Glycoconjugate Journal (1998) **15:** 209–216

REVIEW

Computational carbohydrate chemistry: what theoretical methods can tell us

Robert J. Woods

Complex Carbohydrate Research Center, Department of Biochemistry, 220 Riverbend Road, University of Georgia, Athens, Georgia, 30602, USA

Computational methods have had a long history of application to carbohydrate systems and their development in this regard is discussed. The conformational analysis of carbohydrates differs in several ways from that of other biomolecules. Many glycans appear to exhibit numerous conformations coexisting in solution at room temperature and a conformational analysis of a carbohydrate must address both spatial and temporal properties. When solution nuclear magnetic resonance data are used for comparison, the simulation must give rise to ensemble-averaged properties. In contrast, when comparing to experimental data obtained from crystal structures a simulation of a crystal lattice, rather than of an isolated molecule, is appropriate. Molecular dynamics simulations are well suited for such condensed phase modeling. Interactions between carbohydrates and other biological macromolecules are also amenable to computational approaches. Having obtained a three-dimensional structure of the receptor protein, it is possible to model with accuracy the conformation of the carbohydrate in the complex. An example of the application of free energy perturbation simulations to the prediction of carbohydrate-protein binding energies is presented.

***Keywords*: conformational analysis, molecular dynamics, NMR, free energy, perturbation, oligosaccharide, polysaccharide, GLYCAM**

Introduction

The inherent chemical and biochemical properties of oligo- and polysaccharides, which enable their many functional roles *in vivo*, make them a most challenging class of molecules for conformational analysis. In order to probe the significance of many of their structural features, it would be desirable to introduce specific atomic alterations. However, the high number of chiral carbon atoms found in sugars places extreme demands on any but the most straightforward modifications. In contrast with commonplace protein biochemical methods, it is not possible to readily introduce point mutations in the sugar sequence. It is often the case that oligo- and polysaccharide structure-function relationships must be deduced from naturally available molecules. Consequently, few general rules have emerged regarding the influence of sugar composition on oligosaccharide properties.

Synthetic difficulties aside, the continuing interest in carbohydrates has resulted in the development of expression, isolation and purification systems that enable many naturally occurring glycoproteins, oligo- and polysaccharides to be produced in sufficient amounts to be fully analyzed. Concomitant advances in the sensitivities of many analytical methodologies have enabled analysis of smaller and more heterogeneous samples. Promising developments in alternative experimental techniques have been reported, including, atomic-force microscopy [1], fluorescence energy transfer [2] and circular dichroism – optical rotatory dispersion (CD/ORD) measurements [3].

Experimental advances, in turn, provide all the more incentive for theoreticians to devise valid models for data interpretation. Given the challenges associated with specific chemical and biochemical modifications of oligo- and polysaccharides, there is potential for modeling methods to provide valuable insight into the effects of putative structural modifications.

The conformational analysis of carbohydrates differs in several ways from that of other molecules, and proteins in particular. In conformational studies of proteins, much can be learned from modeling based on homology with known X-ray or nuclear magnetic resonance (NMR) structures. Rarely is a conformational analysis undertaken on a protein for which only the primary sequence is known, and for which no homology with other protein structures is available. And yet, this is typically the situation in a carbohydrate

0282–0080

conformational analysis. The lack of homology data for carbohydrates arises from the inherent flexibility of many glycans.

Although there exist relatively rigid oligosaccharides, such as the blood group determinants, it cannot be said that the properties of these compact branched structures are representative of the majority of oligosaccharides. Oligosaccharide flexibility deters homology modeling in at least two ways. The more flexible a molecule is, the harder it is to induce crystallization. Oligosaccharides behave in this regard more as peptides, than as proteins. That is not to imply that only flexibility is responsible for the resistance to crystal formation, other contributing factors include the presumably high degree to which water molecules coordinate to glycans, and the lack of strong inter-residue lipophilic or dipolar interactions.

If a carbohydrate is flexible, it may not have a single characteristic three-dimensional shape. In practice, many glycans appear to exhibit numerous conformations coexisting in solution at room temperature. With such a definition it is clear that a carbohydrates' conformation consists of both spatial and temporal components. With the exception of very modest atomic modifications [4, 5], homology modeling is not likely to be applicable to carbohydrate conformational analysis.

The recognition that a carbohydrate exists as a conformational ensemble in solution places limits on the extent to which solution-derived data, such as nuclear Overhauser intensities (NOE), may be employed to deduce the conformation. NOE data reflect only the time averaged properties of the molecule and rarely enable the constituent conformations to be determined directly. In relatively rigid systems it is possible to relate NOE intensities to inter-proton distances, and so derive a reasonable NMR-consistent conformation. However, as internal flexibility becomes significant, this practice may lead to the generation of virtual conformations, which represent structures that are never physically present [6]. Again in contrast to proteins, oligosaccharides frequently give rise to only two or three inter-residue NOE contacts. The paucity of NOE intensities often leads to an insufficient number of data points to uniquely determine the conformation of the oligosaccharide [7].

