MM modeling: A case study of norovirus capsid protein in complex with Lewis antigen. *Phys. Chem. Chem. Phys.* **2018**, *20*, 4652–4665. [[CrossRef](#)]
130. Tafazzoli, M.; Ghiasi, M. Structure and conformation of  $\alpha$ -,  $\beta$ - and  $\gamma$ -cyclodextrin in solution: Theoretical approaches and experimental validation. *Carbohydr. Polym.* **2009**, *78*, 10–15. [[CrossRef](#)]
131. Ardèvol, A.; Rovira, C. Reaction Mechanisms in Carbohydrate-Active Enzymes: Glycoside Hydrolases and Glycosyltransferases. Insights from *ab Initio* Quantum Mechanics/Molecular Mechanics Dynamic Simulations. *J. Am. Chem. Soc.* **2015**, *137*, 7528–7547. [[CrossRef](#)]
132. Tvaroška, I. Atomistic insight into the catalytic mechanism of glycosyltransferases by combined quantum mechanics/molecular mechanics (QM/MM) methods. *Carbohydr. Res.* **2015**, *403*, 38–47. [[CrossRef](#)] [[PubMed](#)]
133. Johnson, G.P.; Petersen, L.; French, A.D.; Reilly, P.J. Twisting of glycosidic bonds by hydrolases. *Carbohydr. Res.* **2009**, *344*, 2157–2166. [[CrossRef](#)]
134. Chung, L.W.; Sameera, W.M.C.; Ramozzi, R.; Page, A.J.; Hatanaka, M.; Petrova, G.P.; Harris, T.V.; Li, X.; Ke, Z.; Liu, F.; et al. The ONIOM Method and Its Applications. *Chem. Rev.* **2015**, *115*, 5678–5796. [[CrossRef](#)]
135. Burnham, J.F. Scopus database: A review. *Biomed. Digit. Libr.* **2006**, *3*, 1. [[CrossRef](#)]
136. Frank, M.; Schloissnig, S. Bioinformatics and molecular modeling in glycobiology. *Cell. Mol. Life Sci.* **2010**, *67*, 2749–2772. [[CrossRef](#)]
137. Sapay, N.; Nurisso, A.; Imberty, A. Simulation of Carbohydrates, from Molecular Docking to Dynamics in Water. In *Biomolecular Simulations: Methods and Protocols*; Monticelli, L., Salonen, E., Eds.; Humana Press: Totowa, NJ, USA, 2013; pp. 469–483.
138. Pérez, S.; Tvaroška, I. Chapter 1—Carbohydrate–Protein Interactions: Molecular Modeling Insights. In *Advances in Carbohydrate Chemistry and Biochemistry*; Horton, D., Ed.; Academic Press: Cambridge, MA, USA, 2014; Volume 71, pp. 9–136.
139. Fadda, E.; Woods, R.J. Molecular simulations of carbohydrates and protein–carbohydrate interactions: Motivation, issues and prospects. *Drug Discov. Today* **2010**, *15*, 596–609. [[CrossRef](#)]
140. Yongye, A.B.; Gonzalez-Outeiriño, J.; Glushka, J.; Schultheis, V.; Woods, R.J. The Conformational Properties of Methyl  $\alpha$ -(2,8)-Di/Trisialosides and Their N-Acyl Analogues: Implications for Anti-Neisseria meningitidis B Vaccine Design. *Biochemistry* **2008**, *47*, 12493–12514. [[CrossRef](#)]
141. Re, S.; Nishima, W.; Miyashita, N.; Sugita, Y. Conformational flexibility of N-glycans in solution studied by REMD simulations. *Biophys. Rev.* **2012**, *4*, 179–187. [[CrossRef](#)]
142. Patel, D.S.; Pendrill, R.; Mallajosyula, S.S.; Widmalm, G.; MacKerell, A.D. Conformational Properties of  $\alpha$ - or  $\beta$ -(1 $\rightarrow$ 6)-Linked Oligosaccharides: Hamiltonian Replica Exchange MD Simulations and NMR Experiments. *J. Phys. Chem. B* **2014**, *118*, 2851–2871. [[CrossRef](#)]
143. Mishra, S.K.; Kara, M.; Zacharias, M.; Koča, J. Enhanced conformational sampling of carbohydrates by Hamiltonian replica-exchange simulation. *Glycobiology* **2014**, *24*, 70–84. [[CrossRef](#)]
144. Mallajosyula, S.S.; MacKerell, A.D. Influence of Solvent and Intramolecular Hydrogen Bonding on the Conformational Properties of O-Linked Glycopeptides. *J. Phys. Chem. B* **2011**, *115*, 11215–11229. [[CrossRef](#)] [[PubMed](#)]
145. Alibay, I.; Burusco, K.K.; Bruce, N.J.; Bryce, R.A. Identification of Rare Lewis Oligosaccharide Conformers in Aqueous Solution Using Enhanced Sampling Molecular Dynamics. *J. Phys. Chem. B* **2018**, *122*, 2462–2474. [[CrossRef](#)] [[PubMed](#)]
146. Alibay, I.; Bryce, R.A. Ring Puckering Landscapes of Glycosaminoglycan-Related Monosaccharides from Molecular Dynamics Simulations. *J. Chem. Inf. Model.* **2019**, *59*, 4729–4741. [[CrossRef](#)] [[PubMed](#)]
147. Balogh, G.; Komáromi, I.; Bereczky, Z. The mechanism of high affinity pentasaccharide binding to antithrombin, insights from Gaussian accelerated molecular dynamics simulations. *J. Biomol. Struct. Dyn.* **2019**, *38*, 4718–4732. [[CrossRef](#)]

148. Balogh, G.; Gyöngyösi, T.; Timári, I.; Herczeg, M.; Borbás, A.; Fehér, K.; Kövér, K.E. Comparison of Carbohydrate Force Fields Using Gaussian Accelerated Molecular Dynamics Simulations and Development of Force Field Parameters for Heparin-Analogue Pentasaccharides. *J. Chem. Inf. Model.* **2019**, *59*, 4855–4867. [[CrossRef](#)] [[PubMed](#)]
149. Suzuki, T.; Kajino, M.; Yanaka, S.; Zhu, T.; Yagi, H.; Satoh, T.; Yamaguchi, T.; Kato, K. Conformational Analysis of a High-Mannose-Type Oligosaccharide Displaying Glucosyl Determinant Recognised by Molecular Chaperones Using NMR-Validated Molecular Dynamics Simulation. *ChemBioChem* **2017**, *18*, 396–401. [[CrossRef](#)] [[PubMed](#)]
150. Yamaguchi, T.; Sakae, Y.; Zhang, Y.; Yamamoto, S.; Okamoto, Y.; Kato, K. Exploration of Conformational Spaces of High-Mannose-Type Oligosaccharides by an NMR-Validated Simulation. *Angew. Chem. Int. Ed.* **2014**, *53*, 10941–10944. [[CrossRef](#)] [[PubMed](#)]
151. Foley, B.L.; Tessier, M.B.; Woods, R.J. Carbohydrate force fields. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2012**, *2*, 652–697. [[CrossRef](#)]
152. Kozmon, S.; Matuška, R.; Spiwok, V.; Koča, J. Dispersion interactions of carbohydrates with condensate aromatic moieties: Theoretical study on the CH- $\pi$  interaction additive properties. *Phys. Chem. Chem. Phys.* **2011**, *13*, 14215–14222. [[CrossRef](#)]
153. Hudson, K.L.; Bartlett, G.J.; Diehl, R.C.; Agirre, J.; Gallagher, T.; Kiessling, L.L.; Woolfson, D.N. Carbohydrate–Aromatic Interactions in Proteins. *J. Am. Chem. Soc.* **2015**, *137*, 15152–15160. [[CrossRef](#)]
154. Hsu, C.-H.; Park, S.; Mortenson, D.E.; Foley, B.L.; Wang, X.; Woods, R.J.; Case, D.A.; Powers, E.T.; Wong, C.-H.; Dyson, H.J.; et al. The Dependence of Carbohydrate–Aromatic Interaction Strengths on the Structure of the Carbohydrate. *J. Am. Chem. Soc.* **2016**, *138*, 7636–7648. [[CrossRef](#)]
155. Stanković, I.M.; Blagojević Filipović, J.P.; Zarić, S.D. Carbohydrate–Protein aromatic ring interactions beyond CH/ $\pi$  interactions: A Protein Data Bank survey and quantum chemical calculations. *Int. J. Biol. Macromol.* **2020**, *157*, 1–9. [[CrossRef](#)]
156. Nivedha, A.K.; Makeneni, S.; Foley, B.L.; Tessier, M.B.; Woods, R.J. Importance of ligand conformational energies in carbohydrate docking: Sorting the wheat from the chaff. *J. Comput. Chem.* **2014**, *35*, 526–539. [[CrossRef](#)]
157. Nivedha, A.K.; Thieker, D.F.; Hu, H.; Woods, R.J. Vina-Carb: Improving Glycosidic Angles during Carbohydrate Docking. *J. Chem. Theory Comput.* **2016**, *12*, 892–901. [[CrossRef](#)]
158. Pérez, S. 2.11—Molecular Modeling in Glycoscience. In *Comprehensive Glycoscience*; Kamerling, H., Ed.; Elsevier: Oxford, UK, 2007; Volume 2, pp. 347–388.
