We are IntechOpen, the world’s leading publisher of Open Access books
Built by scientists, for scientists

3,800
Open access books available

116,000
International authors and editors

120M
Downloads

154
Countries delivered to

TOP 1%
Our authors are among the most cited scientists

12.2%
Contributors from top 500 universities

WEB OF SCIENCE™
Selection of our books indexed in the Book Citation Index in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com
The Molecular Structure and Conformational Dynamics of Chitosan Polymers: An Integrated Perspective from Experiments and Computational Simulations

Richard A. Cunha, Thereza A. Soares, Victor H. Rusu, Frederico J.S. Pontes, Eduardo F. Franca and Roberto D. Lins

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/51803

1. Introduction

Chitosan is a semi-synthetic linear copolymer composed of a variable number of $\beta$-(1-4) linked units of 2-acetamide-2-deoxy-$\beta$-d-glucopyranose (GlcNAc) and 2-amino-2-deoxy-$\beta$-d-glycopyranose (GlcN) [1]. The two monomers differ with respect to the C2-substituent in the sugar ring, which is either an amino or acetamide group (Figure 1). Chitosan is obtained via the alkaline deacetylation of chitin. However, the deacylation reaction hardly proceeds completely in a normal heterogeneous reaction, leading to a random distribution of GlcNAc and GlcN residues in the chitosan polymer [2, 3]. The degree of acetylation of a polymer is a measure of the average number of GlcNAc per 100 chitosan monomers in percentile unit. The degree of acetylation governs important physical-chemical properties of the chitosan polymer such as solubility and conformation, being critical for the effectiveness of various technological applications [4-6]. The threshold for the conversion of chitin into chitosan depends on the solubility of the oligosaccharide in a slightly acid solution (0.1 mol/L of acetic acid). Conventionally, chitin polymers with a level of acetylation below 50% is considered as chitosan. The level of protonated amino groups in the glucosamine monomers dictates the solubility of chitosan, conferring the cationic nature to the polymer [7].

Chitin is the most abundant amino polysaccharide, being produced in the amount of one hundred billion tons per year in nature [8]. Its main source is the exoskeleton of crabs and shrimps, whose availability in nature makes chitin a renewable source of chitosan. In the last decades chitosan has emerged as a biomaterial with unique properties for advanced applications in green chemistry, biomedical, pharmaceutical, food [9] and agriculture [10].

INTECH © 2012 Cunha et al., licensee InTech. This is an open access chapter distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/3.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.
The variety of applications of chitosan is determined by its chemical structure, which varies with respect to size (average molecular weight; MW), degree of acetylation (DA) and numerous chemical modifications [6, 11-15]. An increasing number of chitosan chemical modifications have been described in the literature [14, 15]. Chitosan is also a highly absorptive material used as heavy-metal chelators in water [16, 17]. The chelation of metals occurs via electrostatic interactions with chitosan reactive groups (hydroxyl, acetamide and primary amino groups). These functional groups are also responsible for properties such as high hydrophilicity, reactivity, and structural flexibility that make chitosan soluble in near-neutral acid solution [17]. The soluble adsorptive properties of chitosan enable its use for removal of pesticide and dyes from water, for adsorption of proteins, as flocculant agent, and even as a catalyst support for biodiesel production [18-24]. Besides being an abundant renewable resource, chitosan stands out due to some unique properties such as exceptional biocompatibility, biodegradability, non-toxicity and the easiness of production of chemically modified forms [14, 25, 26]. Furthermore, chitosan exhibits antimicrobial and antifungal activity [27-30]. These singular properties make chitosan well-suited for a wide range of biomedical applications such as drug delivery, platform for neural stem cell growth, tissue engineering (bone, cartilage, nerve, skin), immunoprophylaxis, gene therapy, wound healing and treatment of infections [31-44].

Figure 1. Schematic representation of a disaccharide formed by 2-amino-2-deoxy-β-d-glycopyranose (GlcN) bound to 2-acetamide-2-deoxy-β-d-glucopyranose (GlcNac). Φ and Ψ dihedral angles are shown in the scheme in red and here are represented by atoms: O5-C1-O1-C4' and C1-O1-C4'-C5', respectively. GlcN and GlcNac are the two units forming chitosan. Solubility is typically achieved when GlcN > 50%.

This short review will cover the structural dynamics of chitosan from a microscopic perspective, focusing on the interplay between its conformational variability and macroscopic properties such as solubility and aggregation. It is not the goal of the review to provide a detailed summary of the extensive literature on carbohydrate structure and its characterization. Excellent reviews on the subject can be found in the literature [45-47]. The text is organized in four main sections. First, we present an overview of chemical interactions between chitosan and biological materials, in particular lipid bilayers. In what follows, we describe the advantages and limitations of experimental and computational techniques used for the structural characterization of oligosaccharides, emphasizing the necessity of combining different approaches in order to obtain high-resolution structural data on chitosan. In this section we also introduce the theoretical principles underlying molecular dynamics (MD) simulations, which has been widely used to study the structural
dynamics of carbohydrates in solution. The third section, we review the types of secondary structure observed for chitin and chitosan in the crystalline state. In the final section, we offer a detailed account of the structure and conformational dynamics of chitosan in solution as unveiled by computational simulations carried out in our group.

