Development of a coarse-grained #-chitin model on the basis of MARTINI forcefield

Citation: Yu, Zechuan, and Denvid Lau. "Development of a coarse-grained α-chitin model on the basis of MARTINI forcefield."Journal of Molecular Modeling (May 2015) 21:128, pp.1-9.

As Published: http://dx.doi.org/10.1007/s00894-015-2670-9

Publisher: Springer Berlin Heidelberg

Persistent URL: http://hdl.handle.net/1721.1/103324

Version: Author's final manuscript: final author's manuscript post peer review, without publisher's formatting or copy editing

Development of a coarse-grained α-chitin model on the basis of MARTINI forcefield

Zechuan Yu[a], *Denvid Lau*[a,b,*]

a Department of Architecture and Civil Engineering, City University of Hong Kong, Hong Kong, China.

b Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA.

* Corresponding author. E-mail: denvid@mit.edu. Tel: +852-3442-6829. Fax: +852-3442-0427.

Abstract

At nanoscale, atomistic simulation is widely used for investigating crystalline chitin fiber, the structural component for many biological materials. However, the longitudinal dimension of naturally occurring chitin fibers exceeds hundreds of nanometer, beyond the investigation range of all-atom simulation due to computing power limitation. Under this context, coarse-grained simulation is a useful alternative that facilitates the investigation of large system. We develop a coarse-grained model for describing the structural and mechanical properties of α-chitin. The developed coarse-grained model can reasonably predict structural and mechanical properties of α-chitin. Moreover, this model is consistent with existing coarse-grained force fields for proteins. The present model of α-chitin possesses good potential and applicability in the investigation of natural chitin-based materials at the length scale of hundreds of nanometers.

Keywords

Coarse-grained model, α-chitin, MARTINI forcefield, molecular dynamics

1. Introduction

Chitin is a linear 1-4 linked polysaccharide of N-acetyl-glucosamine. The N-acetyl and hydroxy methyl groups are very active in forming hydrogen bonds between adjacent chitin monomers [1,2]. The hydrogen bond network gives rise to the formation of stable crystalline chitin fibers. Native chitin adopts three crystalline polymorphs, namely α-, β-, and γ-chitin. Among these three crystalline structures, α-chitin is the most abundant and thermodynamically stable one [3]. Crystalline chitin fiber features a high stiffness to weight ratio [4]. Living organisms take advantage of the stiffness of chitin and produce a number of chitin-based biological materials such as arthropod cuticles, butterfly wings, squid beaks and spider fangs [5-8]. Natural chitin materials contain several typical constituents such as chitin fibers, proteins and minerals [9]. Despite limited types of components, living organisms can produce a variety of versatile biological composites serving for multiple functions. For instance, both the insect cuticle and spider fang are made of chitin, whereas the fang is used as a preying tool to break the protective cuticle of insects [8]. The diverse mechanical performances of chitin composites originate from structural difference. Understanding the relationship between structure and material properties in chitin composites could lead to a novel material design. Atomistic simulation is a useful tool for investigating the material structures at length scales ranging from angstroms to tens of nanometers. There are several all-atom force fields such as CHARMM and GROMOS being used to simulate chitin systems [10-12]. Using these force fields, atomistic simulations of chitin systems have been performed to address questions such as solvated conformation, hydrogen bond pattern, decrystallization and chitin-protein interaction [3,13-15]. However, the all-atom modeling is limited to small length and short time scales. The α-chitin fibers in the cuticle of lobster are 300 nm in length [16], beyond the range of all-atom molecular dynamics simulations.

The coarse-grained (CG) modeling technique is an alternative of all-atom (AA) modeling to perform large-scale simulations. In AA models, each atom is represented by one interactive site, whereas in CG models one interactive site could represent a group of atoms. The interactive site in CG models is called the "bead". The parameters defining the bead-bead interaction are derived in reference to AA simulations. With the concept of beads instead of single atoms, CG models neglect some degrees of freedom but the computational efficiency is significantly improved. Though, the loss of atomistic details may not affect the predictions of bulk properties of the simulated material. For example, the M3B CG approach maps one carbohydrate molecule consisting of around 20 atoms into only 3 beads and can still reproduce properties like the glass transition temperature [17]. The M3B CG simulations are around 7000 times more efficient than the original atomistic model [18]. There are many CG approaches available to grouping atoms into beads and deriving the parameters [19-21]. Among them, the MARTINI force field is a more systematic approach to coarse-graining biomolecules such as proteins, lipids and carbohydrates [22-24]. Within the framework of the MARTINI force field a set of chemical building blocks, which are similar to the concept of functional groups, are predefined [22]. The atomistic model of biomolecule can thereby be divided into different chemical building blocks, which are the beads in the CG model. By adopting the principle of chemical building blocks, the MARTINI force field provides an intrinsically consistent and extendable CG approach.

In the present study, we aim to develop a CG model of crystalline α-chitin that can be used to simulate chitin-protein interface. We base this model within the framework of MARTINI force field in hope of the consistent combination between our model (for chitin) and existing MARTINI models (especially for proteins). The CG model of α-chitin would be helpful for the investigation of ultra-long chitin fibers as well as the chitin-protein interactions at mesoscale.

The CG model should be able to predict the structural properties and the elastic modulus of α-chitin accurately. In addition, the CG model should be consistent with other MARTINI CG models, thereby it can be seamlessly applied within the MARTINI CG framework. We firstly map the chitin monomer into 3 beads. The initial set of the non-bonded and bonded parameters is based on the Boltzmann inversion [25] calculations of atomistic trajectories. We then perform CG simulations and calibrate the parameters against the structure and elastic modulus of α-chitin. Finally, we construct a CG chitin-protein model and measure the adhesion energy between the chitin substrate and a 10-glycine peptide. Results from our CG models are comparable to that of AA simulations, which is a strong evidence for the consistency of our CG α-chitin model within the framework of MARTINI force field.

