Assessment of Glycoproteins Dynamics from Computer Simulations

Laercio Pol-Fachin[a] and Hugo Verli[a,b,*]

[a]*Centro de Biotecnologia, Universidade Federal do Rio Grande do Sul, Av Bento Gonçalves 9500, CP 15005, Porto Alegre 91500-970, RS, Brazil*

[b]*Faculdade de Farmácia, Universidade Federal do Rio Grande do Sul, Av Ipiranga 2752, Porto Alegre 90610-000, RS, Brazil*

Abstract: Glycoproteins are synthesized in most of the living organisms, being major components of the outer surface of mammalian cells and most of the secreted proteins in eukaryotes. Accordingly, a better comprehension of the biological processes in which they are involved requires the characterization of their structure and conformation. As such rationale faces several difficulties from both experimental and theoretical approaches, this review summarizes the current state of the art and methods employed to model and represent glycoproteins, including their carbohydrate moieties, through computer simulations.

Keywords: Conformational analysis, ensemble, glycan, glycoprotein, glycosylation, molecular dynamics, solvation.

1. INTRODUCTION

Glycoproteins are major components of the outer surface of mammalian cells and represent most of the secreted proteins in eukaryotes [1]. In fact, carbohydrate moieties may be covalently attached to polypeptide chains in most of the living organisms [2]. Such oligosaccharides are reported to be able to modify several properties of such molecules (Table **1**) [3-5], including solubility [6,7], folding and conformation [8-10], which, in turn, may influence glycoproteins biological roles. In this context, the comprehension, at the atomic level, of their interaction with solvent and target biomacromolecules passes through the characterization of their structure and conformation through experimental and/or theoretical techniques.

Table 1. Effects of Glycosylation on Proteins and their Surrounding Environment

Property	Effect[a]
Physico-chemical	Resistance to proteolysis
	Resistance to denaturation
	Increasing solution viscosity
	Lowering solution freezing point
	Increasing protein solubility
Enzymological	Shifting optimum pH
	Altering catalytic activity
Folding	Preventing aggregation
	Facilitating interaction with chaperones
	Nucleation of β turns
Biological activity	Altering protein–protein recognition
	Altering protein–carbohydrate recognition
	Altering transport and secretion tax
	Increasing / Decreasing multimerization

[a] Data from [3-5].

From the experimental point of view, the high flexibility of carbohydrate moieties may be a hindering factor for their crystallization, as well as their high degree of coordination to water molecules and the lack of strong lipophilic or dipolar inter-residue interactions [11]. In addition, the crystal environment may be capable to bend their structure, as a recent survey of Protein Data Bank (PDB) entries containing oligosaccharides suggests that about one-third of them contain significant errors in carbohydrate stereochemistry, nomenclature and consistency with the electron density [12,13]. As mentioned by Crispin and co-workers, some of the proposed models contain not only systematic errors in carbohydrate stereochemistry, but also hitherto unreported motifs in the primary structures of the glycans [13]. For example, there are approximately two hundred cases for which the PDB had to assign, by stereochemistry matching, the incorrect 2-(acetylamino)-2-deoxy-α-D-glucopyranose (NDG) rather than the correct 2-(acetylamino)-2-deoxy-β-D-glucopyranose (NAG) [14], which have been considered to differentially affect proteins structure and dynamics [15]. Unlike X-ray crystallography, that determine carbohydrates conformation at the crystal environment, NMR spectroscopy provides a set of spatial constraints representing solution averaged conformations. Unfortunately, NMR techniques may face difficulties in supplying an adequate number of NOE signals for carbohydrates three-dimensional (3D) characterization [11]. In addition, the derived solution averaged conformational states may not correspond to conformations populated in solution [16]. From the theoretical point of view, when properly validated and in conjunction with the above-mentioned experimental

*Address correspondence to this author at the Centro de Biotecnologia, Universidade Federal do Rio Grande do Sul, Av Bento Gonçalves 9500, CP 15005, Porto Alegre 91500-970, RS, Brazil; Tel: +55-51-3308-7770; Fax: +55-51-3308-7309; E-mail: hverli@cbiot.ufrgs.br

1570-193X/11 $58.00+.00

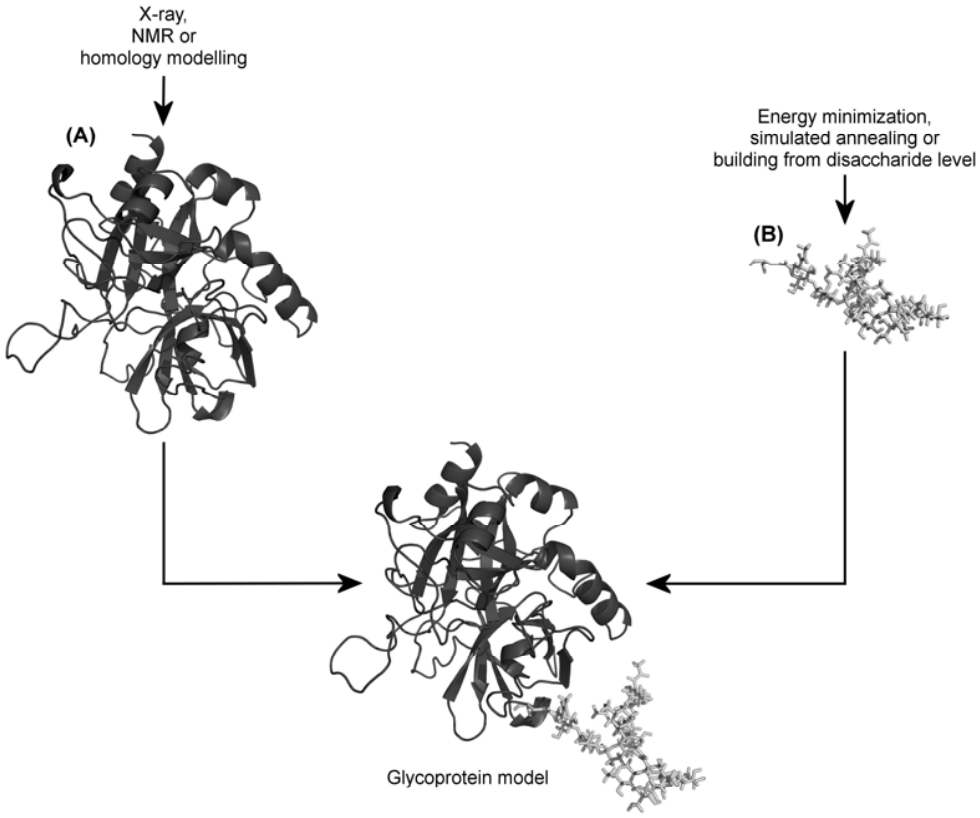

Fig. (1). Schematic representation of building a glycoprotein model.

Table 2. Available Resources for Building Glycoproteins and their Carbohydrate Moieties

Resource	Description	Ref
Online Glycoprotein Builder	Building oligosaccharides and their linking to proteins	[18]
Carbohydrate 3D Structure Predictor	Generation of 3D models for carbohydrates	[18]
Glydict	Prediction of 3D structures for N-oligosaccharides	[19]
Glyprot	Attachment of N-linked oligosaccharides to proteins	[20]

information, as well as relevant biochemical data, computer simulations emerge as a promising tool, capable to describe glycoproteins and mimic biological solutions.

2. GENERAL IDEA ON BUILDING MODELS FOR GLYCOPROTEINS

While the simulation of proteins may employ a starting geometry derived directly from PDB or from comparative modeling, computational studies on glycoproteins require additional steps as, in most cases, their complete forms are still not available in PDB. As a general feature, as illustrated in Fig. (**1**) (using thrombin structure under PDB ID 1DOJ as a model [17]), these steps include: (**I**) obtaining a 3D structure for the protein moiety; (**II**) obtaining a 3D structure for the carbohydrate moiety; and (**III**) attachment of the saccharidic and peptidic segments into a final model. In this context, while web-based resources [18-20] may be employed to build glycan chains and attach them to proteins (Table **2**), each step may be completed separately using distinct approaches and refinements and so supporting a fine tuning of the glycan conformation.

