JACFC (Stability and Aggregation Simulations)

User's Guide

To report any issue / bug, please contact Dr. M. Marucho's lab. Email: csdfts.comphys@gmail.com

CONTENTS

LIST OF FIGURES

LIST OF TABLES

I. OVERVIEW

The Java Application for Cytoskeleton Filament Characterization (JACFC) provides both experts and non-experts in the field suitable tools to determine the polyelectrolyte nature of cytoskeleton filaments. By using comprehensive computer models and high performance algorithms, the software aims to produce valuable data for elucidating the molecular mechanisms modulating the electrical signal propagation, stability and bundle formation of microtubules and F-actin filaments under different molecular (wild type, isoforms, mutants) and environmental (physiological and non-physiological) conditions. As a unique feature, JACFC web application allows users to perform these calculations online without computational restrictions. Ultimately, these studies may figure out whether molecular and cellular alterations substantially alter the equilibrium of interactions and trigger abnormalities in the bundling and signal propagation during various disease states.

JACFC functionality and design build up on our previous JAVA Swing platform CSDFTS [1] and Matlab platform MPBEC [9] graphical user interface software. The first application aims to characterize nanomaterials and molecules of approximately spherical and cylindrical shapes immersed in aqueous electrolyte mixture solutions. Whereas, the second application is able to perform biomolecular electrostatic calculation on biomolecules with irregular shapes. JACFC performs calculations based on our theories and models published in references [7, 10]. The web application version of JACFC was deployed using the webswing application [14].

The version 1.0 of the JACFC offers two modules : 1- Electrical signal propagation, 2- Stability and aggregation. In this guide, we will discuss the second application. The basis for cytoskeleton filaments to overcome electrostatic interactions to form higher-order structures (bundles and networks) appears primarily dominated by the polyelectrolyte nature of these filaments. Typically, when cytoskeleton filaments are immersed in an aqueous medium, their surface reveals a charge-regulated nature due to the protonation/deprotonation reactions of the dissociable functional groups at the solid/liquid interface[17, 18]. Due to the resulting interaction between the surface charge and dissolved ions the later form the electrical double layer (EDL) around the charged filaments[18, 19]. More counterions tend to accumulate, whereas co-ions tend to deplete within the EDL. Another compelling polyelectrolyte feature of cytoskeleton filaments is related to the charge accumulated in the polymerization state, which has been shown to alter its ability to interact electrostatically. Similarly, G-actin isoforms have different charges according to their amino acid sequence. Thus, filaments formed by these isoforms might be more susceptible to bundling at borderline conditions. In a similar vein, mutations affect the charge of monomers and have reportedly caused the mutant filaments to spontaneously bundle. It was also noted that cross-linking is not always required for bundle formation when sufficient filament surface charge neutralization can directly induce bundling due to charge inversion. All these factors have significant effects on the filaments Zeta potential (ZP) and surface charge density (SCD) which dominate their stability and aggregation. Highly positive or negative ZP values introduce large repulsive forces, preventing filaments with

similar electric charge from aggregation and ensuring redispersion in aqueous electrolyte solutions[12, 13, 14]. Beyond the important role of these polyelectrolyte properties of cytoskeleton filaments in biological environments, the relationship between ZP, SCD, and electrolyte solution surrounding the filaments is still poorly understood. JACFC aims to account for extremely sensitive and compelling polyelectrolyte properties of cytoskeleton filaments and environmental conditions playing a fundamental role in their stability and aggregation at infinite dilution.

This JACFC module implements an innovative multi-scale approach able to account for the atomistic details of a protein molecular structure, its biological environment, and their impact on the stability and aggregation of wild type actin filaments. The formulation includes non-trivial contributions to the ZP and SCD coming from the diffuse part of the electrical double layer of G-actins. Additionally, JACFC uses a numerical solver that takes advantage of high performance Fortran90 routines and optimized libraries that enable the user to obtain solutions at low-to-moderate computational cost depending on the distance grid resolution, and the characterization of the actin filament and the electrolyte solution, among other factors. As a unique feature, this multi-scale theory is able to account for molecular structure conformation (mutation) and biological environment (protonations/deprotonations).