2. Methods

2.1. All-atom simulation

All-atom simulation is performed using CHARMM36 [10] force field implemented in LAMMPS package [26]. The simulated α-chitin crystal is composed of four pairs of antiparallel 10-mer polysaccharide. The length and width of the chitin crystal are 19.08 Å and 52.55 Å respectively, exactly the same as the *x* and *z* dimensions of the simulation box. With periodic boundary conditions in all three directions, the periodic images of chitin crystal are aligned continuously along *x* and *z* axes. In the *y* direction, the simulation box (38.066 Å) is larger than the thickness of the chitin crystal (19.615 Å). The α-chitin crystal is placed on the center of the simulation box, while remaining space was filled with TIP3P water molecules [27]. Finally, the atomistic model containing one crystalline chitin fiber (2160 atoms) and 600 water molecules is constructed. After the energy minimization using conjugate gradient algorithm, the atomistic model is heated from 50 K to 300 K in 500 picosecond (ps, 10^{-12} s). It is then equilibrated in a NPT ensemble for 2 nanosecond (ns, 10^{-9} s) with a time step of 2 femtosecond (fs, 10^{-15} s). The pressure is

controlled at 1 bar by anisotropic Nose-Hoover coupling method with a damping time of 1 ps. The temperature is controlled at 300 K by Nose-Hoover thermostats with a damping time of 0.1 ps. The cut-off distances of both short-range van der Waals and Coulombic interactions are 1 nm. Particle-particle Particle Mesh (PPPM) method is applied to calculate to long-range forces. The high-frequency dynamics (energy terms involving hydrogen atoms) are constrained using SHAKE algorithm. After the initial 1-ns equilibration, the root-mean square deviation (RMSD) of atomistic positions is within the range of 2.3 to 2.65 during the latter 1-ns simulation, which indicates that the equilibrated state has been achieved. In the latter 1-ns simulation, coordinates and forces of atoms are recorded every 0.2 ps, so an all-atom trajectory with 5000 frames is obtained. In addition to the aforementioned solvated system, a dehydrated system (pure crystalline chitin fiber without water molecules) and a solvated chitin monosaccharide system (40 chitin monosaccharides dispersed in 1000 water molecules) are also simulated for deriving non-bonded parameters of the CG model. VOTCA package [28] is used to analyze the all-atom trajectory. The procedures of CG model development are described by a flow chart in Fig. 1. We firstly perform AA simulations and obtain the pseudo CG trajectory by mapping atoms into beads. Next, we use Boltzmann inversion [25] to reproduce the potential energy of both bonded (including bonds, angles and dihedrals, or torsions) and non-bonded energy terms. We then fit a harmonic function against the curve of potential energy vs. bond length (also the angle and torsion) and derive the parameters for bonded energy terms. For the non-bonded energy terms, zero-crossing distance and energy well depth are used to derive the parameters. Finally, we set up CG simulations using the derived parameters and test the performance of the newly developed CG force field for α-chitin.

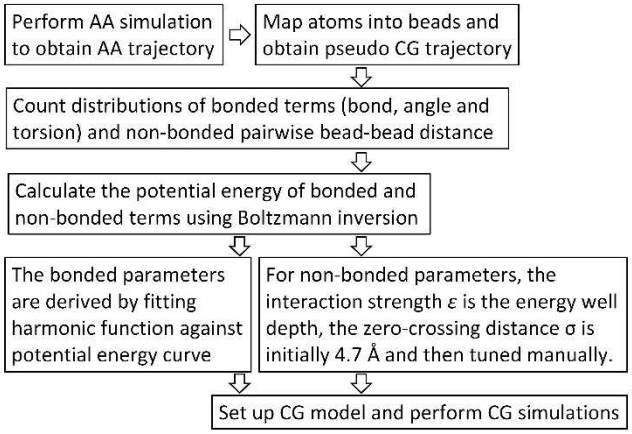


Fig. 1. The flow chart of the development of the CG model for α-chitin

2.2. Mapping strategy

In the chitin monomer molecule, two functional groups (the amide and hydroxymethyl groups as outlined in Fig. 2a stretch out of the pyran ring, making a slim, leaf-like structure. This structure can be sufficiently described by three beads as demonstrated by Fig. 1a. Three groups (including the attached hydrogen atoms), (CT, C, O), (C1, C2, C3, C4, O1, O3, O4, N) and (C6, O6, C5, O5) are mapped into three beads marked by T (tip), P (pyran ring) and B (bottom).

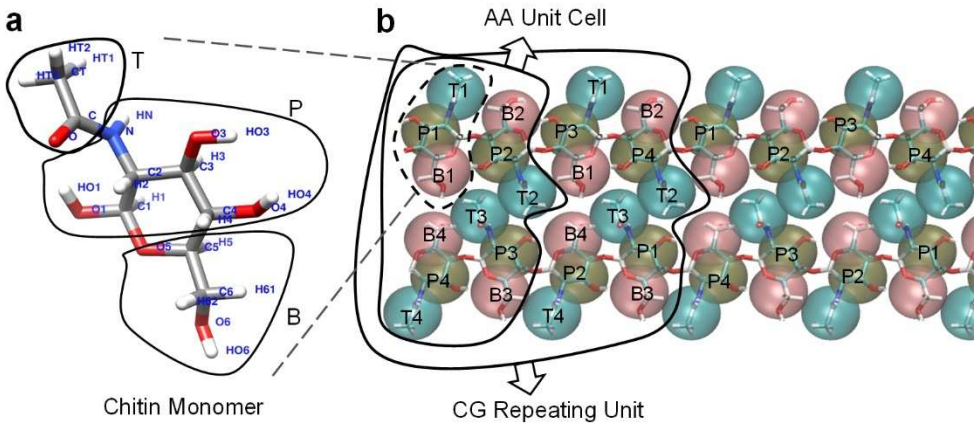


Fig. 2. a The mapping of one chitin monomer into three beads. **b** Two antiparallel chains of α-chitin. The unit cell of α-chitin is composed of four chitin monomers. The repeating unit in CG model is made of two unit cells due to the need to distinguish adjacent backbone beads (P beads)