Regarding the first step, which embraces the obtaining of 3D structures for the protein moiety (Fig. (**1A**)), retrieving an X-ray or NMR structure from PDB data bank is generally the better option. Additionally, comparative modeling may be adequately employed in the absence of previous 3D experimental data to generate realistic models, given that the homologous sequence (target) shares with the experimentally established protein structure (template) significant sequence (30% or more) or structural similarity [21-23]. As well, in order to increase the accuracy of the produced models, the conformational space of such protein may be sampled, which consist in a good test and application for simulations methods as molecular dynamics (MD) [22].

In relation to the second step, comprising the achievement of 3D models for the carbohydrate moiety, theoretical approaches represent the most accessible source of information, seeing that some difficulties may be faced on obtaining meaningful experimental models for glycans (as above-discussed). Yet, X-ray structures, NMR models, *J* coupling constants, ring puckering or NOE signals should be employed, whenever available, to validate the obtained conformations.

Finally, in the third stage, a model for the studied glycoprotein is obtained by linking both protein and carbohydrate moieties in conformity to specific geometrical terms, which are well known for the *N*-glycosidic linkage [24,25], but mostly absent for other types of carbohydrate-amino acid linkages (see further in the text). However, as it will be discussed below, given the dynamical features of carbohydrates, the use of static, single models of glycoproteins for

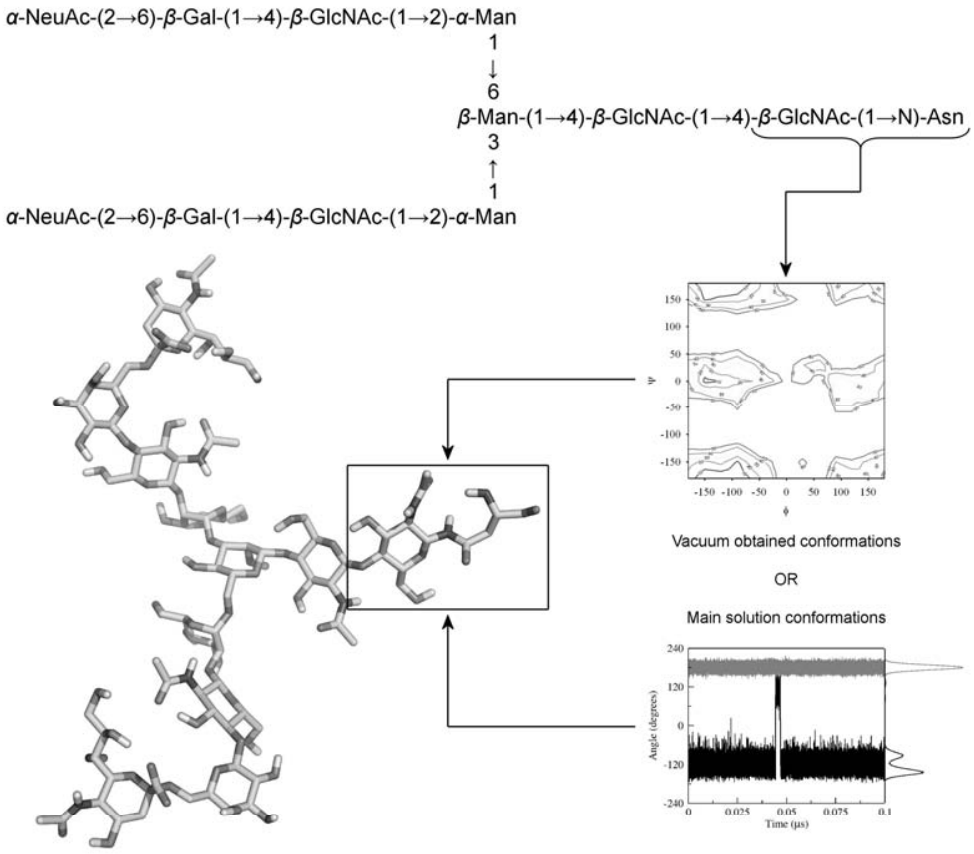

Fig. (2). Schematics of constructing a glycan chain from the disaccharide level. The conformational preferences of each glycosidic linkage composing a given oligosaccharide may be obtained from energy contour plots or solution simulations (β-GlcNAc-(1⟶N)-Asn employed as an example). Such conformational states from disaccharide units may be further combined in order to build the 3D model for the studied carbohydrate moiety.

interpretation of structural, functional and biological aspects of this class of biomolecules, should be carefully evaluated.

3. OBTAINING MODELS FOR GLYCANS

Carbohydrates are considered to have several orders of magnitude higher potential information content than any other biological macromolecule [26], mainly due to their great structural diversity, comprehending: the number of possible monosaccharide units, the formation of linear or branched structures, the two stereochemical possibilities on the linkages between saccharide units (α or β), the two potential isomeric forms (-D or -L) and covalent modifications in sugar residues, as methylation, sulfation, acetylation and phosphorylation [27]. In fact, while mammalian glycans rely on a group of approximately 10 common monosaccharide units, including *N*-acetyl-D-glucosamine (GlcNAc), D-mannose (Man), D-galactose (Gal), L-fucose (Fuc), neuraminic acid (NeuAc), *N*-acetyl-D-galactosamine (GalNAc) and D-glucose (Glc) [28,29], glycans assembled by other organisms may include an indeterminate number of such building blocks, employing a series of additional monosaccharide units, not found in mammalian organisms, such as pentoses [2,27,30]. All these properties make carbohydrates one of the most challenging classes of biomolecules for conformational characterization [11].

3.1. Strategies for Conformational Analysis of Carbohydrates

The methods employed in conformational analysis of compounds may generally fall into three distinct categories: (1) those that are random or stochastic, as MD, Monte Carlo and distance geometry based techniques; (2) those based on heuristics and artificial intelligence methods; and (3) those that are systematic [31]. While Monte Carlo approaches were shown to support a solid search for the conformational space of carbohydrates [32], strategies employed to date, specifically for modeling entire carbohydrate moieties of glycoproteins, include energy minimization [33] and simulated annealing [34]. However, the greater is the glycan moiety complexity, more difficult becomes the sampling of the molecule conformational space. Therefore, alternatively, such oligosaccharides may be built from the disaccharide level [35-37], in which minimum energy conformers from systematic analyses and/or MD derived conformations of isolated disaccharides may be used to build the complete oligosaccharide (Fig. (**2**)). This simplification appears to not impair the assessment of reliable glycoprotein models when compared to experimental data [37,38].

The determination of disaccharides conformational preferences are usually performed by describing their preferred conformations on potential energy surfaces as a function of their glycosidic torsional angles Φ and Ψ [26]. Several methods may be employed for calculating such maps. As it is usually considered that the most important energy variations are those related with the glycosidic dihedral angles [39], the hard sphere potential surfaces approach considers the constituent monosaccharides as rigid spheres, including the exocyclic groups. As a consequence, the glycosidic linkage geometry is determined by the spatial arrangement of such spheres [40]. Whereas simplified, such approach has being recently shown to support a reliable conformational description of complex systems in coarse grained MD simulations [41-43]. In spite of that, important variations in pyranoid ring geometries and orientations of pendent groups associated with Φ and Ψ rotation may be observed, which emphasize the need for a model to include bond length and angle degrees of freedom in some cases [40]. For instance, by al-

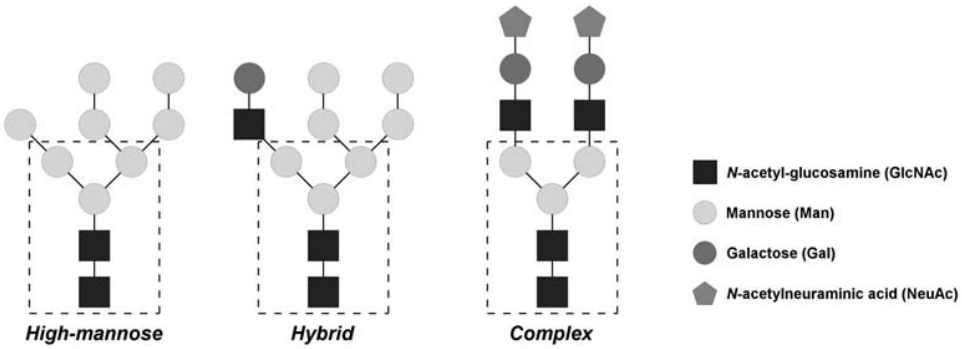

Fig. (3). The three main types of *N*-glycans observed in eukaryotes. The core pentasaccharide is shown inside the dashed lines.

lowing the atom coordinates to be minimized, a relaxed contour plot may be obtained, consisting in a fast approach for describing disaccharides conformational preferences. Such approach usually allows a lowering in both energy barriers between minima and energy of global and local minima, in comparison to maps obtained with hard sphere potential surfaces technique [40]. However, while in relaxed maps only rotations around glycosidic linkage dihedrals are systematically sampled, it should be considered that the orientation of secondary hydroxyl, hydroxymethyl and primary hydroxyl groups, as well as the different degrees of ring puckering, may influence the calculated final energy [39,44]. Therefore, when all of such angles are taken into account for searching the lowest energy of each point in the Φ-Ψ space, an adiabatic map is obtained.

