Structure, Conformation, and Dynamics of Bioactive Oligosaccharides: Theoretical Approaches and Experimental Validations

Anne Imberty* and Serge Pérez

Centre de Recherches sur les Macromolécules Végétales (CERMAV-CNRS, affiliated with Joseph Fourier Université), 601 rue de la Chimie, BP 53, F-38041 Grenoble Cedex 9

Received April 19, 2000

Contents

* To whom correspondence should be addressed. Phone: +33 (0)4 76 03 76 36. Fax: +33 (0)4 76 54 72 03. E-mail: anne.imberty@cermav.cnrs.fr. URL: http://www.cermav.cnrs.fr/home_pages/imberty/imberty_GB.html.

I. Introduction

In nature, carbohydrates form an important family of biomolecules, as simple or complex carbohydrates, either alone or covalently linked to proteins or lipids. Most of the earlier studies on carbohydrates were centered on plant polysaccharides, such as cellulose, starch, pectins, etc., largely because of their wide range of industrial applications. More recently, the role of carbohydrates in biological events has been recognized[1−3] and glycobiology has emerged as a new and challenging research area at the interface of biology and chemistry. The glycoconjugates, i.e., glycolipids, glycoproteins, and glycosaminoglycans present at the cell surface, display diversity in glycosylation pattern between species which appear to be to be driven by evolutionary selection pressures[4] but which also can occur between cell types in the same organism. Modifications of cell glycosylation also occur during cell activation, inflammation, and cancer. Oligosaccharides are known to play structural and physical roles but are also specifically recognized by lectin receptors.[5] Of special interest are the carbohydrate-mediated recognition events that play a role in important biological phenomena involving cell−cell interaction such as fertilization, bacterial infection, inflammatory processes, cell growth, etc.

Determination of the three-dimensional structure of oligosaccharides and understanding the molecular basis of their recognition by receptors represent the main challenges of structural glycobiology. Elucidation of the three-dimensional structures and the dynamical properties of oligosaccharides is a prerequisite for a better understanding of the biochemistry of recognition processes and for the rational design of carbohydrate-derived drugs. Unfortunately, oligo-

10.1021/cr990343j CCC: $35.00
Published on Web 11/17/2000

Anne Imberty graduated in Biology from Ecole Normal Supérieure and University Pierre & Marie Curie in Paris. In 1984, she joined the Centre National de la Recherche Scientifique in Grenoble and received her Ph.D. degree in Chemistry in 1988 for studies concerning the crystal structure of starch. During her postdoctoral stay in Toronto in Jeremy Carver's group she studied protein–carbohydrate interaction and then joined Serge Pérez in Nantes. During five years, she worked in the group of Jean Villieras. Since 1996, she is back at CERMAV in Grenoble where she was awarded a senior research position in 1999. Her research interests include methodology development for molecular modeling of carbohydrates and protein/glycan interactions, structural studies of biologically active oligosaccharides, and nucleotide sugars and their interactions with proteins and particularly lectins and glycosyltransferases.

Serge Pérez was born in Périgueux, France. He received his Ph.D. degree in crystallography from the University of Bordeaux, France. As a postdoctoral fellow he spent one year at the Institute of Molecular Biology of the University of Oregon at Eugene and, subsequently, three years in the Department of Chemistry at the University of Montreal. He returned to France upon accepting a permanent position as a Junior Scientist of the CNRS (Centre National de la Recherche Scientifique) at CERMAV (Centre de Recherches sur les Macromolécules Végétales) at Grenoble, and he spent a sabbatical year working for Eastman Kodak in Rochester, NY. After working in Grenoble for 10 years, he moved on to Nantes to the Institut de la Recherche Agronomique, where he was awarded a position as Senior Research Scientist. In January 1996 he moved back to Grenoble to become Chairman of the CERMAV. His research interests span across the whole area of structural and conformational analysis of oligosaccharides, polysaccharides, glycoconjugates, and protein–carbohydrate interactions in solution and in the solid state. This includes interests in computational chemistry and molecular modeling, crystallography, NMR spectroscopy, along with the structure–function and structure–properties relationship.

saccharides, either in their free form or when part of glycoconjugates, are inherently difficult to crystallize and structural data from X-ray studies are sparse.[6] In solution, the flexibility of certain glycosidic linkages produces multiple conformations which coexist in equilibrium. The use of several spectroscopic methods, with appropriate time resolution, is necessary for analysis of the conformational behavior of such molecules.[7,8] Hence, procedures for molecular modeling of carbohydrates have been devised as an important tool for structural studies of these compounds. Since the pioneering work of Prof. Lemieux and co-workers,[9] various molecular modeling methods have been developed[10] and widely used for the determination of oligosaccharide conformations. The progress made in algorithms and computational power now allows for the simulation of carbohydrates in their natural environment, i.e., solvated in water or in organic solvent, in concentrated solution, or in the binding site of a protein receptor.

The last two decades have witnessed tremendous advances in the elucidation and understanding of the conformations and dynamics of oligosaccharides thanks to the technical developments in nuclear magnetic resonance (NMR), X-ray diffraction, and computer modeling methods. It is the aim of the present paper to review the significant contributions in the field, restricting the examples to some carbohydrate molecules that are considered to possess a significant biological function. We focus essentially on molecules belonging to the class of histo-blood-group antigens and to the glycosaminoglycan family for which a range of conformational characterizations has been reported, in turn providing a consistent understanding of the conformation and conformational changes that these molecules undergo as a function of their chemical and biological environment. Obviously other biologically active oligosaccharides and glycoconjugates have been reported. but none of these has been sufficiently well characterized, in terms of conformation, to be included, thus avoiding the risk of transforming this paper into a catalog. As for the different concepts and tools underlying the assessment of the structural and dynamics features of oligosaccharides, they will be presented using the most significant scientific contribution, whether the studied molecules belong to the classes of biologically active oligosaccharides defined above.

II. Computational Studies of Oligosaccharides

A. Structural Diversity and the Conformational Challenges

Carbohydrates have a potential information content that is several orders of magnitude higher than any other biological macromolecule. The diversity of carbohydrate structures results from the broad range of monomers (>100) of which they are composed and the different ways in which these monomers are joined. Thus, even a small number of monosaccharide units can provide a large number of different oligosaccharides, including branched structures, a unique feature among biomolecules. For example, the number of all possible linear and branched isomers of a hexasaccharide exceeds 10^{12}.[11] The complexity of the topology of oligosaccharides (and polysaccharides) requires the design of dedicated molecular building procedures that can rapidly convert the commonly used sequence information into a preliminary but reliable 3-D model. Particular procedures have been

designed for such tasks[12,13] using libraries of constituent monomers.[14]

Currently, the accepted vision is of oligosaccharides as flexible molecules containing several bonds about which there is free rotation. Therefore, they constitute a particularly challenging class of molecules for conformational analysis, both from the theoretical and experimental point of view. Carbohydrates are also thought to be especially difficult to model because of their highly polar functionality and the differences in electronic arrangements, such as the anomeric, *exo*-anomeric, and gauche effects, that occur during conformational and configurational changes. These effects have been extensively reviewed.[15,16] To address these issues, molecular modeling is required, and this is most effective when it is used in conjunction with diffraction methods, high-resolution NMR spectroscopy, and other spectroscopic methods. In this context, to perform realistic modeling, three basic questions should be answered: (1) What are the most appropriate force fields and concomitant parameters to use? (2) What is the most satisfactory and efficient way to travel through the conformational hyperspace? (3) What is the appropriate way to calculate, from a modeling study, the spectroscopic observables for which experimental data are available?

The low-energy conformers of a disaccharide can be estimated using molecular mechanics. In such compounds the global shape depends mainly on rotations about the glycosidic linkages, because the flexibility of the pyranose ring is rather limited and the different orientations of the pendant groups have a limited influence on the conformational space of the disaccharide. The relative orientations of saccharide units are therefore expressed in terms of the glycosidic linkage torsional angles Φ and Ψ which are defined as $\Phi = O5-C1-O-C'x$ and $\Psi = C1-O-C'x-C'(x-1)$ for a $(1\rightarrow x)$ linkage (Figure 1).

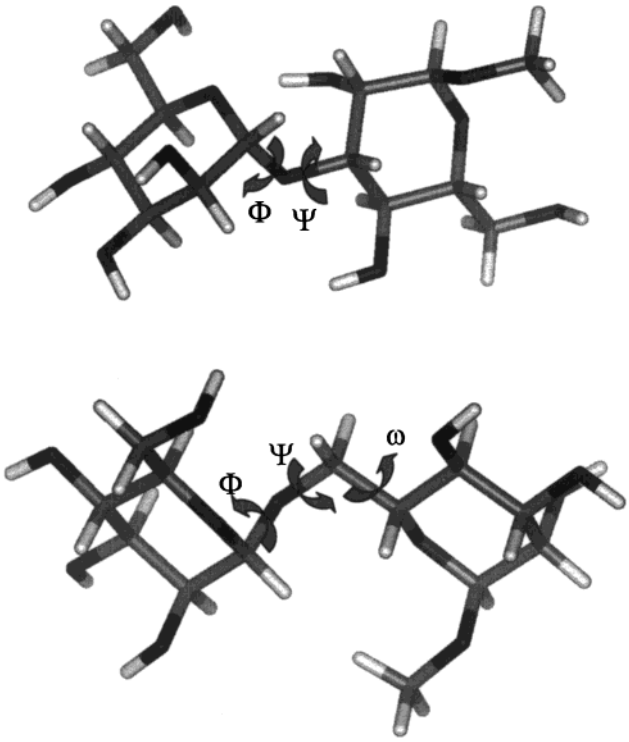

Figure 1. Graphical representation of two disaccharides along with labeling of the torsion angles defining the conformation of the glycosidic linkage. (top) The αMan(1−3)Man disaccharide. (bottom) The αMan(1−6)Man disaccharide.

B. Force Fields for Carbohydrates

Molecular mechanics potential-energy functions have been developed to describe a variety of systems, such as various small molecules, including the important case of water, simple organic compounds, proteins, and nucleic acids. Unfortunately, because the characteristics of a particular functional group may depend on the chemical environment, it is usually not possible to transfer molecular mechanics potential-energy parameters developed for a specific case to the description of the same group in a different environment. For this reason, potential-energy functions and parameter sets developed for proteins or for general organic molecules may not be appropriate for carbohydrate systems. Several carbohydrate potential-energy functions and/or parameter sets are available in the literature, and these have been used extensively in the past. The following force fields are widely used or have been designed especially for carbohydrates.

The MM2 and MM3 force fields are molecular mechanics force fields initially meant for hydrocarbons but now applicable to a wide range of compounds.[17−20] Tvaroska and Pérez published a modified version especially for oligosaccharides called MM2CARB.[21]

The GROMOS force field was developed for molecular dynamics simulations of proteins, nucleotides, or sugars in aqueous or apolar solutions or in crystalline form[22] and has been modified to include the *exo*-anomeric effect.[23]

The CHARMM force field is designed for the modeling (both molecular mechanics and dynamics calculations) of macromolecular systems.[24] Several revisions for carbohydrates have been proposed.[25,26] Kouwijzer and Grootenhuis redeveloped the CHEAT force field: a CHARMM-based force field for carbohydrates in which a molecule in aqueous solution is mimicked by a simulation of the isolated molecule.[27,28]

The AMBER force field was developed for simulations of proteins and nucleic acids.[29] A modification of this, for conformational analysis of oligosaccharides, was made by Homans.[30] Glennon et al.[31] and more recently Momany and Willet[32] presented an AMBER-based force field especially modified for α-(1→4) linkages. Woods et al. developed the GLYCAM parameter set for molecular dynamics simulations of glycoproteins and oligosaccharides that is consistent with AMBER.[10]

The AMBER* version used in the MacroModel package[33] has been expanded with carbohydrate parameters and validated by free-energy calculations on various simple sugars and disaccharides.[34]

The consistent force field (CFF), originally a molecular mechanics force field for cycloalkane and *n*-alkane molecules optimized on structural and vibrational data,[35] has been developed, in later ver-

sions, for other classes of compounds including carbohydrates.[36,37]

The SPACIBA program has been developed as a vibrational force field with particular emphasis on monosaccharides and oligosaccharides.[38]

The TRIPOS molecular mechanics force field is designed to simulate both biomolecules (peptides) and small organic molecules.[39] Additional parameters for conformational analysis of oligosaccharides, including sulfated glycosaminoglycan fragments and glycopeptides, were derived by Imberty et al.[14,40,41]

The CVFF and CFF force fields available from Biosym Technologies have been evaluated for modeling carbohydrates.[42] Recently, methods for deriving class II force fields[43] have been applied to carbohydrates and the parameters incorporated into the CFF force field.

The DREIDING force field, developed for the simulation of organic, biological, and main-group inorganic molecules, is one of the newer force fields in this list.[44]

The OPLS force field[45] has been expanded recently to include carbohydrates.[46]

The Merck Molecular Force Field (MMFF94) has been published.[47] It seeks to achieve MM3-like accuracy for small molecules in a combined "organic/protein" force field equally applicable to proteins and other systems of biological significance.

A comparison and chemometric analysis of 20 of these molecular mechanics force fields and parameter sets applied to carbohydrates has been performed.[48] The applications of these force fields and/or sets of parameters to a series of seven test cases provided a fairly general picture of the potential of these parameter sets to give a consistent image of the structure and energy of carbohydrate molecules. The results derived from a chemometric analysis (principal component analysis) produced a global view of the performances of the force fields and parameters sets for carbohydrates. The analysis (i) provided an identification of the parameters sets which differ from the bulk, (ii) helped to establish the relationship that exists between the different parameters sets, and (iii) provided indications for selecting different parameters sets for exploration of the force field dependency (or the lack of thereof) in a given molecular modeling study.

C. Exploring Hyperdimensional Space

Computational methods are applied extensively for the exploration of the conformational space of oligosaccharides.[49] The determination of conformational preferences of oligosaccharides is first performed by describing their preferred conformations on potential-energy surfaces as a function of their glycosidic torsional angles: Φ and Ψ. It is assumed that each glycosidic linkage displays a conformational behavior that is independent of the structural features in the molecule.

The Φ, Ψ space can be explored in a systematic way. Both torsions are sequentially rotated in small increments over the full 360° range. At each point of the grid the energy according to the force field in use is calculated. It is then possible to represent the energies of all the conformations available as a contour map in Φ, Ψ space. These contour maps enable graphical description of energy changes as a function of the relative orientation of the monosaccharides. They indicate the shape and position of minima, the routes for interconversion between conformers, and the heights of the transitional barriers. There are many different methods for calculating contour maps.

In the rigid residue or hard sphere potential surfaces approach, the constituent monosaccharides are assumed to be rigid with pendant groups fixed. As the Φ, Ψ values are changed, steric interactions occur between the pendant groups which are unable to relax. These steric interactions cause a rapid increase in energy. This effect is especially prominent in sterically crowded molecules. In addition, surveys of a large number of known crystal structures, together with and supported by semiempirical calculations, reveal small but important variations in pyranoid ring geometries and orientations of pendant groups with the Φ, Ψ values. These are dependent upon the anomeric and *exo*-anomeric effects and emphasize the need for a model to include bond length and angle degrees of freedom.

The strain produced by steric interactions inherent from rotation of monosaccharide residues is relieved by the inclusion of adjustments to bond length and angle by minimization of all degrees of freedom of the system (except Φ and Ψ) at each grid point. During minimization, pendant groups move to the nearest minimum downhill from the starting point. In the process of driving the molecule through unfavorable regions of the Φ, Ψ space, large steric interactions can sometimes cause pendant groups to overcome torsional barriers. This results in minimization to a different local well. This relaxed map describes a larger accessible potential-energy surface than rigid maps and leads to a lowering of the energy barriers between minima and a lower energy minimum far removed from the initial starting geometry.[50]

Whereas rigid residue maps represent a two-dimensional cross section of a $3N - 6$ dimensional surface, where N is the number of atoms, relaxed maps represent a larger cross-sectional window of a given potential-energy surface because they allow minimization of the internal coordinates (bond lengths, bond angles, and torsional angles) to local low-energy wells. However, as minimization will only lead to conformations 'downhill' from the starting structure, the torsional dimension where most conformational variation occurs is limited to only one orientational well. It is possible that rotation of pendant groups over torsional barriers could produce lower energy conformations at that point in the Φ, Ψ space. Ideally at each point in this space an investigation of all possible combinations of pendant group orientations is required (i.e., assuming that each pendant group can exist in each of the three idealized staggered orientations, $3n$ different conformations at each point in the Φ, Ψ space, where n is the number of pendant torsions). This results in 3^{12} (531 441) conformations for a simple disaccharide and 3^{19} (1.16×10^9)

conformations for the more complex disaccharide repeat of heparin.

Adiabatic maps attempt to represent the lowest energy of all possible pendant group orientations at each point in the Φ, Ψ space. On comparison with the corresponding relaxed maps, adiabatic maps are flatter, allow greater freedom about the glycosidic bonds, locate additional minima, and reduce the barriers between minima. At present there are several different methods for calculating adiabatic conformational maps: In the method most commonly used, the energy at each point in Φ, Ψ space for several starting orientations of the hydroxyl groups is evaluated systematically and the lowest energy for each point is used to generate the map. This can be very time-consuming, so such a systematic search is only possible for carbohydrates of limited size and flexibility.

Several procedures have been developed to scan the energy surface as a function of the two glycosidic angles in an efficient way. For example, the random molecular mechanics (RAMM)[51] grid method searches the orientation of pendant groups at each point in Φ, Ψ space. At each point, a random walk procedure uses 1000 steps for varying pendant group orientation and evaluating unrelaxed energies. Only the resultant lowest energy structure is optimized and accepted as the energy for that point in the Φ, Ψ space.