In the AA model of α-chitin, four chitin molecules compose one α-chitin unit cell. In the CG model, the four molecules of one unit cell are specified as four types so bead names are suffixed

with numbers 1, 2, 3 and 4 as shown in Fig. 2b, the snapshot is captured by VMD [29]. Also, the P beads are suffixed with 1, 2, 3 and 4 in a sequence along the backbone direction. This treatment is essential because of the need to distinguish angle and dihedral terms. For example, in Fig. 2b, the angles B2-P2-P1 and B2-P2-P3 feature different equilibrium values, if P3 were still marked as P1, these two angles would be wrongly identified. Consequently, adjacent P beads are marked differently and the repeating unit of CG α-chitin model is composed of eight chitin molecules.

2.3. Force field parameterization

The bonded energy terms in the present CG force field can be written as follows.

$$U_{bonded} = k_b(b - b_0)^2 + k_\theta[\cos(\theta) - \cos(\theta_0)]^2 + k_\varphi(\varphi - \varphi_0)^2 \qquad (1)$$

Three harmonic terms in Eq. (1) describe the potential energy of bond, angle and torsion (or dihedral). Boltzmann inversion [25] is used to calculate the potential energy of bonded terms in reference to the all-atom trajectory. The curve of potential energy vs. bond length, angle or torsion can be fitted by a harmonic function. For example, the potential energy of the bond P1-P2 and the corresponding harmonic function are shown in Fig. 3a. The initially obtained parameters are later tuned according to the performance of the CG simulations. For example, the P-P bond and P-P-P angle parameters are found to be closely related to, and therefore tunable against, the chain-direction elastic modulus of chitin fiber. Similar calculations and fittings are performed to derive the parameters of all possible bonds, angles and dihedrals.

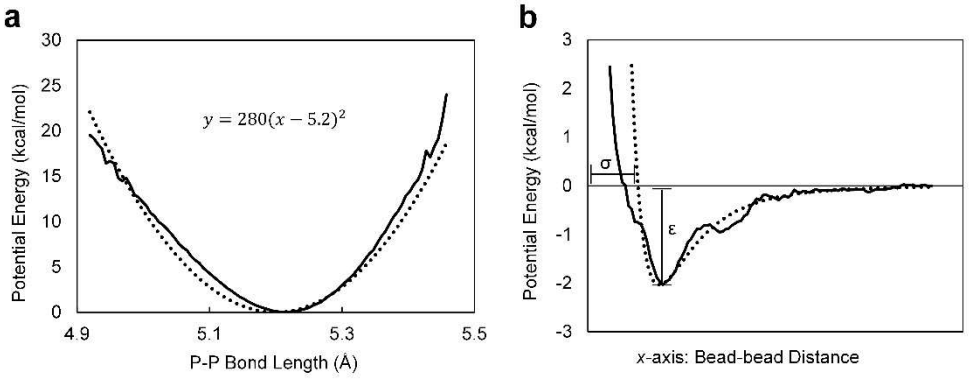


Fig. 3. a The plot of the potential energy against the P-P bond length. Solid and dotted curves are the Boltzmann inversion result and the fitted harmonic function respectively. The harmonic function writes as y=280(x-5.2)^2, so the parameters of bond P1-P2 can be initially determined as k_b=280 kcal/mol/Å^2 and b_0=5.2 Å. **b** The sample curve of the non-bonded potential energy against the bead-bead distance. Solid and dotted curves are Boltzmann inversion result and fitted Lennard-Jones function. Data is obtained from the simulation of chitin monosaccharide solution.

The non-bonded potential is calculated using Boltzmann inversion as well. A Lennard-Jones (LJ) 12-6 interaction potential describes the non-bonded bead-bead interactions.

$$U_{non-bonded} = 4\varepsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right] \quad (2)$$

The ε and σ in Eq. (2) are interaction strength and zero-crossing distance of the LJ potential respectively. The MARTINI force field provides a set of rules to determine the ε and σ for pairwise bead-bead interactions according to the functional groups the bead contains. The beads are classified into four fundamental groups, namely Q (charged), P (polar), N (nonpolar) and C (apolar), which reflect the interaction strength of the bead. Each type is further distinguished by subscripts showing different levels of polar affinity or the hydrogen-bonding capability. Accordingly, the beads B, P and T can be classified as N_a, N_{da} and P_1 respectively (the subscript "a" means the bead can act as hydrogen bond donor, the subscript "da" means the bead P can act as both hydrogen bond donor and acceptor). The non-bonded parameters for every pair of beads can be thereby defined. However, these settings are not useful in reproducing the crystalline structure of α-chitin. Alternatively, we start the CG model development in another way. The

zero-crossing distance, σ is set to 4.7 Å for all non-bonded pairs as the MARTINI force field assumes. The interaction strength, ε, is determined according to the energy well depth in the plot of potential energy against the distance, as shown in Fig. 2b. Noticing that the initial model cannot reproduce satisfactory CG simulation results, we modified several interactions and obtained reasonable CG model. The deriving procedures are stated below.