Moreover, *ab initio* [45,46], density functional theory (DFT) [47,48] and molecular mechanics (MM) [49] have been employed to generate energy contour plots for disaccharides. While *ab initio* and DFT procedures may better describe the aspects determining glycosidic linkages geometry (as the *exo*-anomeric effect), and it may be difficult to properly describe such interactions in MM force fields [11], when such parameters are well validated, MM may provide adequate agreement with experimental data. As well, the many local minima possibly existing in *ab initio* or DFT obtained conformational maps for disaccharides may masquerade each other, reducing the contour plot quality [45]. Also, considering that MM is much less time demanding than quantum mechanical calculations, the former methodology is usually more applied to study the conformational preferences of disaccharides.

3.2. Force Fields for Carbohydrates

In the context of the carbohydrate moieties within glycoproteins, few groups of parameters have been employed for their conformational description, mainly by means of building energy contour plots in vacuum, as CHARMM [50], CVFF [51] and GROMOS96 [52,53]. Nevertheless, the largest amount of data about such disaccharides is based on MM3 [54,55], which is recognized to offer a highly detailed representation of carbohydrates conformational features in vacuum [56,57], whereas not used for representing glycoproteins (see further). The so obtained disaccharide conformations have been used in tools for 3D prediction of N-linked oligosaccharides (Glydict) [19] and for their attachment to proteins PDB structures (Glyprot) [20]. Such topic, including several other force field parameters for carbohydrates, is reviewed elsewhere [58,59].

Conformational search methods have indeed been employed to characterize the conformational behavior of carbohydrates [60], mainly through energy contour plots for disaccharides in vacuum. On the other hand, explicit solvent simulations have been considered as capable to better reproduce the conformational properties of oligosaccharides in comparison to calculation in its absence [32,61-64]. In spite of that, such vacuum derived conformations are not commonly associated with solution MD simulations, in order to achieve a proper conformational ensemble [60].

3.3. Solution Simulations of Carbohydrates

The explicit inclusion of solvent molecules has been described to reveal a distinct set of conformers when compared to calculations in its absence [65]. Specifically in the case of the carbohydrate moiety of glycoproteins, which present a high degree of branching, solvation appeared to be required to disclose conformations closer to those observed by experimental data [35,37]. In this context, while branching has been proposed to stiffen oligosaccharide main chains [11,64], the majority of conformational transitions in the carbohydrate moiety of glycoproteins have been observed to occur at their branching linkages [35]. Yet, the conformation and dynamics of these glycans [35,37] and other branched oligosaccharides [64], obtained from computer MD simulations with diverse force field parameters, have been observed to be in accordance to previous experimental data, mainly NMR, which further support the validity of solution simulations on these systems.

In fact, the role of solvation on the conformational preferences of glycosidic linkages composing the pentasaccharidic central core of *N*-glycans (Fig. **3**), where some branching points are located, is reinforced by several studies [66,67]. Regarding the β-D-Man-(1→4)-β-D-GlcNAc-(1→4)-β-D-GlcNAc unit, a more entropically favored conformation could be observed after solvation, in comparison to the vacuum structure, by means of a glycosidic linkage geometry modification and further interaction with water molecules [67]. In addition, residency times for water molecules were observed to be highly prevalent around the central core pentasaccharide, influencing the flexibility and overall topology of an oligomannose *N*-glycan [66]. These data reinforce the importance in submitting carbohydrate structural models to simulation techniques, if possible under explicit solvent conditions, as a strategy to obtain reliable solution models for such class of molecules.

4. SETTLING AND REFINING MODELS FOR GLYCOPROTEINS

After suitable models for both protein and carbohydrate moieties are obtained, the glycosylated amino acid residue and the first monosaccharide of the glycan must be correctly attached, as the geometry of such glycosidic linkages determines the exposure of the glycan chains on the protein surface. To date, about five different glycosylation types are identified, comprising ~40 glycosidic linkages occurring in nature [2], distributed in essentially all living organisms (Table **3**). In spite of such diversity, experimental data on the conformational preferences of such linkages are mostly restricted to the *N*- glycosyl bond (β-GlcNAc-(1→N)-Asn) [24,25], while only ~10 *N*- and *O*-glycosylation motifs have been hitherto conformationally studied by molecular modeling techniques [38,68-72]. In the absence of an adequate picture of the conformational features associated to such additional types of monosaccharide-amino acid linkages, as C-glycosylation and P-glycosylation, the

Table 3. Phylogenetic Distribution of Glycosidic Linkages Between Monosaccharides and Amino Acids

Glycosylation Type	Distribution[a]		
	Eukaryotes	*Archaea*	Bacteria
N-glycosylation	+	+	+
O-glycosylation	+	+	+
Glypiation	+	+	–
C-glycosylation	+	–	–
P-glycosylation	+	–	–

[a] Based on data from reference [2];

Table 4. Glycoproteins MD Simulations Studies

Glycoprotein	Number of Models	Simulated Time (Each)	Refs
Lectin[a]	1	0.3 ns	[35]
gp120[b]	2	0.1 ns	[74]
MHC Class I	4	0.5 ns	[36]
Human prion protein	1	~ 2 ns	[33]
Antifreeze glycoprotein 8	10	2 ns	[75]
Human coagulation fVII	1	~ 3 ns	[76]
Human mucin[b]	3	1 ns	[77]
Hemagglutinin[b]	2	10 ns	[15]
EGF-*like* domain	1	50 ns	[38]
CD59[c]	3	50 ns	[38]
Human CD2 domain	2	50 ns	[38]
α-subunit of hCG[d]	2	50 ns	[38]
Ovine ciclooxygenase-1	1	50 ns	[37]
Murine ciclooxygenase-2	1	50 ns	[37]

[a] From *Erythrina corallodendron*;
[b] Glycopeptides derived from such molecules;
[c] Human complement regulatory protein CD59;
[d] Human chorionic gonadotropin.

derived glycoprotein models may probably reveal biologically irrelevant protein-carbohydrate interactions, thus failing to properly describe one of its main structural aspects.

Subsequent to the obtaining of a glycoprotein model, refinement techniques may be employed to enhance its biological meaning. While some studies make use of energy minimization, for instance, to minimize bad steric interactions [73], solution simulations are often employed for this purpose [15,33,35-38,74-77], also supporting an investigation of the conformational space associated to such macromolecules. As MD simulations are a usual choice, it carries the challenge of covering sufficiently long time scales, necessary to adequately represent biological phenomena and/or experimental data [78]. For instance, reproduction of atomic force microscopy (AFM) data based on steered molecular dynamics (SMD) is not readily achievable due to the gap between the time scales of computer simulations (up to ~1 μs) and AFM measurements (~1 s) [79]. In spite of that, the recent advances regarding both hardware and software have supported a progressive increase in timescales accessible through MD simulations. The computational cost associated with carbohydrate simulations, for instance, had supported conformational samplings in the microsecond timescale, providing detailed information on such compounds dynamics, partially inaccessible to experiment, at the monosaccharide [80], disaccharide [81] and oligosaccharide [82] levels. Conversely, in the context of glycoproteins, the difficulties in achieving higher time scales are proportional to the systems size, being usually in the order of tens of nanosecond [37, 38].

4.1. Conformational Ensemble and Multiple Models

Conformational fluctuations are considered essential to the functions of proteins [78] and, consequently, of glycoproteins. In this context, simulation methods, as MD simulations, are capable to provide data on macromolecules conformational ensemble at the atomic level [83]. In order to obtain a proper conformational sampling during MD simulations, the energy barriers associated to the system degrees of freedom should be overcome in order to allow its configurational space to be explored [84]. While this sampling may be achieved by means of long time simulations, conversely, multiple simulations may be performed for the same system, starting from diverse initial conformations, which may support divergent [85] and convergent conformational descriptions [86] at the achieved time scales. Accordingly, multiple starting geometries had been employed when simulating glycoproteins, usually being obtained from multiple NMR models, multiple starting geometries for a given glycan moiety or multiple glycan compositions for a same glycoprotein, as demonstrated in Table **4** [15,33,35-38,74-77].