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, all contribute to the need for complementary experimental and theoretical methods.

The successful application of modeling methods depends greatly on the degree to which the inter-atomic properties of the molecules to be simulated can be approximated by a mathematical description. In molecular modeling, such a description is called a force field, and is frequently based on classical mechanical equations of force. Both carbohydrates and proteins may be described by the same force field, provided their pertinent inter-atomic properties are parametrized into the force field. Computational methods have had a long history of application to carbohydrate systems and their development in this regard has for many years been driven by questions relating to the interpretation of NMR data [8]. The unique demands of carbohydrate modeling led to the initial development of modeling protocols and force field equations that were specific for carbohydrates [9–15]. Although some of these early approaches continue to be useful for predicting the solution conformations of relatively rigid oligosaccharides [16–18], they were not, until recently, extendible to glycoconjugates, or carbohydrate protein complexes [19, 20]. With mounting interest in these areas, approaches to carbohydrate modeling have begun to appear, which are compatible with established protein modeling methods [21–27].

Carbohydrate modeling methods: a brief historical overview

Despite numerous proposed force fields for carbohydrate modeling, very few have gained widespread recognition and support. This is due, in part, to a lack of suitable experimental data to use as benchmarks in the testing of the force fields. In many cases, reasonable force fields have been proposed and applied, only later to have their weaknesses exposed as more data became available. This is a natural evolutionary process in the development of appropriate theoretical models and each diagnosed limitation in a force field provides far more insight into the model than does any given success. Nonetheless, the goal of an accurate and generally applicable modeling method, one that provides ensemble averaged data and that can be applied to glycoconjugates and complexes, has at times appeared elusive [28].

The tremendous diversity in force fields that are applied to modeling carbohydrate systems stems from a fundamental divergence early in the evolution of carbohydrate modeling. There are currently two general philosophies [8]. The first, and earliest, assumes that the conformation of the oligosaccharide is determined exclusively by van der Waals interactions, with the singular exception of the orientation of the ϕ-angle, for which a torsion potential that mimics the *exo*-anomeric effect is included. This model is known as the hard sphere e*xo-a*nomeric (HSEA) force field [29, 30]. The HSEA force field is similar to early force fields for proteins, such as the empirical conformation energy program for peptides (ECEPP) [31], in that it maintains individual residues in a predefined frozen geometry. The rigid residue approach was considered a good first approximation, and resulted in considerably more efficient calculations. However, the lack of molecular relaxation is well known to over-estimate repulsion energies in some conformations [12, 30], and with the improvement in computer performance, a rigid residue treatment is now unnecessary.

A particularly intriguing feature of the HSEA method is that it ignores completely electrostatic effects, such as

hydrogen bonding and dipolar interactions. Further, HSEA omits any treatment of solvent. Despite the very approximate nature of the HSEA force field, it has been in use for over 20 years. Pioneered by the efforts of Lemieux and Bock [29, 32], the HSEA method has been found to generate conformations that are frequently consistent with NMR data [13, 30]. Early HSEA-based predictions of the conformations of the blood-group antigenic determinants remain examples of the success of the HSEA method when applied to molecules that appear to exist predominately in one conformation in solution [12]. One approach to improving the ability of HSEA methods to treat flexible systems has been to introduce Monte Carlo (MC) sampling protocols [15, 18]. Flexibility issues notwithstanding, the successes of HSEA and related methods [14, 20] raise fundamental questions as to the roles played by solvation and electrostatics in determining carbohydrate conformations.

The second general class of force fields applied to carbohydrates is that in which the potential energy is expressed as a sum of energetic contributions from bond stretching, angle bending, torsional rotation and non-bonded interactions. The common macro-molecular force fields fall into this latter category and parameters for carbohydrates have been proposed for most of these, including AMBER [21, 23, 26], CHARMm [27, 33] and GROMOS [34, 35]. General purpose force fields, such as TRIPOS [36] and CVFF [37], have also been employed in carbohydrate modeling [25, 38–41].