2.4. Modification of non-bonded parameters

The initial set of the non-bonded parameters resulted in abnormal α-chitin conformation as shown in Fig. 4, where the black balls are backbone beads (P bead representing the pyran ring) connected to form chitin chains. As assumed in MARTINI force field, the zero-crossing distance is initially set to 4.7 Å for all non-bonded interactions, so the equilibrium distance between each pair of beads is the same. In the Lennard-Jones potential formula, the energy minimum distance is proportional to the zero-crossing distance. Consequently, the inter-chain P beads tend to keep a same distance between each other, so the adjacent chains tend to assemble in a staggered manner as schematically drawn in the middle part of Fig. 4a. In actual α-chitin structure, these chitin chains should align in parallel with each other and form orderly arranged fibers as demonstrated in Fig. 4b. However, after simulation using the initial set of non-bonded parameters, these chains spontaneously altered the parallel alignment to form bent or staggered conformations as shown in Fig. 4c. These abnormal conformations are caused by the undistinguished pairwise interactions between P1, P2, P3 and P4. One solution to avoid these abnormal behaviors is to increase the equilibrium distance between dissimilar P beads such as the P1 and P2 in CG model. When the interactions between similar P beads (P1-P1 and P2-P2) and dissimilar P beads (P1-P2) are distinguished, the adjacent chains can align orderly as schematically shown in the right half of Fig. 4a. We gradually increase the zero-crossing

distance for the dissimilar P-P LJ interactions and eventually find an appropriate value, 6.5 Å. After this modification, the adjacent chains tend to assemble in an ordered manner, as shown in Fig. 4d. The CG model can thereby resemble realistic α-chitin structures. Such biased tuning focuses on the intramolecular chitin-chitin beads interactions, while the interaction between chitin and other molecules (for example proteins) remains unaffected. We can still follow the MARTINI philosophy to define the non-bonded interactions between chitin beads and protein beads. Therefore, the developed CG chitin model can be consistently applied within the MARTINI framework.

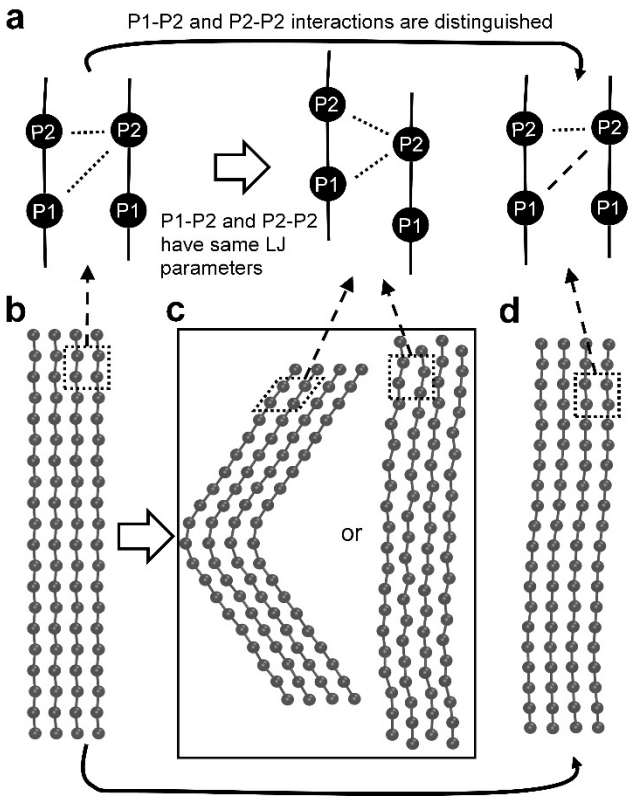


Fig. 4. a Scheme demonstrating the behavior of adjacent chains under different situations. **b** Top view of the initial arrange of four chitin chains. The black balls are the backbone P beads and other beads are eliminated. **c** The bent or staggered conformation with incorrect non-bonded parameters. **d** Modified non-bonded parameters lead to the ordered arrangement of parallel chitin chains

2.5. Coarse-grained simulations

Coarse-grained simulations are performed using LAMMPS package. The short CG chitin sample consists of four pairs of antiparallel 20-mer polysaccharide chains. The simulation box is periodic in all three directions with an initial size of x=19.1 Å, y=19.5 Å and z=106.0 Å. We have also constructed a 300-nm-long chitin fiber, which is composed of six pairs of 600-mer polysaccharide chains, the size of simulation box for long chitin fiber is x=28.7 Å, y=19.5 Å and z=3180.0 Å. Along the backbone direction the terminal chitin monomers are covalently connected with the periodic images so the system can be regarded as a part of an infinitely long chitin fiber. After minimized using conjugate gradient algorithm, the CG model is equilibrated in NPT ensemble with temperature and pressure maintained at 300 K and 1 bar. The thermostat and barostat include a Nose-Hoover thermostat with temperature damping parameter of 100 time steps and an anisotropic Nose-Hoover barostat coupled to x, y and z dimensions with pressure damping parameter of 1000 time steps. The cut-off distance of 12-6 LJ pair interaction is 12 Å. The CG model is simulated for 2 ns with time step of 10 fs. We note that a 20-fs time step can be used at similar coarse-graining level [30]. We have also performed simulation with 1-fs time step and found similar configurations. Depending on the purpose, one can use 1-fs time step for more accuracy or choose 10-fs time step for exploring longer time scale. Lattice parameters of chitin are measured as the average value over the 2-ns CG trajectory. After equilibration the chitin polymer underwent a deformation test. The simulation box elongates along the backbone direction (z axis) by 0.1% every 250 ps. At every elongation period, the pressure average over the latter 125 ps is calculated as the stress of the chitin. The entire deformation lasts for 5 ns and the final strain of the chitin is 2%. Stress-strain curve is obtained from this deformation process.