The use of multiple starting models is particularly important when simulating glycoproteins as a strategy to accurately describe their glycan moieties flexible pattern [11,26]. As illustrated in Fig. (**4**), a single glycoprotein 3D structure is not capable to properly represent the conformational ensemble adopted by any of three glycan chains determined by distinct NMR studies [87-89]. Similarly, a single glycoprotein static model is not representative of its plasticity, requiring an adequate conformational ensemble description, as obtained by means of simulation based methods. In fact, the

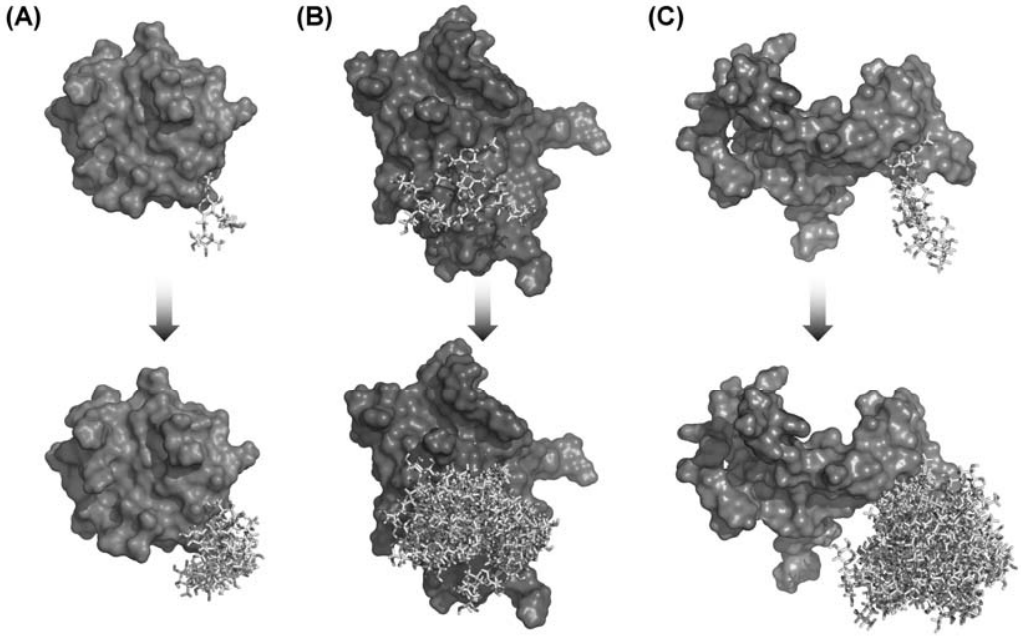

Fig. (4). Comparison between single and multiple models for representing the carbohydrate moiety of glycoproteins.

inclusion of molecular motion in glycoproteins study, as well as protein-carbohydrate complexes, may not be only important to evaluate glycosidic linkages modifications in the carbohydrate moiety, but also alterations in monosaccharide conformations. In this context, while chair conformations are considered the most populated in aqueous solutions [90], different ring puckering conformers have been described as relevant for some systems, both when free in solution [80,81] or bound to a target molecule [91].

4.2. Force Fields for Glycoproteins

As a general feature, in order to allow the study of glycoproteins, both carbohydrate and protein moieties should be described by the same, properly validated, force field [11]. In this context, most MD simulation programs are well-suited for modeling the protein part, while a lower number of them have tools for modeling of oligosaccharides moiety [26,58]. On the other hand, the linkage between them is considered, at times, to not be correctly taken into account [68], which indicates that a proper parameterization of the linkages between amino acids and monosaccharides is an important challenge for future refinements in force fields for glycoproteins. For example, while five different types of connections have been so far observed to occur in living organisms (Table **3**), studies on their conformational preferences have been reported for only two of them, that is, N- and O-glycosidic linkages. Accordingly, advancements in the comprehension of their molecular behavior, by means of new and better force field parameters, may further contribute in increasing our knowledge on glycoproteins biological functions.

Furthermore, to properly mimic glycoproteins biological environment, water molecules and counter ions are usually considered during simulations. In this context, taking into account that some of the parameters most employed to study isolated disaccharides (as the MM3 force field) were developed for the gas phase, their applicability for describing such molecules in solution is unclear [57]. Moreover, the behavior of such force fields when applied to simulate proteins in solution is also a matter of debate [56]. Nevertheless, other sets of parameters have been customized for studying glycoproteins, including AMBER [92], AMBER – GLYCAM [93,94], CHARMM [50], CVFF [51] and GROMOS96 [38]. Unfortunately, no comparative study between them has been carried out, as there has been for carbohydrates [57,58]. Still, such studies had been able to add insights into glycosylation effects over polypeptide chains.

4.3. Mutual Influence Between Protein and Carbohydrate Moieties

The comprehension of carbohydrates role on the biological function of glycoproteins pass through the comprehension on how protein and carbohydrate moieties interact and mutually influence each other in a single biomacromolecule. Such analyses, frequently based on root mean square deviation (RMSD) measurements, are mainly focused on the carbohydrate influence on the protein flexibility [33,38,77]. Among several studies comprising computer simulations of glycoproteins [33,38,74,75,95], conformational stabilization is frequently characterized as an important influence of glycosylation over polypeptide chains, mainly by means of mobility restriction. Such effect, when occurring close to the glycosylation site, considering *N*-linked glycans, is mostly attributed to the central core pentasaccharide (Fig. **3**) [10]. As well, it is proposed that such effect is raised when the number of oligosaccharides bounded to the protein is increased [9]. Considering that an N-linked oligosaccharide typically show from two to three branches, each presenting up to four monosaccharide residues, glycans may be considered to cover a vast part of the protein surface [1]. As a consequence, such oligosaccharides may influence glycoproteins conformation and dynamics in different regions of the polypeptide sequence, both through direct intra-molecular interactions and mediation of solvent molecules.

Considering the evaluation of the protein influence over the glycan chain, although RMSD is also used to measure oligosaccharides mobility [35,36], analysis on glycosidic linkage geometries may offer a more complete picture of carbohydrates conformational preferences. Nevertheless, the protein scaffold is recognized to influence some properties of their linked oligosaccharides, mainly by reducing the conformational flexibility of glycans [35,36]. Such restriction has been observed to occur due to a reduction on the number of possible conformers assumed by the glycosidic linkages that compose these glycan chains [35], possibly as a result of interactions between the polypeptide chain and their attached carbohydrate moieties [96].

4.4. Validating Simulations

Regarding the protein moiety of glycoproteins, the huge amount of data available on their structure and dynamics usually supports the process of validation. For instance, X-ray structures provide the most complete description of a structure, at the atomic level [97], and can be easily obtained from PDB. However, it is recognized that the crystal environment may produce packing effects, capable to influence protein conformation [98-101]. Moreover, the majority of our knowledge on proteins comes from time- and/or ensemble-averaged experiments [83]. In this context, when such data, as solution NMR, are used for comparison and validation, computer simulations must also consider ensemble-average properties, thus treating the obtained results in a manner that is analogous to what happens experimentally [83]. Besides, different protein force fields are recognized to behave comparably during MD simulations in relation to several structural and dynamical properties, such as solvent-accessible surface area, radius of gyration, deviation from their respective experimental structures and secondary structure (in this case, in the absence of glycan chains) [102].

While data related to the structure and dynamics of proteins are abundant, conformational information on carbohydrate moieties of glycoproteins, to be used for validation, are quite incipient. This occurs mainly due to the resistance of oligosaccharides to crystallization, the frequently low number of inter-residue NOEs and the difficulties associated with interpreting NOEs in terms of conformation [11]. When available, NOE signals should be employed as a measure of internuclear distances for validation. Nevertheless, some of such glycans were well described by NMR methods [87-89], providing multiple models and a realistic set of solution-averaged conformations to be compared to computer simulations. However, the glycosidic linkages composing glycoproteins carbohydrate moieties have been barely analyzed in most of the studies to date, which impairs a better depiction of such glycans conformational preferences.