The prudent ascent method moves through the Φ, Ψ space in a way that is dependent on previous minimizations. Large steric interactions are minimized by dealing with the most favorable geometries first in a way similar to the local relaxed map. It makes use of inelastic deformations that decrease the energy by recalculating the energies of surrounding geometries using the new lower energy structure as the starting geometry. On average the energy for each point in the Φ, Ψ space is calculated twice.[52]

With the CICADA method (channels in conformational space analyzed by driver approach),[53] the potential-energy surface is explored using single-coordinate driving approach:[54] each selected torsion angle is driven with a concomitant full-geometry optimization at each increment (except for the driven angle). A comparative application of this type of heuristic search, with an exhaustive systematic search for disaccharides,[55] has established the reliability of the CICADA method. It displays several advantages over other conformational searches: (i) it has polynomial dependence of dimensions on computer time, in contrast to the systematic grid searches which have exponential dependence; (ii) the conformations found are free of artificial harmonic constraint potentials; (iii) it overcomes all barriers among families of conformations on the conformational hypersurface but spends almost all of its time in the essential highly populated areas; (iv) the inherent properties of the algorithm make rigorous optimization superfluous and provide good convergence behavior; (v) it provides low-energy conversion pathways that can be used for estimating adiabatic rotational barriers.

The Monte Carlo method is essentially a random search method. The method is more efficient for atomic or simple molecular systems than for complex (macro)molecular systems, since a random displacement in the latter case will generally lead to such distortions of a molecule so that the energy of a new configuration will usually be very high. Metropolis Monte Carlo methods have been applied to the conformational analysis of oligosaccharides.[56]

In molecular dynamics simulations, an ensemble of configurations is generated by applying the laws of motion to the atoms of the molecular system. The two major simulation techniques are molecular dynamics in which Newton's equations of motion are integrated over time and stochastic dynamics in which the Langevin equation for Brownian motion is integrated over time. Several algorithms have been developed for molecular dynamics simulations. Such simulations follow a system for a limited time. Physically observed properties are computed as the appropriate time averages through the collective behavior of individual molecules. For the results to be meaningful, the simulations must be sufficiently long so that the important motions are statistically well sampled. Experimentally accessible spectroscopic and thermodynamic quantities can be computed, compared, and related to microscopic interactions. It should be noted that molecular dynamics is severely limited by the available computer power. With presently available computers, it is feasible to perform a simulation with several thousand explicit atoms for a total time of up to a few nanoseconds. To explore the conformational space adequately, it is necessary to perform many such simulations. In addition, it may be possible that carbohydrate molecules undergo dynamical events on longer time scales. These motions cannot be investigated with standard molecular dynamics techniques.

It is important to recognize that most quantum mechanical and molecular mechanical procedures are designed to treat molecules in the isolated state. Omission of the effect of the environment from the calculation results in a neglect of the fraction of the energy contribution that arises from these interactions. For example, a carbohydrate in an aqueous or crystalline environment will usually form hydrogen bonds only to neighboring molecules, while the simulation of the molecule in vacuo is dominated by conformations with energetically favorable intramolecular hydrogen bonds.[28]

Several different approaches have been proposed to treat solvation effects.[57] In one of the simplest, the effect of the solvent is achieved by increasing the dielectric constant for calculations of electrostatic interactions or by the use of a distance-dependent dielectric constant. Unfortunately this affects all electrostatics. An alternative approach is to treat the solvent as a dielectric continuum. The conformational free energy of a given conformer in a particular solvent may be described as arising from the contribution of the energy of the isolated state and the solvation free energy.[58,59] A computational method that is very efficient is the use of the CHEAT95 force field, which is parametrized in such a way that the

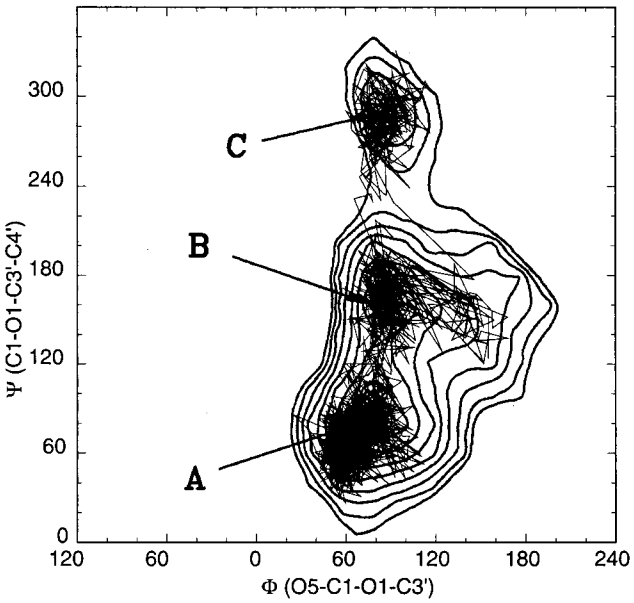

Figure 2. Potential-energy surface of the αGal(1−3)Gal disaccharide calculated with the MM3 force field as a function of Φ and Ψ.[143] The isoenergy contours are drawn by interpolation of 1 kcal/mol above the absolute minimum A. The broken line represents the history of Φ and Ψ torsion angles for a 2.5-ns molecular dynamics trajectory of the αGal(1−3)βGal(1−4)GlcNAc trisaccharide in water solution.

simulation of isolated carbohydrates mimics the behavior of the molecule in aqueous solution.[28] An alternative method, using Langevin dynamics simulation in which the water is simulated by a frictional model, has also been used for modeling disaccharides[60] and trisaccharides.[61]

At present the best approach is the inclusion of the environment in the simulation, viz. a molecular dynamics simulation with explicit water molecules or other surrounding molecules. By applying periodic boundary conditions a true but still very small system is simulated (Figure 2). Of course this is very time-consuming for an oligosaccharide in water, and only a few thorough investigations have been reported so far in practice for the disaccharides: maltose,[62,63] sucrose,[64] trehalose,[65] xylobiose,[66] neocarrabiose,[67] αMan(1−3)Glc,[68] and fragments of hyaluronan.[69,70]

D. Probing the Structures

The experimental data available for the atomic structures and energies of carbohydrates primarily comprise geometries and crystal packing in the solid phase, anomeric equilibria in solution, translational and rotational diffusion in solution, and molecular vibrations in the condensed phase. All of these experimental properties can be modeled, but they require consideration of a number of choices for modeling the environment and for sampling the statistically representative conformers that contribute to the experimental data.

1. Crystal Structures

More than 3600 crystal structure determinations of carbohydrates are now listed in the Cambridge Crystallographic Data Base. X-ray analysis provides the most accurate data concerning the conformation of a carbohydrate. Precise atomic coordinates are obtained, along with an explicitly defined environment. Although the crystalline state is often dismissed as irrelevant to biological processes, comparisons with crystal structures are among the most precise test of modeling available for carbohydrate molecules, provided that packing forces are taken into account. By molecular dynamics simulations of crystal structures, both force fields and methods can be validated.[71]

A more common way to use crystallographic data to test computer simulations is the superimposition of conformations found in crystal structures on a calculated potential-energy map. However, it should be kept in mind that the conformations found in crystals can be influenced by packing effects, so that they may differ from the preferred conformation(s) in aqueous solution and in vacuo. An interesting example of the problems that can arise with in vacuo calculations is given by sucrose. The MM3 calculation predicts a high potential energy (5.5 kcal/mol higher that the global minimum) for the conformation of the sucrose link found in raffinose. The simulation in water shows that this is an artifact of the MM3 force field; the calculations with the CHEAT95 force field perform much better in this respect. The lowest energy conformation is stabilized by an intramolecular hydrogen bond which cannot be formed in the raffinose conformation. Nevertheless, this conformation appears to be stabilized by surrounding molecules. An extensive study of the energy contributions of this glycosidic link showed that the problem was not due to the overlapping anomeric sequence, as suggested, but to a very high barrier for one of the torsions.

2. Conformational Behavior in Solution

In solution, the method of choice to study the three-dimensional structure of saccharides is NMR, through the parameters represented by chemical shifts, coupling constants, nuclear Overhauser effects (nOe), and also relaxation time measurements. The use of NMR experiments for the conformational analysis of oligosaccharides differs significantly from that of proteins. Isotopically labeled oligosaccharides are not readily available, and therefore, many heteronuclear NMR experiments are not easily applicable to the study of oligosaccharides. The conformational dependence of the carbon[72] and proton[73] chemical shifts is far from being understood; nevertheless, the use of chemical shifts of hydroxyl proton as conformational probes for studies in aqueous solution is emerging.[74] The magnitude of the torsion angles at the glycosidic linkages can be evaluated using a Karplus-type relationship established for vicinal proton−carbon[75] or carbon−carbon[76] coupling constants. Measurements of nOes can provide estimations of distances between protons that are close together in space. In contrast to proteins, distance geometry calculations for oligosaccharides are either nonapplicable or complicated because of the flexible nature of these molecules. NMR data alone, therefore, can rarely define the structural or dynamic properties of oligo-

saccharides unambiguously. In addition, relaxation time measurements give information on the mobility and the behavior of molecules in solution. Here again the assessment is complicated by the fact that the internal motions often occur on the same time scale as the overall tumbling time of the molecules. Therefore, interpretation of the experimental relaxation rates requires a dynamic model which is incorporated in the expression for the spectral densities. In the model-free approach,[77] internal motion is defined by only two variables: the order parameters, S^2, which describe its spatial restriction or amplitude, and τ_e, which is the corresponding effective correlation time. Despite its inherent complexity, such an approach is being used in the area of oligosaccharide conformational studies.[78] Additional problems are expected when the overall molecular tumbling time is anisotropic. Detailed studies in this respect are being performed on large oligosaccharides such as a synthetic heparin pentasaccharide.[79] These show that the effects of anisotropic motion generally cannot be neglected.

A major difficulty in the determination of the conformation of an oligosaccharide from NMR data arises from the flexibility of carbohydrates, especially flexibility arising from the glycosidic links. When multiple conformations are present in solution, NMR data will represent a time-averaged conformation. Since the geometrical parameters are usually related in a nonlinear way to the experimental data, these data can be very misleading. Consider, for example, an oligosaccharide in solution that occupies two distinct conformations, one with a relatively short distance between the protons at both sides of the link and one (which is preponderant) with a rather large distance. Since the measured nOe is an average, its value could easily lead the interpreter to a single nonexisting conformation. Even when it is known that two conformations are present, errors can easily be made since the predominant conformation will produce only a small contribution to the resulting nOe.

A very promising and yet simple approach lies in the combination of molecular dynamics simulation, with the explicit consideration of water molecules, to help resolve both the structural and dynamical features of oligosaccharides as seen by NMR spectroscopy. A nanosecond time scale molecular dynamics simulation of sucrose in aqueous solution was used to interpret NMR data from this disaccharide.[64] From these simulations the glycosidic heteronuclear coupling constants and NOESY volumes were calculated to be in good agreement with the experimental values, in contrast to the results derived from a rigid model. Because the internal motions of the disaccharide occurred on the same time scale as overall tumbling, a new motional model for spectral densities was devised. The calculated rotational tumbling time, translational diffusion coefficient, and radius of gyration also agreed well with experiments. In the MD simulations, the disaccharide exhibited significant conformational flexibility, as suggested from its relaxed conformation energy map, derived either from CHARMM[80] or from MM3.[81] In particular, the two inter-ring hydrogen bonds of the crystal did not persist in solution. Although they were present intermittently for a small fraction of the simulation time, most of the time they were exchanged for hydrogen bonds to solvent molecules. A study, utilizing rotating frame Overhauser enhancement spectroscopy (ROESY) experiments in supercooled water, indicated that direct exchange occurred between the hydroxyl groups belonging to the glucose and the fructose moiety, respectively.[82] It was concluded that a corresponding intramolecular hydrogen bond exists transiently in sucrose in solution. This does not necessarily contradict the above results because the experimental conditions in supercooled solutions certainly represent an extreme situation.

In recent years ^{13}C isotopic enrichment has proved to be a way for assessing new experimental NMR parameters not previously measurable.[83] For oligosaccharides the most interesting advances arose from the combination of isotope labeling with alignment of the molecules in strong magnetic fields.[84] Measurement of C−H (or H−H) residual dipolar couplings determined in dilute aqueous liquid-crystal media can complement or replace nOes for the assessment of the average conformation of oligosaccharides.[85,86] In some cases of protein/oligosaccharide interaction, the increase in dipolar coupling constant arising from ligand binding and the alignment in the magnetic field can be evaluated to provide data for the binding conformation of the oligosaccharide.[87]

There are not many experimental means, other than NMR, that are suitable for probing carbohydrate conformations and evaluating calculated potential-energy surfaces. However, the optical activity of saccharides depends on their chemical composition, configuration, and conformation. Optical rotations can be calculated from 3D structures using the semiempirical theory of Stevens and Sathyanarayana.[88] These studies result in the location of preferred regions in conformational space rather than in determination of well-defined energy minima. Numerous disaccharides, including sucrose[89] α(1−2) and α(1−3) linked mannobiose[90,91] and rhamnobiose,[92] have been investigated so far. For oligosaccharides, the technique is not widely applicable but is a useful complement to NMR methods.

III. Structure of Oligosaccharides: What Can We Learn from Crystal Structures?

A. Crystal Structures of Oligosaccharides

The Cambridge Structural Data Base (CSDB) contains over 160 000 entries in the form of structural data related to geometry, configuration, conformation, and packing of molecular crystals for organic and metal−organic compounds.[93] As a consequence of the automation of experimental measurements along with the development of crystal structure analysis for noncentrosymmetric space groups, there has been a significant increase in the number of reported crystal structures of carbohydrates between 1980 and 1995. The number of entries dealing with carbohydrate crystal structures in the CSDB that can be considered as structurally informative in terms of conformations and configurations

Table 1. Crystal Structure of Di- and Trisaccharides Related to Oligosaccharides of Biological Interest

name	formula	ref
Lewis X trisaccharide	βGal(1−4)[αFuc(1−3)]βGlcNAC-*O*-Me	95
blood-group B trisaccharide	αFuc(1−2)βGal(1−3) βGlcNAc-*O*-Me, $9H_2O$	94
N-glycan fragment	αMan(1−3)βMan(1−4)GlcNAc	97
H-type 1 disaccharide	αFuc(1−2)βGal, $^1/_2$ H_2O	96
mannobiose	αMan(1−2)αMan-*O*-Me	100
galabiose	αGal(1−4)Gal	104
N-acetyllactosamine	βGal(1−4)αGlcNAc, H_2O	98
lactose	βGal(1−4)αGlc, H_2O	107
	βGal(1−4)αGlc, $CaCl_2$ $7H_2O$	105
	βGal(1−4)βGlc	106
chitobiose	βGal(1−4)βGlcNAc	99
KDO disaccharide	αKdo(2−8) αKdo-*O*-allyl, $2Na^+$ H_2O	108

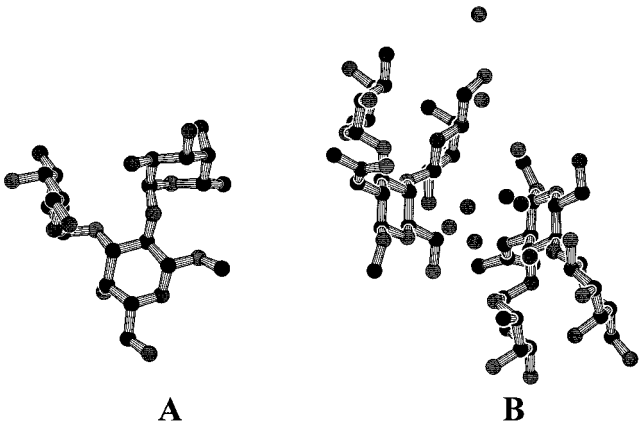

Figure 3. Asymmetric unit content of the crystal structure of (A) blood-group A trisaccharide[94] and (B) hydrated Le^x trisaccharide.[95] Graphical representations are drawn with PLATON.[247]

amounts to 3600. Among these there are 55 crystal structures of unsubsituted disaccharides and a dozen crystal structures of the fully (or almost fully) acetylated disaccharides. In the case of larger compounds, there are 17 trisaccharide crystal structures and only 4 tetrasaccharides.

Among these crystal structures, no more than 10 disaccharidic fragments of glycoconjugates and three trisaccharides are present (Table 1). The structures of histo-blood-group B[94] and Lewis x[95] trisaccharides have been obtained recently (Figure 3) together with H-type disaccharide.[96] As for the *N*-glycans, one trisaccharide[97] as well as some of the constituent disaccharides[98−100] have been crystallized. Structural data are also available for the *N*-linkage between GlcNAc and asparagine, both for small model compounds[101,102] and for glycoproteins.[103] Fragments of glycolipids such as galabiose[104] and lactose[105−107] have also been crystallized. The KDO disaccharide[108] is the only representative of the acidic oligosaccharide family. Very few structural data have been gathered on glycosaminoglycan structures. Only one disaccharide crystal structure has been solved,[109] and the only relevant sulfated structures obtained relate to monosaccharides (see refs 110 and 111 and references therein).

