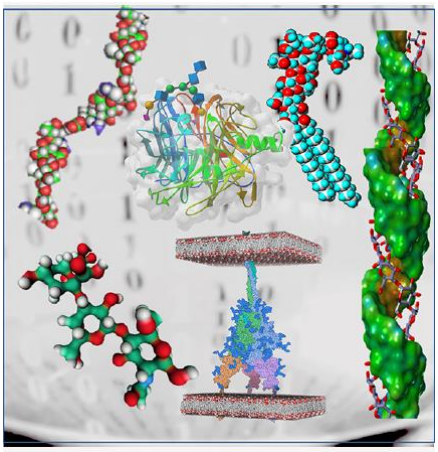
Multifaceted Computational Modeling in Glycoscience

Serge Perez, Olga Makshakova

Multifaceted Computational Modeling in Glycoscience

Serge Perez* and Olga Makshakova

ABSTRACT: Glycoscience assembles all the scientific disciplines involved in studying various molecules and macromolecules containing carbohydrates and complex glycans. Such an ensemble involves one of the most extensive sets of molecules in quantity and occurrence since they occur in all microorganisms and higher organisms. Once the compositions and sequences of these molecules are established, the determination of their three-dimensional structural and dynamical features is a step toward understanding the molecular basis underlying their properties and functions. The range of the relevant computational methods capable of addressing such issues is anchored by the specificity of stereoelectronic effects from quantum chemistry to mesoscale modeling throughout molecular dynamics and mechanics and coarse-grained and docking calculations. The Review leads the reader through the detailed presentations of the applications of computational modeling. The illustrations cover carbohydrate−carbohydrate interactions, glycolipids, and N- and O-linked glycans, emphasizing their role in SARS-CoV-2. The presentation continues with the structure of polysaccharides in solution and solid-state and lipopolysaccharides in membranes. The full range of protein-carbohydrate interactions is presented, as exemplified by carbohydrate-active enzymes, transporters, lectins, antibodies, and glycosaminoglycan binding proteins. The final section features a list of 150 tools and databases to help address the many issues of structural glycobioinformatics.



CONTENTS

1. INTRODUCTION

Glycoscience is a rapidly emerging area of research that deals with carbohydrates' fundamental and applied investigations. Plants, algae, and photosynthetic bacteria produce glucose from sunlight, carbon dioxide, and water. As a result of photosynthesis, carbohydrates are the most abundant biological macromolecules and materials on Earth. This biomass helps reduce carbon footprints while offering a sustainable alternative to fossil resources for energy and raw materials. The simplest carbohydrates are transformed into complex elements such as polysaccharides, complex glycans, and glycoconjugates. Glycans are further complexed with proteins and lipids as cellular components of tissues and organs. However, unlike nucleic acids and proteins, these biosynthesis events do not follow a simple template. Their transformations are driven and influenced by many factors, including cellular metabolism, cell type, developmental stage, and environmental factors, such as nutrients. These factors and their resulting structures allow for considerable diversity and make studying glycans challenging.

Glycoscience is an interdisciplinary field involving informatics, biochemistry, polymer chemistry, materials science engineering, physiology, developmental biology, microbiology, medicine, and ecology. These disciplines are confronted with the unique complexity of the glycans. The prefix "glyco" (sweet in Greek) is often used before the designation of the discipline to indicate the field's specificity. In such a broad field of discipline and the expected advances in fundamental knowledge and translational applications, the determination of the threedimensional (3D) structure of complex carbohydrates, carbohydrate polymers, and glycoconjugates are critical. It aims to understand the molecular basis underlying the structures, associations, and glycan interactions. Glycans, being expressed in all microorganisms and all higher organisms, have a profound impact on biology and medicine. The dramatic roles of spike protein glycans in mediating virus−receptor interactions and antibody production illustrate protein glycosylation's paramount importance. Several other facets, such as vaccines for antimicrobial resistance, the development of innate and adaptive immunity, the growth and spread of cancer, modulation of the gut microbiome and health, responses to bacterial and viral infections, and the development of diseases such as diabetes, and the development of rationally designed carbohydrate-derived drugs enlighten how understanding and exploiting structural glycoscience and glycomics have become necessary for modern molecular diagnostic precision medicine. With an interest in sustainable bioenergy, much effort is devoted to studying cellulose and other polysaccharides. The so-called biomass recalcitrance, i.e., the resistance of the biomass to degradation to individual fermentable glucose residues, generates fundamental and applied investigations about the 3D structure of cellulosic material and the enzymes and microorganisms capable of achieving such degradation. Developing bioinspired sustainable materials from carbohydrates defines a strategy to produce new materials with unique properties from waste resources. Polysaccharides such as cellulose, chitin, starch, and pectins have properties for various applications developed in the circular bioeconomy.

Once the compositions and sequences are established, the main goal of structural glycoscience is to determine the 3D structural and dynamical features of complex carbohydrates, carbohydrate polymers, and glycoconjugates, along with an understanding of the molecular basis of their associations and interactions. However, except for a few cases, these molecules have a high propensity for ample conformational flexibility, producing a multiplicity of conformations that coexist in equilibrium in solution. The use of several spectroscopic methods, such as NMR, with appropriate temporal or spatial resolution, provides invaluable experimental data that require the contribution of molecular modeling to be fully interpreted. In the case of polysaccharides and polysaccharide-based materials, the lack of sufficient diffraction data prevents a direct determination of their 3D structure. Ironically, their elucidation was part of the first steps of computer modeling in carbohydrate research, from which there was the realization that carbohydrates require particular needs on top of those usually recognized when modeling proteins or other organic molecules. These needs are embedded in major computational applications arising from several stereoelectronic effects. As such, they provide realistic simulations of complex carbohydrates in their natural environment, solvated in water or organic solvent, as well as in concentrated solutions or interactions with protein receptors. The large-scale structures incorporating glycan require multiscale computational levels, affording quantum chemical models, molecular mechanics and molecular dynamics models, coarse-grained models, and mesoscale models. Figure 1 shows a time scale versus length scale representation of various computational methods applied to investigate a wide range of glycans and glycan-containing biomacromolecules.

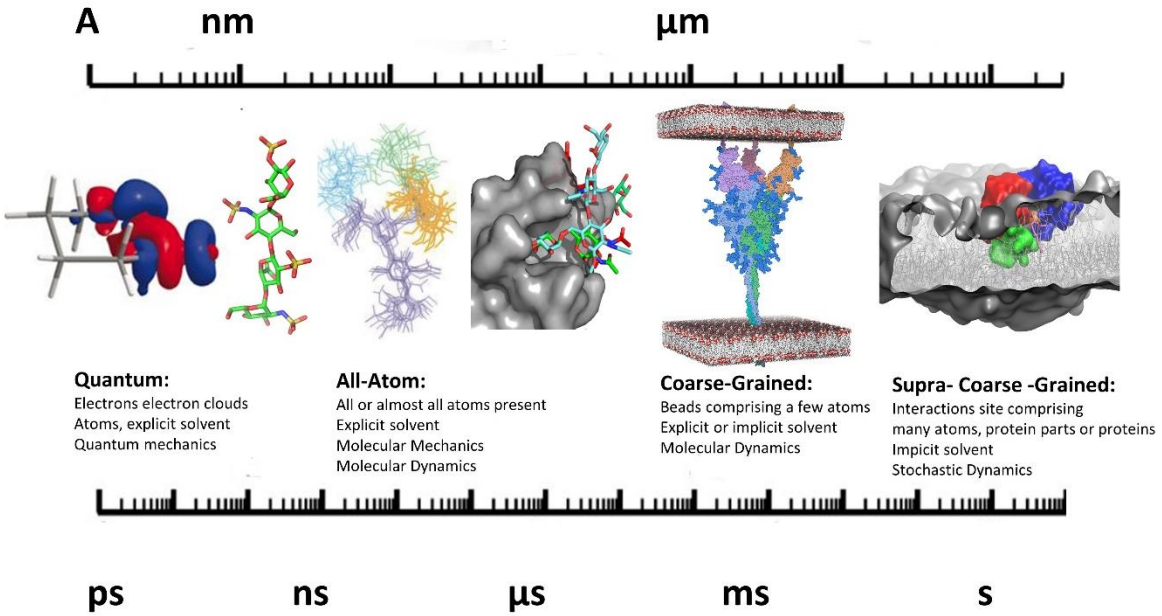


Figure 1. Spatial and temporal multiscale modeling in glycoscience. Applications range from molecular modeling at different resolutions: quantum; all atoms; coarse-grained, and supra-coarse-grained. The axes display approximate ranges of time scales and system sizes.

This chapter is one of a series of contributions that have, over the recent years, reviewed and analyzed the role of simulation and theoretical methods applied to carbohydrates.[1–8] Recent developments have broadened the nature and size of the substrates. In composing this Review, we considered the growing interest in glycans on the part of communities that had remained aloof from this often-considered complex discipline. The present Review is part of the thematic issue on glycobiology/glycochemistry. It describes how the particular features that characterize carbohydrates, in general, can be investigated through several computational applications and are becoming part of the major software and suites used in the general modeling of biomacromolecules and related systems. With a focus on glycochemistry and glycobiology, the Review covers glycans structures at all levels, from monosaccharides to polysaccharides at the mesoscale, including chemical reactivity. The full range of protein–carbohydrate interactions is presented, as exemplified by carbohydrate-active enzymes, transporter, lectins, antibodies, and glycosaminoglycans. We conclude our review by presenting the new developments in glyco-bioinformatics that guide, using a rich tool-box, any practitioner to display, navigate, and quest for correlations between structure and function in glycobiology.

2. STRUCTURAL DIVERSITY AND CONFORMATIONAL CHALLENGES

2.1. Nomenclature and Structural Representation

The individual units from which carbohydrates are made are the monosaccharides. They are the monomeric constituents of an assembly of sugars, either free or attached to another molecule or macromolecule. Throughout this Review, the term "glycan" will designate the several constituting members: mono-, oligo-, polysaccharides, and glycoconjugates. The term oligosaccharide refers to molecules having between 2 and 10 monosaccharides linked together either linearly or branched. Polysaccharide stands for glycan chains built up from more than ten monosaccharides. The glycoconjugate term implies the existence of a covalent linkage between the glycan chains and the proteins (glycoproteins), lipids (glycolipids), and naturally occurring aglycones (e.g., alkaloids, saponins, antibiotics).

The glycans family is rich in several hundred monosaccharide components that characterize it as the most diverse biological family. Monosaccharides, having a suitable number of carbon atoms (>5), can exist as open and cyclic structures. Pentoses form five-membered furanose rings, whereas hexoses form mostly six-membered pyranose rings. In an aqueous medium, monosaccharides with hydroxyl and carbonyl functional groups undergo an intramolecular cyclic hemiacetal formation. Under weak acidic or alkaline conditions, the equilibrium accelerates the formation and favors cyclic forms, and "open chain" forms occur only in trace amounts. The anomeric carbon atom, denoted as C1 carbon, becomes chiral upon cyclization, giving rise to either an α- or β-anomer. Interconversion between unbound monosaccharide α- and β- forms can occur via the acyclic forms. The hydroxyl groups of monosaccharides can undergo a series of chemical modifications: methylation, esterification (acyl esters, phosphate, sulfate esters, N-acetamido groups), and deoxygenation to form deoxysugars. The description and management of many constituting units require well-established, robust nomenclatures (IUPAC Nomenclature home page, http://www.chem.qmw.ac.uk/iupac; IUBMB Nomenclature home page, http:// www.chem.qmw.ac.uk/iubmb) used for representations. Among the several descriptions, a graphical representation called SNFG (Symbol Nomenclature for Glycans) is becoming widely used, resulting from a joint international agreement.9 Such an extension and utilization of the graphical depiction of glycans is a remarkable milestone that offers a way to unify a community in exchange for communicating with other communities and describing the diversity of glycan structures pictorially.

The structural information encoded in the SNFG representation is insufficient to characterize, construct, and manipulate 3D structures. For a given (D) or (L) configuration, monosaccharides can occur as α-pyranose, β-pyranose, α-furanose, β- furanose, i.e., eight different 3D structures. A way to cope with such a requirement follows an extension initially presented in Glycopedia.10 It requires a limited set of rules (as illustrated in Figure 2) that provide the required extension to construct 3D structures and allow the encoding of bioinformatics manipulations while maintaining IUPAC nomenclature.11,12

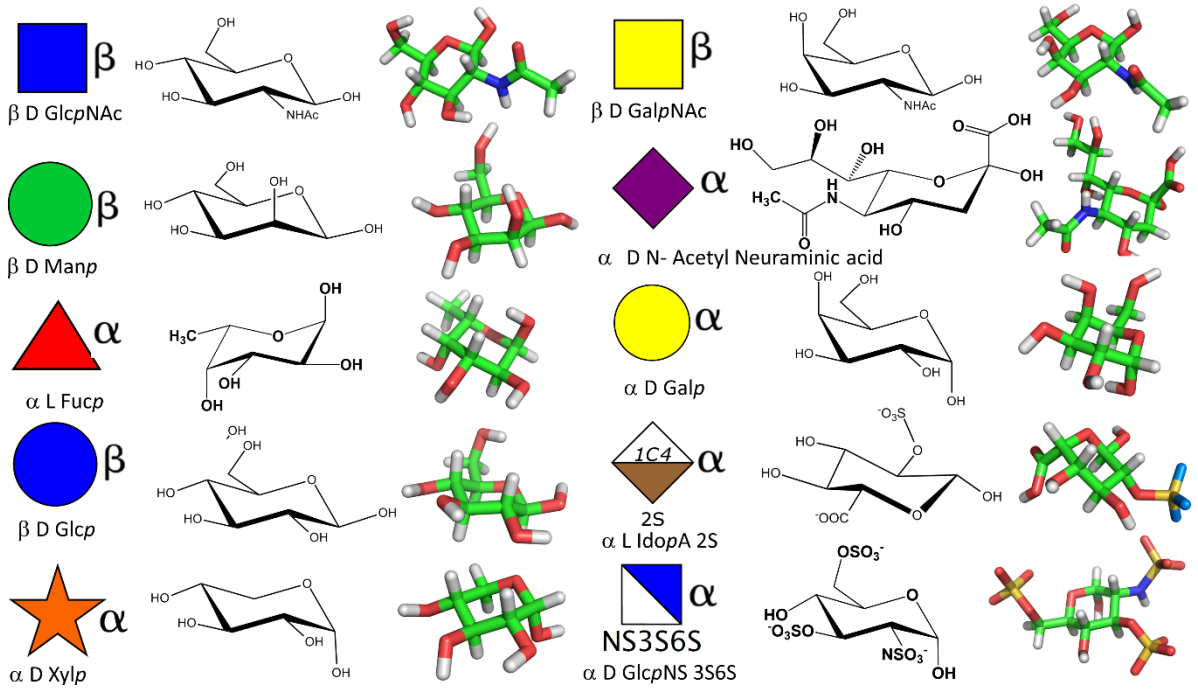


Figure 2. From symbol representation to 3D structures and its extension to 3D structures. Extension of the SNFG cartoons includes the nature of the absolute (D or L) and anomeric (α or β) configurations and the O-esters and ethers, which are attached to the symbol with a number (e.g., 3S for 3-O sulfate groups). All pyranoses in the D configuration are assumed to have the $^4C_1$ chair conformation, whereas those in the L configuration are assumed to have the $^1C_4$ chair conformation. The descriptors of the ring conformations adopted by idopyranoses ($^1C_4$, $^4C_1$, and $^2S_0$) appear within the monosaccharide symbol. The addition of the ring conformation in the monosaccharide symbol is the only discrepancy with the version of the SNFGnomenclature (drawings made using ChemDraw13 and Pymol14).

2.2. Stereoelectronic Effects: Nature and Origins

Because of the differences in the anomeric configurations (α or β), stereochemistry (D or L), and variety of possible linkages (regiochemistry), the resulting number of different glycans and complex polysaccharides with the same sequence is staggering. Interestingly, not all possible sequences occur naturally, indicating that specific glycoforms may have distinct biological functions.

Although less frequently addressed by an average practitioner, critical difficulties are the different ring shapes that might take either five- or six-membered rings due to conformational flexibility. A set of two or three parameters (the so-called puckering parameters) describe five- and six-membered rings, respectively. In the case of five-membered rings, the pseudo-rotational wheel represents the 20 twists and envelope shapes. As for six-membered rings, three parameters control the shape of the ring, which are most conveniently plotted on the Cremer–Pople sphere (Figure 3).

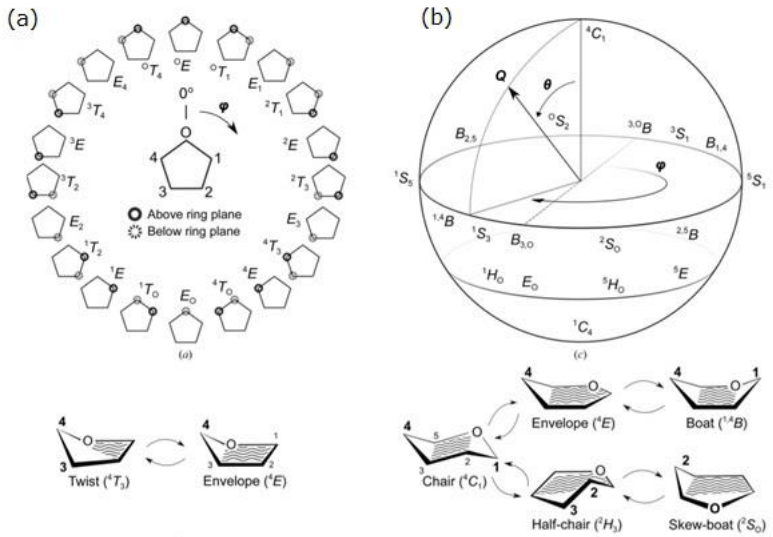


Figure 3. Puckering parameters. (A) Conformational free energy landscape of an isolated β-D-glucose monosaccharide in Stoddard representations. (B) Stoddard represents the puckering sphere for six-membered rings, depicting three low-energy conformations: $^4C_1$, $^1C_4$, and $^2S_0$. (Reproduced with permission from ref 16. Copyright 2017 International Union of Crystallography).

Several challenges for modeling carbohydrate molecules lie in their highly polar functionalities and existing differences in electronic arrangements. The anomeric and exoanomeric effects and the gauche effects dictate conformational and configurational changes (Figure 4). The anomeric effect is the tendency of electronegative substituents at anomeric carbon atoms to be more populated at an axial position than anticipated from the analogy with cyclohexane derivatives.

The origin of the anomeric effect occurs through the hyperconjugation between the molecular np orbital of the O5 ring oxygen atom and the σ^* of the C1–O1 covalent bond (see review by Perez and Tvaroska2 for detailed analysis and calculation of the anomeric effect). The exoanomeric effect dictates the preference of the aglycon around the C1–O1 glycosidic bond. It results from the hyperconjugation between the np of O1 and the σ^* of the O5–C1 bond. There exists a reverse anomeric effect

. In contrast, in a molecular segment X–C–C–Y (where X and Y are electronegative atoms), the so-called gauche effect arises from the tendency of a molecule to adopt the structure with several synclinal interactions between adjacent electron pairs or polar bonds. This effect occurs in conjunction with solvation interactions in the case of the primary hydroxyl group in hexopyranoses.15 These orbital effects and electrostatic interactions are indispensable in the chemical synthesis of carbohydrates since they influence the reactivity of building blocks.

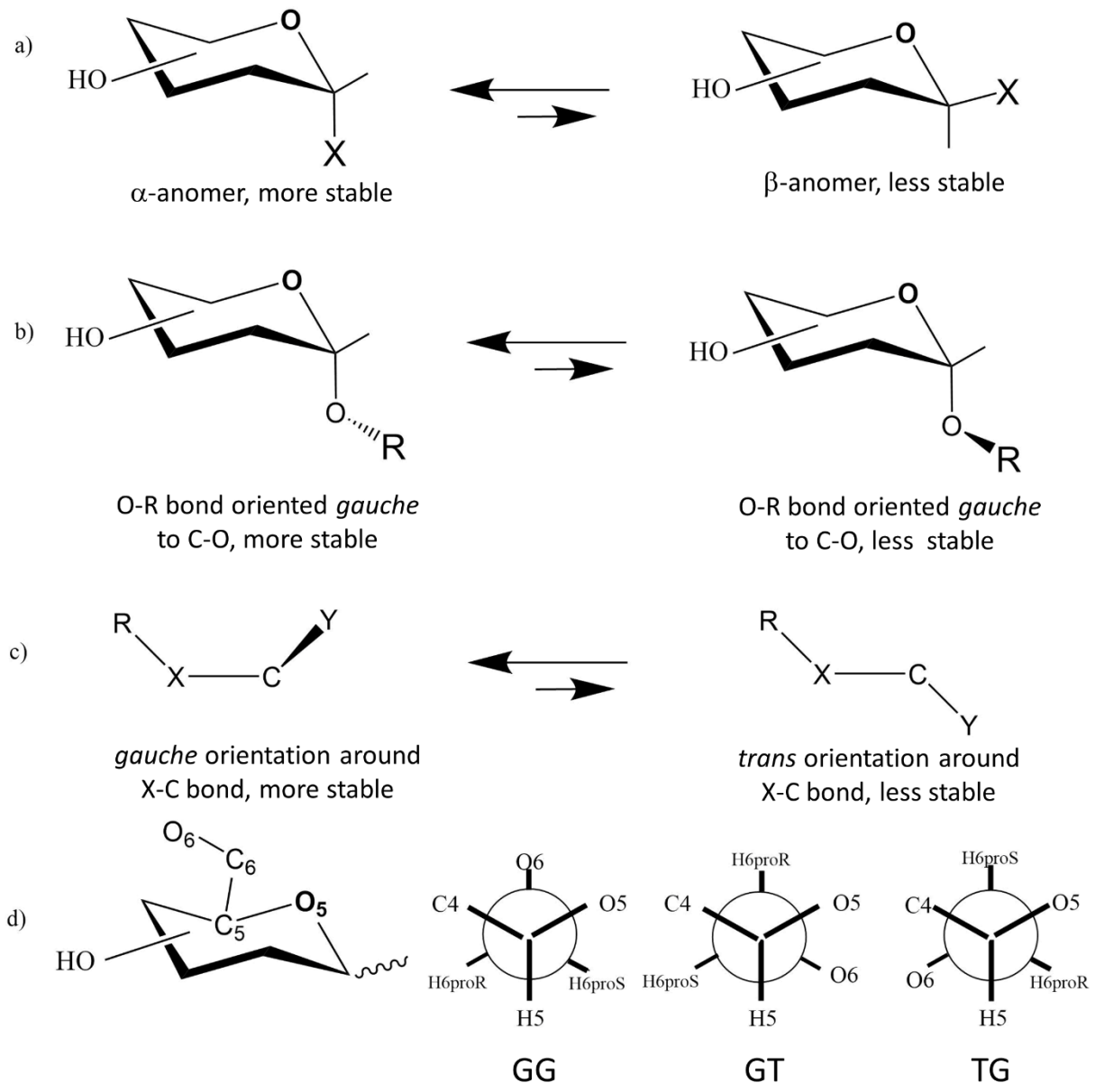


Figure 4. Schematic representation of stereoelectronic effects in carbohydrates. (a) Anomeric effect, (b) exoanomeric effect, and (c) gauche effect. (d) Impact on the conformations of a pyranose's exocyclic primary hydroxyl group is bound to the C6 carbon atom (whereas the secondary hydroxyl groups are linked to the ring carbon atoms). Three staggered situations about the CH−CH_2OH bond (the C5−C6 bond in aldohexoses) can be considered for such an exocyclic hydroxymethyl group. The CH−CH_2OH bond is prochiral; the two hydrogen atoms must be differentiated based on the R/S system. The description of the three rotamers orientations of O6 concerning O5 and C4 is GG (gauche−gauche), GT (gauche−trans), and (trans−gauche) depending on the choice of an atom of ref 17 (drawings made using ChemDraw13).

2.3. Conformational Flexibility: Stereoelectronic Effects and Others

Both the gauche effect and the exoanomeric effect extend to disaccharides and higher oligo- and polysaccharides and impact their conformational preferences. Other effects contribute to van der Waals, hydrogen bonding, and electrostatic interactions, solvation effects, and CH$-\pi$ interactions while interacting with proteins.

The configurational feature of the glycosidic linkage between two monosaccharides is characterized by the axial and equatorial nature of the bonds (for the pyranose ring) that constitute the glycosidic linkage. The bonds' axial and equatorial orientations creating a glycosidic linkage determine a unique stereochemical structure. For example, two hexopyranoses in the $^4C_1$ ring shape can be linked in a 1-1, 1-2, 1-3, 1-4, and 1-6 linkage. Each bond involved in the linkage can occur in either an axial or equatorial configuration. Twenty different types exist for disaccharides made up of $^4C_1$ monosaccharides. Within a disaccharide, the relative orientation of two contiguous monosaccharides linked by a glycosidic bond is characterized by two torsion angles and three in the case of α (1-6), α (2-8), and some other linkages. The energetically favorable conformations of a disaccharide resulting from such values are plotted on potential energy maps called (Φ, Ψ) maps. In contrast to the Ramachandran plot used to visualize the backbone dihedral angles of the constituent amino acids in proteins, (Φ, Ψ) maps have to be established for each disaccharide. They feature multiple minima, separated by 5$-$10 kcal/mol energy barriers (Figure 5)

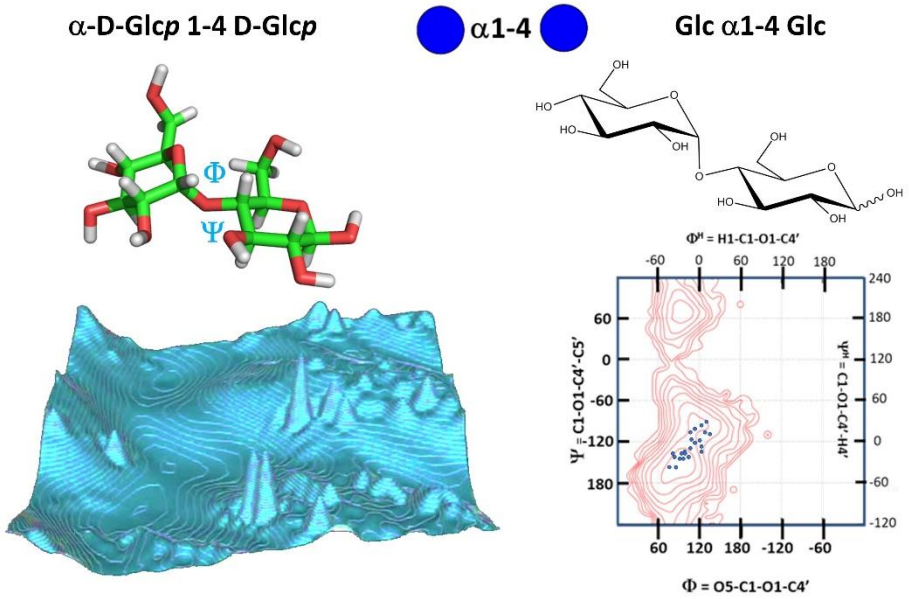


Figure 5. Molecular representation of the structural descriptors of a disaccharide exemplified on maltose. The bottom part of the figure displays two representations of the potential energy landscape as computed from molecular mechanics calculations and projected onto the Φ and Ψ potential energy surface. The iso-energy contours are calculated at 1 kcal/mol intervals regarding the lowest energy minimum. Experimentally observed conformations in the crystalline state are indicated as dots (drawn with ChemDraw,13 PyMol,14 CICADA,18 and CAT19).

Besides the ring conformational flexibility, many hydroxyl groups are the source of several rotatable bonds. Their orientations relative to the sugar ring generate hydrophilic patches (formed by polar hydrogens) and hydrophobic patches (formed by nonpolar aliphatic hydrogens), resulting in an anisotropic solvent density around the molecules.

The considerable conformational flexibility of carbohydrates results in various possible relaxation processes. The time scale of the processes is in the order of 100 ns for interconversion between the various ring puckering transition, 1 ns to 100 ps for the primary and secondary hydroxyl groups, respectively. In contrast, the time of interconversion of the rotation about the glycosidic linkages is about 10 ns.7

3. COMPUTATIONAL CONCEPTS AND TOOLS

3.1. From Quantum Chemistry to Coarse-Grained Calculations

The characterization of carbohydrate structural and dynamic features constitutes a challenge from both the theoretical and experimental points of view. The following theoretical models and their approximations delineate the scope of applications of computational methods to carbohydrate structures and dynamic features. These methods encompass different approaches, ranging from ab initio to coarse-grained methods and from deterministic to heuristic approaches.

They are as follows: ·

- Ab initio simulations based on density functional theory (DFT) ·
- Hybrid quantum mechanics/molecular mechanics (QM/ MM) and QM/QM ·
- Semiempirical methods ·
- Molecular mechanics (MM) and molecular dynamics (MD) calculations ·
- Heuristic methods (Monte Carlo, genetic algorithm-based methods) ·
- Coarse-grained (CG) methods ·
- Docking calculations

Computational methods are applied extensively to explore the conformational space of glycans and polysaccharides based on energy functions and parameter sets appropriate for carbohydrate specificity. Some of them can treat carbohydrates in interactions with proteins considering solvation, and others are more amenable to addressing polysaccharide-based materials.3 Over the last two decades, molecular dynamics became the choice method in their all-atom and coarse-grained representations. Theoretical and technological advances are often used with diffraction, high-resolution spectroscopy, and other spectroscopic methods. They provide a way to reconcile the experimentally available data and predict structural and dynamic features that might not be accessible yet. In such a context, molecular modeling should address those questions: ·

- What are the most appropriate force fields and concomitant parameters to use? ·
- What is the most satisfactory and efficient way to travel through conformational hyperspace? ·
- MD simulation: How long is long enough? ·
- What is the appropriate way to calculate, from a modeling study, the spectroscopic observables for which experimental data are available?
-

3.2. Quantum Chemical Methods

It is beyond the scope of the present Review to provide a detailed description of several QM methods successfully applied to carbohydrates: molecular orbitals, DFT, and QM/MM.

The molecular orbital theory describes the electronic structure of molecules using QM, where a wave function, Ψ, describes the behavior of an electron in a molecule. Electrons around atoms in molecules are limited to discrete energies. The molecular orbital is the region of space where a valence electron in a molecule is likely to be found. The first QM calculations on carbohydrates dealt with conformational studies of small model compounds. They focused on understanding the stereoelectronic effects. The results obtained yielded appropriate values for parametrization force-field methods for carbohydrates. The ab initio QM calculations depend on (i) the quality of the atomic orbitals used to build the molecular orbitals and (ii) the inclusion of electron correlation effects.

The molecular orbital methods have been progressively replaced by the DFT method. The DFT method is a widely used method to investigate biomolecular systems. DFT uses electron density to describe many-electron systems instead of the wave function used in molecular orbital methods. The accuracy of a DFT calculation relies on the quality of the exchange-correlation functional. Among the most popular functionals for calculating low molecular weight conformational properties is the B3LYP functional. The geometry optimization and conformational sampling are usually performed at the B3LYP/6-31+G* level. A single-point calculation at the B3LYP/6-311++G** level follows for the final conformers.

By considering the electronic structure of a system, QM methods offer a description of the chemical reactions, which is computationally demanding. Their use is restricted to relatively simple systems, such as the

active sites of enzymes. The combination of QM and MM approaches provides reliable electronic structure calculations with a realistic description of the enzyme environment. The critical points of such a QM/MM treatment are the boundaries between the QM and MM regions and the size of the QM region. As QM/MM methods provide, the energy value for a given structure can be combined with the existing techniques that map the configurational space available to (bio)molecular systems. They can be molecular dynamics, metadynamics, or heuristic methods such as Monte Carlo. Because of sampling and accuracy limitations, the obtained results need to be considered with caution.

3.3. The All-Atom Representation in Molecular Mechanics and Molecular Dynamics Computations

Within the approximation that molecular structure and dynamics can be accurately described by Newtonian mechanics principles, atoms are represented as hard, impenetrable spheres characterized by mass and size (van der Waals radius) and electrostatic point charge, different for each specific atom type. The conformational features specific to carbohydrates (stereoelectronic effects, ring puckering, glycosidic linkage flexibility, hydroxyl groups, hydrogen bonds, pH) must be considered to obtain a precise analysis, either in vacuo or in solution.

The force field parameter values determine the molecule's potential energy corresponding to each configuration. The choice of a carefully developed and thoroughly tested force field is at the very core of the validity of any Molecular Mechanics or Molecular Dynamics simulation. Once the choice of force field sets the system's potential energy, the conformational space accessible to the glycan is studied over time through MD simulations. Based on the Newtonian motion equation, simulations generate an ensemble of configurations of the molecule's atoms, calculating a molecular system's time displacement coordinates (trajectory) at a given temperature. The positions and velocities of the atoms are done by integrating Newton's equation of motion in time throughout numerical integration. To guarantee energy conservation, the integration time step is 1 fs, which in practical terms is 1 order of magnitude lower than the fastest bond vibration of the C−H bond stretching mode.[20] Constraining the distance of all C−H bonds and water molecules with schemes such as SHAKE[21] or LINCS[22] allows a time step of 2 fs to be used, thereby increasing the sampling capability. There exist several numerical ways to conduct the integrations. The most common ensembles used to perform molecular simulations are the canonical (constant NVT) and the microcanonical (constant NVE).

After establishing the mechanistic framework for the propagation of the system's dynamics, the crucial issue about any simulation is "*how long is long enough*?" In many instances, the answer to this question is, unfortunately, dependent on the availability of computational resources. Glycans can occupy many conformational states because of their flexibility and dynamic behavior. An exhaustive sampling of the conformational space is needed to obtain a correct statistical distribution of the different 3D structures. Nevertheless, running different MD simulations of the same glycan from all potentially relevant conformations is highly advisable, especially if very flexible bonds are present.

Meanwhile, when studying the dynamics of glycans linked to glycoproteins or bound to lectins/enzymes, it is paramount to realize that the starting conformation may be biasing sampling due to strong protein-carbohydrate interactions that may prevent transition to other relevant conformations.

Coverage of the conformational energy hypersurface may require enhanced molecular dynamics sampling protocols. To enhance sampling relative to a standard molecular dynamics simulation, replica-exchange molecular dynamics (REMD)[23] offers a practical solution to sample the whole conformational space of complex bio-macromolecular systems that can be trapped in local energy states while keeping reasonable simulation times. In REMD, a group of simulations or replicas where variables (e.g., potential energy or temperature) have different values run parallel. Combining the REMDmethod with MD simulation with the Monte Carlo algorithm provides overcoming high-energy barriers and sufficient sampling of the conformational space. Other enhanced sampling methods are available, among which are those with bias potentials, such as metadynamics (MTD, using a history-dependent bias potential) and replica state exchange (RSE-MTD), Hamiltonian replica exchange (HREX),[24] multidimensional swarms enhanced sampling MD.[25] The combination of replica-exchange umbrella sampling (REUS) with Gaussian accelerated molecular dynamics (GaMD),[26] referred to as Gaussian accelerated replica-exchange umbrella sampling (GaREUS), applies to free-energy calculations.

The progress in computational resources allows longer and longer simulations to be performed on several thousand explicit atoms for a total time of up to the microsecond scale. It is a prerequisite to exploring conformational space adequately. Nevertheless, limitations persist, and it is still difficult to compare with

experimental data that occur on a much longer time scale than those derived from NMR experiments (primarily NOE and residual dipolar coupling-based experiments). Most molecular dynamics simulations include water molecules. Ions or other surrounding molecules can also be part of the computations.

3.4. Force Fields for Carbohydrates

Through a pioneer contribution in the broad field of molecular mechanics applied to a wide range of organic molecules, the series of Allinger's molecular mechanics force fields and programs offered early access to carbohydrate conformational analysis. The universal force field MM3[27,28] remains appropriate for modeling, essentially low molecular carbohydrates. However, in the development of force fields for nucleic acids and proteins (AMBER, CHARMM, GROMACS, Tinker), particular adaptation to carbohydrate specificities gave glycan modeling a chance to address modern questions. (Figure 6). The commonly used force fields are CHARMM36, GLYCAM06, GROMOS, and OPLS-AA-SEI.[1,29–35]

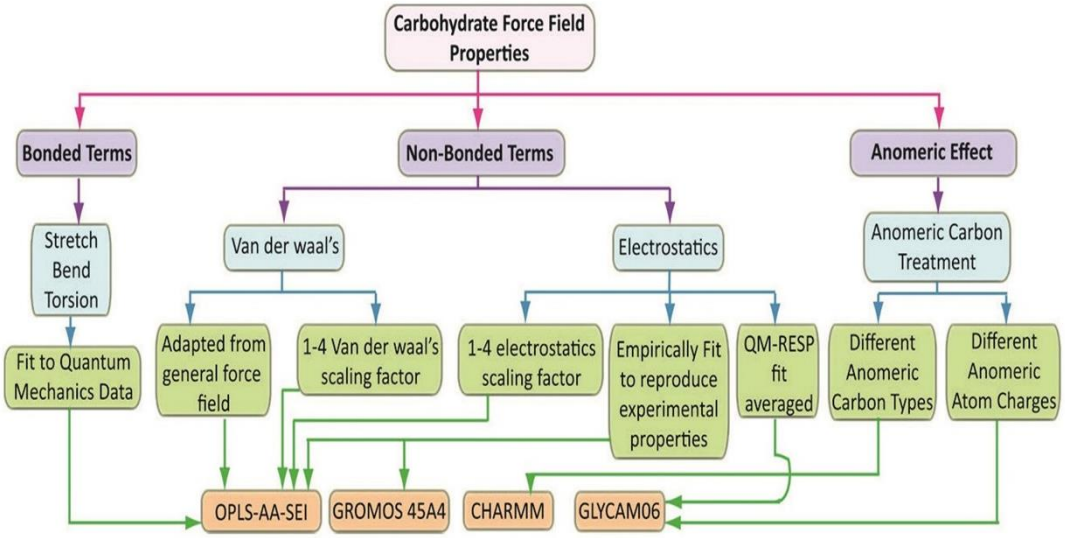


Figure 6. Principle of calculation of the potential energy of a molecule for a molecular mechanics (MM)/molecular dynamics (MD) investigation, along with some parametrization protocol comparison between the carbohydrate force fields. In the MM formalism, the molecule potential energy V(r) is a function of the positions of the N atoms that make up the system. It is represented by an empirical force field, which general functional form is given by

$$V(r) = \sum_{bonds} k_b(r - r_0)^2 + \sum_{angles} k_a(\theta - \theta_0)^2 + \sum_{torsions} \sum_{n} \frac{1}{2} V_n[1 + \cos(n\omega - \gamma)] + \sum_{j=1}^{N-1} \sum_{i=j+1}^{N} f_{ij} \left\{ 4\varepsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right] + \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}} \right\}$$

The first three terms correspond to the potential energy contributions from bonded interactions: covalent bonds, bond angles, and torsion angles. These are all represented by Hooke-type potentials, where classical spring constants are ruled by bond vibrations, angle bending, and dihedral angle potentials (kb/ka). They modulate the stretch from an equilibrium distance (r0) and an equilibrium angle (q0). Fourier series generally represent torsion potentials. The last term is the contribution to the total potential energy from nonbonded (or noncovalent) interactions, dispersion, and electrostatic interactions, represented by Lennard-Jones (LJ) and Coulomb potentials, respectively. The LJ term ε is the energy well's depth at the equilibrium distance σ, while in the Coulombterm q values indicate the partial charges on the atoms i and j, and e0 is the permittivity of the vacuum. The treatment of dispersion interactions with alternative formalisms to the LJ potential, such as the Buckingham potential. It occurs in some software packages and can be used, providing that adhoc parameters are available. For most biomolecular structure and dynamics applications, the LJ potential is sufficiently accurate and generally preferred to the Buckingham more rigorous dispersion treatment for computational speed. Long-range electrostatic forces are treated within periodic boundary conditions within the particle-mesh Ewald (PME) framework. PME electrostatics eliminates artifacts due to the abrupt truncation of the Coulomb potential by introducing cutoff values and/or effects due to the finite size of the simulation box. Reprinted with permission from ref 5. Copyright 2021 Elsevier.

The derivation of the CHARMM all-atom biomolecular force field to carbohydrates into the CHARMM36 force field offers a way to treat monosaccharides in their pyranose[36] and furanose forms,[37] including sulfate and phosphate derivatives,[38] all types of glycosidic linkages, glycans, and glycoproteins, with full provision for dynamic simulation in aqueous media.[39]

Through a proper parametrization of AMBER, the GLYCAM06 force field can cope with a wide range of monosaccharides, including those occurring in glycosaminoglycans. [40] Carbohydrates of all sizes and conformations[41] can be built, and their structures can be investigated. Extensions of the parameters set to glycoproteins, glycolipids,[42,43] lipopolysaccharides, [44] lipids,[45] proteins, and nucleic acids have been reported.

GROMOS, based on a united atom framework, represents a broad family of carbohydrate force field hexopyranose-based saccharides. With the GROMOS engine, several implementations have provided a series of improved parametrizations to account for particular important carbohydrate features. The following describes the main steps of such endeavors and their areas of application: GROMOS 45A4 and GROMOS 53A6GLYC hexopyranose in an explicit solvent,[46] GROMOS 56ACARBO,[47,48] conformational ring equilibria in hexopyranose, [49] GROMOS 56ACARBO_CHT chitosan and its derivatives,[50] GROMOS 56ACARBO/CARBO_R furanose based carbohydrates,[51] and GROMOS96 43A1 glycan structure simulation in glycoproteins.[52,53]

The continuous developments of the all-atom optimized potentials for liquid simulations (OPLS)[54,55] force field provide improvement to correct the performances in estimating the conformational changes around the glycosidic torsion angles and treatment of the carbohydrate-carbohydrate interactions in solution and explicit-water simulations.[56] Developments of a polarizable empirical force field for hexopyranose based on CHARMM Drude are bringing significant improvements in treating a series of monosaccharides[57−59] and their glycosidic linkages.[60−62]

3.5. Coarse-Grained Calculations

Simulating large macromolecular systems or materials for millisecond-scale times requires reducing the atomistic level of representation to reach longer simulation times and larger spatial domains. CG models offer a way to accelerate such a reduction by grouping a given functional group or a molecule into a single "particle". The first coarse-grained model proposed for proteins almost half a century ago.[63] The pioneering CG model of hexopyranose glucose was presented in 2004 by Molinero and Goddard.[64] Terms such as "bead", "pseudo atom", the particle", "interaction site", or "super atoms" are synonyms. As a result, the description of molecules or systems requires a much lower number of components than in an all-atom (AA) representation. The choice of beads is not unique. Chemical intuition or personal experience often guides the choice. Different mapping schemes may exist for the same molecular system. A mapping operator helps determine the coordinates of a CG bead. This determination is made from a linear combination of the coordinates of the atoms that will form the particular bead. The construction comes at the cost of the nonuniqueness of the mapping scheme. It might differ from one system to another or from one researcher to another (Figure 7).

The determined "beads" are usually chosen as the center of mass or the geometric center of a particular group of atoms from the atomistic model. With the progress of computing techniques and machine learning, new methods for constructing and mapping atoms in the CG bead have been developed. This could help eliminate human bias and potentially develop better CG models. The merging schemes vary depending on the specific approach used and the precision required. The least invasive approach is to group the nonpolar hydrogen atoms with the heavy atoms they are bound to. A "united atoms" framework yields up to a 10-fold reduction of the computational effort. Merging schemes can accelerate simulations by three to four orders of magnitude compared to classical all-atom MD simulations.

After mapping the atomistic structure into the desired CG representation, potential functions determining the bonded and nonbonded interactions of the model should be assigned prior to validating these functions against atomistic simulation or experimental results. The use of atomistic simulations to develop CG potentials is known as a bottom-up approach, while the use of experimental data to fit the potentials is called a top-down approach. Several techniques are often used to fit a CG potential based on all-atom simulations, including the inverse Monte Carlo method,[68,69] Boltzmann inversion, relative entropy,[70] force matching (multiscale magnification),[71,72] and the realistic extension algorithm via the Hessian covariance (REACH).[73]

The straightforward way to construct a CG force field is to import the respective all-atom force field expressions. The equations are analogous to those used in classic empirical force field representation. Evaluating

the covalent interactions between beads uses harmonic functions to calculate the potential energy for the virtual bond length bond angles and the virtual torsion angles. Dispersion interaction through a Lennard–Jones term and electrostatic interaction through a Coulomb term account for the noncovalent term. Additional terms in the form of multibody terms are introduced to account for the internal correlations between groups of atoms.74 The effect of such an addition is to smooth the potential energy landscape, thus accelerating the sampling.75–77 Some local energy terms that describe the spatial correlation between pseudo bond vectors need to be considered and processed with arbitrarily chosen functions, as they do not follow the classical harmonic behavior. Finally, water is implicitly represented through continuum electrostatic.78

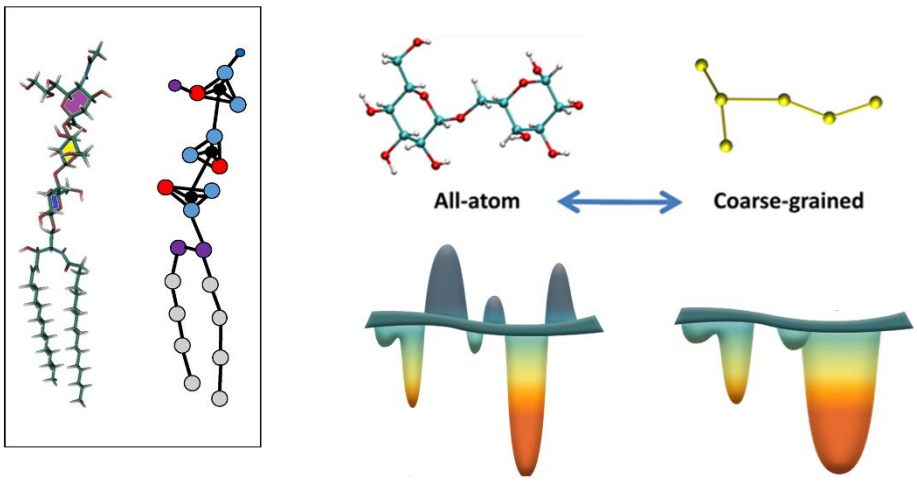


Figure 7. Graining multiple atoms together into “pseudo atoms” or “beads”. Atomistic (on the left) and CG (on the right) representations of the GM3 ganglioside. All-atoms versus coarse-grained energy landscape, showing the effect of smoothing the energy surface and its flattening, enables efficient energy landscape exploration.65–67 An empirical function represents the force field, which general functional form is given by

$$U = \sum_i \frac{1}{2}k_i^d(d_i - d_i^\circ)^2 + \sum_i \frac{1}{2}k_i^\theta(\theta_i - \theta_i^\circ)^2 + \sum_i \sum_n a_i^{(n)}[1 + \cos(n\gamma_i)] + b_i^{(n)}[1 + \sin(n\gamma_i)] + 332\sum_{i<j}\frac{q_i q_j}{Dr_{ij}} + 4\varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] + V_{solv}$$

In the equation, di, di°, and kid are the length, equilibrium length, and force constant of the ith virtual bond, qi, qj, and kid are the actual and equilibrium value and the force constant of the ith virtual bond angle. γi is the ith virtual bond dihedral angle, and ai (n) and bi (n) are the coefficients in the expressions for the torsional potentials. qi is the partial charge on the ith site, D is the relative dielectric permittivity, σij and εij are the constants of the Lennard–Jones potential for the interaction of site i with site j, and rij is the distance between these two sites. The coefficient of 332 in the expression for theCoulombic energy is introduced to express the energy in kcal/mol (if the distance is expressed in ångstroms and the charges are expressed in electroncharge units).