An exception to this coarse division of carbohydrate force fields is the MM2 [42] (and MM3 [43] force field). MM2/3 has been extensively used in predicting the conformations of mono- and disaccharides, and is sometimes employed to refine the carbohydrate geometries derived from rigid-residue conformational searches. Unlike AMBER, CHARMm or GROMOS, both MM2 and particularly MM3 use relatively sophisticated mathematical expressions for many of the force field bonded terms, such as bond stretching and angle bending, and coupled cross-terms. MM2/3 has earned a reputation for accurately reproducing the fine details of molecular structures, including sugar ring puckering and bond length variations arising from the anomeric effect. One of the benefits of the mathematical forms of the MM2 and MM3 potential energy functions is that they provide very flexible descriptions of the potential energy surface. However, MM2/3 differs from the other macro-molecular force fields in an aspect that is perhaps more relevant to conformational analysis of oligosaccharides. In the calculation of electrostatic interaction energies, all other macro-molecular force fields employ partial atomic charges on each of the atoms in the molecule. In contrast, MM2/3 utilizes bond dipole moments rather than partial atomic charges. Although these two formulations are in principle equivalent, in practice, they differ in their abilities to treat correctly certain dipolar interactions, such as hydrogen bonding. In order to overcome weaknesses in this regard, MM2/3 incorporates an additional potential term to describe the attraction between two hydrogen bonded atoms. This approach may correct for structural inaccuracies, however, it is not consistent with any current model for water. Consequently, MM2/3 calculations have been limited to predominantly gas-phase, or miniature crystal, energy minimizations. It should be noted that MM2/3 has the ability to use partial atomic charges, however suitable charges have not yet been developed for its use. Moreover, MM3 now has the option to perform molecular dynamics, and so may be suitable for non-rigid residue MD simulations, which might be complimentary to the HSEA-based MC methods.

It should be noted that the incorporation of electrostatic effects into HSEA-type force fields (through the use of partial atomic charges) has been attempted [6], but led to a reduction in the accuracy of the calculations. In part, this reflects the general observation that inclusion of electrostatic effects, in the absence of explicit water molecules, over-estimates the significance of hydrogen bonds. This behavior in simulations performed *in vacuo* may be attenuated through the use of a bulk dielectric constant. The bulk dielectric constant of water is ~ 80 Debye, and provides significant charge screening. Notably, much smaller values of the dielectric constant have occasionally been shown to give better agreement with experimental data [44, 45]. The appropriateness of treating the bulk dielectric constant as an adjustable parameter has not been established.

Each of the macro-molecular force fields are similar and they may be described in general terms. Within these force fields the form of the energy equation is similar to that present in MM2/3 with the exception that simple harmonic potentials are employed, and that partial atomic charges, rather than bond dipoles, describe the dipolar and electrostatic interactions. Moreover, unlike either HSEA-type or MM2/3 models, the macro-molecular force fields were developed with the option to employ explicit water molecules in MD simulations. These conditions are extremely computationally demanding, however, they provide a model that bears an appealing similarity to the conditions of the experimental system. In particular, explicit solvation enables the MD simulations to be performed at constant pressure (ΔG) or constant volume (ΔA), at room temperature, and in the presence of counter-ions. These force fields were developed for application to proteins and nucleic acids and were not initially parametrized for sugars. Consequently, many researchers with an interest in carbohydrate modeling developed carbohydrate parameter sets. In retrospect it is clear that some of these parameter developments were not consistent with the protein, or even solvent, models associated with the relevant force fields. For example, if the partial atomic charge parameters are not computed with similar protocols for both the solvent and the solute, an internal inconsistency results. More recently, however, consistent parameter sets have been

developed for each force field. Despite the internal consistency, each parameter set may treat carbohydrates with varying levels of approximation. For example in the GLYCAM [21] parameter set for AMBER, we compute unique partial atomic charges for each atom in each sugar, whereas the majority of carbohydrate parameter sets assume that certain atom types will be equivalent and transferable between sugars. It is only through extensive simulations that the strengths and weaknesses of each model may become apparent.

Regardless of the mathematical formulation of a force field, its elegance, sophistication or simplicity, the ultimate test is its ability to accurately predict the conformational properties of a carbohydrate. That is, without biasing the calculation through the use of experimental data as restraints, does the simulation result in agreement with experimental data. To measure the success of any method, it is necessary to compare with data obtained under appropriate conditions. It is generally recognized that when solution NMR data are used for comparison, the simulation must give rise to ensemble-averaged properties. If data are obtained from crystal structure forms, then a simulation of a crystalline environment is appropriate [46, 47], and for the correct reproduction of subtle crystal motions, such as hydrogen bond flipping, ensemble averaged data are again necessary [48, 49]. Carbohydrate force field validation is one of the most essential and challenging features in carbohydrate modeling. In the case of MD simulations, significant strides toward this goal have been made. Although the application of MD simulations to carbohydrates is not yet routinely successful, the use of improved force fields, longer simulation times, and explicit solvation has been shown to lead to ensembles of oligosaccharide conformations that reasonably reproduce experimental NOE intensities [50, 51].