2. The molecular interactions underlying chitosan bioactivity

Chitosan is a very promising material with wide range of biomedical applications. This oligosaccharide incorporates highly sought properties for biomedical applications (biocompatibility, biodegradability and bioresorbability) with the easy processing into gels, membranes, nanofibers, beads, microparticles, nanoparticles, scaffolds and sponges forms [48-52]. Yet, chitosan has a flexible, hydrophilic helical structure with reactive amine groups, which offers a multitude of possible inter- and intra-molecular interactions. A detailed understanding of the effects of different materials and environmental conditions on such interactions can enable the design of novel chitosan-based technologies.

Chitosan amino groups are the major players in metal chelating processes. However, it has been previously shown that electrostatic interactions involving the protonated amino groups per se in low pH are not sufficient to explain the biological behavior of chitosan in presence of biological membranes [53]. A comparative study between chitosan and a fully cationic polymer has shown that hydrophobic interactions play an important role in the polysaccharide action. A study on the effect of the pH and the molecular weight of chitosan in multilamellar vesicles of dipalmitoylphosphatidylcholine (DPPC) has shown that increasing the biopolymer molecular weight (213 kDa) and decreasing of pH can lead to disruption of the membrane [54]. In contrast, another study has shown that the interactions between chitosan and DPPC lipids in liposomes led to an increase in thermodynamic stability of the composite. It has been proposed that this stabilization results from a shielding mechanism based on the electrostatic interactions between the chitosan chains and phospholipids polar heads [55]. In contrast, anti-fungal and bactericidal activity of chitosan has been attributed to the ability of chitosan to disrupt the inner and outer membranes of cells [56, 57]. The contribution of electrostatic, hydrophobic and hydrogen bond interactions between chitosan and three different lipids have been evaluated using Langmuir films to mimic the interaction of chitosan and bacterial membranes [29]. It has been shown that chitosan had a negligible effect on DPPC monolayers but it distinctly affected dipalmitoylphosphatidylglycerol (DPPG) and cholesterol monolayers [29]. The effect on DPPG was found to decrease with increasing pH, ascribed to the charge-mediating action of chitosan, whereas the pH did not affect the cholesterol monolayers where interactions occurs mainly via hydrogen bonding. A recent study has suggested that sensitivity of fungi to chitosan depends of the membrane fluidity and dynamics [58]. The same group has also suggested in a previous work that chitosan kills fungal cells by an unknown mechanism that does not involve endocytosis [59]. Although chitosan has been used in a variety of biologically relevant applications involving interactions with lipids, proteins, inorganic and organic compounds, a microscopic picture of these interactions remains lacking. The conformational flexibility of chitosan has hampered the acquisition of high-resolution
structural data through of X-ray crystallography and NMR spectroscopy [46, 47]. However, current molecular modeling techniques can be used to bridge the gap of experimental resolution, thus providing complementary information to measurements.

3. Structural characterization of polysaccharides

Several analytical techniques have been applied to the characterization of a variety of oligosaccharide properties [60-63]. Among them, the molecular geometry is one of the most important properties that experimental data can provide on carbohydrates. Its characterization is critical for the understanding of the function and recognition mechanisms of carbohydrates in living organisms. However, sugars are inherently flexible, undergoing conformational changes in response to chemical modifications, complexation to biomolecules, changes in the pH, ionic strength and solvent type [2]. In solution, oligosaccharides tend to adopt a coiled conformation, which fluctuates between local and overall conformations, adopting an enormous variety of spatial arrangements around glycosidic linkages.

As a first approach to the complexity of the conformational flexibility of polysaccharides, let us assume that its monosaccharide units have a rigid ring structure. Thus, the determination of the conformation of oligosaccharide structures is reduced to the characterization of the glycosidic linkages between rigid monosaccharide monomers, i.e., the description of two or three torsion angles for each glycosidic linkage would suffice to characterize the conformations of the oligosaccharide chain. However, the description of these torsion angles faces two major issues [46, 64]. First, the motions associated to each glycosidic linkage range across large-scale vibrations of a single well-defined conformation to transitions between several different conformations. Therefore, the accurate characterization of a given glycosidic linkage requires information on the number of conformations adopted, the time spent in each conformation and the flexibility of each conformation [46]. An additional difficulty is given by the fact that the conformational transitions in different linkages of an oligosaccharide chain are coupled. Second, the two experimental techniques most effective in providing atomic level structural information on biomolecules, namely X-ray diffraction and nuclear magnetic resonance (NMR) spectroscopy, have appreciable limitations when applied to oligosaccharides. In this section, we briefly overview the strengths and limitations of the most representative experimental techniques used for structural characterization of chitin and chitosan.