159. Stortz, C.A.; French, A.D. Disaccharide conformational maps: Adiabaticity in analogues with variable ring shapes. *Mol. Simul.* **2008**, *34*, 373–389. [[CrossRef](#)]
160. Allinger, N.L.; Yuh, Y.H.; Lii, J.H. Molecular mechanics. The MM3 force field for hydrocarbons. 1. *J. Am. Chem. Soc.* **1989**, *111*, 8551–8566. [[CrossRef](#)]
161. Allinger, N.L.; Rahman, M.; Lii, J.H. A molecular mechanics force field (MM3) for alcohols and ethers. *J. Am. Chem. Soc.* **1990**, *112*, 8293–8307. [[CrossRef](#)]
162. Stortz, C.A. Comparative performance of MM3(92) and two TINKER™ MM3 versions for the modeling of carbohydrates. *J. Comput. Chem.* **2005**, *26*, 471–483. [[CrossRef](#)]
163. Stortz, C.A.; Johnson, G.P.; French, A.D.; Csonka, G.I. Comparison of different force fields for the study of disaccharides. *Carbohydr. Res.* **2009**, *344*, 2217–2228. [[CrossRef](#)] [[PubMed](#)]
164. Taha, H.A.; Richards, M.R.; Lowary, T.L. Conformational analysis of furanoside-containing mono- and oligosaccharides. *Chem. Rev.* **2013**, *113*, 1851–1876. [[CrossRef](#)]
165. Stortz, C.A. Additive effects in the modeling of oligosaccharides with mm3 at high dielectric constants: An approach to the ‘multiple minimum problem’. *Carbohydr. Res.* **2006**, *341*, 663–671. [[CrossRef](#)]
166. Stortz, C.A. mm3 Potential energy surfaces of trisaccharide models of  $\lambda$ -,  $\mu$ -, and  $\nu$ -carrageenans. *Carbohydr. Res.* **2006**, *341*, 2531–2542. [[CrossRef](#)]
167. Xiong, X.M.; Chen, Z.Q.; Cossins, B.P.; Xu, Z.J.; Shao, Q.; Ding, K.; Zhu, W.L.; Shi, J.Y. Force fields and scoring functions for carbohydrate simulation. *Carbohydr. Res.* **2015**, *401*, 73–81. [[CrossRef](#)]
168. CHARMM Force Field Files. Available online: <https://www.charmm.org/charmm/resources/charmm-force-fields/#charmm> (accessed on 31 July 2020).

169. MacKerell, A.D.; Bashford, D.; Bellott, M.; Dunbrack, R.L.; Evanseck, J.D.; Field, M.J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; et al. All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. *J. Phys. Chem. B* **1998**, *102*, 3586–3616. [[CrossRef](#)]
170. Mackerell, A.D., Jr.; Feig, M.; Brooks, C.L., III. Extending the treatment of backbone energetics in protein force fields: Limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. *J. Comput. Chem.* **2004**, *25*, 1400–1415. [[CrossRef](#)]
171. Guvench, O.; Hatcher, E.; Venable, R.M.; Pastor, R.W.; MacKerell, A.D. CHARMM Additive All-Atom Force Field for Glycosidic Linkages between Hexopyranoses. *J. Chem. Theory Comput.* **2009**, *5*, 2353–2370. [[CrossRef](#)]
172. Guvench, O.; Greene, S.N.; Kamath, G.; Brady, J.W.; Venable, R.M.; Pastor, R.W.; Mackerell, A.D., Jr. Additive empirical force field for hexopyranose monosaccharides. *J. Comput. Chem.* **2008**, *29*, 2543–2564. [[CrossRef](#)] [[PubMed](#)]
173. Raman, E.P.; Guvench, O.; MacKerell, A.D. CHARMM Additive All-Atom Force Field for Glycosidic Linkages in Carbohydrates Involving Furanoses. *J. Phys. Chem. B* **2010**, *114*, 12981–12994. [[CrossRef](#)] [[PubMed](#)]
174. Guvench, O.; Mallajosyula, S.S.; Raman, E.P.; Hatcher, E.; Vanommeslaeghe, K.; Foster, T.J.; Jamison, F.W.; MacKerell, A.D. CHARMM Additive All-Atom Force Field for Carbohydrate Derivatives and Its Utility in Polysaccharide and Carbohydrate–Protein Modeling. *J. Chem. Theory Comput.* **2011**, *7*, 3162–3180. [[CrossRef](#)]
175. Mallajosyula, S.S.; Guvench, O.; Hatcher, E.; MacKerell, A.D. CHARMM Additive All-Atom Force Field for Phosphate and Sulfate Linked to Carbohydrates. *J. Chem. Theory Comput.* **2012**, *8*, 759–776. [[CrossRef](#)]
176. Cloutier, T.; Sudrik, C.; Sathish, H.A.; Trout, B.L. Kirkwood–Buff-Derived Alcohol Parameters for Aqueous Carbohydrates and Their Application to Preferential Interaction Coefficient Calculations of Proteins. *J. Phys. Chem. B* **2018**, *122*, 9350–9360. [[CrossRef](#)] [[PubMed](#)]
177. Kirschner, K.N.; Yongye, A.B.; Tschampel, S.M.; González-Outeiriño, J.; Daniels, C.R.; Foley, B.L.; Woods, R.J. GLYCAM06: A generalizable biomolecular force field. Carbohydrates. *J. Comput. Chem.* **2008**, *29*, 622–655. [[CrossRef](#)]
178. Tessier, M.B.; DeMarco, M.L.; Yongye, A.B.; Woods, R.J. Extension of the GLYCAM06 biomolecular force field to lipids, lipid bilayers and glycolipids. *Mol. Simul.* **2008**, *34*, 349–364. [[CrossRef](#)]
179. DeMarco, M.L.; Woods, R.J. Atomic-resolution conformational analysis of the GM3 ganglioside in a lipid bilayer and its implications for ganglioside–protein recognition at membrane surfaces. *Glycobiology* **2009**, *19*, 344–355. [[CrossRef](#)]
180. DeMarco, M.L. Molecular Dynamics Simulations of Membrane- and Protein-Bound Glycolipids Using GLYCAM. In *Glycoinformatics*; Lütke, T., Frank, M., Eds.; Humana Press: New York, NY, USA, 2015; pp. 379–390.