Application to larger systems, such as plant or bacterial polysaccharides, will necessitate not only advances in computer performance, but also in theoretical approaches. Conformational sampling over a nanosecond time-scale, which is the current limit of solvated MD simulations, is still far too short to adequately sample the conformational space available to a polysaccharide in solution. MC simulations may achieve considerably more sampling, and, to the extent that the absence of explicit water is in some way compensated for within the force field, these may be applicable to polysaccharides [52]. Crystalline environments may be adequately sampled on much shorter time-scales and MD simulations of solid state polysaccharides have already been useful in their structural and energetic interpretation [53]. Alternative techniques, such as Langevin dynamics have been applied elegantly in macroscopic simulations of DNA super-coiling, involving as many as 1000 base-pairs [54], and may be applicable to polysaccharide simulations. However, preliminary Langevin dynamics on small molecules related to sugars indicate that it may be necessary to include at least a first shell of explicit waters [55]. Longer time-scale motions may also be approached by Brownian dynamics. In these simulations the level of detail has to be sacrificed, however the benefits of obtaining a macroscopic model on the micro- or even millisecond time-scale would be significant [56, 57].

Modeling protein-carbohydrate interactions

Carbohydrate-protein interactions are of profound importance to the correct functioning of living organisms. Biological phenomena as diverse as cell adhesion, targeting and differentiation, tumor invasion and metastasis, immune response, bacterial adhesion to host cells, glycoprotein metabolic clearance, and protein folding, depend in part on carbohydrate recognition [58–63]. Many structural features, such as hydrogen bonding and lipophilic interactions, associated with carbohydrate-protein interactions have been deduced from experimental studies [64, 65], however, it is only recently that the energetics of these interactions have begun to be elucidated [66–69].

Despite the importance of carbohydrate recognition in biology, carbohydrate-protein interactions are relatively weak. The binding of carbohydrates to lectins increases with increasing contact area, from typical values of $K_a = 10^2$–$10^3\ M^{-1}$ for monosaccharides [70] to $K_a = 10^3$–$10^5\ M^{-1}$ for oligosaccharides [68, 69, 71]. In the case of large oligo- or polysaccharide-antibody interactions the binding constants ($K_a = 10^4$–$10^5\ M^{-1}$) [72, 73] are often weaker than typical protein-antibody interactions ($K_a \geqslant 10^7$–$10^8\ M^{-1}$) [74]. In the case of carbohydrate-protein interactions, the carbohydrate is highly accessible and this feature may allow it to make contact and interact with receptors, albeit weakly, where a tighter protein–protein interaction is not desired. An example is the homing of neutrophils to activated endothelial cells, a system characterized by transient carbohydrate-protein contacts [75, 76].

Binding affinities may display greater or lesser dependencies on enthalpy, entropy, solvation effects and stoichiometry for each class of receptor, and even for individual carbohydrate ligands interacting with the same receptor [66, 67, 69]. Experimental techniques, such as titration microcalorimetry, NMR spectroscopy and X-ray crystallography are able to provide detailed energetic and conformational information. However, each relies inherently on the availability of suitable samples.

Due to the issues presented in the previous section, the computation of the solution conformation of an oligosaccharide is extremely challenging. However, in many cases it is the binding of the carbohydrate to a lectin or antibody that is under examination. In these cases the sugar, at least in the vicinity of the binding region, may exist in a single, stable conformation. Thus, provided a reliable model for the protein is available, bound carbohydrate conformations may be derived through the application of less time-consuming approaches, such as distance geometry methods and

two-body docking strategies [25]. A further advantage offered by these systems is that the complex may crystallize, whereas the free oligo- or polysaccharide frequently may not [77, 78]. Carbohydrate-protein complexes offer the possibility of comparing computed and experimental oligosaccharide conformations and of determining the precise interactions that enhance binding. Unfortunately, theoretical quantification of the energies of these interactions is more difficult, again necessitating an accurately parametrized force field and consideration of solvent effects.

That solvent makes a considerable contribution to the energies associated with carbohydrate-protein interactions has been clearly shown by titration microcalorimetric measurements [69, 70, 79, 80]. A key aspect of these measurements is that they indicate that hydrophobic interactions between the protein and the carbohydrate are significant. This at first appears incongruous, given the hydrophilic character of all carbohydrates. However, areas of the sugar's surface arising from aliphatic protons generate hydrophobic patches. The stacking of these regions with aromatic amino acid residues is a characteristic feature of carbohydrate-protein complexes. Both microcalorimetric and NMR [81] data suggest that desolvation entropy may be significant, if not dominant, in complex formation. The enthalpic gain associated with the formation of a hydrogen bond between a sugar hydroxyl group and a polar protein atom may be as much as 6 kcal/mol. However, both of these atoms are fully solvated prior to complexation, and must lose solvent-hydrogen bonds, or desolvate at least partially, to form the complex. Thus, there may be little net enthalpic benefit to complex formation from hydrogen bonding. Bearing in mind the insight provided by the HSEA model, it is perhaps not surprising that the enthalpic contributions arising from strong electrostatic interactions might not be as relevant as anticipated. In contrast to the common assumption in protein-protein complexation, solvent reorganization is important for both hydrophobic and hydrophilic interactions in the formation of protein–sugar complexes. Moreover, the presence of water molecules that coordinate mutually to the carbohydrate and the protein may lead to over-estimates of the size of combining sites, when these estimates are inferred from epitope mapping strategies based on related carbohydrate ligands [82].