The straightforward way to construct a CG force field is to import the respective all-atom force field expressions. The equations are analogous to those used in classic empirical force field representation. Evaluating the covalent interactions between beads uses harmonic functions to calculate the potential energy for the virtual bond length bond angles and the virtual torsion angles. Dispersion interaction through a Lennard–Jones term and electrostatic interaction through a Coulomb term account for the noncovalent term. Additional terms in the form of multibody terms are introduced to account for the internal correlations between groups of atoms.74 The effect of such an addition is to smooth the potential energy landscape, thus accelerating the sampling.75–77 Some local energy terms that describe the spatial correlation between pseudo bond vectors need to be considered and processed with arbitrarily chosen functions, as they do not follow the classical harmonic behavior. Finally, water is implicitly represented through continuum electrostatic.78

The MARTINI force field is undoubtedly the most popular CG force field.79 It applies to lipid systems, proteins, nucleic acids, glycans and polysaccharides,80 and materials science. It is appropriate for simulating glycolipid membranes and monotopic and transmembrane proteins. The MARTINI force field uses universal modular blocks for which representation can be applied to various biomolecular systems once translated from all-atom to CG. The force field uses a one-to-four mapping approach where a single bead stands for a group of four heavy atoms. In such a scheme, one bead represents four water molecules. A slightly higher resolution of

up to two heavy atoms per bead is needed for small ringlike fragments (such as monosaccharides, aromatic amino acid side chains, and cholesterol). Then according to their nature, the chemical property is attributed to these four coarse-grained blocks as polar (P), nonpolar (N), apolar (C), and charged (Q). Based on hydrogen bonding capability, an additional attribute subdivides them into five classes: donor, acceptor, both or none, and polarity. Their combination gives a total of 18 unique "building blocks".67,81,82 The MARTINI force field follows the CG parameters' iterative fitting that defines bonded interaction potentials to atomistic trajectory data. The development of the parametrization strategy combines both structure-based (or topdown) and thermodynamic-based (or bottom-up) parametrization strategies67,80−84 (see Figure 10). There is no electrostatic field and polarization effect within the selected water treatment (e.g., four water molecules devoted to charges per bead). A uniform dielectric constant compensates for neglecting an explicit polarization, which may be a reasonable approximation for bulk water. Such an implicit screening affects the strength of the interactions between polar substances, which is underestimated in nonpolarizable solvents. Introducing a three-bead model to represent four water molecules provides a polarizable water model, as this effect accounts for a reorientation of the water molecules. Q type represents ions. The first hydration shell of single ions is included in the CG representation, to which the total charge is assigned. The CG ion force field is only qualitatively accurate.

Some caveats result from the losses of the directionality of hydrogen bonding interactions. They prevent using CG and MARTINI in protein folding studies. The loss of atomistic resolution and torsional structure within the CG beads does not allow proper treatment of puckering of the carbohydrate ring or to obtain information on stereoisomers in glycan. An imbalance of the nonbonded solute−solute, solute−water, and water− water interactions can result in nonphysical aggregations of glycans treated by MARTINI.85 In the most recent version of MARTINI 3, some imbalances have been improved, using new bead types and the expanded ability to include specific interactions, for example, hydrogen bonding and electronic polarizability.86 Nevertheless, CG models can efficiently explore complex biomolecular systems' conformational space and time-averaged properties.87

The CG representation perturbs the evaluation of a modeled system's thermodynamic properties, particularly the balance between entropic and enthalpic contributions. The entropic evaluation is perturbed by reducing the degree of freedom inherent to any CG-based modeling, which is somehow counterbalanced by reducing the evaluation of the enthalpic terms. Therefore, a CG model may adequately yield free energy differences with inaccurate entropy and enthalpy values. When applied to carbohydrates, one of the main advantages of a CG model lies in its flexibility of the representation, i.e., the number of "beads" chosen per monosaccharide unit.

Upon going to coarser representations, the more detailed features of the model are sacrificed. The regulation of the thermodynamic contribution can be achieved by appropriately scaling the terms describing the nonbonded interactions.

However, the transferability of the CG model to a wide range of concentrations and molecule types, i.e. glycolipids and glycoproteins, to describe diverse and biologically relevant systems remains essential. Except for the generic MARTINI model, none of the proposed CG force field schemes can describe the precise arrangements of real biosystems.88 Through an intrinsically consistent CG approach, MARTINI can extend the simulation to an extensive range of molecules, including different lipid types, sterols, sugars, peptides, and polymers. Such flexibility is required to investigate complex carbohydrate-based systems (such as glycoconjugates, functionalized glycomaterials, and glycoconjugates) or explore carbohydrate−protein interactions.89

3.6. Heuristic Algorithms

Parallel to the application of deterministic algorithms, several heuristic-based algorithms have been employed in software that aims to explore the complex potential energy space of glycan conformations to extract the location(s) of the most probable conformations and provide a low energy conversion pathway.

The development of a method entitled "Channels in Conformational Space Analyzed by Driver Approach" (CICADA) offered an alternative way to explore complex potential energy surfaces using a single coordinate driving approach.90,91 Despite its relevance and efficiency, this application deals with the difficulties arising from the conformational flexibility of carbohydrates.18,92 Such a heuristic approach has not gained popularity among the developers and practitioners of carbohydrate computational modeling.

The application of a random search method based on the Monte Carlo algorithm, was explored to provide average ensemble parameters. The repetition of the procedure results in many conformations, which should represent the system. Its application to oligosaccharides provided satisfactory results to be used in conjunction with experimental data.93,94 It was advocated that the method would be more efficient for simple molecular systems instead of random displacements.

The field of genetic algorithms (GAs), evolutionary programming, and similar areas of computer science take their inspiration from biological evolution: survival of the fittest, the inheritance of traits from parents to offspring, mutation, crossover, and population genetics. Lamarckian evolution is based on the idea that the parent can pass on acquired traits to the offspring. One genetic algorithm implementation offers different searches. Standard GA and parallel GA with Lamarkian and natural evolution)95,96 addressed the automatic conformation prediction of carbohydrates.97,98 It has been applied to complex bacterial polysaccharides.99−101

3.7. Free Energy Calculations

The results derived from molecular dynamics simulations offer a way of comparison with experimental thermodynamics and dynamic properties. The most critical thermodynamic quantities are the fundamental free energies to compute the accurate estimates of how chemical entities recognize each other, associate and react. There are differences between free energy calculations depending on whether a completely free energy pathway is looked for or a pure free energy difference is present. Therefore the methods calculating equilibrium free energy range from more accurate based on the perturbation theory methods to less accurate end-point methods.

The first groups of methods include free energy perturbation (FEP) and thermodynamic integration (TI). The difference between the initial and the final state is described in terms of intermolecular or intramolecular coordinate changes. They may comprise differences in both geometry and chemical "mutation" of one molecule into the other. The free energy map of such a transition is described by a potential of mean force (PMF). In the case of geometry changes, the PMF will characterize the reaction between a reactant and product and provide a change in coordinate indicative of a conformational pathway with a detailed depiction of barriers, heights, and energy changes. The rigorous simulations of the ligand−protein binding free energy can require an extremely high computational cost because of sufficient sampling of the minimum free energy pathway between the unbound and bound states, including entropic differences. Alchemical binding free energy calculations consider the fundamental principle that free energy is a state function and does not depend on the pathway that connects the end states. In practice, a thermodynamic cycle is constructed utilizing an "alchemical" pathway that is computationally optimal. To this end, a coupling parameter (so-called lambda, λ) is used to make the chemical transformation's thermodynamic path from the initial state to the end state smoother by bringing the neighboring states closer. The TI requires the Boltzmann averaged potential energy calculations at each intermediate state λ. Although alchemical simulations cannot provide a complete kinetic characterization of the binding process, they are highly efficient in predicting the binding affinities critical in drug discovery. Nonetheless, these approaches remain computationally expensive, and their use becomes rational only when the initial-to-final state perturbation is minimal (or negligible), i.e. when it is not accompanied by large conformational changes upon the complex formation.

The principal scheme of the binding energy calculations between a ligand and its host is given in Figure 8.

The second group of methods includes molecular mechanics with generalized Born and surface area solvation (MM/GBSA), molecular mechanics with Poisson−Boltzmann surface area MM/PBSA when the Poisson−Boltzmann method is used instead of GB, and linear interaction energy (LIE). This range can extend to the scoring function widely used to rapidly estimate binding energy in molecular docking.

The development of free energy calculations performed on unbound carbohydrate molecules in vacuo and solution7 addressed essential issues such as the interconversion of sugar ring conformations, rotations of the primary and secondary hydroxyl groups, and rotations about the various glycosidic linkages.102−105 These studies aimed at developing and improving force-field parameters.48,49,51,106,107 The free perturbation methods and advanced replica-exchange sampling methods can enhance the sampling.

The computation of the relative and absolute free energy of noncovalent protein−substrate binding is primarily covered in the literature describing the methods commonly used: MMPBSA, LIE, and absolute alchemical methods. Most investigations dealt with lectins since these proteins bind to specific carbohydrates without modifying them.108−114

Classical force fields for calculations of protein−carbohydrate complexes were selected for evaluation in predicting lectin-carbohydrate binding free energies. 115

Their accuracy is hampered by technical difficulties, such as the conformational flexibility of carbohydrates and the highly polar nature of the complexes. A more extensive and detailed comparison of the three prominent families of force fields (CHARMM, GROMOS, and GLYCAM/AMBER) indicates the conservation of the qualitative patterns of the structural descriptors of the interaction, particularly the reproduction of the CH−π interactions. However, the computed values of the unbinding free energies display deviated differently from the experimental data.116

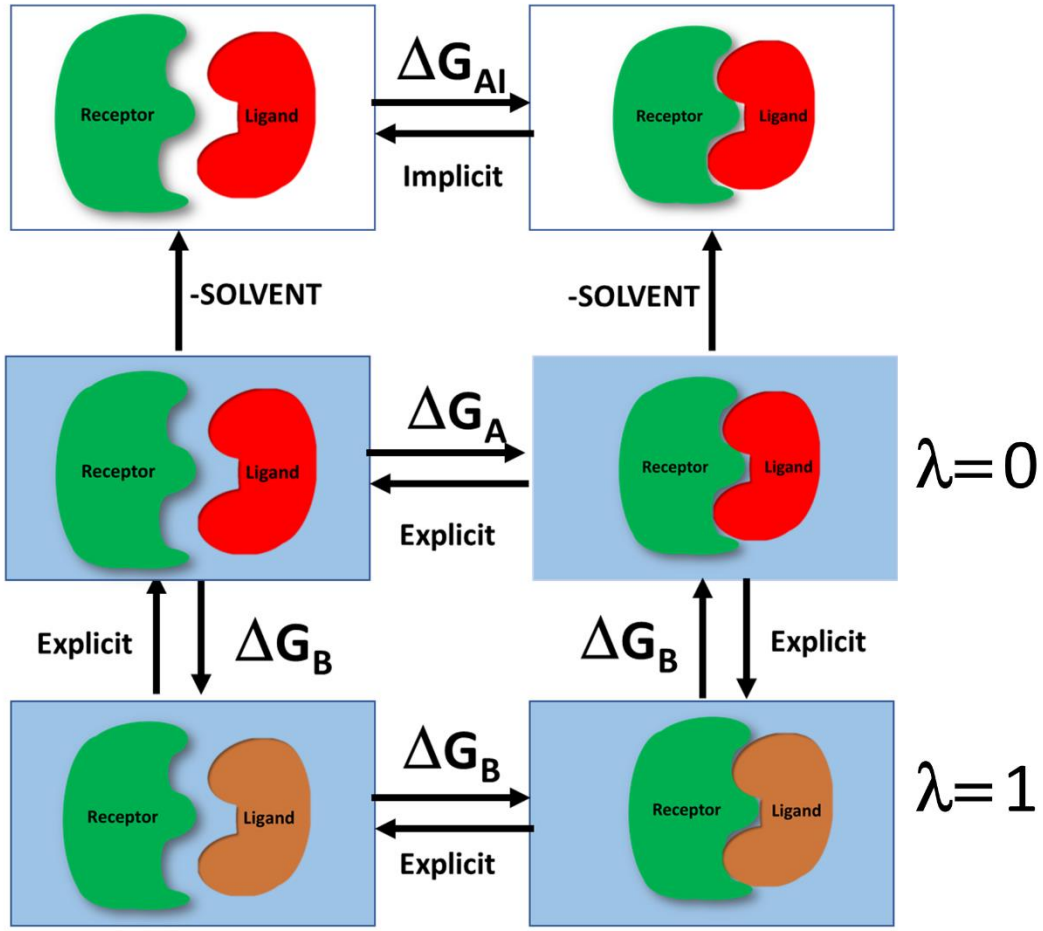


Figure 8. The thermodynamic cycle calculates the free energy of binding between the receptor and ligand. The absolute free energy of binding in an implicit solvent environment is calculated by molecular mechanics Poisson−Boltzmann surface area (MM-PBSA) (ΔGAI). Thermodynamics integration methods evaluate the free-energy binding difference between closely related systems (A and B) (ΔΔG = ΔGC − ΔGD =ΔGA − ΔGB). A gradual transformation of the initial state A to the final state B is described by an additional nonspatial coordinate, λ. The following equation ΔGTI =∫ 01 dλ⟨δV(λ)/δ(λ)⟩λV(A) for λ = 0 and V(B) for λ = 1 is used to evaluate the free energy difference between the states.

3.8. Molecular Docking

Molecular docking aims to explore the ligand's preferred orientation to a receptor to form a stable complex. Each snapshot of the pair is referred to as a pose. Exploring the whole conformational space of the interacting molecules is an essential part of binding free energy simulations. It requires sufficient sampling to explore all accessible poses with MD, a prohibitive computational cost. It is essential to obtain a reasonable initial pose, and in cases where the best starting pose is ambiguous, multiple poses should be searched. Docking programs utilize molecular modeling approaches for protein−glycan complexes. They rely on initial geometry generations, conformational sampling, grafting, and active site mapping and estimate binding affinity.117−121

Choosing the appropriate software for the problem to be studied remains critical to the proposed solution. It is mainly the case when small ligands interact in large and ill-defined binding sites. Docking is a computational method that places a small molecule (ligand) in its macromolecular target's (receptor's) binding site and estimates the binding affinity. Molecular docking requires (at least partial) three-dimensional knowledge of the ligand and receptor of interest. Glycans are usually constructed using molecular mechanics methods or directly from structural databases. Several energy parameters tailored to the energy minimization of protein-glycan complexes are available from different force fields.122

X-ray crystallography and NMR spectroscopy previously solved the receptor structures in most investigated cases. Otherwise, homology modeling, threading, and de novo methods provide the receptor coordinates in the absence of experimentally determined data.

Several docking programs are performing well in the docking of glycans to proteins. They are available as stand-alone software or/and as online services: BALLDock/SLICK123,124 (https:// ball-project.org/download/), GlycoTorch Vina125 (http:// ericboittier.pythonanywhere.com/); HADDOCK Modeling of biomolecular complexes with support of glycosylated proteins126 (https://wenmr.science.uu.nl/), Cluspro Sulfated GAG docking127,128 (https://cluspro.bu.edu/login.php), Vina- Carb129 (http://glycam.org/docs/), and GlycanDock130(https://new.rosettacommons.org/docs/latest/application_ documentation/carbohydrates/GlycanDock).

They work in slightly different ways. They all share two main features: sampling and scoring. Sampling concerns the conformational position and orientation of the ligand in the receptor-binding site. Flexible docking methods account for the pendant groups' possible influence in evaluating the glycans' preferred orientations in the binding sites. Typically, they are concerned with the orientation of the hydroxyl and hydroxymethyl groups or sulfate groups, the network of hydrogen bonds they direct, and the conformational flexibility occurring at each glycosidic linkage.

Most docking programs treat the ligand as flexible while the protein conformation is kept rigid. However, some programs can perform a type of "soft docking". Correctly accounting for receptor flexibility is much more computationally expensive and is not yet common practice. The high conformational flexibility is one of the essential features of complex carbohydrates, along with CH−π interactions. Some programs have provisions for treating such features.125,129−132

Docking algorithms are grouped into deterministic approaches that ensure reproducibility and stochastic approaches. The algorithm includes random factors that do not allow full reproducibility in the latter case. The incremental construction algorithms consist of splitting a ligand into rigid fragments. One of the fragments is selected and placed in the protein binding site. The ligand reconstruction is performed in situ by adding the remaining fragments (as implemented in the DOCK program).133 Among the stochastic search approaches, the genetic algorithm (inspired by evolutionary biology) is implemented in AutoDock. Various other sampling methods are used in docking programs. These include simulated annealing protocols and Monte Carlo simulations. Glide uses a hierarchical algorithm.134

Scoring functions evaluate the best conformation, translation, and orientation (called poses) and rank the ligands. Energy scoring functions evaluate the binding of free energy between proteins and glycans. The affinity is evaluated using the Gibbs−Helmholtz equation. Finally, empirical scoring functions use parametrized terms describing properties critical in protein-glycan binding and formulate an equation for predicting affinities. These terms typically describe polar−apolar interactions, loss of ligand flexibility and desolvation effects.

The hydrogen bonds formed between the carbohydrate ligand and the protein determine the selectivity. Their impact on the binding energy might not be crucial for the free energy of binding135 since these interactions can be substituted by water in the unbound state. In addition, the modest gain of enthalpy can be compensated by the changes in entropy due to water displacement.136,137 Charged residues, ions and

structural water molecules are essential for the carbohydrate-binding.138,139 A correct charge treatment is another crucial point for accurate scoring.1

A distinctive feature of protein−glycan recognition is the interaction between the aromatic side chains of proteins and the C−H bonds of the hydrophobic faces of the constituting monosaccharides, which results in the formation of critical CH−π contacts.124,140−142 However, in the docking procedure, which does not take the solvent into account explicitly, the electrostatic interactions in the protein−carbohydrate complex may contribute to the scoring dominantly when compared to the CH−π interactions.123 As such an effect accounts differently from other interactions, some programs may not adequately consider them.

Widely used docking programs, which take different accounts of these interactions, may not work for the protein−glycan complexes. Various docking programmes and scoring functions perform differently for other targets. Such variable performances could occur for different types of ligands. An accurate determination of carbohydrate−protein complexes remains a nontrivial issue.130,131 The biased methods offer a tool to improve carbohydrate-binding mode prediction.143 The solvent- structure-biased docking considers the similarity in the location of water molecules in the binding sites and the interactions of carbohydrate's OH-groups with the protein in the complex. The solvent binding sites, where the probability of finding a probe atom is higher than that in the bulk solvent, are used as structural knowledge to place the OH− sugar groups.144,145

The protein "mapping" using small molecules of a different physicochemical nature allows guiding of the docking toward pharmacophoric interactions, namely, hydrogen bonds and hydrophobic or aromatic interactions.146−148

In carbohydrate−lectin complexes, difficulties arise from many lectins' shallow, multi-chamber binding sites. The extension of several docking software packages to handle glycans has improved the comprehension of several systems. Many docking studies have been performed in the case of lectin−glycan interactions compared to those complexes determined experimentally by crystallography.149 In contrast, relatively few published docking studies on carbohydrate antibody recognition have been reported150 (http://glycam. org/ad). This situation reflects the limited number of suitable validation tests (i.e., high-resolution carbohydrate−antibody crystal structure complexes) and the inherent difficulty in modeling such systems

. Despite these difficulties arising from the challenges posed by protein−glycan complexes, molecular docking has begun to produce reliable and insightful results. The extension of several docking software packages to handle glycosaminoglycans (GAGs)125,127,151 has improved the comprehension of GAG−protein interactions.152,153

The complete characterization of glycan recognition by proteins requires quantitative measurements of the strength and specificity of the interactions, which can be assessed through the combined use of biophysical methods and computer simulation. Together with docking simulation, NMR methods have addressed several classes of interactions at recognition sites.155

NMR techniques can be based on either protein156,157 or ligand observations.154,158 The glycan receptor binding illustrates the latter by mumps virus hemagglutinin-neuraminidase. This is an example of the combination of an NMR, docking, and molecular modeling search, together with CORCEMA-ST calculations, to unravel the structural features of sialoglycan/MuV-HN complexes. When compared to unbranched ligands, the different activity of the enzyme for longer and more complex substrates has been examined by NMR kinetic analysis154 (Figure 9).

However, many challenges remain, and it remains a nontrivial exercise far from being a turnkey tool. In particular, the ability of docking programs to correctly score docking positions (especially in cases of small ligands in large and ill-defined binding sites) requires a critical inspection of the results.

4. LOW MOLECULAR WEIGHT CARBOHYDRATES

4.1. 3D Structures

Over the years, many simulations of carbohydrates have been carried out. They concern carbohydrates in isolation or the solvated state. Quantum mechanics methods deal with relatively simple model compounds. They offer a proper quantified description of the main stereoelectronic effects and interactions that determine the behavior and properties of these molecules, e.g., mutarotations in solution, optical rotations, glassy states of sugar, the conformational preferences in the gas phase, and aqueous solution and the chemical reactivity. In

a continuous dialogue between quantum mechanics and molecular mechanics methods, evaluating the conformational features of many disaccharides has attracted much attention, and the comparison with experimental data has ascertained the credibility and the limitations of these methods.159 Large oligosaccharides, such as cycloamylose with 26 glucose residues per cycle,160 and RhamnoGalacturonan II, a 30-mer mega-oligosaccharide,161 stand among the examples where molecular modeling revealed insightful architectural features. The application of molecular mechanics methods, often in conjunction with solution NMR experiments, has been repeatedly applied for complex glycans and their fragments.

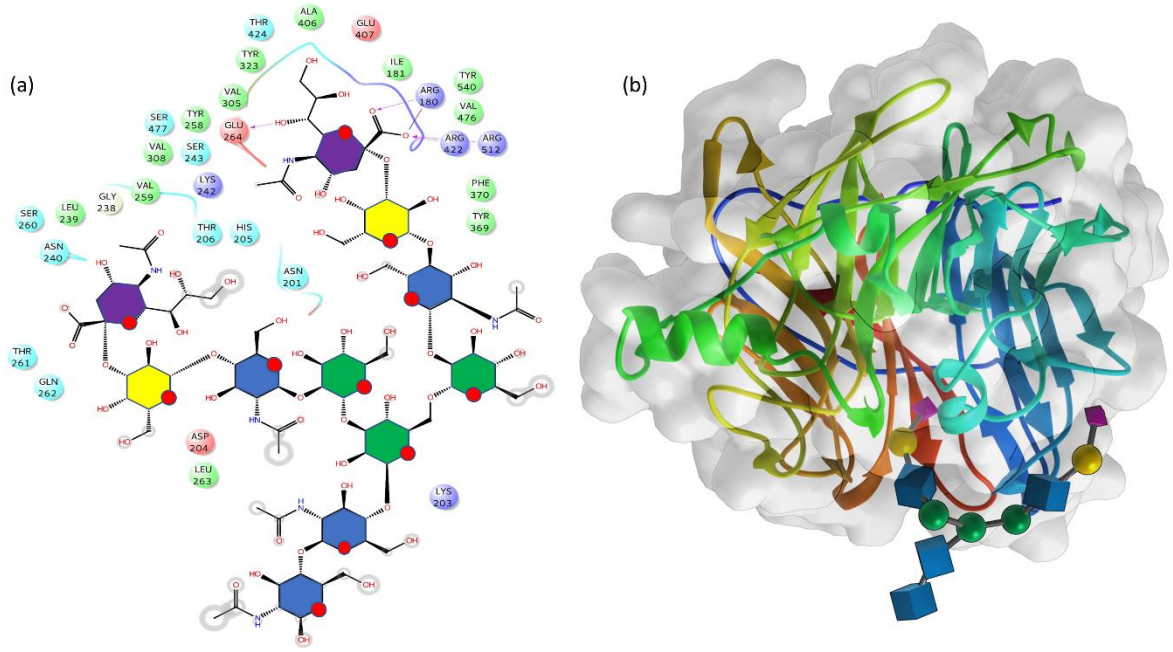


Figure 9. Structural features of the glycan receptor binding by mumps virus hemagglutinin-neuraminidase. (a) 2D plot illustrating the interaction of the sialylated undecasaccharide with the residues in the binding pocket. (b) 3D representation of the undecasaccharide in interactions with the receptor. (Drawn with OneAngstrom, SAMSON from structural data kindly provided by the authors.154)

4.2. Chemical Reactivity

The assembly of two carbohydrate building blocks connected by a glycosidic linkage occurs throughout a glycosylation reaction, a substitution reaction between a nucleophile and an electrophile. In a glycosylation reaction called oxocarbenium ions, the reactive intermediates appear as glycosyl cations. These glycosyl cations are notably different from the parent glycosides due to the change in the hybridization state of an anomeric carbon atom from sp3 to sp2 hybrids and the acquisition of a positive charge. The electron-withdrawing substituents strongly influence the shape and reactivity of the glycosyl donor. The archetypal mechanisms of glycosylation, SN2- and SN1-pathways, are found as extreme cases on the continuum. Stepwise transformations can transform the oxocarbenium ion to a reactive covalent α- or β- intermediate.

The general mechanism for the most commonly employed glycosylation reaction is depicted in Figure 10, showing the step of activating a donor molecule by a promotor system.162−166 Such an activation leads to an array of reactive intermediates. Covalent species can undergo SN2-like substitutions, while cationic intermediates can partake in SN1-like reactions in two possible envelope oxocarbenium ion conformers.

Several computational studies have been undertaken to gauge the effects of multiple substituents on the conformational behavior, reactivity, and stability of pyranosyl oxocarbenium ions.167−174 Whitfield and co-workers have investigated the conformational preference of tetra-O-methyl-gluco- and tetra- O-methyl-mannopyranosyl oxocarbenium ions and their 4,6-Obenzylidene analogues. They studied the conformational behavior of the oxocarbenium ions formed upon dissociation of the triflate-leaving group.170 Hansen and co-

workers have developed a computational approach to investigate the stability and reactivity of glycosyl cations as a function of their shape175 (Figure 11). The developed computational application probes the entire conformational space available to the six-membered ring cations and their representation on the Cremer−Pople sphere. The aim was to establish a conformational energy landscape (CEL) that could be displayed. They constructed and computed at the DFT level all the different conformers arising from a systematic scan of the three intracyclic torsional angles (C1−C2−C3−C4, C3−C4−C5−O5, and C5−O5−C1−C2). All energy is computed at the PCM(CH2CL2)-B3LYP/6- 311G(d,p) level of theory and expressed as the solution-phase electronic energy. The scheme was applied to more than 30 glycosyl oxocarbenium displaying different substitution patterns. The selectivity of the product formed in an SN1 glycosylation reaction could be assessed based on the conformational preference of the intermediate ions. The computed CEL maps could fully display these glycosyl oxocarbenium ions' stability, reactivity, and conformational mobility. The conformational preference of the cations could be directly related to the experimental stereochemical outcome of addition reactions with typical SN1-nucleophiles.

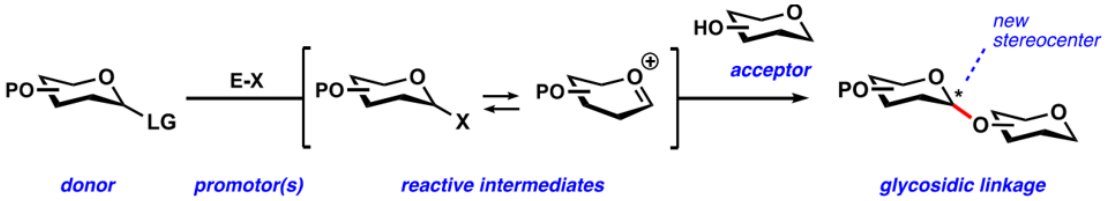


Figure 10. General overview of the glycosylation reaction. Glycosylation reactions are best considered as occurring at a continuum between two formal extremes of the mechanism, the SN1- and SN2-mechanisms (LG, leaving group; P, protective group, E-X promotor system). Reproduced with permission from ref 166, Copyright 2021 Elsevier.

The examination of the CEL maps highlights the influence of the various substituents' electronic effect on the cations' overall shape. Such an influence drives the subsequent stereochemical course of the reactions. Decreasing electronegativity and increasing the size of the substituent translate to a preference to adopt an equatorial position (3H4-like conformations) to minimize steric interactions. This trend was also found for substituents at the C3 position and electronegative substituents at the C2 position. They occur preferentially in a pseudoequatorial position to enable the hyperconjugative stabilization of the oxocarbenium ion by the pseudoaxial C2− H2 bond.

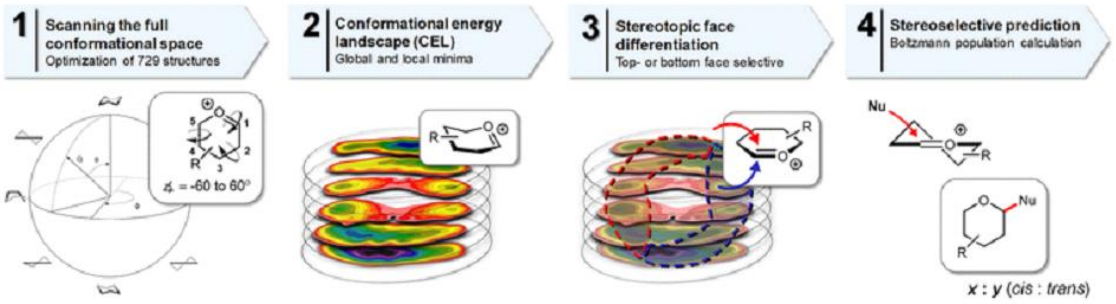


Figure 11. Overview of the workflow to map the conformational and stereoselective preference of pyranosyl oxocarbenium. (1) The entire conformational space of the six-membered ring was established by computing 729 prefixed structures. A few canonical conformations (chair, half-chair, envelope, and boat) are depicted. (2) The associated energies are displayed on slices dividing the Cremer−Popple sphere. (3) Top- and bottom-face of selective conformers in separate areas of the sphere. The family of the top-face-selective (3E, 3H4, E4, and B2,5)-like structures occurs in the area contoured with the red-dashed line, while the bottom face-selective family (4E,4H3, E3, and 2,5B)-like conformers occurs on the opposite side of the sphere, grouped with the blue-dashed line. (4) Based on the Boltzmann distribution of the top- and bottom-face selective structures, the stereochemical outcome of nucleophilic addition to reactions to pyranosyl oxocarbenium ions can be computed. Reproduced from ref 175. Copyright 2019 AmericanChemical Society.

Hansen and co-workers studied the formation of dioxolenium ions by remote acyl groups mounted on the donor molecules C- 3, C-4, or C-6.176 They were established to follow the order: 3- Ac-Man "4-Ac-Gal > 3-Ac-Glc ∼ 3-Ac-Gal > 4-Ac-Glc > 4-Ac- Man ∼ 6-Ac-Glc/Gal/Man. The CEL maps quantitatively explained the experimentally found trends, and the stability of the dioxolenium ions could be linked to the strength of longrange participation. The central role of long-range participation in C-3 acyl mannosyl systems could be traced back to the strong preference for the formation of a dioxolenium ion.

4.3. Carbohydrate–Water Interactions

Carbohydrate hydration results from the balance between intramolecular hydrogen bonding and hydrogen bonding to water molecules, with the water molecules being distributed within several strata, the so-called hydration shells. The shape of some carbohydrates can fit into a network structure of icosahedral clusters of water molecules, with hydrogen bonding, by substituting a chair-form water hexamer in a cluster. Carbohydrate hydration is still not fully established despite its importance in determining the structure, interaction, and function. An ab initio investigation of glucose's first hydration shell structure has been reported. Both α- and β-glucose reach complete hydration shells of 15 water molecules with 28 hydrogen bonds.177 Each hydroxyl group participates in about two hydrogen bonds, forming a weaker acceptor and a stronger donor to water molecules which fit poorly within a locally tetrahedral network.177 This feature may lead to water and glucose clusters by hydrogen bonds in aqueous glucose solutions.177 The equatorial/axial configuration of the hydroxyl groups fundamentally influences the hydration scheme. The occurrence of intramolecular hydrogen-bonding reduces the hydration of the carbohydrate, with a resulting increase in the nonpolar character and its subsequent impact on the biological recognition177

Considering the first hydration shell leads to the question of the total number of water molecules inside the first hydration shell, and whether water molecules can be shared and are a dominant feature of the solute structure,179 molecular dynamics simulation offers straightforward access to the hydration scheme's essential features. It identifies the number of water molecules and their contacts on the first hydration shell (Figure 12). Nevertheless, due to the structural complexity of the solute, the large number of water molecules, and the high degree of mobility, ad hoc treatment requires a statistical approach for an accurate description and analysis. An adequate tool uses a radial pair distribution function that evaluates the probability of finding two atoms at a distance, r, relative to the probability expected for a random distribution at the same density. The radial pair distribution function identifies the significant differences in the first and second hydration shells of different solute atoms. In the case of flexible disaccharides (where the first hydration shell number of water molecules per solute molecule is about 27), the normalized 2D radial pair distribution function is particularly adequate for investigating bridging water molecules between two monosaccharide rings in carbohydrates. 179,180 The calculation of the average residence time for all water molecules around solute atoms of interest is a descriptor of the dynamics that occurred within the first hydration shell. Typically, a residence time of about 0.5(2) ps is characteristic of water−water interactions, which is slightly lower than the average residence time of (0.6 (2)−0.700 (2) ps) for water molecules in their interactions with hydroxyl oxygen of carbohydrates. Notorious exceptions exist, as in sucrose and trehalose, where the interaction between water molecules and some hydroxyl groups is 2.7−3.0 ps.178

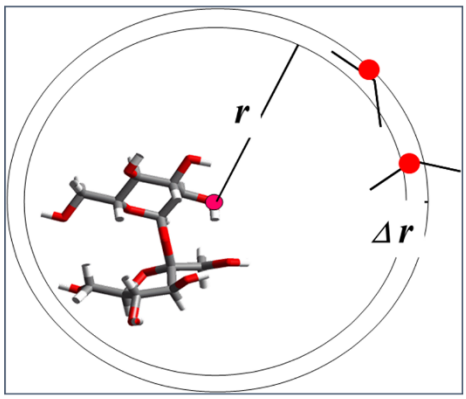
4.4. Carbohydrate−Carbohydrate Interactions

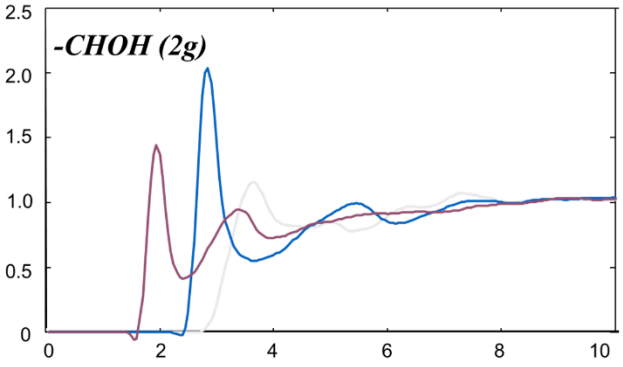
The crystal morphology of low molecular weight carbohydrates is directly related to how these molecules organize in three dimensions, as in the case of sucrose. These features result from the symmetry elements composing the crystalline unit cell and the carbohydrate−carbohydrate interactions. The same principles apply to polysaccharides in the solid state with a supplementary contribution of a multivalent effect since the interactions reproduce along the polysaccharide chains. Such noncovalent interactions between individual polysaccharide chains have been recognized and documented.

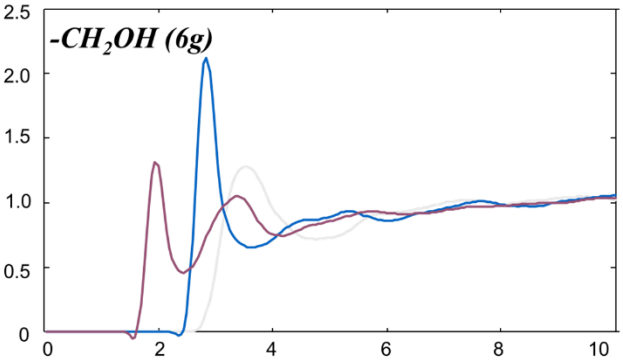
The most prominent surface-exposed structure, carbohydrates, must play a reliable and versatile role in cell adhesion and recognition. Nevertheless, carbohydrate−carbohydrate interactions have been less considered for other biochemically and biologically relevant situations181 since their measurements are challenging to perform. Nevertheless, such interactions result from embryonal cell compaction and aggregation,

myelin compaction, and melanoma cell adhesion.182,183 The adhesion of freshwater sponge cells results from carbohydrate−carbohydrate interactions in a low-calcium environment.184

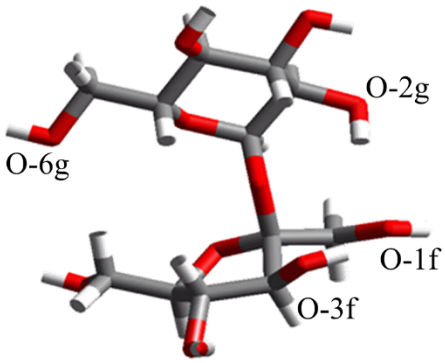
Using extensive atomistic molecular dynamics simulations, the computational characterization of the homophilic carbohydrate carbohydrate-carbohydrate interactions of the trisaccharide LewisX leads to quantifying the association constants and the adhesion energy of the anchored glycolipid in the membrane, which is in good agreement with experimental results, previously reported. The authors observed fuzzy, weak c s

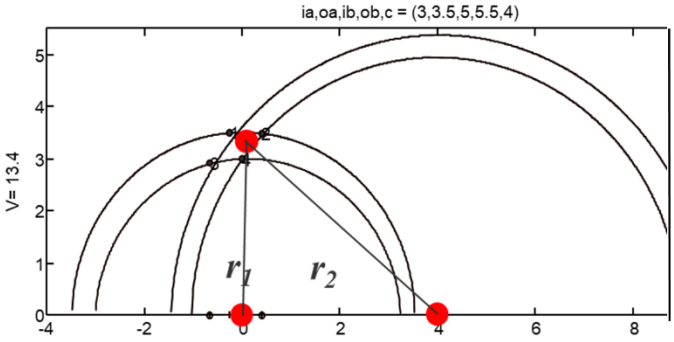






$$g(r)=\frac{N(r)}{\rho_w \times \frac{4}{3}\pi \times [(r+\Delta r)^3 - r^3]}$$





$$g(r_1,r_2)=\frac{N(r_1,r_2)}{\rho_w \times V_{intersect}(r_1,r_2,\Delta r)}$$

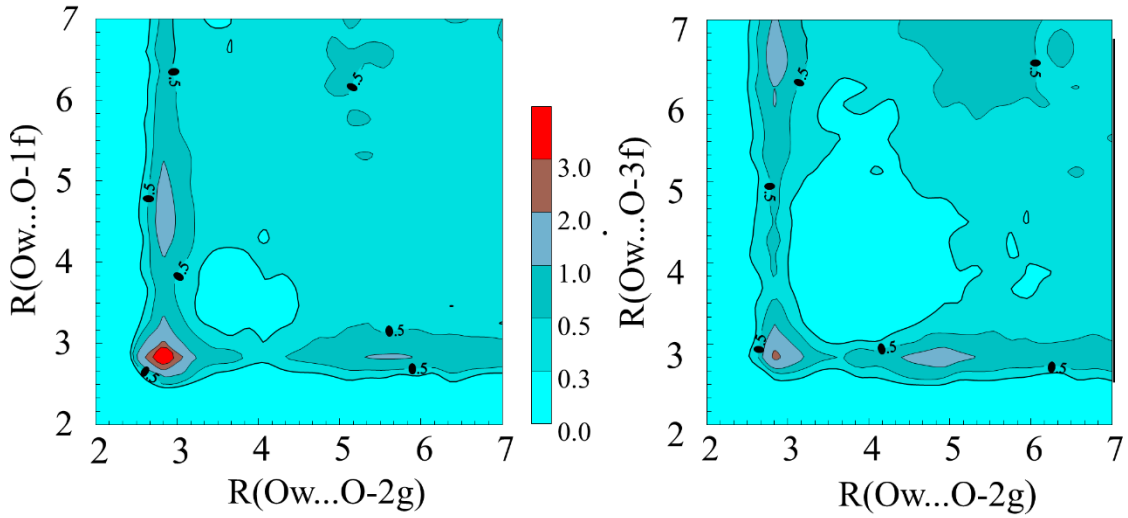


Figure 12. Radial pair distribution function calculated from molecular dynamics trajectory of sucrose. Top: HO-2g···Ow (gray); O-2g···Ow (blue) C-2g···Ow (thin dark). Bottom: O-6g . . . Ow(gray); O-6g···Ow (blue) and C6g···Ow (thin dark). Two-dimensional radial pair distribution functions for sucrose. Left: neighboring intra-ring water density O-2g···Ow···O-3g. Right: inter-ring bridging water density between O-2g···Ow···O-3f. (Adapted from ref 178.)

5. GLYCANS AND GLYCOCONJUGATES

5.1. N- and O-Linked Glycans

Protein glycosylation is one of the major post-translational modifications. More than half of human proteins have the potential to be glycosylated or proved to be glycosylated.186 There are two types of glycosylation where the glycans are linked to N- or O-atoms in the protein side chains. In the N-linked glycosylation, the oligosaccharide is enzymatically transferred from a lipid-linked oligosaccharide to the asparagine residue of a nascent protein in an en bloc manner. Oligosaccharyltransferases operate at the endoplasmic reticulum membrane. The asparagine is located in a consensus sequon Asn-X-Thr/Ser, where X is any residue except proline.187,188 In the O-linked protein glycosylation, a single GalNAc (one of the many Oglycans) is enzymatically attached (by N-acetyl galactosaminyltransferases) to serine or threonine residue, and then it can be further extended or modified on the site. There is no well-defined amino acid sequence for O-glycosylation189 (Figure 13). N-Glycans have a common core architecture, namely, Manα1→6(Manα1→3)Manβ1→4GlcNAcβ1→4GlcNAcβ1→ attached to Asn. The three mannose residues initiate branches or antennae that form arms, linear or branched. The antennae consist of oligomannose glycan, complex glycans (GlcNAc, Gal, Neu5Ac, and Fuc residues), and a mixture (a hybrid type). The antennae length, branching, and composition may vary within each type of glycan. Such a functionalization is case-dependent and specific to health and disease states.190,191 N-Glycan modification regulates the flexibility of the protein to which the glycan is attached and the protein−glycan recognition.

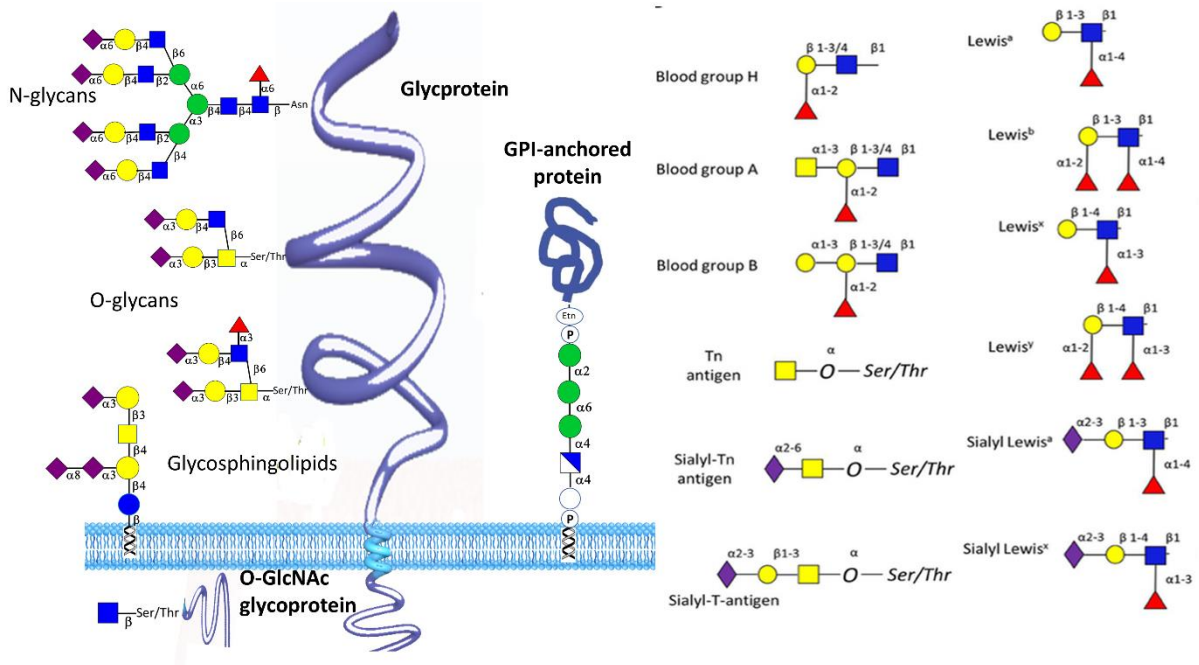


Figure 13. Major types of mammalian glycosylation (glycoproteins, N- and O-glycans, etc.). Schematic depiction of some major antigen determinants displayed according to the SNFG representation.