B. Oligosaccharides in Protein Crystal Structures

Since the beginning of the 1990s, an increasing number of crystal structures have been reported for glycoproteins and protein−carbohydrate complexes. The resolution of the first reported structures was rarely sufficient to provide reliable conformational information. However, significant and rapid progress, arising from the use of synchrotron radiation, is being made toward the provision of highly resolved structures. Among proteins that interact noncovalently with carbohydrates, lectins bind mono- and oligosaccharides reversibly and specifically while displaying no catalytic or immunological activity. More than 200 crystal structures of lectins have been solved, most of them as complexes with carbohydrate ligands. From a database of three-dimensional structures of lectins (http://www.cermav.cnrs.fr/databank/lectine), Table 2 has been compiled to illustrate the extraordinary wealth of information which is available regarding interaction between lectins and oligosaccharides. Only glycan structures at least as large as trisaccharides have been included in Table 2.

Many of these complexes involved biologically important oligosaccharides for which no structural information was available. Cocrystallization of oligosaccharides with lectins therefore appears to be the method of choice for the study of the conformation of glycan moieties such as (i) the sialic acid containing oligosaccharides (sialyllactose, sialoglycopeptide etc.),[112−116] (ii) important antigens (histo-blood-group antigen such as Lewis antigens and their sialylated or sulfated forms),[117,118] (iii) moieties of glycolipids such as GM1 and GM3,[119−122] and (iv) oligosaccharides belonging to N-linked glycans, either of the oligomannose type[123−129] or of the complex type.[114,130−135]

Whereas the study of the conformations about the glycosidic torsion angles indicated somewhat limited flexibility in molecular crystals (vide infra), a different picture emerges for the flexibility of oligosaccharides interacting with lectins. Figure 4 depicts the distributions of glycosidic torsion angles within three disaccharide segments: αMan(1−3)Man, βGlcNAc(1−2) Man, and αNeuAc(2−3)Gal. In the case of the αMan(1−3)Man segment, the observed conformations are located essentially around a Φ value of 80°, with an excursion in the vicinity of 140°. Perhaps more interesting is the observation that a remote low-energy area (located at $\Phi = 90°$ and $\Psi = 310°$) can be occupied, as observed in the crystalline complex between *Lathyrus ochrus* and a biantennary glycan.[130] The study of the dispersion of conformations observed for the disaccharide segment βGlcNAc(1−2)Man provides another illustration of the occurrence

Table 2. Crystal Structures of Complexes between Oligosaccharides and Lectins and Other Protein Receptors

glycan	protein	PDB code and ref
	oligomannose fragments from *N*-glycan	
αMan(1−3)βMan(1−4)GlcNAc	*Lathyrus ochrus* isolectin I	1LOG[123]
αMan(1−6)[αMan(1−3)]αMan	Concanavalin A	1CVN[126] 1ONA[127]
	Dioclea grandiflora lectin	1DGL[128]
mannopentaose	*Galanthus nivalis* agglutinin	1JPC[125]
phosphorylated mannopentaose	*Bos taurus* mannose 6-P receptor	1C39[129]
biantennary glycopeptide	*Rattus rattus* mannose binding protein A	2MSB[124]
complex fragments from *N*-glycan		
pentasaccharide from *N*-glycan	concanavalin A	1TEI[135]
biantennary oligosaccharide	*Bos taurus* galectin-1	1SLA, 1SLB,1SLC[133]
	Lathyrus ochrus isolectin I	1LOF[130]
glycopeptide	*Lathyrus ochrus* isolectin II	1LGC[132]
N2 fragment of lactotransferine		1LGB[132]
sialoglycopeptide	wheat germ agglutinin I	2CWG[131]
disialylated oligosaccharide	murine polyomavirus coat protein	1SIE[114], 1VPS[134]
	fragments of ABH and Lewis histo-blood-group oligosaccharides	
fucosyllactose	*Ulex europaeus* isolectin II	1QOT[154]
Lewis B	*Griffonia simplicifolia* isolectin IV	1LED[117]
Lewis Y		1GSL[117]
3′-sialyllactose	wheat germ agglutinin I	1WGC[112]
3′-sialyllactose	wheat germagglutinin II	2WGC[112]
3′-sialyllactose	*Maackia amurensis* leukoagglutinin	1DBN[116]
3′-sialyllactose	*Mus musculus* sialoadhesin	1QFO[115]
3′-sialyllactose	influenzae virus hemagglutinin	1HGG[113]
3′-sialyllactose	murine polyomavirus coat protein	1SID[114]
3'sialyl Lewis X	*Rattus norvegicus* MBP-A (CL-K3)	2KMB[118]
3′-sulfo Lewis X		3KMB[118]
4′-sulfo Lewis X		4KMB[118]
	fragments of glycolipids	
GM1 pentasaccharide	*Vibrio cholerae* cholera toxin	1CHB[119], 2CHB, 1CT1[120]
receptor GB3	*Escherichia coli* verotoxin-1	1BOS[122]
GM3 trisaccharide	*Staphylococcus aureus* enterotoxin B	1SE3[121]
	fragment from glycosaminoglycans	
pentasaccharide from heparin	human antithrombin-III	1AZX[138]
heparin tetrasaccharide	basic fibroblast growth factor	1BFB[136]
heparin hexasaccharide		1BFC[136]
heparin hexasaccharide	acidic fibroblast growth factor	1AXM, 2AXM[137]

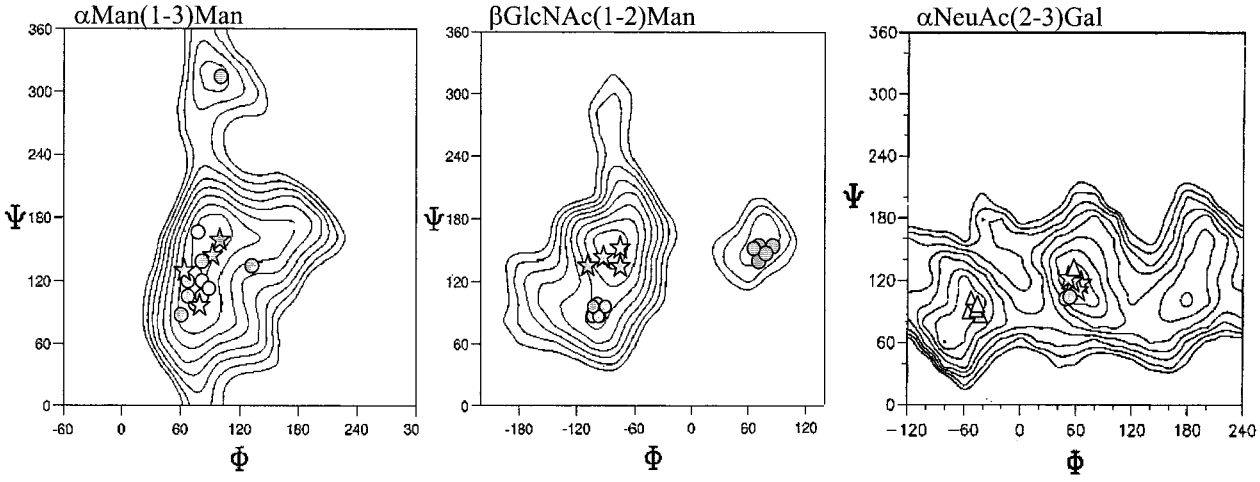

Figure 4. Potential-energy maps of three disaccharide fragments constituting *N*-glycan oligosaccharides. The energy maps have been calculated with the MM3 force field; the isoenergy contours are drawn by interpolation of 1 kcal/mol above the absolute minimum. Φ and Ψ are defined as O-5−C-1−O-1−C′-*x* and C-1−O-1−C′-*x*−C′-x_{+1} (O-6−C-2−O-2−C′-3 and C-2−O-2−C′-3−C′-4 for αNeuAc(2−3)Gal). Conformations observed in crystal structures of lectin/oligosaccharide complexes have been reported for the corresponding energy maps using circle, star, and triangle symbols, respectively, for complexes with plant, animal, and microorganism lectins.

of conformations in a remote energy well of the potential-energy surfaces. The location of this well is 120° away from what would correspond to the stable conformation driven by the *exo*-anomeric effect for an equatorial type of linkage. Such examples are observed in crystalline complexes involving the isolectin II of *Lathyrus ochrus*, complexed with high molecular oligosaccharides such as a biantennary octasaccharide,[130] a glycopeptide, or a N2 fragment of lactotransferrin.[132] The αNeuAc(2−3)Gal offers an

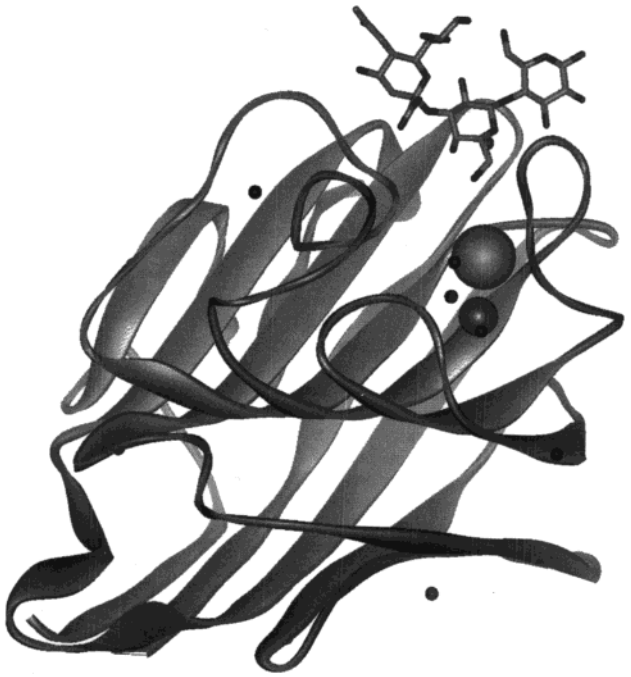

Figure 5. Graphical representation of the monomer of *Maackia amurensis* leukoagglutinin complexed with sialyllactose as observed in the crystal structure (code PDB 1DBN).[116] Protein, carbohydrate, and cations are displayed using ribbon, stick, and space-filling representations, respectively, using the WebLab ViewerLite program.[248]

extreme case of conformational flexibility as this can be 10 times more flexible than the other disaccharides. Here again, the conformation corresponding to the establishment of the *exo*-anomeric effect, with Φ at about 60°, is adopted in the binding site of several proteins including wheat germ agglutinin,[112] influenza virus hemagglutinin,[113] and murine sialoadhesin.[115] The same conformation is observed for sialyllactose cocrystallized with a lectin from *Maackia amurensis*, the only legume lectin that is specific for sialylated oligosaccharide (Figure 5).[116] Conformational stabilization due to the *exo*-anomeric effect can be easily overridden as exemplified by the occurrence of several low-energy conformations having Φ in the vicinity of −60°, as observed for the GM1 pentasaccharide in the combining site of cholera toxin.[120]

The structures of three proteins interacting with heparin fragments[136−138] are also included in Table 2 since, even if these are not usually included in the lectin families, these complexes yield very interesting data in terms of structural biology. The conformation of heparin fragments when bound to protein receptors is discussed in another paragraph of the present review.

IV. Histo-Blood-Group Oligosaccharides

The human histo-blood-group ABH(O) systems were the first major human alloantigens to be identified. The carbohydrate nature of the A, B, H, and Lewis antigens was established in the 1950s (see review by Kabat[139] and Watkins).[140] Figure 6 shows a schematic representation of current histo-blood-group antigens although additional rare types also exist.[141]

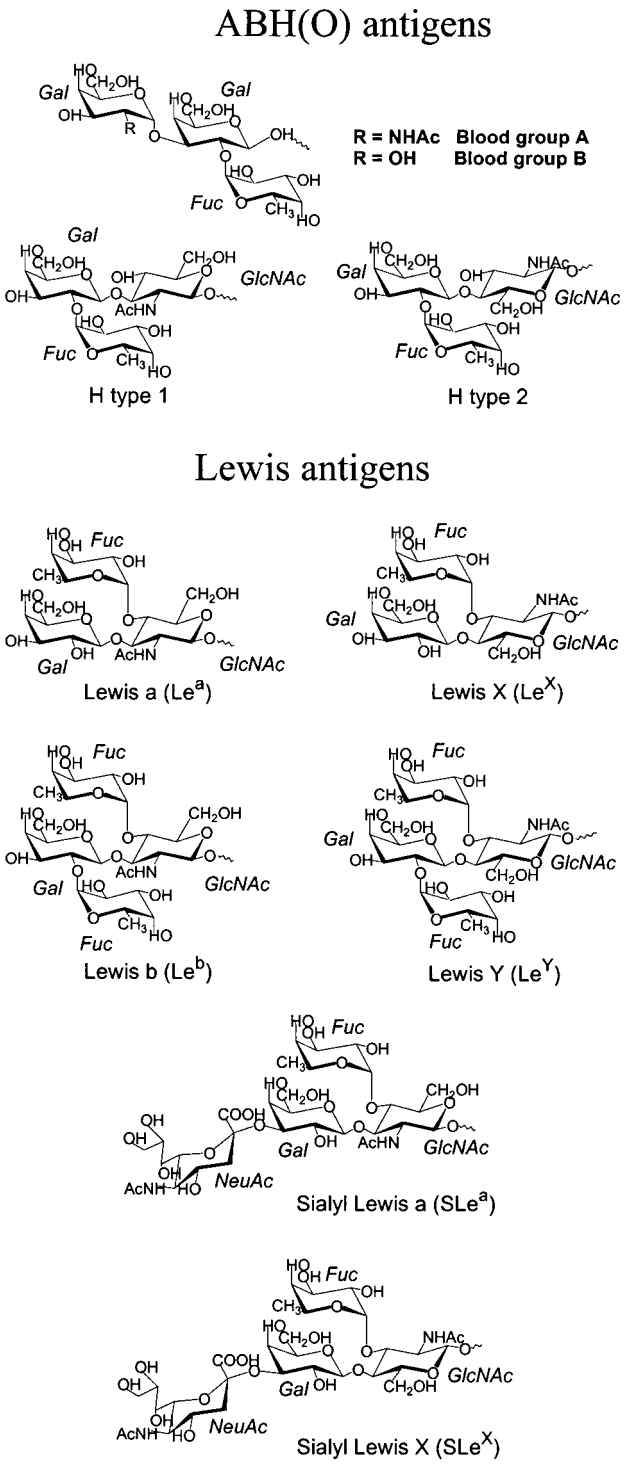

Figure 6. Schematic representation of histo-blood-group oligosaccharides.

The oligosaccharidic epitope of ABH(O) and Lewis histo-blood groups have been the subject of many structural investigations. The energy maps of all of the disaccharide fragments have been established by molecular mechanics methods.[142,143] In addition to the classical NMR and molecular modeling studies, some of the oligosaccharides have been crystallized. Furthermore, the "bioactive conformation", i.e., the one adopted by the glycan when interacting with a protein receptor, has been assessed either by protein crystallography or by use of the transferred nOe method.

A. Blood-Group A, B, and H Oligosaccharides in Different Environments

1. Theoretical Studies

Early theoretical and experimental studies on ABH oligosaccharides predicted them to be rigid molecules.[9,144] However, when systematic conformational search methods were applied to these molecules,[143,145] different conformational behavior was predicted to occur, depending on the nature of the oligosaccharides. While H type 1 and H type 2 trisaccharides were predicted to adopt one main low-energy conformation, the conformational analysis, when applied to A and B blood-group determinants, demonstrated much higher flexibility. The tetrasaccharide determinants of blood-group A type 1, A type 2, and B type 2 can occupy three families of conformations of which two are highly populated. The differences between these conformations are due mainly to the H-type trisaccharide, the terminal αGal(1−3), or αGalNAc-(1−3) linkage being less flexible. The same conclusion has also been reached about the higher flexibility of the αFuc1−2Gal moiety by means of a molecular dynamics study.[146]

2. Solution Conformational Behavior

The first NMR studies performed on H trisaccharides in solution led to the conclusion that these compounds were rigid.[147] It should be borne in mind that these studies usually use NMR data as constraints that are applied to the energy minimizations or to the molecular dynamics trajectories so that the identification of one conformational family only is built into the results. In recent studies conducted on analogues of H-type trisaccharides such as αFuc(1−2)βGal(1−3)βGalNAc (H type 4),[148] α(Fuc1−2)βGal-(1−4)βGlc (H type 6),[149] and αFuc(1−2)lactitol,[150] NMR data were not used as constraints; this allowed the identification of two conformational families. In every case the flexibility of the compound arose from the Ψ angle of the αFuc1−2Gal linkage which can adopt two values differing by about 80°.

Blood-group A (and blood-group B) oligosaccharides contain a terminal nonreducing Gal (GalNAc) residue that may interact with the fucose, thus limiting its conformational freedom. In the late 1970s, pioneering ^{1}H NMR work, performed at 270 MHz, had already identified a nOe contact between H-1 of Fuc and H-3 of GalNAc in the blood-group A trisaccharide[9] Additional NMR studies on blood-group A oligosaccharides[151] and, more recently, on blood-group B trisaccharide[94] identified short distances between the anomeric proton of the fucose residue and H-3 and H-5 protons of the αGalNAc (or αGal) residue. The occurrence of two such short distances (about 3 Å) between two noncovalently linked residues led to the conclusion that blood-group A and blood-group B trisaccharide are relatively rigid in solution.