Mass spectrometry (MS) can be used to determine the total mass of a carbohydrate or differentiate distinct oligosaccharides as function of the respective weight [65]. Although MS cannot offer a detailed description of the oligosaccharide structure, it can identify the location of branch points [66-69]. Further fragmentation will not result in additional information, because the fragments can be alike. Despite the inadequacy of MS to generate information on the molecular geometry of oligosaccharides, MS can be coupled with separation techniques such as high performance liquid chromatography (HPLC) to differentiate between solutions containing different types of carbohydrates [65, 70-73].
The techniques of X-ray diffraction and NMR spectroscopy determine time and spatial averages of molecular properties in atomic coordinates measured from an ensemble of molecules corresponding the Avogadro’s number [74-76]. Yet, the two techniques differ significantly with respect to the spatial distribution of molecules and time scale accessible to each one [74-76]. X-ray data represents an average over molecules arranged in a periodic crystal lattice over the second to hour timescale whereas NMR data represents an average over semi-randomly oriented molecules in solution over the nanosecond to second timescale. Despite the robustness of single X-ray crystallography in protein structure determination, the technique is not easily applicable to oligosaccharides. The difficulty to obtain highly crystalline samples for oligosaccharides imposes limits on the quality of the diffraction pattern. X-ray diffraction of carbohydrate structures larger than tetramers are rare and only seen when co-crystallized with proteins [18]. Due to the difficulty to obtain single crystals from oligosaccharides, oriented fibers have often been used for X-ray diffraction studies. Fibers exhibit helical symmetry rather than the three-dimensional symmetry seen in single crystals. Analysis of the diffraction pattern from orientated fibers allows deducing the helical symmetry of the molecule, in some cases also the structure. This is possible by constructing a model of the fiber and calculating the expected diffraction pattern. By comparing the calculated and observed diffraction patterns one eventually arrives at a better model. However, oligosaccharides in crystalline fibers can be affected by intra-molecular and crystal lattice packing, which may lock the structure in a conformation not representative of the conformational ensemble in solution.

On the other hand, NMR spectroscopy can provide information on the covalent structure and the complex conformational equilibria of oligosaccharides in solution [46, 47, 64, 77-79]. Moreover, it is the only available technique that can determine both the anomericities and linkages of a novel glycan. Another practical advantage of NMR spectroscopy is the possibility of measuring relative dilute solutions of oligosaccharides. Sample requirement amounts to as little as 1 mg, which is within the limits of enzyme-assisted synthesis [47]. NMR spectroscopy is probably the most used experimental tool to characterize the atomic structure of carbohydrates. For this reason, it has been the subject of numerous reviews [46, 47, 64, 78]. Biomolecular NMR spectroscopy has progressed appreciably in the last decades. Developments in the instrumentation, pulse sequences and spectral interpretation associated to molecular modeling techniques led to great advances in the determination of primary and three-dimensional structures of biomolecules in solution [17]. Such progress has been more noticeable in the structural characterization of proteins and nucleic acids. Notwithstanding, carbohydrates are not too far behind despite difficulties in proton assignment of each atom due to the structural similarity of the monosaccharide units [19-29].

The NMR spectroscopy data reflect primarily short-range through-bond interactions (J-coupling constants), short-range through-space interactions via the nuclear Overhauser effect (NOEs) or local perturbations to electronic shielding (chemical shifts). NOE is the main source of conformational information on carbohydrates. The strength of the NOE signal between two nuclei is proportional to the inverse sixth power of the distance between the atoms. However, a given distance between two protons is often consistent with a range
of distinct conformations that will be represented by a set of NOE-derived distance constraints. The larger the number of available NOE constraints, the more consistent a single structure will be with this collection of spatial constraints. Nonetheless, the use of NOE constraints in structure determination of oligosaccharides is beset by a few issues (reviewed in [46]). For instance, the number of NOE constraints may not suffice to unambiguously define a conformation, particularly around the glycosidic linkage. In addition, NOE is sensitive only to short-distance nuclei (< 5-6 Å). For that reason, NOE constraints are obtained only between nuclei within a monosaccharide unit or across a glycosidic linkage. Due to the lack of long-range information, the accurate determination of the whole structure of oligosaccharides depends on combining the local conformations of the individual linkages. Such procedure leads to the cumulative addition of any uncertainties or errors in the local structure and its dissemination to the whole oligosaccharide structure (except if in the presence of sufficient sequential NOEs). A last issue concerns space-time ensemble averaging effects. Accordingly, different NMR parameters are averaged over time scales ranging from 50 ms to 1 s. In the case of oligosaccharides transitioning between several conformations, NOE constraints will represent average values that cannot be easily decomposed into each of the single conformation contributing to the average constraints. The conformational uncertainty ensuing from these issues can be minimized to some degree by the use of additional conformational constraints such as scalar coupling constants (J values), which are simple linear averages over the ensemble of individual conformers. \(^{1}H-^{1}H\) J-coupling constants can be used to define ring conformations and dihedral angles. This information can also be obtained via NMR residual dipolar couplings measurements. Non-anomeric protons can be assigned through 2D homonuclear correlation COSY and/or TOCSY experiments. NOESY experiments can be used to provide monosaccharide sequence information due to the absence of coupling over the glycosidic linkage of the COSY and TOCSY spectra. \(^{1}H-^{13}C\) HSQC and HMQC experiments provide important correlations in the determination of repeating units of polysaccharides [2]. Yet, the identification of distinct carbohydrate conformations requires combining complementary techniques. These techniques vary from other experimental methods to atomistic molecular dynamics simulations [46, 64, 80-82].