Aside from the simulation of pure chitin fiber, a CG model of chitin-protein interface is constructed and simulated as well. A short 10-glycine α-helix is placed over the <010> surface of the chitin substrate referring to a previous study on chitin-protein interface [14]. The CG protein model is built on the basis of MARTINI force field directly. The N-terminus, backbone amino acids and C-terminus are classified as N_a, N_{da} and P_3 respectively. The helix conformation can be maintained by selectively control the equilibrium value of harmonic angle (96 degree) and torsion (60 degree) terms. The non-bonded interactions between chitin and protein are also determined following MARTINI force field principles. Using the same settings as that of pure chitin simulation, the system is equilibrated in NPT ensemble for 2 ns as the RMSD of the system reaches to a stable value. After equilibration, steered molecular dynamics (SMD) simulation is performed to measure the adhesion energy between chitin substrate and the protein filament. The SMD program simulates the detachment of peptide filament from the substrate by pulling a virtual spring tethered to one end of the peptide filament. The pulling speed and the spring constant were 10 Å/ns and 100 kcal/mol/Å^2, same as those in the previous studies [14,15]. The energy input to the system during the detachment is recorded. Using Jarzynski equality [31], the adhesion strength (the free energy difference between attached and detached states) can be calculated as a result of statistical analysis over 10 SMD trajectories.

3. Results and Discussion

3.1. Elastic modulus

After the preliminary calibration of the non-bonded parameters, the CG model can correctly reflect the structure of the α-chitin fiber. Now we want the CG model to be able to reproduce reasonable mechanical properties of α-chitin. Prediction of the elastic properties of material is an imperative utility of molecular dynamics simulation, a valid and useful CG model must reproduce reasonable elastic modulus. Here we use the chain-direction elastic modulus to

validate the CG model of α-chitin. When the chitin fiber is subjected to elongation along chain direction, the P-P bonds are stretched and the P-P-P angles are enlarged as schematically shown in Fig. 5a. This phenomenon suggests that the chain-direction elastic modulus of CG chitin model depends on the force constants of the P-P bonds and P-P-P angles. A series of CG simulations of 10-nm chitin fiber with varied force constants confirm the suggestion and show that the chain-direction elastic modulus of the CG α-chitin model are amendable to these two force constants. In order to obtain a valid CG model, we calibrate the force constants of P-P bonds and P-P-P angles against the correct elastic modulus. Initially the force constants are set to 280 kcal/(mol·Å^2) for P-P bonds and 400 kcal/mol for P-P-P angles respectively. These initial settings lead to a chain-direction elastic modulus of 200 GPa, around twice as much as the results from other numerical studies (ranging from 88 GPa to 120 GPa) [4,14]. We reduce the force constants gradually to 75% and 50% of the original values and find that the chain-direction elastic modulus is nearly proportional to the force constants. Selected sets of parameters, the stress-strain relationship as well as the elastic moduli are shown in Fig. 5b. The chain-direction elastic modulus is tuned to a reasonable value, 92.2 GPa, which corresponds to other simulation results. Also, the tensile backbone elastic modulus of 300-nm long chitin fiber is 92.1 GPa, the same as the 20-nm-long one.

It should be noted that the linkage between properties and parameters of CG model is elusive. When adjusting the force constants of the backbone bonds and angles, we find that variations of these force constants can also influence the structural properties (such as the unit cell size) of the CG α-chitin model. This phenomenon indicates the complex relationship between properties and parameters of the CG model, *i.e.*, single parameter can determine multiple properties. Meanwhile, the elastic modulus is not exactly proportional to the backbone force constants,

which suggests that the elastic modulus could be influenced by other bonded or non-bonded parameters as well. That is to say, single property may depend on multiple parameters. The elusive relationship between properties and parameters would preclude the fine tuning of the CG model. Unlike the chain-direction elastic modulus which is a conveniently amendable property, some properties are hard to calibrate. For example, we find that the unit cell size is hardly amendable due to the multi-parameter dependency and the non-linear relationship with the parameters of CG model. Though, our CG model can still reproduce reasonable unit cell size, which will be discussed in the next section.

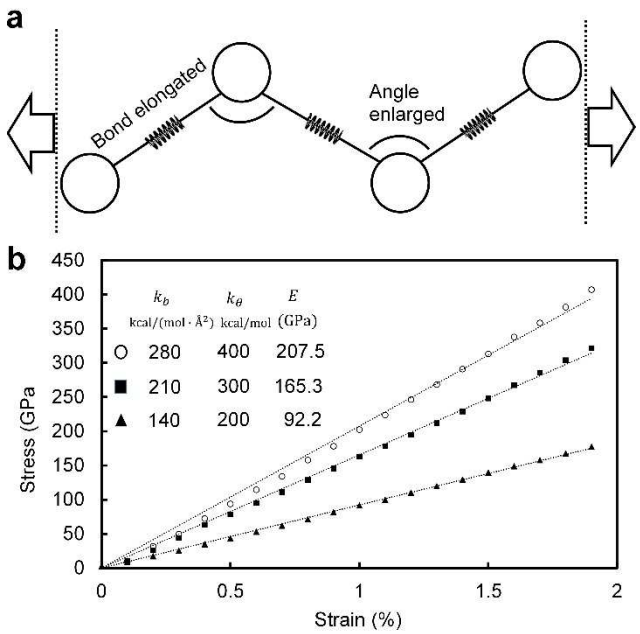


Fig. 5. **a** Scheme of backbone beads subjected to extension. **b** The stress-strain relationship obtained from chain-direction tensile tests with varied bond and angle constants

