

# Chitin Builder: a VMD tool for the generation of structures of chitin molecular crystals for atomistic simulations

David C. Malaspina\*, Jordi Faraudo\*

ICMAB-CSIC, Campus UAB Bellaterra, Barcelona, Spain.

---

## PROGRAM SUMMARY

*Program Title:* Chitin Builder

*Licensing provisions:* GPLv3

*Programming language:* Tcl/Tk

*Nature of problem:* Build-up of structures of chitin organic crystals to be used as input for molecular dynamics simulations.

*Solution method:* Chitin builder VMD plugin generates crystal coordinates and structural/topology files of  $\alpha$  or  $\beta$  chitin polymorphs of desired sizes for use in MD simulations.

*Additional comments including Restrictions and Unusual features:* The Chitin Crystal Builder tool is not a standalone program but a plugin that enhances VMD's modelling features (version 1.9.3 or higher)

## 1. Program description, use and features

### 1.1. Brief description of the program

The code is a plugin for VMD written in tcl/tk v8.4 programming language. The source code contains two main parts: the code that obtains the atomic coordinates of the atoms of the crystal and the structure and topology of the crystal (see Methods section) and a code responsible for the graphical user interface (gui). The program generates two main outputs: a coordinate file (PDB) containing the position of all the atoms in the generated structure and a topology file (PSF) containing all the bonds, angles and dihedrals according to CHARMM36 carbohydrate section [2]. These output files are named crystal-alpha-psf.pdb / psf or crystal-beta-psf.pdb / psf depending on the allomorph, and are stored in the working folder chosen by the user. This two files, plus the included CHARMM36 parameters file (located in the /ForceField/ folder) allows the user to easily start a molecular dynamics simulation using the NAMD simulation program [3] that accompanies VMD. Also, using VMD, the users can easily export the data from these two files (PDB and PSF) to other coordinates formats or convert the topology to the formats required by other programs such as GROMACS [4, 5] using the topotools plugin included in VMD.

---

\*E-mail addresses: dmalaspina@icmab.es (D.C. Malaspina), jfaraudo@icmab.es (J. Faraudo)