However, these conclusions were based on a qualitative analysis of the nOe data and did not take into account the possibility that other conformations might participate in an equilibrium. When a quantitative comparison between the experimental and calculated nOe values of blood-group A trisaccharide was performed, a different conclusion was reached.[152] A thorough study by molecular mechanics predicted the occurrence of the two low-energy conformations displayed in Figure 7. Experimental interresidue nOe build-up curves were compared to theoretical curves calculated from each conformational family or from a population consisting of a 50/50 mixture of each conformation. A possible conclusion is that the calculated interresidue nOes are smaller than expected for conformation A and larger for conformation B (displayed in Figure 7). The best agreement is obtained when both families are given the same weight in the calculations, suggesting that this trisaccharide is flexible in solution.

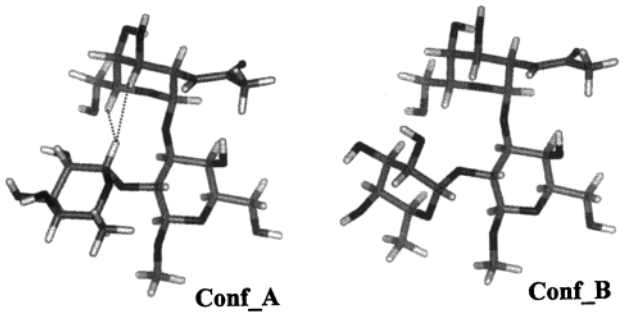

Figure 7. Two solution conformations of blood-group A trisaccharide.[152] Dotted lines represent the major nOes observed between the two nonadjacent carbohydrate rings.

3. Conformations Observed in the Solid State

It was only very recently that the H-type disaccharide was crystallized[96] and, shortly after, that the crystal structure of the blood-group B trisaccharide[94] solved. Interestingly, the αFuc(1−2)βGal linkage adopts different conformations in the two crystal structures. As a fragment of blood-group B trisaccharide, the observed (Φ,Ψ) values are (−66°, −91°), i.e., 52° and 28° for Φ_H and Ψ_H, respectively. This conformation exhibits short contacts between the anomeric proton of Fuc and the nonreducing αGal residue. On the other hand, in the H-type disaccharide crystal,[96] the same glycosidic linkage adopts (Φ,Ψ) values of (−93°, −175°), i.e., 28° and −57° for Φ_H and Ψ_H, respectively. Hence, in the solid state, the αFuc(1−2)βGal linkage can adopt two conformations that differ by almost 90° in Ψ. These two different conformations correspond to those predicted previously by molecular dynamics[146] and molecular mechanics[143] and also to the two families of conformation of the αFuc(1−2)βGal moiety that were detected in the previously mentioned NMR study of blood-group A trisaccharide.[152]

4. Conformations Observed in the Protein-Bound State

Fucosyllactose is the only related ABH blood-group oligosaccharide that has been cocrystallized with a protein receptor. Indeed, isolectin II of *Ulex europaeus* is classified as a GlcNAc/chitobiose-specific lectin, but its highest affinity is for H-type 2 trisaccharide and related H-type 6 trisaccharide or fucosyllactose (αFuc(1−2)βGal(1−4)Glc).[153] In the recently determined crystal structure of the complex between UEA-II and fucosyllactose,[154] the trisaccha-

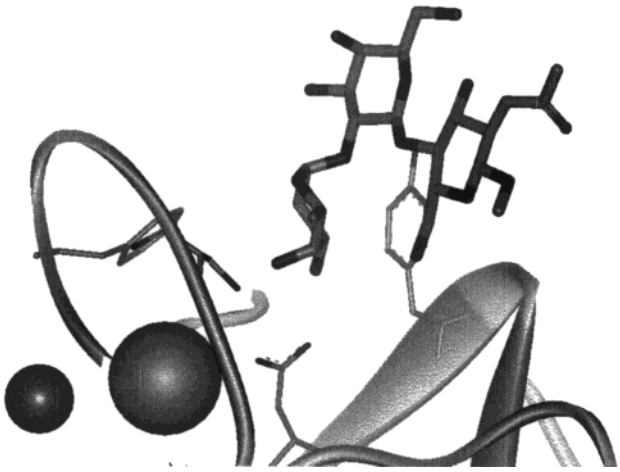

Figure 8. Model of the H-type 2 trisaccharide in the binding site of isolectin 1 from *Ulex europaeus*. The spheres represent the cations Ca^{2+} and Mn^{2+}.

ride adopts a conformation close to that of lowest energy. The αFuc(1−2)βGal linkage displays (Φ,Ψ) values of (−123°, −132°), somewhat between the two low-energy conformations previously reported. In fact, according to the (Φ,Ψ) energy map,[143] this conformation is located in the low-energy plateau that contains these two minima.

In the absence of crystal structures, other methods can provide information on the conformation of ligands bound to protein receptors.[155] Transferred nOes can be monitored, provided that the exchange between the complexed and uncomplexed states is sufficiently fast, a condition that appears to be satisfied frequently by sugar-binding proteins.[156] Interaction of blood-group A trisaccharide with the lectin from *Dolichos biflorus* seeds was characterized by negative transferred nOes that corresponded to one conformation only of the trisaccharide. Theoretical transferred nOes, calculated using an extended procedure for the complete relaxation matrix analysis of multispin exchanging systems,[152] confirmed that only one conformation is selected upon binding to the lectin. This result was not obvious from molecular modeling alone, since both solution conformations could be docked satisfactorily in the binding site of the *Dolichos biflorus* lectin.

Starting from the modeled structure of isolectin I from *Ulex europaeus*, several docking modes were predicted for the orientation of fucose in the primary binding site of the lectin,[157] two of them being in agreement with binding data obtained previously with chemically modified monosaccharides.[158] It was then possible to model the interaction of this lectin with its best known ligand, H-type 2 trisaccharide. The final optimized model of the interaction between H-type 2 trisaccharide and isolectin I from *Ulex europaeus* is shown in Figure 8. The model is in agreement with (i) that H type 1 is a less potent ligand since its docking will result in steric conflict between the *N*-acetyl group of GlcNAc and the protein surface and (ii) that methylation of O-6 of GlcNAc enhances binding[159] since this group interacts with an hydrophobic area of the protein.

B. Lewis x: Conformation and Association

1. NMR and Modeling Study of the Lewis Blood-Group Oligosaccharides

As early as 1980, modeling indicated that Lewis oligosaccharides are basically rigid molecules.[9,144] In each molecule a strong nonbonded interaction occurs between Gal and Fuc residues, both of which are linked to the GlcNAc moiety. The main conclusion of this early investigation was the striking similarities exhibited by Le^a and Le^x trisaccharides on the one hand and by Le^b and Le^y tetrasaccharides on the other hand. These similarities can be correlated with the previously reported similarities of the global minima of the βGal(1−3)βGlcNAc (Le^c) and βGal(1−4)βGlcNAc (LacNAc) disaccharides. The only difference arises from the position of the *N*-acetyl group which is either on one side or the other side of the molecule, depending on whether the neighboring linkage is 1−3 or 1−4. Recently, an exhaustive exploration of the conformational space of the Lewis oligosaccharides was performed using the CICADA approach.[143] It was concluded that the Le^a and Le^b oligosaccharides are slightly less rigid than their Le^x and Le^y counterparts. In theory, the first two could form a second conformational family even if not energetically favored This study also predicted that the main low-energy conformation can display variations of about 20° for each torsion angle. From another modeling study that took solvent effect into account, it was concluded that for the Lewis oligosaccharides, several conformers can occur, the abundance of each conformational family depending strongly on the solvent.[160]

Despite their well-recognized biological role, the first crystal structure of a histo-blood-group carbohydrate-dependent antigen, i.e., Le^x (βGal(1−4)[αFuc(1−3)] βGlcNAc), was reported only in 1996.[95,161] The two crystallographically independent Le^x molecules differ in their overall conformations, and these differences are found principally in the glycosidic torsion angles at the βGal(1−4)GlcNAc linkage for which Φ differs by 10°. Neither of the trisaccharides exhibits an intramolecular hydrogen bond. A strong interaction occurs between the fucose and galactose residue, but only nonpolar van der Waals contacts are involved, each ring presenting its most hydrophobic face to the other. The same conformation is observed in the more bulky Le^y tetrasaccharide when cocrystallized with the isolectin IV of *Griffonia simplicifolia*[117] or with an antibody.[162]

The Le^x molecule, as well as several analogues, has been the subject of many NMR studies, from nOe[163] to NOESY and ROESY experiments[164] and more sophisticated procedures for measurement of relaxation rates[165,166] On the basis of the nOe and dynamics data, it has been shown that no conformational exchange occurs, the molecule existing as a unique conformational family, corresponding to that observed by X-ray crystallography.[95] However, when the ^{13}C NMR parameters are interpreted using the Lipari−Szabo[77] model-free approach, the fucose ring possesses smaller order parameters than the *N*-

acetyllactosamine core, indicating that this ring displays more extensive local motion.[166]

2. Le^x–Le^x Interaction: A Carbohydrate–Carbohydrate-Specific Recognition?

The role of specific carbohydrate−carbohydrate interactions as a biologically significant molecular recognition phenomenon is only just beginning to be recognized. It is now accepted that in some cases (i) the glycosphingolipids in the cell membrane can establish side-by-side contacts that create microdomains and (ii) microdomains from different membranes result in interaction between cells through head-to-head contacts.[167,168] Among the known examples of carbohydrate−carbohydrate interactions, only aggregation of cells through homotypic interaction between Le^x glycosphingolipids have been proven unambiguously.[169−171]

Crystal structures of oligosaccharides yield information related not only to the conformation of the molecule, but also to their packing, i.e., their preferred modes for optimizing intermolecular interactions. In the recently reported crystal structure of Le^x,[95] analysis of the packing indicates a well-defined hierarchy of intermolecular contacts, of which some may be indicative of the manner in which Le^x−Le^x interactions may occur in biological situations, thus providing a molecular basis for a cell−cell recognition event. Along one crystal axis the trisaccharides are aligned in a row with hydrogen bonds between fucose and galactose in neighboring molecules. Indeed, by molecular modeling it was possible to modify this Le^x-row into a glycosphingolipid by addition of the relevant carbohydrate and ceramide moieties. Side-by-side association of such rows leads to the design of a sphingolipid microdomain. This "glycolandscape" model provides a representation of the sphingolipid microdomain in which the accessible surface is totally covered by fucose and galactose.[172]

Head-to-head interactions between Le^x molecules, shown to have a role in cell aggregation,[169−171] have also been confirmed by physical methods. Because of the very weak interaction between two isolated Le^x molecules, it is not possible to observe directly the dimerization by NMR spectroscopy. However, the transferred nOes method, classically used for studying the interaction of a small ligand with a protein receptor, can also detect and quantify the affinity of such ligand for a membrane.[173] When applied to the interaction between two species of Le^x trisaccharide, free and membrane-bound, the affinity constant was found to be 2−3 m^{-1}.[174] In both biochemical and NMR assays, the Le^x−Le^x interaction is Ca^{2+} dependent. From the NMR study, two LacNAc moieties are supposed to dimerize through coordination of the calcium ion. This sugar−calcium interaction is further supported by a NMR study that makes use of the paramagnetic properties of divalent cations to determine their site of interaction with a Le^x pentasaccharide derivative.[175] In the crystal structure of Le^x,[95] head-to-head interactions between facing rows of trisaccharide were also observed. However, these close contacts were directly mediated by hydrogen bonds between opposed fucose and galactose residues and no cations were involved. The interaction in the condensed solid state and in the absence of divalent cations may therefore be different to that postulated between biological membranes.

C. Sialyl Lewis: The Flexible NeuAc-Gal Linkage

Since the discovery of its importance as a recognition ligand in leukocyte adhesion to endothelial cells, the tetrasaccharide sialyl Le^x (αNeuAc(2−3)βGal(1−4)[αFuc(1−3)]βGlcNAc) has been the subject of much interest.

1. Solution Conformation of Sle^x

The sialyl Le^x tetrasaccharide has been the subject of many conformational analysis studies. The difficulty in such theoretical conformational analysis lies in the number of degrees of freedom. Different methods have been used for exploring the multidimensional conformational space of this tetrasaccharide: rigid-body systematic variations around the torsion angles followed by energy minimization,[176,177] molecular dynamics simulation,[178,179] random molecular mechanics (RAMM),[160] and heuristic conformational search (CICADA).[180] The predicted conformations of the tetrasaccharide, as obtained by these different theoretical approaches, are collected in Table 3. All theoretical approaches agree on the relative rigidity of the Le^x backbone compared to the flexibility of the αNeuAc(2−3)βGal linkage.

On the basis of the observation of a nOe between the sialyl H3ax and the galactosyl H3 resonances, the first NMR studies concluded that only conformations conf_B and conf_B′ (see Table 3 and Figure 9) occur in solution.[176,181] When molecular dynamics and NMR data are combined for a quantitative analysis, it becomes clear that the conformation reported here as conf_A is also present in solution.[179,182] The most recent NMR data allowed the collection of more experimental data on SLe^x in solution, either through the investigation of hydroxyl protons in conditions of slow chemical exchange[183] or with the use of homonuclear[184] or heteronuclear[185] 3D NMR. These recent investigations all concluded that several conformations of the sialic acid residue were involved in a dynamic exchange that could not be analyzed from the NMR data in a straightforward manner.

2. Protein-Bound Conformation of SLe^x

Although the crystal surface of the carbohydrate binding domain of human E-selectin has been solved,[186] the structure of the complex with the SLe^x is not known. The conformation of SLe^x tetrasaccharide bound to E-selectin was studied by the use of transferred nOe experiments. The first attempts demonstrated that the transfer nOes observed in the bound state are different from those characteristic of the solution conformations.[187,188] Recent studies have established the conformation of the tetrasaccharide in the bound state. The conformations proposed by the three most recent studies[183−185] all belong to the conf_A family. However, they differ significantly from each other, not only in the orienta-

Table 3. Comparison of Literature Data on SLe^x Conformation[a]

	αNeuNAc(2−3)Gal		αFuc(1−3)GlcNAc		βGal(1−4)GlcNAc	
	Φ	Ψ	Φ	Ψ	Φ	Ψ
			conformation A			
Grid/GESA (GESA-C)[176]	41	127	−72	145	−65	−113
Grid/CHARMm[177]	55	118	−63	142	−65	−116
RAMM/MM2[160]	62	130	−85	141	−72	−60
CICADA/MM3[180]	65	114	−81	150	−76	−104
MD/CHARMM[182]	52	126	−85	140	−60	−113
MD/Amber/NMR[179]	50	125	−70	140	−70	−100
transfer nOe (E-selectin)[184]	44	126	−82	146	−81	−108
transfer nOe (E-selectin) [183]	62	98	−49	134	−96	−86
heteronuclear transfer NOE (E-selectin)[185]	77	108	−91	161	−75	−101
transfer nOe (P-selectin)[183]	35	116	−59	146	−75	−102
crystallography (mutated MBP)[118]	50	94	−68	133	−87	−95
(trimer)	68	82	−70	85	−72	−97
	51	101	−70	140	−75	−109
			conformation B			
Grid/GESA (GESA-B)[176]	−50	116	−72	144	−66	−111
Grid/CHARMM [177]	−44	100	−63	142	−64	−117
RAMM/MM2[160]	−49	105	−86	142	−72	−107
CICADA/MM3[180]	−57	115	−79	149	−77	−103
MD CHARMM[182]	−22	93	−51	143	−57	−120
MD Amber/NMR[179]	−40	100	−70	140	−70	−100
			conformation B′			
Grid/GESA (GESA-A)[176]	−77	63	−72	145	−66	−112
CICADA/MM3[180]	−84	63	−81	150	−73	−105
			conformation C			
Grid/GESA (GESA-D)[176]	−172	100	−72	144	−66	−111
Grid/CHARMM[177]	−178	108	−62	143	−63	−119
CICADA/MM3[180]	−175	108	−80	150	−76	−103

[a]The torsion angles have been all referred to the same convention (Φ = O6−C2−O2−C3′ and Ψ = C2−O2−C3′-C4′) by adding ±120°.

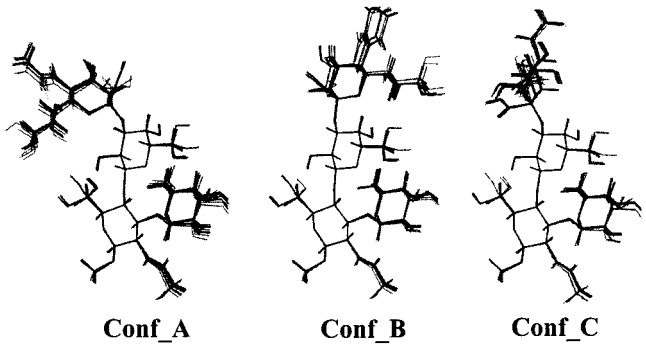

Figure 9. Three lowest energy conformational families of SLe^X as calculated with the CICADA program coupled with the MM3 force field.[180] The lowest energy conformation in each family is represented by black lines, whereas the others, in a energy window of 5 kcal/mol, are represented as gray lines.

tion at the αNeuAc(2−3)βGal linkage but also for the αFuc(1−3)βGlcNAc moiety. At the present time, those discrepancies cannot be explained or solved. Indeed, the protein protons may be sufficiently close to the carbohydrate protons to cause significant spin diffusion. In the absence of a 3D structure or of a reliable model of the complex between SLe^x and the E-selectin surface, the bound conformation and the orientation of the sugar at the protein surface cannot be completely elucidated.

Because it was not possible to determine directly the structures of SLe^x bound to the selectins, an indirect approach has been adopted. The rat serum mannose-binding protein (MBP-A), which is homologous to selectins, has been mutated in order to mimic the essential aspects of carbohydrate recognition by E-selectin.[189] The crystal structure of the K3 mutant of MBP-A has been solved as a complex with several carbohydrate ligands.[118] In this complex, the SLe^x adopts a conformation close to the one predicted by the several nOe studies. The hydroxyl groups of the fucose residue interact directly with the calcium in the binding site. In the complex, there is no direct contact between the sialic acid moiety and the protein surface.