Overall, JACFC does not require specialized training and expertise in computational and theoretical biology, which is often an obstacle for many researchers, experimentalists, even students lacking these requirements. By simply holding the mouse pointer over the corresponding text or blank box, the user will find in each screen helpful information about how to fill out the input data. The user will also find default values for key input parameters and preselected algorithms to speed up the setup of the input data. However, they may be easily changed at any time. Moreover, JACFC tests all the input data before running the application to avoid the incorrect use of the software and prevent meaningless results. At the end of the calculations, JACFC generates two-dimension plots of selected output files to provide graphical visualization of the electrical potential, effective charge, and the ionic distribution around the Actin filament. Finally, all the output data files are properly saved and organized in a user-designated folder for post-analysis purposes. This user's guide also includes examples to illustrate the solver performance and applicability. Further information including a tutorial video and the web application can be found in our website

http://marucholab.myqnapcloud.com/WordPress/stability-and-aggregation/

II. STABILITY AND AGGREGATION MODULE

A. Description

In this section we describe the screen sequence generated by the web application. The first, second and third screens correspond to "*project window*", "*model window*", and "*results visualization window*", respectively. Each screen provides information to help the user fill out the input data by moving the mouse pointer over the corresponding text or blank box (see Figure 1).

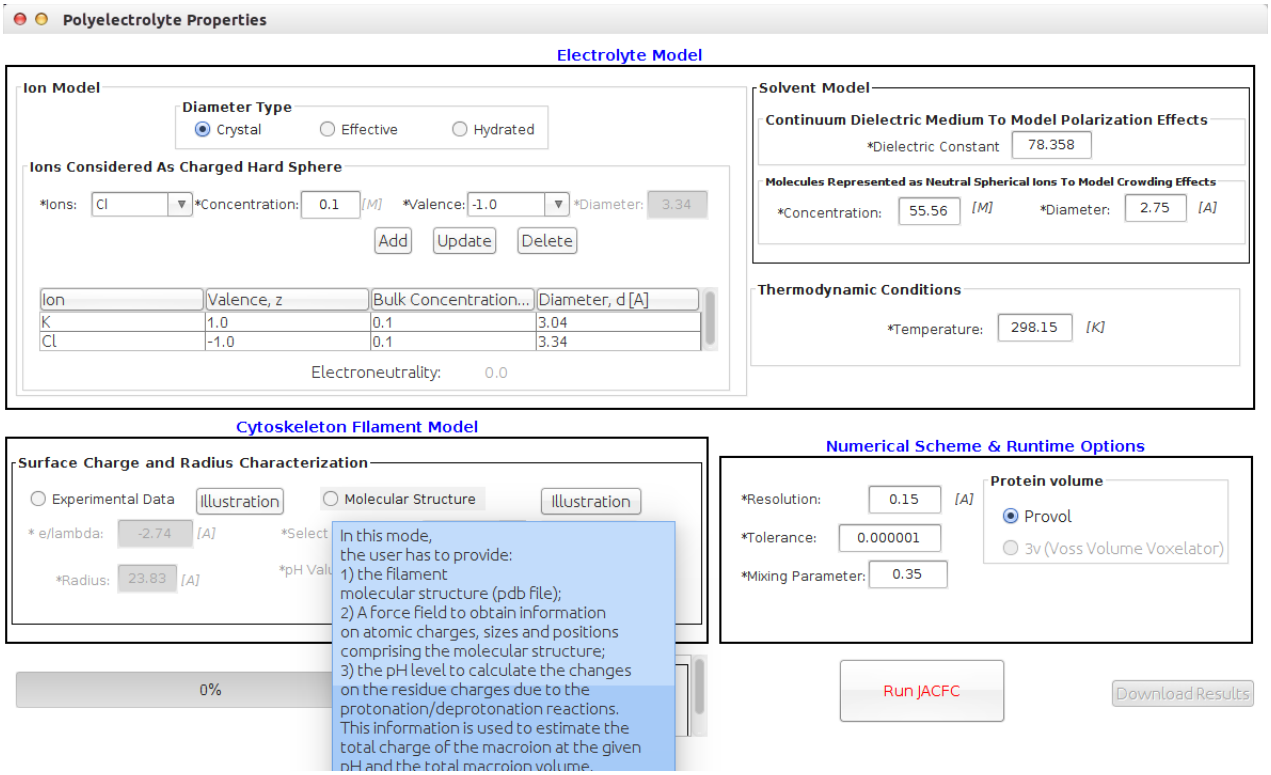


Figure 1: Help Messages.