O- and N-glycans are highly flexible, challenging their conformational sampling. The number of glycosidic torsion angle pairs for a typical biantennary N-glycan of complex type is 22.192. The linkage position and

branching may affect the glycosidic bond and pyranose ring flexibility, implying the variability of glycan motions and their functional diversity.193 Conformational flexibility and numerous intramolecular hydrogen bonds require extensive sampling to explore the conformational landscape of glycans, to which biased methods are applied. Replica exchange molecular dynamics (REMD) is widely used for N-glycan calculations.194–196 The Gaussian accelerated replica-exchange umbrella sampling (GaREUS) applies to free-energy calculations. It has been developed and applied to N-glycans.197 The improvement of glycosidic linkage sampling and the PMF calculations were reported for Hamiltonian replica exchange (H-REX), which was combined with biasing potentials in one-dimensional and two-dimensional grid-based correction maps (bpCMAP).198,199 An overview describes the recent challenges of enhanced sampling of N-glycans. 200

According to the unbiased simulation protocol validated (experimentally consistent) for N-glycan conformational sampling in explicit water, the simulations of 10 μs long allow coverage of all aqueous conformers.193 To plot the glycan conformational ensemble as an extension of carb-Rama plots (f Φ, Ψ, and ω angles).201 Andréet al. proposed a fingerprint-like presentation as a single plot in polar coordinates that describe the conformational dependence of many glycosidic linkages.192 For the past decades, these computational approaches have been extensively applied to both N- and O-glycans in solution and those attached to their aglycones, proteins, and lipids.202–205

The analysis of N-glycan's structural dynamics at the atomistic level reveals two features determining glycan's structural complexity, e.g., the monosaccharide type and linkages (Figure 14). The functionalization has a relatively local effect that is case-dependent and specific to health and disease states.190,191 N-Glycan modification regulates the flexibility of the protein to which the glycan is attached and the protein–glycan recognition in its vicinity. An alternative vision of glycans made of glycoblocks has been proposed, integrating the monosaccharide and linkage types to rationalize the structural disorder.206

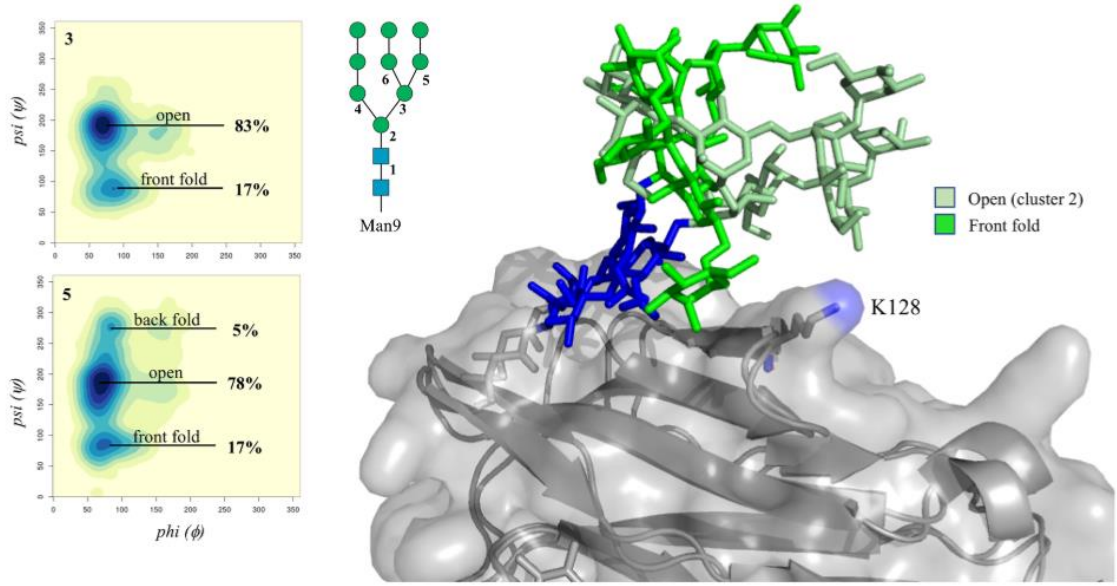


Figure 14. Conformational analysis of the (1-3) arm and (1-6) branch of the N-linked Man9 in terms of φ and ψ torsion angles, obtained from the 2 μsof cumulative MD sampling of the Man9 glycosylated FcγRIIIa. Heat maps are labeled in the top-left corner according to the Man9 numbering in the sketch. The two dominant conformations of the N-linked Man9 are shown on the right-hand side, with the protein represented by the solvent-accessible surface and underlying cartoons in grey and the mannose residues with different shades of green, as described in the legend. Heat maps were made with RStudio,207, and the structure was rendered with PyMol.14 N-Glycans are colored according to the SNFG convention. Reproduced from ref208. Copyright 2021 American Chemical Society

Mammalian glycans contain about ten types of monosaccharides. Sialylation and fucosylation are the most common types of functionalization of N-glycans. Sialic acidic monosaccharides often terminate N-glycans and play a role in cell recognition interacting with selectins. In humans, sialic acids are limited by N-

acetylneuraminic acid (Neu5Ac). Sialylation serves for protein recognition and regulates the conformational dynamics of glycans. Desialylation influences glycan's flexibility, shifts its preferential conformation and allows a larger opening of the antennae.

In contrast, the core fucosylation does not disturb the overall glycan conformation. However, it may block some antennae.209 The stacked rings are highly rigid in sialyl Lewis antennae, while other pyranose rings in the oligosaccharide exhibit flexible conformational equilibrium.193 The addition of bisecting GlcNAc to biantennary N-glycans changes the glycan conformation and shifts the linear extension to back folding of α1−6 antennae and the stacking of GlcNAc with Man or GlcNAc neighboring residues.210 Miyashita et al. distinguish five conformations for N-glycans. Some conformers' population change is due to the bisecting GlcNAc addition. It regulates the selection of a particular "key" from the "bunch of keys" .194

N-linked glycans may have different conformational behavior when attached to a protein than solutions due to direct interactions to protein surface and water- or ion-mediated contacts. Glycans may obtain increased conformational diversity on the protein surface.211,212 Glycans influence protein folding and dynamics, playing a multifaceted role in regulating protein functioning and cell signaling.213−216 The impact of glycosylation may propagate to other protein regions; MD simulations visualized the allosteric mechanism.217,218

An essential role of glycan decorating the protein surface is shielding the peptide core from potentially aggressive surroundings. For example, on the HIV gp120/ gp41 envelope trimer surface, the shield is conformationally heterogeneous with a glycan−glycan interaction network. Such a network may regulate the interactions with neutralizing antibodies.219 Glycosylation modulates the flexibility of the V1/V2 and V3 loops, regulating the entropy cost of antibody recognition.220,221 The roles of glycans in HIV-1 gp120 binding to the neutralizing antibodies PG9 and PGT128 revealed multivalent interactions between glycans and the protein surface, enhancing the binding.221 The influence of N-glycans on antibodies of IgG class are considered in ref 222. The CG parameters for N-glycans within the Martini force field have been optimized and applied to glycan-lectin interactions.223

5.2. Structure and Dynamics of SARS-CoV-2

The recognition is an essential step preceding the cell membrane fusion and pathogen penetration into the cell. In SARS-CoV-2, the spike protein (S) decorates the viral cell surface (Figure 15). It binds to the angiotensin-converting enzyme 2 (ACE2) membrane on the host cell's surface.224,225 Both S and ACE2 are highly N-glycosylated.226−230 As viruses hijack the host biosynthetic machinery, viral proteins acquire the glycan shield similar to the endogenous host proteins. As already reported for HIV, the highly dense glycan coat masks the virus from the host immune system.231 Furthermore, the glycan epitopes may elicit specific neutralizing antibody responses.232 As established by cryo-EM studies, the spike protein of SARS-CoV-2 occurs as a homotrimer.233 Each protomer consists of two subunits: S1 and S2. It carries 22 canonical sequons for N-linked glycosylation and a few for O-linked glycosylation sites.228 The subunit S1 comprises (i) N terminal domain (NTD); (ii) receptor-binding domain (RBD). The subunit S2 includes (i) one fusion peptide; (ii) two heptad repeats (HR) 1 and 2, (iii) one transmembrane domain (TM); and (iv) a cytoplasmic domain (CT).

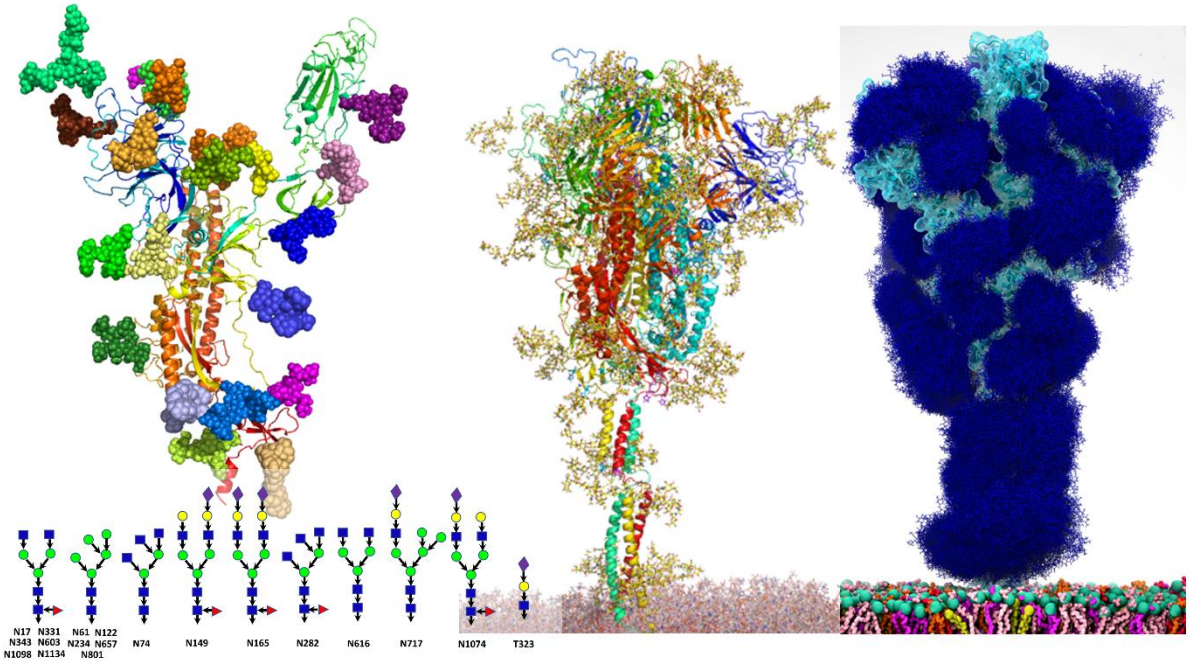


Figure 15. Three-dimensional representation of the SARS-CoV-2 spike protein assembly: (Left) Single amino-acid polypeptide chain with N- and Oglycans shown in ball representation (the N- and O-glycans, along with their amino acid linkage to the protein, are depicted according to the SNFGrepresentation). (Middle) Trimeric assembly of the left panel with glycans shown as stick representation. (Right) molecular representation of the middle panel where glycans at several frames are shown with blue lines. The right panel is reproduced from ref 234. Copyright 2020 AmericanChemical Society.

Cryo-EM diffraction provided the structures of the trimer in open (PDB 6vxx) and closed (PDB 6vsb) conformations.233 In conjunction with those derived from mass spectrometry,226,228 these data became the source of computational studies, from which several all-atom models were constructed. Of particular relevance was the construction of some missing domains: the "stalk" comprising the HR2 and TM domains and CT domain; positioning of missing loops and their glycans; addition of full-length glycans, which were not detected in the 3D structures, to GlcNAc or GlcNAc β1−4 GcNAc. Despite bearing some similarities, all-atom models show several differences. Some of these models considered the S protein inserted into the membrane. A trajectory of 10 μs long was reported as calculated on Anton 2 supercomputers and became available online (D. E. Shaw Research. https://www.deshawresearch.com/ downloads/download_trajectory_sarscov2.cgi/). Another structural model of S protein was derived from all-atom microsecond MD simulations calculated with the Amber/ Glycam force field.235 Five different glycoforms were considered: (1) glycosylated paucimannose (Man3), (2) the high mannose (Man9), (3) biantennary complex (Complex), (4) core-fucosylated biantennary complex (Complex Core F), and (5) one reported for site-specific glycosylation of the protein produced recombinantly in HEK293 (human embryonic kidney 293) cells.228 Another group of authors236 reported the complete model of glycoprotein S inserted into the lipid membrane. The carbohydrate moiety was adjusted to the composition with the highest abundance estimated by the experiment.226,228 The authors also included palmitoylation of a certain Cys in the CP domain (CYSP for a palmitoylated Cys residue in the CHARMM force field, still missing in Amber force fields). This addition was built up using CHARMM-GUI and equilibrated in the CHARMM force field for 600 ns. All the initial structures are available in the COVID-19 Archive of CHARMM-GUI (http://www.charmm-gui.org/docs/archive/ covid19), and the input files for CHARMM,237 NAMD,238 GROMACS,239 AMBER,240 GENESIS,241 and OpenMM242,243 are provided.

A thorough analysis of the structural role of glycosylation, which is beyond shielding, was performed in ref 234. The authors reported an essential role of N-glycans at positions N165 and N234 in regulating the conformational dynamics of S1 in the part of the RBD opening. The opening of the S protein is needed for the ACE2 binding. The MD simulations showed that, in the absence of glycosylation at position N234 and at both N234 and N165 positions, the RBD exhibits larger conformational freedom, which leads to its instability. A bilayer interferometry experiment supports that S binding to ACE2 is remarkably reduced for the mutant N234A and slightly impaired for N165A. Intriguingly, the presence of Man9 and Man5, respectively, at these positions, stabilizes the RBD in the "up" conformation due to their particular interactions with the protein, as seen from MD simulations. Man9 at the N234 position located on NTD (B subunit) "crawls" and deeply extends into the large cavity created by the opening of the RBD of the adjacent protomer (A subunit). It props the latter up, forming several stable hydrogen bonds with RBD residues. The N-glycan at N165 position are more exposed to the solvent and extensively interacts with the "up" conformation of RBD.

Man9 is a nascent form of glycan, which can be processed further or remains as such, in which case steric crowding occurs. A further report244 explored the role of glycosylation at N234 using MD techniques. They concluded that the N234-Man5 model is competent compared to the N234-Man9 model regarding its exposure to the open RBD, regardless of its higher dynamics. The decrease of the mannose level in the N234-Man3 model leads to a "more closed" conformation of RBD since the paucimannose is insufficient to form stable interactions within the pocket. Therefore, finetuning of N-glycans is essential in lifting the RBD and stabilizing open conformation.

Another critical position is N370, which, due to a recent mutation T372A in the NST sequon, is nonglycosylated in SARS-CoV-2 and unlikely in SARS-CoV and MERS. To investigate the impact of this lack of

glycosylation, N370 can be "artificially" glycosylated in the molecular modeling of the S protein. This glycan does not interfere with the binding pocket's glycan chains at N234, N165, and N343. It may contribute to stabilizing the RBD "up" conformation. Additionally, it may contribute to stabilizing the closed trimer interacting with the RBD of the adjacent subunit in a specific cleft. In contrast, in the absence of the glycan chain, this cavity on the close conformation remains available for other glycans such as glycosaminoglycans245,246 or glycolipids.247 The role of glycosylation, in the modulation of the lifetime of open and closed states, has been further explored in ref 248.

Point mutations modify the efficacy of RBD binding to ACE2. In the case of D614G (most dominant around the world249), the mutation favors the open conformation state.250 The N439K mutation enhances RBD affinity for ACE2.251 Glycans are broadly involved in interactions between S and ACE2. The glycan-mediated interactions occur between the S protein and the glycans located at N090, N322, and N546 of ACE2 (Figure 16).229 Steered MD results in agreement with an experimental force spectroscopy study support N-linked glycan's importance on N090 of ACE2. It prolongs the interaction lifetime, mostly absent in the SARS-CoV-1 RBD-ACE2 complex. After removing N-linked glycans on ACE2, its mechanical binding strength with SARS-CoV-2 RBD decreases to a similar level to that of the SARS-CoV-1 RBD-ACE2 interaction.252 In addition to being important in protein receptor interactions, N-glycan may stabilize the global allosteric interaction network.253,254 Several ways to enhance temporal dimension are using a CG level of representation253,255 or increasing the computer power, as exemplified throughout the Folding@home experience.256

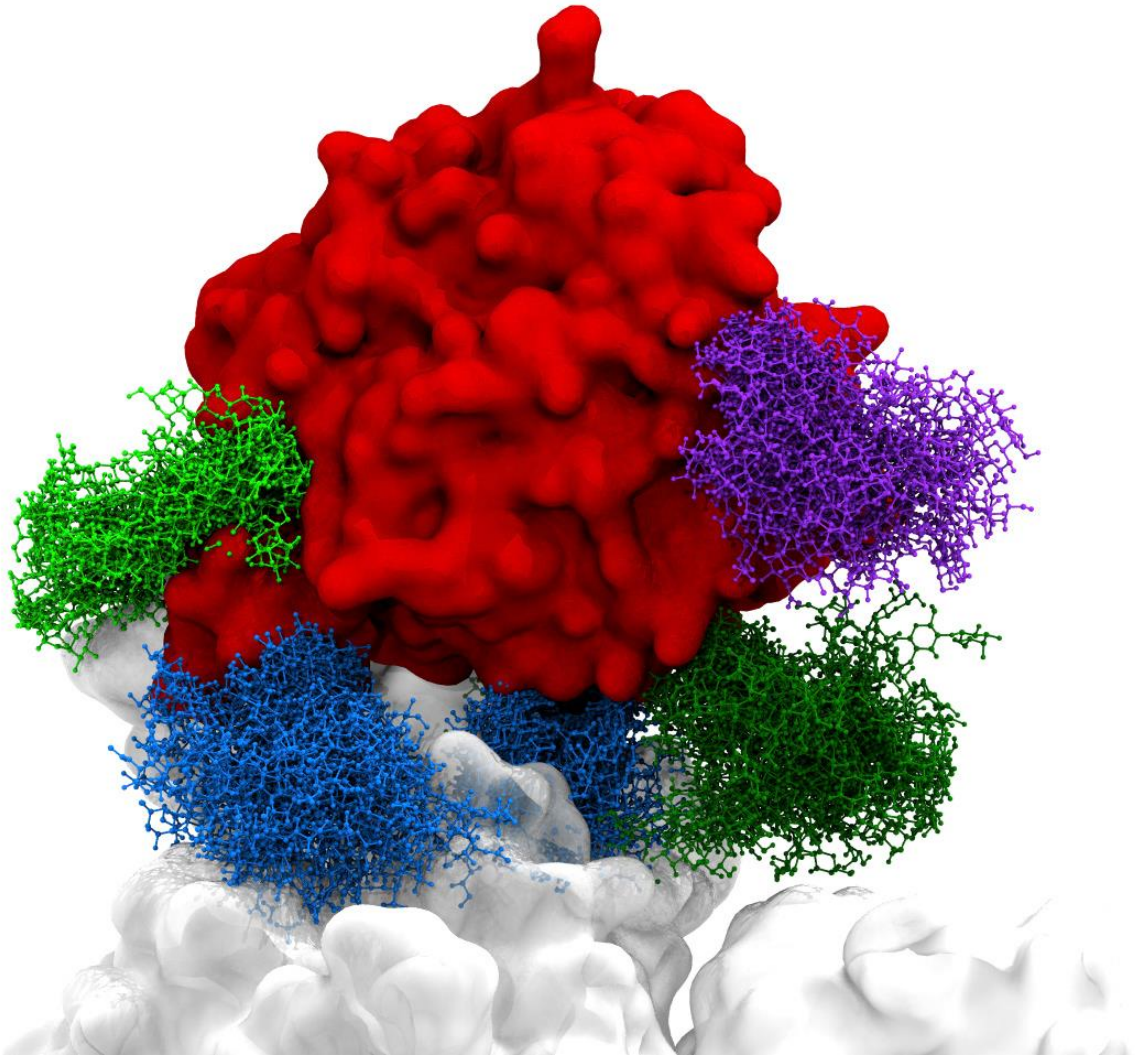


Figure 16. Interactions of glycosylated soluble human angiotensin-converting enzyme 2 (ACE2) and glycosylated SARS-CoV-2 S trimer. ACE2 is colored red, with ACE2 glycans interacting with the spike protein's upper part. Reproduced with permission from ref 229 Copyright 2020 Elsevier.

5.3. Structure, Function, and Dynamics of Glycolipids

Many vital processes occur in the vicinity of cell membranes, including host−pathogen recognition, cell−cell communication, and signalling. These processes are regulated by glycoconjugates tethered to the lipid membrane. In this regard, computational modeling of influenced transmembrane protein conform glycoproteins and glycolipids in the membrane environment provides an insightful vision of such structural complexity.

The atomistic simulations demonstrated glycosylation cruciallyation and spatial orientation. In the epidermal growth factor receptor (EGFR),257, the Man3GlcNAc2 glycan altered the receptor subdomain orientation and its orientation to the membrane. The glycans attached at positions N151 (DI), N172 (DII), N389 and N420 (DIII), and N504 (DIV) lifted these domains from the membrane plane. Such an elevated position was in agreement with experimental FRET studies.257 In the case of CD2, a small cell adhesion receptor expressed by T-cells and natural killer cells, the ectodomain orientation is determined by glycosylation and the local lipid composition of the membrane. The glycans attached at positions Asn141 and Asn150, located close to the membrane surface, create a shield preventing the electrostatic interactions with charged lipids.258

About 1% of human proteins are anchored to the cell membranes via glycolipids called glycosylphosphatidylinositol (GPI) anchors. GPI has a flexible259 pseudo-pentasaccharide glycan core Man-α(1-2)-Man-α(1-6)-Man-α(1→4)-GlcN-α(1- 6)-Myo-inositol, which is connected to a lipid tail. The conformational behavior of GPI and GPI-anchored green fluorescent protein (GFP) in the membrane surrounding has been studied using CG MD.260 In addition to the novel CG parameters of GPI, consistent with the Martini force field, the authors outlined the importance of polarizable water model usage.

In membranes, glycolipids play an essential role as receptors for protein binding. Lipids are not uniformly distributed across the membrane, forming micro- and nanosized domains. Many signaling molecules accumulated in regions enriched with sphingomyelin and cholesterol in the plasma membrane, so-called rafts. Furthermore, the proteins may sort the glycolipids. Computer simulations of such partitioning and trafficking are in high demand. Some simplified models of lipid patches can be calculated using atomistic molecular representation in terms of computer power. However, the events occurring in the highly heterogeneous membranes at a large temporal scale are demanding and can be grasped only at the CG level of calculations.79

Gangliosides are glycolipids that contain one or more sialic acids in the oligosaccharide headgroup attached to a ceramide. They are involved in many neuronal processes, particularly monosialotetrahexosylganglioside (GM1) and monosialodihexosylganglioside (GM3). These cell-surface receptors are located at the outer membrane leaflet; they may be available for binding by their respective ligands or hidden from this, a phenomenon termed "crypticity". Many studies have been devoted to ganglioside headgroup conformation on the membrane surface.261−263 The conformations of linkages GalNAc-β1-4-Gal and Neu5Ac-α2-3-Gal were reported to be restricted when GM1 is embedded in the micelle.261 The GM1 headgroup can form charge pairs with the choline group of the DOPC molecules with a lifetime in the order of a few nanoseconds, which is much longer than DOPC−DOPC interactions.262 The conformation of GM1 is mainly similar in the Lo and Ld phases.264 Nonetheless, in the presence of cholesterol, the GM1 adopts a conformation tilted toward the membrane plane.265

Another question addressed is the influence of gangliosides on membrane curvature.266,267 The spatial distribution of lipids strongly influences membrane curvature. The appearance of glycolipids raft domains was detected in a membrane leaflet with negative curvatures.266,267 Interestingly, atomistic simulations showed that GM3 clusters are slightly larger and more ordered than GM1 clusters because of the smaller headgroup of GM3.268

The nanoscale dynamics of bilayers mimicking plasma membranes with the complexity of lipid composition increasing up to dozens of lipid species have been evaluated.270−274 From the combination of 14 headgroups and 11 types of tails, 63 lipid species were asymmetrically distributed across the two leaflets. 274 The MD analysis revealed nonideal lipid mixing with transient domain formation, which disappeared at the microsecond time scale. The domains on two membrane leaflets were coupled and enriched with gangliosides on the outer leaflet.

Gangliosides are used as receptors for recognition by pathogen lectins, i.e., Cholera toxin B subunit264,275 and SV40s Viral Capsid Protein VP1.276 Shiga toxin B subunit and LecA from Pseudomonas aeruginosa share Globotriaosyl ceramide (Gb3), also known as the PK blood group, for the host cell

recognition.277,278 A comparative study has demonstrated that lipids randomly distributed across the membrane revealed different modes of Gb3 clustering upon virulence factor binding279 (Figure 17).

Glycolipids play a significant role in embedding transmembrane proteins. The thylakoid membranes are where the galactolipids may bind specifically to the photosystem II complex.280,281 CG modeling unveiled the role of glycolipids in forming supercomplexes between photosystem II and the light-harvesting complex.282,283 Some other relevant investigation studies relate to the diffusion of small molecules and ligands in the photosystem II embedded into the thylakoid membrane.284,285 The role of lipid composition in their mixing was studied through a comparative analysis of the thylakoid membrane of cyanobacteria and higher plants.286 The lipid composition was closely related to the real membranes and included up to five galactolipids plus phosphatidylglycerol in each case. The analysis showed the nanoscale heterogeneities existing, particularly in the plant membrane, while the fluidity of the cyanobacterial membrane was markedly reduced compared to the plant membrane. For more details about MD simulations in photosynthesis, see ref 287.

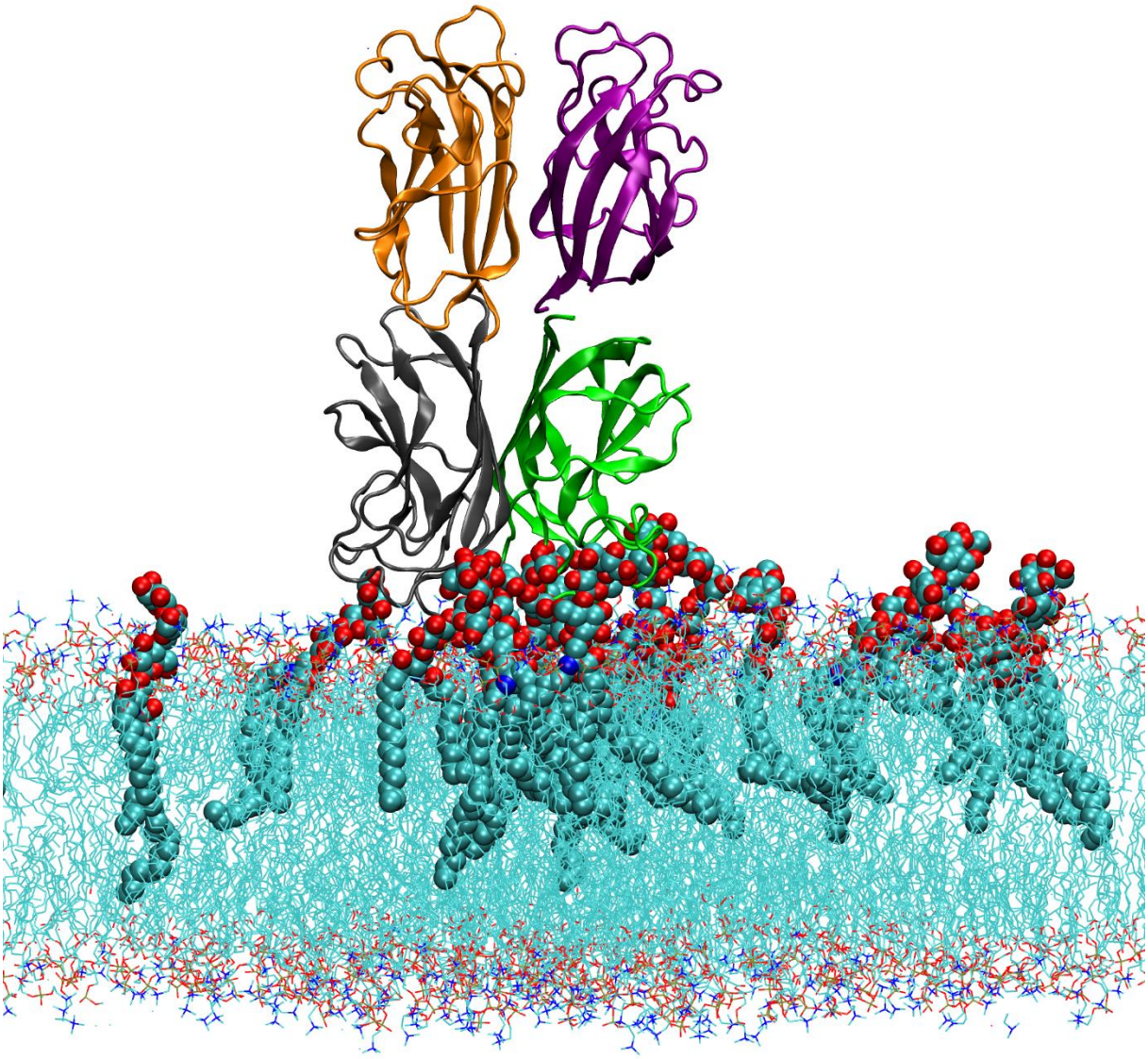


Figure 17. Globotriaosyl ceramide molecules (Gal-α1-4Gal-β1-4Glcβ1-Cer, shown as van der Waals spheres) clustering in the 1,2-dioleoyl-sn-glycero-3-

6. POLYSACCHARIDES

Polysaccharides constitute the most diverse and abundant class of biopolymers. Several hundred known examples offer a great diversity of chemical structures whose number and nature of constituents exceed those usually found in mammalian glycans.288 Polysaccharides range from linear homopolymers to complex heteropolymers where the repeating units may be as large as octasaccharides. It is the case for bacterial polysaccharides, including capsular and exopolysaccharides, lipopolysaccharides, and peptidoglycans. Some of these polysaccharides can be branched, a unique feature among natural macromolecules. Depending on their primary structures, polysaccharide chains adopt distinctive shapes which characterize their secondary structures: ribbons, extended helices, hollow helices, and multiple helices. Metastable structures may appear whenever crystallization and biosynthetically driven polymerization are concomitant. Some of these features may persist locally in the dilute state and direct the solution properties of the polysaccharides. Polysaccharides exhibit various structures and architectural organizations depending on their origin. They can develop over several orders of magnitude, as observed for cellulose,289 chitin,290 or starch.291 The computational methods initially dealt with small- and medium-sized polysaccharides; they must be extended to cope with various situations. Simulations based on the principle of all-atoms representation remain used.

Nevertheless, the characterization of the structural characteristics of polysaccharides has benefited from developing a CG model in its ability to simulate biomolecular systems on large scales and possibly time scales, inaccessible to all-atom models. These applications mainly concern polysaccharides: (i) From biomass, i.e., cellulose292−301 nanocellulose,302−304 cellulose-polysaccharide interactions,305,306 and chitin.307,308 The development of an accurate CG model for chitosans offers a characterization of the microscopic and mesoscopic structural properties of large polysaccharides in solution for a wide range of solvent pHs and ionic strengths. Of particular interest is the investigation of the effect of polymer length and degree and pattern of deacetylation on the polymer properties.309 (ii) From the extracellular matrix, GAGs,6,310−313 and other polysaccharides such as α(1-3) glucan.314,315 In most cases, these applications study the interactions of these polysaccharides with proteins.

6.1. Polysaccharides in Solution

Polymer chains in solution adopt disordered structures that fluctuate between local and global configurations. Polysaccharides adopt various spatial arrangements around glycosidic bonds because these molecules have much conformational freedom. Theoretical models of polysaccharides are established on studies of the relative abundance of different conformations, following the statistical theory of polymer chain configuration.316

The possible interactions between the polysaccharide chain residues other than nearest neighbors are ignored in the first approximation. A Monte Carlo sampling reflects the range of conformations of polymer molecules. The dilute solution conformation properties, such as persistence and mass per unit lengths, are derived over the full range of conformations available to the chain. These properties correspond to the equilibrium state of the chain and refer to unperturbed chain dimensions that ignore the long-range excluded volume effect. They are computed from the conformational states derived from the potential energy surfaces Φ and Ψ of consecutive disaccharide fragments.

A Monte Carlo algorithm has been developed that operates on the MD trajectories; it offers a more accurate description of the systems since all physical interactions between even distant residues are computed.317 Despite an ad hoc application,239, the dynamic simulation of molecular motion and polysaccharides' conformation is not frequent. The dynamic simulations of the molecular motion and conformation of the triple helix conformation (1→3)-β-D-glucans having β-D-glucopyranosyl units attached by (1→6) linkages extracted from a black fungus (Auricularia auricula) were reported.318

The following example illustrates the application of a series of computational tools to decipher the structural features and calcium-mediated interactions of the exopolysaccharide produced by the deep-sea hydrothermal bacterium Alteromonas infernus: Infernan319 (Figure 18).

6.2. Polysaccharides in the Solid-State

Half of the annual biomass produced is cellulose, the most abundant being bio-organics on the planet. The understanding and the "intelligent use" of cellulose and complex architectures of components are attracting

considerable attention in setting up computational methods. The quest to overcome cellulose recalcitrance for producing biofuels and sustainable biomaterials, on the one hand, and the establishment of structure-properties relationships, on the other hand, have paved the way for the most recent developments.

Plants have evolved complex nanofibril-based cell walls. The nanoscale to mesoscale organization meets the strength and the extensibility of growing cell walls and can be approached throughout the development of CG models.320 The model reveals some biomaterial design principles encompassing stiffness while yielding extensibility. As a complement, modeling the interactions in a plant's primary and secondary cell walls requires elucidating the interactions between cellulose and hemicelluloses such as xylan and xyloglucan.321 A molecular dynamics study suggests that the adsorption of xyloglucan to cellulose surfaces is driven by entropy.322 The presentation of a CG model describing xylan and its interactions with crystalline cellulose highlighted how the complementarity of the chains directs the interaction. Extended modeling revealed that the adsorbed xylan could adopt coiled structures, especially on hydrophobic cellulose surfaces.305

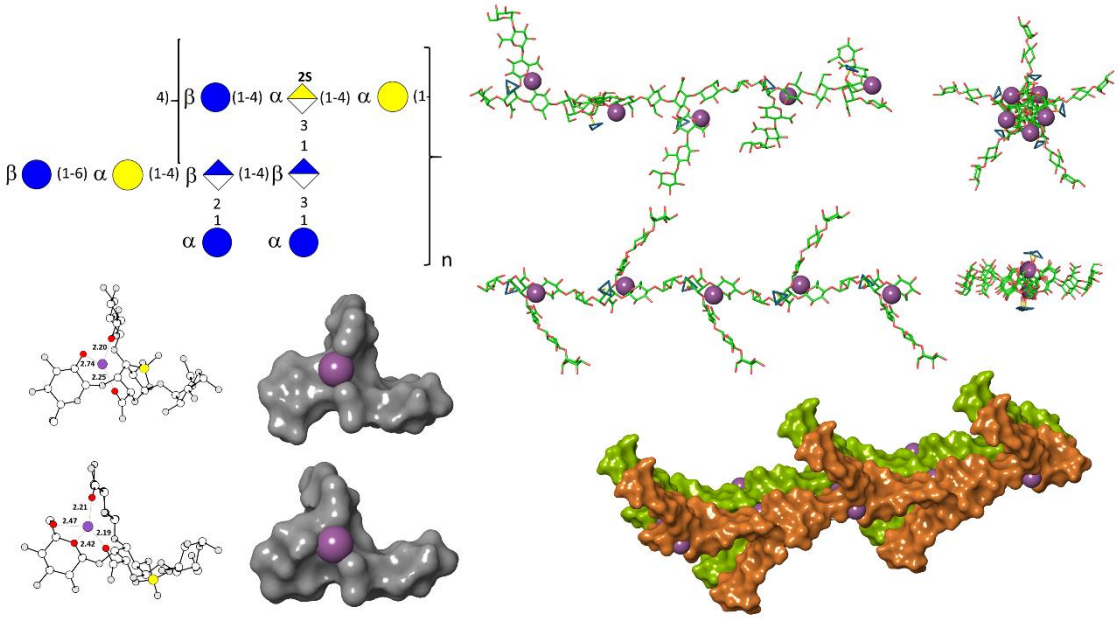


Figure 18. The exopolysaccharide Infernan is produced by the bacterial strain Alteromonas infernus from collected samples of vent fluids at a depth of 2000 m. (a) Infernan has a complex repeating unit of nine monosaccharides established on a double-layer of side chains. A uronic and sulfated monosaccharide cluster confers to Infernan functional and biological activities. (b) Molecular mechanics and dynamics along molecular dynamics trajectories clustered the conformations in extended 2-fold and 5-fold helical structures. (c) The electrostatic potential distribution over all the structures revealed negatively charged cavities explored for Ca2+ binding through quantum chemistry computation. (d) The ribbonlike shape of 2-fold helices brings neighboring chains in proximity without steric clashes. The cavity chelating of the Ca2+ of one chain is completed by the interaction of a sulfate group from the neighboring chain. The resulting “junction zone” is based on unique chain−chain interactions governed by a heterotypic binding mode. Drawn with PyMol.14 Reprinted with permission from ref 319. Copyright 2022 Elsevier.

Several CG models using distinct ”beads” representations have provided important insights (Figure 19). For example, a simplified representation using ”bead” for each monomeric glucose subunit described the intrinsic conformational transition of long cellulose macro fibrils between crystalline and amorphous phases on long time scales.301 The investigation was further extended to evaluate the significance of considering an explicit solvent. The results showed that the cellulose fibril's persistence length in the transition region between fully crystalline and amorphous corresponds to that of native cellulose fibrils. As indicated by the analysis of the individual energetic contributions, the polysaccharide−water interactions contribute to the crystalline to amorphous transition of the cellulose fibril.292,293,295,299

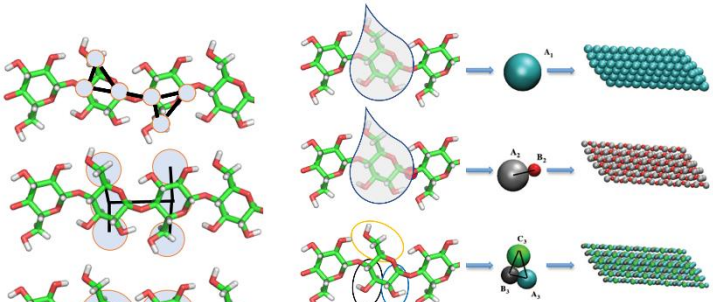


Figure 19. Coarse-grained schemes used for cellulose. In many studies, three beads have represented each glucose unit.64,75,292−294,301 Reproduced with permission from ref 5. Copyright 2021 Elsevier

A group of authors derived one set of coarse-grained MARTINI force field parameters to simulate crystalline cellulose fibers. The model is adapted to reproduce different physicochemical and mechanical properties of native cellulose. The model can handle the transition between cellulose allomorphs and capture the physical response to temperature and mechanical bending of longer cellulose nanofibers.294 These few examples are among the many investigations where computational methods decipher complex macromolecular architectures related to physical and mechanical properties. At the same time, they lay the foundations for further exploring the relationship between the crystal morphology in topo-enzymology and topo-chemistry.

As biodegradable nanomaterials, cellulose nanocrystals (CNCs) and cellulose microfibrils offer many applications with unique physical, chemical, and mechanical properties. CG simulations allow the construction of reliable structural models that can be extended to predict mechanical properties. For example, a CG model based on a "bead" representation of the cellobiose unit indicated that well-aligned CNCs lead to a more brittle and catastrophic failure mechanism, while naturally twisted interfaces favor hardening mechanisms that help achieve optimal mechanical performance.303 The same model was applied to investigate the effect of interfacial energy and torsion on mechanical performance. It showed that elastic modulus, strength, and toughness are more sensitive to the torsion angle than interfacial energy.304,306

6.3. Lipolysaccharides in Membranes

Bacterial membranes are complex in the chemical composition of their constituents that form the cell envelope surrounding the Gram-negative bacteria cytoplasm, e.g., the inner membrane, the periplasm, and the outer membrane. The outer membrane is an asymmetrical bilayer where zwitterionic and anionic phospholipids compose the inner leaflet.323

Lipopolysaccharides (LPSs) have the most extensive and complex chemical diversity that composes the outer leaflet (Figure 20). They offer a challenge to establish their structure and dynamics via computational methods.79 They offer a way to investigate membrane biophysics and properties resulting from the distribution and behavior of LPS in the outer membrane. 324−329

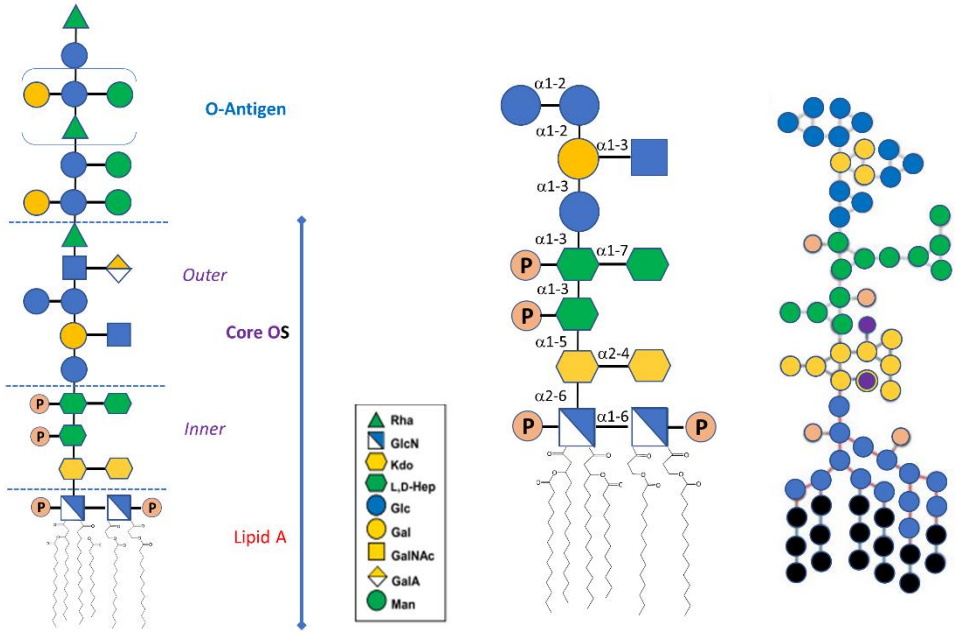


Figure 20. Schematic representation of the general structure of lipopolysaccharides. The zigzag lines in lipid A represent fatty acid chains. Abbreviations: Rha, 6-deoxy-mannose (rhamnose); GlcN, 2-amino-2-deoxy-glucose (glucosamine); Kdo, 3-deoxy-α-D-manno-oct-2-ulopyranosonic acid; L,D-Hep, L-glycero-D-manno-heptose; Glc, glucose; Gal, galactose; GalNAc, 2-acetamido-2-deoxy-galactose (N-acetylgalactosamine); GalA, galacturonic acid; Man, mannose; P, phosphate. Structural comparison between a core-type LPS and corresponding Martini model. The constituting monosaccharides are depicted following the SNFG representation. Phosphates and Kdo carboxyl groups are colored in tan and dark blue, respectively. The Martini model of LPS follows a 4-to-1 mapping scheme of the Martini force field.330

The size and complexity of LPSs can benefit from the application of CG methods, provided that parameter sets for a MARTINI representation are developed from atomistic simulation and compared to experimental data. Developing such a parameter set requires several approximations, mainly concerning the glycan part of the LPS, which contains several anomeric centers. Information about the many stereoisomers is lost because the CG beads lack atomic resolution. The application of this model to 27 different membrane compositions was carried out for a 100 μs cumulative MD simulation. It revealed critical structural evolutions, including a significant result when studying the effect of increasing the composition of LPS in 1,2-dipalmitoyl-3-phosphatidylethanolamine (DPPE) in the outer leaflet. The change in composition induces a decrease in the packing of the LPS molecules. It is accompanied by an increase in the membrane's surface area per lipid. Thus, the phase transition values decrease from 346 to 290 K. This simulation describes how chemical heterogeneity in membrane composition can lead to significant membrane properties, such as fluidity and phase transition temperatures, varying by ±15° in the experimental systems.

While losing resolution compared to all-atom simulations, CG simulations provide a better understanding of the architecture and physicochemical properties resulting from the curvature and permeability of membranes.87,331 The combined use of all-atom and CG methods provided a description of the outer membrane architecture of Pseudomonas aeruginosa, from which the effects of its LPS on membrane permeability and dynamics and their impact on antimicrobial resistance were discovered.87,332,333

The main results indicate that the LPS layers are arranged in a closely interacting mesh within the outer leaflet, forming a gel phase, which offers resistance to mechanical breakage. The observation of limited lateral mobility of the outer leaflet may impact the movement of the molecules through the membrane during their translocation to the upper leaflet.

Another study revealed how divalent cations, such as Ca2+ and Mg2+, help maintain the integrity of the outer membrane structure by forming stable lamellae from forming strong complexes with phosphate groups.332,334 The stability of lipid A bilayers for different acylated structures was studied by GC-MD modeling.335 The effect of O-antigen on the lipid bilayer enlightens the tendency to pack and the subsequent role in increasing mechanical strength and stability. When the outer membrane sheets contain only LPS, their firm, cohesive intermolecular interactions generate tight packing of O-antigen chains. However, when phospholipids and LPS are present in the outer leaflet, the O-antigen chains are tilted and less tightly packed. The diversity of O-antigen chain packing affects lipid mobility and the mechanical strength of the membrane.327

Gram-negative membranes with outer leaflets of LPS alone withstand surface tensions that cause the membrane to rupture much more readily. Insights into understanding the bacterial cell envelope are accumulating from CG modeling.8 The motions of the LPS molecules are highly correlated with each other and the outer proteins embedded with the membrane.

The discovery of the tlr4 gene encoding for a receptor able to bind LPS, the Toll-like receptor 4 (TLR4), remodeled the LPS immunology. TLR4, along with its accessory protein myeloid differentiation factor 2 (MD-2), builds a heterodimeric complex that recognizes LPS explicitly. This discovery initiated, among others, structural investigations to decode the specific sequence of events from when and how the immune system discerns the LPS to how and when it triggers a response from it.336−341

A series of molecular modeling and computational studies have provided insights into the mechanism regulating the activation/inactivation of the TLR4/MD-2 system receptor and the fundamental interactions modulating the molecular recognition process by agonist and antagonist ligands (Figure 21).[342−352]

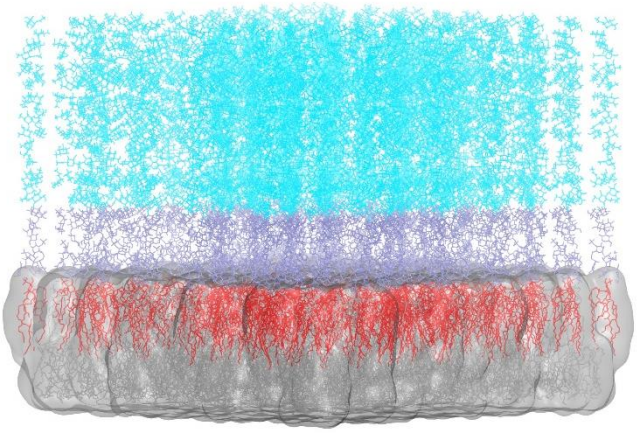


Figure 21. All-atom representation of model membrane of Escherichia coli. DPPE lipids are shown as gray. For LPS moieties, the followingcolor coding was used: Lipid A (Type 1), red; core oligosaccharide (R1) (−3)βDGlc(1-3)[αDGal(1-2)αDGal(1-2)]αDGlc(1-3)αDGlc- (1-3)[αLDHep(1-7)]αLDHep(1-3)αLDHep(1-5)[αDKdo(2-4)]- 2)]αLRha(1- 2)αLRha(1-2)αDGal(1-3)βDGlcNAc(1-), cyan.

7. PROTEIN CARBOHYDRATE INTERACTIONS

7.1. Presentation: Synopsis of the Protein Families

All naturally occurring glycan structures and conjugates result from complex actions involving several critical steps in which interaction with proteins occurs. It is beyond the scope of this review to cover these events, and we consider broad classes of biological events and protein actions. Those interactions driving these different biological events involve the classes of enzymes having catalytic activity: biosynthesis, modification and hydrolysis. Other important events are mediated by or involved in recognition, such as transporters and an essential group of carbohydrate-binding proteins: lectin, antibodies, carbohydrate-binding modules, and glycosaminoglycan-binding proteins. Prediction experimentally based on the availability of elucidation of protein−carbohydrate crystalline complexes[353− 355] or protein sequences.[356] Figure 22 is a schematic description of the events involving protein−carbohydrate interactions.