Classical molecular dynamics (MD) simulations can be used to complement incomplete experimental data and to provide detailed conformational distributions in time and space that experimental measurements can only obtain as averages [80, 82, 83]. It can also be used to predict properties under environmental conditions that may not be accessible to experimental measurements. As such, MD is an indispensable tool to interpret experimental data. However, the accuracy of MD simulations is intrinsically dependent on the quality of the empirical potential energy functions and the force-field parameters used. Robust force fields for MD simulations of carbohydrate-based systems are available. Some of the most used are CHARMM[84-87], GLYCAM/AMBER[88, 89], GROMOS[90] and OPLS-AA[91]. These force fields offer a realistic description of the structural dynamics of oligosaccharides within the limitations of the experimental data available, making MD simulations a reliable procedure for the prediction of molecular interactions [92-95]. Therefore, the importance of
accurate measurements of the spatial arrangements of carbohydrates can never be overestimated as they are the principal component in the development of physical chemical parameters (force fields) governing molecular simulations. The availability of high-quality experimental data is critical for biomolecular modeling [80, 81]. Classical force fields used for simulations of biomolecules are built from quantum chemistry calculations and/or experimental measurements. Without experimental measurements, the development of classical force fields would be extremely difficult as the expensive costs of quantum-chemical theoretical models limit their use in force-field construction [80, 90, 96-100]. In addition, quantum-mechanical data is not an ideal validation target as it only yields gas-phase quantities. Model validation and comparisons of biomolecular simulations are often best done against condensed-phase experimental data [82, 98, 101]. The availability of structural data on carbohydrates has made possible the creation of several databases like the SUGABASE, CarbBank, EUROCarbDB, Glycoconjugate DB, GLYCOSCIENCES.de, GlycoSuiteDB, JCGGDB, KEGG-Glycan, CFG-Glycan Database and GlycoBase. These databases represent a convenient tool for the building of molecular models as well as for comparison of atomistic simulations of carbohydrates against experimental data.

3.1. Theoretical foundations of molecular dynamics simulations

The MD method has its foundations in the laws of classical mechanics [102, 103]. It allows the simulation of the time-dependent behavior of molecular systems according to Newton’s laws of motion. The atom nuclei are treated classically as spheres connected to each other through a set of springs emulating chemical bonds. The forces acting on each atom, necessary to simulate their motion, are derived from a set of force field parameters, and the set of coordinates and velocities that mapped during the whole process comprise the phase space. In a simulation, the force $F$ on each atom is expressed as a function of time, and is equal to the negative gradient of the potential energy $V$ with respect to the position $r_i$ of each atom, in a distinct expression of the more common form of the equation $F = ma$:

$$\frac{dV}{dr_i} = m \frac{d^2r_i}{dt^2}$$

The MD method integrates iteratively and numerically the classical equations of motion for every atom in the system at time increments ($\Delta t$ – time step) defined by the user. A number of algorithms exist for this purpose and are implemented in different computational codes [104]. There are several algorithms available for performing the numerical integration of the equations of motion. The Verlet-type algorithms (Verlet, velocity-Verlet and leap-frog) are widely used because it requires a minimum amount of computer memory and CPU time [105, 106]. The velocity Verlet, for instance, uses positions, velocities and accelerations at the current time step, which gives a more accurate integration than the original Verlet algorithm. Other algorithms, as the Beeman gives better energy conservation at the expense of computer memory and CPU time [107]. The Gear predictor-corrector algorithm predicts the next set of positions and accelerations, and then compares the accelerations to the
predicted ones to compute a correction for the step \([108]\). Each step can thus be refined iteratively. Predictor-corrector algorithms give an accurate integration but are seldom used due to their large computational cost. In the classical Verlet algorithm, the positions in the next time step \((\Delta t)\) are calculated from a given set of particles with coordinates \(r_i\) using a Taylor expansion:

\[
\begin{align*}
    r_{i+1} &= r_i + \frac{\partial r}{\partial t} (\Delta t) + \frac{1}{2} \frac{\partial^2 r}{\partial t^2} (\Delta t)^2 + \frac{1}{6} \frac{\partial^3 r}{\partial t^3} (\Delta t)^3 + \ldots \\
    r_{i+1} &= r_i + \mathbf{v}_i (\Delta t) + \frac{1}{2} \mathbf{a}_i (\Delta t)^2 + \frac{1}{6} \mathbf{b}_i (\Delta t)^3 + \ldots
\end{align*}
\]

where the last equation links the spatial coordinates with the velocities \(v_i\) (the first derivative of the positions in respect to time \((dr/dt)\) at time \(t_i\). The accelerations \(a_i\) (the second derivatives \((d^2r/dt^2)\) at time \(t_i\) and so on. If the goal is to determine the positions for a small time step \((\Delta t)\) earlier, the equation becomes:

\[
r_{i-1} = r_i - \mathbf{v}_i (\Delta t) + \frac{1}{2} \mathbf{a}_i (\Delta t)^2 - \frac{1}{6} \mathbf{b}_i (\Delta t)^3 + \ldots
\]