Consequently, the accurate calculation of interaction energies will necessitate an explicit treatment of solvent. A computational method, related to MD, called free energy perturbation (FEP) simulation is perhaps the most accurate method for computing relative binding affinities [83–86].

The technique of FEP simulation has been applied in the calculation of both absolute and relative binding free energies for protein–protein interactions and, *in well suited systems*, may predict energies that are in excellent agreement with experimental data [83, 87]. This powerful technique has been used recently to probe biotin–streptavidin binding [88], but has only recently been applied to oligosaccharide–protein interactions. A thorough review of the FEP methodology has recently been published, and the details of the approach will not be presented here. However, it should be noted that FEP benefits from the cancellation of systemic errors when applied to simulations aimed at determining relative binding energies. Computed relative binding energies are likely to be considerably more accurate than absolute binding energies. Moreover, FEP is only valid for extremely modest mutations.

We have applied MD and FEP simulations to the well-characterized oligosaccharide–protein complex between the O-antigen of *Salmonella paratyphi* B and its monoclonal antibody Fab and Fv fragments (Figure 1) [89]. The comparatively rigid nature of the Salmonella O-antigen facilitates accurate FEP calculation of the relative binding energies in this system. The results of preliminary FEP simulations for two reported analogs (Figure 2) of the wild-type ligand are presented in Table 1.

The complexes consisted of approximately 22500 atoms including over 6000 water molecules and required more than 500 h of real time per ligand on a dedicated IBM 590 (R6000) computer. Despite the excellent agreement with experimental data it should be noted that the values of ΔG are all close to zero. Consequently, further simulations employing this protocol will have to be performed before

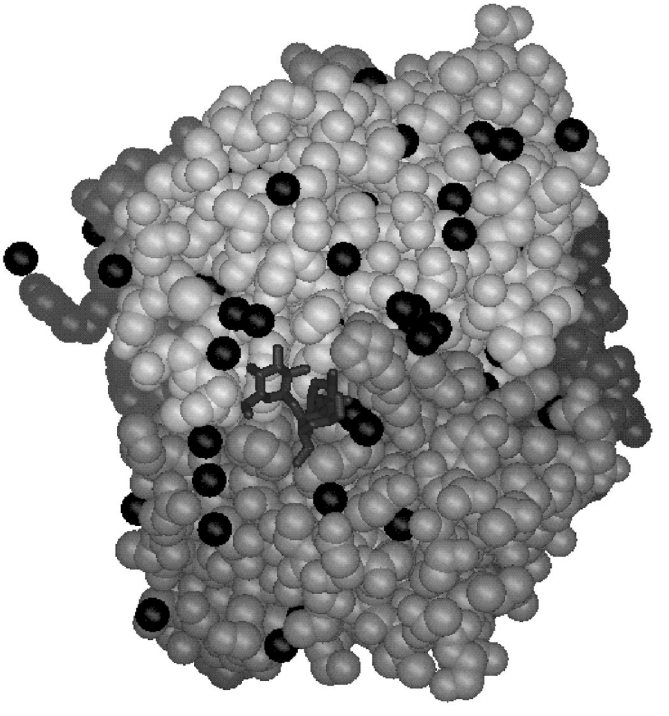

Figure 1. X-ray structure of the Fv-trisaccharide complex. The trisaccharide binds in a cavity between the VL (light gray) and VH (dark gray) domains. Waters of crystallization are illustrated in black.

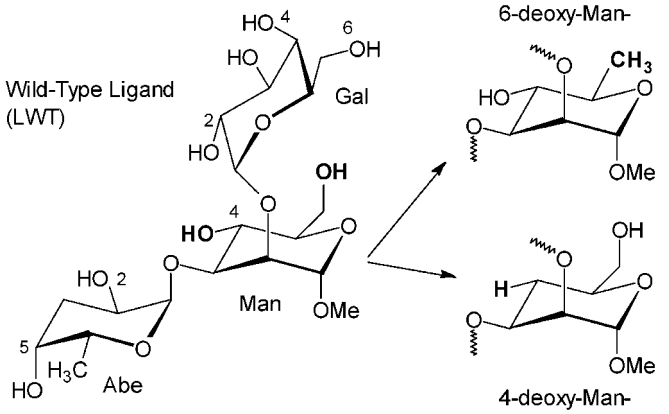

Figure 2. O-antigen mutations.