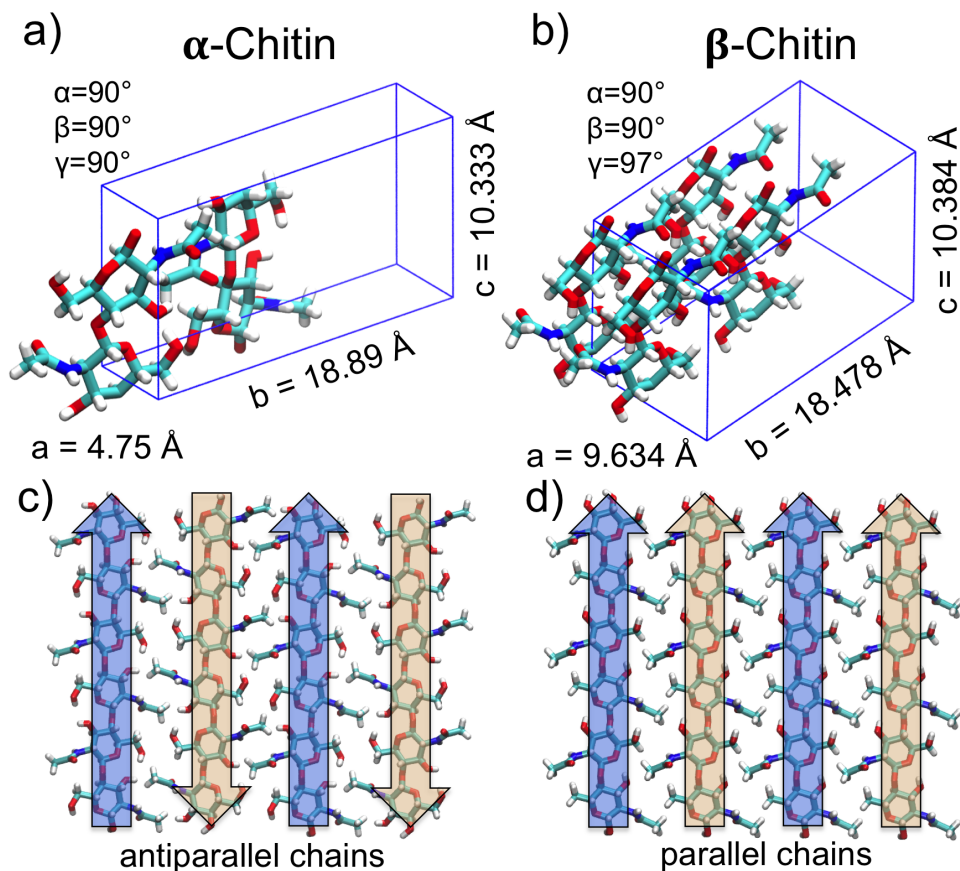


Figure 1: Structure of chitin crystal allomorphs  $\alpha$  and  $\beta$ . a)  $\alpha$  chitin primitive unit cell. b)  $\beta$  chitin expanded unit cell (corresponding to 2 primitive unit cells). c) scheme of the antiparallel organization of the chains in  $\alpha$  chitin, indicated by arrows over a crystal generated with the input for the number of replicas  $a=1$   $b=2$   $c=3$ . d) scheme of the parallel chain organization in  $\beta$  chitin, indicated by arrows over a crystal generated with the input for the number of replicas  $a=1$   $b=2$   $c=3$  (note that an extra plane of molecules is below the image in this case). Chitin molecules are shown in bond representation in all cases (O in red, H in white, N in blue and H in white). Figure made with VMD. [1].

## 1.2. Installation

As we mentioned before the code is a plugin for VMD. In the Readme.md file we detailed the installation process. Basically, the plugin can be easily installed by simply moving the downloaded plugin folder inside the standard plugin folder of VMD (the folder `/plugins/noarch/tcl/` inside the folder where VMD is installed). Usually administrator privileges are needed to copy into the VMD installation folder. Also note that the specific location at which VMD is installed vary depending on the operating system and the user VMD installation preferences, as explained in VMD documentation. We also included a bash script that performs an automatic installation for Ubuntu Linux system.

### 1.3. Graphical user interface

All the options of our plugin are controlled from a graphical user interface (GUI), shown in figure 2. This GUI allows the user to switch between the different options for the chitin crystal generation (Figure 2). It also provides a quick view to the unit cell data as well as the cell vectors describing the generated crystal, to be employed in a MD simulation.

In Figure 2, we also indicate the different sections of the GUI which are the following:

Section 1) of the gui correspond to the "help" menu. It contains two items: "About", that contains information of the developers and "Documentation", that redirects the user to the Bitbucket repository code webpage where a Readme.md file explains the main features of the code. Also, the last version of the code is available at this repository.

Section 2) of the gui corresponds to the selection of the crystal allomorph. It contains a button that allows the user to choose between  $\alpha$  and  $\beta$  chitin allomorphs as explained in the Method section. Under this button a description of the unit cell data including  $a$ ,  $b$  and  $c$  cell vectors and  $\alpha$ ,  $\beta$  and  $\gamma$  angles is shown (corresponding to the unit cells in Figures 1 a and c) as well as the reference to the primary publication where the experimental data (from which this unit cell was build) can be found.

Section 3) of the gui allows the user to input the number of replicas in  $a$ ,  $b$  and  $c$  crystallographic directions of the new crystal that is going to be generated. The inputs need to be an integer and the maximum number that the calculation allow is such that the number of replicas do not exceed  $9.999 \times 10^3$  residues (the maximum allowed by VMD). Under the input for the number of replicas, a button allows the user to select between a structure with ("yes") or without ("no") periodic bonds in the longitudinal direction of the polysaccharide chain. This type of bonds allow the simulation of infinitely long molecular crystals by using periodic boundary conditions. At the bottom of this section the button "Generate Chitin Structure" will prompt dialog box where the user need to choose the destination folder and after this the code will initialize for the replication and topology generation. Once the calculation is done the generated structure will appear in the VMD display and the output files crystal-alpha-psf.psf/pdb or crystal-beta-psf.psf/pdb will be generated on the destination folder selected by the user. Data of the crystal cell will be printed on the gui, on the main console and in the log file "crystal.log".