181. Kirschner, K.N.; Lins, R.D.; Maass, A.; Soares, T.A. A Glycam-Based Force Field for Simulations of Lipopolysaccharide Membranes: Parametrization and Validation. *J. Chem. Theory Comput.* **2012**, *8*, 4719–4731. [[CrossRef](#)] [[PubMed](#)]
182. Singh, A.; Tessier, M.B.; Pederson, K.; Wang, X.; Venot, A.P.; Boons, G.-J.; Prestegard, J.H.; Woods, R.J. Extension and validation of the GLYCAM force field parameters for modeling glycosaminoglycans. *Can. J. Chem.* **2016**, *94*, 927–935. [[CrossRef](#)]
183. Lins, R.D.; Hünenberger, P.H. A new GROMOS force field for hexopyranose-based carbohydrates. *J. Comput. Chem.* **2005**, *26*, 1400–1412. [[CrossRef](#)] [[PubMed](#)]
184. Hansen, H.S.; Hünenberger, P.H. A reoptimized GROMOS force field for hexopyranose-based carbohydrates accounting for the relative free energies of ring conformers, anomers, epimers, hydroxymethyl rotamers, and glycosidic linkage conformers. *J. Comput. Chem.* **2011**, *32*, 998–1032. [[CrossRef](#)]
185. Pontes, F.J.S.; Rusu, V.H.; Soares, T.A.; Lins, R.D. The Effect of Temperature, Cations, and Number of Acyl Chains on the Lamellar to Non-Lamellar Transition in Lipid-A Membranes: A Microscopic View. *J. Chem. Theory Comput.* **2012**, *8*, 3830–3838. [[CrossRef](#)]
186. Pol-Fachin, L.; Rusu, V.H.; Verli, H.; Lins, R.D. GROMOS 53A6GLYC, an Improved GROMOS Force Field for Hexopyranose-Based Carbohydrates. *J. Chem. Theory Comput.* **2012**, *8*, 4681–4690. [[CrossRef](#)] [[PubMed](#)]
187. Pol-Fachin, L.; Verli, H.; Lins, R.D. Extension and validation of the GROMOS 53A6glyc parameter set for glycoproteins. *J. Comput. Chem.* **2014**, *35*, 2087–2095. [[CrossRef](#)] [[PubMed](#)]

188. Plazinski, W.; Lonardi, A.; Hünenberger, P.H. Revision of the GROMOS 56A6CARBO force field: Improving the description of ring-conformational equilibria in hexopyranose-based carbohydrates chains. *J. Comput. Chem.* **2016**, *37*, 354–365. [[CrossRef](#)]
189. Naumov, V.S.; Ignatov, S.K. Modification of 56ACARBO force field for molecular dynamic calculations of chitosan and its derivatives. *J. Mol. Model.* **2017**, *23*, 244. [[CrossRef](#)]
190. Panczyk, K.; Gaweda, K.; Drach, M.; Plazinski, W. Extension of the GROMOS 56a6CARBO/CARBO\_R Force Field for Charged, Protonated, and Esterified Uronates. *J. Phys. Chem. B* **2018**, *122*, 3696–3710. [[CrossRef](#)] [[PubMed](#)]
191. Nester, K.; Gaweda, K.; Plazinski, W. A GROMOS Force Field for Furanose-Based Carbohydrates. *J. Chem. Theory Comput.* **2019**, *15*, 1168–1186. [[CrossRef](#)]
192. Pol-Fachin, L.; Fernandes, C.L.; Verli, H. GROMOS96 43a1 performance on the characterization of glycoprotein conformational ensembles through molecular dynamics simulations. *Carbohydr. Res.* **2009**, *344*, 491–500. [[CrossRef](#)] [[PubMed](#)]
193. Fernandes, C.L.; Sachett, L.G.; Pol-Fachin, L.; Verli, H. GROMOS96 43a1 performance in predicting oligosaccharide conformational ensembles within glycoproteins. *Carbohydr. Res.* **2010**, *345*, 663–671. [[CrossRef](#)] [[PubMed](#)]
194. Kony, D.; Damm, W.; Stoll, S.; Van Gunsteren, W.F. An improved OPLS-AA force field for carbohydrates. *J. Comput. Chem.* **2002**, *23*, 1416–1429. [[CrossRef](#)] [[PubMed](#)]
195. Damm, W.; Frontera, A.; Tirado-Rives, J.; Jorgensen, W.L. OPLS all-atom force field for carbohydrates. *J. Comput. Chem.* **1997**, *18*, 1955–1970. [[CrossRef](#)]
196. Jorgensen, W.L.; Maxwell, D.S.; Tirado-Rives, J. Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. *J. Am. Chem. Soc.* **1996**, *118*, 11225–11236. [[CrossRef](#)]
197. Jamali, S.H.; Westen, T.V.; Moulton, O.A.; Vlugt, T.J.H. Optimizing Nonbonded Interactions of the OPLS Force Field for Aqueous Solutions of Carbohydrates: How to Capture Both Thermodynamics and Dynamics. *J. Chem. Theory Comput.* **2018**, *14*, 6690–6700. [[CrossRef](#)]
198. Patel, D.S.; He, X.; MacKerell, A.D. Polarizable Empirical Force Field for Hexopyranose Monosaccharides Based on the Classical Drude Oscillator. *J. Phys. Chem. B* **2015**, *119*, 637–652. [[CrossRef](#)]
199. Yang, M.; Aytenfisu, A.H.; MacKerell, A.D. Proper balance of solvent-solute and solute-solute interactions in the treatment of the diffusion of glucose using the Drude polarizable force field. *Carbohydr. Res.* **2018**, *457*, 41–50. [[CrossRef](#)]
200. Jana, M.; MacKerell, A.D. CHARMM Drude Polarizable Force Field for Aldopentofuranoses and Methyl-aldopentofuranosides. *J. Phys. Chem. B* **2015**, *119*, 7846–7859. [[CrossRef](#)] [[PubMed](#)]
201. Pandey, P.; Aytenfisu, A.H.; MacKerell, A.D.; Mallajosyula, S.S. Drude Polarizable Force Field Parametrization of Carboxylate and N-Acetyl Amine Carbohydrate Derivatives. *J. Chem. Theory Comput.* **2019**, *15*, 4982–5000. [[CrossRef](#)] [[PubMed](#)]
202. He, X.; Lopes, P.E.M.; MacKerell, A.D., Jr. Polarizable Empirical Force Field for Acyclic Polyalcohols Based on the Classical Drude Oscillator. *Biopolymers* **2013**, *99*, 724–738. [[CrossRef](#)] [[PubMed](#)]
203. Aytenfisu, A.H.; Yang, M.; MacKerell, A.D. CHARMM Drude Polarizable Force Field for Glycosidic Linkages Involving Pyranoses and Furanoses. *J. Chem. Theory Comput.* **2018**, *14*, 3132–3143. [[CrossRef](#)]
204. López, C.A.; Rzepiela, A.J.; De Vries, A.H.; Dijkhuizen, L.; Hünenberger, P.H.; Marrink, S.J. Martini Coarse-Grained Force Field: Extension to Carbohydrates. *J. Chem. Theory Comput.* **2009**, *5*, 3195–3210. [[CrossRef](#)]
205. Schmalhorst, P.S.; Deluweit, F.; Scherrers, R.; Heisenberg, C.-P.; Sikora, M. Overcoming the Limitations of the MARTINI Force Field in Simulations of Polysaccharides. *J. Chem. Theory Comput.* **2017**, *13*, 5039–5053. [[CrossRef](#)]
206. Shivgan, A.; Marzinek, J.K.; Huber, R.G.; Krah, A.; Henchman, R.H.; Matsudaira, P.T.; Verma, C.S.; Bond, P.J. Extending the Martini Coarse-Grained Forcefield to N-Glycans. *J. Chem. Inf. Model.* **2020**, *60*, 3864–3883. [[CrossRef](#)]
207. López, C.A.; Sovova, Z.; van Eerden, F.J.; De Vries, A.H.; Marrink, S.J. Martini Force Field Parameters for Glycolipids. *J. Chem. Theory Comput.* **2013**, *9*, 1694–1708. [[CrossRef](#)]
208. Rusu, V.H.; Baron, R.; Lins, R.D. PITOMBA: Parameter Interface for Oligosaccharide Molecules Based on Atoms. *J. Chem. Theory Comput.* **2014**, *10*, 5068–5080. [[CrossRef](#)]



209. Spiwok, V.; Lipovová, P.; Skálová, T.; Vondráčková, E.; Dohnálek, J.; Hašek, J.; Králová, B. Modelling of carbohydrate–aromatic interactions: *Ab initio* energetics and force field performance. *J. Comput. Aided Mol. Des.* **2005**, *19*, 887–901. [[CrossRef](#)]
210. Wimmerová, M.; Kozmon, S.; Nečasová, I.; Mishra, S.K.; Komárek, J.; Koča, J. Stacking interactions between carbohydrate and protein quantified by combination of theoretical and experimental methods. *PLoS ONE* **2012**, *7*, e46032. [[CrossRef](#)]
211. Makeneni, S.; Thieker, D.F.; Woods, R.J. Applying Pose Clustering and MD Simulations To Eliminate False Positives in Molecular Docking. *J. Chem. Inf. Model.* **2018**, *58*, 605–614. [[CrossRef](#)] [[PubMed](#)]
212. Huang, S.-Y.; Grinter, S.Z.; Zou, X. Scoring functions and their evaluation methods for protein–ligand docking: Recent advances and future directions. *Phys. Chem. Chem. Phys.* **2010**, *12*, 12899–12908. [[CrossRef](#)]
213. Vandebussche, S.; Díaz, D.; Fernández-Alonso, M.C.; Pan, W.; Vincent, S.P.; Cuevas, G.; Cañada, F.J.; Jiménez-Barbero, J.; Bartik, K. Aromatic–Carbohydrate Interactions: An NMR and Computational Study of Model Systems. *Chem. Eur. J.* **2008**, *14*, 7570–7578. [[CrossRef](#)]
214. Hill, A.D.; Reilly, P.J. A Gibbs free energy correlation for automated docking of carbohydrates. *J. Comput. Chem.* **2008**, *29*, 1131–1141. [[CrossRef](#)] [[PubMed](#)]
215. Malik, A.; Baig, M.H.; Manavalan, B. Protein–Carbohydrate Interactions. In *Encyclopedia of Bioinformatics and Computational Biology*; Ranganathan, S., Gribskov, M., Nakai, K., Schönbach, C., Eds.; Elsevier: Oxford, UK, 2019; Volume 3, pp. 666–677.