V. Fragments of Glycosaminoglycans

The glycosaminoglycans comprise a family of complex anionic polysaccharides including glucosaminoglycans (heparin, heparan sulfate), galactosylaminoglycans (chondroitin sulfate and dermatan sulfate), hyaluronic acid, and keratan sulfate. In addition to their participation in the physicochemical properties of the extracellular matrix, glycosaminoglycan fragments are specifically recognized by protein receptors and play a role in the regulation of many processes, such as hemostasis, growth factor control, anticoagulation, and cell adhesion.[190] Given the importance of protein−glycosaminoglycan interactions, oligosaccharide fragments are important targets for drug design. Considerable effort has been invested in establishing the structure of glycosaminoglycan fragments with biological activity, either in solution or in their bioactive conformation, i.e., in interaction with a protein receptor.

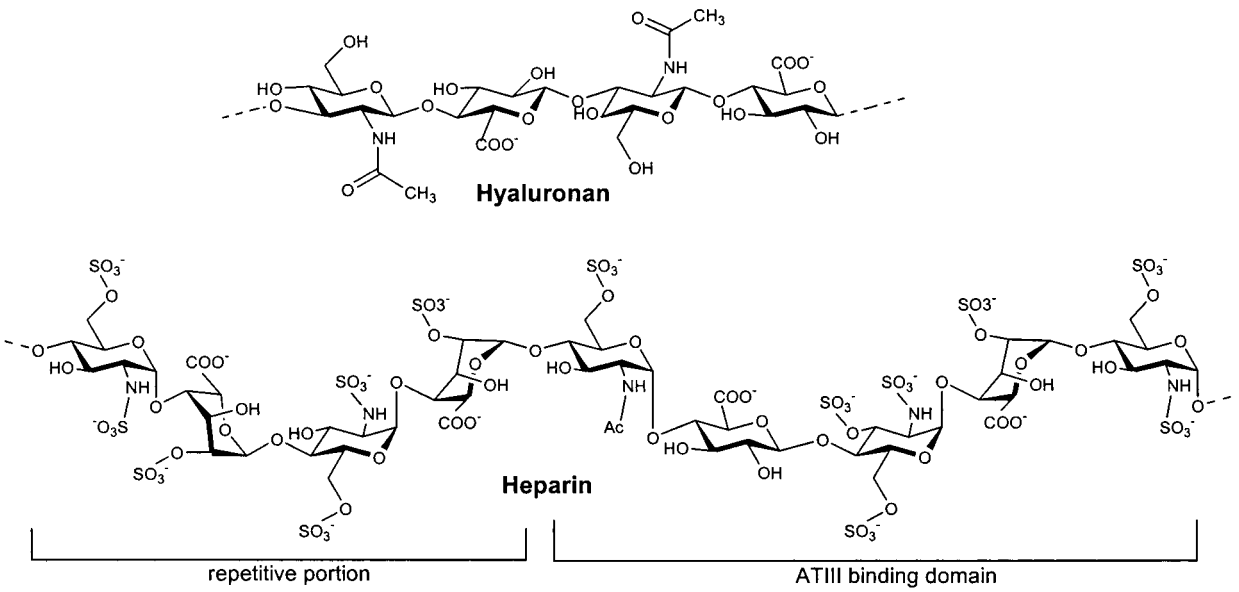

Figure 10. Schematic representation of two glycosaminoglycans.

A. Heparin Fragments: Highly Sulfated Biologically Active Oligosaccharides

The heparin polymer[191] is made up of disaccharide repeat units consisting of α1−4 linked D-glucosamine and L-iduronic acid carrying three sulfate groups. Heparin is recognized by a variety of protein receptors,[192] such as growth factors, chemokines, and protease inhibitors of the blood coagulation cascade, etc. Some proteins recognize the most regular regions of heparin, while others have affinity only for unique irregularities in this structure. For example, the anticoagulant activity of heparin is due to a highly specific interaction between the plasma protein antithrombin III (AT-III) and a unique pentasaccharide sequence present in some of the molecules (Figure 10).[193,194]

1. Conformational Studies of the Anticoagulant Heparin Pentasaccharide

Following the chemical synthesis of the specific pentasaccharide,[193,194] systematic chemical modifications were made to the functional groups of the individual constituent monosaccharides to assess the effect of these on biological activity.[195] The first conformational studies[196] highlighted several difficulties that are specific to heparin fragments: (a) compared to other pyranose sugars the iduronate ring displays great flexibility and can adopt a variety of conformations viz., $^4C_1$, $^1C_4$, and ^2S_O, (b) parametrization for the *O*- and *N*-sulfate groups is not readily available in most force fields, and (c) the polyanionic character of the pentasaccharide requires special consideration for solvent and counterion effects. Calculations of potential-energy surfaces of each constituent disaccharide, with the use of the MM2 force field, demonstrated that each glycosidic linkage is potentially flexible and can adopt an average of two conformations.[196] The authors proposed a simplified model, comprising two elongated conformations corresponding to the $^1C_4$ and ^2S_O forms of the iduronate residue, which reproduced most of the observed coupling constants and nOe data. A conformational study was conducted on a related trisaccharide using the MM3 program and nOe measurements.[197] In agreement with binding data on antithrombin III, it was demonstrated that 2-*O*-sulfated substitutions do not affect the conformational equilibrium observed in the case of 2-*N*-sulfated. Longitudinal and transverse relaxation times, measured at different magnetic fields, indicated that the synthetic pentasaccharide displays relatively complex motion in solution: [79] its overall molecular tumbling is anisotropic due to its elongated shape, and in addition interpretation of the experimental data requires the inclusion of internal motion. A more recent study on a synthetic tetrasaccharide confirmed the extended shape of this sulfated oligosaccharide.[198] In this latter molecule, the iduronate residue is at the reducing position and, from the analysis of the coupling constants, the estimated ratio of conformers $^1C_4$ to ^2S_O was found to be approximately 75/25.

The study of the interaction between the heparin pentasaccharide and its receptor antithrombin III is not a simple task, due to the fact that a conformational change occurs in the protein upon binding.[199] The first model,[200] obtained using homology modeling for the protein and hand docking of the pentasaccharide, allowed the determination of the basic amino acids involved in the recognition of the sulfate and carboxylate groups. A more recent study,[201] making use of several recently developed docking programs, arrived at the same prediction for the binding site. Indeed, in the 2.9 Å crystal structure of the complex between antithrombin III and synthetic pentasaccharide,[138] the sulfates and carboxylates interact with a cluster of basic amino acids. In the complex, the pentasaccharide is not as fully extended as observed in solution (Figure 11). All of the linkages are in their lowest energy conformation except one: the Φ angle of the glucosamine−iduronate linkage which deviates by about 30°. This deviation, together with the ^2S_O conformation of the iduronate ring, creates a kink in the overall shape of the pentasaccharide. Interest-

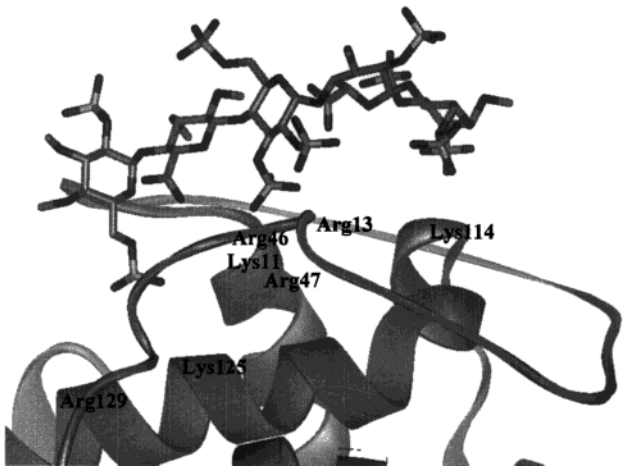

Figure 11. Graphical representation of the heparin pentasaccharide with anticoagulant properties in the binding site of antithrombin III as observed in the crystal structure of the complex (PDB code 1AZX).[138] Protein and oligosaccharide are displayed as ribbon and stick representations, respectively, using the WebLab ViewerLite program.[248] Amino acids directly involved in salt bridges with the pentasaccharide sulfate or carboxyl groups are labeled.

ingly, a recent NMR study[198] of the interaction between a related tetrasaccharide and the protein did not lead to the same conclusion. On the basis of coupling constants and measurements of transferred nOes, it was inferred that the iduronate ring adopts a $^1C_4$ conformation and that the linkage between the glucosamine and the glucuronate residues is the one distorted upon binding. It is difficult to determine whether the discrepancies between the solution and the crystal studies arose from the difference in the ligand or from the limitations in the accuracy of both methods.

2. Conformational Studies of the Repeating Region of Heparin

NMR study of the heparin polysaccharide has demonstrated that its solution conformation is ribbonlike, with a cluster of sulfate groups on each side of the ribbon.[202] Shorter oligosaccharides have also been studied by means of NMR and molecular dynamics.[203,204] It is of interest to note that the internal iduronate rings adopt a $^2S_0$ conformation in the shorter fragment,[203] a tetrasaccharide, but undergoes a change in the hexasaccharidic fragment[204] or in the polysaccharide[202] with almost equal proportions of $^2S_0$ and $^1C_4$ conformations being found. Although there are many heparin receptors with important biological roles, very few have been cocrystallized. Only the basic and acidic fibroblast growth factors (FGF-1 and FGF-2) have been solved as complexes with fragments of heparin.[136,137] The specific interaction between FGF-2 and tetra- and hexasaccharides utilize the negative groups of heparin and, in particular, the *N*- and *O*-sulfates.[136] The oligosaccharide is not buried in a binding site but is wrapped around the protein surface. In its crystalline form, FGF-1 exists as a dimer exclusively bridged by a decasaccharide of heparin with no protein−protein contact (Figure 12). These complexes served as a template, on the one hand, for understanding the role

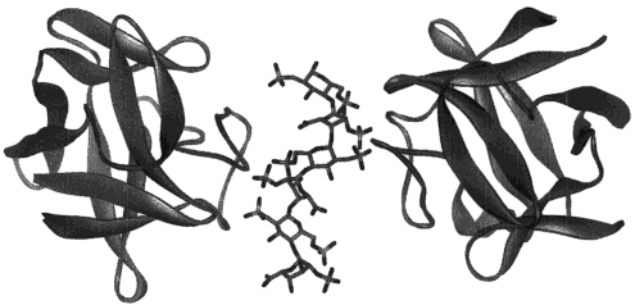

Figure 12. Graphical representation of the heparin hexasaccharide sandwiched between two monomers of acidic fibroblast growth factor as observed in the crystal structure of the complex (PDB code 2AXM).[137] Protein and oligosaccharide are displayed using ribbon and stick representation, respectively, using the WebLab ViewerLite program.[248]

of sulfated glycosaminoglycans in FGF function[205] and, on the other, for modeling ternary interactions involving FGF receptors. However, it should not be forgotten that modeling the interaction of GAGs with proteins is a challenging task because of the weak surface complementarity, the high charge density of the binding areas, and the highly flexible nature of the polysaccharide.[201] Nevertheless, a ternary model has been proposed for the interaction between FGF-2, heparin, and FGF receptor.[206] On the basis of the crystal structure of the complex between interleukin-1 and its receptor, FGF-1 and its receptor were predicted to form an "electric sandwich" in which the heparin adopts an S shape.[207] Finally, modeling efforts have also been made to understand the interaction of heparin with platelet factor 4 and with interleukin-8.[201,208]

B. Fragments of Hyaluronan: A Non-sulfated Glycosaminoglycan

Hyaluronan is a negatively charged nonsulfated glycosaminoglycan that exhibits a wide variety of biological effects mediated by binding to cell-surfaces and therefore involved in cell adhesion. Its viscoelastic properties make it an important component of synovial fluid, cartilage, vitreous humor, and other tissues. The polymer is built up from repeating −βGlcNAc(1−4)βGlcA(1−3)− disaccharide units (Figure 10). A combined NMR and molecular dynamics study of octasaccharide fragments demonstrated that the β1−3 linkage is less flexible than the β1−4 and can adopt two main conformers.[209] The β1−3 linkage was further characterized by a molecular modeling study in the absence and presence of counterions.[210] It was demonstrated that the presence of Na^+ influences the conformational preference of this linkage. This was later confirmed by ^1H and ^{13}C NMR studies in different cationic environments[211,212] in which it appeared that each of the different cations influences the linkage mobility in a different manner. Such specific interactions will complicate the modeling studies.

Specific interactions also occur between the carbohydrate and the water molecules. Molecular dynamics simulations of the two disaccharide repeats of hyaluronan, performed in explicit water with the use

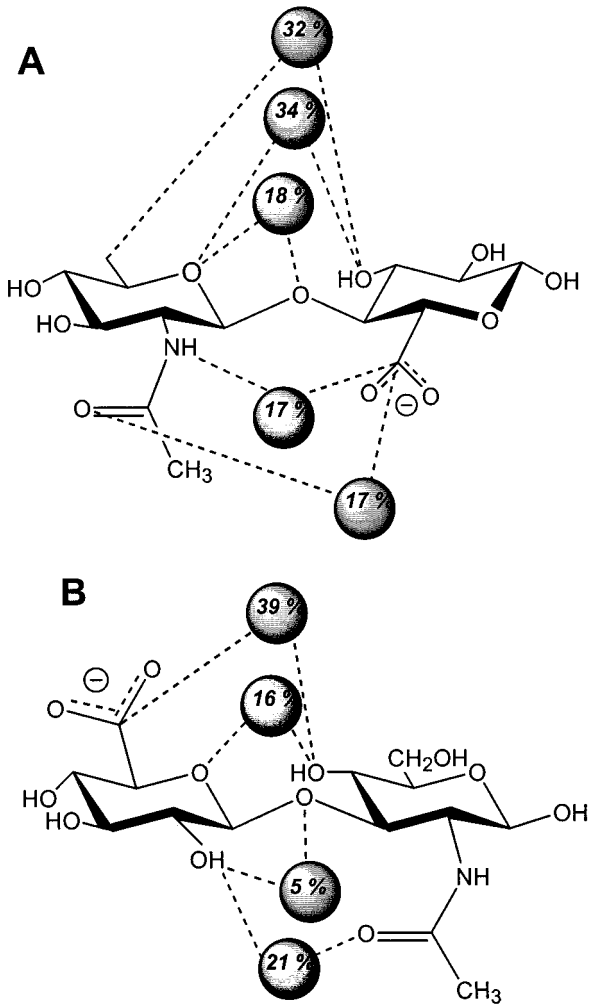

Figure 13. Schematic representation of the two disaccharides constituting hyaluronan along with the probability of the various water-bridged inter-ring hydrogen bonds formed during a 500 ps molecular dynamics simulation in water. (Adapted with permission from ref 69. Copyright 1997 Oxford University Press.)

of the CHARMm force field, demonstrated the role of "water bridges" at both linkages.[69] A combination of intramolecular hydrogen bonds and "water bridged" hydrogen bonds were suggested to play an important role in the conformational behavior of the oligosaccharides. In particular, the water molecules, which are frequent residents around the glycosidic linkages, may screen the electron interactions that are responsible for the *exo*-anomeric effect and decrease its influence. Indeed, another molecular dynamics study, performed with the GROMOS force field, confirmed the unusual high density of water around hyaluronan fragments.[70] The reorientation of water molecules is a highly dynamic process, but on average, there are between 10 and 15 hydrogen bonds between a disaccharide and the surrounding water (Figure 13). This specific behavior may explain the viscoelastic properties of the proteoglycan aggrecan that make up cartilage.

Discussion of composite proteoglycans such as aggrecan leads to a brief consideration of the conformations of the constituent polysaccharide chains. From X-ray fiber diffraction studies it has been known for many years that these chains show conformational versatility and can exist in a variety of polymorphic forms in the solid state, depending on factors such as hydration, temperature, and counterion. This versatility has been verified in modeling studies[210] which have shown that these chains may indeed adopt a large number of conformations having similar energies. Furthermore, the capacity for the formation of chain−chain interactions and multiple helices has also been confirmed, and this is yet another factor that needs to be attended to when dealing with molecules at the polysaccharide level.