Additional sources of experimental information on carbohydrates conformation may be obtained through *J* coupling constants [103]. In this context, the glycosidic linkage angles may be determined through inter-glycosidic $^{3}J_{HCOC}$, $^{3}J_{CCOC}$ and $^{2}J_{COC}$ couplings by employing Karplus-type correlation curves [104-106]. As well, vicinal $^{3}J_{HH}$ may be useful when analyzing ring puckering coordinates [107,108], that is, for determining endocyclic torsion angles required for describing ring conformations (for instance, in terms of chairs, boats or skew-boats), mainly by means of Cremer-Pople puckering parameters [109].

5. FUTURE PROSPECTS

While it has been proposed that more than half of known proteins can be potentially glycosylated [110,111], no more than twenty of them have been studied so far by molecular modeling techniques. As well, future advances in hardware and software, which allow further increase in reachable timescales for glycoproteins simulations, may be expected to provide further progress in relation to the comprehension of biological processes involving glycoproteins. An important aspect, for instance, comprise evaluating whether point modifications in carbohydrate or amino acid residues are capable to modify glycoproteins conformation, dynamics and, ultimately, function.

Whereas protein-protein interactions are considered vital to almost all cellular processes [112], the direct recognition between proteins and glycan chains are also considered important to a variety of biological processes [113,114]. While molecular docking is usually the main strategy employed for understanding such contacts, additional parameterization efforts are still required for both residues (amino acids and monosaccharides) flexibility and on the available scoring functions, in order to circumvent the high number of false positive results [115,116] and to support a more accurate description for the pertinent complexes. Moreover, free energy calculations may be expected to contribute in such goals, as previously achieved with protein-protein [117] and protein-carbohydrate [118,119] complexes. Unfortunately, the currently available data pointing to the importance of explicit solvent models for the obtaining of accurate conformational descriptions of carbohydrates [62] suggest a potential limitation of continuum solvent methods as MM-PBSA, even so efforts have been made recently in this area [120,121]. As well, force field parameterization and adequate conformational sampling may represent important challenges to a far-reaching use of Free Energy Perturbation [122,123] and Linear Interaction Energy [124,125] strategies in such systems.

Additionally, coarse graining approach emerge allowing micro and millisecond timescales to be sampled, with low computational cost, as well as the study of multi-macromolecular systems, comprising protein and carbohydrates separately [42,43]. As new parameters are developed and included in the currently available force fields, or new and more complete set of parameters are developed, including topologies for different types of carbohydrate-amino acid linkages and biologically important monosaccharide residues, important advances may expected to occur in comprehending biological processes involving glycoproteins.

ACKNOWLEDGEMENTS

This work was supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq#472174/2007-0), MCT, by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), MEC, Brasília, DF, Brazil.

REFERENCES

[1] Dwek, R.A. Glycobiology: toward understanding the function of sugars. *Chem. Rev.*, **1996**, *96*, 683-720.

[2] Spiro, R.G. Protein glycosylation: nature, distribution, enzymatic formation, and disease implications of glycopeptide bonds. *Glycobiology*, **2002**, *12*, 43R-56R.

[3] Varki, A. Biological roles of oligosaccharides: all of the theories are correct. *Glycobiology*, **1993**, *3*, 97-130.

[4] Sears, P.; Wong, C.-H. Enzyme action in glycoprotein synthesis. *Cell. Mol. Life Sci.*, **1998**, *54*, 223-252.

[5] Skropeta, D. The effect of individual N-glycans on enzyme activity. *Bioorg. Med. Chem.*, **2009**, *17*, 2645-2653.

[6] Tams, J.W.; Vind, J.; Welinder, K.G. Adapting protein solubility by glycosylation. N-Glycosylation mutants of Coprinus cinereus peroxidase in salt and organic solutions. *Biochim. Biophys. Acta*, **1999**, *1432*, 214-221.

[7] Price, N.J.; Pinheiro, C.; Soares, C.M.; Ashford, D.A.; Ricardo, C.P.; Jackson, P.A. A biochemical and molecular characterization of LEP1, an extensin peroxidase from lupin. *J. Biol. Chem.*, **2003**, *278*, 41389-41399.

[8] Wyss, D.F.; Wagner, G. The structural role of sugars in glycoproteins. *Curr. Opin. Biotechnol.*, **1996**, *7*, 409-416.

[9] Shental-Bechor, D.; Levy, Y. Effect of glycosylation on protein folding: A close look at thermodynamic stabilization. *Proc. Natl. Acad. Sci. U. S. A.*, **2008**, *105*, 8256-8261.

[10] Hanson, S.R.; Culyba, E.K.; Hsu, T.L.; Wong, C.H.; Kelly, J.W.; Powers, E.T. The core trisaccharide of an N-linked glycoprotein intrinsically accelerates folding and enhances stability. *Proc. Natl. Acad. Sci. U. S. A.*, **2009**, *106*, 3131-3136.

[11] Woods, R.J. Computational carbohydrate chemistry: what theoretical methods can tell us. *Glycoconjug. J.*, **1998**, *15*, 209-216.

[12] Lütteke, T.; Frank, M.; von der Lieth, C.-W. Data mining the protein data bank: automatic detection and assignment of carbohydrate structures. *Carbohydr. Res.*, **2004**, *339*, 1015-1020.

[13] Crispin, M.; Stuart, D.I.; Jones, E.Y. Building meaningful models of glycoproteins. *Nat. Struct. Mol. Biol.*, **2007**, *14*, 354.

[14] Berman, H.M.; Henrick, K.; Nakamura, H.; Markley, J. Building meaningful models of glycoproteins. *Nat. Struct. Mol. Biol.*, **2007**, *14*, 354-355.

[15] Bosques, C.J.; Tschampel, S.M.; Woods, R.J.; Imperiali, B. Effects of glycosylation on peptide conformation: a synergistic experimental and computational study. *J. Am. Chem. Soc.*, **2004**, *126*, 8421-8425.

[16] Cumming, D.A.; Carver, J.P. Virtual and solution conformations of oligosaccharides. *Biochemistry*, **1987**, *26*, 6664-6676.

[17] Recacha, R.; Costanzo, M.J.; Maryanoff, B.E.; Carson, M.; DeLucas, L.; Chattopadhyay, D. Structure of human alpha-thrombin complexed with

RWJ-51438 at 1.7 A: unusual perturbation of the 60A-60I insertion loop. *Acta Crystallogr. D Biol. Crystallogr.*, **2000**, *56*, 1395-1400.

[18] Complex Carbohydrate Research Center, University of Georgia, Athens, GA. Woods Group (2005-2010). GLYCAM http:// www.glycam.com(Accessed April 28, 2010)

[19] Frank, M.; Bohne-Lang, A.; Wetter, T.; von der Lieth, C.-W. Rapid generation of a representative ensemble of N-glycan conformations. *In Silico Biol.*, **2002**, *2*, 427-439.

[20] Bohne-Lang, A.; von der Lieth, C.-W. GlyProt: in silico glycosylation of proteins. *Nucleic Acids Res.*, **2005**, *33*, W214-W219.

[21] Vitkup, D.; Melamud, E.; Moult, J.; Sander, C. Completeness in structural genomics. *Nat. Struct. Biol.*, **2001**, *8*, 559-566.

[22] Baker, D.; Sali, A. Protein structure prediction and structural genomics. *Science*, **2001**, *294*, 93-96.

[23] Cavasotto, C.N.; Phatak, S.S. Homology modeling in drug discovery: current trends and applications. *Drug Discov. Today*, **2009**, *14*, 676-683.

[24] Imberty, A.; Peréz, S. Stereochemistry of the *N*-glycosylation sites in glycoproteins. *Protein Eng.*, **1995**, *8*, 699-709.

[25] Petrescu, A.J.; Milac, A.L.; Petrescu, S.M.; Dwek, R.A.; Wormald, M.R. Statistical analysis of the protein environment of N-glycosylation sites: implications for occupancy, structure, and folding. *Glycobiology*, **2004**, *14*, 103-114.

[26] Imberty, A.; Pérez, S. Structure, conformation, and dynamics of bioactive oligosaccharides: theoretical approaches and experimental validations. *Chem. Rev.*, **2000**, *100*, 4567-4588.