216. Samsonov, S.A.; Teyra, J.; Pisabarro, M.T. Docking glycosaminoglycans to proteins: Analysis of solvent inclusion. *J. Comput. Aided Mol. Des.* **2011**, *25*, 477–489. [[CrossRef](#)] [[PubMed](#)]
217. Samsonov, S.A.; Gehrcke, J.-P.; Pisabarro, M.T. Flexibility and Explicit Solvent in Molecular-Dynamics-Based Docking of Protein–Glycosaminoglycan Systems. *J. Chem. Inf. Model.* **2014**, *54*, 582–592. [[CrossRef](#)]
218. Gerlits, O.O.; Coates, L.; Woods, R.J.; Kovalevsky, A. Mannobiose Binding Induces Changes in Hydrogen Bonding and Protonation States of Acidic Residues in Concanavalin A As Revealed by Neutron Crystallography. *Biochemistry* **2017**, *56*, 4747–4750. [[CrossRef](#)] [[PubMed](#)]
219. Yuriev, E.; Holien, J.; Ramsland, P.A. Improvements, trends, and new ideas in molecular docking: 2012–2013 in review. *J. Mol. Recognit.* **2015**, *28*, 581–604. [[CrossRef](#)]
220. Mishra, N.K.; Kříž, Z.; Wimmerová, M.; Koča, J. Recognition of selected monosaccharides by *Pseudomonas aeruginosa* Lectin II analyzed by molecular dynamics and free energy calculations. *Carbohydr. Res.* **2010**, *345*, 1432–1441. [[CrossRef](#)] [[PubMed](#)]
221. Mishra, S.K.; Adam, J.; Wimmerová, M.; Koča, J. in silico Mutagenesis and Docking Study of *Ralstonia solanacearum* RSL Lectin: Performance of Docking Software to Predict Saccharide Binding. *J. Chem. Inf. Model.* **2012**, *52*, 1250–1261. [[CrossRef](#)] [[PubMed](#)]
222. Raghuraman, A.; Mosier, P.D.; Desai, U.R. Finding a Needle in a Haystack: Development of a Combinatorial Virtual Screening Approach for Identifying High Specificity Heparin/Heparan Sulfate Sequence(s). *J. Med. Chem.* **2006**, *49*, 3553–3562. [[CrossRef](#)] [[PubMed](#)]
223. Samsonov, S.A.; Pisabarro, M.T. Computational analysis of interactions in structurally available protein–glycosaminoglycan complexes. *Glycobiology* **2016**, *26*, 850–861. [[CrossRef](#)] [[PubMed](#)]
224. Gehrcke, J.-P.; Pisabarro, M.T. Identification and characterization of a glycosaminoglycan binding site on interleukin-10 via molecular simulation methods. *J. Mol. Graph. Modell.* **2015**, *62*, 97–104. [[CrossRef](#)]
225. Agostino, M.; Sandrin, M.S.; Thompson, P.E.; Yuriev, E.; Ramsland, P.A. in silico analysis of antibody–carbohydrate interactions and its application to xenoreactive antibodies. *Mol. Immunol.* **2009**, *47*, 233–246. [[CrossRef](#)]
226. Lee, J.; Cheng, X.; Swails, J.M.; Yeom, M.S.; Eastman, P.K.; Lemkul, J.A.; Wei, S.; Buckner, J.; Jeong, J.C.; Qi, Y.; et al. CHARMM-GUI Input Generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM Simulations Using the CHARMM36 Additive Force Field. *J. Chem. Theory Comput.* **2016**, *12*, 405–413. [[CrossRef](#)]
227. Lee, J.; Hitznerberger, M.; Rieger, M.; Kern, N.R.; Zacharias, M.; Im, W. CHARMM-GUI supports the Amber force fields. *J. Chem. Phys.* **2020**, *153*, 035103. [[CrossRef](#)]
228. Labonte, J.W.; Adolf-Bryfogle, J.; Schief, W.R.; Gray, J.J. Residue-centric modeling and design of saccharide and glycoconjugate structures. *J. Comput. Chem.* **2017**, *38*, 276–287. [[CrossRef](#)]
229. Whitmore, E.K.; Vesenka, G.; Sihler, H.; Guvench, O. Efficient Construction of Atomic-Resolution Models of Non-Sulfated Chondroitin Glycosaminoglycan Using Molecular Dynamics Data. *Biomolecules* **2020**, *10*, 537. [[CrossRef](#)]

230. Park, S.J.; Lee, J.; Qi, Y.F.; Kern, N.R.; Lee, H.S.; Jo, S.; Joung, I.; Joo, K.; Lee, J.; Im, W. CHARMM-GUI Glycan Modeler for modeling and simulation of carbohydrates and glycoconjugates. *Glycobiology* **2019**, *29*, 320–331. [[CrossRef](#)]
231. Lemmin, T.; Soto, C. Glycosylator: A Python framework for the rapid modeling of glycans. *BMC Bioinf.* **2019**, *20*, 513. [[CrossRef](#)] [[PubMed](#)]
232. Alford, R.F.; Leaver-Fay, A.; Jeliazkov, J.R.; O'Meara, M.J.; DiMaio, F.P.; Park, H.; Shapovalov, M.V.; Renfrew, P.D.; Mulligan, V.K.; Kappel, K.; et al. The Rosetta All-Atom Energy Function for Macromolecular Modeling and Design. *J. Chem. Theory Comput.* **2017**, *13*, 3031–3048. [[CrossRef](#)] [[PubMed](#)]
233. Roy Burman, S.S.; Nance, M.L.; Jeliazkov, J.R.; Labonte, J.W.; Lubin, J.H.; Biswas, N.; Gray, J.J. Novel sampling strategies and a coarse-grained score function for docking homomers, flexible heteromers, and oligosaccharides using Rosetta in CAPRI rounds 37–45. *Proteins Struct. Funct. Bioinf.* **2020**, *88*, 973–985. [[CrossRef](#)] [[PubMed](#)]
234. Leman, J.K.; Weitzner, B.D.; Lewis, S.M.; Adolf-Bryfogle, J.; Alam, N.; Alford, R.F.; Aprahamian, M.; Baker, D.; Barlow, K.A.; Barth, P.; et al. Macromolecular modeling and design in Rosetta: Recent methods and frameworks. *Nat. Methods* **2020**, *17*, 665–680. [[CrossRef](#)] [[PubMed](#)]
235. Arroyuelo, A.; Vila, J.A.; Martin, O.A. Azahar: A PyMOL plugin for construction, visualization and analysis of glycan molecules. *J. Comput. Aided Mol. Des.* **2016**, *30*, 619–624. [[CrossRef](#)]
236. Rosen, J.; Miguet, L.; Pérez, S. Shape: Automatic conformation prediction of carbohydrates using a genetic algorithm. *J. Cheminf.* **2009**, *1*, 16. [[CrossRef](#)]
237. Frank, M.; Bohne-Lang, A.; Wetter, T.; Von Der Lieth, C.-W. Rapid Generation of a Representative Ensemble of N-Glycan Conformations. *In Silico Biol.* **2002**, *2*, 427–439.
238. Nahmany, A.; Strino, F.; Rosen, J.; Kemp, G.J.L.; Nyholm, P.-G. The use of a genetic algorithm search for molecular mechanics (MM3)-based conformational analysis of oligosaccharides. *Carbohydr. Res.* **2005**, *340*, 1059–1064. [[CrossRef](#)]
239. Xia, J.; Daly, R.P.; Chuang, F.-C.; Parker, L.; Jensen, J.H.; Margulis, C.J. Sugar Folding: A Novel Structural Prediction Tool for Oligosaccharides and Polysaccharides 1. *J. Chem. Theory Comput.* **2007**, *3*, 1620–1628. [[CrossRef](#)]
240. Xia, J.; Daly, R.P.; Chuang, F.-C.; Parker, L.; Jensen, J.H.; Margulis, C.J. Sugar Folding: A Novel Structural Prediction Tool for Oligosaccharides and Polysaccharides 2. *J. Chem. Theory Comput.* **2007**, *3*, 1629–1643. [[CrossRef](#)]
241. Xia, J.; Margulis, C. A tool for the prediction of structures of complex sugars. *J. Biomol. NMR* **2008**, *42*, 241–256. [[CrossRef](#)]
242. Xia, J.; Margulis, C.J. Computational Study of the Conformational Structures of Saccharides in Solution Based on J Couplings and the “Fast Sugar Structure Prediction Software”. *Biomacromolecules* **2009**, *10*, 3081–3088. [[CrossRef](#)] [[PubMed](#)]
243. Bohne-Lang, A.; Von Der Lieth, C.-W. GlyProt: in silico glycosylation of proteins. *Nucleic Acids Res.* **2005**, *33*, W214–W219. [[CrossRef](#)]
244. Liebschner, D.; Afonine, P.V.; Baker, M.L.; Bunkoczi, G.; Chen, V.B.; Croll, T.I.; Hintze, B.; Hung, L.-W.; Jain, S.; McCoy, A.J.; et al. Macromolecular structure determination using X-rays, neutrons and electrons: Recent developments in Phenix. *Acta Crystallogr. Sect. D: Struct. Biol.* **2019**, *75*, 861–877. [[CrossRef](#)] [[PubMed](#)]
245. Tessier, M.B.; Grant, O.C.; Heimburg-Molinaro, J.; Smith, D.; Jadey, S.; Gulick, A.M.; Glushka, J.; Deutscher, S.L.; Rittenhouse-Olson, K.; Woods, R.J. Computational screening of the human TF-glycome provides a structural definition for the specificity of anti-tumor antibody JAA-F11. *PLoS ONE* **2013**, *8*, e54874. [[CrossRef](#)]
246. Grant, O.C.; Smith, H.M.; Firsova, D.; Fadda, E.; Woods, R.J. Presentation, presentation, presentation! Molecular-level insight into linker effects on glycan array screening data. *Glycobiology* **2014**, *24*, 17–25. [[CrossRef](#)] [[PubMed](#)]