VI. Can We Use Conformational Knowledge for the Design of Oligosaccharidic Ligands?

A. Structural Mimetics

Rational drug design makes use of available information about a particular naturally active compound and its receptor. Accordingly, knowledge of three-dimensional structure and conformational behavior of the ligand may help in the design of new active compounds. For example, the crystal structure of the complex between the Shiga-like toxin and its GB3 glycolipid acceptor[122] has been used to design a multivalent carbohydrate ligand with subnanomolar inhibitory activity.[213]

The search for analogues of SLe^x has been an active area of research and provided several examples where conformational studies have assisted in the provision of lead compounds. This field has been reviewed recently,[214] and what follows is a summary of the strategies that have been used. For example, the inherent pseudo-C_2 symmetry of the Lewis antigens described previously has been used as a template to design tetrasaccharides having full C_2 symmetry.[215] The resulting product, bis-α-fucosylated Galβ,β-trehalose, has a similar affinity for E-selectin as the SLe^x tetrasaccharide. Thus, targeting a compound having the same overall shape and conformational behavior as Le^b and Le^y tetrasaccharide appears to be a promising strategy for the design of new class of Lewis mimetics. Some analogues that mimic the major conformers of SLe^x, as determined by NMR spectroscopy, have also been designed. However, this approach has been less successful[216] since, as was confirmed later, the selectin-bound conformation is noticeably different from the most populated conformer in solution. Conformations of SLe^x bound to the selectins have been determined by using transfer nOe techniques,[183−185] and these have assisted in the design of many small analogues. Comparison between the shape of SLe^x and its analogues was achieved using the "internal coordinate" procedure and led to the chemical synthesis of a potent SLe^x mimic.[217,218]

B. *S*- and *C*-Glycoside Analogs

The search for ligand analogues that are not susceptible to hydrolytic attack led to the synthesis of heteroanalogues in which the glycosidic oxygen of several disaccharides was replaced by a S or a CH_2 group. The goal of these studies was to obtain compounds displaying the conformational attributes

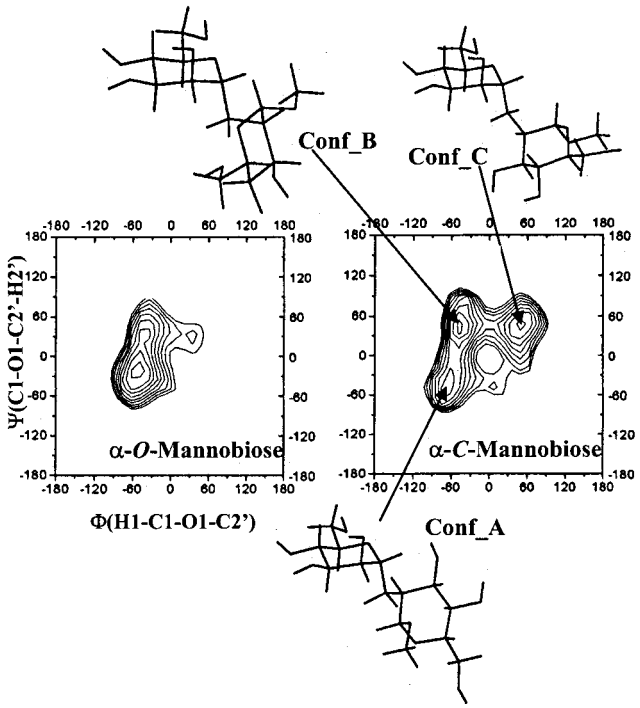

Figure 14. Comparison of the (Φ,Ψ) energy maps of αMan(1−2)αMan-*O*-Me and its *C*-analogue as calculated using the MM3* force field, together with the low-energy conformations of the *C*-disaccharide. (Adapted with permission from ref 223. Copyright 1999 Wiley-VCH.)

of the *O*-linked parent compounds but with improved chemical and biochemical stability. Initially it was thought that *C*-disaccharides were true structural mimetics of the natural compounds,[219] but subsequently, contradictory results were obtained. First it was demonstrated that *S*-disaccharides were more flexible than the natural compounds, so that although the thioglycoside analogues of galabiose,[220] αFuc(1−3)GlcNAc,[221] and maltoside[222] displayed the same global energy minima as the native disaccharides, they also exhibited secondary energy minima that differed both in Φ and Ψ torsion angles and that corresponded to significant populations in solution. The *C*-analogues were found to be even more flexible, and for both an α-linkage (αMan(1−2)Man)[223] (Figure 14) and a *β*-linkage (lactose),[224,225] it was demonstrated that the *C*-analogue can adopt additional conformations. The global energy minimum of *C*-lactose corresponds to an anti conformation which is a minor conformational family for *O*-lactose.[224] It is now accepted that protein receptors can select conformations other than that of lowest energy. Thus, *C*-lactose, ricin B,[226] and bovine heart galectin-1[227] select the *syn*-conformation, i.e., the lowest energy conformation of natural lactose, whereas *E. coli* *β*-galactosidase[228] selects the *gauche*−*gauche* (or anti-Φ) conformation corresponding to a local minimum.

The flexibility of *S*- and *C*-analogues may result in a high entropy cost for protein binding so that the affinity for the receptors is decreased. In any case, even if it turns out that these compounds have limitations as therapeutic agents, they have been shown to be excellent compounds for studying mechanisms of glycohydrolase action in structural biology.[229]

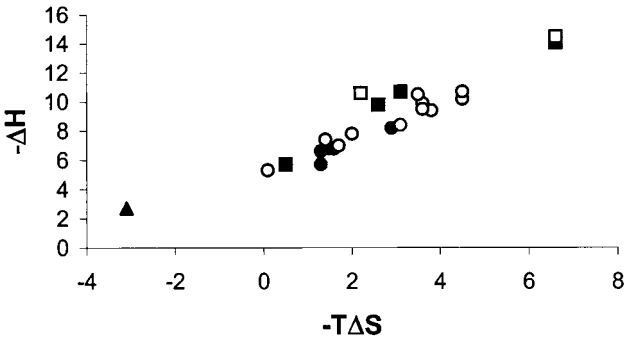

Figure 15. Enthalpy−entropy compensation plot for the interaction between concanavalin A and mannoside and mannose-containing oligosaccharides:[241,249−252] (●) monosaccharides, (○) disaccharides, (■) trisaccharide; (□) biantennary pentasaccharides, and (▲) cyclic trisaccharide (**3**).

C. The Search for Higher Affinity Ligands: Conformationally Constrained Molecules

Carbohydrate−protein interactions are characterized by the weak affinities involved in their association. It is thought that higher affinities may be obtained by factors related to the presentation of the protein or by multivalent associations.

1. Thermodynamics of Protein/Carbohydrate Interactions

Calorimetric studies have provided information on the thermodynamics of complex formation in many protein−carbohydrate associations.[230] Generally, although there are counter examples, protein−carbohydrate associations are typified by favorable enthalpy terms which are offset by unfavorable entropy contributions. Thus, larger ligands usually have correspondingly more negative values of ΔH and $T\Delta S$ leading to the phenomenon of "entropy/enthalpy compensation". This is illustrated in Figure 15, which summarizes the thermodynamic parameters measured for the interaction of concanavalin A with mannose and mannose-containing oligosaccharides of different size. It is widely accepted that the enthalpy term arises from both numerous hydrogen bonds and extensive van der Waals interactions. The origin of the entropy barrier is more controversial and has been considered to arise either from solvation effects[231] or from the loss of conformational flexibility of the carbohydrate ligand.[232]

In the first hypothesis it is considered that protein−carbohydrate binding brings together 'polyamphiphilic' surfaces, i.e., surfaces containing both polar and nonpolar regions. The studies of Lemieux and coworkers show that at such surfaces the water structure is perturbed. Complex formation triggers desolvation and the return of these perturbed water molecules to the bulk,; a process which may be entropically favorable. Support for this interpretation has come from thermodynamics, in the form of solvent isotope effects determined by titration microcalorimetry.[233] Controversially, microcalorimetry experiments, coupled with osmotic stress strategy, led to the conclusion that as many as five water molecules were taken from the bulk solvent upon binding of mannose by concanavalin A.[234] However, the 'polyamphiphilic burial' model has been attacked and

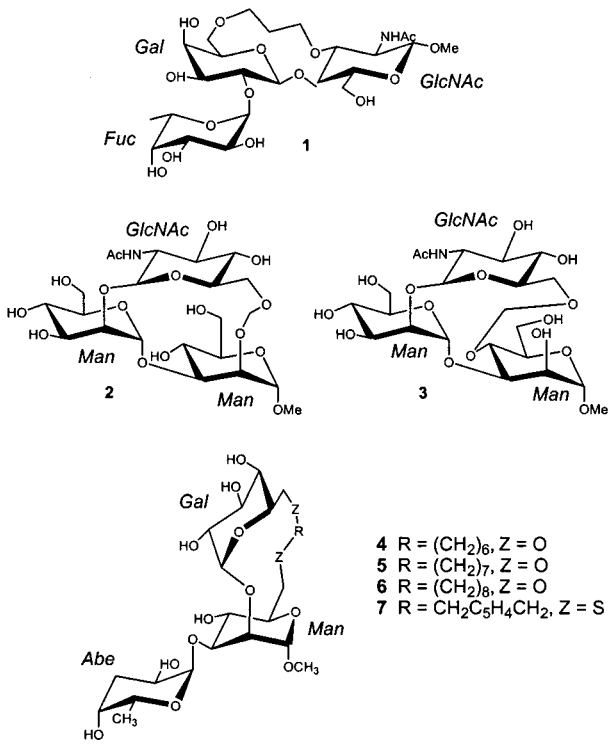

Figure 16. Conformationally constrained analogues of H type 2 blood-group trisaccharide (**1**),[237] *N*-linked glycan trisaccharide (**2** and **3**),[240,241] and trisaccharide epitope of *Salmonella* cell-surface carbohydrate (**4**−**7**).[238,239]

it has been proposed that highly favorable interactions between the lectin and the carbohydrate drive the enthalpy of binding.[235]

In the second hypothesis, the entropy barrier is due to the freezing of flexible oligosaccharides or carbohydrate side chains upon binding by the protein. The estimates for freezing single-bond rotamers vary from 0.6[236] to 2 kcal/mol.[232]

2. Rational Design of Conformationally Constrained Oligosaccharides.

The low affinities that characterize protein/oligosaccharide interactions is a major barrier to the development of carbohydrate-derived drugs. To get higher-affinity ligands, some attempts have been made to lower the conformational entropy by locking the oligosaccharide in its bioactive conformation (Figure 16). According to Bundle,[237] the following criteria should be fulfilled in the design of tethered oligosaccharides: (i) the tether should not be attached to key polar groups that are involved in the interaction with the host, (ii) the bioactive conformation should be targeted, and (iii) the tether should not create steric hindrance to the binding.

A constrained H-type 2 blood-group trisaccharide that displays slightly lower affinity for *Ulex europaeus* lectin I than the native trisaccharide has been synthesized,[237] but no thermodynamic data are yet available for this binding. The same group has designed tethered analogues of the trisaccharide epitope αGal(1−2)[αAbe(1−3)]αMan to study its interaction with antibody fragments.[238] Several molecules, carrying variations in the nature of the tether, were synthesized, but none exhibits free-energy

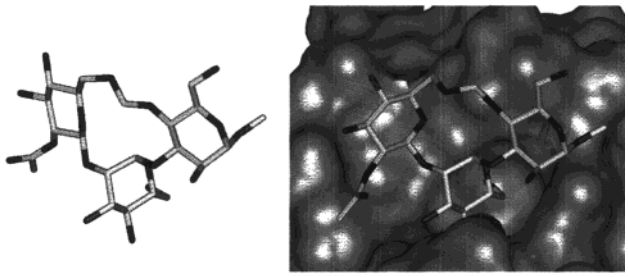

Figure 17. Calculated lowest energy conformation for cyclic trisaccharide **3** alone and in the binding site of concanavalin A.[241]

changes larger than 0.5 kcal/mol when compared to the native trisaccharide.[239] The authors concluded that interresidue flexibility is not a major contributor to the low affinity observed in protein/carbohydrate interaction.

However, a recent study, conducted on an oligosaccharide with a more severe flexibility restraint, arrived at different conclusion. Two cyclic analogues of βGlcNAc(1−2)αMan(1−3)αMan were designed and synthesized,[240,241] with the aim of mimicking the two different conformations observed in the binding site of two crystal structures of *Lathyrus ochrus* complexed with oligosaccharides.[130,132] Of the two cyclic trisaccharides, one was shown to have markedly reduced flexibility as demonstrated by a full characterization by NMR and molecular modeling studies.[241] Indeed, the expected gain in entropy was confirmed in microcalorimetric titration experiments since there is a 1.6 kcal/mol ΔH value difference in binding between the linear and cyclic trisaccharide. Unfortunately the gain in the entropic barrier is lost in the enthalpic term and the free energy of binding of the tethered molecule is less favorable than for the linear one. It could be hypothesized that the lowest energy conformation of the cyclic trisaccharide in solution is still too different from that predicted to produce the lowest energy of interaction with the protein (Figure 17).

Thus, it seems that the target of a perfect analogue, i.e., one that mimics perfectly the bioactive conformation but results in no conformational entropy, is still a challenge. Recently Schmidt demonstrated that it is possible to constrain a glycosidic linkage in its anti conformation,[242] suggesting that the approach described here remains one of the most promising for the production of high-affinity carbohydrate-derived drugs

VII. Conclusions

In the past few years we have seen that theoretical and experimental studies are becoming mutually reliant for the elucidation of structural and dynamical data in the oligosaccharide field. The next frontier will most likely focus on the problems associated with hydration. Molecular dynamics simulations of carbohydrates with the inclusion of explicit water molecules has proven to be a powerful tool for reconciling theoretical and experimental conformational data.[64] Furthermore, calculation of free-energy perturbations is a promising approach for the prediction of oligosaccharide-receptor binding affinities.[243−244] Detailed

analyses of hydration patterns for carbohydrates in water[245] or in aqueous solvent mixtures[246] can now be undertaken. Understanding the role of water as a structural but dynamic element of the solvated oligosaccharide will allow modeling of the interaction between charged carbohydrates and counterions and for this new and improved experimental and theoretical procedures will be needed. It is only by the merging of theory and experiment that understanding of the role of water in the conformation and dynamics and ultimately the biology of many oligosaccharides will be achieved.

VIII. Acknowledgments

The authors are indebted to Professor W. Mackie (University of Leeds) for fruitful scientific discussions and corrections of the manuscript.