Adding the last two equations makes it possible to find a new equation that predicts the position at a chosen time step using the current and previous atom positions and current acceleration. The latter can be calculated from the force or potential.

\[
\begin{align*}
    r_{i+1} &= (2r_i - r_{i-1}) + \mathbf{a}_i (\Delta t)^2 + \ldots \\
    a_i &= \frac{\mathbf{F}_i}{m_j} = -\frac{1}{m_j} \frac{dV}{dr_i}
\end{align*}
\]

At the beginning of the simulation, when the previous positions are not available, this quantity can be estimated from the following approximation:

\[
r_{i-1} = r_i - \mathbf{v}_i \Delta t
\]

The time increment in MD simulation should be sufficiently small that errors in the integration equations keep small, preserving the conservation of the energy. Normally \(\Delta t\) is on the order of femtoseconds \((10^{-15} \text{s})\). This time order is one order of magnitude smaller than the fastest molecular process. Furthermore, because the forces \(F_i\) should be recalculated for every step, MD is a computation intensive task. Currently achieved timescale of MD simulations is on the order of multi-nanoseconds to a few microseconds. This time is shorter than many relevant phenomena, for this reason MD results should be analyzed under the point of view of the sample of the phase space close to the starting condition, in spite of this capacity of sampling different configurations. One strategy is to increase the time step value and to allow longer simulation times is to freeze the bond lengths related to hydrogen atoms. The fastest processes in molecules are stretching vibrations, especially involving hydrogen atoms. As these degrees of freedom have little influence on many properties,
some algorithms were developed to keep frozen these chemical bonds as the SHAKE [109], RATTLE [110] and LINCS [111]. Alternative MD-based methodologies have also been recently developed aiming to partly overcome this limitation. The so-called enhanced sampling techniques artificially drive a given system according a set of pre-defined rules that result in a larger sampling of the configurational phase space within the same simulation time (e.g., simulated annealing, replica-exchange, parallel-tempering, local elevation search, metadynamics) [112, 113].

An accurate description of the aqueous medium that shapes the structure, dynamics and function of biological molecules is essential for the realistic reproduction of its kinetics and thermodynamics properties. It is known that simulations of a small arrangement of atoms do not reproduce satisfactorily the properties of bulk liquids due to surface effects suffered by a large fraction of the molecules. The obvious solution for this problem, which would be to increase the number of solvent molecules, can lead to issues in the evaluation of the force between the atoms. An alternative solution to treat explicit solvent molecules in MD simulations is the use of periodic boundary conditions [104]. In this approach the simulation box is replicated throughout the space to form an infinite lattice, where the number of molecules entering or leaving the simulation box is kept constant during the simulation and as a consequence, surface effects are canceled. There are currently numerous water models used in MD simulations. Some of the models currently implemented in major classical MD softwares are the SPC model [114] and the TIP3P, TIP4P, and TIP5P models [115, 116]. These models were parameterized assuming that a cut-off is applied to nonbonded interactions and treat water as a rigid molecule. Although bond stretching and bond-angle bending [117], or polarization effects and many-body interactions [118], have been introduced into water models, they involve a large increase of computational expense, which has limited their use as widely as the SPC or TIP models. The water models are usually parameterized at a single temperature (ca. 298 K) and therefore may not capture correctly the temperature dependence of properties such as the solvent density or diffusion coefficients [119].

The basic principle underlying the MD theory is that if one allows the system to evolve in time indefinitely, it will eventually pass through all possible states. Thus, MD simulations should cover time scales sufficiently long to generate enough representative conformations to satisfy this principle. In other words, the simulations must sample a sufficient amount of the phase space corresponding to the system in consideration. In that case, experimentally relevant information concerning structural, dynamic and thermodynamic properties can be calculated using a feasible amount of computational resources. The connection between theoretical results and experiments is made through the use of the Ergodic hypothesis. This fundamental axiom of statistical mechanics states that the average obtained by following a small number of particles over a long time is equivalent to averaging over a large number of particles for a short time. Exploring the limit of a sufficient large time scale, the Ergodic hypothesis implies that the time average over a single particle is equivalent to the average over a large number of particles at any given time. This theoretical justification in the scope of a MD simulation validates the calculation of thermodynamic averages for molecular systems if finite molecular dynamics trajectories are “long enough” in the ergodic sense.
In the solid state, chitosan is characterized by an ordered fibrillar structure with a high degree of crystallinity, and polymorphism [120, 121]. X-ray measurements of the chitosan polymer have shown an extended two-fold helix in a zigzag structure [122, 123]. The crystal packing is mainly formed by chitosan chains arranged in an antiparallel fashion (Figure 2A), and similar to the anhydrous form of the α-chitin structure. The structure of the α and β forms differ only in the arrangement of the piles of chains, which is alternately antiparallel in α-chitin and all parallel in β-chitin [92, 124]. The crystallographic structure of chitin and chitosan has also revealed that although both biopolymers exhibit a hydrated and anhydrous forms, chitin occurs exclusively in the conventional extended 2-fold helical conformation (Figure 2A) [123, 125-128]. The presence of free amino groups in the structure of chitosan gives rise to different types of helical conformations in acid (Figure 2) [128]. These structures can be classified in four main types: type I (anhydrous), type II (hydrated), type IIa (hydrated) and type III (anhydrous), which adopt a helical conformations in a two-fold helix, relaxed two-fold helix, a 4/1 helix and a five-fold helix, respectively (Figure 2) [128, 129]. The diversity of chitosan structural types depend on the experimental conditions (kind and concentration of acid, temperature and salt preparation) used for the conversion of chitin into chitosan [128]. The helical structure propensities can be determined according to the repeating unit and helical symmetry as observed in chitosan crystal structures [121, 123, 128, 130, 131]. A less common motif, classified as 3-fold, has also been identified (Figure 2B).