247. Grant, O.C.; Woods, R.J. Recent advances in employing molecular modelling to determine the specificity of glycan-binding proteins. *Curr. Opin. Struct. Biol.* **2014**, *28*, 47–55. [[CrossRef](#)]
248. Grant, O.C.; Xue, X.; Ra, D.; Khatamian, A.; Foley, B.L.; Woods, R.J. Gly-Spec: A webtool for predicting glycan specificity by integrating glycan array screening data and 3D structure. *Glycobiology* **2016**, *26*, 1027–1028. [[CrossRef](#)] [[PubMed](#)]
249. Grant, O.C.; Tessier, M.B.; Meche, L.; Mahal, L.K.; Foley, B.L.; Woods, R.J. Combining 3D structure with glycan array data provides insight into the origin of glycan specificity. *Glycobiology* **2016**, *26*, 772–783. [[CrossRef](#)]



250. Jo, S.; Kim, T.; Im, W. Automated builder and database of protein/membrane complexes for molecular dynamics simulations. *PLoS ONE* **2007**, *2*, e880. [[CrossRef](#)]
251. Jo, S.; Lim, J.B.; Klauda, J.B.; Im, W. CHARMM-GUI Membrane Builder for mixed bilayers and its application to yeast membranes. *Biophys. J.* **2009**, *97*, 50–58. [[CrossRef](#)] [[PubMed](#)]
252. Wu, E.L.; Cheng, X.; Jo, S.; Rui, H.; Song, K.C.; Dávila-Contreras, E.M.; Qi, Y.; Lee, J.; Monje-Galvan, V.; Venable, R.M.; et al. CHARMM-GUI Membrane Builder toward realistic biological membrane simulations. *J. Comput. Chem.* **2014**, *35*, 1997–2004. [[CrossRef](#)]
253. Gao, Y.; Lee, J.; Widmalm, G.; Im, W. Modeling and Simulation of Bacterial Outer Membranes with Lipopolysaccharides and Enterobacterial Common Antigen. *J. Phys. Chem. B* **2020**, *124*, 5948–5956. [[CrossRef](#)]
254. Baltoumas, F.A.; Hamodrakas, S.J.; Iconomidou, V.A. The gram-negative outer membrane modeler: Automated building of lipopolysaccharide-rich bacterial outer membranes in four force fields. *J. Comput. Chem.* **2019**, *40*, 1727–1734. [[CrossRef](#)] [[PubMed](#)]
255. Krüger, D.M.; Kamerlin, S.C.L. Micelle Maker: An Online Tool for Generating Equilibrated Micelles as Direct Input for Molecular Dynamics Simulations. *ACS Omega* **2017**, *2*, 4524–4530. [[CrossRef](#)]
256. Dashti, H.; Westler, W.M.; Wedell, J.R.; Demler, O.V.; Eghbalnia, H.R.; Markley, J.L.; Mora, S. Probabilistic identification of saccharide moieties in biomolecules and their protein complexes. *Sci. Data* **2020**, *7*, 210. [[CrossRef](#)]
257. Woods, R. GlyFinder and GlyProbity: New Online Tools for Locating and Curating Carbohydrate Structures in wwPDB. In *Time-Proof Perspectives on Glycoscience—Beilstein Glyco-Bioinformatics Symposium, Limburg, Germany, 25–27 June 2019*; Beilstein-Institut: Frankfurt, Germany, 2019; pp. 82–83.
258. Woods, R.J.; Montgomery, D.W.; Young, J.; Grant, O.C.; Wentworth, D.; Foley, B.L. Tools to Find Glycoproteins in the Protein Data Bank and Generate Realistic 3D Structures for Them. *FASEB J.* **2020**, *34*, 1. [[CrossRef](#)]
259. Lütteke, T.; Frank, M.; von der Lieth, C.-W. Data mining the protein data bank: Automatic detection and assignment of carbohydrate structures. *Carbohydr. Res.* **2004**, *339*, 1015–1020. [[CrossRef](#)]
260. Jo, S.; Song, K.C.; Desaire, H.; MacKerell, A.D., Jr.; Im, W. Glycan reader: Automated sugar identification and simulation preparation for carbohydrates and glycoproteins. *J. Comput. Chem.* **2011**, *32*, 3135–3141. [[CrossRef](#)]
261. Park, S.-J.; Lee, J.; Patel, D.S.; Ma, H.; Lee, H.S.; Jo, S.; Im, W. Glycan Reader is improved to recognize most sugar types and chemical modifications in the Protein Data Bank. *Bioinformatics* **2017**, *33*, 3051–3057. [[CrossRef](#)]
262. Danne, R.; Poojari, C.; Martínez-Seara, H.; Rissanen, S.; Lolicato, F.; Róg, T.; Vattulainen, I. doGlycans—Tools for Preparing Carbohydrate Structures for Atomistic Simulations of Glycoproteins, Glycolipids, and Carbohydrate Polymers for GROMACS. *J. Chem. Inf. Model.* **2017**, *57*, 2401–2406. [[CrossRef](#)]
263. Bohne, A.; Lang, E.; von der Lieth, C.-W. W3-SWEET: Carbohydrate Modeling by Internet. *J. Mol. Model.* **1998**, *4*, 33–43. [[CrossRef](#)]
264. Bohne, A.; Lang, E.; von der Lieth, C.W. SWEET—WWW-based rapid 3D construction of oligo- and polysaccharides. *Bioinformatics* **1999**, *15*, 767–768. [[CrossRef](#)] [[PubMed](#)]
265. Chernyshov, I.Y.; Toukach, P.V. RESTLESS: Automated translation of glycan sequences from residue-based notation to SMILES and atomic coordinates. *Bioinformatics* **2018**, *34*, 2679–2681. [[CrossRef](#)] [[PubMed](#)]
266. Engelsens, S.B.; Cros, S.; Mackie, W.; Pérez, S. A molecular builder for carbohydrates: Application to polysaccharides and complex carbohydrates. *Biopolymers* **1996**, *39*, 417–433. [[CrossRef](#)]
267. Engelsens, S.B.; Hansen, P.I.; Pérez, S. POLYS 2.0: An open source software package for building three-dimensional structures of polysaccharides. *Biopolymers* **2014**, *101*, 733–743. [[CrossRef](#)]
268. Kuttel, M.; Mao, Y.; Widmalm, G.; Lundborg, M. CarbBuilder: An Adjustable Tool for Building 3D Molecular Structures of Carbohydrates for Molecular Simulation. In *Proceedings of the 2011 IEEE Seventh International Conference on eScience, Stockholm, Sweden, 5–8 December 2011*; pp. 395–402.
269. Kuttel, M.M.; Stahle, J.; Widmalm, G. CarbBuilder: Software for building molecular models of complex oligo- and polysaccharide structures. *J. Comput. Chem.* **2016**, *37*, 2098–2105. [[CrossRef](#)]
270. Clerc, O.; Mariethoz, J.; Rivet, A.; Lisacek, F.; Pérez, S.; Ricard-Blum, S. A pipeline to translate glycosaminoglycan sequences into 3D models. Application to the exploration of glycosaminoglycan conformational space. *Glycobiology* **2019**, *29*, 36–44. [[CrossRef](#)]
271. Singh, A.; Montgomery, D.; Xue, X.; Foley, B.L.; Woods, R.J. GAG Builder: A web-tool for modeling 3D structures of glycosaminoglycans. *Glycobiology* **2019**, *29*, 515–518. [[CrossRef](#)]

272. Kerzmann, A.; Neumann, D.; Kohlbacher, O. SLICK—Scoring and Energy Functions for Protein–Carbohydrate Interactions. *J. Chem. Inf. Model.* **2006**, *46*, 1635–1642. [[CrossRef](#)]
273. Kerzmann, A.; Fuhrmann, J.; Kohlbacher, O.; Neumann, D. BALLDock/SLICK: A New Method for Protein–Carbohydrate Docking. *J. Chem. Inf. Model.* **2008**, *48*, 1616–1625. [[CrossRef](#)] [[PubMed](#)]
274. van Zundert, G.C.P.; Rodrigues, J.P.G.L.M.; Trellet, M.; Schmitz, C.; Kastiris, P.L.; Karaca, E.; Melquiond, A.S.J.; van Dijk, M.; de Vries, S.J.; Bonvin, A.M.J.J. The HADDOCK2.2 Web Server: User-Friendly Integrative Modeling of Biomolecular Complexes. *J. Mol. Biol.* **2016**, *428*, 720–725. [[CrossRef](#)] [[PubMed](#)]
275. Mottarella, S.E.; Beglov, D.; Beglova, N.; Nugent, M.A.; Kozakov, D.; Vajda, S. Docking Server for the Identification of Heparin Binding Sites on Proteins. *J. Chem. Inf. Model.* **2014**, *54*, 2068–2078. [[CrossRef](#)]
276. Sankaranarayanan, N.V.; Nagarajan, B.; Desai, U.R. So you think computational approaches to understanding glycosaminoglycan–protein interactions are too dry and too rigid? Think again! *Curr. Opin. Struct. Biol.* **2018**, *50*, 91–100. [[CrossRef](#)]
277. Griffith, A.R.; Rogers, C.J.; Miller, G.M.; Abrol, R.; Hsieh-Wilson, L.C.; Goddard, W.A. Predicting glycosaminoglycan surface protein interactions and implications for studying axonal growth. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 13697–13702. [[CrossRef](#)] [[PubMed](#)]
278. Eric, B.; Jed, B.; Neha, G.; Vito, F. GlycoTorch Vina: Improved Docking of Sulfated Sugars Using QM-derived Scoring Functions. *ChemRxiv* **2020**. [[CrossRef](#)]
279. Frank, M. Conformational Analysis of Oligosaccharides and Polysaccharides Using Molecular Dynamics Simulations. In *Glycoinformatics*; Lütteke, T., Frank, M., Eds.; Humana Press: New York, NY, USA, 2015; pp. 359–377.