3.2. Unit cell size

Revealed from *ab initio* and experimental studies, unit cell is the fundamental construction block of the crystalline fiber. Normally the model of α-chitin can be constructed by replicating the unit cell along the *x*, *y* and *z* directions because the unit cell of α-chitin is almost orthorhombic. We

use the standard unit cell dimensions to build CG α-chitin model, while during the equilibrium process the unit cell size is adapted according to the interaction parameters between beads. The equilibrated unit cell size of the CG model can reflect the resemblance of CG model to AA model. Here, we use the dimensions of the unit cell as the indicator of the accuracy of the CG model for α-chitin. The unit cell parameters of AA and CG models together with two references are listed in Table 1. Results from two fiber samples (10-nm and 300-nm fibers) are close. All the parameters especially *b* and *c* from our CG model are in good agreement with that of references, whereas the parameter *a* still has room for improvement. When calibrating the elastic modulus of the CG model, we have noticed the association between unit cell size and the backbone terms, P-P bond and P-P-P angle. Increasing the P-P bond length leads to the increase of chain-direction size. Altering the force constants can vary the lateral dimensions. The variations of unit cell size are non-linear to the value of the parameters. Meanwhile, the non-bonded parameters also affect the unit cell size. Attempts have been made to tune CG model parameters for getting an accurate unit cell size but no one has succeeded in reproducing all three unit cell dimensions exactly. Too much calibration against one parameters is always accompanied with the loss of general accuracy, *i.e.*, the values of *b* or *c* would drift away from correct values while efforts are made to improve the accuracy in predicting *a*. We realized that the linkage between unit cell size and CG parameters is very complicated and the fine tuning is hard to achieve, same as the situation in the development of CG model for native cellulose [30]. Fortunately, the initial set of the CG parameters can give reasonable results in terms of the unit cell size and we just need to maintain the good performance when modifying the model against other properties such as the elastic modulus. Eventually, the tuned CG model can predict both the elastic modulus and the unit cell parameters with satisfactory accuracy.

	a (Å)	*b* (Å)	*c* (Å)
Ab initio study [2]	4.98	19.32	10.45
Experiment [1]	4.74 ± 0.01	18.86 ± 0.02	10.32 ± 0.02
Molecular dynamics	4.69 ± 0.16	19.02 ± 0.19	10.32 ± 0.09
Coarse-grained (10 nm)	5.26 ± 0.05	18.86 ± 0.05	10.07 ± 0.07
Coarse-grained (300 nm)	5.27 ± 0.03	18.85 ± 0.07	10.08 ± 0.05

Table 1. The unit cell parameters from references and our coarse-grained model

Robustness of a CG model indicates the ability of the model to predict multiple material properties and the applicability of the model to diverse situations. Our CG α-chitin model shows its robustness in predicting both the chain-direction elastic modulus and the unit cell parameters. Furthermore, this model can also be embedded into the framework of MARTINI force field to investigate the interface between chitin and protein.

3.3. Adhesion of α-helix at chitin substrate

The above two discussion sections have proven the accuracy of the CG model in reproducing the crystalline structure and the elastic modulus of α-chitin. In this section, the consistency of the CG model with the MARTINI force field is tested. We construct a chitin substrate using our CG model and a peptide filament using the MARTINI coarse-graining principles for protein. When initially aligned along the *z* axis over the (010) surface of α-chitin substrate, the short α-helix peptide would gradually alter its orientation and lie in the groove, as demonstrated by Fig. 6a and b. At the chitin surface, there exist “hills” and “valleys” formed by amide and hydroxymethyl groups, as pointed out in Fig. 6a. This configuration provides a specific place (the groove in between the “hills” as shown by the dashed line in Fig. 6b favored by the short α-helix peptide. The specific binding of the peptide filament to the groove of the chitin surface corresponds to the observation from AA simulations [14]. After attachment, the peptide is detached by an upright pulling force and the adhesion strength can be calculated using Jarzynski equality [31]. The pulling process is illustrated in Fig. 6c, where the peptide, pulled by the upright force on one

terminal, detaches from the chitin substrate like a peeled-off bandage. The work done by the pulling force and the product of Jarzynski equality are plotted as gray and black curves respectively as shown in Fig. 6d. The free energy difference between attached state and detached state is 43.1 kcal/mol, in good agreement with AA simulation result (46.6 kcal/mol) [14]. These similar results indicate the accuracy of the CG model in predicting interactive behavior of chitin and protein, as well as the consistency of the CG model within the framework of MARTINI force field.