IX. References

(1) Kobata, A. *Eur. J. Biochem*. **1992**, *209*, 483.
(2) Varki, A. *Glycobiology* **1993**, *3*, 97.
(3) Dwek, R. A. *Chem. Rev.* **1996**, *96*, 683.
(4) Gagneux, P.; Varki, A. *Glycobiology* **1999**, *9*, 747.
(5) Lis, H.; Sharon, N. *Chem. Rev.* **1998**, *98*, 637.
(6) Pérez, S.; Gautier, C.; Imberty, A. In *Oligosaccharides in Chemistry and Biology: A Comprehensive Handbook*; Ernst, B., Hart, G., Sinay, P., Eds.; Wiley/VCH: Weinheim, 2000; pp 969−1001.
(7) Peters, T.; Pinto, B. M. *Curr. Opin. Struct. Biol.* **1996**, *6*, 710.
(8) Rice, K. G.; Wu, P.; Brand, L.; Lee, Y. C. *Curr. Opin. Struct. Biol.* **1993**, *3*, 669.
(9) Lemieux, R. U.; Bock, K.; Delbaere, L. T. J.; Koto, S.; Rao, V. S. R. *Can. J. Chem*. **1980**, *58*, 631.
(10) Woods, R. J. In *Reviews of Computational Chemistry*; Lipkowitz, K., Boyd, D. B., Eds.; VCH Publishers: New York, 1996; Vol. 9, pp 129−165.
(11) Laine, R. A. *Glycobiology* **1994**, *4*, 759.
(12) Engelsen, S. B.; Cros, S.; Mackie, W.; Pérez, S. *Biopolymers* **1996**, *39*, 417.
(13) Bohne, A.; Lang, E.; von der Lieth, C. W. *J. Mol. Model.* **1998**, *4*, 33.
(14) Imberty, A.; Bettler, E.; Karababa, M.; Mazeau, K.; Petrova, P.; Pérez, S. In *Perspectives in Structural Biology*; Vijayan, M., Yathindra, N., Kolaskar, A. S., Eds.; Indian Academy of Sciences and Universities Press: Hyderabad, 1999; pp 392−409.
(15) Lemieux, R. U.; Koto, S.; Voisin, D. In *Anomeric effect: origin and consequence*; Szarek, W. A., Horton, D., Eds.; ACS Symposium Series 87; American Chemical Society: Washington, DC, 1979; pp 17−29.
(16) Tvaroska, I.; Bleha, T. *Adv. Carbohydr. Chem. Biochem*. **1989**, *47*, 45.
(17) Allinger, N. L. *J. Am. Chem. Soc.* **1977**, *99*, 8127.
(18) Allinger, N. L.; Yuh, Y. H.; Lii, J.-H. *J. Am. Chem. Soc.* **1989**, *111*, 8551.
(19) Allinger, N. L.; Rahman, M.; Lii, J.-H. *J. Am. Chem. Soc.* **1990**, *112*, 8293.
(20) Allinger, N. L.; Zhu, Z.-Q.; Chen, K. *J. Am. Chem. Soc.* **1990**, *112*, 6120.
(21) Tvaroska, I.; Pérez, S. *Carbohydr. Res.* **1986**, *149*, 389.
(22) van Gunsteren, W. F. *GROMOS: Gröningen Molecular Simulation Program Package*; University of Groningen, 1987.
(23) Ott, K.-H.; Meyer, B. *J. Comput. Chem*. **1996**, *17*, 1068.
(24) Brooks, B.; Bruccoleri, R.; Olafson, B.; States, D.; Swaminathan, S.; Karplus, M. *J. Comput. Chem*. **1983**, *4*, 187.
(25) Ha, S. N.; Giammona, A.; Field, M.; Brady, J. W. *Carbohydr. Res.* **1988**, *180*, 207.
(26) Reiling, S.; Schlenkrich, M.; Brickmann, J. *J. Comput. Chem*. **1996**, *17*, 450.
(27) Grootenhuis, P. D. J.; Haasnoot, C. A. G. *Mol. Simul.* **1993**, *10*, 75.
(28) Kouwijzer, M. L. C. E.; Grootenhuis, P. D. J. *J. Phys. Chem*. **1995**, *99*, 13426.
(29) Weiner, S. J.; Kollman, P. A.; Case, D. A.; Singh, U. C.; Ghio, C.; Alagona, G.; Profeta Jr, S.; Weiner, P. *J. Am. Chem. Soc.* **1984**, *106*, 765.
(30) Homans, S. W. *Biochemistry* **1990**, *29*, 9110.
(31) Glennon, T. M.; Zheng, Y.-J.; Le Grand, S. M.; Schutzberg, B. A.; Merz Jr., K. M. *J. Comput. Chem*. **1994**, *15*, 1019.
(32) Momany, F. A.; Willet, J. L. *Carbohydr. Res.* **2000**, *326*, 194.
(33) Mohamadi, F.; Richards, N. G. H.; Guida, W. C.; Liskamp, R.; Lipton, M.; Caufield, C.; Chang, G.; Hendrickson, T.; Still, W. C. *J. Comput. Chem*. **1990**, *11*, 440.
(34) Senderowitz, H.; Still, W. C. *J. Org. Chem*. **1997**, *62*, 1427.
(35) Lifson, S.; Warshel, A. *J. Chem. Phys.* **1968**, *49*, 5116.
(36) Kildeby, K.; Melberg, S.; Rasmussen, K. *Acta Chem. Scand.* **1977**, *A31*, 1.
(37) Fabricius, J.; Engelsen, S. B.; Rasmussen, K. *J. Carbohydr. Chem*. **1997**, *16*, 751.
(38) Dauchez, M.; Derreumaux, P.; Lagant, P.; Vergoten, G. *J. Comput. Chem*. **1995**, *16*, 188.
(39) Clark, M.; Cramer, R. D. I.; van den Opdenbosch, N. *J. Comput. Chem*. **1989**, *8*, 982.
(40) Imberty, A.; Hardman, K. D.; Carver, J. P.; Pérez, S. *Glycobiology* **1991**, *1*, 456.
(41) Pérez, S.; Meyer, C.; Imberty, A. In *Modelling of Biomolecular Structures and Mechanisms*; Pullman, A., Jortner, J., Pullman, B., Eds.; Kluwer Academic Publishers: Dordrecht, 1995; pp 425−454.
(42) Martin-Pastor, M.; Espinosa, J. F.; Asensio, J. L.; Jiménez-Barbero, J. *Carbohyd. Res.* **1997**, *298*, 15.
(43) Hwang, M.-J.; Ni, X.; Waldman, M.; Ewig, C. S.; Hagler, A. T. *Biopolymers* **1998**, *45*, 435.
(44) Mayo, S. L.; Olafson, B. D.; Goddard, W. A. I. *J. Phys. Chem*. **1990**, *94*, 8897.
(45) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1995**, *118*, 11225.
(46) Damm, W.; Frontera, A.; Tirado-Rives, J.; Jorgensen, W. L. *J. Comput. Chem*. **1997**, *18*, 1955.
(47) Halgren, T. A. *J. Comput. Chem*. **1996**, 490.
(48) Pérez, S.; Imberty, A.; Engelsen, S. B.; Gruza, J.; Mazeau, K.; Jiménez-Barbero, J.; Poveda, A.; Espinosa, J. F.; van Eyck, B. P.; Jonhson, G.; French, A. D.; Kouwijzer, M. L. C. E.; Grootenhuis, D. J.; Bernardi, A.; Raimondi, L.; Senderowitz, H.; Durier, V.; Vergoten, G.; K., R. *Carbohydr. Res.* **1998**, *314*, 141.
(49) Von der Lieth, C.-W.; Kozar, T.; Hull, W. E. *J. Mol. Struct. (THEOCHEM)* **1997**, *395*−*396*, 225.
(50) French, A. D. *Biopolymers* **1988**, *27*, 1519.
(51) Kozár, T.; Petrák, F.; Gálová, Z.; Tvaroska, I. *Carbohydr. Res.* **1990**, *204*, 27.
(52) Hooft, R. W. W.; Kanters, J. A.; Kroon, J. *J. Comput. Chem*. **1991**, *12*, 943.
(53) Koca, J. *J. Mol. Struct. (THEOCHEM)* **1994**, *308*, 13.
(54) Koca, J. *Prog. Biophys. Mol. Biol.* **1998**, *70*, 137.
(55) Engelsen, S. B.; Koca, J.; Braccini, I.; Hervé du Penhoat, C.; Pérez, S. *Carbohydr. Res.* **1995**, *276*, 1.
(56) Peters, T.; Meyer, B.; Stuike-Prill, R.; Somorjai, R.; Brisson, J. R. *Carbohydr Res* **1993**, *238*, 49.
(57) Orozco, M.; Alhambra, C.; Barril, X.; Lopez, J. M.; Busquets, M. A.; Luque, F. J. *J. Mol. Model.* **1996**, *1*, 1.
(58) Tvaroska, I.; Kozar, T. *J. Am. Chem. Soc.* **1980**, *102*, 6929.
(59) Still, W. C.; Tempczyk, A.; Hawley, R. C.; Hendrickson, T. *J. Am. Chem. Soc.* **1990**, *112*, 6127.
(60) Höög, C.; Widmalm, G. *Glycoconjugate J.* **1998**, *15*, 183.
(61) Stenutz, R.; Widmalm, G. *Glycoconjugate J.* **1998**, *15*, 415.
(62) Brady, J. W.; Schmidt, R. K. *J. Phys. Chem*. **1993**, *97*, 958.
(63) Ott, K. H.; Meyer, B. *Carbohydr. Res.* **1996**, *281*, 11.
(64) Engelsen, S. B.; Hervé du Penhoat, C.; Pérez, S. *J. Phys. Chem*. **1995**, *99*, 13334.
(65) Liu, Q.; Schmidt, R. K.; Teo, B.; Karplus, P. A.; Brady, J. W. *J. Am. Chem. Soc.* **1997**, *119*, 7851.
(66) Naidoo, K. J.; Brady, J. W. *J. Am. Chem. Soc.* **1999**, *121*, 2244.
(67) Ueda, K.; Brady, J. W. *Biopolymers* **1996**, *38*, 461.
(68) Vishnyakov, A.; Widmalm, G.; Kowalewski, J.; Laaksonen, A. *J. Am. Chem. Soc.* **1999**, *121*, 5403.
(69) Almond, A.; Sheehan, J. K.; Brass, A. *Glycobiology* **1997**, *7*, 597.
(70) Kaufmann, J.; Möhle, K.; Hofmann, H.-J.; Arnold, K. *J. Mol. Struct. (THEOCHEM)* **1998**, *422*, 109.
(71) Kouwijzer, M. L. C. E.; van Eijck, B. P.; Kroes, S. J.; Kroon, J. *J. Comput. Chem*. **1993**, *14*, 1281.
(72) Bock, K.; Brignole, A.; Sigurskjold, B. W. *J. Chem. Soc., Perkin Trans. 2* **1986**, 1711.
(73) De Bruyn, A. *J. Carbohydr. Chem*. **1991**, *10*, 159.
(74) Sandström, C.; Baumann, H.; Kenne, L. *J. Chem. Soc., Perkin Trans. 2* **1998**, 2385.
(75) Tvaroska, I.; Hricovini, M.; Petrakova, E. *Carbohydr. Res.* **1989**, *189*, 359.
(76) Bose, B.; Zhao, S.; Stenutz, R.; Clorean, F.; Bondo, P. B.; Hertz, B.; Carmichael, I.; Serianni, A. S. *J. Am. Chem. Soc.* **1998**, *120*, 11158.
(77) Lipari, G.; Szabo, A. *J. Am. Chem. Soc.* **1982**, *104*, 4546.
(78) Catoire, L.; Braccini, I.; Bouchemal-Chibani, N.; Jullien, L.; Hervé du Penhoat, C.; Pérez, S. *Glycoconjugate J.* **1997**, *14*, 935.
(79) Hricovini, M.; Torri, G. *Carbohydr. Res.* **1995**, *268*, 159.
(80) Tran, V.; Brady, J. W. *Biopolymers* **1990**, 961.
(81) French, A. D.; Dowd, M. K. *J. Mol. Struct. (THEOCHEM)* **1993**, *1993*, 183.
(82) Sheng, S.; van Halbeek, H. *Biochem. Biophys. Res. Commun.* **1995**, *215*, 504.

(83) Homans, S. W. *Biochem. Soc. Trans.* **1998**, *26*, 551.
(84) Jiménez-Barbero, J.; Asensio, J. L.; Cañada, F. X.; Poveda, A. *Curr. Opin. Struct. Biol.* **1999**, *9*, 549.
(85) Rundolf, T.; Landersjo, C.; Lycknert, K.; Maliniak, A.; Widmalm, G. *Magn. Reson. Chem.* **1998**, *36*, 773.
(86) Bolon, P. J.; Prestegard, J. H. *J. Am. Chem. Soc.* **1998**, *120*, 9366.
(87) Shimizu, H.; Donohue-Rolfe, A.; Homans, S. W. *J. Am. Chem. Soc.* **1999**, *121*, 5815.
(88) Stevens, E. S.; Sathyanarayana, B. K. *J. Am. Chem. Soc.* **1989**, *111*, 4149.
(89) Stevens, E. S.; Duda., C. A. *J. Am. Chem. Soc.* **1991**, *113*, 8622.
(90) Stevens, E. G. *Biopolymers* **1994**, *34*, 1403.
(91) Stevens, E. G. *Biopolymers* **1994**, *34*, 1395.
(92) Hardy, B. J.; Bystricky, S.; Kovac, P.; Widmalm, G. *Biopolymers* **1996**, *41*, 83.
(93) Allen, F. H.; Kennard, O. *Chem. Des. Autom. News* **1993**, *8*, 31.
(94) Otter, A.; Lemieux, R. U.; Ball, R. G.; Venot, A. P.; Hindsgaul, O.; Bundle, D. R. *Eur. J. Biochem.* **1999**, *259*, 295.
(95) Pérez, S.; Mouhous-Riou, N.; Nifant'ev, N. E.; Tsvetkov, Y. E.; Bachet, B.; Imberty, A. *Glycobiology* **1996**, *6*, 537.
(96) Watt, D. K.; Brasch, D. J.; Larsen, D. S.; Melton, L. D.; Simpson, J. *Carbohydr. Res.* **1996**, *285*, 1.
(97) Warin, V.; Baert, F.; Fouret, R.; Strecker, C.; Spik, C.; Fournet, B.; Montreuil, J. *Carbohydr. Res.* **1979**, *76*, 11.
(98) Longchambon, F.; Ohanessian, J.; Gillier-Pandraud, H.; Duchet, D.; Jacquinet, J.-C.; Sinay, P. *Acta Crystallogr.* **1981**, *B37*, 601.
(99) Mo, F. *Acta Chem. Scand.* **1979**, *A33*, 207.
(100) Srikrishnan, T.; Chowdhary, M. S.; Matra, K. L. *Carbohydr. Res.* **1989**, *186*, 167.
(101) Delbaere, L. T. J. *Biochem. J.* **1974**, *143*, 197.
(102) Sriram, D.; Lakshmanan, T.; Loganathan, D.; Srinivasan, S. *Carbohydr. Res.* **1998**, *309*, 227.
(103) Imberty, A.; Pérez, S. *Protein Eng.* **1995**, *8*, 699.
(104) Svensson, G.; Albertsson, J.; Svensson, C.; Magnusson, C.; Dahmen, J. *Carbohydr. Res.* **1986**, *146*, 29.
(105) Bugg, C. E. *J. Am. Chem. Soc.* **1973**, *95*, 908.
(106) Hirotsu, K.; Shimada, A. *Bull. Chem. Soc. Jpn.* **1974**, *47*, 1872.
(107) Noordik, J. H.; Beurskens, P. T.; Bennema, P.; Visser, R. A.; Could, R. O. *Z. Kristallogr.* **1984**, *168*, 59.
(108) Mikol, V.; Kosma, P.; Brade, H. *Carbohydr. Res.* **1994**, *263*, 35.
(109) Senma, M.; Taga, T.; Osaka, K. *Chem. Lett.* **1974**, 1415.
(110) Yates, E. A.; Mackie, W.; Lamba, D. *Int. J. Biol. Macromol.* **1995**, *17*, 219.
(111) Ojala, W. H.; Albers, K. E.; Gleason, W. B.; Choo, C. G. *Carbohydr. Res.* **1995**, *275*, 49.
(112) Wright, C. S. *J. Mol. Biol.* **1990**, *215*, 635.
(113) Sauter, N. K.; Hanson, J. E.; Glick, G. D.; Brown, J. H.; Crowther, R. L.; Park, S. J.; Skehel, J. J.; Wiley, D. C. *Biochemistry* **1992**, *31*, 9609.
(114) Stehle, T.; Harrison, S. C. *Structure* **1996**, *4*, 183.
(115) May, A. P.; Robinson, R. C.; Vinson, M.; Crocker, P. R.; Jones, E. Y. *Mol. Cell.* **1998**, *1*, 719.
(116) Imberty, A.; Gautier, C.; Lescar, J.; Pérez, S.; Wyns, L.; Loris, R. *J. Biol. Chem.* **2000**, *275*, 17541.
(117) Delbaere, L. T.; Vandonselaar, M.; Prasad, L.; Quail, J. W.; Wilson, K. S.; Dauter, Z. *J. Mol. Biol.* **1993**, *230*, 950.
(118) Ng, K. K.; Weis, W. I. *Biochemistry* **1997**, *36*, 979.
(119) Merritt, E. A.; Sarfaty, S.; Van Den Akker, F.; L'hoir, C.; Martial, J. A.; Hol, W. G. J. *Protein Sci.* **1994**, *3*, 166.
(120) Merritt, E. A.; Sarfaty, S.; Jobling, M. G.; Chang, T.; Holmes, R. K.; Hirst, T. R.; Hol, W. G. *Protein Sci.* **1997**, *6*, 1516.
(121) Swaminathan, S.; Furey, W.; Pletcher, J.; Sax, M. *Nature Struct. Biol.* **1995**, *2*, 680.
(122) Ling, H.; Boodhoo, A.; Hazes, B.; Cummings, M. D.; Armstrong, G. D.; Brunton, J. L.; Read, R. J. *Biochemistry* **1998**, *37*, 1777.
(123) Bourne, Y.; Roussel, A.; Frey, M.; Rougé, P.; Fontecilla-Camps, J. C.; Cambillau, C. *Proteins* **1990**, *8*, 365.
(124) Weis, W. I.; Drickamer, K.; Hendrickson, W. A. *Nature* **1992**, *360*, 127.
(125) Wright, C. S.; Hester, G. *Structure* **1996**, *4*, 1339.
(126) Naismith, J. H.; Field, R. A. *J. Biol. Chem.* **1996**, *271*, 972.
(127) Loris, R.; Maes, D.; Poortmans, F.; Wyns, L.; Bouckaert, J. *J. Biol. Chem.* **1996**, *271*, 30614.
(128) Rozwarski, D. A.; Swami, B. M.; Brewer, C. F.; Sacchettini, J. C. *J. Biol. Chem.* **1998**, *273*, 32818.
(129) Olson, L. J.; Zhang, J.; Lee, Y. C.; Dahms, N. M.; Kim, J. J. P. *J. Biol. Chem.* **1999**, *274*, 29889.
(130) Bourne, Y.; Rougé, P.; Cambillau, C. *J. Biol. Chem.* **1992**, *267*, 197.
(131) Wright, C. S.; Jaeger, J. *J. Mol. Biol.* **1993**, *232*, 620.
(132) Bourne, Y.; Mazurier, J.; Legrand, D.; Rougé, P.; Montreuil, J.; Spik, G.; Cambillau, C. *Structure* **1994**, *2*, 209.
(133) Bourne, Y.; Bolgiano, B.; Liao, D. I.; Strecker, G.; Cantau, P.; Herzberg, O.; Feizi, T.; Cambillau, C. *Nature Struct. Biol.* **1994**, *1*, 863.
(134) Stehle, T.; Harrison, S. C. *EMBO J.* **1997**, *16*, 5139.
(135) Moothoo, D. N.; Naismith, J. H. *Glycobiology* **1998**, *8*, 173.
(136) Faham, S.; Hileman, R. E.; Fromm, J. R.; Linhardt, R. J.; Rees, D. C. *Science* **1996**, *271*, 1116.
(137) DiGabriele, A. D.; Lax, I.; Chen, D. I.; Svahn, C. M.; Jaye, M.; Schlessinger, J.; Hendrickson, W. A. *Nature* **1998**, *393*, 812.
(138) Jin, L.; Abrahams, J. P.; Skinner, R.; Petitou, M.; Pike, R. N.; Carrell, R. W. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *23*, 14683.
(139) Kabat, E. A. In *Carbohydrates in Solution*; Isbell, H., Ed.; American Chemical Society: Washington, DC, 1973; Vol. 117, pp 334−355.
(140) Watkins, W. M. In *Advances in human genetics*; Harris, H., Hirsch-horn, K., Eds.; Plenum Press: New York, 1980; Vol. 10, pp 1−136.
(141) Hakomori, S. I. *Biophys. Biochim. Acta* **1999**, *1473*, 247.
(142) Bizik, F.; Tvaroska, I. *Chem. Pap.* **1995**, *49*, 202.
(143) Imberty, A.; Mikros, E.; Koca, J.; Mollicone, R.; Oriol, R.; Pérez, S. *Glycoconjugate J.* **1995**, *12*, 331.
(144) Biswas, M.; Rao, V. S. R. *Biopolymers* **1980**, *19*, 1555.
(145) Koca, J.; Pérez, S.; Imberty, A. *J. Comput. Chem.* **1995**, *16*, 796.
(146) Yan, Z. Y.; Bush, C. A. *Biopolymers* **1990**, *29*, 799.
(147) Rao, B. N. N.; Dua, V. K.; Bush, C. A. *Biopolymers* **1985**, *24*, 2207.
(148) Toma, L.; Ciuffreda, P.; Colombo, D.; Ronchetti, F.; Lay, L.; Panza, L. *Helv. Chim. Acta* **1994**, *77*, 668.
(149) Ejchart, A.; Dabrowski, J.; von der Lieth, C. W. *Magn. Res. Chem.* **1992**, *30*, S105.
(150) Widmalm, G.; Venable, R. M. *Biopolymers* **1994**, *34*, 1079.
(151) Bush, C. A.; Yan, Z. Y.; Rao, B. N. N. *J. Am. Chem. Soc.* **1986**, *108*, 6168.
(152) Casset, F.; Peters, T.; Etzler, M.; Korchagina, E.; Nifant'ev, N.; Pérez, S.; Imberty, A. *Eur. J. Biochem.* **1996**, *239*, 710.
(153) Pereira, M. E. A.; Gruezo, F.; Kabat, E. A. *Arch. Biochem. Biophys.* **1979**, *194*, 511.
(154) Loris, R.; De Greve, H.; Dao-Thi, M.-H.; Messens, J.; Imberty, A.; Wyns, L. *J. Mol. Biol.*, **2000**, *301*, 987.
(155) Ni, F. *Prog. Nucl. Magn. Reson. Spectrosc.* **1994**, *26*, 517.
(156) Poveda, A.; Jiménez-Barbero, J. *Chem. Soc. Rev.* **1998**, *27*, 133.
(157) Gohier, A.; Espinosa, J. F.; Jiménez-Barbero, J.; Carrupt, P. A.; Pérez, S.; Imberty, A. *J. Mol. Graph.* **1996**, *14*, 322.
(158) Hindsgaul, O.; Khare, D. P.; Bach, M.; Lemieux, R. U. *Can. J. Chem.* **1985**, *63*, 2653.
(159) Hindsgaul, O.; Norberg, T.; Le Pendu, J.; Lemieux, R. U. *Carbohydr. Res.* **1982**, *109*, 109.
(160) Bizik, F.; Tvaroska, I. *Chem. Pap.* **1996**, *50*, 84.
(161) Yvelin, F.; Zhang, Y.-M.; Mallet, J. M.; Robert, F.; Jeannin, Y.; Sinay, P. *Carbohydr. Lett.* **1996**, *1*, 475.
(162) Jeffrey, P. D.; Bajorath, J.; Chang, C. Y.; Yelton, D.; Hellstrom, I.; Hellstrom, K. E.; Sheriff, S. *Nat. Struct. Biol.* **1995**, *2*, 466.
(163) Miller, K. E.; Mukhopadhyay, C.; Cagas, P.; Bush, C. A. *Biochemistry* **1992**, *31*, 6703.
(164) Poveda, A.; Asensio, J. L.; Martin-Pastor, M.; Jiménez-Barbero, J. *Carbohydr. Res.* **1997**, *300*, 3.
(165) Berthault, P.; Birlirakis, N.; Rubinstenn, G.; Sinay, P.; Desvaux, H. *J. Biomol. NMR* **1996**, *8*, 23.
(166) Poveda, A.; Asensio, J. L.; Martin-Pastor, M.; Jiménez-Barbero, J. *J. Biomol. NMR* **1997**, *10*, 29.
(167) Bovin, N. In *Glycosciences: status and perspectives*; Gabius, H. J., Gabius, S., Eds.; Chapman & Hall: London, 1997; pp 277−289.
(168) Hakomori, S.; Handa, K.; Iwabuchi, K.; Yamamura, S.; Prinetti, A. *Glycobiology* **1998**, *8*, xi.
(169) Eggens, I.; Fenderson, B.; Toyokuni, T.; Dean, B.; Stroud, M.; Hakamori, S. *J. Biol. Chem.* **1989**, *264*, 9476.
(170) Kojima, N.; Fenderson, B. A.; Stroud, M. R.; Goldberg, R. I.; Habermans, R.; Toyokuni, T.; Hakomori, S. *Glycoconjugate J.* **1994**, *11*, 238.
(171) Boubelik, M.; Floryk, D.; Bohata, J.; Draberova, L.; Macak, J.; Smid, F.; Draber, P. *Glycobiology* **1998**, *8*, 139.
(172) Imberty, A.; Breton, C.; Oriol, R.; Mollicone, R.; Pérez, S. *Adv. Macromol. Carbohydr. Res.*, in press.
(173) Bersch, B.; Koehl, P.; Nakatani, Y.; Ourisson, G.; Milon, A. *J. Biomol. NMR* **1993**, *3*, 443.
(174) Geyer, A.; Gege, C.; Schmidt, R. R. *Angew. Chem., Int. Ed. Engl.* **1999**, *38*, 1466.
(175) Henry, B.; Desvaux, H.; Pristchepa, M.; Berthault, P.; Zhang, Y. M.; Mallet, J. M.; Esnault, J.; Sinay, P. *Carbohydr. Res.* **1999**, *315*, 48.
(176) Ichikawa, Y.; Lin, Y.-C.; Dumas, D. P.; Shen, G.-J.; Garcia-Junceda, E.; Williams, M. A.; Bayer, R.; Ketcham, C.; Walker, L. E.; Paulson, J. C.; Wong, C.-H. *J. Am. Chem. Soc.* **1992**, *114*, 9283.
(177) Hirayama, N.; Yoda, N.; Nishi, T. *Chem. Lett.* **1994**, 1479.
(178) Mukhopadhyay, C.; Miller, K. E.; Bush, C. A. *Biopolymers* **1994**, *34*, 21.
(179) Rutherford, T. J.; Spackman, D. G.; Simpson, P. J.; Homans, S. W. *Glycobiology* **1994**, 59.
(180) Imberty, A.; Pérez, S. In *Carbohydrate mimics. Concepts and Methods*; Y., C., Ed.; Wiley VCH: Weinheim, 1998; pp 349−364.
(181) Ball, G. E.; O'Neill, R. A.; Schultz, J. E.; Lowe, J. B.; Weston, B. W.; Nagy, J. O.; Brown, E. G.; Hobbs, C. J.; Bednarski, M. D. *J. Am. Chem. Soc.* **1992**, *114*, 5449.