280. Makeneni, S.; Foley, B.L.; Woods, R.J. BFMP: A Method for Discretizing and Visualizing Pyranose Conformations. *J. Chem. Inf. Model.* **2014**, *54*, 2744–2750. [[CrossRef](#)]
281. Chalmers, G.; Glushka, J.N.; Foley, B.L.; Woods, R.J.; Prestegard, J.H. Direct NOE simulation from long MD trajectories. *J. Magn. Reson.* **2016**, *265*, 1–9. [[CrossRef](#)]
282. Lee, H.S.; Jo, S.; Mukherjee, S.; Park, S.-J.; Skolnick, J.; Lee, J.; Im, W. GS-align for glycan structure alignment and similarity measurement. *Bioinformatics* **2015**, *31*, 2653–2659. [[CrossRef](#)]
283. Lütteke, T.; Frank, M.; von der Lieth, C.-W. Carbohydrate Structure Suite (CSS): Analysis of carbohydrate 3D structures derived from the PDB. *Nucleic Acids Res.* **2005**, *33*, D242–D246. [[CrossRef](#)]
284. Rojas-Macias, M.A.; Lütteke, T. Statistical Analysis of Amino Acids in the Vicinity of Carbohydrate Residues Performed by GlyVicinity. In *Glycoinformatics*; Lütteke, T., Frank, M., Eds.; Humana Press: New York, NY, USA, 2015; pp. 215–226.
285. Marchetti, R.; Perez, S.; Arda, A.; Imberty, A.; Jimenez-Barbero, J.; Silipo, A.; Molinaro, A. Rules of Engagement of Protein–Glycoconjugate Interactions: A Molecular View Achievable by using NMR Spectroscopy and Molecular Modeling. *ChemistryOpen* **2016**, *5*, 274–296. [[CrossRef](#)] [[PubMed](#)]
286. Yu, Y.; Delbianco, M. Conformational Studies of Oligosaccharides. *Chem. Eur. J.* **2020**, *26*, 9814–9825. [[CrossRef](#)]
287. Imberty, A.; Pérez, S. Structure, Conformation, and Dynamics of Bioactive Oligosaccharides: Theoretical Approaches and Experimental Validations. *Chem. Rev.* **2000**, *100*, 4567–4588. [[CrossRef](#)]
288. Wormald, M.R.; Petrescu, A.J.; Pao, Y.-L.; Glithero, A.; Elliott, T.; Dwek, R.A. Conformational Studies of Oligosaccharides and Glycopeptides: Complementarity of NMR, X-ray Crystallography, and Molecular Modelling. *Chem. Rev.* **2002**, *102*, 371–386. [[CrossRef](#)]
289. Lütteke, T. Analysis and validation of carbohydrate three-dimensional structures. *Acta Crystallogr. Sect. D Struct. Biol.* **2009**, *65*, 156–168. [[CrossRef](#)] [[PubMed](#)]
290. Blanco Capurro, J.I.; Di Paola, M.; Gamarra, M.D.; Martí, M.A.; Modenutti, C.P. An efficient use of X-ray information, homology modeling, molecular dynamics and knowledge-based docking techniques to predict protein–monosaccharide complexes. *Glycobiology* **2019**, *29*, 124–136. [[CrossRef](#)]
291. Coxon, B. Chapter 3 Developments in the Karplus Equation as they Relate to the NMR Coupling Constants of Carbohydrates. In *Advances in Carbohydrate Chemistry and Biochemistry*; Academic Press: Cambridge, MA, USA, 2009; Volume 62, pp. 17–82.
292. Widmalm, G. A perspective on the primary and three-dimensional structures of carbohydrates. *Carbohydr. Res.* **2013**, *378*, 123–132. [[CrossRef](#)]
293. Slynko, V.; Schubert, M.; Numao, S.; Kowarik, M.; Aebi, M.; Allain, F.H.T. NMR Structure Determination of a Segmentally Labeled Glycoprotein Using *In vitro* Glycosylation. *J. Am. Chem. Soc.* **2009**, *131*, 1274–1281. [[CrossRef](#)] [[PubMed](#)]

294. Soares, P.A.G.; Queiroz, I.N.L.; Pomin, V.H. NMR structural biology of sulfated glycans. *J. Biomol. Struct. Dyn.* **2017**, *35*, 1069–1084. [[CrossRef](#)] [[PubMed](#)]
295. van Beusekom, B.; Lutteke, T.; Joosten, R.P. Making glycoproteins a little bit sweeter with PDB-REDO. *Acta Crystallogr. Sect. F Struct. Biol. Commun.* **2018**, *74*, 463–472. [[CrossRef](#)]
296. Frenz, B.; Ramisch, S.; Borst, A.J.; Walls, A.C.; Adolf-Bryfogle, J.; Schief, W.R.; Veessler, D.; DiMaio, F. Automatically Fixing Errors in Glycoprotein Structures with Rosetta. *Structure* **2019**, *27*, 134–139. [[CrossRef](#)]
297. Bagdonas, H.; Ungar, D.; Agirre, J. Leveraging glycomics data in glycoprotein 3D structure validation with Privateer. *Beilstein J. Org. Chem.* **2020**, *16*, 2523–2533. [[CrossRef](#)]
298. Casañal, A.; Lohkamp, B.; Emsley, P. Current developments in Coot for macromolecular model building of Electron Cryo-microscopy and Crystallographic Data. *Protein Sci.* **2020**, *29*, 1055–1064. [[CrossRef](#)] [[PubMed](#)]
299. Agirre, J.; Davies, G.; Wilson, K.; Cowtan, K. Carbohydrate anomalies in the PDB. *Nat. Chem. Biol.* **2015**, *11*, 303. [[CrossRef](#)] [[PubMed](#)]
300. Hendrickx, J.; Tran, V.; Sanejouand, Y.-H. Numerous severely twisted N-acetylglucosamine conformations found in the protein databank. *Proteins Struct. Funct. Bioinf.* **2020**, *88*, 1376–1383. [[CrossRef](#)]
301. Atanasova, M.; Bagdonas, H.; Agirre, J. Structural glycobiology in the age of electron cryo-microscopy. *Curr. Opin. Struct. Biol.* **2020**, *62*, 70–78. [[CrossRef](#)]
302. Agirre, J. Strategies for carbohydrate model building, refinement and validation. *Acta Crystallogr. Sect. D Struct. Biol.* **2017**, *73*, 171–186. [[CrossRef](#)]
303. Pallesen, J.; Murin, C.D.; De Val, N.; Cottrell, C.A.; Hastie, K.M.; Turner, H.L.; Fusco, M.L.; Flyak, A.I.; Zeitlin, L.; Crowe, J.E.; et al. Structures of Ebola virus GP and sGP in complex with therapeutic antibodies. *Nat. Microbiol.* **2016**, *1*, 16128. [[CrossRef](#)]
304. Lee, J.H.; Ozorowski, G.; Ward, A.B. Cryo-EM structure of a native, fully glycosylated, cleaved HIV-1 envelope trimer. *Science* **2016**, *351*, 1043–1048. [[CrossRef](#)]
305. Pinto, D.; Park, Y.-J.; Beltramello, M.; Walls, A.C.; Tortorici, M.A.; Bianchi, S.; Jaconi, S.; Culap, K.; Zatta, F.; De Marco, A.; et al. Cross-neutralization of SARS-CoV-2 by a human monoclonal SARS-CoV antibody. *Nature* **2020**, *583*, 290–295. [[CrossRef](#)]
306. Walls, A.C.; Park, Y.-J.; Tortorici, M.A.; Wall, A.; McGuire, A.T.; Veessler, D. Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell* **2020**, *181*, 281–292.e6. [[CrossRef](#)]
307. Bubb, W.A. NMR spectroscopy in the study of carbohydrates: Characterizing the structural complexity. *Concepts Magn. Reson. Part A* **2003**, *19A*, 1–19. [[CrossRef](#)]
308. Ardá, A.; Coelho, H.; de Toro, B.F.; Galante, S.; Gimeno, A.; Poveda, A.; Sastre, J.; Unione, L.; Valverde, P.; Cañada, F.J.; et al. Recent advances in the application of NMR methods to uncover the conformation and recognition features of glycans. In *Carbohydrate Chemistry*; The Royal Society of Chemistry: London, UK, 2017; Volume 42, pp. 47–82.