Finally, section 4) of the gui displays the cell data of the generated chitin structure. The cell data is displayed as a matrix of the cell vector data(the employed units are Å). The first row correspond to the 3 cartesian components (i, j, k) of the  $a$  vector of the cell, second row correspond to the 3 components of the  $b$  vector of the cell and third row to the 3 components of the  $c$  vector of the cell. This data is provided in this format so the user can use it directly in a configuration file for a MD simulation (as in a simulation with NAMD for example). This useful data is also printed on the main console and on the log file.

## 2. Examples

Using the chitin builder plugin we have generated several structures as is displayed in Figure 3. All the generated structures (PDB + PSF files) are included in the /examples/ folder as well as a screen capture of the gui that generated the respective structure.

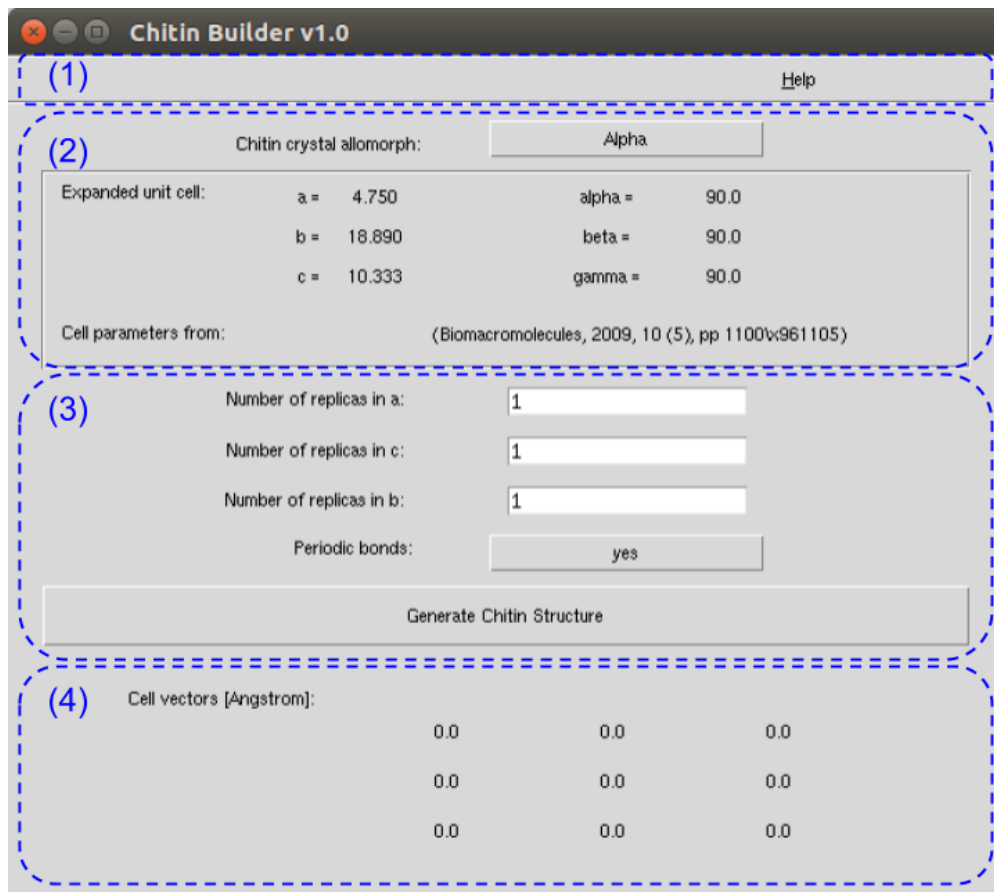


Figure 2: Chitin builder graphical user interface sections

In the first example (Figure 3a), we have created an  $\alpha$  chitin crystal with the input parameters for the replicas  $a=4$ ,  $b=2$  and  $c=4$  and the option of periodic bonds was turned to "no". In Figure 3a we show the resulting crystal produced by the program. Using this input, the program produces a crystal cell of 19.0 Å in the  $a$  direction, 37.78 Å in the  $b$  direction and 41.33 Å in the  $c$  direction. This structure is provided to the user in the folder /examples/example-a-alpha-4-2-4/.