(182) Mukhopadhyay, C.; Miller, K. E.; Bush, C. A. *Biopolymers* **1994**, *34*, 21.
(183) Poppe, L.; Brown, G. S.; Philo, J. S.; Nikrad, P. V.; Shah, B. H. *J. Am. Chem. Soc.* **1997**, *119*, 1727.
(184) Scheffler, K.; Brisson, J. R.; Weisemann, R.; Magnani, J. L.; Wong, W. T.; Ernst, B.; Peters, T. *J. Biomol. NMR* **1997**, *9*, 423.
(185) Harris, R.; Kiddle, G. R.; Field, R. A.; Ernst, B.; Magnani, J. L.; Homans, S. W. *J. Am. Chem. Soc.* **1999**, *121*, 2546.
(186) Graves, B. J.; Crowther, R. L.; Chandran, C.; Rumberger, J. M.; Li, S.; Huang, K. S.; Presky, D. H.; Familletti, P. C.; Wolitzky, B. A.; Burns, D. K. *Nature* **1994**, *367*, 532.
(187) Cooke, R. M.; Hale, R. S.; Lister, S. G.; Shah, G.; Weir, M. P. *Biochemistry* **1994**, *33*, 10591.
(188) Hensley, P.; McDevitt, P. J.; Brooks, I.; Trill, J. J.; Feild, J. A.; McNulty, D. E.; Connor, J. R.; Griswold, D. E.; Kumar, N. V.; Kopple, K. D.; Carr, S. A.; Dalton, B. J.; Johanson, K. J. *J. Biol. Chem.* **1994**, *269*, 23949.
(189) Blanck, O.; Iobst, S. T.; Gabel, C.; Drickamer, K. *J. Biol. Chem.* **1996**, *271*, 7289.
(190) Kjellen, L.; Lindahl, U. *Annu. Rev. Biochem.* **1991**, *60*, 443.
(191) Lane, D. A.; Lindhal, U. *Heparin*; Edward Arnold: London, 1989.
(192) Conrad, H. E. *Heparin-binding proteins*; Academic Press: San Diego, 1998.
(193) Choay, J.; Petitou, M.; Lormeau, J.-C.; Sinay, P.; Casu, B.; Gatti, G. *Biochem. Biophys. Res. Commun.* **1983**, *116*, 492.
(194) Thunberg, L.; Bäckström, G.; Lindahl, U. *Carbohydr. Res.* **1982**, *100*, 393.
(195) van Boeckel, C. A. A.; Petitou, M. *Angew. Chem., Int. Ed. Engl.* **1993**, *32*, 1671.
(196) Ragazzi, M.; Ferro, D. R.; Perly, B.; Sinay, P.; Petitou, M.; Choay, J. *Carbohydr. Res.* **1990**, *195*, 169.
(197) Cros, S.; Petitou, M.; Sizun, P.; Pérez, S.; Imberty, A. *Bioorg. Med. Chem.* **1997**, *5*, 1301.
(198) Hricovini, M.; Guerrini, M.; Bisio, A. *Eur. J. Biochem.* **1999**, *261*, 789.
(199) Skinner, R.; Abrahams, J. P.; Whisstock, J. C.; Lesk, A. M.; Carrell, R. W.; Wardell, M. R. *J. Mol. Biol.* **1997**, *266*, 601.
(200) Grootenhuis, P. D. J.; van Boeckel, C. A. A. *J. Am. Chem. Soc.* **1991**, *113*, 2743.
(201) Bitomsky, W.; Wade, R. C. *J. Am. Chem. Soc.* **1999**, *121*, 3004.
(202) Mulloy, B.; Forster, M. J.; Jones, C.; Davies, D. B. *Biochem. J.* **1993**, *293*, 849.
(203) Mikhailov, D.; Mayo, K. H.; Vlahov, I. R.; Toida, T.; Pervin, A.; Linhardt, R. J. *Biochem. J.* **1996**, *318*, 93.
(204) Mikhailov, D.; Linhardt, R. J.; Mayo, K. H. *Biochem. J.* **1997**, *328*, 51.
(205) Faham, S.; Linhardt, R. J.; Rees, D. C. *Curr. Opin. Struct. Biol.* **1998**, *8*, 578.
(206) Lam, K.; Rao, V. S. R.; Qasba, P. K. *J. Biomol. Struct. Dyn.* **1998**, *15*, 1009.
(207) Huhtala, M. T.; Pentikainen, O. T.; Johnson, M. S. *Structure* **1999**, *7*, 699.
(208) Stuckey, J. A.; St Charles, R.; Edwards, B. F. *Proteins* **1992**, *14*, 277.
(209) Holmbeck, S. M. A.; Petillo, P. A.; Lerner, L. E. *Biochemistry* **1994**, *33*, 14246.
(210) Miertus, S.; Bella, J.; Toffanin, R.; Matulova, M.; Paoletti, S. *J. Mol. Struct. (THEOCHEM)* **1997**, *395*−*396*, 437.
(211) Cowman, M. K.; Hittner, D. M.; Feter-Davis, J. *Macromolecules* **1996**, *29*, 2894.
(212) Sicinska, W.; Lerner, L. E. *Carbohydr. Res.* **1996**, *286*, 151.
(213) Kitov, P. I.; Sadowska, J. M.; Mulvey, G.; Armstrong, G. D.; Ling, H.; Pannu, N. S.; Read, R. J.; Bundle, D. R. *Nature* **2000**, *403*, 669.
(214) Simanek, E. E.; McGarvey, G. J.; Jablonowski, J. A.; Wong, C.-H. *Chem. Rev.* **1998**, *98*, 833.
(215) Geyer, A.; Reinhardt, S.; Bendas, G.; Rothe, U.; Schmidt, R. R. *J. Am. Chem. Soc.* **1997**, *119*, 11707.
(216) Thoma, G.; Schwarsenbach, F.; Duthaler, R. O. *J. Org. Chem.* **1996**, *61*, 514.
(217) Janhke, W.; Kolb, H. C.; Blommers, M. J. J.; Magnani, J. L.; Ernst, B. *Angew. Chem., Int. Engl.* **1997**, *36*, 2603.
(218) Kolb, H. C.; Ernst, B. *Chem. Eur. J.* **1997**, *3*, 1571.
(219) Kishi, Y. *Pure Appl. Chem.* **1993**, *65*, 313.
(220) Nilsson, U.; Johansson, R.; Magnusson, G. *Chem. Eur. J.* **1996**, *2*, 295.
(221) Aguilera, B.; Jiménez-Barbero, J.; Fernadez-Mayoralas, A. *Carbohydr. Res.* **1998**, *308*, 19.
(222) Weimar, T.; Kreis, U. C.; Andrews, J. S.; Pinto, B. M. *Carbohydr. Res.* **1999**, *315*, 222.
(223) Espinosa, J.-F.; Bruix, M.; Jarreton, O.; Skrydstrup, T.; Beau, J.-M.; Jiménez-Barbero, J. *Chem. Eur. J.* **1999**, *5*, 442.
(224) Espinosa, J. F.; Canada, F. J.; Asensio, J. L.; Martin-Pastor, M.; Dietrich, H.; Martin-Lomas, M.; Schmidt, R. R.; Jiménez-Barbero, J. *J. Am. Chem. Soc.* **1996**, *118*, 10862.
(225) Rubinstenn, G.; Sinaÿ, P.; Berthault, P. *J. Phys. Chem. A* **1997**, *101*, 2536.
(226) Ravishankar, R.; Surolia, A.; Lim, M. V. S.; Kishi, Y. *J. Am. Chem. Soc.* **1998**, *120*, 11297.
(227) Asensio, J. L.; Espinosa, J. F.; Dietrich, H.; Cañada, F. J.; Schmidt, R. R.; Martín-Lomas, M.; André, S.; Gabius, H.-J.; Jiménez-Barbero, J. *J. Am. Chem. Soc.* **1999**, *121*, 8995.
(228) Espinosa, J. F.; Montero, E.; Vian, A.; Garcia, J. L.; Dietrich, H.; Schmidt, R. R.; Martin-Lomas, M.; Imberty, A.; Canada, F. J.; Jiménez-Barbero, J. *J. Am. Chem. Soc.* **1998**, *120*, 1309.
(229) Driguez, H. *Top. Curr. Chem.* **1997**, *187*, 85.
(230) Toone, E. J. *Curr. Opinion Struct. Biol.* **1994**, *4*, 719.
(231) Lemieux, R. U. *Acc. Chem. Res.* **1996**, *29*, 373.
(232) Carver, J. P. *Pure Appl. Chem.* **1993**, *65*, 763.
(233) Chervernak, M. C.; Toone, E. J. *J. Am. Chem. Soc.* **1994**, *116*, 10533.
(234) Swaminathan, C. P.; Surolia, N.; Surolia, A. *J. Am. Chem. Soc.* **1998**, *120*, 5153.
(235) Garcia-Hernandez, E.; Hernandez-Arana, A. *Protein Sci.* **1999**, *8*, 1075.
(236) Searle, S.; Willia, D. H. *J. Am. Chem. Soc.* **1992**, *114*, 10690.
(237) Wacowich-Sgarbi, S. A.; Bundle, D. R. *J. Org. Chem.* **1999**, *64*, 9080.
(238) Alibes, R.; Bundle, D. R. *J. Org. Chem.* **1998**, *63*, 6288.
(239) Bundle, D. R.; Alibès, R.; Nilar, S.; Otter, A.; Warwas, M.; Zhang, P. *J. Am. Chem. Soc.* **1998**, *120*, 5317.
(240) Navarre, N.; van Oijen, A. H.; Boons, G. J. *Tetrahedron Lett.* **1997**, *38*, 2023.
(241) Navarre, N.; Amiot, N.; Van Oijen, A.; Imberty, A.; Poveda, A.; Jiménez-Barbero, J.; Cooper, A.; Nutley, M. A.; Boons, G. J. *Chem. Eur. J.* **1999**, *5*, 2281.
(242) Geyier, A.; Müller, M.; Scmidt, R. R. *J. Am. Chem. Soc.* **1999**, *121*, 6312.
(243) Liang, G.; Schmidt, R. K.; Yu, H. A.; Cumming, D. A.; Brady, J. W. *J. Phys. Chem.* **1996**, *100*, 2528.
(244) Pathiaseril, A.; Woods, R. J. *J. Am. Chem. Soc.* **2000**, *122*, 331.
(245) Andersson, C.; Engelsen, S. B. *J. Mol. Graph. Model.* **1999**, *17*, 101.
(246) Vishnyakov, A.; Widmalm, G.; Laaksonen, A. *Angew. Chem., Int. Ed.* **2000**, *39*, 140.
(247) Spek, A. L. *J. Appl. Crystallogr.* **1988**, *21*, 578.
(248) MSI, 3.5 ed.; http://www.msi.com, 1999.
(249) Schwarz, F. P.; Puri, K. D.; Bhat, R. G.; Surolia, A. *J. Biol. Chem.* **1993**, *268*, 7668.
(250) Williams, B. A.; Chervenack, M. C.; Toone, E. J. *J. Biol. Chem.* **1992**, *267*, 22907.
(251) Mandal, D. K.; Kishore, N.; Brewer, C. F. *Biochemistry* **1994**, *33*, 1149.
(252) Chervenak, M. C.; Toone, E. J. *Biochemistry* **1995**, *34*, 5685.

CR990343J