The type I salts are the anhydrous form of the unreacted chitosan crystal. The polysaccharide chains in these crystals have a 2/1 helical symmetry with a repeating pattern of 1.0 nm. This conformation is similar to that of chitin, and characterizes the two-fold helix (Figure 2A) [92, 132, 133]. Type II chitosan exhibits a hydrated crystal with a fiber repeat of about 4.08 nm long and an asymmetric repeating units consisting of tetrasaccharides. In this type, the helical conformation is composed of eight glucosamine residues with repeating units related by a 2/1 helical symmetry. This pattern suggests a two-fold helix even though the corresponding asymmetric unit is rather distinct from that of type I where the asymmetric unit has only one glucosamine residue. The main difference between the type I and type II conformations is that the latter is almost four times longer than chitosan, and originated the designation of relaxed two-fold helix (Figure 2E) [92, 134-136]. A type II salt variant, called Type IIa, has a similar fiber repeat (4.05 nm), but with an asymmetric unit consisting of a glucosamine dimer in a 4/1 helical symmetry. This right-handed helix comprised of four asymmetric subunits is classified as 4/1-helix conformation, being also called four-fold helix (Figure 2C) [121, 129]. The most recently discovered type III form has a chain repeat of 2.55 nm, a 5/3 helical symmetry, and an asymmetric unit of a single glucosamine residue. Type III helical conformation is classified as five-fold helix, and displays a less symmetric helicoidal conformation (Figure 2D) [129, 137].
In solid state, the two-fold helix pattern is stabilized by O3-HO3•••O5' intra-chain hydrogen bonds across the glycosidic linkages [120]. In order to verify these helical properties in aqueous solution, MD simulations were carried out for chitin and chitosan [92, 93]. These simulations have shown that chitin chains assume exclusively a two-fold helix conformation which indeed is stabilized by the O3-HO3•••O5' intra-chain hydrogen bonds [92]. However, chitosan chains can adopt several distinct conformations, including all the helical conformation observed in solid state. Helical preferences and conformational interchangeability were shown to be affected by the level of acetylation of the chitosan chains [92].

5. Structural dynamics of chitin and chitosan biopolymers

Structural characterization of chitin and chitosan conformations and their underlying interactions (intra- or inter-chain) has been largely determined by X-ray crystallography. The high flexibility of these oligosaccharides in solution has limited the acquisition of high-resolution structural data almost exclusively to X-ray diffraction of solid states (fiber, powder and tablet) (see section 3). Although NMR techniques are more suitable for structural characterization in solution, the flexibility of oligosaccharides makes NMR-
derived geometrical constraints scant and limits the application of NMR spectroscopy to the
determination of chitosan tridimensional structure [138]. Experimental data describing
dynamic processes such as solvation, particle formation and aggregation remain limited to a
macroscopic view, which is based on the measurement of chain stiffness and intrinsic
viscosity [139]. Transmission electron microscopy has also been used as a complementary
technique. Combining the latter with uranyl staining, electrostatic interactions involving
chitosan protonated amino groups were attributed a major role on chitin and chitosan
agglomeration in solution [140]. Therefore, the role of intra- and inter-chain hydrogen
bonds, ionic strength and temperature on the structural dynamics of chitosan cannot be
addressed exclusively by the means of experimental techniques [141]. Towards this end, MD
simulations can be used to obtain information on the time-evolution of carbohydrate
conformations at the atomic level and under varied environmental conditions that can be
complementary to experimental measurements [92-95].

Chitosan conformational diversity influences its solubility/physical state (soluble, gel,
aggregate), porosity, particle size and shape (fiber, nanoparticle, hollow fiber), ability to
chelate metal ions and organic compounds, biodegradability and consequently its biological
activity. The transition between these distinct conformational states is modulated by the
percentage and distribution of acetyl groups. The level of chitosan acetylation and the
distribution of N-acetyl groups along the chain have been shown to influence properties such
as solubility [142, 143], biodegradability [144] and apparent pK\textsubscript{a} values [145, 146]. Therefore,
the percentage and distribution of acetyl groups are key parameters for determining if
chitosan can effectively interact with biological systems [147]. The degree of acetylation can be
experimentally determined by infra-red spectroscopy [148, 149], enzymatic reaction [150],
ultra-violet spectroscopy [151], \textsuperscript{1}H liquid-state NMR [152], and solid-state \textsuperscript{13}C NMR [63, 153].
However, the interplay between chitosan acetylation and conformational transitions in
solution cannot be characterized at high-resolution by experimental techniques. In these cases,
atomistic MD simulation is a more suitable approach.