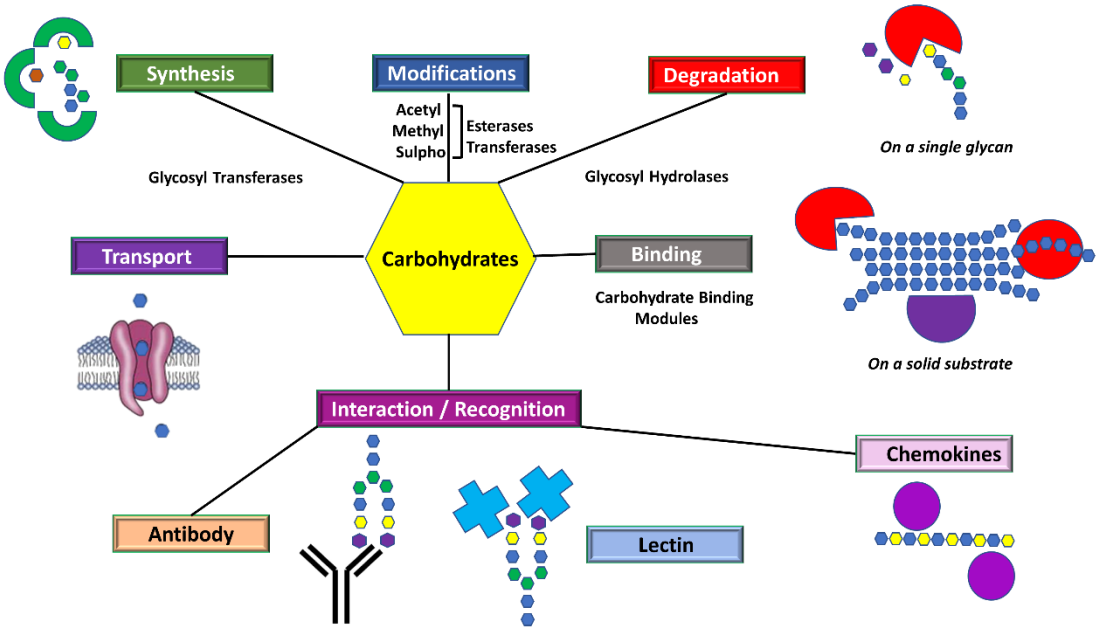


Figure 22. Synopsis of the families of the principal families of proteins interacting with carbohydrates along with their functions: transport; synthesis (glycosyltransferase) modifications (auxiliaries enzymes); degradation (glycosyl hydrolase, on single glycan and semicrystalline and crystalline glycan); carbohydrate-binding modules; antibodies, lectins, and chemokines. Reproduced with permission from ref 5. Copyright 2021 Elsevier.

These are glycosyltransferases (GT) and glycoside (glycosyl) hydrolases (GH). Glycosylation proceeds in a stepwise manner. The enzyme's expression, location, and specificity constitute regulatory elements in generating the repertoire of biosynthesized glycans and glycoconjugates. The covalent additions of the glycans to peptides, lipids, and proteins represent the most abundant post-translational modifications, which generate a paramount number of structural diversities. These structural changes in cell surface glycans guide physiological and pathological cellular processes. The same holds for carbohydrate polymers and their resulting functional properties. Computational modeling has contributed to the atomistic understanding of some origins of catalysis, and its machinery, from ideal to highly relevant complex cases. Consequently, it is interesting to decipher the mechanisms these carbohydrate-acting enzymes use.

The combined quantum mechanics-molecular mechanics QM/MM approaches are most suited to provide a realistic atomistic description of the enzyme environment and offer reliable electronic structure calculations. Combined with molecular dynamics simulations that scan the configuration space available, estimates computational methods have been chosen to characterize the breaking/formation of chemical bonds as they occur in a fully hydrated enzyme. These methods offer access to other essential features, such as representing the reaction transition state (TS) or obtaining an estimate of the free energy barrier (ΔG). Computation can also be used to study ligand association−dissociation or processivity.

7.2. Insights into Glycosyltransferases

Glycosyltransferases (GTs) comprise a group of processing enzymes (EC 2.4) that transfer glycosyl residues from a donor to other molecules.357−360 The catalysis transfers donor substrates (mainly sugar nucleotides, such as UDP-GlcNAc, UDP-Gal, GDP-Man, and unsubstituted glycosyl phosphates and lipid-linked sugars) to a nucleophilic glycosyl acceptor. Other molecules, such as proteins, lipids, DNA, antibiotics, or other small molecules, can act as acceptors. GTs require a specific metal ion cofactor for catalysis. These divalent metal ion cofactors are Mn2+ and Mg2+. In the presence of other ions, the catalysis is often impaired361,362 (Figure 23).

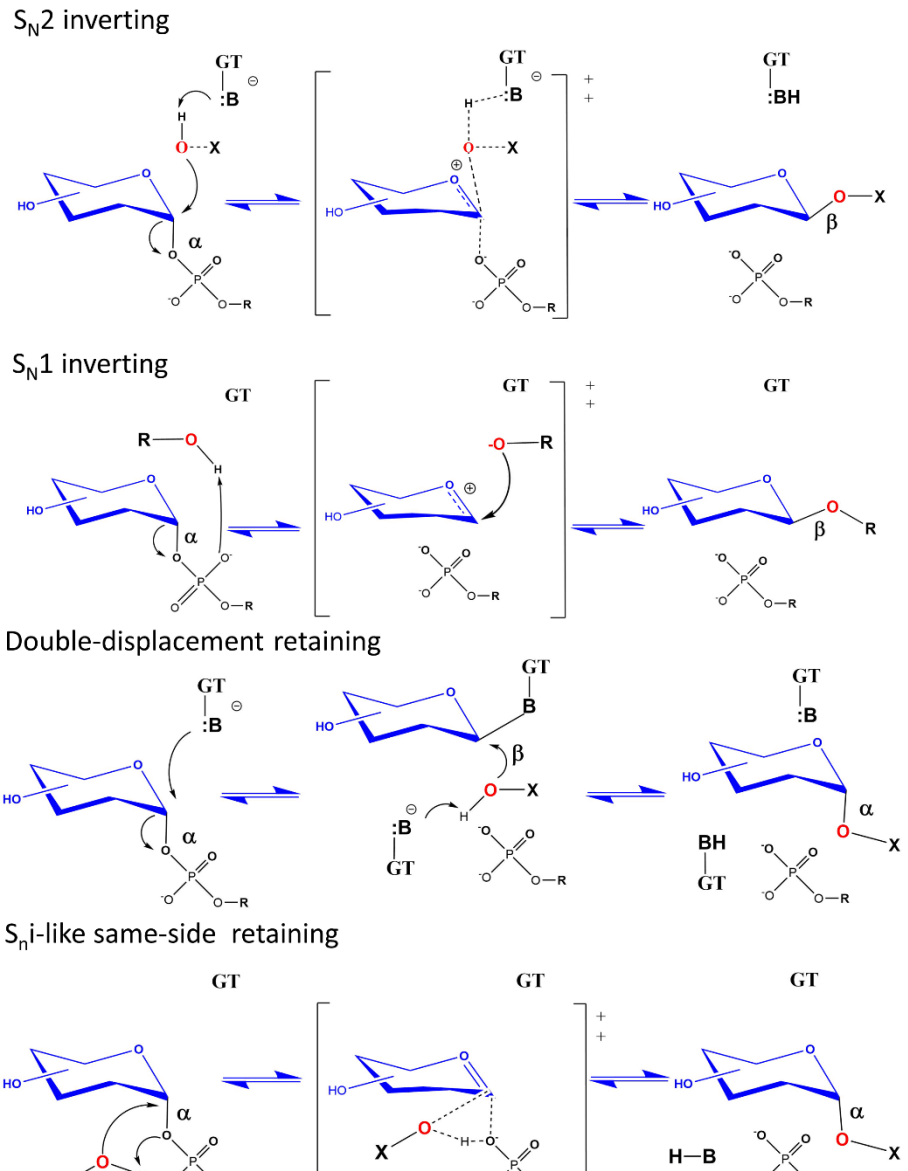


Figure 23. Schematic diagram of overall reaction catalyzed by GTs. The reaction mechanism proposed for inverting GTs proceeds in a single displacement SN_2-like mechanism, forming an oxocarbenium ion transition state. A catalytic amino acid (B) serves as a general base that deprotonates the nucleophile hydroxyl group of the acceptor. The reaction mechanisms proposed for retaining glycosyltransferases imply double-displacement mechanisms. They involve two successive SN_2-like steps with a nucleophilic attack of an amino acid of a glycosyltransferase on the anomeric center of the donor substrate, leading to the formation of a covalent glycosyl−enzyme intermediate. In the second step, the glycosyl−enzyme intermediate is attacked by a hydroxyl group of the acceptor facilitated by its deprotonation by a catalytic base, resulting in overall net retention of configuration. For the SN_i-like mechanism, the front-side nucleophilic attack proceeds in a single step to form an enzyme-stabilized oxocarbenium ion. The interaction with the departing phosphate and the incoming acceptor nucleophile facilitates the deprotonation of the acceptor nucleophile (drawn with ChemDraw13).

GTs display low sequence homology.363 They are classified into more than 114 families based on amino acid sequence similarities of about 100,000 sequences (http://www.cazy.org/GlycosylTransferases.html).364,365 As assessed from X-ray crystallography, the three-dimensional structures of GT exhibit a small range of folds366 as derived from the Structural Classification of Proteins (http://scop.berkeley.edu/). They are referred to as GT-A, GT-B, and GT-C. The GT-A architecture comprises two domains involving nucleotide binding and acceptor binding. GT-A enzymes contain a single Rossmann-like fold, a domain found in enzymes that bind nucleotides. Most, but not all, GT-A fold enzymes require divalent cations for activity and contain a “DxD” motif involved in coordinating the metal and nucleotide sugar. GT-B fold enzymes do not require divalent cations for activity. GT-C fold enzymes, such as oligosaccharyltransferases, contain multiple transmembrane α-helices and employ lipid-linked sugar donors.367

The chemistry of the catalytic reaction involves a nucleophilic displacement of the substituted phosphate leaving group (for example, the UDP functional group) at the anomeric C-1 carbon of the transferred sugar residue of a donor by a hydroxyl group of a specific acceptor.

Each GT family can invert or retain the stereochemistry at the anomeric linkage during transfer. They belong to the ”retaining” or ”inverting” class of enzymes according to whether the stereochemistry of the anomeric bond of the donor is retained (α→α) or inverted (α→β or β→α) after the transfer. Several detailed reviews have been published on mechanistic and structural studies of glycosyltransferases.368−378

Based on experimental and theoretical calculations, the inversion of the anomeric configuration obeys an SN2-like direct displacement. The amino acid side chain in the active site acts as a catalytic base that deprotonates the acceptor nucleophile.379 As with the retaining GTs, a dual displacement mechanism involves a covalent glycosyl-enzyme intermediate. However, fundamental questions about the reaction mechanism remain to be answered, which has motivated an active field of computational research380−387 with, sometimes, the knowledge of experimental observations.388

Structural and kinetic data have provided information on mechanistic strategies these enzymes employ. Nevertheless, molecular modeling studies remain essential for understanding the reaction catalyzed by GT at the atomistic level. QM/MM methods have become crucial to establishing different chemical reaction mechanisms. These methods allow enzymatic reactions to be modeled using quantum mechanical methods to calculate the electronic structure of the active site models and treat the remaining enzymatic environment with faster molecular mechanical methods. There is still a long road to travel to fully understand the role of conformational dynamics in enzyme activity and disclose the various reaction mechanisms these enzymes employ. Several reports on applying QM/MM methods to GT-catalyzed reactions highlight the insight gained from modeling glycosyl transfer into the mechanisms and transition states structures of both inverting and retaining GTs.389−394

However, only a few studies consider the physiological context of the reaction.395−400 Chloroplast provides a challenging example to investigate the enzymatic catalytic event in a complex organization.401 The

chloroplast converts the collected photons into chemical energy. Within the chloroplast, thylakoids are membrane-bound compartments and sites of the light-dependent reactions of photosynthesis. Chloroplast thylakoids form stacks of membraneous disks (or grana) connected by unstacked stroma membranes (Figure 24). Within such photosynthetic machinery, there exists a unique spatial architecture that results from the presence and organization of two uncharged galactoglycerolipids (monogalactosyldiacylglycerol (MGDG) and digalactosyldiacylglycerol (DGDG)). Their content reaches 80% of the total amount of lipids. A monotopic monogalactosyldiacylglycerol synthase (MGD1) is embedded in the outer leaflet of the inner envelope membrane of chloroplasts and synthesizes MGDG. It adds galactose from the water-soluble donor substrate, UDP-α-D-galactose, to the hydrophobic acceptor substrate, diacylglycerol (DAG). The transfer needs an anionic lipid and proceeds with the inversion of the anomeric configuration to the donor substrate and the reaction product. The intricate organization and process of assembling and synthesizing the complexes were studied at coarse-grained and all-atom computer simulation levels.401,402 It allowed for a large covering of the temporal and spatial scales.

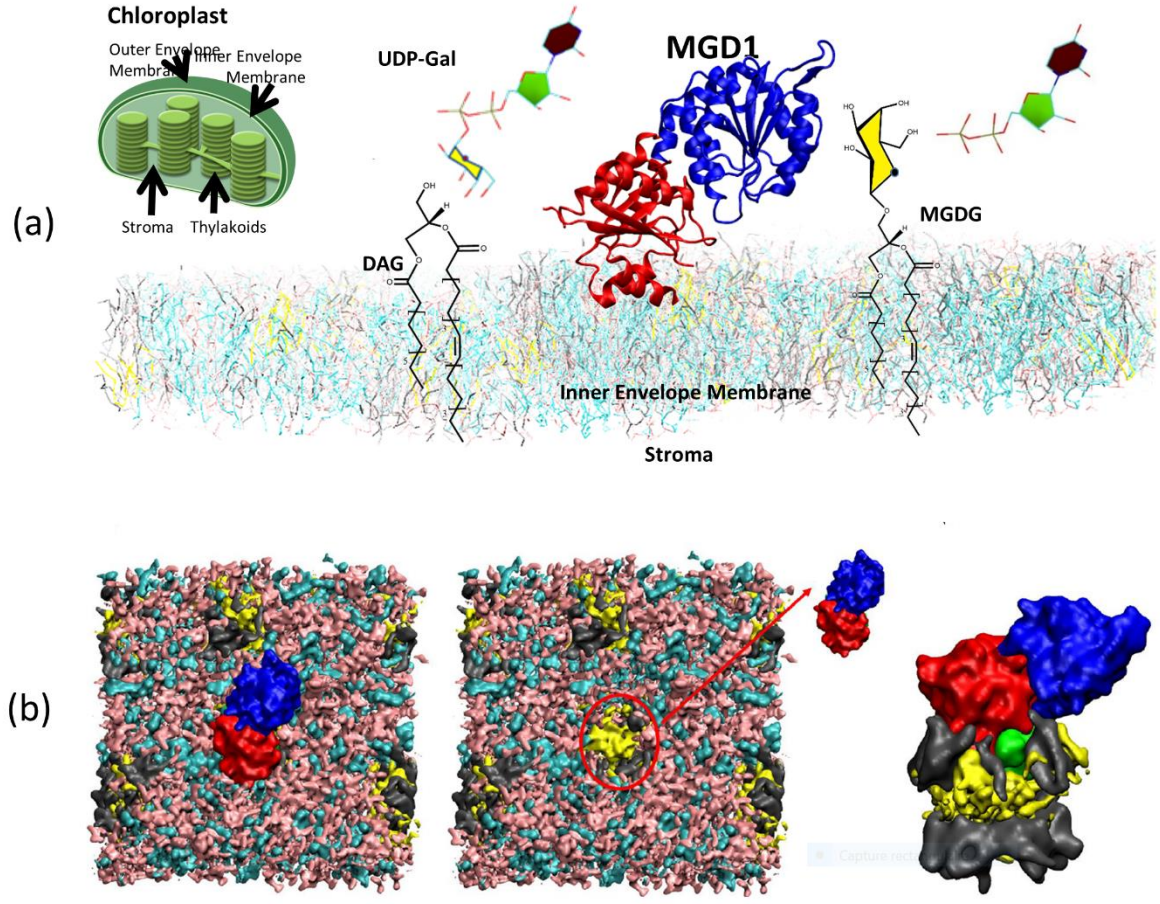


Figure 24. (a) Schematic representation of the main actors for synthesising monogalactosyldiacylglycerol (MGDG) in the chloroplast's inner envelopemembrane (IEM). The N- and C-domains of MGD1 are highlighted in red and blue, respectively; the galactosyl residue is shown in yellow. (b) Compelling features of interactions between MGD1 and lipid bilayers and lipid capture reveal the reciprocal influence of the membrane and the protein. A snapshot ofMGDG (cyan)/DGDG (pink)/PG (gray)/DAG (yellow) lipid bilayer and MGD1 (N-domain is in red, C-domain is in blue, and LOOP is in green). View from the top (left). The PG/DAG cluster with MGD1 bound, frontal view (right). The protein induces lipid reorganization in the membrane, facilitating the activator's and substrate's capture. This reorganization is concomitant to the intrinsic dynamics of the proteins, which is essential for enzyme activity. Reproduced with permission from ref 401. Copyright 2020 Nature Publishing Group.

7.3. Insights into Glycosyl Hydrolases

Glycosidases or glycoside hydrolases (GH) hydrolyze glycosidic bonds in carbohydrates, polysaccharides, glycoproteins, glycolipids, etc. These enzymes are classified as endo- and exo-types. Exotype glycosidases attack and hydrolyze monoglycosides into free sugar and aglycon. When acting on oligo- or polysaccharides, they liberate a monosaccharide unit from the nonreducing end. Endotype glycosidases act on oligo- and polysaccharides and catalyze the hydrolysis of an internal glycosidic linkage, thereby liberating two carbohydrate moieties or releasing an oligosaccharide (or polysaccharide) and monoglycoside of the reducing end. Some glycosidases can act as both exo- and endotypes. Based on the similarities in their sequences, GH has been clustered into over 128 families.403

Members of the same family function throughout a similar catalytic mechanism and share the same general fold. For more than 90 families, at least one 3D structure is available. From the availability of crystallographic data for more than 90 families, it has been observed that 3D structures are more conserved than primary sequences.

Consequently, 15 superfamilies called GH clans regroup different families sharing similar structures. The catalytic cleavage of glycosidic linkages by GH proceeds according to two distinct general canonical mechanisms. They yield either retaining or inverting of the anomeric configuration, a double nucleotide substitution for the former, and a single displacement for the latter. Both mechanisms involve an oxocarbenium-like transition state. A pair of carboxylic acids promotes the aglycon's departure as a leaving group, followed by a nucleophilic water attack at the anomeric center. The study of these mechanisms has been the subject of many computational investigations.368,369,374,375,378,394,404−408 A stepwise mechanism drives the catalytic cleavage of retaining glycosidases (Figure 25).

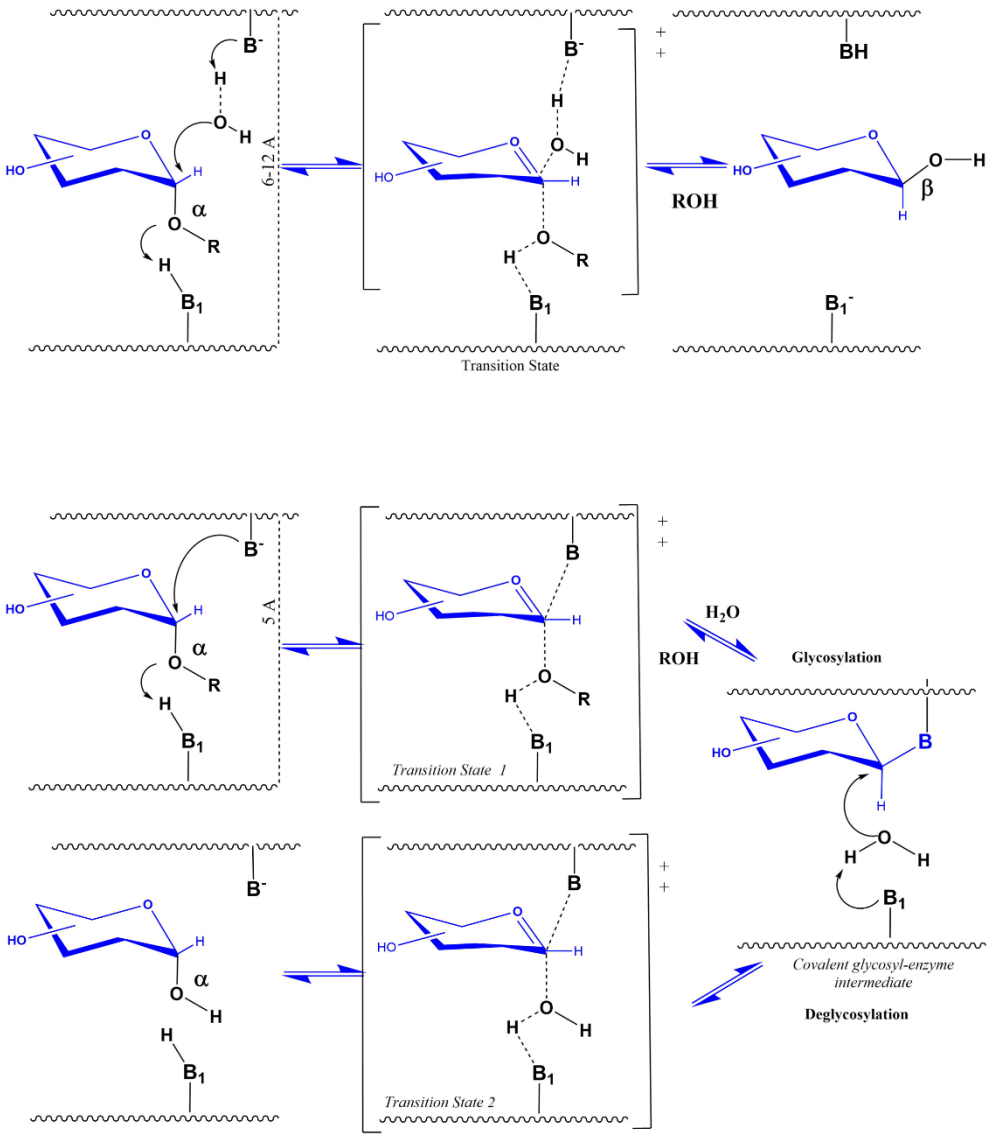


Figure 25. Schematic representation of the two nucleophilic displacement mechanisms of inverting glycosyl hydrolases exemplified by the hydrolysis of an α-D-glycoside. The hydrolysis involves a single step via the direct displacement of the aglycone. The transition state displays oxocarbeniumion-like features. One carboxylic acid (B) acts as the general base and activates a water molecule for nucleophilic attack at the anomeric center of the substrate. Simultaneously, the second carboxylic acid (B1) facilitates the departure of the leaving group via the general acid catalysis mechanism. A distance between 5 and 10 Å separates the two catalytic residues. Schematic representation of the two nucleophilic displacement mechanisms of retaining glycosylhydrolases exemplified by the hydrolysis of an α-D-glycoside. The first glycosylation step involves the formation of a glycosyl enzyme intermediate via an oxocarbenium ion-like transition state. The general acid catalyst is deprotonated in the following deglycosylation step and acts as a general base. It occurs by activating a water molecule for nucleophilic attack at the anomeric center of the glycosyl−enzyme intermediate. This step also proceeds via an oxocarbenium ion-like transition state. A distance of approximately 5 Å separates the two carboxylic acid residues: B and B1 (drawn with ChemDraw13).

Hydrolysis requires two nucleophilic displacement steps. A covalent glycoside−enzyme intermediate is formed in the first step via an oxo-carbenium ion-like transition state. In the second step, the general acid catalyst is deprotonated and acts as a general base. It activates a water molecule for nucleophilic attack at the anomeric center of the glycosyl-enzyme intermediate. This step also occurs via an oxocarbenium ion-like transition state. Most glycoside hydrolases operate by these two steps. The main exceptions are those that act on 2- acetamido sugars. The acetamido carbonyl's oxygen atom acts as an intramolecular nucleophile in a substrate-assisted nucleophile. In glycosidases' catalytic mechanism, the pyranose ring at subsite-1 occurs in a distorted

conformation instead of the more stable 4C1 conformation. First proposed for hen eggwhite lysozyme,409–411, this feature was repeatedly observed in several retaining and inverting glycosidase complexes.405,412–418

Computational methods contributed significantly to elucidating this substrate preactivation showing that monosaccharide distortion determines the pathway of the glycosidase reaction. This is supported by reported calculations of conformational free energy surfaces that allowed a description of the catalytic routes used by glycosidases.418–421

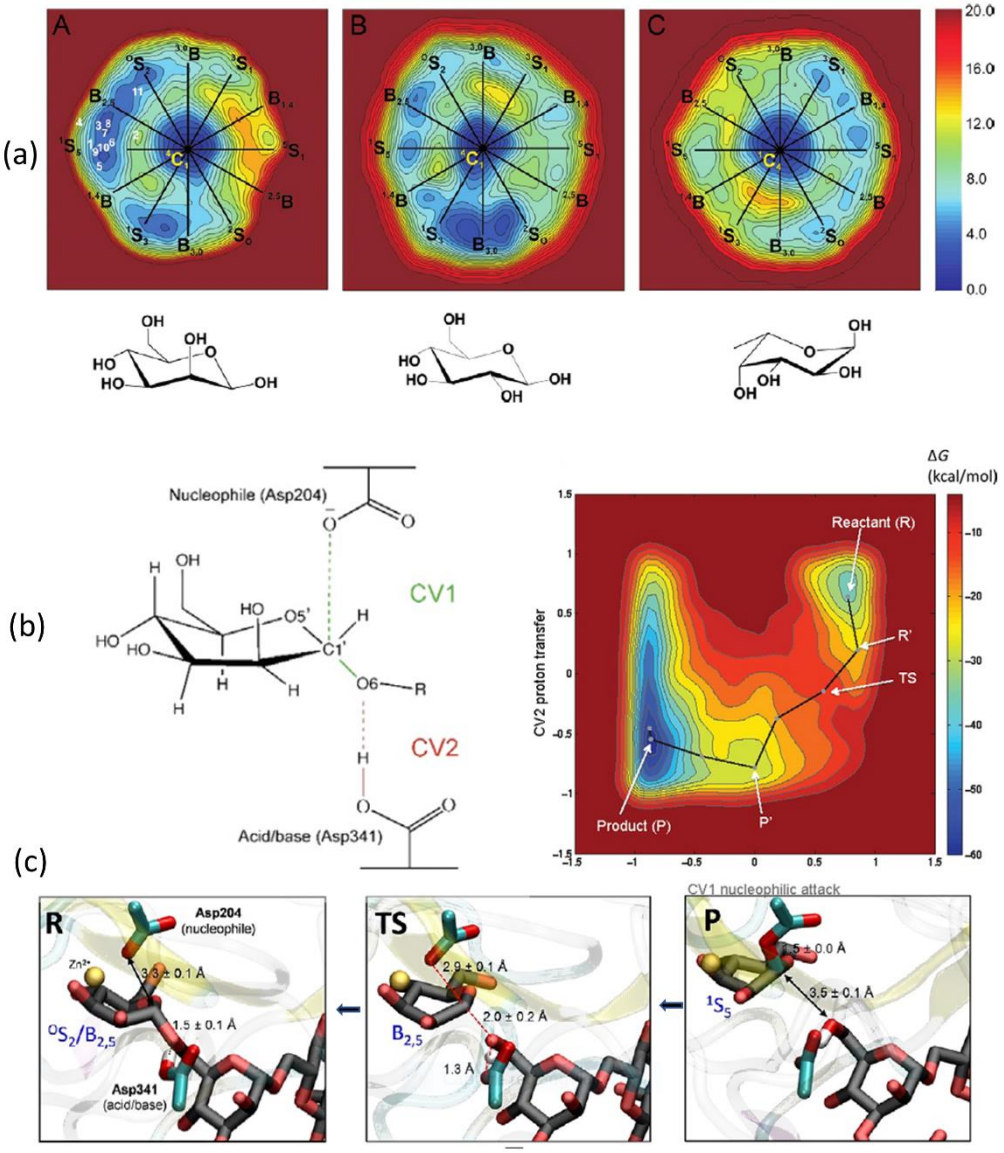


Figure 26. (a) Distribution of the canonical conformations on the computed free-energy surface of (A) β-D-mannopyranose, (B) β-D-glucopyranose, and (C) α-L-fucopyranose (southern hemisphere). Reproduced from ref 419. Copyright 2010 American Chemical Society (b) Results obtained for the QM/MM metadynamics simulation of glycosidic bond cleavage in α-mannosidase II. Collective variables definition. Evolution of the CVs during the metadynamics simulation. Structures of the reactants (R), transition state (TS), and glycosyl-enzyme intermediate (P). Reproduced with permission from ref 425. Copyright 2016 Elsevier

The free energy surfaces for α-D-glucopyranose,420 α-Lfucopyranose, 422 and β-D-mannopyranose419 were calculated using metadynamics simulation using the CPMD formalism and Cremer and Pople puckering.423,424 Figure 24 displays the calculated free energy maps and the evident qualitative differences. Low-energy regions (the most stable minima) are located on one side of the diagram but shifted from the southwest (α-D-glucopyranose) to the northeast (α-L-fucopyranose) and northwest (β-D-mannopyranose). The agreement between the experimental data derived from X-ray structures of the complexes corroborates the

computation results and suggests an evolution of the corresponding glycosidases to select those conformations that require less energy for ring distortion (Figure 26).

As for inverting the glycoside hydrolases, the catalytic mechanism underlying the inversion of configuration proceeds via a single displacement mechanism. Because of its simplicity, a limited number of reported investigations are reported for inverting enzymes (Figure 25).426−429 One dealing with endoglucanase from GH8 was investigated with the DFT QM/MM method. Another group characterized the nature of the reaction's transition state and the conformations adopted by the glycan xylopyranosyl ring along the reaction pathway under the action of an inverting β-xylosidase.426,427 A series of articles have been reported describing the application of computational methods in retaining glycosidases.430−440

7.4. Enzymatic Degradation of Polysaccharides in the Solid-State

Some polysaccharides occur in very compact three-dimensional arrangements, resulting from extensive networks of inter and intramolecular hydrogen bonds and van der Waals interactions. These characteristics make the structures utterly insoluble in water (e.g., cellulose, chitin) and highly resistant to attack by most enzymes. The degradation of cellulose to glucose requires the cooperative action of three enzymes, collectively known as cellulases. Spatial models of cellulose degradation must consider effects such as enzyme crowding and surface heterogeneity, which result in reduced hydrolysis rates (Figure 27).

The first QM/MM study of the cellulose hydrolysis mechanism focused on the mechanism of endoglucanase (GH8 family). The free energy landscape's chemical reaction path and the obtention were obtained through DFT calculation using QM atoms and metadynamics. A reduced number of collective variables was used. The quantification of the free energy barrier of the reaction could not be assessed.427 The simulations reproduced the concerted one-step general inversion mechanism. They confirmed the identity of the general base residue and the boatlike conformation of the transition state. The complete reaction pathway for the hydrolysis of cellulose was investigated using transition pathway sampling. The results of the investigation indicate that deglycosylation proceeds via a product-assisted mechanism in which cellobiose interacts with a water molecule for nucleophilic attack on the anomeric carbon atom of the glycosyl−enzyme intermediate.441

Recent years have seen the discovery and characterization of lytic polysaccharide monooxygenases (LPMOs) that significantly improve cellulose degradation.446 The first QM/MM study of H2O2-dependent activity in the AA9 LPMO family shows that the catalysis involves forming a caged hydroxyl radical and a Cu(II)-oxyl intermediate that oxidizes the H C4 bond of the polysaccharide substrate.447,448 In parallel to these investigations aimed at understanding the mechanisms of GH and capturing the complex reaction coordinates and conformational route that substrates follow throughout the catalytic pathway by the quantum chemical method, another research direction is exploring the potential of the CG method. Indeed, the size of the complete system to be studied requires the development and applications of such low-resolution CG representations.449 The authors use the MARTINI representation450 and reparameterized cellobiose so that the nonbonded interactions are fitted to reproduce the water-cyclohexane partitioning free energies for a series of cello-oligomers. The extrapolation of the model to longer cello-oligomers and the assignment of the specific nonbonding interactions to cellulose resulted in a model that yields a stable, ordered structure in water that closely resembles the crystal structure of cellulose Iβ. The potential of the methods is illustrated by the successful simulation of the motion of the carbohydrate-binding domain of Cel7A (the fungal cellulase from *Trichoderma reseei*) on the surface of crystalline cellulose.

The representation of the cellulose surface layer as a twodimensional grid is another way to illustrate the significant events during the enzymatic degradation of cellulose at the mesoscopic level. The calculation considers the free and bound states of endo- and exo-cellulases with explicit surface reactive terms (e.g., hydrogen bond breaking, covalent bond cleavage) and corresponding reaction rates.451

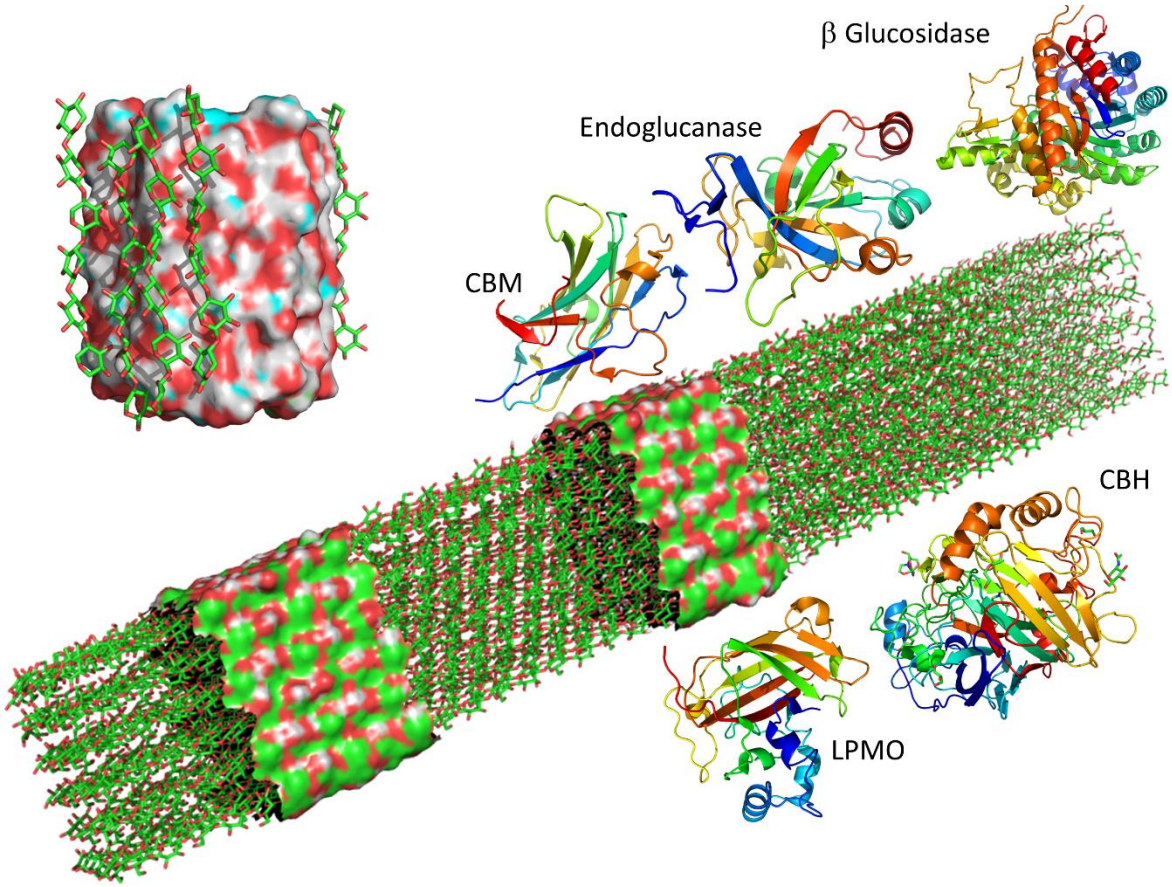


Figure 27. Interaction and deconstruction of cellulose. The middle portion of the figure shows an idealized construction of the parallel arrangement of 36 cellulose chains constructed for an MD simulation with full consideration of hydration. The water molecules are not displayed. Two sections of the microfibrils display the van der Waals surface to highlight the potential for interactions. The left-hand side panel displays the interactions between the cellulose microfibril and xylan (one of the constituents of the plant cell) resulting from a molecular modeling study.305 The figure shows the three-dimensional structures of the representative cellulolytic enzymes degrading cellulose down into glucose: endoglucanase, cellobiohydrolase, and β-glucosidase. They follow the action of the carbohydrate-binding module and lytic polysaccharide monooxygenases (LPMOs), which significantly enhance cellulose breakdown. CBM, PDB: 4B96; endoglucanase, PDB: 5XBU;442 β-glucosidase PDB 3WH5;443 CBH, PDB: 505D;444LPMO, PDB: 6RW7445. Reproduced from ref 5. Copyright 2021 Elsevier

Cellulose interactions with other macromolecules, such as proteins,449 enzymes,450,452 and polysaccharides, have also been investigated by CG simulations. Enzymes, such as *Trichoderma reesei* cellobiohydrolase (CBH) I (Cel7A), are responsible for cellulose hydrolysis and thus are industrially crucial for cellulose decomposition. The translational motion and thermodynamic driving forces of the carbohydrate-binding module (CBM) of Cel7A on the cellulose surface were investigated with the Martini CG model450 and mixed atomistic−CG simulations 452, respectively. Specific CG models for cellulose interacting with xylan305 and ionic liquids453 were also developed and provided fundamental insight into the interactions between the different species.

7.5. Transport

The transport of carbohydrates is critically essential for the proper functioning of many cellular processes. These molecules diffuse poorly across membranes and must travel across the channel and pores. Such a role is achieved by a family of membrane protein secondary transporters belonging to the Major Facilitator Superfamily (MSF).454 Instead of using ATP directly for transport, MSF uses an existing electrochemical gradient. Materials can move against the concentration gradient. Transport can concern just one substrate (uniporter) or two

substrates in the same direction (symporter) or in the opposite direction (antiporter). These transmembrane proteins allow the permeation of low molecular weight carbohydrates.

The subject of intense research is elucidating their three-dimensional structures and modeling the mechanistic transport model. During the past decade, a combination of high-resolution structures of transporters in multiple functional states and biochemical analysis, biophysical analysis, and in silico MD computations led to interpret more specific details of the fundamental transport process, to the point that the functional motions of membrane transporters can be visualized with molecular dynamics simulations.455 Glucose transporters (GluTs) are among the most prominent families of membrane transporters. They occur in all the kingdoms of life. They provide the pathway to transport mono and disaccharide across the membrane. Molecular dynamics simulations have been performed on the human glucose transporter GluT1,456, a neuronic protein glucose transporter 3.457. The molecular basis of the Glucose Transport mechanics in plants was also investigated.458 The mechanism of the conformational transition of the human glucose transporter was studied in the absence and presence of glucose. It has been investigated through extensive MD simulations of the GLUT1 transporter.459 One of the salient features resulting from this investigation is the characterization of the behavior of glucose while residing in the central cavity of the protein. Glucose interacts with the surrounding amino acids through multiple hydrogen bonds upon entering the protein cavity, contributing to a favorable enthalpy. At the same time, glucose undergoes numerous rotations and axial movements that contribute to favorable entropy. Within the protein cavity, the free energy of glucose binding may be lower than in other ligand positions. These continuous movements affect the orientation of the side chains, which impacts the mechanics of the protein, allowing the progression of glucose and slowing down the process of glucose transfer to the intracellular medium.

Resolution of the crystal structure of E. coli maltoporin460 suggested that maltose and malto-oligosaccharide could be translocated across the membrane by guided diffusion through the track. The mechanistic aspects of maltotriose translocation across the maltoporin membrane channel were deciphered by molecular dynamics simulation. Like the general diffusion transport mechanism, the movements involve continuous conformational changes.461 A ”greasy slide” within the channel is formed by a continuous stretch of aromatic residues forming a left-hand path. The first event is the binding of maltotriose to the first residue of the ”greasy slide”, which occurs via van der Waals interactions with the hydrophobic side of the glucosyl ring. Deeper penetration into the channel occurs throughout the guided diffusion of the oligosaccharide with the ”greasy slide”. Due to the progressive dehydration of the oligosaccharide, short hydrogen bonds promote interactions between the hydroxyl groups of the maltotriose and the surrounding amino acids. It is due to the conformational flexibility of the glycosidic bonds and primary hydroxyl groups. The presence of charged side chains (called ”polar tracks”) mimics the first hydration shell of maltotriose, providing hydrogen bonds to its hydroxyl groups. Alongside the ”greasy slide”, the polar tracks are divided into donor and acceptor lanes. The movement of maltotriose to the following binding site of the “greasy slide” occurs with a rearrangement of hydrogen bonds. The authors describe such an arrangement as the “register shift”. The movement of maltotriose through the porin occurs capillary fashion due to hydrogen bonds’ continuous creation and breaking.

Characterizing the machinery around the maltose transporter is still a matter of computational investigation. The analysis of the free energy landscapes for the opening-closing of the maltose transporter ATPase MalK(2) has been conducted using enhanced-sampling molecular dynamics.462 In contrast, an all-atom molecular dynamics simulation dealt with the mechanism of nucleotide-binding domain dimerization in the intact Maltose Transporter.463

7.6. Lectins

Lectins are proteins of nonimmune origin that bind mono- and oligosaccharides reversibly and specifically while not exhibiting any catalytic or immunological activity.464 The established roles of lectins in biology relate to fertilization, embryogenesis, inflammation, and metastasis. Other essential functions deal with parasite−symbiote recognition and invertebrates and plants. Throughout their interactions with carbohydrates, lectins play a crucial role in innate as they facilitate discrimination between self and non-self.465 Within one of the ongoing concepts of the glycocode, lectins could be considered molecular readers deciphering the high-density bio-encoding of complex carbohydrates.

While lectins display a very high specificity toward monosaccharides, the affinity describing the interaction in an individual binding site is not high (typically in the millimolar range). Affinity expresses the thermodynamic principles which govern any reversible biomolecular interaction. Lectins exhibit diverse folds and display several modes of “architectural multivalence”. Such a feature results in a strong avidity. It reflects the overall binding strength of the interaction between a carbohydrate and a multivalent lectin. Avidity cannot be described by thermodynamic terms but by kinetic measurements. 466

Rich literature deals with the three-dimensional structures of lectins, most coming from X-ray crystallography (Figure 28). At the end of 2021, among the 2400 structures deposited in the PDB,471 1500 are reported to be protein−carbohydrate complexes.472 Such an ensemble of highly characterized structures offers a rich portfolio to test and develop computational methods. A wide variety of lectins have been studied using molecular docking. Nevertheless, accurate determination of carbohydrate−lectin complexes remains a nontrivial problem because of many lectins' shallow and multichambered binding sites.

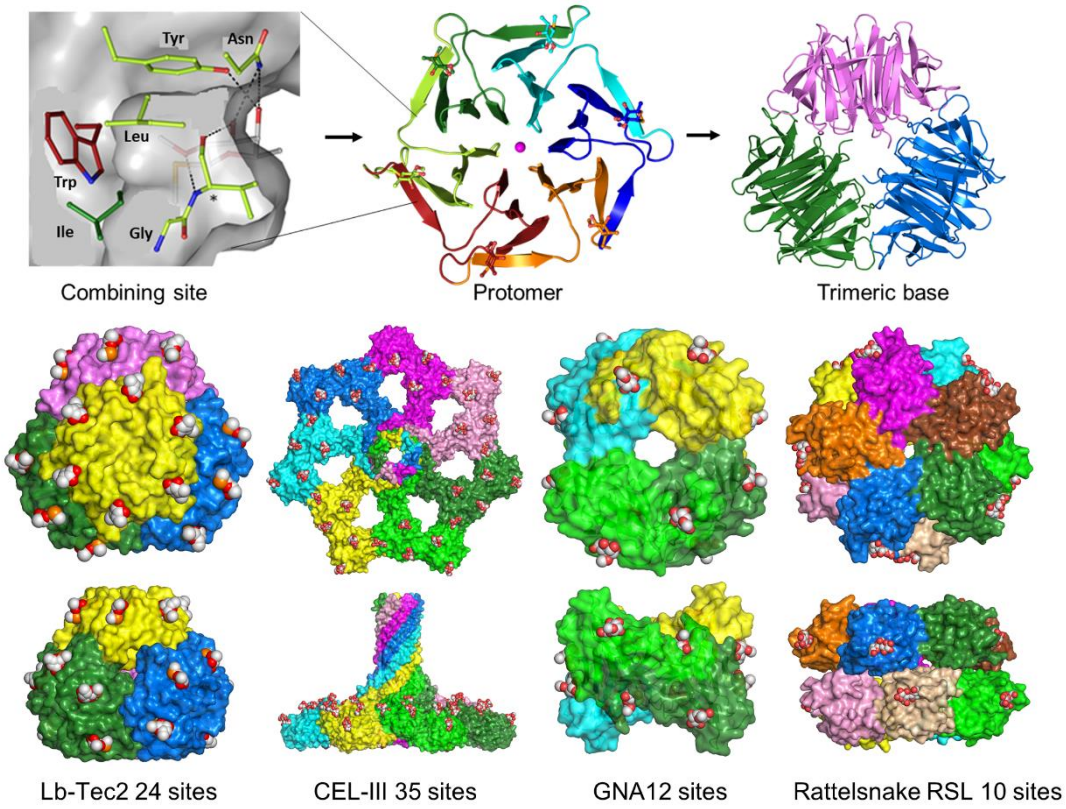


Figure 28. Topography of lectin structural arrangement: from carbohydrate combining to a trimeric base. Top and side surface representations of highly multivalent lectins with their binding-site distribution. From left to right: Lb-Tec2 (PDB: 5FCB) CEL-III from sea cucumber with lactulose (PDB: 3W9T),467 GNA from Galanthus nivalis with α-methyl-D-mannoside (PDB: 1MSA),468 and rattlesnake venom; galactose-specific C-type lectin (RSL) with lactose (PDB: 1JZN)469 (drawn with PyMol14). Reproduced with permission from ref 470. Copyright 2018 American Chemical Society.

When used in glycan conformational studies, molecular modeling methods provide ways to extend the study of protein-glycan interactions. These simulations are among the most potent methods to obtain atomistic-level information on the molecular recognition of highly dynamic systems, such as glycans. They offer some avenues to explore whether recognition pathway processes between two opposed processes are described as ”conformational selection” and ”induced fit”. Following the conformational selection theory, the receptor will bind only the conformers corresponding to the final bound conformation, which defines an adequate substrate concentration in practice. The induced-fit theory explains that recognition occurs regardless of the substrate’s specific 3D conformation. The substrate will fold in place to match the spatial constraints of the receptor-binding site. Borrowing from the “intrinsically disordered proteins” field, an intermediate case scenario describes that

the substrate can form local 3D motifs, known as molecular recognition features (MoRFs)473, that the receptor recognizes. MoRFs act as nucleation sites initiating a folding-in-place process. Because of their highly dynamic architecture, glycans are intrinsically disordered biomolecules.

Consequently, it is rather difficult, if not impossible, to experimentally determine how lectins or glycan-processing enzymes recognize them. Significant steps have occurred, most notably the development and implementation of force fields capable of accounting for glycan specificity and compatibility with those developed for proteins. The conformational flexibility of glycans must be characterized and considered at every research step.