1. *Main Window: Module Selection*

The Project window shown in Figure 2 is the main screen which provides user access to the *Menu*, *Information* and *Model Sections*.

The menu section, located at the top left corner of the window, contains the *File*, *Tools* and *Help menu*s. The *File Menu* contains the *Results Visualization* and *Exit* options (see Figure 3(a)). The *Results Visualization* option may be used after running the simulation. This option allows users to select output file(s) and visualize the solution(s) in two dimensional plot(s).

The *Tools Menu* contains the *Browse PDB files*, *Crystal Ion Library, Effective Ion Library* and Hydrated Ion Library options (see Figure 3(b)). The tool *Browse PDB files* is used for biophysical applications only. It allows users to open a web browser and download protein molecular structure(s) from the protein data base web page (http://www.rcsb.org/). The user may use this tool at anytime. The *Ion Library* tools provide the user access to the tables pretabulated with specific information on ion species, valences and diameters that are required to characterize the electrolyte aqueous solution. These tools can be used to redefine existing ion species and define new ion species as well. The last menu in Figure 3(c) is the *Help Menu* which provides access to the user's guide, software license and contact information.

The *Select the research study* section is located at the central-bottom part of the window. To select the Stability and Aggregation module, the user has two options: *Dilute conc. (single filament characterization)* and *High conc. (Bundle and Network formation).* Only the former option is available in this JACFC version.

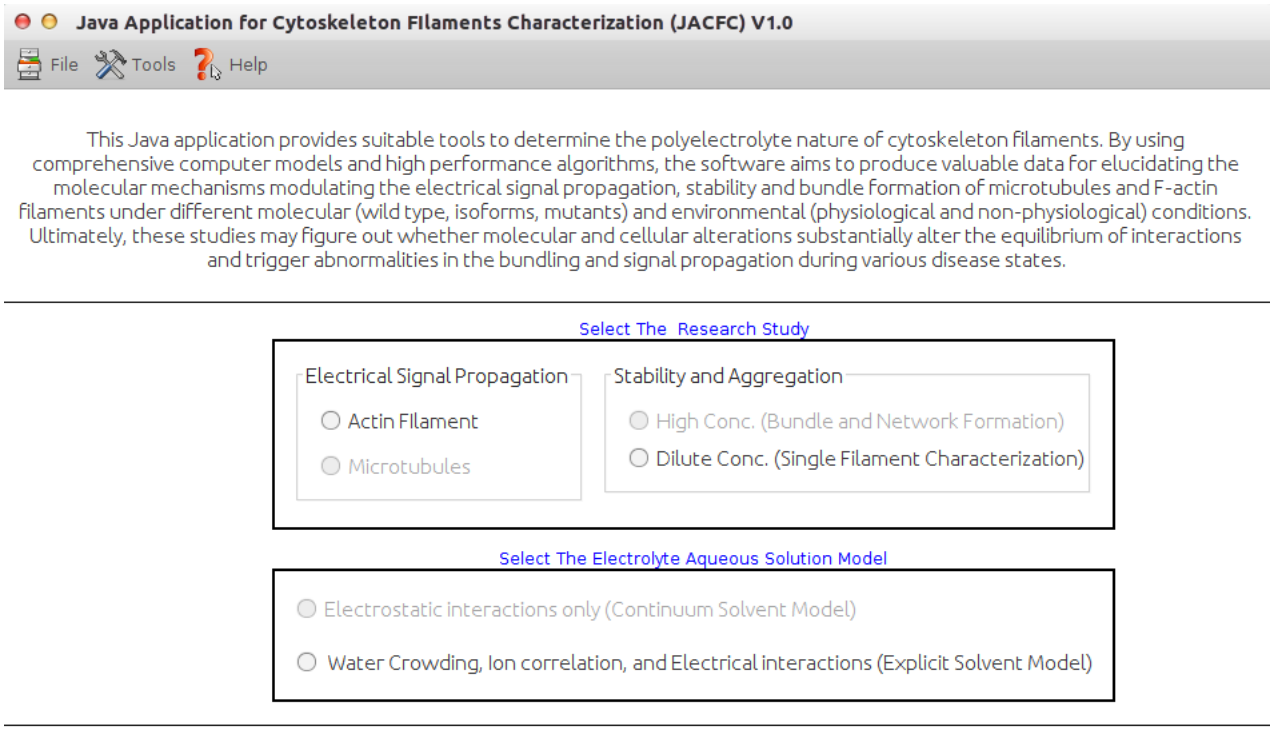


Figure 2: Project Window - main screen.