Table 1. Computed and experimental relative binding energies for the interactions between trisaccharides related to salmonella O-antigen and the antibody Fab fragment

Ligand	ΔΔG *Computed*	ΔΔG *Experimental*[a]
Wild-type	0.00	0.00
4-deoxy-Man-	0.37 ± 0.5	0.85 ± 0.2
6-deoxy-Man-	0.26 ± 0.5	0.50 ± 0.2

[a] From [73].

reaching an accurate assessment of the sensitivity of the method.

FEP simulations may be extremely powerful if the ligands differ only slightly, however, experimental carbohydrate–protein binding affinities are often determined as a function of increasing oligosaccharide size [90] and perturbations of that magnitude cannot be treated accurately by FEP methods.

Concluding remarks

The shape of an oligosaccharide may be defined in terms of the glycosidic torsion angles, however, internal motion of these angles results in an ensemble of conformations. If these motions occur over a narrow range of values (± 15 °) the average conformation of the carbohydrate may represent reasonably the ensemble behavior. Moreover, in some oligosaccharides the torsional motions may be correlated, leading to compensatory oscillations, which enable the oligosaccharide to maintain a stable conformation despite the presence of internal motions [51]. However, for larger variations, or for multiple distinct conformations it is inappropriate to think of the oligosaccharide conformation in static terms. Modeling methods such as MD simulations or MC sampling have the ability to account correctly for carbohydrate flexibility. Although branching may stiffen an oligosaccharide, general rules governing the relationship between sequence and conformational behavior remain to be determined. The role of water in influencing carbohydrate conformational properties remains elusive, however, the debate over explicit versus implicit solvation models may become moot as advances in computer performance enable the treatment of larger systems in shorter times.

Simulations may be applied in two general ways, either to generate structures that are consistent with known experimental data, or to predict structural properties. In the former case, NOE-derived distance restraints are often applied, and, provided there are sufficient restraints, a single unambiguous conformation may emerge. In such cases the restraints dominate the calculation, which becomes relatively independent of the particular force field. If the goal is to apply the force field in a predictive capacity, every attempt should be made to ensure that the chosen method is sufficiently accurate. Validation of computer predictions requires a close interplay between experimentalists and theoreticians alike.

Interactions between carbohydrates and other biological macromolecules are also amenable to computational approaches. With a three-dimensional structure of the receptor molecule in hand, it is possible to model with accuracy the conformation of the carbohydrate in the complex. The modeling may also be used to predict the relative binding energies for closely related ligands.

Several reviews of carbohydrate modeling have appeared in recent years and indicate clearly that many advances have been made both in methodology and in approaches to validation [8, 91–94]. Given the sophistication of these methods, it may be tempting to think of the simulation as being truly representative of the properties of the carbohydrate. But many aspects of the real system, such as proton exchange, pH effects, anomerization, induced-polarization and other quantum effects are absent. Despite these limitations modern computational approaches can provide insight into physical properties, which may not be accessible experimentally.

References

1 McIntire TM, Penner RM, Brant DA (1995) *Macromolecules* **28**: 6375–7.
2 Rice KG, Wu P, Brand L, Lee YC (1991) *Biochemistry* **30**: 6646–55.
3 Stevens ES (1994) *Biopolymers* **34**: 1395–401.
4 Fernández P, Jiménez-Barbero J (1994) *J Carbohydr Chem* **13**: 207–33.
5 Fernández P, Jiménez-Barbero J (1993) *Carbohydr Res* **248**: 15–36.
6 Cumming DA, Carver JP (1987) *Biochemistry* **26**: 6664–76.