The resolution of the crystal structure of the β propeller lectin from Ralstonia solanacearum (RSL) opened up a very demanding case with the observation of an unexpected ”open” conformation of the Lewis X trisaccharide (LeX)474 (Figure 29). The structure's high resolution established the observation's validity, away from the observations gathered for the unbound conformation of Lewis X (LeX) in the solid-state, solution, or other lectin complexes.475,476 A conventional (nonaccelerated) molecular dynamics simulation of two LeXRSL complexes of 1 and 0.85 μs trajectory followed the detailed crystallographic analysis of the open RSL-LeX complex. This investigation determined both the binding and unbinding pathways of LeX to and from the lectin.477 Moreover, the analysis of thirty 1 μs long trajectories from uncorrelated conformations probed the dynamic of the conformations of LeX in solution. One lesson from this analysis shows that enhanced sampling protocols provide results in structure, dynamics, and energetics compared with those obtained from microsecond-long multiple trajectories. For the time being, the application of exhaustive sampling is limited to tri- and tetra-saccharides, even if it faces a high-energy transition. However, it rapidly becomes unattainable as the glycan’s complexity and size increase.

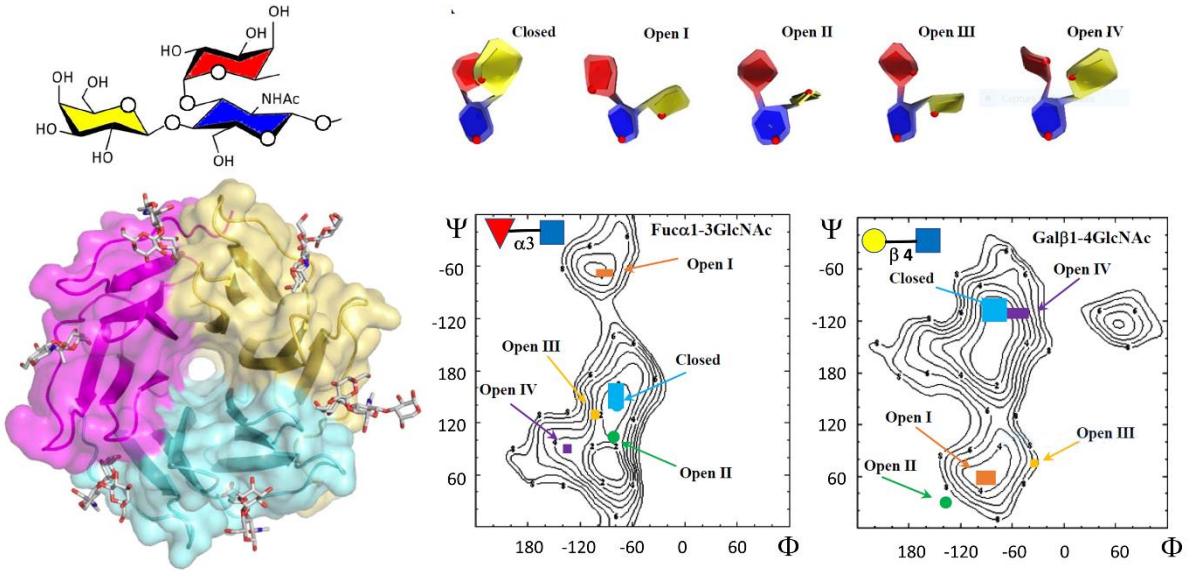


Figure 29. Hidden conformations of LewisX (LeX): β-D-Gal (1-4) (α-L-Fuc (1-3), β-D-GlcNAc). The conformations of fucosylated Lewis oligosaccharides are considered rigid, adopting a single shape referred to as the “closed” conformation. This rigid shape is due to stacking between fucose (Fuc), and galactose (Galp) rings by a nonconventional−Ohydrogen bond and by steric hindrance of the theN-acetyl group of GlcpNAc. The crystal structure of the crystalline conformation of LeX trisaccharide, together with NMR and modeling data, confirmed that the trisaccharide presents only limited conformational fluctuations around the closed shape. When bound to the Ralstonia solanacearum lectin, the LeX core adopts several distinct conformations (open). Extensive molecular dynamics simulations confirm rare transient LeX openings in solution, frequently assisted by distortion of the central N-acetyl glucosamine ring. Additional directed molecular dynamic trajectories revealed the role of a conserved tryptophan residue in guiding the fucose into the binding site. The conformations taken by these different conformations are displayed on the corresponding (Φ, Ψ) maps (drawn with SweetUnityMol478)

The results from comprehensive molecular dynamics simulations confirmed the stability of the LeX open conformation in its interaction with RSL. They provided a detailed understanding of the ensemble of specific interactions that stabilize this open form. The lengths of the MD simulations of LeX were long enough to capture how the opening events occur. The opening pathway involves a concerted change of the GlcpNAc ring puckering and two glycosidic linkages. The LeX structure's opening happens in solution but is too rare an occurrence to be picked up by NMR. The enthalpy of binding is estimated to be around 10.6 kJ/mol. The nature of the interactions and the morphology of the lectin binding site compensate for such a cost (Figure 29). A series of contributions addressed predicting rare events in solution through computational methods. 193,196,479−481 In the case of LeX and sialyl LeX, the evaluation of the conformational free-energy surfaces was conducted under four different computational protocols:482 (i) multidimensional variant of the swarm-enhanced sampling molecular dynamics method (msesMD), (ii) accelerated MD,483 (iii) microseconds unbiased MD, and (iv) umbrella sampling. The authors discussed the potential of the msesMD simulation in its capacity to analyze glycan conformations.

7.7. Antibodies

Many pathogens and aberrant malignant cells display unique glycans on their surfaces. Minimal carbohydrate epitopes occur at the terminal end of more complex polysaccharide chains, experiencing various contexts and environments or surface densities. Consequently, antibodies with similar specificities for individual glycan epitopes may display a different cell-selective profile depending on their unique presentations in targeted cells. It is typically the case for many tumor-associated carbohydrate antigens expressed by tumors or normal tissues at a lower level.484

New carbohydrate-based vaccines have been designed, and some have reached clinical phase studies, such as *Streptococcus pneumoniae, Neisseria meningitides, Haemophilus influenza* type b, or *Salmonella enterica*. The extension of such developments offers new opportunities for cancer immunotherapy. Carbohydrate antigens recognized by preformed and elicited antibodies are essential in blood group transfusion and organ transplantation. In these cases, the antibody-recognized carbohydrate determinants are expressed on the cell surface as glycolipids and glycoproteins.

New challenges aim to target carbohydrates on the surface of bacteria, protozoa, helminths, viruses, fungi, and cancer cells for vaccination purposes as the identification and evaluation of unique carbohydrate epitopes on a plethora of pathogens and malignant cells becomes available.485

Although monoclonal antibodies are clinically effective tools, those against carbohydrates tend to have low affinity and complex or mixed specificity. Glycan antigen recognition may depend on glycan density, valence, presentation, and flexibility. 486 There is an important limitation in exploiting glycans as disease markers or therapeutic targets due to the scarcity of high specificity and high affinity of specific antibodies against carbohydrate targets. In such complex environments, experimental methods cannot establish glycan conformations and must be complemented by systematic computational modeling, the antibody and the glycan-antibody epitope. Elucidating the molecular basis of the complexes' formation and the balance between the enthalpic and entropic contribution involved in the binding are required.487 Recognizing that computational approaches do not lead to one plausible model, a confrontation with experimental remains essential to select the most likely plausible models. An example of how such a computational, experimental approach can lead to a rational design of potent antibodies targeting glycans is given by elucidating the structural origin of antibody recognition of sialyl-Tn antigen, a tumor-associated carbohydrate antigen (TACA). The information from a glycan microarray screening and STD-NMR experiment helped select an optimal three-dimensional model of the glycan antibody (Sialyl-Tb-mAbTKH2) by molecular dynamics docking simulations.488 Compared with many docking studies on carbohydrate−lectin and carbohydrate−enzyme recognition, there are few published computational-aided carbohydrate-antibody recognition studies.

Most life-threatening septicaemia, meningitis, and pneumonia cases occur from the deleterious action of surface capsular polysaccharides on bacteria. While these bacterial polysaccharides may have similar carbohydrate sequences, they differ in immunogenicity, antigenicity, virulence, and geographical dispersion. Moreover, diseases might occur from several strains that circulate simultaneously in a region, such as *N. meningitidis, Shigella flexneri, S. pneumoniae, Klebsiella pneumoniae*, etc. The broad protection conferred by vaccines often relies on several serogroups/serotypes (vaccine valency). The subsequent task is to design a vaccine with maximal coverage against prevalent bacterial strains with minimal vaccine valency. The valency of

a vaccine can be reduced if the selected vaccine serotype(s) can confer protection against closely related nonvaccine serotypes. 489

The contribution of molecular modeling aims at providing detailed descriptions of the dynamic motion of the polysaccharide chain to enlighten the effects of substituents (or lack of) on the backbone conformation in chemically similar carbohydrate chains.490−493

Molecular dynamics simulations proceed with progressively extending chains while increasing the simulation length. This approach establishes prevalent conformations and starts with initial carbohydrate structures with low-energy conformations. At this stage, the simulation trajectory will likely converge quickly and explore the solution's most general conformations of the native carbohydrate chains. Two metrics of chain flexibility are needed to assess the convergence of the MD simulations. They are the end-to-end distance, t, and the radius of gyration of the polysaccharide chains. A block-averaging analysis assesses the simulation convergence. This algorithm divides a simulation trajectory, having N frames, into a set of M ”blocks” of length n frames, N = Mn. Then, the selected metric (e.g., the radius of hydration, end-to-end distance) is averaged in each block. The block (n) length is progressively increased, and the block averages are recalculated at each value of n. The standard deviation in the block set yields the blocked standard error for each value of n. The convergence of the simulation is assessed whenever the current estimate of the blocked standard error asymptotes to a plateau. It represents the standard error in the estimate of the mean.

S. fLexneri is the primary causal agent of the endemic form of bacillary dysentery (Figure 30). The O-antigen is the polysaccharide moiety of the lipopolysaccharide; it is the principal target of the serotype-specific protective humoral response elicited upon host infection by *S. flexeneri*. Among the 30 serotypes reported, 29 with a common repeat unit differentiate through variations in the lateral substituents. Whereas a quadrivalent serotype-based vaccine could enhance broader protection, the question arises as to whether other serotypes should be considered. The application of the above-described computational method aims at identifying whether the quadrivalent vaccine containing serotype could provide coverage against other prevalent serotypes.495

The central role of carbohydrates in blood group transfusion and organ transplantation dramatically underscores the importance of glycan-protein interactions in major biological processes. The families of antigens called ABH(O) and Lewis determinants are among the significant carbohydrate determinants of blood groups. Most ABO antigens are expressed on human erythrocytes at the ends of long polylactosamine chains, a minority of the epitope occurring on neutral glycosphingolipids. Despite the critical role played by these determinants, the description at the three-dimensional level of the interactions between antigens and antibodies is slowly accumulating. Without detailed structures, a comprehensive study of the cross-reaction patterns of 9 antibodies against 12 carbohydrate antigens was conducted using computational methods. Three-dimensional descriptors of the molecular properties of carbohydrate antigens were used in comparative molecular field analysis (COMFA). Processing the quantitative structure−activity relationship (QSAR) data provided insights into the carbohydrate epitopes required for antibody recognition while also providing insight into the nature of molecular recognition.496,497

The limited number of antibody−carbohydrate docking studies reflects the paucity of crystal structure complexes at high resolution. The structural analysis of these available structural data indicates some general trends about how such antibodies recognize different types of carbohydrates. Those antibodies which recognize a terminal carbohydrate motif display a cavity-like binding site, reminiscent of the combining sites found in lectins, where the insertion of one or more monosaccharide residues occurs at is ”end-on”. Antibodies that recognize an internal carbohydrate motif (a single or several repeat units) of a polysaccharide display, in general, groove-like binding sites or vast cavities which are ”open” at both ends of the sites allowing for ”side-on” entry of the antigen. While displaying some structural similarities with the catalytic sites of glycoside hydrolases, these features allow the carbohydrate to be recognized as a ”conformational epitope” as shown in the case of group B meningococcal α-2-8 linked sialic acid polysaccharide.498

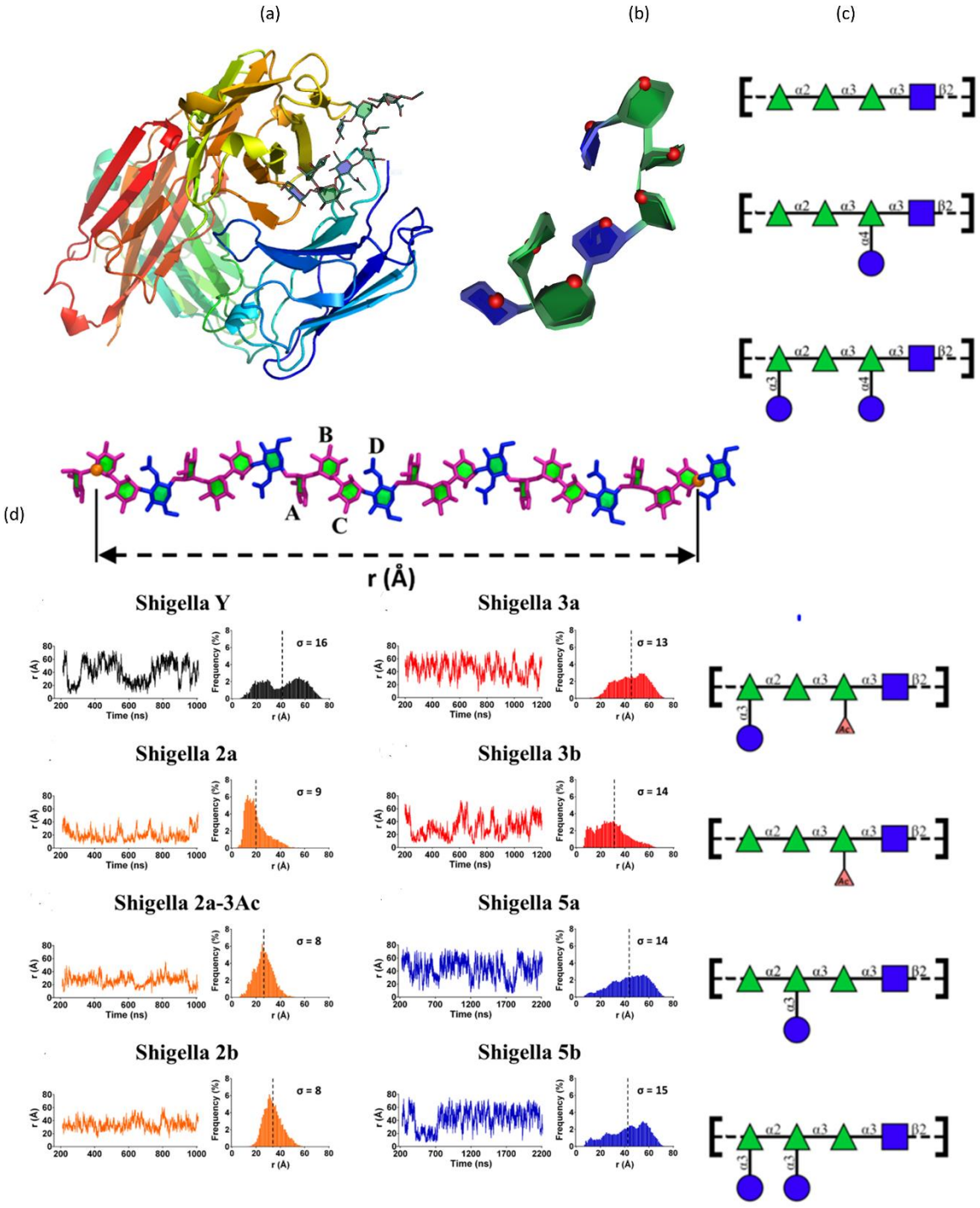


Figure 30. (a) Depiction of the three-dimensional structure of a decasaccharide fragment of *S. flexeneri* from serotype 2a in complex with a protective monoclonals antibody Fab F22-4 (PDB: 3BZ4494) (drawn with PyMol14 and SweetUnityMol478). (b) Structures All serogroups share the same serotype backbone but are distinguished with substitution. Results of the exploration of the effect of glucosylation on the conformation and dynamics of the S. flexeneri O-Ags, using descriptor r, as a simple measure of the molecular extension and flexibility, over a 1000 ns molecular dynamics simulation, as computed for 8 serotypes. The differences in the r time series (left column for each serotype) and corresponding distribution captured the significant influence arising from the variations in substitutions on a common backbone repeating unit. Adapted with permission from ref 495.

7.8. Glycosaminoglycans and Signaling

Molecular modeling of the structure, dynamics, and interactions of glycosaminoglycans (GAGs) assembles most of the difficulties of glycoscience, as they combine the challenges of glycans and polyelectrolyte polysaccharides. GAGs are a family of complex anionic polysaccharides, including (i) glucosaminoglycans (heparin and heparan sulfate), (ii) galactosylaminoglycans (chondroitin sulfate and dermatan sulfate), and (iii) hyaluronic acid and keratan sulfate (Figure 31).

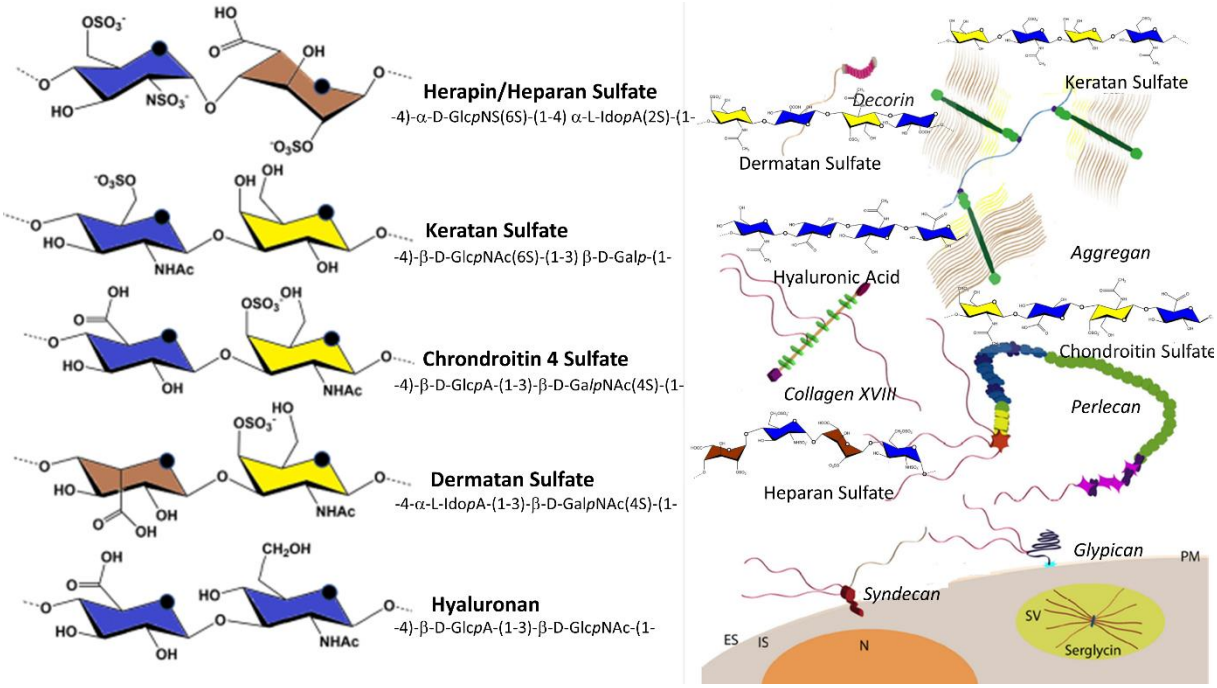


Figure 31. Cartoon representation of the chemical constitution of the five families of GAGs and of six categories of proteoglycans (aggrecan; decorin, perlecan, and collagen; glypican; and syndecan and serglycin). ES, extracellular; IS, intracellular; N, nucleus; SV, secretory vesicle. Reprinted in part from ref 499. Copyright 2008 Wiley-Liss. Reproduced with permission from ref 500. Copyright 2021 MDPI.

Several features pose difficulties in the computational modeling of GAGs and their interactions with proteins. GAGs exhibit structural and chemical heterogeneity. Various sulfation motifs and their distribution along the chain contribute to their sequence, conformational diversity, ring puckering,479 and high charge density. Many torsional angles between glycosidic angles and side chains create high conformational flexibility. Despite the limited number of constituting elements, the conformational flexibility of idpoyranoses increases to 32, which is the set of possible low-energy conformations to be considered elementary building blocks. Despite the large number of unique GAG disaccharides, which amounts to 202,501 macromolecular builders to deal with nonsulfated or sulfated glycosaminoglycans, allow grasping of their 3D features.502−504

It is difficult to characterize the impact of solvation and desolvation on GAG structures. Their size and heterogeneity necessitate multiscale modeling of glycosaminoglycans, from disaccharide fragments to polysaccharides.6 Computational studies have been performed on the GAG fragments, usually no longer than 5−10 monosaccharide units, shorter than the significantly longer natural polysaccharides. Attempts to accelerate and improve the construction and the simulation of GAGs have been developed.502,503,505,506 Moreover, they offer new perspectives on the computational simulations of GAGs.507

Characterization of a heparin disaccharide in optimized conformation by quantum methods, application of molecular modeling to hyaluronan decasaccharide in water, and multi-microsecond aqueous simulation of heparan and heterogeneous proteoglycans313 and glycosaminoglycans provide some illustrations of spatio-temporal realizations. CG simulation is particularly relevant to characterize the dynamic features of GAGs on long-time scales, despite the low resolution resulting from this kind of computation.508 GAGs display important physicochemical properties.509

GAGs bind to numerous proteins on the cell surfaces, including growth factors, cytokines, proteases, and coagulation enzymes, and within the extracellular matrix. For example, heparin and heparin-sulfate interact with several hundred proteins and numerous pathogens. These interactions mediate their biological activities and play essential roles in physiopathological processes such as growth factor control, anticoagulation, hemostasis, and cell adhesion.510,511 New data are being integrated to complement the investigation of the integration of glycosaminoglycans.501 As part of the interactions between proteins and GAGs imply electrostatic interactions, it is indispensable to consider solvent for the interactions, but the impact of solvation/desolvation on complex formation is difficult to assess.128,508 The challenge remains identifying well-defined and predicted GAG-binding pockets on bound proteins.151,512

GAGs interacting with proteins bear a high negative charge, and electrostatic interactions partly drive their binding onto protein surfaces.509 A preliminary step is the calculation of the electrostatic potential isosurfaces for the protein that indicates the location of the GAG binding region (Figure 32).513 A complementary way to assist the determination of the binding site may result from rigid or multiple docking of a GAG fragment(s).514 Such a local docking approach guide the initial search on grossly determined binding sites. Nevertheless, the challenge remains to predict a binding pose, even for short GAGs, whenever a putative binding site on a protein is known.

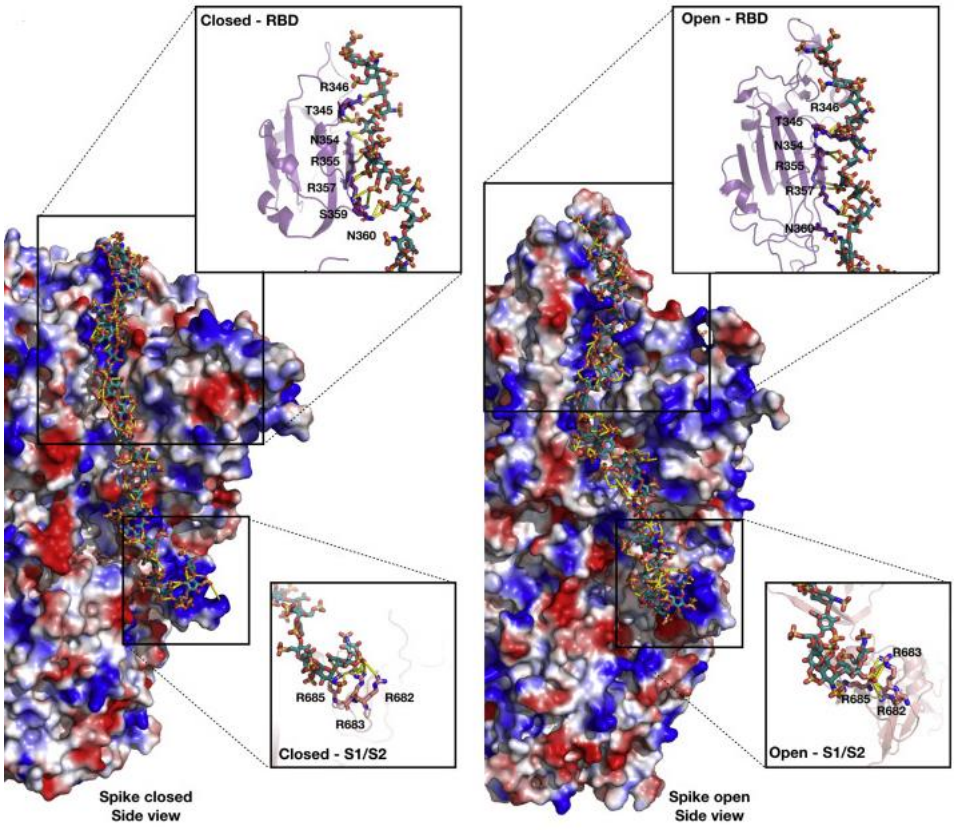


Figure 32. Stability of the spike−heparin complexes. Representative closed (left), and open (right) structures obtained after MD simulation of spike bound to three heparin chains are displayed as molecular surfaces with electrostatic potential mapped onto them to show the partially grooved positively charged path occupied by a heparin 31mer. Heparin is shown in a stick representation colored by elements with cyan carbons. The yellow dashed lines show H-bonds between the spike and heparin. The insets for the closed and open conformations highlight the H-bonding interactions between heparin and the residues in the RBD (T345, R346, N354, R355, N360) and S1/S2 (R682, R683, R685) HBDs shown in stick representation with carbons colored according to the subunits to which they are bound. Reproduced with permission from ref 513. Copyright 2021 MDPI.

In the quest to evaluate the performances of docking protein-GAG complexes with several widely used docking programs, a comparison established that the commercial program Glide and the free docking software Autodock515−517 provided reasonably good performance.153 The authors conclude that an adequate prediction of the protein−GAG biding pose is best achieved when clustering and experimental data are applied.518 Nevertheless, the situation worsens when constituting monosaccharides are higher than six.153 Other investigations using conventional docking protocols and tools also showed limited success in investigating protein-GAGs systems with modest fragments having lengths between four and six monosaccharides.519−521

The adaptation of the program for docking flexible molecules522 to the GAG-protein system first used a "coarse6docking" approach to predict the binding site, followed by a "fine-grained" calculation to further minimization. As for docking more extended GAGs into a protein, a molecular dynamics docking method based on a steered MD simulation proved to perform better than other approaches at the cost of requiring more computational resources.523

A fragment-based docking approach has been explored in the quest for new computational tools capable of dealing with more extended GAGs. The GAGs fragments are separately docked on the proteins, and the results are used to rebuild a profile of massive conformations of the entire GAG.524−526 As applied to GAGs, the fragment-based docking protocol uses an incremental ligand construction starting from the docking seed pose of one fragment. The accuracy of the seed pose determines the success of the operation. Another fragment docking approach has been used for binding site prediction throughout evaluating the number of contacts with docking poses made by the protein residues.514,527 Starting from a coarsely identified protein binding site, a novel fragment-based method with a fully flexible long GAG ligand was validated on a benchmark of 13 experimentally known GAG−protein structures.528

There is a continuous exploration of approaches to tackle the complexity of solving the interactions of proteins with GAGs of suitable length. In particular, molecular docking coupled to MD simulations provides a way to explore the induced-fit mechanism of protein−GAG binding. The evaluation of complex stability, as well as the refinement and the rescoring of docking poses.125,153,528,529

GAG−protein interactions constitute an intense area of research to understand fundamental biological problems and develop new bioactive molecules of therapeutic value. In experimental and computational studies, almost half of the studies concern heparin, frequently used as a structural analogue of heparin and heparan sulfate proteoglycan.512 It is easy to understand given the wide range of biological functions heparin performs and the importance of designing heparin-like drugs to treat coagulation disorders, abnormal inflammatory or immune responses, and angiogenesis-dependent diseases. It explains the relatively low number devoted to studying chondroitin sulfate, despite its biological relevance. A few computational investigations aim to design innovative functional materials, to control and promote application processes in bone and skin regeneration. Over 20 protein GAG systems have been characterized in the last 15 years.

The development of new force fields40,506, scoring functions and databases has contributed to understanding GAGs in conjunction with data from experimental techniques. These have proved relevant and valuable for the detailed characterization of structure−function and structure−property relationships.29 A flowchart describing the use of computational approaches to be addressed and critical questions about GAG-protein interactions have been proposed to understand a series of critical points, such as the following: (i) Is this protein interacting with GAGs? (ii) Where does the binding occur? (iii) Is there an optimum GAG sequence for the binding? (iv) Is the GAG−protein complex stable?128

Many computational modeling challenges remain: (i) to investigate the role of solvation and counterions, (ii) to elucidate the contribution of the polyelectrolyte nature of GAGs to the interaction, (iii) to characterize GAG chains with a degree of polymerization greater than 10, which exhibit a high degree of conformational flexibility, (iv) to develop new docking protocols based on the fragment-based approach or coarse-grained techniques, (v) to elucidate the specificity of the interactions with proteins and the role of GAG degree of polymerization and sulfation patterns, (vi) to explore whether GAG molecules induce allosteric effects on their target proteins, (vii) to assess the thermodynamic and kinetic features of protein−-GAG systems, and (viii) to decipher how GAG−protein interactions take place in the context of the multicomponent nature of the peri- and extracellular matrix.

8. STRUCTURAL GLYCOBIOINFORMATICS. TOOLS AND DATABASES

Future scientific output in glycoscience represents the publication of some 70,000 research articles annually. The availability of powerful high-performance computing and search capabilities dramatically amplifies the space of accessible information. In addition to offering the generalization of network analysis of a semantic corpus to study the dynamics of science, technology, innovation, and knowledge production, such data mining opens up new opportunities for discovery. This new vision emphasizes the value of analyzing and curating the raw data of published works. Untapped wealth can be harvested from the collected data sets from which the extracted information is translated into knowledge. Such advances benefit the entire field of structural glycoscience by developing tools and databases in various online resources. They cover the structures of glycans and glycoconjugates, the enzymes responsible for their biosynthesis and degradation, and the binding of glycans to human pathogens, glyco-epitopes and their antibodies as typical examples. The development of enabling technologies has led to a rich computational toolbox for solving the 3D structures of complex glycans and their identification and manipulation. Simultaneously, several computational tools and databases have been developed through the activities of several independent research groups worldwide. By early 2021, more than 150 entries were available on the Internet as open-source applications.

8.1. Applications to Building and Visualization

8.1.1. Structure Modeling.

CHARMM-GUI Glycan Modeler: in silico N-/O-glycosylation of proteins; modeling of carbohydrate-only systems, web service[530] (http://www. charmm-gui.org/?doc=input/glycan). CHARMM-GUI Glycolipid/LPS Modeler

CHARMM-GUI Glycolipid/LPS Modeler: Glycolipid and lipoglycan structure modeling web service[530] (http://charmmgui. org/?doc=input/glycolipid; http://charmm-gui.org/?doc= input/lps).

MARTINI General Purpose Coarse-Grained Force Field: The Martini force field is a CG force field suited for molecular dynamics simulations of biomolecular systems[86] (http:// cgmartini.nl/).

Glycosylator: Rapid modeling of glycans and glycoproteins (including glycosylation) based on a CHARMM force field, Python framework[531] (https://github.com/tlemmin/glycosylator).

Rosetta Carbohydrate: Modeling various saccharide and glycoconjugate structures (including loop modeling, glycoligand docking, and glycosylation), Python framework[131,532–534] (https://www.rosettacommons.org/docs/latest/application_documentation/carbohydrates/WorkingWithGlycans).

Azahar: Monte Carlo conformational search and trajectory analysis of glycans, Python framework, PyMol plugin[535] (https://github.com/BIOS-IMASL/Azahar).

Shape: Carbohydrate-dedicated fully automated MM3-based conformation simulation, stand-alone software[98] (https:// sourceforge.net/projects/shapega/)

Glydict: MM3-based N-glycan structure prediction based on MD simulations, web service[536] (http://www.glycosciences.de/modeling/glydict/).

GLYGAL: MM3-based conformational analysis of oligosaccharides, stand-alone software[97].

Fast Sugar Structure Prediction Software (FSPS): Automatic structure prediction tool for oligo- and polysaccharides in solution.[537]

8.1.2. Structure Building and Model Preparation.

doGlycans: Preparing carbohydrate structures (including polysaccharides, glycolipids and glycoproteins) for GROMACS atomistic simulations[538] (https://bitbucket.org/biophys-uh/doglycans/src/master/).

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)[41] (https://glycam.org).

SWEET-II: Rapid 3D model construction of oligo- and polysaccharides with MM3 optimization[539] (http://www. glycosciences.de/modeling/sweet2/).

Polys-Glycan Builder: Structure generation of polysaccharides and complex oligosaccharides from MM3-precalculated glycosidic linkage torsions and energy minimization[504] (http://glycan-builder.cermav.cnrs.fr/).

8.1.3. Glycosylation Modeling and Grafting.

GLYCAMWeb Glycoprotein Builder: Attaching a glycan (user input) to a protein (PDB file), web service (http://glycam.org/gp).

GlyProt: in silico generation of N-glycosylated 3D models of protein[540] (http://www.glycosciences.de/modeling/glyprot/ php/main.php).

Phenix CarboLoad: Loading a carbohydrate structure into protein model and PDB file generation[541] (https://www. phenix-online.org/documentation/reference/carbo_load. html**).**

GLYCAM-Web GlySpec (Grafting): Prediction of glycan specificity by integrating glycan array screening data and 3D structure[3,542–545] (https://glycam.org).

Noteworthy stand-alone programming frameworks for structure modeling are Glycosylated (modeling of glycans, glycoproteins and glycosylation) and Rosetta Carbohydrate (loop modeling,[131] glycan-to-protein docking, and glycosylation modeling).

8.1.4. Biological Membranes and Micelles.

Three applications are available to construct biological membranes and micelles.

CHARMM-GUI Membrane Builder: Constructs complex glycolipid, lipopolysaccharides and lipo-oligosaccharides inserted in biological membranes[334,530,546–548] (http://www.charmm-gui.org/?doc=input/membrane.bilayer).

GNOMM: A lipopolysaccharide-riched bacterial outer membranes can be constructed (http://thalis.biol.uoa.gr/ GNOMM/, prior MD simulations in GROMACS).[549]

Micelle Maker: The web-based Micelle Maker application provides a broad range of starting lipids and glycolipids readily compatible with AMBER and the GLYCAM library[550] (http:// www.micellemaker.net).

8.1.5. Polysaccharide Builders.

To build diverse saccharide 3D models online, one can use such tools as REStLESS and SWEET-II. doGlycans stand-alone framework can be used to prepare the atomistic models of glycopolymers, glycolipids and glycoproteins. Complex polysaccharide 3D models can be generated via POLYS[551,552] (http://www.models.life.ku.dk/polys,http://glycan-builder.cermav.cnrs.fr/) and CarbBuilder[553] (https://people.cs.uct.ac.za/~mkuttel/ Downloads.html). Another particular class of polysaccharide builders is dedicated to glycosaminoglycans (GAGs) which can be accessed using POLYS GAG-builder (http://matrixdb.univlyon1.fr/)[501] and GLYCAM-Web GAG-builder[506] (http://glycam.org/gag). Another approach for building GAG molecules was reported.[503] Unfortunately, the application scope of most of the existing structure building and modeling services is limited to a rigidly defined set of supported sugar residues and lacks non-carbohydrate moiety support.

8.1.6. Visualization.

The majority of molecular modeling software can achieve visualization of carbohydrates based on atomic coordinates. Nevertheless, several applications have been developed over the years to cope with the complexity of carbohydrate structures.[269,478,539,554–568]

The first attempt to display carbohydrate molecules in a meaningful and straightforward way was achieved by developing PaperChain and Twister graphics algorithm,[562] and these have been implemented in CarboHydra[569] and Visual Molecular Dynamics[269] packages. The most recent developments have emerged intending to convert the input of the 2D SNFG representation into a 3D-SNFG illustration.

SweetUnityMol:[478] https://sourceforge.net/projects/ unitymol/files/UnityMol_1.0.37/. With an extension: Umbrella visualization for N-glycan structures.[570]

Azahar plugin for PyMol[535] http://www.pymolwiki.org/ index.php/Azahar.

3D-SNFG VMD interface:[269] http://glycam.org/docs/othertoolsservice/2016/06/03/3d-symbol-nomenclature-forglycans-3d-snfg/. Tangram plugin: https://github.com/insilichem/tangram_ snfg.[566]

LiteMol (https://www.litemol.org/ and https://github.com/dsehnal/LiteMol) and its successor Mol* (https:// molstar.org/viewer/).[558]

SAMSON – OneAngstrom: Molecular simulation in the cloud. Provide a complete series of monosaccharides conformations for 3-dimensional constructions, optimization, visualization (https://www.oneangstrom.com).

8.1.7. Structural Data Analysis.

Conformational Analysis Tool (CAT): Analysis of carbohydrate molecular trajectory data derived from MD simulations, stand-alone software19 (http:// www.md-simulations.de/CAT/).

Best-fit, Four-Membered Plane (BFMP): Analysis of conformational data from crystal structures andMDsimulations of carbohydrates571 (http://glycam.org/docs/ othertoolsservice/download-docs/publication-materials/ bfmp/).

Distance Mapping: Estimation of nuclear Overhauser effects in disaccharides (http://www.glycosciences.de/modeling).

MD2NOE: Calculation of Nuclear Overhauser effect buildup curves from long MD trajectories572 (http://glycam.org/docs/othertoolsservice/download-docs/publication-materials/ md2noe/).

GS-align: Glycan structure alignment and similarity calculation, stand-alone software217 (http://www.glycanstructure. org/gsalign).

GlyTorsion and GlyVicinity analyze: Respectively, the torsion angles in carbohydrates and the amino acids in the vicinity of carbohydrates from Protein Data Bank573,574 (http:// www.glycosciences.de/tools/).

8.1.8. Tools for Structural Validation of Glycans.

CNS: Macromolecular structure determination and refinement (including carbohydrates and glycoproteins) based on X-ray and NMR data, stand-alone software575–578 (http://cns-online. org/v1.3/).

pdb-care: Identification and assigning carbohydrate structures using atom types and coordinates from PDB files, web tool579 (http://www.glycosciences.de/tools/pdb-care/).

CARP: Glycoprotein 3D quality evaluation based on the analysis of glycosidic torsion angles from PDB, web tool573 (http://www.glycosciences.de/tools/carp/).

GlyProbity: Accuracy and internal consistency check of carbohydrate 3D structures, part of a web service pipeline580 (https://dev.glycam.org/portal/gf_home/).

PDB2Glycan: 3D structure analysis and validation of glycoprotein PDB entries, part of a web service pipeline581 (https://glyconavi.org/TCarp/) (https://gitlab.com/ glyconavi/pdb2glycan).

PDB-REDO: Glycoprotein structure model improvement and validation, web service and stand-alone software582,583 (https://pdb-redo.eu**/).**

Coot: Refinement and validation of glycoprotein 3D structure from cryoEM and X-ray crystallography data, stand-alone software584,585 (https://www2.mrc-lmb.cam.ac.uk/personal/ pemsley/coot/).

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 framework586 https://www.rosettacommons.org/docs/latest/application_documentation/ carbohydrates/WorkingWithGlycans). Privateer: Automated validation of carbohydrate conformation data based on 3D structure analysis, stand-alone software587,588 (https://smb.slac.stanford.edu/facilities/ software/ccp4/html/privateer.html).

Privateer: The case for postpredictional modifications in the AlphaFold Protein Structure Database.589 According to the authors, the latest versions of Privateer can even graft glycans from other PDB structures onto AlphaFold models.

Phenix: Determination, refinement, and validation of macromolecular structure (including carbohydrates and glycoproteins) from cryoEM, X-ray diffraction and neutron diffraction crystallography data, stand-alone software541 (http://phenixonline. org/).

Motive Validator: Automatic custom residue validation in biomolecules, including carbohydrates, web service590 (https:// webchem.ncbr.muni.cz/Platform/MotiveValidator).

ValidatorDB: Precomputed validation results of ligands and nonstandard residues in PDB (including carbohydrates), web service591 (http://webchem.ncbr.muni.cz/Platform).

8.2. Structural Databases

In parallel with the development of methods in molecular modeling, there has been a revival in the number and scope of databases, websites, and both real and virtual glycan libraries that address the information needs in glycosciences. These efforts intend to (1) assess "primary data: (covalent and 3D structures of glycans and glycoconjugates) and (2) organize these primary data into databases, which can (a) speed up the production of primary data, (b) predict new features, and (c) characterize structure−activity or structure−function relationships, throughout their integration into meta-databases. Most existing repositories for glycan 3D structures are accessible via the web interface. They offer a variety of data covering oligosaccharides, polysaccharides, glycoproteins, and protein−carbohydrate complexes. These data are associated with experimentally resolved structures, NMR, X-ray and neutron crystallography,592 cryoEM, and small angle X-ray scattering. Some other databases contain data derived from molecular mechanics or molecular dynamics simulations. A complete list of glycan databases with 3D structure support, as existing in 2020, is given by Scherbinina and Toukach.4

8.2.1. 3D Structure Centric.

Carbohydrate Structure Database (CSDB):288,593,594 http://csdb.glycoscience.ru/ database.

Glycosciences.de:595−597 http://www.glycosciences.de/.

Glyco3D: A portal for structural glycosciences598,599 that includes a family of databases covering the 3D features of monosaccharides, disaccharides, oligosaccharides, polysaccharides, glycosyltransferases, lectins, monoclonal antibodies, and glycosaminoglycan binding proteins that have been developed with nonproprietary software and are freely available to the scientific community (http://glyco3d.cermav.cnrs.fr).

POLYSAC3DB:600 PolySac3DB is an annotated database that contains the 3D structural information and original fiber diffraction data of 157 polysaccharide entries that result from an extensive screening of scientific literature (http://www. polysac3db.cermav.cnrs.fr/).

EPS Database601 The EPS database provides access to detailed structural (1D−3D) taxonomic and bibliographic information on bacterial EPS (http://www.epsdatabase.com/).

EK3D: An E. coli K antigen 3-dimensional structure database602 (https://www.iith.ac.in/EK3D/).

3DSDSCAR: A three-dimensional structural database for sialic acid-containing carbohydrates through molecular dynamics simulations603,604.

MATRIX-DB: A biological database focused on molecular interactions between extracellular proteins and polysaccharides. It contains protein−protein interactions (PPIs) and also protein−glycosaminoglycan interactions and 3D structures of glycosaminoglycans605−607 (http://matrixdb.univ-lyon1.fr/).

GFDB: A glycan fragment database of PDB-based glycan 3D structures608 (http://www.glycanstructure.org).

8.2.2. Glycoproteomics.

UniLectin3D: The UniLectin platform is a dedicated portal of databases and tools to study the lectins472,609 (https://unilectin.eu/).

CBM 3DB: A curated database containing 3D structures of carbohydrate-binding modules of proteins that have been crystallized with their ligands610 (https://cbmdb.glycopedia. eu/).

GAG 3DB: A curated database that classifies the threedimensional features of the six mammalian GAGs (chondroitin sulfate, dermatan sulfate, heparin, heparan sulfate, hyaluronan, and keratan sulfate) and their oligosaccharides complexed with proteins505 (https://gagdb.glycopedia.eu/).

ProCarbDB: A comprehensive database containing over 5200 3D X-ray crystal structures of protein−carbohydrate complexes611 (http://www.procarbdb.science/procarb/).

GBSDB GLYCAN: GBS-predictor; GS-align. GlyMDB. Statistics; How To Use; Reference; Search Glycan Microarray DB.612 (http://www.glycanstructure.org/gbs-db/pdb/)

ProCaff: Protein−carbohydrate complex binding affinity database.613

StackCBPred: A stacking-based prediction of protein-carbohydrate binding sites from a sequence.614

SAbDab: Database of antibody structures annotated with several properties (experimental data, antibody nomenclature, curated affinity, and sequence annotations)615 (http://opig. stats.ox.ac.uk/webapps/newsabdab/sabdab/).

9. CONCLUSIONS

In the last 15 years, there has been a significant increase in the development and applications of computational methods. These methods established many structural and dynamic features of complex carbohydrates, either in isolation or in complex with other biomolecules. Quantum chemical methods provided the theoretical basis for understanding the stereoelectronic effects that characterize carbohydrates. The developments and implementations of force fields integrating these effects in several “generic” software packages provide many users with a comprehensive way to carry on computational explorations of their endeavors in association with their experiments. The use of quantum chemical methods to rationalize the reactivity principle in carbohydrate chemical synthesis is a remarkable example of such an achievement. Molecular simulation techniques have reached the sampling power and sophistication to provide the missing structural insight necessary to interpret or support experiments and be a primary scientific discovery tool. Complementing these developments, advances in high-performance computing have enabled molecular simulation methods to play a more significant role in supporting experiments and transcending this mandate to guide experimental design and make scientific discoveries independently. Untouched spatial and temporal dimensions are accessible in a reasonable computational time. As such, atomistic MD simulations provide a unique insight, accurately describing the actual 3D structure and dynamic patterns at the real-time scale where molecular events occur. Not only can structure−function relationships be proposed in some cases, but the characterization of physicochemical and mechanical properties paves the way for establishing new structure−property relationships.

The development of coarse-grained simulations allows the application of these techniques to most glycoscience systems. The applications are manifold and concern polysaccharides (mammalian, bacterial, microbial, plant and marine polysaccharides, lipopolysaccharides in membranes), N-linked and oxygen-linked glycans, and glycolipids. As the coarse-graining procedure embraces a wide range of length scales, it is expected that the modeling of complex carbohydrate-embedded materials will develop, an essential step in the quest for a biomass-based economy.

Due to the reduced complexity inherent in the model, coarse-grained computer simulations can offer an alternative and bridge the gap between experiment and all-atom calculations for systems of high complexity. In particular, they are well suited to investigate the dynamic glycolipid nano-aggregation, in which carbohydrate−carbohydrate interactions play a crucial role. The question remains whether such methods help decipher the multivalent effect that drives protein-carbohydrate assembly.

Indeed, as most carbohydrate-binding proteins, especially adhesins and lectins, have a relatively low affinity and generally narrow carbohydrate recognition domains, their intrinsic specificities often lie in their valence and various topologies. The simultaneous presentation of several proper and identical glycoside units converts relatively weak interactions into specific recognition effects. Therefore, one needs to consider some physicochemical principles that underline such associations as patches of glycolipids and glyco-surfaces that would define a “glyco landscape: or “glycotope”. If this concept develops, computational tools should consider the glyco-surface generated by the side-by-side arrangement of glycolipids and consider the interaction with the glyco-surface or glyco-canopy (by analogy with the crown-canopy as the uppermost layer formed by the crown of trees in a forest).