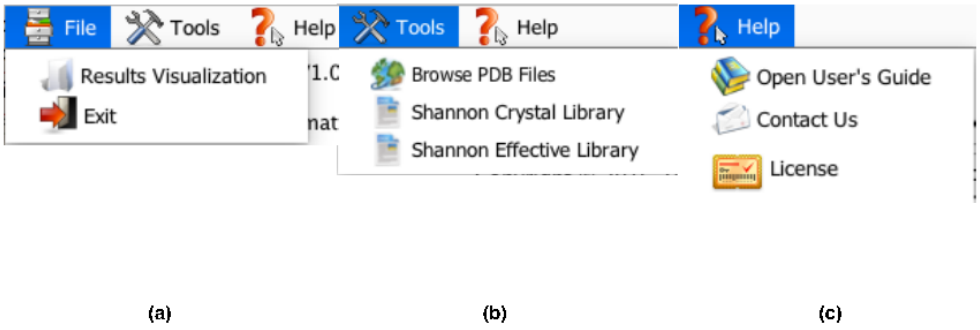


Figure 3: (a) File Menu. (b) Tools Menu. (c) Help Menu.

The *Electrolyte Aqueous Solution Model* is located right below the *Select the research study* section. JACFC offers two electrolyte aqueous solution theories. NLPB uses an implicit solvent model and considers electric interactions only, whereas, CSDFT uses an explicit solvent model (SPM) and considers not only the electric, but also the entropic and ion-ion correlation interactions (see comparative Figure 4).The selection of the *Electrolyte Aqueous Solution Model* model mainly depends on the computational resources, the electrolyte conditions, polyelectrolyte properties of the filament, and the accuracy required for the numerical solution (illustrative examples are provided below). As a rule of thumb, NLPB is more efficient but less accurate than CSDFT. It is worth mentioning that CSDFT is a unique feature of the software, whereas NLPB theory is included for testing and comparison purposes, mostly. Indeed, there are other efficient programs based on implicit solvent models such as APBS (http://www.poissonboltzmann.org/) and MPBEC [9] who provide the solution for both symmetric and asymmetric macroion shapes.