[27] Pilobello, K.T.; Mahal, L.K. Deciphering the glycocode: the complexity and analytical challenge of glycomics. *Curr. Opin. Chem. Biol.*, **2007**, *11*, 300-305.

[28] Helenius, A.; Aebi, M. Intracellular functions of N-Linked glycans. *Science*, **2001**, *291*, 2364-2369.

[29] Kukuruzinska, M.A.; Lennon, K. Protein N-Glycosylation: molecular genetics and functional significance. *Crit. Rev. Oral Biol. Med.*, **1998**, *9*, 415-448.

[30] Stenutz, R.; Weintraub, A.; Widmalm, G. The structures of *Escherichia coli* O-polysaccharide antigens. *FEMS Microbiol. Rev.*, **2006**, *30*, 382-403.

[31] Beusen, D.D.; Shands, E.F.B. Systematic search strategies in conformational analysis. *Drug Discov. Today*, **1996**, *1*, 429-437.

[32] Kifli, N; Htar, T.T.; De Clercq, E.; Balzarini, J.; Simons, C. Novel bicyclic sugar modified nucleosides: synthesis, conformational analysis and antiviral evaluation. *Bioorg. Med. Chem.*, **2004**, *12*, 3247-3257.

[33] Zuegg, J.; Gready, J.E. Molecular dynamics simulation of human prion protein including both N-linked oligosaccharides and the GPI anchor. *Glycobiology*, **2000**, *10*, 959-974.

[34] Mukhopadhyay, C. Molecular dynamics simulation of glycoprotein-glycans of Immunoglobulin G and Immunoglobulin M. *Biopolymers*, **1998**, *45*, 177-190.

[35] Naidoo, K.J.; Denysyk, D.; Brady. J.W. Molecular dynamics simulations of the N-linked oligosaccharide of the lectin from *Erythrina corallodendron*. *Protein Eng.*, **1997**, *10*, 1249-1261.

[36] Mandal, T.K.; Mukhopadhyay, C. Effect of glycosylation on structure and dynamics of MHC Class I glycoprotein: a molecular dynamics study. *Biopolymers*, **2001**, *59*, 11-23.

[37] Fernandes, C.L.; Sachett, L.G.; Pol-Fachin, L.; Verli, H. GROMOS96 43a1 performance in predicting oligosaccharides conformational ensemble within glycoproteins. *Carbohydr. Res.*, **2010**, *345*, 663-667.

[38] 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.

[39] Stortz, C.A.; Cerezo, A.S. Disaccharide conformational maps: 3D contours or 2D plots? *Carbohydr. Res*., **2002**, *337*, 1861-1871.

[40] Pérez, S.; Kouwijzer, M.; Mazeau, K.; Engelsen, S.B. Modeling polysaccharides: Present status and challenges. *J. Mol. Graph.*, **1996**, *14*, 307-321.

[41] Liwo, A.; Khalili, M.; Czaplewski, C.; Kalinowski, S.; Ołdziej, S.; Wachucik, K.; Scheraga, H.A. Modification and Optimization of the United-Residue (UNRES) potential energy function for canonical simulations. i. temperature dependence of the effective energy function and tests of the optimization method with single training proteins. *J. Phys. Chem. B*, **2007**, *111*, 260-285.

[42] 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.

[43] Lopez, C.A., Rzepiela, A.J., de Vries, A.H., Dijkhuizen, L., Hünenberger, P.H., and Marrink, S.J. Martini coarse-grained force field: extension to carbohydrates. *J. Chem. Theory Comput.*, **2009**, *5*, 3195-3210.

[44] Stortz, C.A. Disaccharide conformational maps: how adiabatic is an adiabatic map? *Carbohydr. Res*., **1999**, *322*, 77-86.

[45] da Silva, C.O.; Nascimento, M.A. Ab initio conformational maps for disaccharides in gas phase and aqueous solution. *Carbohydr. Res*., **2004**, *339*, 113-122.

[46] French, A.D.; Kelterer, A.-M.; Johnson, G.P.; Dowd, M.K.; Cramer, C.J. HF/6-31G* energy surfaces for disaccharide analogs. *J. Comput. Chem.*, **2001**, *22*, 65-78.

[47] Schnupf, U.; Willett, J.L.; Bosma, W.B.; Momany, F.A. DFT studies of the disaccharide, α-maltose: relaxed isopotential maps. *Carbohydr. Res*., **2007**, *342*, 2270-2285.

[48] Yousfi, N.; Sekkal-Rahal, M.; Sayede, A.; Springborg, M. Relaxed energetic maps of κ-Carrabiose: A DFT study. *J. Comput. Chem.*, **2010**, *31*, 1312-1320.

[49] Tvaroška, I.; Pérez, S. Conformational-energy calculations for oligosaccharides: a comparison of methods and a strategy of calculation. *Carbohydr. Res*., **1986**, *149*, 389-410.

[50] Ha, S.N.; Giammona, A.; Field, M.; Brady, J.W. A revised potential-energy surface for molecular mechanics studies of carbohydrates. *Carbohydr. Res.*, **1988**, *180*, 207-221.

[51] Hwang, M.J.; Ni, X.; Waldman, M.; Ewig, C.S.; Hagler, A.T. Derivation of class II force fields. VI. Carbohydrate compounds and anomeric effects. *Biopolymers*, **1998**, *45*, 435-468.

[52] Scott, W.R.P.; Hunenberger, P.H.; Tironi, I.G.; Mark, A.E.; Billeter, S.R.; Fennen, J.; Torda, A.E.; Huber, T.; Kruger, P.; van Gunsteren, W.F. The GROMOS biomolecular simulation program. *J. Phys. Chem. A*, **1999**, *103*, 3596-3607.

[53] Verli, H.; Guimarães, J.A. Molecular dynamics simulation of a decasaccharide fragment of heparin in aqueous solution. *Carbohydr. Res.*, **2004**, *339*, 281-290.

[54] Allinger, N.L.; Yuh, Y.H.; Lii, J.-H. Molecular mechanics. The MM3 Force Field for Hydrocarbons. 1. *J. Am. Chem. Soc.*, **1989**, *111*, 8551-8582.

[55] 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.

[56] MacKerell, A.D.Jr.; Bashford, D.; Bellott, M.; Dunbrack, R.L.Jr.; Evanseck, J.D.; Field, M.J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F.T.K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D.T.; Prodhom, B.; Reiher, W.E.3rd.; Roux, B.; Schlenkrich, M.; Smith, J.C.; Stote, R.; Straub, J.; Watanabe, M.; Wiórkiewicz-Kuczera, J.; Yin, D.; Karplus, M. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B*, **1998**, *102*, 3586-3616.

[57] Mackerell, A.D.Jr. Empirical force fields for biological macromolecules: overview and issues. *J. Comput. Chem*., **2004**, *25*, 1584-1604.

[58] Pérez, S.; Imberty, A.; Engelsen, S.B.; Gruza, J.; Mazeau, K.; Jimenez-Barbero, J.; Poveda, A.; Espinosa, J.-F.; van Eyck, B.P.; Johnson, G.; French, A.D.; Kouwijzer, M.L.C.E.; Grootenuis, P.D.J.; Bernardi, A.; Raimondi, L.; Senderowitz, H.; Durier, V.; Vergoten, G.; Rasmussen, K. A comparison and chemometric analysis of several molecular mechanics force fields and parameter sets applied to carbohydrates. *Carbohydr. Res.*, **1998**, *314*, 141-155.

[59] Hemmingsen, L.; Madsen, D.E.; Esbensen, A.L.; Olsen, L.; Engelsen, S.B. Evaluation of carbohydrate molecular mechanics force fields by quantum mechanical calculations. *Carbohydr. Res.*, **2004**, *339*, 937-948.

[60] Hong, X.; Hopfinger, A.J. Construction, molecular modeling, and simulation of *Mycobacterium tuberculosis* cell walls. *Biomacromolecules*, **2004**, *5*, 1052-1065.

[61] Kirschner, K.N.; Woods, R.J. Solvent interactions determine carbohydrate conformation. *Proc. Natl. Acad. Sci. U. S. A.*, **2001**, *98*, 10541-10545.

[62] Almond, A.; Sheehan, J.K. Predicting the molecular shape of polysaccharides from dynamic interactions with water. *Glycobiology*, **2003**, *13*, 255-264.

[63] Elklund, R.; Widmalm, G. Molecular dynamics simulations of an oligosaccharide using a force field modified for carbohydrates. *Carbohydr. Res.*, **2003**, *338*, 393-398.