This concept applies to research dealing with the solid-state degradation of crystalline or semicrystalline polysaccharides by enzymes. The computational exploration of these systems is far from complete, despite the fundamental contribution that computer simulations have made to understanding the chemical mechanisms underlying the actions of glycosyl hydrolases. They have allowed the identification of catalytic residues, the discovery of complex conformational pathways, and the identification of mechanistic details that escape experimental probes. This field of research takes advantage of the availability of an extensive collection of crystal structures of proteins and their carbohydrate complexes.

Structural data collected by X-ray crystallography offers the possibility of organizing well-targeted databases using PDB information, with appropriate treatment of the topology of the glycans. Nevertheless, there are many documented cases where the high degree of structural disorder that characterizes unbound glycans makes them inferior targets for experimental structural biology studies. Meanwhile, molecular simulations are a powerful complementary tool to standard structural biology techniques that can provide valuable high-resolution structural and energetic information on glycans. However, despite its extensive use of computational resources, molecular modeling does not have a place where documented results could be deposited and made available to the public. Therefore, it is tempting to envisage creating and organizing such a 3D data repository. The stored data would correspond to the most populated conformers derived from the simulation with

annotated details on energetics and relative populations. The constitution of such a structural database would allow data reproduction, a presently absent feature. It would give users the unique ability to monitor the actual volume of space-specific glycosylation patterns on a membrane or a protein surface, allowing or preventing recognition from other receptors. Its integration into meta-databases would enrich the computational and experimental data available for training machine learning endeavors to enable rapid progress and contribute to deciphering the many complex processes involving these complex biomolecules.

AUTHOR INFORMATION

Corresponding Author

Serge Perez − Centre de Recherche sur les Macromolecules Vegetales, University of Grenoble-Alpes, Centre National de la Recherche Scientifique, Grenoble F-38041, France; orcid.org/0000-0003-3464-5352; Email: serge.perez@ cermav.cnrs.fr, spsergeperez@gmail.com

Author

Olga Makshakova − FRC Kazan Scientific Center of Russian Academy of Sciences, Kazan Institute of Biochemistry and Biophysics, Kazan 420111, Russia Complete contact information is available at: https://pubs.acs.org/10.1021/acs.chemrev.2c00060

Notes The authors declare no competing financial interest.

Biographies

Serge Perez is a Research Director within the Centre National de la Recherche Scientifique at the Centre de Recherches sur les Macromolecules Végétales, to which he has held the position of chairman for 11 years. Among other residences in different scientific institutions, he was Scientific Director at the European Synchrotron Radiation Facility, in Grenoble, in charge of the Life Sciences, Chemistry and Imaging programs. He has been the President of the European Carbohydrate Organization and the chairman and/or partner in several national, European, and international projects dealing with glycoscience's structural and functional aspects. He is the author and co-author of more than 320 scientific articles and reviews. He has a keen interest in biotech startup companies and the bioeconomy of glycoscience, with the publication of two books. He is the founder of the Internet site glycopedia.eu, an initiative aimed at promoting the field of glycosciences and providing open-access material for educational purposes

Olga Makshakova is a leading researcher at the Kazan Institute of Biochemistry and Biophysics of the Federal Research Center "Kazan Scientific Center of the Russian Academy of Sciences". After graduating in physics from Kazan State University, she got her PhD in Biophysics in 2010 with a thesis on the joint application of FTIR spectroscopy and quantum chemical (DFT) calculations to biopolymers and hydration water. Further, she applied both experimental and theoretical physics methods to study biological macromolecules' spatial structure and interactions. Since 2015, due to the collaborative projects with Centre National de la Recherche Scientifique, she has been involved in glycoscience with a particular focus on polysaccharide structure and glycoconjugate assembly on the lipid membrane interface. She is the author of 36 peer-reviewed articles. She contributed to 18 as a principal coauthor (first or last corresponding author). Currently, her work focuses on bio-macromolecular assembly and structure-induced fit under protein−carbohydrate interactions for innovative bioinspired materials.

ACKNOWLEDGMENTS

The following funding supported this work: The Cross-Disciplinary Program Glyco@Alps within the framework "Investissement d'Avenir" program [ANR-15IDEX-02]. According to research project no, the study was partly funded by RFBR & CNRS. 21-54-15008. S.P. and O.M. acknowledge the support of the CNRS and the government assignment for FRC Kazan Scientific Center of the Russian Academy of Sciences, respectively.

REFERENCES

10 REFERENCES

(1) Fadda, E.; Woods, R. J. Molecular Simulations of Carbohydrates and Protein-Carbohydrate Interactions: Motivation, Issues and Prospects. *Drug Discov. Today* **2010**, *15*, 596-609.

(2) Perez, S.; Tvaroska, I. Carbohydrate-Protein Interactions: Molecular Modeling Insights. Adv. Carbohydr. Chem. Biochem. 2014, 71, 9-136.

(3) 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.

(4) Scherbinina, S. I.; Toukach, P. V. Three-Dimensional Structures of Carbohydrates and Where to Find Them. Int. J. Mol. Sci. 2020, 21, 7702.

(5) Perez, S.; Fadda, E.; Makshakova, O. Computational Modeling in Glycoscience. In Comprehensive Glycoscience (Second Edition), Barchi, J. J. Ed.; Elsevier, 2021; pp 374-404.

(6) Almond, A. Multiscale Modeling of Glycosaminoglycan Structure and Dynamics: Current Methods and Challenges. Curr. Opin. Struct. Biol. 2018, 50, 58-64.

(7) M. Kuttel, M. The Conformational Free Energy of Carbohydrates. Mini-Reviews in Organic Chemistry 2011, 8, 256-262.

(8) Khalid, S.; Piggot, T. J.; Samsudin, F. Atomistic and Coarse Grain Simulations of the Cell Envelope of Gram-Negative Bacteria: What Have We Learned? Acc. Chem. Res. 2019, 52, 180-188.

(9) 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.

(10) Perez, S. The Symbolic Representation of Monosaccharides in the Age of Glycobiology. In Glycopedia; 2014.

(11) McNaught, A. D. International Union of Pure and Applied Chemistry and International Union of Biochemistry and Molecular Biology. Joint Commission on Biochemical Nomenclature. Nomenclature of Carbohydrates. Carbohydr. Res. 1997, 297, 1-92.

(12) McNaught, A. D. Nomenclature of Carbohydrates (Recommendations 1996). Adv. Carbohydr. Chem. Biochem. 1997, 52, 43-177.

(13) Chemdraw (Rrid:Scr_016768); http://www.perkinelmer.co.uk/category/chemdraw (accessed 2022-01-15).

(14) The Pymol Molecular Graphics System, Version 2.0 Schrödinger, Llc (accessed 2022-01-15).

(15) Kirschner, K. N.; Woods, R. J. Solvent Interactions Determine Carbohydrate Conformation. Proc. Natl. Acad. Sci. USA 2001, 98, 10541-10545.

(16) Agirre, J. Strategies for Carbohydrate Model Building, Refinement and Validation. Acta Crystallogr. D 2017, 73, 171-186.

(17) Marchessault, R. H.; Perez, S. Conformations of the Hydroxymethyl Group in Crystalline Aldohexopyranoses. Biopolymers 1979, 18, 2369-2374.

(18) Koča, J.; Pérez, S.; Imberty, A. Conformational Analysis and Flexibility of Carbohydrates Using the Cicada Approach with Mm3. J. Comput. Chem. 1995, 16, 296-310.

(19) Frank, M. Conformational Analysis of Oligosaccharides and Polysaccharides Using Molecular Dynamics Simulations. Methods Mol. Biol. 2015, 1273, 359-377.

(20) Braun, E.; Gilmer, J.; Mayes, H. B.; Mobley, D. L.; Monroe, J. I.; Prasad, S.; Zuckerman, D. M. Best Practices for Foundations in Molecular Simulations [Article V1.0]. Living J. Comput. Mol. Sci. 2019, 1, 5957.

(21) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. Numerical Integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of N-Alkanes. J. Comput. Phys. 1977, 23, 327-341.

(22) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M. Lincs: A Linear Constraint Solver for Molecular Simulations. J. Comput. Chem. 1997, 18, 1463-1472.

(23) Sugita, Y.; Okamoto, Y. Replica-Exchange Molecular Dynamics Method for Protein Folding. Chem. Phys. Lett. 1999, 314, 141-151.

(24) Wang, L.; Friesner, R. A.; Berne, B. J. Replica Exchange with Solute Scaling: A More Efficient Version of Replica Exchange with Solute Tempering (Rest2). J Phys Chem B 2011, 115, 9431-9438.

(25) Burusco, K. K.; Bruce, N. J.; Alibay, I.; Bryce, R. A. Free Energy Calculations Using a Swarm-Enhanced Sampling Molecular Dynamics Approach. ChemPhysChem 2015, 16, 3233-3241.

(26) Miao, Y.; Feher, V. A.; McCammon, J. A. Gaussian Accelerated Molecular Dynamics: Unconstrained Enhanced Sampling and Free Energy Calculation. J. Chem. Theory Comput. 2015, 11, 3584-3595.

(27) 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.

(28) 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.

(29) Foley, B. L.; Tessier, M. B.; Woods, R. J. Carbohydrate Force Fields. Wiley Interdiscip. Rev. Comput. Mol. Sci. 2012, 2, 652-697.

(30) Xiong, X.; Chen, Z.; Cossins, B. P.; Xu, Z.; Shao, Q.; Ding, K.; Zhu, W.; Shi, J. Force Fields and Scoring Functions for Carbohydrate Simulation. Carbohydr. Res. 2015, 401, 73-81.

(31) 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.

(32) 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.

(33) 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.

(34) Stortz, C. A. Mm3 Potential Energy Surfaces of Trisaccharide Models of Lambda-, Mu-, and Nu-Carrageenans. Carbohydr. Res. 2006, 341, 2531-2542.

(35) Taha, H. A.; Richards, M. R.; Lowary, T. L. Conformational Analysis of Furanoside-Containing Mono- and Oligosaccharides. Chem. Rev. 2013, 113, 1851-1876.

(36) 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.

(37) 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.

(38) 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.

(39) 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.

(40) 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.

(41) Kirschner, K. N.; Yongye, A. B.; Tschampel, S. M.; Gonzalez-Outeirino, J.; Daniels, C. R.; Foley, B. L.; Woods, R. J. Glycam06: A Generalizable Biomolecular Force Field. Carbohydrates. J. Comput. Chem. 2008, 29, 622-655.

(42) 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.

(43) DeMarco, M. L. Molecular Dynamics Simulations of Membrane- and Protein-Bound Glycolipids Using Glycam. Methods Mol. Biol. 2015, 1273, 379-390.

(44) 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.

(45) 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, 349-363.

(46) 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.

(47) Lins, R. D.; Hünenberger, P. H. A New Gromos Force Field for Hexopyranose-Based Carbohydrates. J. Comput. Chem. 2005, 26, 1400-1412.

(48) 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.

(49) Plazinski, W.; Lonardi, A.; Hünenberger, P. H. Revision of the Gromos 56a6(Carbo) Force Field: Improving the Description of Ring-Conformational Equilibria in Hexopyranose-Based Carbohydrates Chains. J. Comput. Chem. 2016, 37, 354-365.

(50) Naumov, V. S.; Ignatov, S. K. Modification of 56a(Carbo) Force Field for Molecular Dynamic Calculations of Chitosan and Its Derivatives. J. Mol. Model. 2017, 23, 244.

(51) Nester, K.; Gaweda, K.; Plazinski, W. A Gromos Force Field for Furanose-Based Carbohydrates. J. Chem. Theory Comput. 2019, 15, 1168-1186.

(52) 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.

(53) 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.

(54) Damm, W.; Frontera, A.; Tirado–Rives, J.; Jorgensen, W. L. Opls All-Atom Force Field for Carbohydrates. J. Comput. Chem. 1997, 18, 1955-1970.

(55) 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.

(56) Jamali, S. H.; Westen, T. v.; Moultos, 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.

(57) 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.

(58) 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.

(59) Jana, M.; MacKerell, A. D. Charmm Drude Polarizable Force Field for Aldopentofuranoses and Methyl-Aldopentofuranosides. J. Phys. Chem. B 2015, 119, 7846-7859.

(60) He, X.; Lopes, P. E.; Mackerell, A. D., Jr. Polarizable Empirical Force Field for Acyclic Polyalcohols Based on the Classical Drude Oscillator. Biopolymers 2013, 99, 724-738.

(61) Aytenfisu, A. H.; Yang, M. J.; MacKerell, A. D. Charmm Drude Polarizable Force Field for Glycosidic Linkages Involving Pyranoses and Furanoses. J. Chem. Theory Comput. 2018, 14, 3132-3143.

(62) Kognole, A. A.; Lee, J.; Park, S. J.; Jo, S.; Chatterjee, P.; Lemkul, J. A.; Huang, J.; MacKerell, A. D., Jr.; Im, W. Charmm-Gui Drude Prepper for Molecular Dynamics Simulation Using the Classical Drude Polarizable Force Field. J. Comput. Chem. 2022, 43, 359-375.

(63) Levitt, M.; Warshel, A. Computer Simulation of Protein Folding. Nature 1975, 253, 694-698.

(64) Molinero, V.; Goddard, W. A. M3b: A Coarse Grain Force Field for Molecular Simulations of Malto-Oligosaccharides and Their Water Mixtures. J. Phys. Chem. B 2004, 108, 1414-1427.

(65) Gu, J.; Bai, F.; Li, H.; Wang, X. A Generic Force Field for Protein Coarse-Grained Molecular Dynamics Simulation. Int. J. Mol. Sci. 2012, 13, 14451-14469.

(66) Hadley, K. R.; McCabe, C. Coarse-Grained Molecular Models of Water: A Review. Mol. Simul. 2012, 38, 671-681.

(67) Marrink, S. J.; Tieleman, D. P. Perspective on the Martini Model. Chem. Soc. Rev. 2013, 42, 6801-6822.

(68) Lyubartsev, A. P.; Laaksonen, A. Calculation of Effective Interaction Potentials from Radial Distribution Functions: A Reverse Monte Carlo Approach. Phys. Rev. E Stat. Phys. Plasmas Fluids Relat. Interdiscip. Topics 1995, 52, 3730-3737.

(69) Lyubartsev, A. P.; Naômé, A.; Vercauteren, D. P.; Laaksonen, A. Systematic Hierarchical Coarse-Graining with the Inverse Monte Carlo Method. J. Chem. Phys. 2015, 143, 243120.

(70) Shell, M. S. Coarse-Graining with the Relative Entropy. In Advances in Chemical Physics, 2016; pp 395-441.

(71) Izvekov, S.; Voth, G. A. A Multiscale Coarse-Graining Method for Biomolecular Systems. J. Phys. Chem. B 2005, 109, 2469-2473.

(72) Izvekov, S.; Voth, G. A. Multiscale Coarse Graining of Liquid-State Systems. J. Chem. Phys. 2005, 123, 134105.

(73) Moritsugu, K.; Smith, J. C. Coarse-Grained Biomolecular Simulation with Reach: Realistic Extension Algorithm Via Covariance Hessian. Biophys. J. 2007, 93, 3460-3469.

(74) Scherer, C.; Andrienko, D. Understanding Three-Body Contributions to Coarse-Grained Force Fields. Phys. Chem. Chem. Phys. 2018, 20, 22387-22394.

(75) Liu, P.; Izvekov, S.; Voth, G. A. Multiscale Coarse-Graining of Monosaccharides. J. Phys. Chem. B 2007, 111, 11566-11575.

(76) Ingolfsson, H. I.; Lopez, C. A.; Uusitalo, J. J.; de Jong, D. H.; Gopal, S. M.; Periole, X.; Marrink, S. J. The Power of Coarse Graining in Biomolecular Simulations. Wiley Interdiscip. Rev. Comput. Mol. Sci. 2014, 4, 225-248.

(77) Kar, P.; Feig, M. Recent Advances in Transferable Coarse-Grained Modeling of Proteins. Adv. Protein Chem. Struct. Biol. 2014, 96, 143-180.

(78) Kleinjung, J.; Fraternali, F. Design and Application of Implicit Solvent Models in Biomolecular Simulations. Curr. Opin. Struct. Biol. 2014, 25, 126-134.

(79) Marrink, S. J.; Corradi, V.; Souza, P. C. T.; Ingólfsson, H. I.; Tieleman, D. P.; Sansom, M. S. P. Computational Modeling of Realistic Cell Membranes. Chem. Rev. 2019, 119, 6184-6226.

(80) Lopez, C. A.; Rzepiela, A. J.; de Vries, A. H.; Dijkhuizen, L.; Hunenberger, P. H.; Marrink, S. J. Martini Coarse-Grained Force Field: Extension to Carbohydrates. J. Chem. Theory Comput. 2009, 5, 3195-3210.

(81) Wassenaar, T. A.; Ingolfsson, H. I.; Priess, M.; Marrink, S. J.; Schafer, L. V. Mixing Martini: Electrostatic Coupling in Hybrid Atomistic-Coarse-Grained Biomolecular Simulations. J. Phys. Chem. B 2013, 117, 3516-3530.

(82) Marrink, S. J.; Risselada, H. J.; Yefimov, S.; Tieleman, D. P.; de Vries, A. H. The Martini Force Field: Coarse Grained Model for Biomolecular Simulations. J. Phys. Chem. B 2007, 111, 7812-7824.

(83) Monticelli, L.; Kandasamy, S. K.; Periole, X.; Larson, R. G.; Tieleman, D. P.; Marrink, S.-J. The Martini Coarse-Grained Force Field: Extension to Proteins. J. Chem. Theory Comput. 2008, 4, 819-834.

(84) Uusitalo, J. J.; Ingólfsson, H. I.; Akhshi, P.; Tieleman, D. P.; Marrink, S. J. Martini Coarse-Grained Force Field: Extension to DNA. J. Chem. Theory Comput. 2015, 11, 3932-3945.

(85) 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.

(86) Souza, P. C. T.; Alessandri, R.; Barnoud, J.; Thallmair, S.; Faustino, I.; Grünewald, F.; Patmanidis, I.; Abdizadeh, H.; Bruininks, B. M. H.; Wassenaar, T. A.; et al. Martini 3: A General Purpose Force Field for Coarse-Grained Molecular Dynamics. Nat. Methods 2021, 18, 382-388.

(87) Ma, H.; Irudayanathan, F. J.; Jiang, W.; Nangia, S. Simulating Gram-Negative Bacterial Outer Membrane: A Coarse Grain Model. J. Phys. Chem. B 2015, 119, 14668-14682.

(88) Johnson, M. E.; Head-Gordon, T.; Louis, A. A. Representability Problems for Coarse-Grained Water Potentials. J. Chem. Phys. 2007, 126, 144509.

(89) Qi, Y.; Ingólfsson, H. I.; Cheng, X.; Lee, J.; Marrink, S. J.; Im, W. Charmm-Gui Martini Maker for Coarse-Grained Simulations with the Martini Force Field. J. Chem. Theory Comput. 2015, 11, 4486-4494.

(90) Koča, J. Computer Program Cicada — Travelling Along Conformational Potential Energy Hypersurface. J. Mol. Struct. THEOCHEM 1994, 308, 13-24.

(91) Koca, J. Travelling through Conformational Space: An Approach for Analyzing the Conformational Behaviour of Flexible Molecules. Prog. Biophys. Mol. Biol. 1998, 70, 137-173.

(92) Engelsen, S. B.; Koca, J.; Braccini, I.; Hervé du Penhoat, C.; Pérez, S. Travelling on the Potential Energy Surfaces of Carbohydrates: Comparative Application of an Exhaustive Systematic Conformational Search with an Heuristic Search. Carbohydr. Res. 1995, 276, 1-29.

(93) Weimar, T.; Meyer, B.; Peters, T. Conformational Analysis of Alpha-D-Fuc-(1-->4)-Beta-D-Glcnac-Ome. One-Dimensional Transient Noe Experiments and Metropolis Monte Carlo Simulations. J. Biomol. NMR 1993, 3, 399-414.

(94) Peters, T.; Meyer, B.; Stuike-Prill, R.; Somorjai, R.; Brisson, J. R. A Monte Carlo Method for Conformational Analysis of Saccharides. Carbohydr. Res. 1993, 238, 49-73.

(95) Mitchell, T. M. Machine Learning; McGraw-Hill Education, 1997.

(96) Lamarck, J. B. Philosophie Zoologique; Dentu, 1809.

(97) Nahmany, A.; Strino, F.; Rosen, J.; Kemp, G. J.; 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.

(98) Rosen, J.; Miguet, L.; Perez, S. Shape: Automatic Conformation Prediction of Carbohydrates Using a Genetic Algorithm. J. Cheminform. 2009, 1, 16.

(99) Rosen, J.; Robobi, A.; Nyholm, P. G. Conformation of the Branched O-Specific Polysaccharide of Shigella Dysenteriae Type 2: Molecular Mechanics Calculations Show a Compact Helical Structure Exposing an Epitope Which Potentially Mimics Galabiose. Carbohydr. Res. 2002, 337, 1633-1640.

(100) Rosen, J.; Robobi, A.; Nyholm, P. G. The Conformations of the O-Specific Polysaccharides of Shigella Dysenteriae Type 4 and Escherichia Coli O159 Studied with Molecular Mechanics (Mm3) Filtered Systematic Search. Carbohydr. Res. 2004, 339, 961-966.

(101) Strino, F.; Nahmany, A.; Rosen, J.; Kemp, G. J.; Sa-correia, I.; Nyholm, P. G. Conformation of the Exopolysaccharide of Burkholderia Cepacia Predicted with Molecular Mechanics (Mm3) Using Genetic Algorithm Search. Carbohydr. Res. 2005, 340, 1019-1024.

(102) Hansen, H. S.; Hünenberger, P. H. Using the Local Elevation Method to Construct Optimized Umbrella Sampling Potentials: Calculation of the Relative Free Energies and Interconversion Barriers of Glucopyranose Ring Conformers in Water. J. Comput. Chem. 2010, 31, 1-23.

(103) Gaweda, K.; Plazinski, W. The Systematic Influence of Solvent on the Conformational Features of Furanosides. Org. Biomol. Chem. 2019, 17, 2479-2485.

(104) Plazinska, A.; Plazinski, W. Comparison of Carbohydrate Force Fields in Molecular Dynamics Simulations of Protein-Carbohydrate Complexes. J. Chem. Theory Comput. 2021, 17, 2575-2585.

(105) Plazinski, W.; Drach, M. The Dynamics of the Conformational Changes in the Hexopyranose Ring: A Transition Path Sampling Approach. RSC Advances 2014, 4, 25028-25039.

(106) Balogh, G.; Gyöngyösi, T.; Timári, I.; Herczeg, M.; Borbás, A.; Sadiq, S. K.; Fehér, K.; Kövér, K. E. Conformational Analysis of Heparin-Analogue Pentasaccharides by Nuclear Magnetic Resonance Spectroscopy and Molecular Dynamics Simulations. J. Chem. Inf. Model. 2021, 61, 2926-2936.

(107) 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.

(108) Kadirvelraj, R.; Foley, B. L.; Dyekjaer, J. D.; Woods, R. J. Involvement of Water in Carbohydrate-Protein Binding: Concanavalin a Revisited. J. Am. Chem. Soc. 2008, 130, 16933-16942.

(109) Adam, J.; Kríz, Z.; Prokop, M. P.; Wimmerová, M.; Koca, J. In Silico Mutagenesis and Docking Studies of Pseudomonas Aeruginosa Pa-Iil Lectin Predicting Binding Modes and Energies. J. Chem. Inf. Model. 2008, 48, 2234-2242.

(110) Mishra, S. K.; Sund, J.; Aqvist, J.; Koca, J. Computational Prediction of Monosaccharide Binding Free Energies to Lectins with Linear Interaction Energy Models. J. Comput. Chem. 2012, 33, 2340-2350.

(111) Mishra, S. K.; Calabro, G.; Loeffler, H. H.; Michel, J.; Koca, 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.

(112) Topin, J.; Arnaud, J.; Sarkar, A.; Audfray, A.; Gillon, E.; Perez, S.; Jamet, H.; Varrot, A.; Imberty, A.; Thomas, A. Deciphering the Glycan Preference of Bacterial Lectins by Glycan Array and Molecular Docking with Validation by Microcalorimetry and Crystallography. PLoS One 2013, 8, e71149.

(113) Koppisetty, C. A. K.; Frank, M.; Lyubartsev, A. P.; Nyholm, P. G. Binding Energy Calculations for Hevein-Carbohydrate Interactions Using Expanded Ensemble Molecular Dynamics Simulations. J. Comput. Aided Mol. Des. 2015, 29, 13-21.

(114) Liu, W.; Jia, X. Y.; Wang, M. T.; Li, P. F.; Wang, X. H.; Hu, W. X.; Zheng, J.; Mei, Y. Calculations of the Absolute Binding Free Energies for Ralstonia Solanacearum Lectins Bound with Methyl-Alpha-L-Fucoside at Molecular Mechanical and Quantum Mechanical/Molecular Mechanical Levels. RSC Advances 2017, 7, 38570-38580.

(115) 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.

(116) Plazinski, W.; Plazinska, A. Molecular Dynamics Simulations of Hexopyranose Ring Distortion in Different Force Fields. Pure Appl. Chem. 2017, 89, 1283-1294.

(117) 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.

(118) 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, 2013; pp 469-483.

(119) Spiwok, V.; Lipovová, P.; Skálová, T.; Vondráčková, E.; Dohnálek, J.; Hašek, J.; Králová, B. Modelling of Carbohydrate–Aromatic Interactions: Ab Initio Energeticsand Force Field Performance. J. Comput. Aided Mol. Des. 2005, 19, 887-901.

(120) 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.

(121) 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.

(122) Sarkar, A. S.; Perez, S. Protein Carbohydrate Interactions: Computational Aspects. In Structural Glycobiology, 1 ed.; Yuriev, E., Ramsland, P. A. Eds.; CRC Press, Boca Raton, 2012; pp 71-110.

(123) 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.

(124) Kerzmann, A.; Neumann, D.; Kohlbacher, O. Slick-Scoring and Energy Functions for Protein-Carbohydrate Interactions. J. Chem. Inf. Model. 2006, 46, 1635-1642.

(125) Boittier, E. D.; Burns, J. M.; Gandhi, N. S.; Ferro, V. Glycotorch Vina: Docking Designed and Tested for Glycosaminoglycans. J. Chem. Inf. Model. 2020, 60, 6328-6343.

(126) van Zundert, G. C. P.; Rodrigues, J.; Trellet, M.; Schmitz, C.; Kastritis, P. L.; Karaca, E.; Melquiond, A. S. J.; van Dijk, M.; de Vries, S. J.; Bonvin, A. The Haddock2.2 Web Server: User-Friendly Integrative Modeling of Biomolecular Complexes. J. Mol. Biol. 2016, 428, 720-725.

(127) 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.

(128) 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.

(129) Nivedha, A. K.; Thieker, D. F.; Makeneni, S.; Hu, H.; Woods, R. J. Vina-Carb: Improving Glycosidic Angles During Carbohydrate Docking. J. Chem. Theory Comput. 2016, 12, 892-901.

(130) Nance, M. L.; Labonte, J. W.; Adolf-Bryfogle, J.; Gray, J. J. Development and Evaluation of Glycandock: A Protein–Glycoligand Docking Refinement Algorithm in Rosetta. J. Phys. Chem. B 2021, 125, 6807-6820.

(131) 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.

(132) 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.

(133) Ewing, T. J. A.; Makino, S.; Skillman, A. G.; Kuntz, I. D. Dock 4.0: Search Strategies for Automated Molecular Docking of Flexible Molecule Databases. J. Comput. Aided Mol. Des. 2001, 15, 411-428.

(134) Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T. A.; Klicic, J. J.; Mainz, D. T.; Repasky, M. P.; Knoll, E. H.; Shelley, M.; Perry, J. K.; et al. Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy. J. Med. Chem. 2004, 47, 1739-1749.

(135) Sigurskjold, B. W.; Bundle, D. R. Thermodynamics of Oligosaccharide Binding to a Monoclonal Antibody Specific for a Salmonella O-Antigen Point to Hydrophobic Interactions in the Binding Site. J. Biol. Chem. 1992, 267, 8371-8376.

(136) Shimokhina, N.; Bronowska, A.; Homans, S. W. Contribution of Ligand Desolvation to Binding Thermodynamics in a Ligand-Protein Interaction. Angew. Chem. Int. Ed. Engl. 2006, 45, 6374-6376.

(137) Williams, B. A.; Chervenak, M. C.; Toone, E. J. Energetics of Lectin-Carbohydrate Binding. A Microcalorimetric Investigation of Concanavalin a-Oligomannoside Complexation. J. Biol. Chem. 1992, 267, 22907-22911.

(138) Boraston, A. B.; Bolam, D. N.; Gilbert, H. J.; Davies, G. J. Carbohydrate-Binding Modules: Fine-Tuning Polysaccharide Recognition. Biochem. J. 2004, 382, 769-781.

(139) Quiocho, F. A. Carbohydrate-Binding Proteins: Tertiary Structures and Protein-Sugar Interactions. Annu. Rev. Biochem. 1986, 55, 287-315.

(140) 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.

(141) Nishio, M.; Umezawa, Y.; Fantini, J.; Weiss, M. S.; Chakrabarti, P. Ch-Π Hydrogen Bonds in Biological Macromolecules. Phys. Chem. Chem. Phys. 2014, 16, 12648-12683.

(142) Fernández-Alonso, M. d. C.; Cañada, F. J.; Jiménez-Barbero, J.; Cuevas, G. Molecular Recognition of Saccharides by Proteins. Insights on the Origin of the Carbohydrate−Aromatic Interactions. J. Am. Chem. Soc. 2005, 127, 7379-7386.

(143) Arcon, J. P.; Turjanski, A. G.; Martí, M. A.; Forli, S. Biased Docking for Protein-Ligand Pose Prediction. Methods Mol. Biol. 2021, 2266, 39-72.

(144) Gauto, D. F.; Petruk, A. A.; Modenutti, C. P.; Blanco, J. I.; Di Lella, S.; Marti, M. A. Solvent Structure Improves Docking Prediction in Lectin-Carbohydrate Complexes. Glycobiology 2013, 23, 241-258.

(145) López, E. D.; Arcon, J. P.; Gauto, D. F.; Petruk, A. A.; Modenutti, C. P.; Dumas, V. G.; Marti, M. A.; Turjanski, A. G. Watclust: A Tool for Improving the Design of Drugs Based on Protein-Water Interactions. Bioinformatics 2015, 31, 3697-3699.

(146) Modenutti, C. P.; Capurro, J. I. B.; Di Lella, S.; Martí, M. A. The Structural Biology of Galectin-Ligand Recognition: Current Advances in Modeling Tools, Protein Engineering, and Inhibitor Design. Front. Chem. 2019, 7, 823.

(147) Arcon, J. P.; Defelipe, L. A.; Modenutti, C. P.; López, E. D.; Alvarez-Garcia, D.; Barril, X.; Turjanski, A. G.; Martí, M. A. Molecular Dynamics in Mixed Solvents Reveals Protein–Ligand Interactions, Improves Docking, and Allows Accurate Binding Free Energy Predictions. J. Chem. Inf. Model. 2017, 57, 846-863.

(148) Arcon, J. P.; Modenutti, C. P.; Avendaño, D.; Lopez, E. D.; Defelipe, L. A.; Ambrosio, F. A.; Turjanski, A. G.; Forli, S.; Marti, M. A. Autodock Bias: Improving Binding Mode Prediction and Virtual Screening Using Known Protein-Ligand Interactions. Bioinformatics 2019, 35, 3836-3838.

(149) Dingjan, T.; Gillon, É.; Imberty, A.; Pérez, S.; Titz, A.; Ramsland, P. A.; Yuriev, E. Virtual Screening against Carbohydrate-Binding Proteins: Evaluation and Application to Bacterial Burkholderia Ambifaria Lectin. J. Chem. Inf. Model. 2018, 58, 1976-1989.

(150) 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.

(151) 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.

(152) Gehrcke, J.-P.; Pisabarro, M. T. Identification and Characterization of a Glycosaminoglycan Binding Site on Interleukin-10 Via Molecular Simulation Methods. J. Mol. Graphics Model. 2015, 62, 97-104.

(153) Samsonov, S. A.; Pisabarro, M. T. Computational Analysis of Interactions in Structurally Available Protein-Glycosaminoglycan Complexes. Glycobiology 2016, 26, 850-861.

(154) Forgione, R. E.; Di Carluccio, C.; Kubota, M.; Manabe, Y.; Fukase, K.; Molinaro, A.; Hashiguchi, T.; Marchetti, R.; Silipo, A. Structural Basis for Glycan-Receptor Binding by Mumps Virus Hemagglutinin-Neuraminidase. Sci. Rep. 2020, 10, 1589.

(155) 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.

(156) Joseph, P. R. B.; Mosier, P. D.; Desai, U. R.; Rajarathnam, K. Solution NMR Characterization of Chemokine CXCL8/Ll-8 Monomer and Dimer Binding to Glycosaminoglycans: Structural Plasticity Mediates Differential Binding Interactions. Biochem. J. 2015, 472, 121-133.

(157) Makshakova, O. N.; Safarova, E. R.; Zuev, Y. F. Structural Insights in Interactions between Rnase from Bacillus Intermedius and Rhamnogalacturonan I from Potato. Carbohydr. Polym. 2021, 251, 117038.

(158) Forgione, R. E.; Di Carluccio, C.; Guzmán-Caldentey, J.; Gaglione, R.; Battista, F.; Chiodo, F.; Manabe, Y.; Arciello, A.; Del Vecchio, P.; Fukase, K.; et al. Unveiling Molecular Recognition of Sialoglycans by Human Siglec-10. iScience 2020, 23, 101231.

(159) French, A. D. Energy Maps for Glycosidic Linkage Conformations. Methods Mol. Biol. 2015, 1273, 333-358.

(160) French, A. D. Computerized Models of Carbohydrates. In Polysaccharides, Ramawat, K. G., Mérillon, J.-M. Eds.; Springer International Publishing, 2014; pp 1-38.

(161) Rodríguez-Carvajal, M. A.; Hervé du Penhoat, C.; Mazeau, K.; Doco, T.; Pérez, S. The Three-Dimensional Structure of the Mega-Oligosaccharide Rhamnogalacturonan II Monomer: A Combined Molecular Modeling and NMR Investigation. Carbohydr. Res. 2003, 338, 651-671.

(162) Bohé, L.; Crich, D. A Propos of Glycosyl Cations and the Mechanism of Chemical Glycosylation. C. R. Chim. 2011, 14, 3-16.

(163) Bohé, L.; Crich, D. A Propos of Glycosyl Cations and the Mechanism of Chemical Glycosylation; the Current State of the Art. Carbohydr. Res. 2015, 403, 48-59.

(164) Adero, P. O.; Amarasekara, H.; Wen, P.; Bohé, L.; Crich, D. The Experimental Evidence in Support of Glycosylation Mechanisms at the S(N)1-S(N)2 Interface. Chem. Rev. 2018, 118, 8242-8284.

(165) Mydock, L. K.; Demchenko, A. V. Mechanism of Chemical O-Glycosylation: From Early Studies to Recent Discoveries. Org. Biomol. Chem. 2010, 8, 497-510.

(166) Hansen, T.; van der Vorm, S.; Tugny, C.; Remmerswaal, W. A.; van Hengst, J. M. A.; van der Marel, G. A.; Codée, J. D. C. Stereoelectronic Effects in Glycosylation Reactions. In Comprehensive Glycoscience (Second Edition), Barchi, J. J. Ed.; Elsevier, 2021; pp 83-102.

(167) Hosoya, T.; Kosma, P.; Rosenau, T. Contact Ion Pairs and Solvent-Separated Ion Pairs from D-Mannopyranosyl and D-Glucopyranosyl Triflates. Carbohydr. Res. 2015, 401, 127-131.

(168) Hosoya, T.; Kosma, P.; Rosenau, T. Theoretical Study on the Effects of a 4,6-O-Diacetal Protecting Group on the Stability of Ion Pairs from D-Mannopyranosyl and D-Glucopyranosyl Triflates. Carbohydr. Res. 2015, 411, 64-69.

(169) Hosoya, T.; Takano, T.; Kosma, P.; Rosenau, T. Theoretical Foundation for the Presence of Oxacarbenium Ions in Chemical Glycoside Synthesis. J. Org. Chem. 2014, 79, 7889-7894.

(170) Whitfield, D. M. Dft Studies of the Ionization of Alpha and Beta Glycopyranosyl Donors. Carbohydr. Res. 2007, 342, 1726-1740.

(171) Whitfield, D. M. Plausible Transition States for Glycosylation Reactions. Carbohydr. Res. 2012, 356, 180-190.

(172) Whitfield, D. M. In a Glycosylation Reaction How Does a Hydroxylic Nucleophile Find the Activated Anomeric Carbon? Carbohydr. Res. 2015, 403, 69-89.

(173) Nukada, T.; Bérces, A.; Whitfield, D. M. Can the Stereochemical Outcome of Glycosylation Reactions Be Controlled by the Conformational Preferences of the Glycosyl Donor? Carbohydr. Res. 2002, 337, 765-774.

(174) Satoh, H.; Nukada, T. Computational Chemistry on Chemical Glycosylations. Trends Glycosci. Glycotechnol. 2014, 26, 11-27.

(175) Hansen, T.; Lebedel, L.; Remmerswaal, W. A.; van der Vorm, S.; Wander, D. P. A.; Somers, M.; Overkleeft, H. S.; Filippov, D. V.; Désiré, J.; Mingot, A.; et al. Defining the SN1 Side of Glycosylation Reactions: Stereoselectivity of Glycopyranosyl Cations. ACS Cent. Sci. 2019, 5, 781-788.

(176) Hansen, T.; Elferink, H.; van Hengst, J. M. A.; Houthuijs, K. J.; Remmerswaal, W. A.; Kromm, A.; Berden, G.; van der Vorm, S.; Rijs, A. M.; Overkleeft, H. S.; et al. Characterization of Glycosyl Dioxolenium Ions and Their Role in Glycosylation Reactions. Nat. Commun. 2020, 11, 2664.

(177) Dashnau, J. L.; Sharp, K. A.; Vanderkooi, J. M. Carbohydrate Intramolecular Hydrogen Bonding Cooperativity and Its Effect on Water Structure. J. Phys. Chem. B 2005, 109, 24152-24159.

(178) Engelsen, S. B.; Monteiro, C.; Herve de Penhoat, C.; Perez, S. The Diluted Aqueous Solvation of Carbohydrates as Inferred from Molecular Dynamics Simulations and NMR Spectroscopy. Biophys. Chem. 2001, 93, 103-127.

(179) Engelsen, S. B.; Pérez, S. The Hydration of Sucrose. Carbohydr. Res. 1996, 292, 21-38.

(180) Engelsen, S. B.; Pérez, S. Unique Similarity of the Asymmetric Trehalose Solid-State Hydration and the Diluted Aqueous-Solution Hydration. J. Phys. Chem. B 2000, 104, 9301-9311.

(181) Rojo, J.; Morales, J. C.; Penadés, S. Carbohydrate-Carbohydrate Interactions in Biological and Model Systems. In Host-Guest Chemistry: Mimetic Approaches to Study Carbohydrate Recognition, Penadés, S. Ed.; Springer Berlin Heidelberg, 2002; pp 45-92.

(182) Handa, K.; Takatani-Nakase, T.; Larue, L.; Stemmler, M. P.; Kemler, R.; Hakomori, S.-i. Lex Glycan Mediates Homotypic Adhesion of Embryonal Cells Independently from E-Cadherin: A Preliminary Note. Biochem. Biophys. Res. Commun. 2007, 358, 247-252.

(183) Seah, N.; Basu, A. Carbohydrate– Carbohydrate Interactions. In Wiley Encyclopedia of Chemical Biology, Begley, T. Ed.; 2008.

(184) Vilanova, E.; Ciodaro, P. J.; Bezerra, F. F.; Santos, G. R. C.; Valle-Delgado, J. J.; Anselmetti, D.; Fernàndez-Busquets, X.; Mourão, P. A. S. Adhesion of Freshwater Sponge Cells Mediated by Carbohydrate–Carbohydrate Interactions Requires Low Environmental Calcium. Glycobiology 2020, 30, 710-721.

(185) Kav, B.; Grafmüller, A.; Schneck, E.; Weikl, T. R. Weak Carbohydrate–Carbohydrate Interactions in Membrane Adhesion Are Fuzzy and Generic. Nanoscale 2020, 12, 17342-17353.

(186) Apweiler, R.; Hermjakob, H.; Sharon, N. On the Frequency of Protein Glycosylation, as Deduced from Analysis of the Swiss-Prot Database. Biochim. Biophys. Acta 1999, 1473, 4-8.

(187) Weerapana, E.; Imperiali, B. Asparagine-Linked Protein Glycosylation: From Eukaryotic to Prokaryotic Systems. Glycobiology 2006, 16, 91R-101R.

(188) Lizak, C.; Gerber, S.; Numao, S.; Aebi, M.; Locher, K. P. X-Ray Structure of a Bacterial Oligosaccharyltransferase. Nature 2011, 474, 350-355.

(189) Van den Steen, P.; Rudd, P. M.; Dwek, R. A.; Opdenakker, G. Concepts and Principles of O-Linked Glycosylation. Crit. Rev. Biochem. Mol. Biol. 1998, 33, 151-208.

(190) Varki, A. Biological Roles of Glycans. Glycobiology 2017, 27, 3-49.

(191) Reily, C.; Stewart, T. J.; Renfrow, M. B.; Novak, J. Glycosylation in Health and Disease. Nat. Rev. Neurosci. 2019, 15, 346-366.

(192) Andre, S.; Kozar, T.; Kojima, S.; Unverzagt, C.; Gabius, H. J. From Structural to Functional Glycomics: Core Substitutions as Molecular Switches for Shape and Lectin Affinity of N-Glycans. Biol. Chem. 2009, 390, 557-565.

(193) 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.

(194) Re, S.; Miyashita, N.; Yamaguchi, Y.; Sugita, Y. Structural Diversity and Changes in Conformational Equilibria of Biantennary Complex-Type N-Glycans in Water Revealed by Replica-Exchange Molecular Dynamics Simulation. Biophys. J. 2011, 101, L44-L46.

(195) Nishima, W.; Miyashita, N.; Yamaguchi, Y.; Sugita, Y.; Re, S. Effect of Bisecting Glcnac and Core Fucosylation on Conformational Properties of Biantennary Complex-Type N-Glycans in Solution. J. Phys. Chem. B 2012, 116, 8504-8512.

(196) 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.

(197) Oshima, H.; Re, S.; Sugita, Y. Replica-Exchange Umbrella Sampling Combined with Gaussian Accelerated Molecular Dynamics for Free-Energy Calculation of Biomolecules. J. Chem. Theory Comput. 2019, 15, 5199-5208.

(198) Yang, M.; Huang, J.; MacKerell, A. D. Enhanced Conformational Sampling Using Replica Exchange with Concurrent Solute Scaling and Hamiltonian Biasing Realized in One Dimension. J. Chem. Theory Comput. 2015, 11, 2855-2867.

(199) Yang, M.; MacKerell, A. D. Conformational Sampling of Oligosaccharides Using Hamiltonian Replica Exchange with Two-Dimensional Dihedral Biasing Potentials and the Weighted Histogram Analysis Method (Wham). J. Chem. Theory Comput. 2015, 11, 788-799.

(200) Re, S.; Yamaguchi, Y.; Sugita, Y. Molecular Dynamics Simulation of Glycans. Trends Glycosci. Glycotechnol. 2020, 32, E113-E118.

(201) Salisburg, A. M.; Deline, A. L.; Lexa, K. W.; Shields, G. C.; Kirschner, K. N. Ramachandran-Type Plots for Glycosidic Linkages: Examples from Molecular Dynamic Simulations Using the Glycam06 Force Field. J. Comput. Chem. 2009, 30, 910-921.

(202) Kern, N. R.; Lee, H. S.; Wu, E. L.; Park, S.; Vanommeslaeghe, K.; MacKerell, A. D., Jr.; Klauda, J. B.; Jo, S.; Im, W. Lipid-Linked Oligosaccharides in Membranes Sample Conformations That Facilitate Binding to Oligosaccharyltransferase. Biophys. J. 2014, 107, 1885-1895.

(203) Taylor, E. S.; Pol-Fachin, L.; Lins, R. D.; Lower, S. K. Conformational Stability of the Epidermal Growth Factor (Egf) Receptor as Influenced by Glycosylation, Dimerization and Egf Hormone Binding. Proteins 2017, 85, 561-570.

(204) Turupcu, A.; Oostenbrink, C. Modeling of Oligosaccharides within Glycoproteins from Free-Energy Landscapes. J. Chem. Inf. Model. 2017, 57, 2222-2236.

(205) Ellis, C. R.; Maiti, B.; Noid, W. G. Specific and Nonspecific Effects of Glycosylation. J. Am. Chem. Soc. 2012, 134, 8184-8193.

(206) Fogarty, C. A.; Harbison, A. M.; Dugdale, A. R.; Fadda, E. How and Why Plants and Human N-Glycans Are Different: Insight from Molecular Dynamics into the "Glycoblocks" Architecture of Complex Carbohydrates. Beilstein J. Org. Chem. 2020, 16, 2046-2056.

(207) (accessed.

(208) Fogarty, C. A.; Fadda, E. Oligomannose N-Glycans 3d Architecture and Its Response to the Fcγriiia Structural Landscape. J. Phys. Chem. B 2021, 125, 2607-2616.

(209) Guillot, A.; Dauchez, M.; Belloy, N.; Jonquet, J.; Duca, L.; Romier, B.; Maurice, P.; Debelle, L.; Martiny, L.; Durlach, V.; et al. Impact of Sialic Acids on the Molecular Dynamic of Bi-Antennary and Tri-Antennary Glycans. Sci. Rep. 2016, 6, 35666.

(210) Andre, S.; Unverzagt, C.; Kojima, S.; Frank, M.; Seifert, J.; Fink, C.; Kayser, K.; von der Lieth, C. W.; Gabius, H. J. Determination of Modulation of Ligand Properties of Synthetic Complex-Type Biantennary N-Glycans by Introduction of Bisecting Glcnac in Silico, in Vitro and in Vivo. Eur. J. Biochem. 2004, 271, 118-134.

(211) Jo, S.; Qi, Y.; Im, W. Preferred Conformations of N-Glycan Core Pentasaccharide in Solution and in Glycoproteins. Glycobiology 2016, 26, 19-29.

(212) Frank, M.; Walker, R. C.; Lanzilotta, W. N.; Prestegard, J. H.; Barb, A. W. Immunoglobulin G1 Fc Domain Motions: Implications for Fc Engineering. J. Mol. Biol. 2014, 426, 1799-1811.

(213) Pang, P.-C.; Chiu, P. C. N.; Lee, C.-L.; Chang, L.-Y.; Panico, M.; Morris, H. R.; Haslam, S. M.; Khoo, K.-H.; Clark, G. F.; Yeung, W. S. B.; et al. Human Sperm Binding Is Mediated by the Sialyl-Lewis(X) Oligosaccharide on the Zona Pellucida. Science 2011, 333, 1761-1764.

(214) Barthel, S. R.; Gavino, J. D.; Descheny, L.; Dimitroff, C. J. Targeting Selectins and Selectin Ligands in Inflammation and Cancer. Expert Opin. Ther. Targets 2007, 11, 1473-1491.

(215) Mitra, N.; Sinha, S.; Ramya, T. N.; Surolia, A. N-Linked Oligosaccharides as Outfitters for Glycoprotein Folding, Form and Function. Trends Biochem. Sci. 2006, 31, 156-163.