309. Ardá, A.; Jiménez-Barbero, J. The recognition of glycans by protein receptors. Insights from NMR spectroscopy. *Chem. Commun.* **2018**, *54*, 4761–4769. [[CrossRef](#)] [[PubMed](#)]
310. Valverde, P.; Quintana, J.I.; Santos, J.I.; Ardá, A.; Jiménez-Barbero, J. Novel NMR Avenues to Explore the Conformation and Interactions of Glycans. *ACS Omega* **2019**, *4*, 13618–13630. [[CrossRef](#)] [[PubMed](#)]
311. Yang, M.; Angles d’Ortoli, T.; Säwén, E.; Jana, M.; Widmalm, G.; MacKerell, A.D. Delineating the conformational flexibility of trisaccharides from NMR spectroscopy experiments and computer simulations. *Phys. Chem. Chem. Phys.* **2016**, *18*, 18776–18794. [[CrossRef](#)]
312. Säwén, E.; Hinterholzinger, F.; Landersjö, C.; Widmalm, G. Conformational flexibility of the pentasaccharide LNF-2 deduced from NMR spectroscopy and molecular dynamics simulations. *Org. Biomol. Chem.* **2012**, *10*, 4577–4585. [[CrossRef](#)]
313. Turupcu, A.; Blaukopf, M.; Kosma, P.; Oostenbrink, C. Molecular Conformations of Di-, Tri-, and Tetra- $\alpha$ -(2→8)-Linked Sialic Acid from NMR Spectroscopy and MD Simulations. *Int. J. Mol. Sci.* **2020**, *21*, 30. [[CrossRef](#)]
314. Frank, M.; Collins, P.M.; Peak, I.R.; Grice, I.D.; Wilson, J.C. An unusual carbohydrate conformation is evident in *Moraxella catarrhalis* oligosaccharides. *Molecules* **2015**, *20*, 14234–14253. [[CrossRef](#)]
315. Burley, S.K.; Berman, H.M.; Kleywegt, G.J.; Markley, J.L.; Nakamura, H.; Velankar, S. Protein Data Bank (PDB): The single global macromolecular structure archive. In *Protein Crystallography*; Wlodawer, A., Dauter, Z., Jaskolski, M., Eds.; Humana Press: New York, NY, USA, 2017; pp. 627–641.

316. Groom, C.R.; Bruno, I.J.; Lightfoot, M.P.; Ward, S.C. The Cambridge Structural Database. *Acta Crystallogr. Sect. B Struct. Sci. Cryst. Eng. Mater.* **2016**, *72*, 171–179. [[CrossRef](#)]
317. De Meirelles, J.L.; Nepomuceno, F.C.; Peña-García, J.; Schmidt, R.R.; Pérez-Sánchez, H.; Verli, H. Current Status of Carbohydrates Information in the Protein Data Bank. *J. Chem. Inf. Model.* **2020**, *60*, 684–699. [[CrossRef](#)]
318. Lütteke, T.; von der Lieth, C.W. Data Mining the PDB for Glyco-Related Data. In *Glycomics: Methods and Protocols*; Packer, N.H., Karlsson, N.G., Eds.; Humana Press: Totowa, NJ, USA, 2009; pp. 293–310.
319. Agirre, J.; Davies, G.J.; Wilson, K.S.; Cowtan, K.D. Carbohydrate structure: The rocky road to automation. *Curr. Opin. Struct. Biol.* **2017**, *44*, 39–47. [[CrossRef](#)]
320. Crispin, M.; Stuart, D.I.; Jones, E.Y. Building meaningful models of glycoproteins. *Nat. Struct. Mol. Biol.* **2007**, *14*, 354. [[CrossRef](#)]
321. Joosten, R.P.; Lütteke, T. Carbohydrate 3D structure validation. *Curr. Opin. Struct. Biol.* **2017**, *44*, 9–17. [[CrossRef](#)]
322. Speciale, G.; Thompson, A.J.; Davies, G.J.; Williams, S.J. Dissecting conformational contributions to glycosidase catalysis and inhibition. *Curr. Opin. Struct. Biol.* **2014**, *28*, 1–13. [[CrossRef](#)]
323. Fushinobu, S. Conformations of the type-1 lacto-N-biose I unit in protein complex structures. *Acta Crystallogr. Sect. F Struct. Biol. Commun.* **2018**, *74*, 473–479. [[CrossRef](#)] [[PubMed](#)]
324. Zardecki, C.; Shao, C.; Feng, Z.; Westbrook, J.D.; Persikova, I.; Burley, S.K.; Young, J.Y. Collaborating with Glycoscience Community To Improve Data Representation of Carbohydrates in the Protein Data Bank. *FASEB J.* **2020**, *34*, 1. [[CrossRef](#)]
325. Van Beusekom, B.; Wezel, N.; Hekkelman, M.L.; Perrakis, A.; Emsley, P.; Joosten, R.P. Building and rebuilding N-glycans in protein structure models. *Acta Crystallogr. Sect. D Struct. Biol.* **2019**, *75*, 416–425. [[CrossRef](#)]
326. Lütteke, T.; von der Lieth, C.W. pdb-care (PDB carbohydrate residue check): A program to support annotation of complex carbohydrate structures in PDB files. *BMC Bioinf.* **2004**, *5*, 69. [[CrossRef](#)] [[PubMed](#)]
327. Brunger, A.T.; Adams, P.D.; Clore, G.M.; DeLano, W.L.; Gros, P.; Grosse-Kunstleve, R.W.; Jiang, J.-S.; Kuszewski, J.; Nilges, M.; Pannu, N.S.; et al. Crystallography & NMR System: A New Software Suite for Macromolecular Structure Determination. *Acta Crystallogr. Sect. D Struct. Biol.* **1998**, *54*, 905–921.
328. Brunger, A.T. Version 1.2 of the Crystallography and NMR system. *Nat. Protoc.* **2007**, *2*, 2728–2733. [[CrossRef](#)]
329. Emsley, P.; Brunger, A.T.; Lütteke, T. Tools to Assist Determination and Validation of Carbohydrate 3D Structure Data. In *Glycoinformatics*; Lütteke, T., Frank, M., Eds.; Humana Press: New York, NY, USA, 2015; pp. 229–240.
330. Feng, Y. Compatible topologies and parameters for NMR structure determination of carbohydrates by simulated annealing. *PLoS ONE* **2017**, *12*, e0189700. [[CrossRef](#)] [[PubMed](#)]
331. Emsley, P.; Crispin, M. Structural analysis of glycoproteins: Building N-linked glycans with Coot. *Acta Crystallogr. Sect. D Struct. Biol.* **2018**, *74*, 256–263. [[CrossRef](#)] [[PubMed](#)]
332. Agirre, J.; Iglesias-Fernandez, J.; Rovira, C.; Davies, G.J.; Wilson, K.S.; Cowtan, K.D. Privateer: Software for the conformational validation of carbohydrate structures. *Nat. Struct. Mol. Biol.* **2015**, *22*, 833–834. [[CrossRef](#)] [[PubMed](#)]
333. Vařeková, R.S.; Jaiswal, D.; Sehnal, D.; Ionescu, C.-M.; Geidl, S.; Pravda, L.; Horský, V.; Wimmerová, M.; Koča, J. MotiveValidator: Interactive web-based validation of ligand and residue structure in biomolecular complexes. *Nucleic Acids Res.* **2014**, *42*, W227–W233. [[CrossRef](#)]
334. Sehnal, D.; Svobodová Vařeková, R.; Pravda, L.; Ionescu, C.-M.; Geidl, S.; Horský, V.; Jaiswal, D.; Wimmerová, M.; Koča, J. ValidatorDB: Database of up-to-date validation results for ligands and non-standard residues from the Protein Data Bank. *Nucleic Acids Res.* **2015**, *43*, D369–D375. [[CrossRef](#)]