7 Wooten EW, Edge CJ, Bazzo R, Dwek RA, Rademacher TW (1990) *Carbohydr Res* **203**: 13–17.
8 Woods RJ (1996) In *Reviews in Computational Chemistry* (Lipkowitz KB, Boyd DB, eds) pp 129–65. New York: VCH Publishers Inc.
9 Melberg S, Rasmussen K (1980) *Carbohydr Res* **78**: 215–24.
10 Bock K, Lemieux RU (1982) *Carbohydr Res* **100**: 63–74.
11 Bock K, Josephson S, Bundle DR (1982) *J Chem Soc Perkin Trans II*: 59–70.
12 Lemieux RU, Bock K, Delbaere LTJ, Koto S, Rao VS (1980) *Can J Chem* **58**: 631–53.
13 Lemieux RU, Bock K (1983) *Arch Biochem Biophys* **221**: 125–34.
14 Paulsen H, Peters T, Sinnwell V, Lebhun R, Meyer B (1984) *Liebigs Ann Chem* 951–76.
15 Peters T, Meyer B, Stuike-Prill R, Somorjai R, Brisson J-R (1993) *Carbohydr Res* **238**: 49–73.
16 Levery SB (1991) *Glycoconjugate J* **8**: 484–92.
17 Poppe L, Stuike-Prill R, Meyer B, van Halbeek H (1992) *J Biomol NMR* **2**: 109–36.
18 Weimar T, Meyer B, Peters T (1993) *J Biomol NMR* **3**: 399–414.
19 Engelsen SB, Rasmussen K (1993) *Int J Biol Macromol* **15**: 56–62.
20 Stuike-Prill R, Meyer B (1990) *Eur J Biochem* **194**: 903–19.
21 Woods RJ, Dwek RA, Edge CJ, Fraser-Reid B (1995) *J Phys Chem* **99**: 3832–46.
22 Imberty A, Pérez S (1994) *Glycobiology* **4**: 351–66.
23 Glennon TM, Zheng Y-J, LcGrand SM, Shutzberg BA, Merz KM Jr (1994) *J Comput Chem* **15**: 1019–40.
24 Grootenhuis PDJ, Haasnoot CAG (1993) *Molec Simulation* **10**: 75–95.
25 Imberty A, Hardman KD, Carver JP, Pérez S (1991) *Glycobiology* **1**: 631–42.
26 Homans SW (1990) *Biochemistry* **29**: 9110–18.
27 Ha SN, Giammona A, Field M, Brady JW (1988) *Carbohydr Res* **180**: 207–21.
28 Rutherford TJ, Partridge J, Weller CT, Homans SW (1993) *Biochemistry* **32**: 12715–24.
29 Bock K (1983) *Pure Appl Chem* **55**: 605–22.
30 Thogersen H, Lemieux RU, Bock K, Meyer B (1982) *Can J Chem* **60**: 44–57.
31 Momany FA, McGuire RF, Burgess AW, Scheraga HA (1975) *J Phys Chem* **79**: 2361–81.
32 Lemieux RU, Koto S (1974) *Tetrahedron* **30**: 1933–44.
33 Rasmussen K (1982) *Acta Chem Scand* **A36**: 323–7.
34 Koehler JEH, Saenger W, van Gunsteren WF (1987) *Eur Biophys J* **15**: 197–210.
35 Prabhakaran M, Harvey SC (1987) *Biopolymers* **26**: 1087–96.
36 Clark M, Cramer III R D, Van Opdenbosch N (1989) *J Comput Chem* **10** 982–1012.
37 Hagler AT, Lifson S, Dauber P (1979) *J Am Chem Soc* **101**: 5122–30.
38 Casset F, Hamelryck T, Loris R, Brisson J-R, Tellier C, Dao-Thi M-H, Wyns L, Poortmans F, Pérez S, Imberty A (1995) *J Biol Chem* **270**: 25619–28.
39 Dauchez M, Mazurier J, Montreuil J, Vergoten G (1992) *Biochemie* **74**: 63–74
40 Balaji PV, Qasba PK, Rao VSR (1993) *Biochemistry* **32**: 12599–611.
41 Asensio JL, López R, Fernández-Mayoralas A, Jiménez-Barbero J (1994) *Tetrahedron* **50**: 6417–32.
42 Allinger NL (1977) *J Am Chem Soc* **99**: 8127–34.
43 Allinger NL, Yuh YH, Lii J-H (1989) *J Am Chem Soc* **111**: 8551–66.
44 Asensio JL, Jimenez-Barbero J (1995) *Biopolymers* **35**: 55–73.
45 Widmalm G, Venable RM (1994) *Biopolymers* **34**: 1079–88.
46 French AD, Miller DP, Aabloo A (1993) *Int J Biol Macromol* **15**: 30–6.
47 Kouwijzer MLCE, van Eijck BP, Kroes SJ, Kroon J (1993) *J Comput Chem* **14**: 1281–9.
48 Koehler JEH, Saenger W, van Gunsteren WF (1988) *J Biomolec Struct Dynam* **6**: 181–97.
49 Kouwijzer MLCE, van Eijck BP, Kooijman H, Kroon J (1995) *Acta Cryst Sect B* **51**: 209–20.
50 Woods RJ, Edge CJ, Wormald MR, Dwek RA (1993) In *Complex Carbohydrates in Drug Research* (Bock K, Clausen H, Krogsgaard-Larsen P, Kofod H, eds) pp 15–36. Copenhagen: Munksgaard.
51 Woods RJ, Pathiaseril A, Wormald MR, Edge CJ, Dwek RA (1996) *Eur J Biochem* (submitted).
52 Stokke BT, Talashek TA, Brant DA (1994) *Macromolecules* **27**: 1124–35.
53 Heiner AP, Sugiyama J, Teleman O (1995) *Carbohydr Res* **273**: 207–23.
54 Ramachandran G, Schlick T (1996) *DIMACS Series in Discrete Methematics and Theoretical Computer Science* **23**: 215–31.
55 Widmalm G, Pastor RW (1992) *J Chem Soc* **88**: 1747–54.
56 Andrew SM, Thomasson KA, Northrup SA (1993) *J Am Chem Soc* **115**: 5516–21.
57 Northrup SH, Thomasson KA, Miller CM (1993) *Biochemistry* **32**: 6613–23.
58 Dwek RA (1995) *Science* **269**: 1234–5.
59 Hart GW (1992) *Curr Opin Cell Biol* **4**: 1017–23.
60 Sharon N, Lis H (1989) *Science* **246**: 227–46.
61 Rademacher TW, Parekh RB, Dwek RA (1988) *Ann Rev Biochem* **57**: 785–838.
62 Varki A (1993) *Glycobiology* **3**: 97–130.
63 Springer TA, Lasky LA (1991) *Nature* **349**: 196–7.
64 Quiocho FA (1986) *Ann Rev Biochem* **55**: 287–315.
65 Lemieux RU (1989) *Chem Soc Rev* **18**: 347–74.
66 Chervenak MC, Toone EJ (1994) *J Am Chem Soc* **116**: 10533–9.
67 Chervenak MC, Toone EJ (1995) *Biochemistry* **34**: 5685–95.
68 Mandal DK, Bhattacharyya L, Koenig SH, Brown RD III, Oscarson S, Brewer CF (1994) *Biochemistry* **33**: 1157–62.
69 Mandal DK, Kishore N, Brewer CF (1994) *Biochemistry* **33**: 1149–56.
70 Schwarz FP, Puri KD, Bhat RG, Surolia A (1993) *J Biol Chem* **268**: 7668–77.
71 Ramkumar R, Surolia A, Podder SK (1995) *Biochem J* **308**: 237–41.
72 Evans SV, Sigurskjold BW, Jennings HJ, Brisson J-R, To R, Tse WC, Altman E, Frosch M, Weisgerber C, Kratzin HD, Klebert S, Vaesen M, Bitter-Suermann D, Rose DR, Young NM, Bundle DR (1995) *Biochemistry* **34**: 6737–44.
73 Bundle DR, Eichler E, Gidney MA, Meldal M, Ragauskas A, Sigurskjold BW, Sinnot B, Watson DC, Yaguchi M, Young NM (1994) *Biochemistry* **33**: 5172–82.
74 Ito W, Kurosawa Y (1993) *J Biol Chem* **268**: 16639–47.