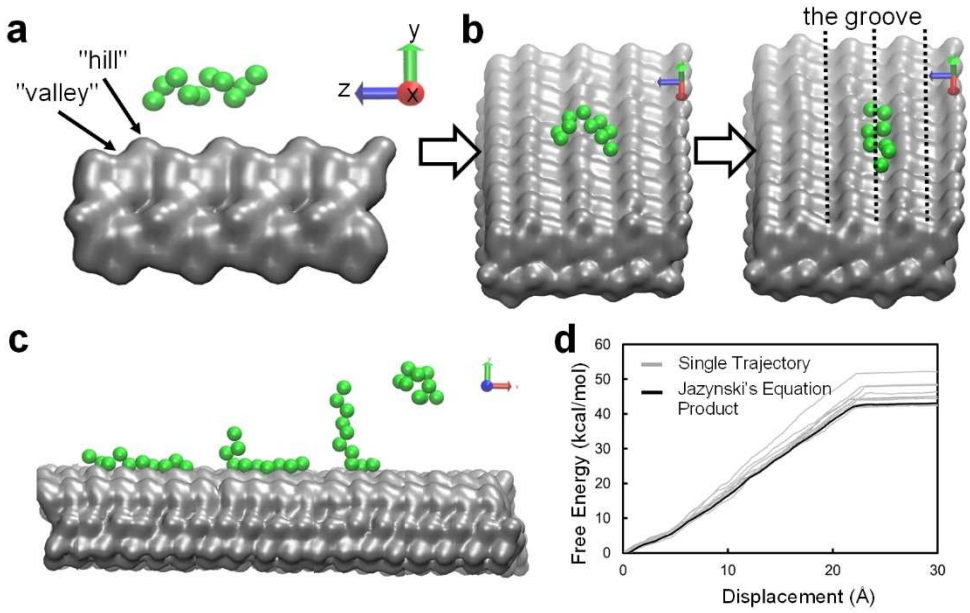


Fig. 6. a The initial alignment of the α-helix over the chitin substrate. The "hill" and "valley" are pointed out by black arrows. **b** The peptide filament gradually alters its orientation in parallel to the groove during the attachment. It is noted that the groove (pointed out by dotted lines) is a favored place for the peptide filament to stay. **c** The pulling process of the peptide away from the chitin substrate. **d** The free energy profile during the pulling. The adhesion strength is 43.1 kcal/mol

3.4 Potential applications and future studies

The case study on chitin-protein interface demonstrates that the developed CG model reasonably resembles a crystalline chitin scaffold and can be used in the simulations of chitin-protein system. Proteins, especially chitin-binding proteins, may behave specifically when interacting with chitin

scaffold. Some proteins are composed of more than 1000 amino acids, causing heavy computing burden to full atomistic simulations. Under this context, the CG approach serves as an efficient tool to estimating the adhesion strength, binding configuration, diffusion coefficient, etc. Furthermore, the chitin model can be used in investigating natural chitin-based materials, such as chitin-protein fibers [4,5] and chitin-silica composites [32,33]. Using coarse-graining model, it is envisioned that we can construct more complicated structures and explore the relationship between material properties and structure at the length scale of several hundred nanometers.

4. Conclusion

We have developed a CG model of α-chitin and manually calibrated several parameters to reproduce realistic crystalline structures as well as reasonable elastic modulus. The developing procedures constitute as useful references for further coarse-graining other crystalline biopolymers including β, γ-chitin and cellulose, so as to facilitate mesoscale simulations. The developed CG model resembles a crystalline α-chitin scaffold, which can be used in combination with existing MARTINI models so as to simulate interactions between α-chitin and proteins. In natural chitin composite system, there exist structures made of 18-25 α-chitin chains with a length of 300 nm, out of the investigation range of AA simulations. The present CG modeling technique makes such long α-chitin fiber touchable in molecular dynamics simulations. There are also limitations for the present CG model. Firstly, the CG model is developed in reference to atomistic trajectories under room temperature and pressure (T=300 K and p=1 atm), so it may not be not universally applicable to other environmental conditions. Meanwhile, the CG model is modified on the purpose of reproducing crystalline structures, therefore the accuracy could diminish when the crystalline form is broken. Despite the limitations, our CG model is conveniently amendable towards the calibration data, so it could be flexible in various

circumstances after specific calibration. It is envisioned that following studies with the CG modeling technique could enrich our understanding of the chitin-protein system at the submicron scale.

Supplementary Material

A file containing the coarse-grained parameters of all the bonds, angles dihedrals as well as the non-bonded interactions is available online along with the electronic version. The settings of coarse-grained protein models are referred from online coarse-graining tutorial (http://www.ks.uiuc.edu/Training/Tutorials/martini/rbcg-tutorial.pdf).

Acknowledgement

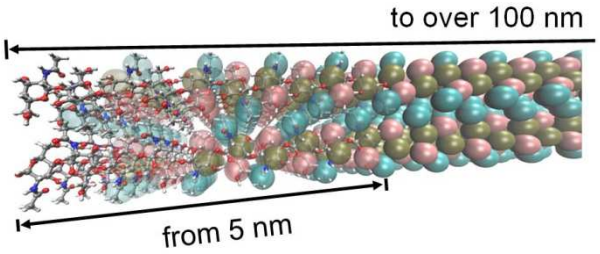
The authors are grateful to the support from Croucher Foundation through the Start-up Allowance for Croucher Scholars with the Grant No. 9500012, and the support from the Research Grants Council (RGC) in Hong Kong through the Early Career Scheme (ECS) with the Grant No. 139113.

References

1. Sikorski P, Hori R, Wada M (2009) Revisit of α-chitin crystal structure using high resolution X-ray diffraction data. Biomacromolecules 10 (5):1100--1105
2. Petrov M, Lymperakis L, Friák M, Neugebauer J (2013) Ab Initio Based conformational study of the crystalline α-chitin. Biopolymers 99 (1):22--34
3. Franca EF, Lins RD, Freitas LC, Straatsma T (2008) Characterization of chitin and chitosan molecular structure in aqueous solution. Journal of Chemical Theory and Computation 4 (12):2141-2149
4. Nikolov S, Petrov M, Lymperakis L, Friák M, Sachs C, Fabritius H-O, Raabe D, Neugebauer J (2010) Revealing the Design Principles of High-Performance Biological Composites Using Ab initio and Multiscale Simulations: The Example of Lobster Cuticle. Advanced Materials 22 (4):519--526
5. Vincent JF, Wegst UG (2004) Design and mechanical properties of insect cuticle. Arthropod Structure & Development 33 (3):187--199
6. Saranathan V, Osuji CO, Mochrie SG, Noh H, Narayanan S, Sandy A, Dufresne ER, Prum RO (2010) Structure, function, and self-assembly of single network gyroid (I4132) photonic crystals in butterfly wing scales. Proceedings of the National Academy of Sciences 107 (26):11676-11681