[64] González-Outeiriño, J.; Kadirvelraj, R.; Woods, R.J. Structural elucidation of type III group B *Streptococcus* capsular polysaccharide using molecular dynamics simulations: the role of sialic acid. *Carbohydr. Res.*, **2005**, *340*, 1007-1018.

[65] Yoneda, J.D.; Albuquerque, M.G.; Leal, K.Z.; Seidl, P.R.; Wheeler, R.A.; Boesch, S.E.; de Alencastro, R.B.; de Souza, M.C.B.V.; Ferreira, V.F. Molecular dynamics simulations of a nucleoside analogue of 1,4-dihydro-4-oxoquinoline-3-carboxylic acid synthesized as a potential antiviral agent: conformational studies in vacuum and in water. *J. Mol. Struct.*, **2006**, *778*, 97-103.

[66] Woods, R.J.; Pathiaseril, A.; Wormald, M.R.; Edge, C.J.; Dwek, R.A. The high degree of internal flexibility observed for an oligomannose oligosaccharide does not alter the overall topology of the molecule. *Eur. J. Biochem.*, **1998**, *258*, 372-386.

[67] Stanca-Kaposta, E.C.; Gamblin, D.P.; Cocinero, E.J.; Frey, J.; Kroemer, R.T.; Fairbanks, A.J.; Davis, B.G.; Simons, J.P. Solvent interactions and conformational choice in a core n-glycan segment: gas phase conformation of the central, branching trimannose unit and its singly hydrated complex. *J. Am. Chem. Soc.*, **2008**, *130*, 10691-10696.

[68] Ali, M.M.; Aich, U.; Varghese, B.; Pérez, S.; Imberty, A.; Loganathan, D. Conformational preferences of the Aglycon moiety in models and analogs of Glcnac-asn linkage: crystal structures and ab initio quantum chemical calculations of N-(β-D-Glycopyranosyl)haloacetamides. *J. Am. Chem. Soc.*, **2008**, *130*, 8317-8325.

[69] Corzana, F.; Busto, J.H.; Jiménez-Osés, G.; Asensio, J.L.; Jiménez-Barbero, J.; Peregrina, J.M.; Avenoza, A. New Insights into α-GalNAc-Ser Motif: influence of hydrogen bonding versus solvent interactions on the preferred conformation. *J. Am. Chem. Soc.*, **2006a**, *128*, 14640-14648.

[70] Corzana, F.; Busto, J.H.; Engelsen, S.B.; Jiménez-Barbero, J.; Asensio, J.L.; Peregrina, J.M.; Avenoza, A. Effect of β-O-Glucosylation on L-Ser and L-Thr Diamides: A Bias toward α-Helical Conformations. *Chemistry*, **2006b**, *12*, 7864-7871.

[71] Corzana, F.; Busto, J.H.; Jiménez-Osés, G.; García de Luis, M.; Asensio, J.L.; Jiménez-Barbero, J.; Peregrina, J.M.; Avenoza, A. Serine versus Threonine Glycosylation: the methyl group causes a drastic alteration on the carbohydrate orientation and on the surrounding water shell. *J. Am. Chem. Soc.*, **2007**, *129*, 9458-9467.

[72] Fernández-Tejada, A.; Corzana, F.; Busto, J.H.; Jiménez-Osés, G.; Jiménez-Barbero, J.; Avenoza, A.; Peregrina, J.M. Insights into the geometrical features underlying β-*o*-glcnac glycosylation: water pockets drastically modulate the interactions between the carbohydrate and the peptide backbone. *Chemistry*, **2009**, *15*, 7297-7301.

[73] Zhu, X.; Borchers, C.; Bienstock, R.J.; Tomer, K.B. Mass spectrometric characterization of the glycosylation pattern of HIV-gp120 expressed in CHO cells. *Biochemistry*, **2000**, *39*, 11194-11204.

[74] Huang, X.; Barchi, J.J.Jr; Lung, F.D.; Roller, P.P.; Nara, P.L.; Muschik, J.; Garrity, R.R. Glycosylation affects both the three-dimensional structure and antibody binding properties of the HIV-1IIIB GP120 Peptide RP135. *Biochemistry*, **1997**, *36*, 10846-10856.

[75] Nguyen, D.H.; Colvin, M.E.; Yeh, Y.; Feeney, R.E.; Fink, W.H. The Dynamics, structure, and conformational free energy of proline-containing antifreeze glycoprotein. *Biophys. J.*, **2002**, *82*, 2892-2905.

[76] Perera, L.; Darden, T.A.; Pedersen, L.G. Predicted solution structure of zymogen human coagulation FVII. *J. Comput. Chem.*, **2002**, *23*, 35-47.

[77] Rubinstein, A.; Kinarsky, L.; Sherman, S. Molecular dynamics simulations of the O-glycosylated 21-residue MUC1 peptides. *Int. J. Mol. Sci.*, **2004**, *5*, 119-128.

[78] Lyman, E.; Zuckerman, D.M. Ensemble-Based convergence analysis of biomolecular trajectories. *Biophys. J.*, **2006**, *91*, 164-172.

[79] Lu, Z.; Hu, H.; Yang, W.; Marszalek, P.E. Simulating force-induced conformational transitions in polysaccharides with the SMD replica exchange method. *Biophys. J.*, **2006**, *91*, L57-L59.

[80] Kräutler, V.; Müller, M.; Hünenberger, P.H. Conformation, dynamics, solvation and relative stabilities of selected β-hexopyranoses in water: a molecular dynamics study with the GROMOS 45A4 force field. *Carbohydr. Res.*, **2007**, *342*, 2097-2124.

[81] Pol-Fachin, L.; Verli, H. Depiction the forces participating in the 2-*O*-sulfo-α-L-iduronic acid conformational preference in heparin sequences within aqueous solutions. *Carbohydr. Res.*, **2008**, *343*, 1435-1445.

[82] Landström, J.; Widmalm, G. Glycan flexibility: insights into nanosecond dynamics from a microsecond molecular dynamics simulation explaining an unusual nuclear Overhauser effect. *Carbohydr. Res.*, **2010**, *345*, 330-333.

[83] Zagrovic, B.; Pande, V.S. How does averaging affect protein structure comparison on the ensemble level? *Biophys. J.*, **2004**, *87*, 2240-2246.

[84] van Gunsteren, W.F.; Berendsen, H.J.C. Computer simulation of molecular dynamics: methodology, applications, and perspectives in chemistry. *Angew. Chem. Int. Ed. Engl.*, **1990**, *29*, 992-1023.

[85] Koller, A.N.; Schwalbe, H.; Gohlke, H. starting structure dependence of NMR order parameters derived from MD simulations: implications for judging force-field quality. *Biophys. J.*, **2008**, *95*, L04-L06.

[86] Brigo, A.; Lee, K.W.; Iurcu Mustata, G.; Briggs, J.M. Comparison of multiple molecular dynamics trajectories calculated for the drug-resistant HIV-1 Integrase T66I/M154I catalytic domain. *Biophys. J.*, **2005**, *88*, 3072-3082.

[87] Fletcher, C.M.; Harrison, R.A.; Lachmann, P.J.; Neuhaus, D. Structure of a soluble, glycosylated form of the human complement regulatory protein CD59. *Structure*, **1994**, *2*, 185-199.

[88] Wyss, D.F.; Choi, J.S.; Li, J.; Knoppers, M.H.; Willis, K.J.; Arulanandam, A.R.N.; Smolyar, A.; Reinherz, E.L.; Wagner, G. Conformation and function of the N-linked glycan in the adhesion domain of human CD2. *Science*, **1995**, *269*, 1273-1278.

[89] Erbel, P.J.A.; Karimi-Nejad, Y.; van Kuik, J.A.; Boelens, R.; Kamerling, J.P.; Vliegenthart, J.F.G. Effects of the N-Linked glycans on the 3D structure of the free r-subunit of human chorionic gonadotropin. *Biochemistry*, **2000**, *39*, 6012-6021.

[90] Reeves, R.E. The shape of pyranose rings. *J. Am. Chem. Soc.*, **1950**, *72*, 1499-1506.

[91] Verli, H.; Guimarães, J.A. Insights into the induced fit mechanism in antithrombin-heparin interaction using molecular dynamics simulations. *J. Mol. Graph. Model.*, **2005**, *24*, 203-212.