75 Varki A (1992) *Curr Opin Cell Biol* **4**: 257–66.
76 Butcher EC (1991) *Cell* **67**: 1033–6.
77 Weis WI, Drickamer K, Hendrickson WA (1992) *Nature* **360**: 127–34.
78 Bourne Y, Rougé P, Cambillau C (1992) *J Biol Chem* **267**: 197–203.
79 Williams BA, Chervenak MC, Toone EJ (1992) *J Biol Chem* **267**: 22907–11.
80 Sigurskjold BW, Bundle DR (1992) *J Biol Chem* **267**: 8371–6.
81 Jiménez-Barbero J, Junquera E, Martín-Pastor M, Sharma S, Vicent C, Penadés S (1995) *J Am Chem Soc* **117**: 11198–204.
82 Lemieux RU (1992) *Carbohydrate Antigens* **ACS Symposium Series # 519** 6–18.
83 Pearlman DA, Kollman PA (1989) In *Computer Simulation of Biomolecular Systems*: *Theoretical and Experimental Applications* (van Gunsteren WF, Weiner PK, eds) pp 101–19. The Netherlands: ESCOM.
84 Singh UC, Brown FK, Bash PA, Kollman PA (1987) *J Am Chem Soc* **109**: 1607–14.
85 van Gunsteren F (1989) In *Computer Simulation of Biomolecular Systems*: *Theoretical and Experimental Applications* (van Gunsteren WF, Weiner PK, eds) pp 27–59 The Netherlands: ESCOM.
86 Straatsma TP (1996) In *Reviews in Computational Chemistry* (Boyd DB, Lipkowitz K, eds) New York: VCH, pp 81–127.
87 Mizushima N, Spellmeyer D, Hirono S, Pearlman D, Kollman P (1991) *J Biol Chem* **266**: 11801–9.
88 Miyamoto S, Kollman P (1993) *Proteins* **16**: 226–45.
89 Woods RJ, Pathiaseril A (1996) (unpublished data).
90 Wessels MR, Pozsgay V, Kasper DL, Jennings HJ (1987) *J Biol Chem* **262**: 8262–7.
91 Peters T, Pinto BM (1996) *Curr Opin Struct Biol* **6**: 710–20.
92 Woods RJ (1995) *Curr Opin Struct Biol* **5**: 591–8.
93 Pérez S, Imberty A, Carver JP (1994) *Advan Comput Biol* **1**: 147–202.
94 French AD, Brady JW (1990) *Computer Modeling of Carbohydrate Molecules* (American Chemical Society Washington DC).

Received 9 September 1996, revised 25 January 1997, accepted 11 March 1997