(216) Sakae, Y.; Satoh, T.; Yagi, H.; Yanaka, S.; Yamaguchi, T.; Isoda, Y.; Iida, S.; Okamoto, Y.; Kato, K. Conformational Effects of N-Glycan Core Fucosylation of Immunoglobulin G Fc Region on Its Interaction with Fcγ Receptor LIIa. Sci. Rep. 2017, 7, 13780.

(217) 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.

(218) Pan, D.; Song, Y. Role of Altered Sialylation of the I-Like Domain of Beta1 Integrin in the Binding of Fibronectin to Beta1 Integrin: Thermodynamics and Conformational Analyses. Biophys. J. 2010, 99, 208-217.

(219) Yang, M.; Huang, J.; Simon, R.; Wang, L. X.; MacKerell, A. D., Jr. Conformational Heterogeneity of the Hiv Envelope Glycan Shield. Sci. Rep. 2017, 7, 4435.

(220) Wood, N. T.; Fadda, E.; Davis, R.; Grant, O. C.; Martin, J. C.; Woods, R. J.; Travers, S. A. The Influence of N-Linked Glycans on the Molecular Dynamics of the Hiv-1 Gp120 V3 Loop. PLoS One 2013, 8, e80301.

(221) Qi, Y.; Jo, S.; Im, W. Roles of Glycans in Interactions between Gp120 and Hiv Broadly Neutralizing Antibodies. Glycobiology 2016, 26, 251-260.

(222) Kiyoshi, M.; Tsumoto, K.; Ishii-Watabe, A.; Caaveiro, J. M. M. Glycosylation of IgG-Fc: A Molecular Perspective. Int. Immunol. 2017, 29, 311-317.

(223) Shivgan, A. T.; Marzinek, J. K.; Huber, R. G.; Krah, A.; Henchman, R. H.; Matsudaira, P.; Verma, C. S.; Bond, P. J. Extending the Martini Coarse-Grained Force Field to N-Glycans. J. Chem. Inf. Model. 2020, 60, 3864-3883.

(224) Hoffmann, M.; Kleine-Weber, H.; Schroeder, S.; Kruger, N.; Herrler, T.; Erichsen, S.; Schiergens, T. S.; Herrler, G.; Wu, N. H.; Nitsche, A.; et al. SARS-CoV-2 Cell Entry Depends on ACE2 and Tmprss2 and Is Blocked by a Clinically Proven Protease Inhibitor. Cell 2020, 181, 271-280.e8.

(225) Li, W.; Moore, M. J.; Vasilieva, N.; Sui, J.; Wong, S. K.; Berne, M. A.; Somasundaran, M.; Sullivan, J. L.; Luzuriaga, K.; Greenough, T. C.; et al. Angiotensin-Converting Enzyme 2 Is a Functional Receptor for the Sars Coronavirus. Nature 2003, 426, 450-454.

(226) Shajahan, A.; Supekar, N. T.; Gleinich, A. S.; Azadi, P. Deducing the N- and O-Glycosylation Profile of the Spike Protein of Novel Coronavirus SARS-CoV-2. Glycobiology 2020, 30, 981-988.

(227) Sanda, M.; Morrison, L.; Goldman, R. N- and O-Glycosylation of the SARS-CoV-2 Spike Protein. Anal. Chem. 2021, 93, 2003-2009.

(228) Watanabe, Y.; Allen, J. D.; Wrapp, D.; McLellan, J. S.; Crispin, M. Site-Specific Glycan Analysis of the SARS-CoV-2 Spike. Science 2020, 369, 330-333.

(229) Zhao, P.; Praissman, J. L.; Grant, O. C.; Cai, Y.; Xiao, T.; Rosenbalm, K. E.; Aoki, K.; Kellman, B. P.; Bridger, R.; Barouch, D. H.; et al. Virus-Receptor Interactions of Glycosylated SARS-CoV-2 Spike and Human ACE2 Receptor. Cell Host Microbe 2020, 28, 586-601.e6.

(230) Watanabe, Y.; Mendonça, L.; Allen, E. R.; Howe, A.; Lee, M.; Allen, J. D.; Chawla, H.; Pulido, D.; Donnellan, F.; Davies, H.; et al. Native-Like SARS-CoV-2 Spike Glycoprotein Expressed by Chadox1 Ncov-19/Azd1222 Vaccine. ACS Cent. Sci. 2021, 7, 594-602.

(231) Wei, X.; Decker, J. M.; Wang, S.; Hui, H.; Kappes, J. C.; Wu, X.; Salazar-Gonzalez, J. F.; Salazar, M. G.; Kilby, J. M.; Saag, M. S.; et al. Antibody Neutralization and Escape by Hiv-1. Nature 2003, 422, 307-312.

(232) 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.

(233) Wrapp, D.; Wang, N.; Corbett, K. S.; Goldsmith, J. A.; Hsieh, C. L.; Abiona, O.; Graham, B. S.; McLellan, J. S. Cryo-Em Structure of the 2019-Ncov Spike in the Prefusion Conformation. Science 2020, 367, 1260-1263.

(234) Casalino, L.; Gaieb, Z.; Goldsmith, J. A.; Hjorth, C. K.; Dommer, A. C.; Harbison, A. M.; Fogarty, C. A.; Barros, E. P.; Taylor, B. C.; McLellan, J. S.; et al. Beyond Shielding: The Roles of Glycans in the SARS-CoV-2 Spike Protein. ACS Cent. Sci. 2020, 6, 1722-1734.

(235) Grant, O. C.; Montgomery, D.; Ito, K.; Woods, R. J. Analysis of the SARS-CoV-2 Spike Protein Glycan Shield Reveals Implications for Immune Recognition. Sci. Rep. 2020, 10, 14991.

(236) Woo, H.; Park, S. J.; Choi, Y. K.; Park, T.; Tanveer, M.; Cao, Y.; Kern, N. R.; Lee, J.; Yeom, M. S.; Croll, T. I.; et al. Developing a Fully Glycosylated Full-Length SARS-CoV-2 Spike Protein Model in a Viral Membrane. J. Phys. Chem. B 2020, 124, 7128-7137.

(237) Brooks, B. R.; Brooks, C. L., 3rd; Mackerell, A. D., Jr.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; et al. Charmm: The Biomolecular Simulation Program. J. Comput. Chem. 2009, 30, 1545-1614.

(238) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K. Scalable Molecular Dynamics with Namd. J. Comput. Chem. 2005, 26, 1781-1802.

(239) Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E. Gromacs: High Performance Molecular Simulations through Multi-Level Parallelism from Laptops to Supercomputers. SoftwareX 2015, 1-2, 19-25.

(240) Case, D. A.; Cheatham, T. E., 3rd; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M., Jr.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. The Amber Biomolecular Simulation Programs. J. Comput. Chem. 2005, 26, 1668-1688.

(241) Jung, J.; Mori, T.; Kobayashi, C.; Matsunaga, Y.; Yoda, T.; Feig, M.; Sugita, Y. Genesis: A Hybrid-Parallel and Multi-Scale Molecular Dynamics Simulator with Enhanced Sampling Algorithms for Biomolecular and Cellular Simulations. Wiley Interdiscip. Rev. Comput. Mol. Sci. 2015, 5, 310-323.

(242) Eastman, P.; Friedrichs, M. S.; Chodera, J. D.; Radmer, R. J.; Bruns, C. M.; Ku, J. P.; Beauchamp, K. A.; Lane, T. J.; Wang, L. P.; Shukla, D.; et al. Openmm 4: A Reusable, Extensible, Hardware Independent Library for High Performance Molecular Simulation. J. Chem. Theory Comput. 2013, 9, 461-469.

(243) Eastman, P.; Swails, J.; Chodera, J. D.; McGibbon, R. T.; Zhao, Y.; Beauchamp, K. A.; Wang, L. P.; Simmonett, A. C.; Harrigan, M. P.; Stern, C. D.; et al. Openmm 7: Rapid Development of High Performance Algorithms for Molecular Dynamics. PLoS Comput. Biol. 2017, 13, e1005659.

(244) Harbison, A. M.; Fogarty, C. A.; Phung, T. K.; Satheesan, A.; Schulz, B. L.; Fadda, E. Fine-Tuning the Spike: Role of the Nature and Topology of the Glycan Shield in the Structure and Dynamics of the SARS-CoV-2 S. Chem. Sci. 2022, 13, 386-395.

(245) Clausen, T. M.; Sandoval, D. R.; Spliid, C. B.; Pihl, J.; Perrett, H. R.; Painter, C. D.; Narayanan, A.; Majowicz, S. A.; Kwong, E. M.; McVicar, R. N.; et al. SARS-CoV-2 Infection Depends on Cellular Heparan Sulfate and ACE2. Cell 2020, 183, 1043-1057.e15.

(246) Schuurs, Z. P.; Hammond, E.; Elli, S.; Rudd, T. R.; Mycroft-West, C. J.; Lima, M. A.; Skidmore, M. A.; Karlsson, R.; Chen, Y. H.; Bagdonaite, I.; et al. Evidence of a Putative Glycosaminoglycan Binding Site on the Glycosylated SARS-CoV-2 Spike Protein N-Terminal Domain. Comput. Struct. Biotechnol. J. 2021, 19, 2806-2818.

(247) Nguyen, L.; McCord, K. A.; Bui, D. T.; Bouwman, K. M.; Kitova, E. N.; Elaish, M.; Kumawat, D.; Daskhan, G. C.; Tomris, I.; Han, L.; et al. Sialic Acid-Containing Glycolipids Mediate Binding and Viral Entry of SARS-CoV-2. Nat. Chem. Biol. 2022, 18, 81-90.

(248) Choi, Y. K.; Cao, Y.; Frank, M.; Woo, H.; Park, S. J.; Yeom, M. S.; Croll, T. I.; Seok, C.; Im, W. Structure, Dynamics, Receptor Binding, and Antibody Binding of the Fully Glycosylated Full-Length SARS-CoV-2 Spike Protein in a Viral Membrane. J. Chem. Theory Comput. 2021, 17, 2479-2487.

(249) Korber, B.; Fischer, W. M.; Gnanakaran, S.; Yoon, H.; Theiler, J.; Abfalterer, W.; Hengartner, N.; Giorgi, E. E.; Bhattacharya, T.; Foley, B.; et al. Tracking Changes in SARS-CoV-2 Spike: Evidence That D614g Increases Infectivity of the Covid-19 Virus. Cell 2020, 182, 812-827.e19.

(250) Mansbach, R. A.; Chakraborty, S.; Nguyen, K.; Montefiori, D. C.; Korber, B.; Gnanakaran, S. The SARS-CoV-2 Spike Variant D614g Favors an Open Conformational State. Sci. Adv. 2021, 7, eabf3671.

(251) Thomson, E. C.; Rosen, L. E.; Shepherd, J. G.; Spreafico, R.; da Silva Filipe, A.; Wojcechowskyj, J. A.; Davis, C.; Piccoli, L.; Pascall, D. J.; Dillen, J.; et al. Circulating SARS-CoV-2 Spike N439k Variants Maintain Fitness While Evading Antibody-Mediated Immunity. Cell 2021, 184, 1171-1187.e20.

(252) Cao, W.; Dong, C.; Kim, S.; Hou, D.; Tai, W.; Du, L.; Im, W.; Zhang, X. F. Biomechanical Characterization of SARS-CoV-2 Spike Rbd and Human ACE2 Protein-Protein Interaction. Biophys. J. 2021, 120, 1011-1019.

(253) Verkhivker, G. M.; Di Paola, L. Dynamic Network Modeling of Allosteric Interactions and Communication Pathways in the SARS-CoV-2 Spike Trimer Mutants: Differential Modulation of Conformational Landscapes and Signal Transmission Via Cascades of Regulatory Switches. J. Phys. Chem. B 2021, 125, 850-873.

(254) Mori, T.; Jung, J.; Kobayashi, C.; Dokainish, H. M.; Re, S.; Sugita, Y. Elucidation of Interactions Regulating Conformational Stability and Dynamics of SARS-CoV-2 S-Protein. Biophys. J. 2021, 120, 1060-1071.

(255) Yu, A.; Pak, A. J.; He, P.; Monje-Galvan, V.; Casalino, L.; Gaieb, Z.; Dommer, A. C.; Amaro, R. E.; Voth, G. A. A Multiscale Coarse-Grained Model of the SARS-CoV-2 Virion. Biophys. J. 2021, 120, 1097-1104.

(256) Zimmerman, M. I.; Porter, J. R.; Ward, M. D.; Singh, S.; Vithani, N.; Meller, A.; Mallimadugula, U. L.; Kuhn, C. E.; Borowsky, J. H.; Wiewiora, R. P.; et al. SARS-CoV-2 Simulations Go Exascale to Predict Dramatic Spike Opening and Cryptic Pockets across the Proteome. Nat. Chem. 2021, 13, 651-659.

(257) Kaszuba, K.; Grzybek, M.; Orłowski, A.; Danne, R.; Róg, T.; Simons, K.; Coskun, Ü.; Vattulainen, I. N-Glycosylation as Determinant of Epidermal Growth Factor Receptor Conformation in Membranes. Proc. Natl. Acad. Sci. USA 2015, 112, 4334-4339.

(258) Polley, A.; Orłowski, A.; Danne, R.; Gurtovenko, A. A.; Bernardino de la Serna, J.; Eggeling, C.; Davis, S. J.; Róg, T.; Vattulainen, I. Glycosylation and Lipids Working in Concert Direct Cd2 Ectodomain Orientation and Presentation. J. Phys. Chem. Lett. 2017, 8, 1060-1066.

(259) Wehle, M.; Vilotijevic, I.; Lipowsky, R.; Seeberger, P. H.; Silva, D. V.; Santer, M. Mechanical Compressibility of the Glycosylphosphatidylinositol (Gpi) Anchor Backbone Governed by Independent Glycosidic Linkages. J. Am. Chem. Soc. 2012, 134, 18964-18972.

(260) Banerjee, P.; Lipowsky, R.; Santer, M. Coarse-Grained Molecular Model for the Glycosylphosphatidylinositol Anchor with and without Protein. J. Chem. Theory Comput. 2020, 16, 3889-3903.

(261) Vasudevan, S. V.; Balaji, P. V. Dynamics of Ganglioside Headgroup in Lipid Environment: Molecular Dynamics Simulations of Gm1 Embedded in Dodecylphosphocholine Micelle. J. Phys. Chem. B 2001, 105, 7033-7041.

(262) Sega, M.; Jedlovszky, P.; Vallauri, R. Molecular Dynamics Simulation of Gm1 Gangliosides Embedded in a Phospholipid Membrane. J. Mol. Liq. 2006, 129, 86-91.

(263) Owen, M. C.; Karner, A.; Sachl, R.; Preiner, J.; Amaro, M.; Vacha, R. Force Field Comparison of Gm1 in a Dopc Bilayer Validated with Afm and Fret Experiments. J. Phys. Chem. B 2019, 123, 7504-7517.

(264) Rissanen, S.; Grzybek, M.; Orłowski, A.; Róg, T.; Cramariuc, O.; Levental, I.; Eggeling, C.; Sezgin, E.; Vattulainen, I. Phase Partitioning of Gm1 and Its Bodipy-Labeled Analog Determine Their Different Binding to Cholera Toxin. Front. Physiol. 2017, 8, 252.

(265) Lingwood, D.; Binnington, B.; Róg, T.; Vattulainen, I.; Grzybek, M.; Coskun, U.; Lingwood, C. A.; Simons, K. Cholesterol Modulates Glycolipid Conformation and Receptor Activity. Nat. Chem. Biol. 2011, 7, 260-262.

(266) Ryu, Y. S.; Jordan, L. R.; Wittenberg, N. J.; Kim, S. M.; Yoo, D.; Jeong, C.; Warrington, A. E.; Rodriguez, M.; Oh, S. H.; Lee, S. D. Curvature Elasticity-Driven Leaflet Asymmetry and Interleaflet Raft Coupling in Supported Membranes. Adv. Mater. Interfaces 2018, 5, 1801290.

(267) Dasgupta, R.; Miettinen, M. S.; Fricke, N.; Lipowsky, R.; Dimova, R. The Glycolipid Gm1 Reshapes Asymmetric Biomembranes and Giant Vesicles by Curvature Generation. Proc. Natl. Acad. Sci. USA 2018, 115, 5756-5761.

(268) Gu, R. X.; Ingolfsson, H. I.; de Vries, A. H.; Marrink, S. J.; Tieleman, D. P. Ganglioside-Lipid and Ganglioside-Protein Interactions Revealed by Coarse-Grained and Atomistic Molecular Dynamics Simulations. J. Phys. Chem. B 2017, 121, 3262-3275.

(269) Humphrey, W.; Dalke, A.; Schulten, K. Vmd: Visual Molecular Dynamics. J. Mol. Graph. 1996, 14, 33-38.

(270) Hedger, G.; Shorthouse, D.; Koldso, H.; Sansom, M. S. P. Free Energy Landscape of Lipid Interactions with Regulatory Binding Sites on the Transmembrane Domain of the Egf Receptor. J. Phys. Chem. B 2016, 120, 8154-8163.

(271) Shorthouse, D.; Hedger, G.; Koldso, H.; Sansom, M. S. Molecular Simulations of Glycolipids: Towards Mammalian Cell Membrane Models. Biochimie 2016, 120, 105-109.

(272) Koldso, H.; Shorthouse, D.; Helie, J.; Sansom, M. S. Lipid Clustering Correlates with Membrane Curvature as Revealed by Molecular Simulations of Complex Lipid Bilayers. PLoS Comput. Biol. 2014, 10, e1003911.

(273) Kociurzynski, R.; Pannuzzo, M.; Bockmann, R. A. Phase Transition of Glycolipid Membranes Studied by Coarse-Grained Simulations. Langmuir 2015, 31, 9379-9387.

(274) Ingólfsson, H. I.; Melo, M. N.; van Eerden, F. J.; Arnarez, C.; Lopez, C. A.; Wassenaar, T. A.; Periole, X.; de Vries, A. H.; Tieleman, D. P.; Marrink, S. J. Lipid Organization of the Plasma Membrane. J. Am. Chem. Soc. 2014, 136, 14554-14559.

(275) Pezeshkian, W.; Nåbo, L. J.; Ipsen, J. H. Cholera Toxin B Subunit Induces Local Curvature on Lipid Bilayers. FEBS Open Bio 2017, 7, 1638-1645.

(276) Kociurzynski, R.; Beck, S. D.; Bouhon, J. B.; Römer, W.; Knecht, V. Binding of Sv40's Viral Capsid Protein Vp1 to Its Glycosphingolipid Receptor Gm1 Induces Negative Membrane Curvature: A Molecular Dynamics Study. Langmuir 2019, 35, 3534-3544.

(277) Pezeshkian, W.; Gao, H. F.; Arumugam, S.; Becken, U.; Bassereau, P.; Florent, J. C.; Ipsen, J. H.; Johannes, L.; Shillcock, J. C. Mechanism of Shiga Toxin Clustering on Membranes. ACS Nano 2017, 11, 314-324.

(278) Pezeshkian, W.; Hansen, A. G.; Johannes, L.; Khandelia, H.; Shillcock, J. C.; Kumar, P. B. S.; Ipsen, J. H. Membrane Invagination Induced by Shiga Toxin B-Subunit: From Molecular Structure to Tube Formation. Soft Matter 2016, 12, 5164-5171.

(279) Kociurzynski, R.; Makshakova, O. N.; Knecht, V.; Romer, W. Multiscale Molecular Dynamics Studies Reveal Different Modes of Receptor Clustering by Gb3-Binding Lectins. J. Chem. Theory Comput. 2021, 17, 2488-2501.

(280) van Eerden, F. J.; Melo, M. N.; Frederix, P.; Marrink, S. J. Prediction of Thylakoid Lipid Binding Sites on Photosystem Ii. Biophys. J. 2017, 113, 2669-2681.

(281) van Eerden, F. J.; van den Berg, T.; Frederix, P.; de Jong, D. H.; Periole, X.; Marrink, S. J. Molecular Dynamics of Photosystem Ii Embedded in the Thylakoid Membrane. J. Phys. Chem. B 2017, 121, 3237-3249.

(282) Lee, C. K.; Pao, C. W.; Smit, B. Psii-Lhcii Supercomplex Organizations in Photosynthetic Membrane by Coarse-Grained Simulation. J. Phys. Chem. B 2015, 119, 3999-4008.

(283) Albanese, P.; Melero, R.; Engel, B. D.; Grinzato, A.; Berto, P.; Manfredi, M.; Chiodoni, A.; Vargas, J.; Sorzano CÓ, S.; Marengo, E.; et al. Pea Psii-Lhcii Supercomplexes Form Pairs by Making Connections across the Stromal Gap. Sci. Rep. 2017, 7, 10067.

(284) van Eerden, F. J.; Melo, M. N.; Frederix, P.; Periole, X.; Marrink, S. J. Exchange Pathways of Plastoquinone and Plastoquinol in the Photosystem Ii Complex. Nat. Commun. 2017, 8, 15214.

(285) Ogata, K.; Hatakeyama, M.; Sakamoto, Y.; Nakamura, S. Investigation of a Pathway for Water Delivery in Photosystem Ii Protein by Molecular Dynamics Simulation. J. Phys. Chem. B 2019, 123, 6444-6452.

(286) van Eerden, F. J.; de Jong, D. H.; de Vries, A. H.; Wassenaar, T. A.; Marrink, S. J. Characterization of Thylakoid Lipid Membranes from Cyanobacteria and Higher Plants by Molecular Dynamics Simulations. Biochim. Biophys. Acta 2015, 1848, 1319-1330.

(287) Liguori, N.; Croce, R.; Marrink, S. J.; Thallmair, S. Molecular Dynamics Simulations in Photosynthesis. Photosynth. Res. 2020, 144, 273-295.

(288) Toukach, P. V.; Egorova, K. S. Carbohydrate Structure Database Merged from Bacterial, Archaeal, Plant and Fungal Parts. Nucleic Acids Res. 2016, 44, D1229-1236.

(289) French, A. D.; Pérez, S.; Bulone, V.; Rosenau, T.; Gray, D. Cellulose. In Encyclopedia of Polymer Science and Technology, 2018; pp 1-69.

(290) Perez, S.; Wertz, J.-L. Chitin and Chitosans in the Bioeconomy; CRC Press, 2021.

(291) Spinozzi, F.; Ferrero, C.; Perez, S. The Architecture of Starch Blocklets Follows Phyllotaxic Rules. Sci. Rep. 2020, 10, 20093.

(292) Bellesia, G.; Chundawat, S. P.; Langan, P.; Redondo, A.; Dale, B. E.; Gnanakaran, S. Coarse-Grained Model for the Interconversion between Native and Liquid Ammonia-Treated Crystalline Cellulose. J. Phys. Chem. B 2012, 116, 8031-8037.

(293) Glass, D. C.; Moritsugu, K.; Cheng, X.; Smith, J. C. Reach Coarse-Grained Simulation of a Cellulose Fiber. Biomacromolecules 2012, 13, 2634-2644.

(294) Lopez, C. A.; Bellesia, G.; Redondo, A.; Langan, P.; Chundawat, S. P.; Dale, B. E.; Marrink, S. J.; Gnanakaran, S. Martini Coarse-Grained Model for Crystalline Cellulose Microfibers. J. Phys. Chem. B 2015, 119, 465-473.

(295) Markutsya, S.; Devarajan, A.; Baluyut, J. Y.; Windus, T. L.; Gordon, M. S.; Lamm, M. H. Evaluation of Coarse-Grained Mapping Schemes for Polysaccharide Chains in Cellulose. J. Chem. Phys. 2013, 138, 214108.

(296) Poma, A. B.; Chwastyk, M.; Cieplak, M. Coarse-Grained Model of the Native Cellulose I Alpha and the Transformation Pathways to the I Beta Allomorph (Vol 23, Pg 1573, 2016). Cellulose 2016, 23, 2247-2247.

(297) Poma, A. B.; Chwastyk, M.; Cieplak, M. Elastic Moduli of Biological Fibers in a Coarse-Grained Model: Crystalline Cellulose and Beta-Amyloids. Phys. Chem. Chem. Phys. 2017, 19, 28195-28206.

(298) Shen, T.; Gnanakaran, S. The Stability of Cellulose: A Statistical Perspective from a Coarse-Grained Model of Hydrogen-Bond Networks. Biophys. J. 2009, 96, 3032-3040.

(299) Srinivas, G.; Cheng, X.; Smith, J. C. Coarse-Grain Model for Natural Cellulose Fibrils in Explicit Water. J. Phys. Chem. B 2014, 118, 3026-3034.

(300) Fan, B.; Maranas, J. Coarse-Grained Simulation of Cellulose Iβ with Application to Long Fibrils. Cellulose 2014, 22, 31-44.

(301) Srinivas, G.; Cheng, X.; Smith, J. C. A Solvent-Free Coarse Grain Model for Crystalline and Amorphous Cellulose Fibrils. J. Chem. Theory Comput. 2011, 7, 2539-2548.

(302) Mirzaeifar, R.; Qin, Z.; Buehler, M. J. Mesoscale Mechanics of Twisting Carbon Nanotube Yarns. Nanoscale 2015, 7, 5435-5445.

(303) Shishehbor, M.; Zavattieri, P. D. Effects of Interface Properties on the Mechanical Properties of Bio-Inspired Cellulose Nanocrystal (CNC)-Based Materials. J. Mech. Phys. Solids. 2019, 124, 871-896.

(304) Ramezani, M. G.; Golchinfar, B. Mechanical Properties of Cellulose Nanocrystal (CNC) Bundles: Coarse-Grained Molecular Dynamic Simulation. J. Compos. Sci. 2019, 3, 57.

(305) Li, L.; Perre, P.; Frank, X.; Mazeau, K. A Coarse-Grain Force-Field for Xylan and Its Interaction with Cellulose. Carbohydr. Polym. 2015, 127, 438-450.

(306) Sauter, J.; Grafmuller, A. Procedure for Transferable Coarse-Grained Models of Aqueous Polysaccharides. J. Chem. Theory Comput. 2017, 13, 223-236.

(307) Benner, S. W.; Hall, C. K. Development of a Coarse-Grained Model of Chitosan for Predicting Solution Behavior. J. Phys. Chem. B 2016, 120, 7253-7264.

(308) Yu, Z.; Lau, D. Development of a Coarse-Grained Alpha-Chitin Model on the Basis of Martini Forcefield. J. Mol. Model. 2015, 21, 128.

(309) Tsereteli, L.; Grafmuller, A. An Accurate Coarse-Grained Model for Chitosan Polysaccharides in Aqueous Solution. PLoS One 2017, 12, e0180938.

(310) Bathe, M.; Rutledge, G. C.; Grodzinsky, A. J.; Tidor, B. A Coarse-Grained Molecular Model for Glycosaminoglycans: Application to Chondroitin, Chondroitin Sulfate, and Hyaluronic Acid. Biophys. J. 2005, 88, 3870-3887.

(311) Samsonov, S. A.; Bichmann, L.; Pisabarro, M. T. Coarse-Grained Model of Glycosaminoglycans. J. Chem. Inf. Model. 2015, 55, 114-124.

(312) Samsonov, S. A.; Lubecka, E. A.; Bojarski, K. K.; Ganzynkowicz, R.; Liwo, A. Local and Long Range Potentials for Heparin-Protein Systems for Coarse-Grained Simulations. Biopolymers 2019, 110, e23269.

(313) Sattelle, B. M.; Shakeri, J.; Cliff, M. J.; Almond, A. Proteoglycans and Their Heterogeneous Glycosaminoglycans at the Atomic Scale. Biomacromolecules 2015, 16, 951-961.

(314) Beltran-Villegas, D. J.; Intriago, D.; Kim, K. H. C.; Behabtu, N.; Londono, J. D.; Jayaraman, A. Coarse-Grained Molecular Dynamics Simulations of A-1,3-Glucan. Soft Matter 2019, 15, 4669-4681.

(315) Ghobadi, A. F.; Jayaraman, A. Effect of Backbone Chemistry on Hybridization Thermodynamics of Oligonucleic Acids: A Coarse-Grained Molecular Dynamics Simulation Study. Soft Matter 2016, 12, 2276-2287.

(316) Brant, D. A. Conformational Theory Applied to Polysaccharide Structure. Q. Rev. Biophys. 1976, 9, 527-596.

(317) Noto, R.; Martorana, V.; Bulone, D.; San Biagio, P. L. Role of Charges and Solvent on the Conformational Properties of Poly(Galacturonic Acid) Chains: A Molecular Dynamics Study. Biomacromolecules 2005, 6, 2555-2562.

(318) Meng, Y.; Shi, X.; Cai, L.; Zhang, S.; Ding, K.; Nie, S.; Luo, C.; Xu, X.; Zhang, L. Triple-Helix Conformation of a Polysaccharide Determined with Light Scattering, Afm, and Molecular Dynamics Simulation. Macromolecules 2018, 51, 10150-10159.

(319) Makshakova, O.; Zykwinska, A.; Cuenot, S.; Colliec-Jouault, S.; Perez, S. Three-Dimensional Structures, Dynamics and Calcium-Mediated Interactions of the Exopolysaccharide, Infernan, Produced by the Deep-Sea Hydrothermal Bacterium Alteromonas Infernus. Carbohydr. Polym. 2022, 276, 118732.

(320) Zhang, Y.; Yu, J.; Wang, X.; Durachko, D. M.; Zhang, S.; Cosgrove, D. J. Molecular Insights into the Complex Mechanics of Plant Epidermal Cell Walls. Science 2021, 372, 706-711.

(321) Gupta, M.; Rawal, T. B.; Dupree, P.; Smith, J. C.; Petridis, L. Spontaneous Rearrangement of Acetylated Xylan on Hydrophilic Cellulose Surfaces. Cellulose 2021, 28, 3327-3345.

(322) Kishani, S.; Benselfelt, T.; Wågberg, L.; Wohlert, J. Entropy Drives the Adsorption of Xyloglucan to Cellulose Surfaces – a Molecular Dynamics Study. J. Colloid Interface Sci. 2021, 588, 485-493.

(323) Di Lorenzo, F.; Duda, K. A.; Lanzetta, R.; Silipo, A.; De Castro, C.; Molinaro, A. A Journey from Structure to Function of Bacterial Lipopolysaccharides. Chem. Rev. 2021.

(324) Piggot, T. J.; Holdbrook, D. A.; Khalid, S. Conformational Dynamics and Membrane Interactions of the E. Coli Outer Membrane Protein Feca: A Molecular Dynamics Simulation Study. Biochim. Biophys. Acta 2013, 1828, 284-293.

(325) Holdbrook, D. A.; Piggot, T. J.; Sansom, M. S. P.; Khalid, S. Stability and Membrane Interactions of an Autotransport Protein: MD Simulations of the Hia Translocator Domain in a Complex Membrane Environment. Biochim. Biophys. Acta 2013, 1828, 715-723.

(326) Shearer, J.; Jefferies, D.; Khalid, S. Outer Membrane Proteins Ompa, Fhua, Ompf, Esta, Btub, and Ompx Have Unique Lipopolysaccharide Fingerprints. J. Chem. Theory Comput. 2019, 15, 2608-2619.

(327) Jefferies, D.; Shearer, J.; Khalid, S. Role of O-Antigen in Response to Mechanical Stress of the E. Coli Outer Membrane: Insights from Coarse-Grained MD Simulations. J. Phys. Chem. B 2019, 123, 3567-3575.

(328) Jefferies, D.; Khalid, S. Atomistic and Coarse-Grained Simulations of Membrane Proteins: A Practical Guide. Methods 2021, 185, 15-27.

(329) Im, W.; Khalid, S. Molecular Simulations of Gram-Negative Bacterial Membranes Come of Age. Annu. Rev. Phys. Chem. 2020, 71, 171-188.

(330) Hsu, P. C.; Bruininks, B. M. H.; Jefferies, D.; Cesar Telles de Souza, P.; Lee, J.; Patel, D. S.; Marrink, S. J.; Qi, Y.; Khalid, S.; Im, W. Charmm-Gui Martini Maker for Modeling and Simulation of Complex Bacterial Membranes with Lipopolysaccharides. J. Comput. Chem. 2017, 38, 2354-2363.

(331) van Oosten, B.; Harroun, T. A. A Martini Extension for Pseudomonas Aeruginosa Pao1 Lipopolysaccharide. J. Mol. Graph. Model. 2016, 63, 125-133.

(332) López, C. A.; Zgurskaya, H.; Gnanakaran, S. Molecular Characterization of the Outer Membrane of Pseudomonas Aeruginosa. Biochim. Biophys. Acta Biomembr. 2020, 1862, 183151.

(333) Matamoros-Recio, A.; Franco-Gonzalez, J. F.; Forgione, R. E.; Torres-Mozas, A.; Silipo, A.; Martín-Santamaría, S. Understanding the Antibacterial Resistance: Computational Explorations in Bacterial Membranes. ACS Omega 2021, 6, 6041-6054.

(334) 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.

(335) De Nicola, A.; Soares, T. A.; Santos, D. E. S.; Bore, S. L.; Sevink, G. J. A.; Cascella, M.; Milano, G. Aggregation of Lipid a Variants: A Hybrid Particle-Field Model. Biochim. Biophys. Acta Gen. Subj. 2021, 1865, 129570.

(336) Ohto, U.; Fukase, K.; Miyake, K.; Satow, Y. Crystal Structures of Human Md-2 and Its Complex with Antiendotoxic Lipid Iva. Science 2007, 316, 1632-1634.

(337) Park, B. S.; Song, D. H.; Kim, H. M.; Choi, B. S.; Lee, H.; Lee, J. O. The Structural Basis of Lipopolysaccharide Recognition by the Tlr4-Md-2 Complex. Nature 2009, 458, 1191-1195.

(338) Resman, N.; Vasl, J.; Oblak, A.; Pristovsek, P.; Gioannini, T. L.; Weiss, J. P.; Jerala, R. Essential Roles of Hydrophobic Residues in Both Md-2 and Toll-Like Receptor 4 in Activation by Endotoxin. J. Biol. Chem. 2009, 284, 15052-15060.

(339) Meng, J.; Lien, E.; Golenbock, D. T. Md-2-Mediated Ionic Interactions between Lipid a and Tlr4 Are Essential for Receptor Activation. J. Biol. Chem. 2010, 285, 8695-8702.

(340) Ohto, U.; Fukase, K.; Miyake, K.; Shimizu, T. Structural Basis of Species-Specific Endotoxin Sensing by Innate Immune Receptor Tlr4/Md-2. Proc. Natl. Acad. Sci. USA 2012, 109, 7421-7426.

(341) DeMarco, M. L.; Woods, R. J. From Agonist to Antagonist: Structure and Dynamics of Innate Immune Glycoprotein Md-2 Upon Recognition of Variably Acylated Bacterial Endotoxins. Mol. Immunol. 2011, 49, 124-133.

(342) Garate, J. A.; Oostenbrink, C. Lipid a from Lipopolysaccharide Recognition: Structure, Dynamics and Cooperativity by Molecular Dynamics Simulations. Proteins 2013, 81, 658-674.

(343) Scior, T.; Alexander, C.; Zaehringer, U. Reviewing and Identifying Amino Acids of Human, Murine, Canine and Equine Tlr4 / Md-2 Receptor Complexes Conferring Endotoxic Innate Immunity Activation by Lps/Lipid a, or Antagonistic Effects by Eritoran, in Contrast to Species-Dependent Modulation by Lipid Iva. Comput. Struct. Biotechnol. J. 2013, 5, e201302012.

(344) Scior, T.; Lozano-Aponte, J.; Figueroa-Vazquez, V.; Yunes-Rojas, J. A.; Zähringer, U.; Alexander, C. Three-Dimensional Mapping of Differential Amino Acids of Human, Murine, Canine and Equine Tlr4/Md-2 Receptor Complexes Conferring Endotoxic Activation by Lipid a, Antagonism by Eritoran and Species-Dependent Activities of Lipid Iva in the Mammalian Lps Sensor System. Comput. Struct. Biotechnol. J. 2013, 7, e201305003.

(345) Artner, D.; Oblak, A.; Ittig, S.; Garate, J. A.; Horvat, S.; Arrieumerlou, C.; Hofinger, A.; Oostenbrink, C.; Jerala, R.; Kosma, P.; et al. Conformationally Constrained Lipid a Mimetics for Exploration of Structural Basis of Tlr4/Md-2 Activation by Lipopolysaccharide. ACS Chem. Biol. 2013, 8, 2423-2432.

(346) Klett, J.; Reeves, J.; Oberhauser, N.; Pérez-Regidor, L.; Martín-Santamaria, S. Modulation of Toll-Like Receptor 4. Insights from X-Ray Crystallography and Molecular Modeling. Curr. Top. Med. Chem. 2014, 14, 2672-2683.

(347) de Aguiar, C.; Costa, M. G.; Verli, H. Dynamics on Human Toll-Like Receptor 4 Complexation to Md-2: The Coreceptor Stabilizing Function. Proteins 2015, 83, 373-382.

(348) Vašl, J.; Oblak, A.; Peternelj, T. T.; Klett, J.; Martín-Santamaría, S.; Gioannini, T. L.; Weiss, J. P.; Jerala, R. Molecular Basis of the Functional Differences between Soluble Human Versus Murine Md-2: Role of Val-135 in Transfer of Lipopolysaccharide from Cd14 to Md-2. J. Immunol. 2016, 1502074.

(349) Billod, J. M.; Lacetera, A.; Guzmán-Caldentey, J.; Martín-Santamaría, S. Computational Approaches to Toll-Like Receptor 4 Modulation. Molecules 2016, 21.

(350) Anwar, M. A.; Choi, S. Structure-Activity Relationship in Tlr4 Mutations: Atomistic Molecular Dynamics Simulations and Residue Interaction Network Analysis. Sci. Rep. 2017, 7, 43807.

(351) Tafazzol, A.; Duan, Y. Key Residues in Tlr4-Md2 Tetramer Formation Identified by Free Energy Simulations. PLoS Comp. Biol. 2019, 15, e1007228-e1007228.

(352) Matamoros-Recio, A.; Franco-Gonzalez, J. F.; Perez-Regidor, L.; Billod, J. M.; Guzman-Caldentey, J.; Martin-Santamaria, S. Full-Atom Model of the Agonist Lps-Bound Toll-Like Receptor 4 Dimer in a Membrane Environment. Chemistry 2021, 27, 15406-15425.

(353) 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.

(354) Zhao, H.; Taherzadeh, G.; Zhou, Y.; Yang, Y. Computational Prediction of Carbohydrate-Binding Proteins and Binding Sites. Curr. Protoc. Protein Sci. 2018, 94, e75.

(355) Tsai, K. C.; Jian, J. W.; Yang, E. W.; Hsu, P. C.; Peng, H. P.; Chen, C. T.; Chen, J. B.; Chang, J. Y.; Hsu, W. L.; Yang, A. S. Prediction of Carbohydrate Binding Sites on Protein Surfaces with 3-Dimensional Probability Density Distributions of Interacting Atoms. PLoS One 2012, 7, e40846.

(356) Banno, M.; Komiyama, Y.; Cao, W.; Oku, Y.; Ueki, K.; Sumikoshi, K.; Nakamura, S.; Terada, T.; Shimizu, K. Development of a Sugar-Binding Residue Prediction System from Protein Sequences Using Support Vector Machine. Comput. Biol. Chem. 2017, 66, 36-43.

(357) Beyer, T. A.; Sadler, J. E.; Rearick, J. I.; Paulson, J. C.; Hill, R. L. Glycosyltransferases and Their Use in Assessing Oligosaccharide Structure and Structure-Function Relationships. Adv. Enzymol. Relat. Areas Mol. Biol. 1981, 52, 23-175.

(358) Schachter, H. Enzymes Associated with Glycosylation. Curr. Opin. Struct. Biol. 1991, 1, 755-765.

(359) Kleene, R.; Berger, E. G. The Molecular and Cell Biology of Glycosyltransferases. Biochim. Biophys. Acta 1993, 1154, 283-325.

(360) Montreuil, J.; Vliegenthart, J. F. G.; Schachter, H. Glycoproteins. In New Comprehensive Biochemistry, Elsevier Science B.V. ed.; Neuberger, A., van Deenen, L. L. M., Eds.; Elsevier: Amsterdam, 1995; Vol. 29a, p 641.

(361) Bobovska, A.; Tvaroska, I.; Kona, J. A Theoretical Study on the Catalytic Mechanism of the Retaining Alpha-1,2-Mannosyltransferase Kre2p/Mnt1p: The Impact of Different Metal Ions on Catalysis. Org. Biomol. Chem. 2014, 12, 4201-4210.

(362) Tvaroska, I.; Kozmon, S.; Wimmerova, M.; Koca, J. A Qm/Mm Investigation of the Catalytic Mechanism of Metal-Ion-Independent Core 2 Beta1,6-N-Acetylglucosaminyltransferase. Chemistry 2013, 19, 8153-8162.

(363) Campbell, J. A.; Davies, G. J.; Bulone, V. V.; Henrissat, B. A Classification of Nucleotide-Diphospho-Sugar Glycosyltransferases Based on Amino Acid Sequence Similarities. Biochem. J. 1998, 329 (Pt 3), 719.

(364) Coutinho, P. M.; Deleury, E.; Davies, G. J.; Henrissat, B. An Evolving Hierarchical Family Classification for Glycosyltransferases. J. Mol. Biol. 2003, 328, 307-317.

(365) 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, D233-238.

(366) Singh, S.; Phillips, G. N., Jr.; Thorson, J. S. The Structural Biology of Enzymes Involved in Natural Product Glycosylation. Nat. Prod. Rep. 2012, 29, 1201-1237.

(367) Moremen, K. W.; Haltiwanger, R. S. Emerging Structural Insights into Glycosyltransferase-Mediated Synthesis of Glycans. Nat. Chem. Biol. 2019, 15, 853-864.

(368) Sinnott, M. L. Catalytic Mechanism of Enzymic Glycosyl Transfer. Chem. Rev. 1990, 90, 1171-1202.

(369) Zechel, D. L.; Withers, S. G. Glycosidase Mechanisms: Anatomy of a Finely Tuned Catalyst. Acc. Chem. Res. 2000, 33, 11-18.

(370) Breton, C.; Mucha, J.; Jeanneau, C. Structural and Functional Features of Glycosyltransferases. Biochimie 2001, 83, 713-718.

(371) Breton, C.; Snajdrova, L.; Jeanneau, C.; Koca, J.; Imberty, A. Structures and Mechanisms of Glycosyltransferases. Glycobiology 2006, 16, 29R-37R.

(372) Lairson, L. L.; Withers, S. G. Mechanistic Analogies Amongst Carbohydrate Modifying Enzymes. Chem. Commun. (Camb) 2004, 2243-2248.

(373) Henrissat, B.; Sulzenbacher, G.; Bourne, Y. Glycosyltransferases, Glycoside Hydrolases: Surprise, Surprise! Curr. Opin. Struct. Biol. 2008, 18, 527-533.

(374) Davies, G. J.; Sinnott, M. L.; Withers, S. G. Glycosyl Transfer. In Comprehensive Biological Catalysis : A Mechanistic Reference, Sinnott, M. L. Ed.; Academic Press Limited, 1998; pp 119-208.

(375) Zhang, R.; Yip, V. L. Y.; Withers, S. G. Mechanisms of Enzymatic Glycosyl Transfer. In Comprehensive Natural Products Ii : Chemistry and Biology, Mander, L. N., Liu, H.-w. Eds.; Elsevier, Oxford, 2010; pp 385-422.

(376) Wilson, I. B. H.; Breton, C.; Imberty, A.; Tvaroska, I. Molecular Basis for the Biosynthesis of Oligo- and Polysaccharides. In Glycoscience : Chemistry and Chemical Biology, Fraser-Reid, B. O., Tatsuta, K., Thiem, J. Eds.; Springer, 2008; pp 2267-2323.

(377) Tvaroska, I. Structural Insights into the Catalytic Mechanism and Transition State of Glycosyltransferases Using Ab Initio Molecular Modeling. Trends Glycosci. Glycotechnol. 2005, 17, 177-190.

(378) Zechel, D. L.; Withers, S. G. Glycosyl Transferase Mechanisms. In Comprehensive Natural Products Chemistry, Poulter, C. D. Ed.; Vol. 5; 1999; pp 279-314.

(379) Kozmon, S.; Tvaroska, I. Catalytic Mechanism of Glycosyltransferases: Hybrid Quantum Mechanical/Molecular Mechanical Study of the Inverting N-Acetylglucosaminyltransferase I. J. Am. Chem. Soc. 2006, 128, 16921-16927.

(380) Gomez, H.; Polyak, I.; Thiel, W.; Lluch, J. M.; Masgrau, L. Retaining Glycosyltransferase Mechanism Studied by Qm/Mm Methods: Lipopolysaccharyl-Alpha-1,4-Galactosyltransferase C Transfers Alpha-Galactose Via an Oxocarbenium Ion-Like Transition State. J. Am. Chem. Soc. 2012, 134, 4743-4752.

(381) Ardevol, A.; Rovira, C. The Molecular Mechanism of Enzymatic Glycosyl Transfer with Retention of Configuration: Evidence for a Short-Lived Oxocarbenium-Like Species. Angew. Chem. Int. Ed. Engl. 2011, 50, 10897-10901.

(382) Rojas-Cervellera, V.; Ardevol, A.; Boero, M.; Planas, A.; Rovira, C. Formation of a Covalent Glycosyl-Enzyme Species in a Retaining Glycosyltransferase. Chemistry 2013, 19, 14018-14023.

(383) Gómez, H.; Lluch, J. M.; Masgrau, L. Substrate-Assisted and Nucleophilically Assisted Catalysis in Bovine A1,3-Galactosyltransferase. Mechanistic Implications for Retaining Glycosyltransferases. J. Am. Chem. Soc. 2013, 135, 7053-7063.

(384) Ardevol, A.; Iglesias-Fernandez, J.; Rojas-Cervellera, V.; Rovira, C. The Reaction Mechanism of Retaining Glycosyltransferases. Biochem. Soc. Trans. 2016, 44, 51-60.

(385) Iglesias-Fernandez, J.; Hancock, S. M.; Lee, S. S.; Khan, M.; Kirkpatrick, J.; Oldham, N. J.; McAuley, K.; Fordham-Skelton, A.; Rovira, C.; Davis, B. G. A Front-Face 'Sni Synthase' Engineered from a Retaining 'Double-Sn2' Hydrolase. Nat. Chem. Biol. 2017, 13, 874-881.

(386) Trnka, T.; Kozmon, S.; Tvaroska, I.; Koca, J. Stepwise Catalytic Mechanism Via Short-Lived Intermediate Inferred from Combined Qm/Mm Merp and Pes Calculations on Retaining Glycosyltransferase Ppgalnact2. PLoS Comp. Biol. 2015, 11, 21.

(387) Mendoza, F.; Lluch, J. M.; Masgrau, L. Computational Insights into Active Site Shaping for Substrate Specificity and Reaction Regioselectivity in the Extl2 Retaining Glycosyltransferase. Org. Biomol. Chem. 2017, 15, 9095-9107.

(388) Albesa-Jove, D.; Mendoza, F.; Rodrigo-Unzueta, A.; Gomollon-Bel, F.; Cifuente, J. O.; Urresti, S.; Comino, N.; Gomez, H.; Romero-Garcia, J.; Lluch, J. M.; et al. A Native Ternary Complex Trapped in a Crystal Reveals the Catalytic Mechanism of a Retaining Glycosyltransferase. Angew. Chem. Int. Ed. Engl. 2015, 54, 9898-9902.