In the second example (Figure 3b) we created a  $\beta$  chitin crystal with the input parameters for the replicas  $a=2$ ,  $b=2$  and  $c=4$ . The option periodic bonds was also turned to "no". The crystal generated with these options is shown in Figure 3b. It has a length of 19.276 Å, a height of 36.956 Å, and depth of 41.536 Å. Note that this structure is equivalent to the  $\alpha$  in Figure 3a), but the number of replicas needed to generate it is different due to the differences in the unit cells. This structure is provided as an example to the user in the folder /examples/example-b-beta-2-2-4/.

In the third example (Figure 3c) we repeated the creation of a  $\beta$  chitin crystal with the input parameters for the replicas  $a=2$ ,  $b=2$  and  $c=4$ , but this time the periodic bond option was set to "yes". In Figure 3c) we can observe the effect of setting the periodic bond option to

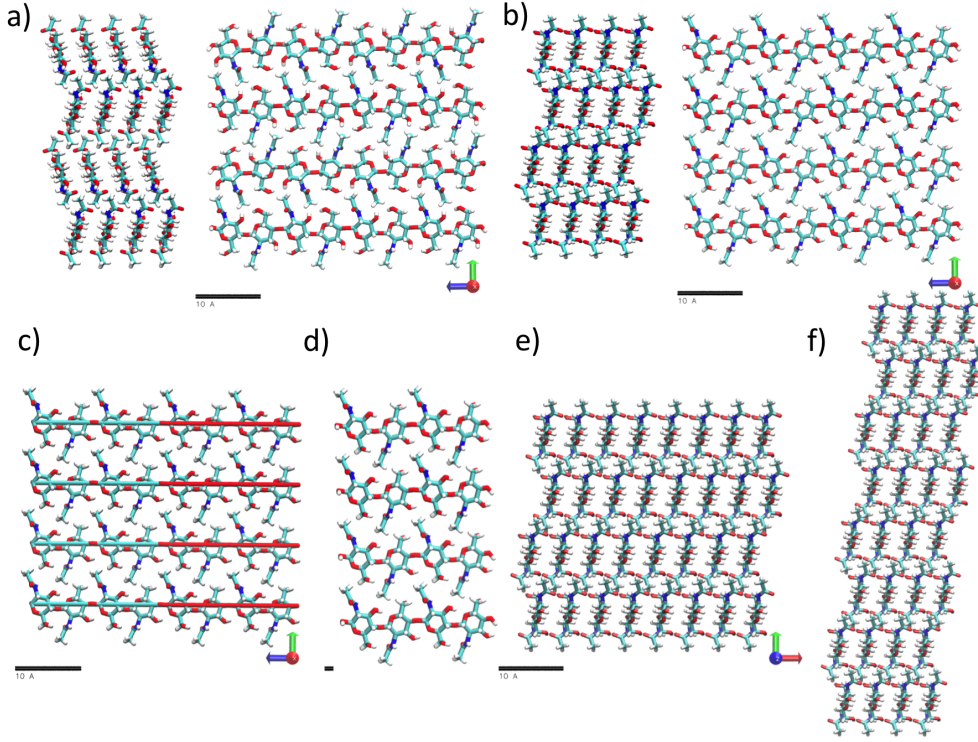


Figure 3: Examples of crystal structures generated by chitin builder. a)  $\alpha$  chitin crystal (input parameters  $a=4$ ,  $b=2$ ,  $c=4$ ), x-y plane on the left and y-z plane on the right. b)  $\beta$  chitin crystal (input parameters  $a=2$ ,  $b=2$ ,  $c=4$ ), x-y plane on the left and y-z plane on the right. c)  $\beta$  chitin crystal (input parameters  $a=2$ ,  $b=2$ ,  $c=4$ ) with periodic bonds, y-z plane. d)  $\beta$  chitin crystal (input parameters  $a=2$ ,  $b=2$ ,  $c=2$ ) y-z plane. e)  $\beta$  chitin crystal (input parameters  $a=4$ ,  $b=2$ ,  $c=4$ ) x-y plane. f)  $\beta$  chitin crystal (input parameters  $a=2$ ,  $b=4$ ,  $c=4$ ) x-y plane. Chitin molecules are shown in bond representation in all cases (O in red, H in white, N in blue and C in grey). Figure made with VMD. [1].