[92] Renouf, D.V.; Hounsell, E.F. Conformational studies of the backbone (poly-N-acetyllactosamine) and the core region sequences of O-linked carbohydrate chains. *Int. J. Biol. Macromol.*, **1993**, *15*, 37-42.

[93] Woods, R.J.; Dwek, R.A.; Edge, C.J.; Fraser-Reid, B. Molecular mechanical and molecular dynamical simulations of glycoproteins and oligosaccharides. 1. GLYCAM-93 parameter development. *J. Phys. Chem.*, **1995**, *99*, 3832-3846.

[94] Kirschner, K.N.; Yongye, A.B.; Tschampel, S.M.; González-Outeiriño, J.; Daniels, C.R.; Foley, B.L.; Woods, R.J. GLYCAM06: A generalizable biomolecular force field. Carbohydrates. *J. Comput. Chem.*, **2008**, *29*, 622.

[95] Bailey, D.; Renouf, D.V.; Large, D.G.; Warren, C.D.; Hounsell, E.F. Conformational studies of the glycopeptide Ac-Tyr-[Man_5GlcNAc-β-(1→4)GlcNAc-β-(1→Nδ)]-Asn-Leu-Thr-Ser-OBz and the constituent peptide and oligosaccharide. *Carbohydr. Res.*, **2000**, *324*, 242-254.

[96] Lommerse, J.P.; Kroon-Batenburg, L.M.; Kamerling, J.P.; Vliegenthart, J.F. Conformational analysis of the Xylose-containing N-Glycan of pineapple stem bromelain as part of the intact glycoprotein. *Biochemistry*, **1995**, *34*, 8196-8206.

[97] Petrescu, A.J.; Wormald, M.R.; Dwek, R.A. Structural aspects of glycomes with a focus on *N*-glycosylation and glycoprotein folding. *Curr. Opin. Struct. Biol.*, **2006**, *16*, 600-607.

[98] Eyal, E.; Gerzon, S.; Potapov, V.; Edelman, M.; Sobolev, V. The limit of accuracy of protein modeling: influence of crystal packing on protein structure. *J. Mol. Biol.*, **2005**, *351*, 431-442.

[99] Andrec, M.; Snyder, D.A.; Zhou, Z.; Young, J.; Montelione, G.T.; Levy, R.M. A large data set comparison of protein structures determined by crystallography and NMR: Statistical test for structural differences and the effect of crystal packing. *Proteins*, **2007**, *69*, 449-465.

[100] Verli, H.; Calazans, A.; Brindeiro, R.; Tanuri, A.; Guimarães, J.A. Molecular dynamics analysis of HIV-1 matrix protein: clarifying differences between crystallographic and solution structures. *J. Mol. Graph. Model.*, **2007**, *26*, 62-68.

[101] Lima, L.M.; Becker, C.F.; Giesel, G.M.; Marques, A.F.; Cargnelutti, M.T.; de Oliveira Neto, M.; Monteiro, R.Q.; Verli, H.; Polikarpov, I. Structural and thermodynamic analysis of thrombin:suramin interaction in solution and crystal phases. *Biochim. Biophys. Acta*, **2009**, *1794*, 873-881.

[102] Price, D.J.; Brooks III, C.L. Modern protein force fields behave comparably in molecular dynamics simulations. *J. Comput. Chem.*, **2002**, *23*, 1045-1057.

[103] Petersen, B.O.; Vinogradov, E.; Kay, W.; Würtz, P.; Nyberg, N.T.; Duus, J.Ø.; Sørensen, O.W. H2BC: a new technique for NMR analysis of complex carbohydrates. *Carbohydr. Res.*, **2006**, *341*, 550-556.

[104] Church, T.; Carmichael, I.; Serianni, A.S. Two-bond ^{13}C-^{13}C spin-coupling constants in carbohydrates: effect of structure on coupling magnitude and sign. *Carbohydr. Res.*, **1996**, *280*, 177-186.

[105] Milton, M.J.; Harris, M.A.; Probert, M.A.; Field, R.A.; Homans, S.W. New conformational constraints in isotopically (^{13}C) enriched oligosaccharides. *Glycobiology*, **1998**, *8*, 147-153.

[106] Bose, B.; Zhao, S.; Stenutz, R.; Cloran, F.; Bondo, F.; Bondo, G.; Hertz, B.; Carmichael, I.; Serianni, A. Three-Bond C−O−C−C spin-coupling constants in carbohydrates: development of a karplus relationship. *J. Am. Chem. Soc.*, **1998**, *120*, 11158-11173.

[107] Haasnoot, C.A.G. Conformational analysis of six-membered rings in solution: ring puckering coordinates derived from vicinal NMR proton-proton coupling constants. *J. Am. Chem. Soc.*, **1993**, *115*, 1460-1468.

[108] Martin-Pastor, M.; Bush, C.A. New strategy for the conformational analysis of carbohydrates based on NOE and ^{13}C NMR coupling constants. Application to the flexible polysaccharide of *Streptococcus mitis* J22. *Biochemistry*, **1999**, *38*, 8045-8055.

[109] Cremer, D.; Pople, J.A. A General definition of ring puckering coordinates. *J. Am. Chem. Soc.*, **1975**, *97*, 1354-1358.

[110] Apweiler, R.; Hermjacob, 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.

[111] Ben-Dor, S.; Esterman, N.; Rubin, E.; Sharon, N. Biases and complex patterns in the residues flanking protein N-glycosylation sites. *Glycobiology*, **2004**, *14*, 95-101.

[112] Gray, J.J.; Moughon, S.E.; Kortemme, T.; Schueler-Furman, O.; Misura, K.M.; Morozov, A.V.; Baker, D. Protein-protein docking predictions for the CAPRI experiment. *Proteins*, **2003**, *52*, 118-122.

[113] Kerzmann, A.; Neumann, D.; Kohlbacher, O. SLICK - scoring and energy functions for protein-carbohydrate interactions. *J. Chem. Inf. Model.*, **2006**, *46*, 1635-1642.

[114] 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.

[115] Camacho, C.J.; Vajda, S. Protein-protein association kinetics and protein docking. *Curr. Opin. Struct. Biol.*, **2002**, *12*, 36-40.

[116] Englebienne, P.; Moitessier, N. Docking ligands into flexible and solvated macromolecules. 4. are popular scoring functions accurate for this class of proteins? *J. Chem. Inf. Model.*, **2009**, *49*, 1568-1580.

[117] Gohlke, H.; Kiel, C.; Case, D.A. Insights into protein-protein binding by binding free energy calculation and free energy decomposition for the Ras-Raf and Ras-RalGDS complexes. *J. Mol. Biol.*, **2003**, *330*, 891-913.

[118] Hill, A.D.; Reilly, P.J. A Gibbs Free energy correlation for automated docking of carbohydrates. *J. Comput. Chem.*, **2008**, *29*, 1131-1141.

[119] Gandhi, N.S.; Mancera, R.L. Free energy calculations of glycosaminoglycan–protein interactions. *Glycobiology*, **2009**, *19*, 1103-1115.

[120] Green, D.F. Optimized parameters for continuum solvation calculations with carbohydrates. *J. Phys. Chem. B*, **2008**, *112*, 5238-5249.

[121] Guimarães, C.R.W.; Mathiowetz, A.M. Addressing limitations with the MM-GB/SA scoring procedure using the watermap method and free energy perturbation calculations. *J. Chem. Inf. Model.*, **2010**, *50*, 547-559.

[122] Zwanzig, R.W. High temperature equation of state by a perturbation method. i. nonpolar gases. *J. Chem. Phys.*, **1954**, *22*, 1420-1426.

[123] Mezei, M. The finite difference thermodynamic integration, tested on calculating the hydration free energy difference between acetone and dimethylamine in water. *J. Chem. Phys.*, **1987**, *86*, 7084-7088.

[124] Åqvist, J.; Medina, C.; Samuelsson, J.E. A new method for predicting binding affinity in computer-aided drug design. *Protein Eng.*, **1994**, *7*, 385-391.

[125] Åqvist, J.; Luzhkov, V.B.; Brandsdal, B.O. Ligand binding affinities from MD simulations. *Acc. Chem. Res.*, **2002**, *35*, 358-365.

Received: March 23, 2010 Revised: May 03, 2010 Accepted: May 12, 2010