(389) 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.

(390) Krupicka, M.; Tvaroska, I. Hybrid Quantum Mechanical/Molecular Mechanical Investigation of the Beta-1,4-Galactosyltransferase-I Mechanism. J. Phys. Chem. B 2009, 113, 11314-11319.

(391) 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.

(392) Trnka, T.; Tvaroska, I.; Koca, J. Automated Training of Reaxff Reactive Force Fields for Energetics of Enzymatic Reactions. J. Chem. Theory Comput. 2018, 14, 291-302.

(393) Janos, P.; Trnka, T.; Kozmon, S.; Tvaroska, I.; Koca, J. Different Qm/Mm Approaches to Elucidate Enzymatic Reactions: Case Study on Ppgalnact2. J. Chem. Theory Comput. 2016, 12, 6062-6076.

(394) Mendoza, F.; Masgrau, L. Computational Modeling of Carbohydrate Processing Enzymes Reactions. Curr. Opin. Chem. Biol. 2021, 61, 203-213.

(395) Lairson, L. L.; Henrissat, B.; Davies, G. J.; Withers, S. G. Glycosyltransferases: Structures, Functions, and Mechanisms. Annu. Rev. Biochem. 2008, 77, 521-555.

(396) Albesa-Jové, D.; Guerin, M. E. The Conformational Plasticity of Glycosyltransferases. Curr. Opin. Struct. Biol. 2016, 40, 23-32.

(397) Ge, C.; Gomez-Llobregat, J.; Skwark, M. J.; Ruysschaert, J. M.; Wieslander, A.; Linden, M. Membrane Remodeling Capacity of a Vesicle-Inducing Glycosyltransferase. FEBS J. 2014, 281, 3667-3684.

(398) Romero-Garcia, J.; Francisco, C.; Biarnes, X.; Planas, A. Structure-Function Features of a Mycoplasma Glycolipid Synthase Derived from Structural Data Integration, Molecular Simulations, and Mutational Analysis. PLoS One 2013, 8, e81990.

(399) Parker, J. L.; Corey, R. A.; Stansfeld, P. J.; Newstead, S. Structural Basis for Substrate Specificity and Regulation of Nucleotide Sugar Transporters in the Lipid Bilayer. Nat. Commun. 2019, 10, 4657.

(400) Romero-Garcia, J.; Biarnes, X.; Planas, A. Essential Mycoplasma Glycolipid Synthase Adheres to the Cell Membrane by Means of an Amphipathic Helix. Sci. Rep. 2019, 9, 7085.

(401) Makshakova, O.; Breton, C.; Perez, S. Unraveling the Complex Enzymatic Machinery Making a Key Galactolipid in Chloroplast Membrane: A Multiscale Computer Simulation. Sci. Rep. 2020, 10, 13514.

(402) Nitenberg, M.; Makshakova, O.; Rocha, J.; Perez, S.; Marechal, E.; Block, M. A.; Girard-Egrot, A.; Breton, C. Mechanism of Activation of Plant Monogalactosyldiacylglycerol Synthase 1 (Mgd1) by Phosphatidylglycerol. Glycobiology 2020, 30, 396-406.

(403) The CAZypedia Consortium. Ten Years of Cazypedia: A Living Encyclopedia of Carbohydrate-Active Enzymes. Glycobiology 2017, 28, 3-8.

(404) Vasella, A.; Davies, G. J.; Böhm, M. Glycosidase Mechanisms. Curr. Opin. Chem. Biol. 2002, 6, 619-629.

(405) Vocadlo, D. J.; Davies, G. J. Mechanistic Insights into Glycosidase Chemistry. Curr. Opin. Chem. Biol. 2008, 12, 539-555.

(406) Legler, G. Glycoside Hydrolases: Mechanistic Information from Studies with Reversible and Irreversible Inhibitors. Adv. Carbohydr. Chem. Biochem. 1990, 48, 319-384.

(407) Heightman, T. D.; Vasella, A. T. Recent Insights into Inhibition, Structure, and Mechanism of Configuration-Retaining Glycosidases. Angew. Chem. Int. Ed. Engl. 1999, 38, 750-770.

(408) Rye, C. S.; Withers, S. G. Glycosidase Mechanisms. Curr. Opin. Chem. Biol. 2000, 4, 573-580.

(409) Blake, C. C.; Johnson, L. N.; Mair, G. A.; North, A. C.; Phillips, D. C.; Sarma, V. R. Crystallographic Studies of the Activity of Hen Egg-White Lysozyme. Proc. R. Soc. Lond. B Biol. Sci. 1967, 167, 378-388.

(410) Blake, C. C.; Koenig, D. F.; Mair, G. A.; North, A. C.; Phillips, D. C.; Sarma, V. R. Structure of Hen Egg-White Lysozyme. A Three-Dimensional Fourier Synthesis at 2 Angstrom Resolution. Nature 1965, 206, 757-761.

(411) Blake, C. C.; Mair, G. A.; North, A. C.; Phillips, D. C.; Sarma, V. R. On the Conformation of the Hen Egg-White Lysozyme Molecule. Proc. R. Soc. Lond. B Biol. Sci. 1967, 167, 365-377.

(412) Davies, G. J.; Mackenzie, L.; Varrot, A.; Dauter, M.; Brzozowski, A. M.; Schülein, M.; Withers, S. G. Snapshots Along an Enzymatic Reaction Coordinate: Analysis of a Retaining Beta-Glycoside Hydrolase. Biochemistry 1998, 37, 11707-11713.

(413) Davies, G. J.; Planas, A.; Rovira, C. Conformational Analyses of the Reaction Coordinate of Glycosidases. Acc. Chem. Res. 2012, 45, 308-316.

(414) Ducros, V. M.; Zechel, D. L.; Murshudov, G. N.; Gilbert, H. J.; Szabó, L.; Stoll, D.; Withers, S. G.; Davies, G. J. Substrate Distortion by a Beta-Mannanase: Snapshots of the Michaelis and Covalent-Intermediate Complexes Suggest a B(2,5) Conformation for the Transition State. Angew. Chem. Int. Ed. Engl. 2002, 41, 2824-2827.

(415) Sulzenbacher, G.; Driguez, H.; Henrissat, B.; Schülein, M.; Davies, G. J. Structure of the Fusarium Oxysporum Endoglucanase I with a Nonhydrolyzable Substrate Analogue: Substrate Distortion Gives Rise to the Preferred Axial Orientation for the Leaving Group. Biochemistry 1996, 35, 15280-15287.

(416) Tailford, L. E.; Offen, W. A.; Smith, N. L.; Dumon, C.; Morland, C.; Gratien, J.; Heck, M. P.; Stick, R. V.; Blériot, Y.; Vasella, A.; et al. Structural and Biochemical Evidence for a Boat-Like Transition State in Beta-Mannosidases. Nat. Chem. Biol. 2008, 4, 306-312.

(417) Tews, I.; Perrakis, A.; Oppenheim, A.; Dauter, Z.; Wilson, K. S.; Vorgias, C. E. Bacterial Chitobiase Structure Provides Insight into Catalytic Mechanism and the Basis of Tay-Sachs Disease. Nat. Struct. Biol. 1996, 3, 638-648.

(418) Thompson, A. J.; Dabin, J.; Iglesias-Fernández, J.; Ardèvol, A.; Dinev, Z.; Williams, S. J.; Bande, O.; Siriwardena, A.; Moreland, C.; Hu, T. C.; et al. The Reaction Coordinate of a Bacterial Gh47 A-Mannosidase: A Combined Quantum Mechanical and Structural Approach. Angew. Chem. Int. Ed. Engl. 2012, 51, 10997-11001.

(419) Ardèvol, A.; Biarnés, X.; Planas, A.; Rovira, C. The Conformational Free-Energy Landscape of B-D-Mannopyranose: Evidence for a (1)S(5) → B(2,5) → (O)S(2) Catalytic Itinerary in B-Mannosidases. J. Am. Chem. Soc. 2010, 132, 16058-16065.

(420) Biarnés, X.; Ardèvol, A.; Planas, A.; Rovira, C.; Laio, A.; Parrinello, M. The Conformational Free Energy Landscape of Beta-D-Glucopyranose. Implications for Substrate Preactivation in Beta-Glucoside Hydrolases. J. Am. Chem. Soc. 2007, 129, 10686-10693.

(421) Biarnés, X.; Nieto, J.; Planas, A.; Rovira, C. Substrate Distortion in the Michaelis Complex of Bacillus 1,3-1,4-Beta-Glucanase. Insight from First Principles Molecular Dynamics Simulations. J. Biol. Chem. 2006, 281, 1432-1441.

(422) Lammerts van Bueren, A.; Ardèvol, A.; Fayers-Kerr, J.; Luo, B.; Zhang, Y.; Sollogoub, M.; Blériot, Y.; Rovira, C.; Davies, G. J. Analysis of the Reaction Coordinate of Alpha-L-Fucosidases: A Combined Structural and Quantum Mechanical Approach. J. Am. Chem. Soc. 2010, 132, 1804-1806.

(423) Cremer, D.; Pople, J. A. General Definition of Ring Puckering Coordinates. J. Am. Chem. Soc. 1975, 97, 1354-1358.

(424) Spiwok, V.; Králová, B.; Tvaroska, I. Modelling of Beta-D-Glucopyranose Ring Distortion in Different Force Fields: A Metadynamics Study. Carbohydr. Res. 2010, 345, 530-537.

(425) Raich, L.; Nin-Hill, A.; Ardèvol, A.; Rovira, C. Enzymatic Cleavage of Glycosidic Bonds: Strategies on How to Set up and Control a Qm/Mm Metadynamics Simulation. Methods Enzymol. 2016, 577, 159-183.

(426) Barker, I. J.; Petersen, L.; Reilly, P. J. Mechanism of Xylobiose Hydrolysis by Gh43 B-Xylosidase. J. Phys. Chem. B 2010, 114, 15389-15393.

(427) Petersen, L.; Ardèvol, A.; Rovira, C.; Reilly, P. J. Mechanism of Cellulose Hydrolysis by Inverting Gh8 Endoglucanases: A Qm/Mm Metadynamics Study. J. Phys. Chem. B 2009, 113, 7331-7339.

(428) Morais, M. A. B.; Coines, J.; Domingues, M. N.; Pirolla, R. A. S.; Tonoli, C. C. C.; Santos, C. R.; Correa, J. B. L.; Gozzo, F. C.; Rovira, C.; Murakami, M. T. Two Distinct Catalytic Pathways for Gh43 Xylanolytic Enzymes Unveiled by X-Ray and Qm/Mm Simulations. Nat. Commun. 2021, 12.

(429) Meelua, W.; Wanjai, T.; Thinkumrob, N.; Oláh, J.; Mujika, J. I.; Ketudat-Cairns, J. R.; Hannongbua, S.; Jitonnom, J. Active Site Dynamics and Catalytic Mechanism in Arabinan Hydrolysis Catalyzed by Gh43 Endo-Arabinanase from Qm/Mm Molecular Dynamics Simulation and Potential Energy Surface. J. Biomol. Struct. Dyn. 2021, 1-11.

(430) Barnes, J. A.; Williams, I. H. Quantum Mechanical/Molecular Mechanical Approaches to Transition State Structure: Mechanism of Sialidase Action. Biochem. Soc. Trans. 1996, 24, 263-268.

(431) Biarnés, X.; Ardèvol, A.; Iglesias-Fernández, J.; Planas, A.; Rovira, C. Catalytic Itinerary in 1,3-1,4-B-Glucanase Unraveled by Qm/Mm Metadynamics. Charge Is Not yet Fully Developed at the Oxocarbenium Ion-Like Transition State. J. Am. Chem. Soc. 2011, 133, 20301-20309.

(432) Brameld, K. A.; Shrader, W. D.; Imperiali, B.; Goddard, W. A., 3rd. Substrate Assistance in the Mechanism of Family 18 Chitinases: Theoretical Studies of Potential Intermediates and Inhibitors. J. Mol. Biol. 1998, 280, 913-923.

(433) Brás, N. F.; Fernandes, P. A.; Ramos, M. J. Qm/Mm Studies on the B-Galactosidase Catalytic Mechanism: Hydrolysis and Transglycosylation Reactions. J. Chem. Theory Comput. 2010, 6, 421-433.

(434) Brás, N. F.; Moura-Tamames, S. A.; Fernandes, P. A.; Ramos, M. J. Mechanistic Studies on the Formation of Glycosidase-Substrate and Glycosidase-Inhibitor Covalent Intermediates. J. Comput. Chem. 2008, 29, 2565-2574.

(435) Brás, N. F.; Ramos, M. J.; Fernandes, P. A. Dft Studies on the B-Glycosidase Catalytic Mechanism: The Deglycosylation Step. J. Mol. Struct. THEOCHEM 2010, 946, 125-133.

(436) Jitonnom, J.; Lee, V. S.; Nimmanpipug, P.; Rowlands, H. A.; Mulholland, A. J. Quantum Mechanics/Molecular Mechanics Modeling of Substrate-Assisted Catalysis in Family 18 Chitinases: Conformational Changes and the Role of Asp142 in Catalysis in Chib. Biochemistry 2011, 50, 4697-4711.

(437) Juers, D. H.; Heightman, T. D.; Vasella, A.; McCarter, J. D.; Mackenzie, L.; Withers, S. G.; Matthews, B. W. A Structural View of the Action of Escherichia Coli (Lacz) Beta-Galactosidase. Biochemistry 2001, 40, 14781-14794.

(438) Passos, Ó.; Fernandes, P. A.; Ramos, M. J. Theoretical Insights into the Catalytic Mechanism of B-Hexosaminidase. Theor. Chem. Acc. 2011, 129, 119-129.

(439) Taylor, N. R.; von Itzstein, M. Molecular Modeling Studies on Ligand Binding to Sialidase from Influenza Virus and the Mechanism of Catalysis. J. Med. Chem. 1994, 37, 616-624.

(440) Nin-Hill, A.; Rovira, C. The Catalytic Reaction Mechanism of the B-Galactocerebrosidase Enzyme Deficient in Krabbe Disease. ACS Catal. 2020, 10, 12091-12097.

(441) Knott, B. C.; Haddad Momeni, M.; Crowley, M. F.; Mackenzie, L. F.; Götz, A. W.; Sandgren, M.; Withers, S. G.; Ståhlberg, J.; Beckham, G. T. The Mechanism of Cellulose Hydrolysis by a Two-Step, Retaining Cellobiohydrolase Elucidated by Structural and Transition Path Sampling Studies. J. Am. Chem. Soc. 2014, 136, 321-329.

(442) Nomura, T.; Iwase, H.; Saka, N.; Takahashi, N.; Mikami, B.; Mizutani, K. High-Resolution Crystal Structures of the Glycoside Hydrolase Family 45 Endoglucanase Eg27ii from the Snail Ampullaria Crossean. Acta Crystallogr. D 2019, 75, 426-436.

(443) Matsuzawa, T.; Jo, T.; Uchiyama, T.; Manninen, J. A.; Arakawa, T.; Miyazaki, K.; Fushinobu, S.; Yaoi, K. Crystal Structure and Identification of a Key Amino Acid for Glucose Tolerance, Substrate Specificity, and Transglycosylation Activity of Metagenomic B-Glucosidase Td2f2. FEBS J. 2016, 283, 2340-2353.

(444) Borisova, A. S.; Eneyskaya, E. V.; Jana, S.; Badino, S. F.; Kari, J.; Amore, A.; Karlsson, M.; Hansson, H.; Sandgren, M.; Himmel, M. E.; et al. Correlation of Structure, Function and Protein Dynamics in Gh7 Cellobiohydrolases from Trichoderma Atroviride, T. Reesei and T. Harzianum. Biotechnol. Biofuels 2018, 11, 5.

(445) Fowler, C. A.; Sabbadin, F.; Ciano, L.; Hemsworth, G. R.; Elias, L.; Bruce, N.; McQueen-Mason, S.; Davies, G. J.; Walton, P. H. Discovery, Activity and Characterisation of an Aa10 Lytic Polysaccharide Oxygenase from the Shipworm Symbiont Teredinibacter Turnerae. Biotechnol. Biofuels 2019, 12, 232.

(446) Vaaje-Kolstad, G.; Westereng, B.; Horn, S. J.; Liu, Z.; Zhai, H.; Sørlie, M.; Eijsink, V. G. An Oxidative Enzyme Boosting the Enzymatic Conversion of Recalcitrant Polysaccharides. Science 2010, 330, 219-222.

(447) Wang, B.; Johnston, E. M.; Li, P.; Shaik, S.; Davies, G. J.; Walton, P. H.; Rovira, C. Qm/Mm Studies into the H2o2-Dependent Activity of Lytic Polysaccharide Monooxygenases: Evidence for the Formation of a Caged Hydroxyl Radical Intermediate. ACS Catal. 2018, 8, 1346-1351.

(448) Wang, B.; Walton, P. H.; Rovira, C. Molecular Mechanisms of Oxygen Activation and Hydrogen Peroxide Formation in Lytic Polysaccharide Monooxygenases. ACS Catal. 2019, 9, 4958-4969.

(449) Poma, A. B.; Chwastyk, M.; Cieplak, M. Polysaccharide–Protein Complexes in a Coarse-Grained Model. J. Phys. Chem. B 2015, 119, 12028-12041.

(450) Wohlert, J.; Berglund, L. A. A Coarse-Grained Model for Molecular Dynamics Simulations of Native Cellulose. J. Chem. Theory Comput. 2011, 7, 753-760.

(451) Asztalos, A.; Daniels, M.; Sethi, A.; Shen, T.; Langan, P.; Redondo, A.; Gnanakaran, S. A Coarse-Grained Model for Synergistic Action of Multiple Enzymes on Cellulose. Biotechnol. Biofuels 2012, 5, 55.

(452) Bu, L.; Beckham, G. T.; Crowley, M. F.; Chang, C. H.; Matthews, J. F.; Bomble, Y. J.; Adney, W. S.; Himmel, M. E.; Nimlos, M. R. The Energy Landscape for the Interaction of the Family 1 Carbohydrate-Binding Module and the Cellulose Surface Is Altered by Hydrolyzed Glycosidic Bonds. J. Phys. Chem. B 2009, 113, 10994-11002.

(453) Reyes, G.; Aguayo, M. G.; Fernández Pérez, A.; Pääkkönen, T.; Gacitúa, W.; Rojas, O. J. Dissolution and Hydrolysis of Bleached Kraft Pulp Using Ionic Liquids. Polymers (Basel) 2019, 11, 673.

(454) Drew, D.; North, R. A.; Nagarathinam, K.; Tanabe, M. Structures and General Transport Mechanisms by the Major Facilitator Superfamily (Mfs). Chem. Rev. 2021, 121, 5289-5335.

(455) Furuta, T. Structural Dynamics of Abc Transporters: Molecular Simulation Studies. Biochem. Soc. Trans. 2021, 49, 405-414.

(456) Park, M. S. Molecular Dynamics Simulations of the Human Glucose Transporter Glut1. PLoS One 2015, 10, e0125361.

(457) Sun, M.; Zheng, Q. Key Factors in Conformation Transformation of an Important Neuronic Protein Glucose Transport 3 Revealed by Molecular Dynamics Simulation. ACS Chem. Neurosci. 2019, 10, 4444-4448.

(458) Selvam, B.; Yu, Y.-C.; Chen, L.-Q.; Shukla, D. Molecular Basis of the Glucose Transport Mechanism in Plants. ACS Cent. Sci. 2019, 5, 1085-1096.

(459) Galochkina, T.; Ng Fuk Chong, M.; Challali, L.; Abbar, S.; Etchebest, C. New Insights into Glut1 Mechanics During Glucose Transfer. Sci. Rep. 2019, 9, 998.

(460) Schirmer, T.; Keller, T. A.; Wang, Y. F.; Rosenbusch, J. P. Structural Basis for Sugar Translocation through Maltoporin Channels at 3.1 a Resolution. Science 1995, 267, 512-514.

(461) Dutzler, R.; Schirmer, T.; Karplus, M.; Fischer, S. Translocation Mechanism of Long Sugar Chains across the Maltoporin Membrane Channel. Structure 2002, 10, 1273-1284.

(462) Hsu, W.-L.; Furuta, T.; Sakurai, M. Analysis of the Free Energy Landscapes for the Opening–Closing Dynamics of the Maltose Transporter Atpase Malk2 Using Enhanced-Sampling Molecular Dynamics Simulation. J. Phys. Chem. B 2015, 119, 9717-9725.

(463) Hsu, W.-L.; Furuta, T.; Sakurai, M. The Mechanism of Nucleotide-Binding Domain Dimerization in the Intact Maltose Transporter as Studied by All-Atom Molecular Dynamics Simulations. Proteins: Struct. Funct. Bioinform. 2018, 86, 237-247.

(464) Lis, H.; Sharon, N. Lectins: Carbohydrate-Specific Proteins That Mediate Cellular Recognition. Chem. Rev. 1998, 98, 637-674.

(465) Kaltner, H.; Gabius, H. J. Animal Lectins: From Initial Description to Elaborated Structural and Functional Classification. Adv. Exp. Med. Biol. 2001, 491, 79-94.

(466) Fasting, C.; Schalley, C. A.; Weber, M.; Seitz, O.; Hecht, S.; Koksch, B.; Dernedde, J.; Graf, C.; Knapp, E. W.; Haag, R. Multivalency as a Chemical Organization and Action Principle. Angew. Chem. Int. Ed. 2012, 51, 10472-10498.

(467) Unno, H.; Goda, S.; Hatakeyama, T. Hemolytic Lectin Cel-Iii Heptamerizes Via a Large Structural Transition from A-Helices to a B-Barrel During the Transmembrane Pore Formation Process. J. Biol. Chem. 2014, 289, 12805-12812.

(468) Hester, G.; Kaku, H.; Goldstein, I. J.; Wright, C. S. Structure of Mannose-Specific Snowdrop (Galanthus Nivalis) Lectin Is Representative of a New Plant Lectin Family. Nat. Struct. Biol. 1995, 2, 472-479.

(469) Walker, J. R.; Nagar, B.; Young, N. M.; Hirama, T.; Rini, J. M. X-Ray Crystal Structure of a Galactose-Specific C-Type Lectin Possessing a Novel Decameric Quaternary Structure. Biochemistry 2004, 43, 3783-3792.

(470) Sommer, R.; Makshakova, O. N.; Wohlschlager, T.; Hutin, S.; Marsh, M.; Titz, A.; Künzler, M.; Varrot, A. Crystal Structures of Fungal Tectonin in Complex with O-Methylated Glycans Suggest Key Role in Innate Immune Defense. Structure 2018, 26, 391-402.e4.

(471) Mir, S.; Alhroub, Y.; Anyango, S.; Armstrong, D. R.; Berrisford, J. M.; Clark, A. R.; Conroy, M. J.; Dana, J. M.; Deshpande, M.; Gupta, D.; et al. Pdbe: Towards Reusable Data Delivery Infrastructure at Protein Data Bank in Europe. Nucleic Acids Res. 2017, 46, D486 - D492.

(472) 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.

(473) Mohan, A.; Oldfield, C. J.; Radivojac, P.; Vacic, V.; Cortese, M. S.; Dunker, A. K.; Uversky, V. N. Analysis of Molecular Recognition Features (Morfs). J. Mol. Biol. 2006, 362, 1043-1059.

(474) Topin, J.; Lelimousin, M.; Arnaud, J.; Audfray, A.; Perez, S.; Varrot, A.; Imberty, A. The Hidden Conformation of Lewis X, a Human Histo-Blood Group Antigen, Is a Determinant for Recognition by Pathogen Lectins. ACS Chem. Biol. 2016, 11, 2011-2020.

(475) Azurmendi, H. F.; Martin-Pastor, M.; Bush, C. A. Conformational Studies of Lewis X and Lewis a Trisaccharides Using NMR Residual Dipolar Couplings. Biopolymers 2002, 63, 89-98.

(476) Pérez, S.; Mouhous-Riou, N.; Nifant'ev, N. E.; Tsvetkov, Y. E.; Bachet, B.; Imberty, A. Crystal and Molecular Structure of a Histo-Blood Group Antigen Involved in Cell Adhesion: The Lewis X Trisaccharide. Glycobiology 1996, 6, 537-542.

(477) Torrie, G. M.; Valleau, J. P. Nonphysical Sampling Distributions in Monte Carlo Free-Energy Estimation: Umbrella Sampling. J. Comput. Phys. 1977, 23, 187-199.

(478) Pérez, 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.

(479) Alibay, I.; Bryce, R. A. Ring Puckering Landscapes of Glycosaminoglycan-Related Monosaccharides from Molecular Dynamics Simulations. J. Chem. Inf. Model. 2019, 59, 4729-4741.

(480) Ikebe, J.; Umezawa, K.; Higo, J. Enhanced Sampling Simulations to Construct Free-Energy Landscape of Protein-Partner Substrate Interaction. Biophys. Rev. 2016, 8, 45-62.

(481) Roy, R.; Ghosh, B.; Kar, P. Investigating Conformational Dynamics of Lewis Y Oligosaccharides and Elucidating Blood Group Dependency of Cholera Using Molecular Dynamics. ACS Omega 2020, 5, 3932-3942.

(482) 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.

(483) Hamelberg, D.; Mongan, J.; McCammon, J. A. Accelerated Molecular Dynamics: A Promising and Efficient Simulation Method for Biomolecules. J. Chem. Phys. 2004, 120, 11919-11929.

(484) Soliman, C.; Pier, G. B.; Ramsland, P. A. Antibody Recognition of Bacterial Surfaces and Extracellular Polysaccharides. Curr. Opin. Struct. Biol. 2020, 62, 48-55.

(485) Soliman, C.; Yuriev, E.; Ramsland, P. A. Antibody Recognition of Aberrant Glycosylation on the Surface of Cancer Cells. Curr. Opin. Struct. Biol. 2017, 44, 1-8.

(486) Gómez-Redondo, M.; Ardá, A.; Gimeno, A.; Jiménez-Barbero, J. Bacterial Polysaccharides: Conformation, Dynamics and Molecular Recognition by Antibodies. Drug Discov. Today Technol. 2020, 35-36, 1-11.

(487) Agostino, M.; Jene, C.; Boyle, T.; Ramsland, P. A.; Yuriev, E. Molecular Docking of Carbohydrate Ligands to Antibodies: Structural Validation against Crystal Structures. J. Chem. Inf. Model. 2009, 49, 2749-2760.

(488) Amon, R.; Grant, O. C.; Leviatan Ben-Arye, S.; Makeneni, S.; Nivedha, A. K.; Marshanski, T.; Norn, C.; Yu, H.; Glushka, J. N.; Fleishman, S. J.; et al. A Combined Computational-Experimental Approach to Define the Structural Origin of Antibody Recognition of Sialyl-Tn, a Tumor-Associated Carbohydrate Antigen. Sci. Rep. 2018, 8, 10786.

(489) Snapper, C. M. Differential Regulation of Polysaccharide-Specific Antibody Responses to Isolated Polysaccharides, Conjugate Vaccines, and Intact Gram-Positive Versus Gram-Negative Extracellular Bacteria. Vaccine 2016, 34, 3542-3548.

(490) 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, Prasad, A. K. Ed.; Acs Symposium Series, Vol. 1290; American Chemical Society, 2018; pp 139-173.

(491) Clément, M. J.; Imberty, A.; Phalipon, A.; Pérez, S.; Simenel, C.; Mulard, L. A.; Delepierre, M. Conformational Studies of the O-Specific Polysaccharide of Shigella Flexneri 5a and of Four Related Synthetic Pentasaccharide Fragments Using NMR and Molecular Modeling. J. Biol. Chem. 2003, 278, 47928-47936.

(492) Theillet, F. X.; Simenel, C.; Guerreiro, C.; Phalipon, A.; Mulard, L. A.; Delepierre, M. Effects of Backbone Substitutions on the Conformational Behavior of Shigella Flexneri O-Antigens: Implications for Vaccine Strategy. Glycobiology 2011, 21, 109-121.

(493) Hlozek, J.; Kuttel, M. M.; Ravenscroft, N. Conformations of Neisseria Meningitidis Serogroup a and X Polysaccharides: The Effects of Chain Length and O-Acetylation. Carbohydr. Res. 2018, 465, 44-51.

(494) Vulliez-Le Normand, B.; Saul, F. A.; Phalipon, A.; Bélot, F.; Guerreiro, C.; Mulard, L. A.; Bentley, G. A. Structures of Synthetic O-Antigen Fragments from Serotype 2a Shigella in Complex with a Protective Monoclonal Antibody. Proc. Natl. Acad. Sci. USA 2008, 105, 9976.

(495) Hlozek, J.; Owen, S.; Ravenscroft, N.; Kuttel, M. M. Molecular Modeling of the Shigella Flexneri Serogroup 3 and 5 O-Antigens and Conformational Relationships for a Vaccine Containing Serotypes 2a and 3a. Vaccines (Basel) 2020, 8, 643.

(496) Imberty, A.; Mikros, E.; Koca, J.; Mollicone, R.; Oriol, R.; Pérez, S. Computer Simulation of Histo-Blood Group Oligosaccharides: Energy Maps of All Constituting Disaccharides and Potential Energy Surfaces of 14 Abh and Lewis Carbohydrate Antigens. Glycoconj. J. 1995, 12, 331-349.

(497) Imberty, A.; Mollicone, R.; Mikros, E.; Carrupt, P. A.; Pérez, S.; Oriol, R. How Do Antibodies and Lectins Recognize Histo-Blood Group Antigens? A 3d-Qsar Study by Comparative Molecular Field Analysis (Comfa). Bioorg. Med. Chem. 1996, 4, 1979-1988.

(498) Brisson, J. R.; Baumann, H.; Imberty, A.; Pérez, S.; Jennings, H. J. Helical Epitope of the Group B Meningococcal Alpha(2-8)-Linked Sialic Acid Polysaccharide. Biochemistry 1992, 31, 4996-5004.

(499) Rodgers, K. D.; San Antonio, J. D.; Jacenko, O. Heparan Sulfate Proteoglycans: A Gaggle of Skeletal-Hematopoietic Regulators. Dev. Dyn. 2008, 237, 2622-2642.

(500) Nikitovic, D.; Pérez, S. Preface for the Special Issue on the Exploration of the Multifaceted Roles of Glycosaminoglycans: Gags. Biomolecules 2021, 11, 1630.

(501) 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.

(502) Whitmore, E. K.; Martin, D.; Guvench, O. Constructing 3-Dimensional Atomic-Resolution Models of Nonsulfated Glycosaminoglycans with Arbitrary Lengths Using Conformations from Molecular Dynamics. Int. J. Mol. Sci. 2020, 21, 7699.

(503) 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.

(504) Perez, S.; Rivet, A. Polys Glycan Builder: An Online Application for Intuitive Construction of 3d Structure of Complex Carbohydrates. In Glycoinformatics : Methods and Protocols Second Edition Series, 2 ed.; Lutteke, T. Ed.; Methods in Molecular Biology Springer, 2021.

(505) Pérez, S.; Bonnardel, F.; Lisacek, F.; Imberty, A.; Ricard Blum, S.; Makshakova, O. Gag-Db, the New Interface of the Three-Dimensional Landscape of Glycosaminoglycans. Biomolecules 2020, 10, 1660.

(506) 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.

(507) Nagarajan, B.; Sankaranarayanan, N. V.; Desai, U. R. Perspective on Computational Simulations of Glycosaminoglycans. Wiley Interdiscip. Rev. Comput. Mol. Sci. 2019, 9, e1388.

(508) 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.

(509) Imberty, A.; Lortat-Jacob, H.; Pérez, S. Structural View of Glycosaminoglycan-Protein Interactions. Carbohydr. Res. 2007, 342, 430-439.

(510) Vallet, S. D.; Clerc, O.; Ricard-Blum, S. Glycosaminoglycan-Protein Interactions: The First Draft of the Glycosaminoglycan Interactome. J. Histochem. Cytochem. 2021, 69, 93-104.

(511) Kjellén, L.; Lindahl, U. Specificity of Glycosaminoglycan-Protein Interactions. Curr. Opin. Struct. Biol. 2018, 50, 101-108.

(512) Paiardi, G.; Milanesi, M.; Wade, R. C.; D'Ursi, P.; Rusnati, M. A Bittersweet Computational Journey among Glycosaminoglycans. Biomolecules 2021, 11, 739.

(513) Paiardi, G.; Richter, S.; Oreste, P.; Urbinati, C.; Rusnati, M.; Wade, R. C. The Binding of Heparin to Spike Glycoprotein Inhibits SARS-CoV-2 Infection by Three Mechanisms. J. Biol. Chem. 2021, 101507.

(514) Agostino, M.; Gandhi, N. S.; Mancera, R. L. Development and Application of Site Mapping Methods for the Design of Glycosaminoglycans. Glycobiology 2014, 24, 840-851.

(515) Morris, G. M.; Goodsell, D. S.; Halliday, R. S.; Huey, R.; Hart, W. E.; Belew, R. K.; Olson, A. J. Automated Docking Using a Lamarckian Genetic Algorithm and an Empirical Binding Free Energy Function. J. Comput. Chem. 1998, 19, 1639-1662.

(516) Morris, G. M.; Huey, R.; Lindstrom, W.; Sanner, M. F.; Belew, R. K.; Goodsell, D. S.; Olson, A. J. Autodock4 and Autodocktools4: Automated Docking with Selective Receptor Flexibility. J. Comput. Chem. 2009, 30, 2785-2791.

(517) Trott, O.; Olson, A. J. Autodock Vina: Improving the Speed and Accuracy of Docking with a New Scoring Function, Efficient Optimization, and Multithreading. J. Comput. Chem. 2010, 31, 455-461.

(518) Hofmann, T.; Samsonov, S. A.; Pichert, A.; Lemmnitzer, K.; Schiller, J.; Huster, D.; Pisabarro, M. T.; von Bergen, M.; Kalkhof, S. Structural Analysis of the Interleukin-8/Glycosaminoglycan Interactions by Amide Hydrogen/Deuterium Exchange Mass Spectrometry. Methods 2015, 89, 45-53.

(519) Pichert, A.; Samsonov, S. A.; Theisgen, S.; Thomas, L.; Baumann, L.; Schiller, J.; Beck-Sickinger, A. G.; Huster, D.; Pisabarro, M. T. Characterization of the Interaction of Interleukin-8 with Hyaluronan, Chondroitin Sulfate, Dermatan Sulfate and Their Sulfated Derivatives by Spectroscopy and Molecular Modeling. Glycobiology 2012, 22, 134-145.

(520) Hintze, V.; Samsonov, S. A.; Anselmi, M.; Moeller, S.; Becher, J.; Schnabelrauch, M.; Scharnweber, D.; Pisabarro, M. T. Sulfated Glycosaminoglycans Exploit the Conformational Plasticity of Bone Morphogenetic Protein-2 (Bmp-2) and Alter the Interaction Profile with Its Receptor. Biomacromolecules 2014, 15, 3083-3092.

(521) Agostino, M.; Mancera, R. L.; Ramsland, P. A.; Yuriev, E. Automap: A Tool for Analyzing Protein–Ligand Recognition Using Multiple Ligand Binding Modes. J. Mol. Graphics Model. 2013, 40, 80-90.

(522) Taylor, J. S.; Burnett, R. M. Darwin: A Program for Docking Flexible Molecules. Proteins 2000, 41, 173-191.

(523) 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.

(524) Brenke, R.; Kozakov, D.; Chuang, G. Y.; Beglov, D.; Hall, D.; Landon, M. R.; Mattos, C.; Vajda, S. Fragment-Based Identification of Druggable 'Hot Spots' of Proteins Using Fourier Domain Correlation Techniques. Bioinformatics 2009, 25, 621-627.

(525) Imai, T.; Oda, K.; Kovalenko, A.; Hirata, F.; Kidera, A. Ligand Mapping on Protein Surfaces by the 3d-Rism Theory: Toward Computational Fragment-Based Drug Design. J. Am. Chem. Soc. 2009, 131, 12430-12440.

(526) Sadjad, B.; Zsoldos, Z. Toward a Robust Search Method for the Protein-Drug Docking Problem. IEEE ACM Trans. Comput. Biol. Bioinform. 2011, 8, 1120-1133.

(527) Babik, S.; Samsonov, S. A.; Pisabarro, M. T. Computational Drill Down on Fgf1-Heparin Interactions through Methodological Evaluation. Glycoconj. J. 2017, 34, 427-440.

(528) Samsonov, S. A.; Zacharias, M.; Chauvot de Beauchene, I. Modeling Large Protein–Glycosaminoglycan Complexes Using a Fragment-Based Approach. J. Comput. Chem. 2019, 40, 1429-1439.

(529) Uciechowska-Kaczmarzyk, U.; Chauvot de Beauchene, I.; Samsonov, S. A. Docking Software Performance in Protein-Glycosaminoglycan Systems. J. Mol. Graph. Model. 2019, 90, 42-50.

(530) Park, S. J.; Lee, J.; Qi, Y.; 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.

(531) Lemmin, T.; Soto, C. Glycosylator: A Python Framework for the Rapid Modeling of Glycans. BMC Bioinform. 2019, 20, 513.

(532) 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.

(533) 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 2020, 88, 973-985.

(534) 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.

(535) 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.

(536) Frank, M.; Bohne-Lang, A.; Wetter, T.; Lieth, C. W. Rapid Generation of a Representative Ensemble of N-Glycan Conformations. In Silico Biol. 2002, 2, 427-439.

(537) 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.

(538) Danne, R.; Poojari, C.; Martinez-Seara, H.; Rissanen, S.; Lolicato, F.; Rog, 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.

(539) Bohne, A. Pdb2multigif: A Web Tool to Create Animated Images of Molecules. J. Mol. Model. 1998, 4, 344-346.

(540) Bohne-Lang, A.; von der Lieth, C.-W. Glyprot: In Silico Glycosylation of Proteins. Nucleic Acids Res. 2005, 33, W214-W219.

(541) Liebschner, D.; Afonine, P. V.; Baker, M. L.; Bunkóczi, 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. D 2019, 75, 861-877.

(542) 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.

(543) Grant, O. C.; Smith, H. M. K.; 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.

(544) 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.

(545) 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.

(546) Jo, S.; Kim, T.; Im, W. Automated Builder and Database of Protein/Membrane Complexes for Molecular Dynamics Simulations. PLoS One 2007, 2, e880.

(547) 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.

(548) 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.

(549) 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.

(550) 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.

(551) Engelsen, S. B.; Cros, S.; Mackie, W.; Pérez, S. A Molecular Builder for Carbohydrates: Application to Polysaccharides and Complex Carbohydrates. Biopolymers 1996, 39, 417-433.

(552) Engelsen, 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.

(553) 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.

(554) Rose, A. S.; Bradley, A. R.; Valasatava, Y.; Duarte, J. M.; Prlic, A.; Rose, P. W. Ngl Viewer: Web-Based Molecular Graphics for Large Complexes. Bioinformatics 2018, 34, 3755-3758.

(555) Rose, A. S.; Hildebrand, P. W. Ngl Viewer: A Web Application for Molecular Visualization. Nucleic Acids Res. 2015, 43, W576-W579.

(556) Sayle, R. A.; Milner-White, E. J. Rasmol: Biomolecular Graphics for All. Trends Biochem. Sci. 1995, 20, 374.

(557) Sehnal, D.; Bittrich, S.; Deshpande, M.; Svobodová, R.; Berka, K.; Bazgier, V.; Velankar, S.; Burley, S. K.; Koča, J.; Rose, A. S. Mol* Viewer: Modern Web App for 3d Visualization and Analysis of Large Biomolecular Structures. Nucleic Acids Res. 2021, 49, W431-W437.

(558) 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.

(559) Willighagen, E.; Howard, M. Fast and Scriptable Molecular Graphics in Web Browsers without Java3d. Nature Precedings 2007.

(560) 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 (San Diego, Calif.) 2020, 173, 94-104.

(561) UniProt Consortium. Uniprot: A Worldwide Hub of Protein Knowledge. Nucleic Acids Res. 2019, 47, D506-D515.

(562) Cross, S.; Kuttel, M. M.; Stone, J. E.; Gain, J. E. Visualisation of Cyclic and Multi-Branched Molecules with Vmd. J. Mol. Graph. Model. 2009, 28, 131-139.

(563) Kuttel, M.; Gain, J.; Burger, A.; Eborn, I. Techniques for Visualization of Carbohydrate Molecules. J. Mol. Graph. Model. 2006, 25, 380-388.

(564) McNicholas, S.; Agirre, J. Glycoblocks: A Schematic Three-Dimensional Representation for Glycans and Their Interactions. Acta Crystallogr. D 2017, 73, 187-194.

(565) McNicholas, S.; Potterton, E.; Wilson, K. S.; Noble, M. E. Presenting Your Structures: The Ccp4mg Molecular-Graphics Software. Acta Crystallographica Section D: Structural Biology 2011, 67, 386-394.

(566) 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.

(567) Sehnal, D.; Grant, O. C. Rapidly Display Glycan Symbols in 3d Structures: 3d-Snfg in Litemol. J. Proteome Res. 2019, 18, 770-774.

(568) 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.

(569) Eborn, I.; Burger, A.; Kuttel, M.; Gain, J. Carbohydra: Rendering Carbohydrate Cartoons; University of Cape Town, 2004.

(570) Besançon, C.; Wong, H.; M Rao, R.; Dauchez, M.; Belloy, N.; Jonquet-Prevoteau, J.; Baud, S. Improved Umbrella Visualization Implemented in Unitymol Gives Valuable Insight on Sugar/Protein Interplay. In Workshop on Molecular Graphics and Visual Analysis of Molecular Data, Byška, J., Krone, M., Sommer, B. Eds.; The Eurographics Association, 2020.

(571) 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.

(572) 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.

(573) 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.

(574) Rojas-Macias, M. A.; Lütteke, T. Statistical Analysis of Amino Acids in the Vicinity of Carbohydrate Residues Performed by Glyvicinity. Methods Mol. Biol. 2015, 1273, 215-226.

(575) Brünger, 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. D 1998, 54, 905-921.

(576) Brunger, A. T. Version 1.2 of the Crystallography and NMR System. Nat. Protoc. 2007, 2, 2728-2733.

(577) Emsley, P.; Brunger, A. T.; Lütteke, T. Tools to Assist Determination and Validation of Carbohydrate 3d Structure Data. Methods Mol. Biol. 2015, 1273, 229-240.

(578) Feng, Y. Compatible Topologies and Parameters for NMR Structure Determination of Carbohydrates by Simulated Annealing. PLoS One 2017, 12, e0189700.

(579) 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 Bioinform. 2004, 5, 69.

(580) Woods, R. J. 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; 2019; pp 82–83

(581) Yamada, I.; Angata, K.; Watanabe, Y.; Ono, T. Databases for Glycoconjugates (Glycosmos Glycoproteins and Glycolipids, Glycoprotdb, Glyconavi:Tcarp, Glycopost). Glycoforum, 2020; Vol. 23, A22.

(582) van Beusekom, B.; Lütteke, T.; Joosten, R. P. Making Glycoproteins a Little Bit Sweeter with Pdb-Redo. Acta. Crystallogr. F Struct. Biol. Commun. 2018, 74, 463-472.

(583) 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. D 2019, 75, 416-425.

(584) 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, 1069-1078.

(585) Emsley, P.; Crispin, M. Structural Analysis of Glycoproteins: Building N-Linked Glycans with Coot. Acta Crystallogr. D 2018, 74, 256-263.

(586) Frenz, B.; Rämisch, S.; Borst, A. J.; Walls, A. C.; Adolf-Bryfogle, J.; Schief, W. R.; Veesler, D.; DiMaio, F. Automatically Fixing Errors in Glycoprotein Structures with Rosetta. Structure 2019, 27, 134-139.e3.

(587) Agirre, J.; Iglesias-Fernández, 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.

(588) Bagdonas, H.; Ungar, D.; Agirre, J. Leveraging Glycomics Data in Glycoprotein 3d Structure Validation with Privateer. Beilstein J. Org. Chem. 2020, 16, 2523-2533.

(589) Bagdonas, H.; Fogarty, C. A.; Fadda, E.; Agirre, J. The Case for Post-Predictional Modifications in the Alphafold Protein Structure Database. Nat. Struct. Mol. Biol. 2021, 28, 869-870.

(590) 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.

(591) 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.

(592) Pérez, S.; de Sanctis, D. Glycoscience@Synchrotron: Synchrotron Radiation Applied to Structural Glycoscience. Beilstein J. Org. Chem. 2017, 13, 1145-1167.

(593) Egorova, K. S.; Toukach, P. V. Csdb_Gt: A New Curated Database on Glycosyltransferases. Glycobiology 2017, 27, 285-290.

(594) Toukach, P. V. Bacterial Carbohydrate Structure Database 3: Principles and Realization. J. Chem. Inf. Model. 2011, 51, 159-170.

(595) Böhm, M.; Bohne-Lang, A.; Frank, M.; Loss, A.; Rojas-Macias, M. A.; Lütteke, T. Glycosciences.Db: An Annotated Data Collection Linking Glycomics and Proteomics Data (2018 Update). Nucleic Acids Res. 2019, 47, D1195-d1201.

(596) Lütteke, T. Glycan Data Retrieval and Analysis Using Glycosciences.De Applications. In A Practical Guide to Using Glycomics Databases, Aoki-Kinoshita, K. F. Ed.; Springer Japan, 2017; pp 335-350.

(597) Lutteke, 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.

(598) Pérez, S.; Sarkar, A.; Rivet, A.; Breton, C.; Imberty, A. Glyco3d: A Portal for Structural Glycosciences. In Glycoinformatics, Lütteke, T., Frank, M. Eds.; Springer New York, 2015; pp 241-258.

(599) 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 Japan, 2017; pp 133-161.

(600) Sarkar, A.; Pérez, S. Polysac3db: An Annotated Data Base of 3 Dimensional Structures of Polysaccharides. BMC Bioinform. 2012, 13, 302.

(601) Birch, J.; Van Calsteren, M. R.; Pérez, S.; Svensson, B. The Exopolysaccharide Properties and Structures Database: Eps-Db. Application to Bacterial Exopolysaccharides. Carbohydr. Polym. 2019, 205, 565-570.

(602) 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.

(603) 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.

(604) 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.

(605) 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.

(606) 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.

(607) 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.

(608) Jo, S.; Im, W. Glycan Fragment Database: A Database of Pdb-Based Glycan 3d Structures. Nucleic Acids Res. 2013, 41, D470-D474.

(609) 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.; Springer US, 2020; pp 1-14.

(610) Ribeiro, D.; Bonnardel, F.; Carvalho, A. L.; Palma, A.; Perez, S. To Be Published. 2022.

(611) 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.

(612) Cao, Y.; Park, S.-J.; Im, W. A Systematic Analysis of Protein–Carbohydrate Interactions in the Protein Data Bank. Glycobiology 2021, 31, 126-136.

(613) Siva Shanmugam, N. R.; Jino Blessy, J.; Veluraja, K.; Michael Gromiha, M. Procaff: Protein-Carbohydrate Complex Binding Affinity Database. Bioinformatics 2020, 36, 3615-3617.

(614) Gattani, S.; Mishra, A.; Hoque, M. T. Stackcbpred: A Stacking Based Prediction of Protein-Carbohydrate Binding Sites from Sequence. Carbohydr. Res. 2019, 486, 107857.

(615) 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.