”yes” bonds as a direct link between the first residue of chitin chain with the last residue. In a periodic boundary condition representation in VMD this would produce an infinitely long chitin chain, as seen in this figure, since it will link the chain with its periodic representation. This structure is located in the folder `/examples/example-c-beta-2-2-4-periodic/`.

In the fourth example (Figure 3d) we created a  $\beta$  chitin crystal with the input parameters for the replicas  $a=2$ ,  $b=2$  and  $c=2$ , resulting in a crystal with the following dimensions: 19.276 Å length, 36.956 Å height and a depth of 20.768 Å. The option periodic bonds was turned to ”no”. Note that this structure is half the size of the second example (Fig. 3b) in the  $c$  direction. The files for this structure are located in the folder `/examples/example-d-beta-2-2-2/`.

In the fifth example (Figure 3e) we created a  $\beta$  chitin crystal with the input parameters for the replicas  $a=4$ ,  $b=2$  and  $c=4$ , resulting in a crystal with dimensions 38.552 Å, 36.956 Å and 41.536 Å. The option periodic bonds was turned to ”no”. Note that this structure is twice the size of the second example (Fig. 3b) in the  $a$  direction. The files for this structure are located in the folder `/examples/example-e-beta-4-2-4/`.

Finally in the sixth example (Figure 3f)) we created a  $\beta$  chitin crystal with the input parameters for the replicas  $a=2$ ,  $b=4$  and  $c=4$  resulting in a crystal with dimensions 19.276 Å, 36.956 Å and 41.536 Å. The option periodic bonds was turned to "no". Note that this structure is twice the size of 3b) in the  $b$  direction. This structure is located in the folder /examples/example-f-beta-2-4-4/.

### 3. Methods

#### 3.1. Crystal structures of Chitin

As we mentioned in the Introduction, there are three known chitin crystal allomorphs,  $\alpha$ ,  $\beta$  and  $\gamma$  chitin. However, we will not consider  $\gamma$  chitin since its structure is not well known and in fact it is sometimes considered as a mixture of  $\alpha$  and  $\beta$  chitin [6, 7, 8]. The crystal structures of  $\alpha$  and  $\beta$  chitin allomorphs are shown in Figure 1. In the case of  $\alpha$  chitin, we show in Figure 1a the primitive cell as determined by X-ray diffraction in Ref. [9]. This is the unit cell employed in our program for  $\alpha$  chitin. Note that  $\alpha$  chitin crystals have a characteristic antiparallel organization of chitin chains, as shown in Figure 1c.

In the case of  $\beta$  chitin, it is important to note that  $\beta$  chitin is partially soluble in water and in fact interaction with water seems to modify the crystal structure [6]. For this reason, we consider its anhydrous form, as determined by X-ray diffraction in Ref. [10]. The structure of the anhydrous  $\beta$  chitin allomorph is shown in Figure 1. The unit cell of  $\beta$  chitin employed in our program is the cell shown in Figure 1b. This is an expanded cell, corresponding to double of the primitive cell in both the  $a$  and  $b$  directions, that facilitates the understanding of the structure and also simplifies the replication algorithm employed in our program. This is due to the characteristic parallel organization of  $\beta$  chitin chains in the crystal (see Figure 1d). This organization of the extended crystal is more easily generated by simple replications of the unit cell shown in Figure 1b, without introducing any practical limitation in simulation studies.

An important feature determined in experiments is the presence of deacetylated groups (chitosan groups) in the fibril structure. This degree of deacetylation in chitin is of approximately between 5 to 15% [11] although there are species that produce chitin of high purity. In our program, we will consider pure chitin so we do not consider the presence of deacetylated groups in the structures. This possibility is too complex to be incorporated in our program at the present time until further modelling studies clarify a realistic way to incorporate this chemical feature.