7. Miserez A, Li Y, Waite JH, Zok F (2007) Jumbo squid beaks: Inspiration for design of robust organic composites. Acta Biomaterialia 3 (1):139-149
8. Politi Y, Priewasser M, Pippel E, Zaslansky P, Hartmann J, Siegel S, Li C, Barth FG, Fratzl P (2012) A Spider's Fang: How to Design an Injection Needle Using Chitin-Based Composite Material. Advanced Functional Materials 22 (12):2519--2528
9. Chen P-Y, Lin AY-M, McKittrick J, Meyers MA (2008) Structure and mechanical properties of crab exoskeletons. Acta biomaterialia 4 (3):587-596. doi:http://dx.doi.org/10.1016/j.actbio.2007.12.010
10. Guvench O, Mallajosyula SS, Raman EP, Hatcher E, Vanommeslaeghe K, Foster TJ, Jamison FW, MacKerell Jr AD (2011) CHARMM Additive All-Atom Force Field for Carbohydrate Derivatives and Its Utility in Polysaccharide and Carbohydrate--Protein Modeling. Journal of chemical theory and computation 7 (10):3162--3180
11. Huang J, MacKerell AD (2013) CHARMM36 all-atom additive protein force field: Validation based on comparison to NMR data. Journal of computational chemistry 34 (25):2135--2145
12. Oostenbrink C, Villa A, Mark AE, Van Gunsteren WF (2004) A biomolecular force field based on the free enthalpy of hydration and solvation: The GROMOS force-field parameter sets 53A5 and 53A6. Journal of computational chemistry 25 (13):1656-1676
13. Beckham GT, Crowley MF (2011) Examination of the α-chitin structure and decrystallization thermodynamics at the nanoscale. The Journal of Physical Chemistry B 115 (15):4516--4522
14. Jin K, Feng X, Xu Z (2013) Mechanical Properties of Chitin–Protein Interfaces: A Molecular Dynamics Study. BioNanoScience 3 (3):312-320. doi:10.1007/s12668-013-0097-2
15. Yu Z, Xu Z, Lau D (2014) Effect of Acidity on Chitin–Protein Interface: A Molecular Dynamics Study. BioNanoScience 4 (3):207-215. doi:10.1007/s12668-014-0138-5
16. Muzzarelli RA (2011) Chitin Nanostructures in Living Organisms. In: Gupta NS (ed) Chitin, vol 34. Topics in Geobiology. Springer Netherlands, pp 1-34. doi:10.1007/978-90-481-9684-5_1
17. Molinero V, Goddard WA (2004) M3B: A coarse grain force field for molecular simulations of malto-oligosaccharides and their water mixtures. The Journal of Physical Chemistry B 108 (4):1414-1427
18. Liu P, Izvekov S, Voth GA (2007) Multiscale coarse-graining of monosaccharides. The Journal of Physical Chemistry B 111 (39):11566-11575
19. Fukunaga H, Takimoto J-i, Doi M (2002) A coarse-graining procedure for flexible polymer chains with bonded and nonbonded interactions. The Journal of chemical physics 116 (18):8183-8190
20. Müller-Plathe F (2002) Coarse-Graining in Polymer Simulation: From the Atomistic to the Mesoscopic Scale and Back. ChemPhysChem 3 (9):754-769
21. Tschöp W, Kremer K, Batoulis J, Bürger T, Hahn O (1998) Simulation of polymer melts. I. Coarse-graining procedure for polycarbonates. Acta Polymerica 49 (2-3):61-74
22. Marrink SJ, Risselada HJ, Yefimov S, Tieleman DP, de Vries AH (2007) The MARTINI force field: coarse grained model for biomolecular simulations. The Journal of Physical Chemistry B 111 (27):7812-7824

23. Monticelli L, Kandasamy SK, Periole X, Larson RG, Tieleman DP, Marrink S-J (2008) The MARTINI coarse-grained force field: extension to proteins. Journal of Chemical Theory and Computation 4 (5):819-834
24. López CA, Rzepiela AJ, De Vries AH, Dijkhuizen L, Hünenberger PH, Marrink SJ (2009) Martini coarse-grained force field: extension to carbohydrates. Journal of Chemical Theory and Computation 5 (12):3195-3210
25. Reith D, Pütz M, Müller-Plathe F (2003) Deriving effective mesoscale potentials from atomistic simulations. Journal of computational chemistry 24 (13):1624-1636
26. Plimpton S (1995) Fast parallel algorithms for short-range molecular dynamics. Journal of Computational Physics 117 (1):1-19
27. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML (1983) Comparison of simple potential functions for simulating liquid water. The Journal of Chemical Physics 79 (2):926-935. doi:doi:http://dx.doi.org/10.1063/1.445869
28. Rühle V, Junghans C, Lukyanov A, Kremer K, Andrienko D (2009) Versatile object-oriented toolkit for coarse-graining applications. Journal of Chemical Theory and Computation 5 (12):3211-3223
29. Humphrey W, Dalke A, Schulten K (1996) VMD: visual molecular dynamics. Journal of molecular graphics 14 (1):33--38
30. Wohlert J, Berglund LA (2011) A coarse-grained model for molecular dynamics simulations of native cellulose. Journal of Chemical Theory and Computation 7 (3):753-760
31. Park S, Khalili-Araghi F, Tajkhorshid E, Schulten K (2003) Free energy calculation from steered molecular dynamics simulations using Jarzynski's equality. The Journal of chemical physics 119:3559
32. Ehrlich H, Worch H (2007) Sponges as natural composites: from biomimetic potential to development of new biomaterials. Museu National, Rio de Janeiro, Brasil:303-312
33. Ehrlich H, Simon P, Carrillo-Cabrera W, Bazhenov VV, Botting JP, Ilan M, Ereskovsky AV, Muricy G, Worch H, Mensch A (2010) Insights into chemistry of biological materials: newly discovered silica-aragonite-chitin biocomposites in demosponges. Chemistry of Materials 22 (4):1462-1471


to over 100 nm
from 5 nm