MD simulations in explicit solvent have been carried out for chitosan single chains and
nanoparticle aggregates with varied percentage and distribution of acetyl groups [92, 93].
Four degrees of acetylation were considered: 0% (fully deacetylated chains), 40% (60% of the
sites having a N-acyl group uniformly distributed), 40%-block (60% of the sites having a N-
acyl group in two spatially located well-defined regions of the particle), 60% (40% of N-acyl
uniformly distributed), 60%-block (40% of N-acyl groups spatially located in two well-
defined regions of the particle), and 100% (fully N-acetylated nanoparticle), i.e., a chitin
nanoparticle. Snapshots of molecular dynamics simulations after 40 ns for a chitin (100%)
and fully deacetylated chitosan nanoparticles (0%) are shown in Figure 3. Both simulations
started from aggregate crystal-like particles. It can be seen that chitin remain insoluble (in an
aggregate form, Figure 3A), while chitosan chains separate apart one from another until
each chain become fully hydrated (Figure 3B). Water molecules are not display in Figure 3
for clarity. These simulations have also shown a strong dependence of chitosan
conformation and solubility with pH and degree of acetylation. An increase in the level of
acetylation was shown to cause a progressive loss of flexibility and conformational
interchangeability (Figure 4). Thus, acetylation promotes a shift from more flexible structural motifs such as 5-fold and relaxed 2-fold towards a 2-fold conformation (Figure 4). It was also shown that the spatial location of the N-acetyl groups influences significantly chitosan conformational preferences, and therefore its solubility (Figure 4). Analyses of the MD trajectories have also shown that a high degree of acetylation and/or an increase in pH leads to a 3-fold increase of the lifetime of O3-HO3•••O5' intra-chain hydrogen bonds across the glycosidic linkages. The increase in the lifetime of this hydrogen bond was associated to a decrease in chitosan solubility. Chitosan with a high degree of acetylation favored the 2-fold conformation, but higher pH values did not affect significantly the secondary structure pattern of this oligosaccharide. In addition, we have also addressed the influence of spatial distribution of N-acetyl groups along the chitosan chain on swelling and the relative solubility of chitosan nanoparticles [93]. Simulations of chitosan with a uniform and block-wise distribution of N-acetyl groups along chains of a nanoparticle have shown that the latter displayed lower solubility [93]. The mechanism was attributed to the fact that 2-fold crystalline-like regions are created by the block distribution of acetyl groups, which is responsible to keep a more stable aggregate than its uniformly distributed counterpart.

Figure 3. Molecular structure representation of nanoparticles of chitin (A) and chitosan in circumneutral pH (B) at the end of 40-ns molecular dynamics simulations in explicit water. Water molecules and hydrogen atoms have been removed for clarity. Structures are represented in stick model and atoms color coded as: green: carbon; red: oxygen, blue: nitrogen.

Analysis of the cumulative average water content around each chain in the nanoparticles illustrates the relative solubility of each system (Figure 5a). On average, there is 0.26 water molecule per monosaccharide a 0.5 nm radial distance from each chitin chain. That corresponds to one water molecule for roughly every four monosaccharides. The number of water molecules increases to one water molecule per monosaccharide within the same radial distance for fully N-deacetylated chitosan. As expected, nanoparticle swelling is directly proportional to its solubility. The relative swelling can be expressed as the average radius of gyration of each chain in a particle as a function of the degree of acetylation (Figure 5b). Chitosan particles with a degree of acetylation ≥ 60% did not display any significant swelling in water. At this level of acetylation, only a small increase in the relative solvation content of chitosan with a uniform distribution (ca. 0.13 water molecules per monosaccharide) than its counterpart with a block distribution was observed. This small difference in solvation did not affect the overall solubility of the particles, supporting the empirical observation of a solubility
threshold around a level of 50% N-acetylation. Unexpectedly, water molecules within the insoluble chitosan particles were identified contributing for the maintenance of the regions in a 2-fold motif. The N-acetyl-glucosamine residues trapped water molecules between the chitosan chains, creating a hydrogen bond network between water molecules and the different chains without direct interaction between sheets. This finding substantiates a mechanism previously postulated by Ogawa and coworkers outlining the role of water molecules in chitin [123].

Figure 4. Secondary structure preferences for chitin and chitosan particles in water as a function of the degree of acetylation and spatial location of the N-acetyl groups. Results are the average of all chains in each nanoparticle, which are averaged over the last 5 ns of a 30-ns molecular dynamics simulation. The sum in each case does not reach 100% due to the increased flexibility and conformational diversity of these polymers in water. To be classified within a given secondary structure, as determined by solid X-ray crystallography, all dihedral angles in a chain should not present a deviation larger than 15% of the experimentally determined value. Data representation has been modified from [93].