#### 3.2. Generation of coordinates and topology of crystals

The program generates atomic coordinates of  $\alpha$  and  $\beta$  chitin crystals with dimensions entered by the user. The dimensions of the crystal must be a multiple of the unit cell shown in Figure 1a for  $\alpha$  chitin and in Figure 1b for  $\beta$  chitin crystals. The atomic coordinates and cell vectors of the cells shown in Figures 1a and 1b were converted by us from the cif files containing X-ray crystallography data supplied in Refs. [9, 10] and are distributed with our program in PDB format in order to be understood by the VMD program. These files are given in the folder /structures (alpha-4residues.pdb and beta-4residues.pdb). The user

specifies how many replicas wants to obtain in each of the three main axis ( $a$ ,  $b$  and  $c$ ) and the program (which is a TCL script running inside VMD) loads the required PDB file with the atomic coordinates of the unit cell and replicates the coordinates. Once the coordinates are calculated, an intermediate file is generated containing the coordinates of the atoms in the new crystal cell in PDB format (crystal-alpha-temp.pdb or crystal-beta-temp.pdb). These coordinate files can be used inside input files of certain calculations (for example, geometry optimizations or structural relaxations with DFT packages) but they do not contain enough information for force-field based molecular dynamics (MD) simulation packages. For these MD simulations, topology and structural information (such as atom type, atom connectivity, partial charges,...) is required. In order to generate this additional information, we employ the topology from the carbohydrates section of CHARMM36 force field [2]. Chitin monomers correspond to the standard CHARMM36 residue BGLCNA (beta-N-acetyl-glucosamine). But here, we modify the name of this residue to BGNA in order to have a 4 characters residue name which is the preferred residue name length for several simulation and analysis packages and it is also the residue length specified in the standard PDB format. The corresponding topology file from CHARMM36 "top\_all36\_carb.rtf", modified with the new residue name, is included in the */structures* folder. Our program automatically calls the psfgen package included in VMD to generate a new PDB coordinate file and a PSF structure file combo containing the same crystal coordinates generated before but now including the BGNA residue name, CHARMM36 atom types, charges, bonds, angles and dihedrals. In the generation of the structure PSF file, each chitin chain in the new crystal cell is treated as an individual segment. Each chitin residue in each chain is linked by the "14bb" patch in the topology file, that correspond to the (1-4)beta glycosidic bond. Finally, an additional bond between the first and last residue of the chain can be optionally included, by selection of the option "periodic bonds" by the user (see details of the user interface below). This possibility allows the simulation of infinitely long chains with the use of periodic boundary conditions. The generated pdb and psf files are finally saved into the working folder and are ready to be employed in force field MD simulations. For completeness, we also provide a parameters file ("par\_all36\_carb.prm" in the ForceField folder) with CHARMM36 carbohydrate standard Lennard-Jones, bonds, angles and dihedrals sections of the CHARMM36 force field needed for the simulation.

#### 4. Compatibility and distribution

The chitin builder tool is written in tcl/tk as VMD plugin; it will therefore work on any of the platforms supported by the VMD program (Linux, OSX, Windows). The source is released under a GNU 3.0 license. The latest version of the package can be obtained from the repository. The repository includes installation instructions for different platforms.

#### Acknowledgments

We acknowledge financial support from the Spanish Government through the RTI2018-096273-B-I00 grant and the Severo Ochoa Grant SEV-2015-0496 for Research Centres of Excellence awarded to ICMA B. D.C. Malaspina is supported by the European Union's Horizon

2020 research and innovation programme under Marie Skłodowska-Curie grant agreement No 6655919. We acknowledge discussions with Prof. Lars A. Berglund and Dr Yamila Garcia about chitin materials.