335. Aoki-Kinoshita, K.F. *A Practical Guide to Using Glycomics Databases*, 1st ed.; Springer: Tokyo, Japan, 2017; 370p.
336. Tsuchiya, S.; Aoki, N.P.; Shinmachi, D.; Matsubara, M.; Yamada, I.; Aoki-Kinoshita, K.F.; Narimatsu, H. Implementation of GlycanBuilder to draw a wide variety of ambiguous glycans. *Carbohydr. Res.* **2017**, *445*, 104–116. [[CrossRef](#)]
337. Mehta, A.Y.; Cummings, R.D. GlycoGlyph: A glycan visualizing, drawing and naming application. *Bioinformatics* **2020**, *36*, 3613–3614. [[CrossRef](#)]
338. Varki, A.; Cummings, R.D.; Aebi, M.; Packer, N.H.; Seeberger, P.H.; Esko, J.D.; Stanley, P.; Hart, G.; Darvill, A.; Kinoshita, T.; et al. Symbol nomenclature for graphical representations of glycans. *Glycobiology* **2015**, *25*, 1323–1324. [[CrossRef](#)]



339. Neelamegham, S.; Aoki-Kinoshita, K.; Bolton, E.; Frank, M.; Lisacek, F.; Lütteke, T.; O'Boyle, N.; Packer, N.; Stanley, P.; Toukach, P.; et al. Updates to the Symbol Nomenclature For Glycans (SNFG) Guidelines. *Glycobiology* **2019**, *29*, 620–624. [CrossRef]
340. Lal, K.; Bermeo, R.; Perez, S. Computational tools for drawing, building and displaying carbohydrates: A visual guide. *Beilstein J. Org. Chem.* **2020**, *16*, 2448–2468. [CrossRef]
341. Damerell, D.; Ceroni, A.; Maass, K.; Ranzinger, R.; Dell, A.; Haslam, S.M. The GlycanBuilder and GlycoWorkbench glycoinformatics tools: Updates and new developments. *Biol. Chem.* **2012**, *393*, 1357–1362. [CrossRef] [PubMed]
342. Akune, Y.; Hosoda, M.; Kaiya, S.; Shinmachi, D.; Aoki-Kinoshita, K.F. The RINGS Resource for Glycome Informatics Analysis and Data Mining on the Web. *OMICS* **2010**, *14*, 475–486. [CrossRef]
343. Alocci, D.; Suchánková, P.; Costa, R.; Hory, N.; Mariethoz, J.; Svobodová Vařeková, R.; Toukach, P.; Lisacek, F. SugarSketcher: Quick and intuitive online glycan drawing. *Molecules* **2018**, *23*, 3206. [CrossRef]
344. Cheng, K.; Zhou, Y.; Neelamegham, S. DrawGlycan-SNFG: A robust tool to render glycans and glycopeptides with fragmentation information. *Glycobiology* **2017**, *27*, 200–205. [CrossRef]
345. Cheng, K.; Pawlowski, G.; Yu, X.; Zhou, Y.; Neelamegham, S. DrawGlycan-SNFG and gpAnnotate: Rendering glycans and annotating glycopeptide mass spectra. *Bioinformatics* **2020**, *36*, 1942–1943. [CrossRef] [PubMed]
346. Hypercube, Inc. HyperChem. Available online: <http://www.hyper.com/?tabid=360> (accessed on 31 July 2020).
347. Schrödinger, Inc. The PyMOL Molecular Graphics System. Available online: <https://pymol.org/2/> (accessed on 31 July 2020).
348. Dalby, A.; Nourse, J.G.; Hounshell, W.D.; Gushurst, A.K.I.; Grier, D.L.; Leland, B.A.; Laufer, J. Description of several chemical structure file formats used by computer programs developed at Molecular Design Limited. *J. Chem. Inf. Comput. Sci.* **1992**, *32*, 244–255. [CrossRef]
349. Callaway, J.; Cummings, M.; Deroski, B.; Esposito, P.; Forman, A.; Langdon, P.; Libeson, M.; McCarthy, J.; Sikora, J.; Xue, D.; et al. Protein Data Bank Contents Guide: Atomic Coordinate Entry Format Description. *Brookhaven Natl. Lab.* 1996. Available online: [https://cdn.rcsb.org/wwpdb/docs/documentation/file-format/PDB\\_format\\_Dec\\_1996.pdf](https://cdn.rcsb.org/wwpdb/docs/documentation/file-format/PDB_format_Dec_1996.pdf) (accessed on 20 December 1996).
350. Bohne, A. PDB2MultiGIF: A Web Tool to Create Animated Images of Molecules. *J. Mol. Model.* **1998**, *4*, 344–346. [CrossRef]
351. Sayle, R.A.; Milner-White, E.J. RASMOL: Biomolecular graphics for all. *Trends Biochem. Sci.* **1995**, *20*, 374–376. [CrossRef]
352. Willighagen, E.; Howard, M. Fast and Scriptable Molecular Graphics in Web Browsers without Java3D. *Nat. Prec.* **2007**. [CrossRef]
353. Rose, A.S.; Hildebrand, P.W. NGL Viewer: A web application for molecular visualization. *Nucleic Acids Res.* **2015**, *43*, W576–W579. [CrossRef]
354. Rose, A.S.; Bradley, A.R.; Valasatava, Y.; Duarte, J.M.; Prlič, A.; Rose, P.W. NGL viewer: Web-based molecular graphics for large complexes. *Bioinformatics* **2018**, *34*, 3755–3758. [CrossRef]
355. Sehnal, D.; Deshpande, M.; Vařeková, R.S.; Mir, S.; Berka, K.; Midlik, A.; Pravda, L.; Velankar, S.; Koča, J. LiteMol suite: Interactive web-based visualization of large-scale macromolecular structure data. *Nat. Methods* **2017**, *14*, 1121–1122. [CrossRef]
356. Sehnal, D.; Rose, A.; Koca, J.; Burley, S.; Velankar, S. Mol: Towards a Common Library and Tools for Web Molecular Graphics. In *Workshop on Molecular Graphics and Visual Analysis of Molecular Data, Brno, Czech Republic, 4 June 2018*; Byska, J., Krone, M., Sommer, B., Eds.; Eurographics Association: Geneva, Switzerland, 2018; pp. 29–33.
357. Consortium, T.U. UniProt: A worldwide hub of protein knowledge. *Nucleic Acids Res.* **2019**, *47*, D506–D515. [CrossRef]
358. Kuttel, M.; Gain, J.; Burger, A.; Eborn, I. Techniques for visualization of carbohydrate molecules. *J. Mol. Graph. Modell.* **2006**, *25*, 380–388. [CrossRef] [PubMed]
359. Cross, S.; Kuttel, M.M.; Stone, J.E.; Gain, J.E. Visualisation of cyclic and multi-branched molecules with VMD. *J. Mol. Graph. Modell.* **2009**, *28*, 131–139. [CrossRef]
360. Eborn, I.; Burger, A.; Kuttel, M.; Gain, J. *Carbohydrate: Rendering Carbohydrate Cartoons*; University of Cape Town: Cape Town, South Africa, 2004.
361. Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual molecular dynamics. *J. Mol. Graph.* **1996**, *14*, 33–38. [CrossRef]

362. Perez, S.; Tubiana, T.; Imberty, A.; Baaden, M. Three-dimensional representations of complex carbohydrates and polysaccharides—SweetUnityMol: A video game-based computer graphic software. *Glycobiology* **2015**, *25*, 483–491. [[CrossRef](#)]
363. Besançon, C.; Guillot, A.; Blaise, S.; Dauchez, M.; Belloy, N.; PrévotEAU-Jonquet, J.; Baud, S. New visualization of dynamical flexibility of N-Glycans: Umbrella Visualization in UnityMol. In Proceedings of the 2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Madrid, Spain, 3–6 December 2018; pp. 291–298.
364. Besançon, C.; Guillot, A.; Blaise, S.; Dauchez, M.; Belloy, N.; PrévotEAU-Jonquet, J.; Baud, S. Umbrella Visualization: A method of analysis dedicated to glycan flexibility with UnityMol. *Methods* **2020**, *173*, 94–104. [[CrossRef](#)] [[PubMed](#)]
365. Sehnal, D.; Grant, O.C. Rapidly Display Glycan Symbols in 3D Structures: 3D-SNFG in LiteMol. *J. Proteome Res.* **2019**, *18*, 770–774. [[CrossRef](#)]
366. Thieker, D.F.; Hadden, J.A.; Schulten, K.; Woods, R.J. 3D implementation of the symbol nomenclature for graphical representation of glycans. *Glycobiology* **2016**, *26*, 786–787. [[CrossRef](#)] [[PubMed](#)]
367. Pettersen, E.F.; Goddard, T.D.; Huang, C.C.; Couch, G.S.; Greenblatt, D.M.; Meng, E.C.; Ferrin, T.E. UCSF Chimera—A visualization system for exploratory research and analysis. *J. Comput. Chem.* **2004**, *25*, 1605–1612. [[CrossRef](#)] [[PubMed](#)]
368. McNicholas, S.; Potterton, E.; Wilson, K.S.; Noble, M.E.M. Presenting your structures: The CCP4mg molecular-graphics software. *Acta Crystallogr. Sect. D Struct. Biol.* **2011**, *67*, 386–394. [[CrossRef](#)] [[PubMed](#)]
369. McNicholas, S.; Agirre, J. Glycoblocks: A schematic three-dimensional representation for glycans and their interactions. *Acta Crystallogr. Sect. D Struct. Biol.* **2017**, *73*, 187–194. [[CrossRef](#)]
370. Pendrill, R.; Jonsson, K.H.M.; Widmalm, G. Glycan synthesis, structure, and dynamics: A selection. *Pure Appl. Chem.* **2013**, *85*, 1759. [[CrossRef](#)]
371. Varki, A.; Cummings, R.D.; Esko, J.D.; Freeze, H.H.; Stanley, P.; Marth, J.D.; Bertozzi, C.R.; Hart, G.W.; Etzler, M.E. Symbol nomenclature for glycan representation. *Proteomics* **2009**, *9*, 5398–5399. [[CrossRef](#)] [[PubMed](#)]
372. Perez, S.; Aoki-Kinoshita, K.F. Development of Carbohydrate Nomenclature and Representation. In *A Practical Guide to Using Glycomics Databases*; Aoki-Kinoshita, K.F., Ed.; Springer: Tokyo, Japan, 2017; pp. 7–25.

**Publisher’s Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).