Chitosan is a polyelectrolyte in acid medium. Its structure, physical state and conformational dynamics are greatly influenced by pH. The net charge of this cationic polyelectrolyte can be altered by its degree of acetylation [154]. Moreover, its apparent p\text{K}_a is directly related to the level of acetylation, varying from 6.1 to 7.32 units accordingly to proton concentration in the milieu [145, 155-158]. Based on these observations, it was proposed that aggregation occurs upon high levels of acetylation due to reduction of the biopolymer net charge [145], implying in a predictable behavior of chitosan chains depending on its electric charge distribution in aqueous solution [145, 146, 159]. It was also proposed that the low tendency of fully deacetylated chitosan to form aggregates is due to electrostatic repulsion among protonated amino groups. As result, chitosan electrostatic behavior was divided in three distinct patterns: i. DA < 20%, where it displays a polyelectrolyte behavior; ii. 20% < DA < 50%, where it is characterized by a counterbalance between hydrophilic and hydrophobic interactions;
and iii. DA > 50%, where associations of chitosan chains lead to the formation of stable aggregates. The results from atomistic molecular dynamics simulations in explicit water offer support to this hypothesis based on the accurate molecular description of the effect of the degree and distribution of N-acetyl groups on the swelling and aggregation stability of chitosan. Calculations of the hydrophobic and electrostatic contributions to the solvation free energy of the central chain in different particles as a function of acetylation are also consistent with the hypothesis (Figure 6) [129].

Figure 5. Relative solvation of chitin and chitosan nanoparticles as a function of the degree of acetylation and spatial location of the N-acetyl groups. A) Average coordination number of water molecules per residue as a function of distance (from 0.1 to 0.5 nm); B) Radius of gyration (and its variance, represented in bars) averaged per chain, over the last 5 ns of a 30-ns molecular dynamics simulation. Data representation has been modified from [93].

Figure 6. Average apolar and electrostatic contribution for the solvation free energy, per sugar residue, in chitin and chitosan nanoparticles as a function of the degree of acetylation and spatial location of the N-acetyl groups. Data representation has been modified from [92] and [93].
These contributions should be examined only as relative values as there are no experimental data for calibration or comparison of the calculated values. The apolar contribution remained nearly unaffected by the presence of water, while the electrostatic contribution is dominant even for insoluble chitin (100% acetylation). This finding suggests that hydrogen bond interactions, either intra-chains or between polymer chains and water molecules, play a more important role in the solubility of chitin and chitosan than hydrophobic interactions. These results have further shown that fine tuning the electrostatic contributions in chitosan can be used to promote remodeling of its physical state. Additional simulations have shown that the overall net charge and solubility of chitosan can be altered by changes in the pH. Comparison of the electrostatic response of a chitosan and chitin chains to pH changes shows a rather distinct surface charge profile for the two polymers. The electrostatic similarity between chitin and chitosan in basic pH aids to explain the loss of solubility of chitosan at high pH values (Figure 7). The positively charged character of chitosan chains in acid pH is shown by patches in blue (Figure 7D). On the other hand, chitin (Figure 7A) and chitosan chains in basic medium (Figure 7B) show a similar electrostatic potential at their molecular surfaces.

Figure 7. Electrostatic potential represented at the molecular surface of chitin (A) and chitosan (B-D) chains. Molecular structures were obtained from 20-ns molecular dynamics simulations. The different
chitosan chains were simulated at different pH values: B) basic (pH > 10); C) circumneutral (pH = 6.5); and, D) acid (pH < 3.5). Data representation has been modified from [92].

6. Final remarks and perspectives

Chitosan-based materials are involved in a plethora of medical, industrial and bioengineering applications such as bioremediation, radionuclide tissue decontamination and bone replacement to name a few. Due to the intrinsic flexibility and conformational variability of chitosan, the development of novel materials has been conducted mostly empirically. In this review, we have summarized the potential of using computer modeling to characterize in details the conformational behavior of chitin and chitosan. Understanding of the molecular properties of a given material allows for a more efficient/rational design. Therefore, this approach can be used to tailor these properties for specific needs. In this case, a systematic use of concerted experimental-theoretical information can provide a much clearer picture of the structural dynamics of polysaccharides and consequently can aid in such endeavor. This is still an emerging field that will benefit in the few years to come from the development of more accurate/extension of parameters for carbohydrate simulations, as well as novel models capable of better bridging the micro- and macroscopic scales.

Author details

Richard A. Cunha and Eduardo F. Franca*
Institute of Chemistry, Federal University of Uberlândia, Uberlândia, MG, Brazil

Thereza A. Soares, Victor H. Rusu, Frederico J.S. Pontes and Roberto D. Lins*
Department of Fundamental Chemistry, Federal University of Pernambuco, Recife, PE, Brazil

Acknowledgement

This work and the data herein contained was supported by the following sponsoring agencies: CNPq, FACEPE, FAPEMIG, INCT-INAMI (Brazilian National Science and Technology Institute for Integrated Markers), NanoBiotec-BR, CAPES, nBioNet and the Swedish Foundation for International Cooperation in Research and Higher Education (STINT). Part of the computational resources were provided by the Environmental Molecular Sciences Laboratory, an U.S. national scientific user facility sponsored by the U.S. Department of Energy located at the Pacific Northwest National Laboratory.

7. References


* Corresponding Authors