- [1] W. Humphrey, A. Dalke, K. Schulten, VMD: Visual molecular dynamics, *Journal of Molecular Graphics* 14 (1) (1996) 33–38. doi:10.1016/0263-7855(96)00018-5.  
URL <https://www.sciencedirect.com/science/article/pii/0263785596000185>
- [2] O. Guvench, S. S. Mallajosyula, E. P. Raman, E. Hatcher, K. Vanommeslaeghe, T. J. Foster, F. W. Jamison, A. D. MacKerell, CHARMM Additive All-Atom Force Field for Carbohydrate Derivatives and Its Utility in Polysaccharide and CarbohydrateProtein Modeling, *Journal of Chemical Theory and Computation* 7 (10) (2011) 3162–3180. doi:10.1021/ct200328p.  
URL <https://pubs.acs.org/doi/10.1021/ct200328p>
- [3] J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kalé, K. Schulten, Scalable molecular dynamics with NAMD, *Journal of Computational Chemistry* 26 (16) (2005) 1781–1802. doi:10.1002/jcc.20289.  
URL <http://doi.wiley.com/10.1002/jcc.20289>
- [4] D. Van Der Spoel, E. Lindahl, B. Hess, G. Groenhof, A. E. Mark, H. J. C. Berendsen, GROMACS: Fast, flexible, and free, *Journal of Computational Chemistry* 26 (16) (2005) 1701–1718. doi:10.1002/jcc.20291.  
URL <http://doi.wiley.com/10.1002/jcc.20291>
- [5] B. Hess, C. Kutzner, D. Van Der Spoel, E. Lindahl, GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation, *Journal of Chemical Theory and Computation* 4 (3) (2008) 435–447. doi:10.1021/ct700301q.  
URL <https://pubs.acs.org/doi/10.1021/ct700301q>
- [6] M.-K. Jang, B.-G. Kong, Y.-I. Jeong, C. H. Lee, J.-W. Nah, Physicochemical characterization of  $\alpha$ -chitin,  $\beta$ -chitin, and  $\gamma$ -chitin separated from natural resources, *Journal of Polymer Science Part A: Polymer Chemistry* 42 (14) (2004) 3423–3432. doi:10.1002/pola.20176.  
URL <http://doi.wiley.com/10.1002/pola.20176>
- [7] M. Kaya, M. Mujtaba, H. Ehrlich, A. M. Salaberria, T. Baran, C. T. Amemiya, R. Galli, L. Akyuz, I. Sargin, J. Labidi, On chemistry of  $\gamma$ -chitin, *Carbohydrate Polymers* 176 (2017) 177–186. doi:10.1016/J.CARBPOL.2017.08.076.  
URL <https://www.sciencedirect.com/science/article/pii/S0144861717309529>
- [8] D. G. Ramírez-Wong, M. Ramírez-Cardona, R. J. Sánchez-Leija, A. Rugerio, R. A. Mauricio-Sánchez, M. A. Hernández-Landaverde, A. Carranza, J. A. Pojman, A. M. Garay-Tapia, E. Prokhorov, J. D. Mota-Morales, G. Luna-Bárcenas, Sustainable-solvent-induced polymorphism in chitin films, *Green Chemistry* 18 (15) (2016) 4303–4311. doi:10.1039/C6GC00628K.  
URL <http://xlink.rsc.org/?DOI=C6GC00628K>
- [9] P. Sikorski, R. Hori, M. Wada, Revisit of  $\alpha$ -Chitin Crystal Structure Using High Resolution X-ray Diffraction Data, *Biomacromolecules* 10 (5) (2009) 1100–1105. doi:10.1021/bm801251e.  
URL <http://dx.doi.org/10.1021/bm801251e>
- [10] Y. Nishiyama, Y. Noishiki, M. Wada, X-ray Structure of Anhydrous  $\beta$ -Chitin at 1 Å Resolution, *Macromolecules* 44 (4) (2011) 950–957. doi:10.1021/ma102240r.  
URL <http://dx.doi.org/10.1021/ma102240r>
- [11] V. Zargar, M. Asghari, A. Dashti, A Review on Chitin and Chitosan Polymers: Structure, Chemistry, Solubility, Derivatives, and Applications, *ChemBioEng Reviews* 2 (3) (2015) 204–226. doi:10.1002/cben.201400025.  
URL <http://doi.wiley.com/10.1002/cben.201400025>