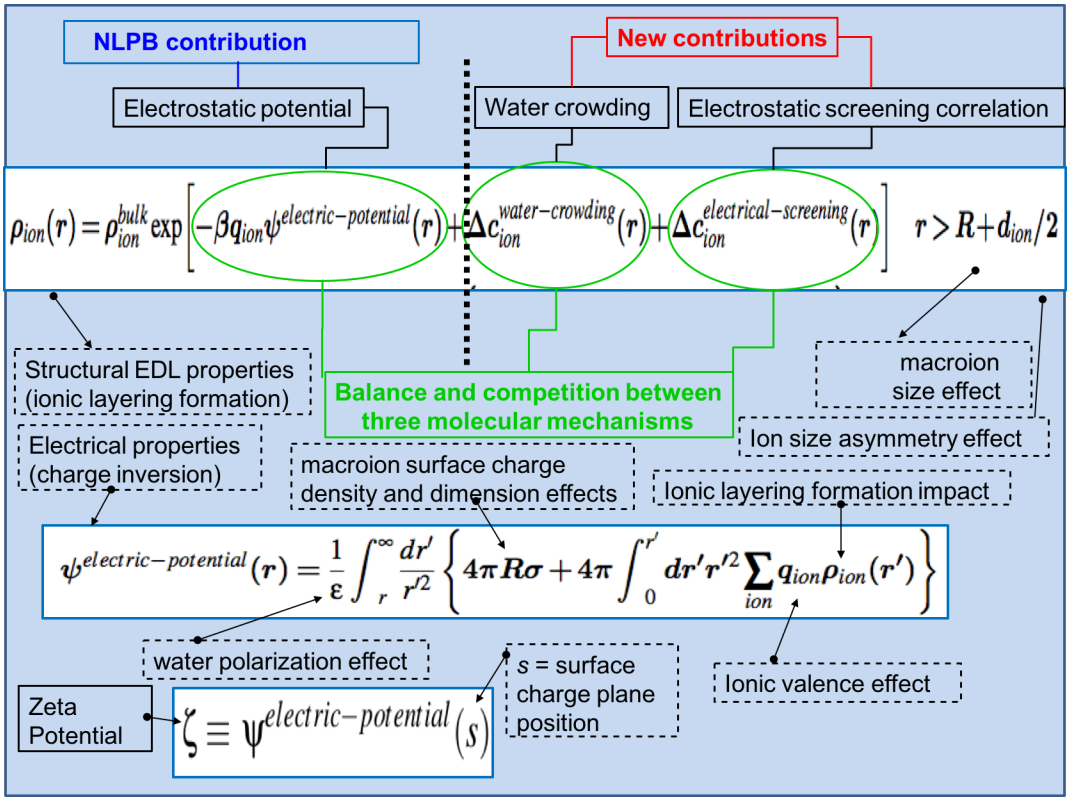


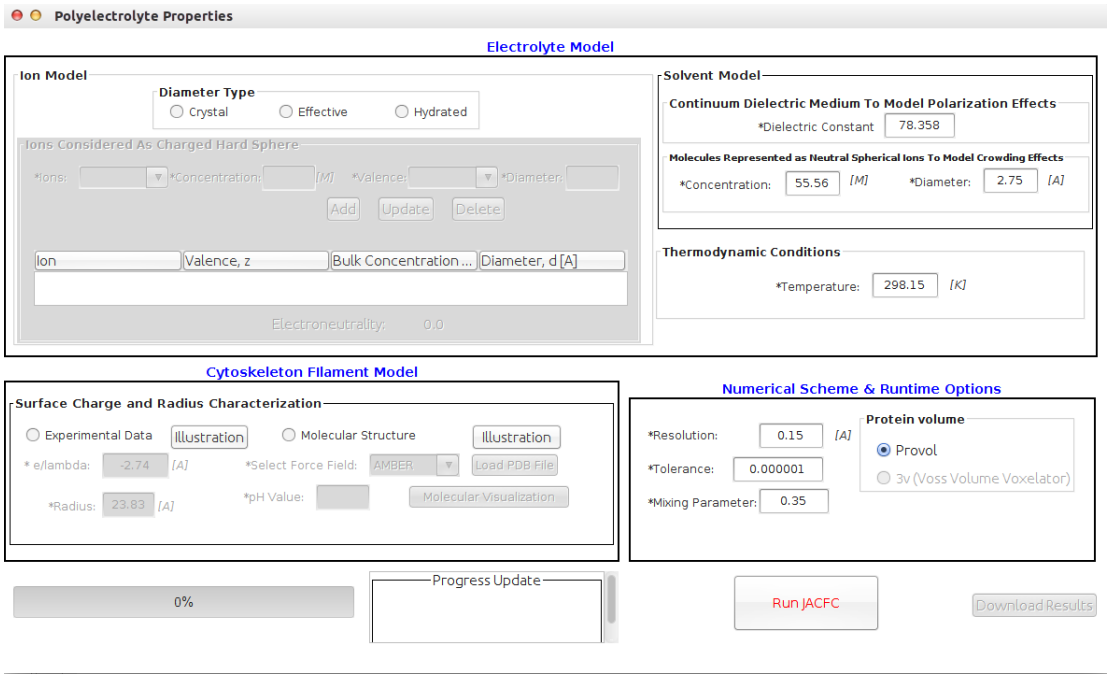
Figure 4: CSDFT theory

2. Model Window: system and solver configuration

In this screen, most of the text based user interaction occurs (see Figure 5). It contains the following modules: *Electrolyte* Model, *Cytoskeleton Filament Model*, and *Numerical Scheme & runtime* Options.

The *Ion Model* section shown in Figure 6 allows users to characterize the electrolyte solution by providing the bulk concentration $[\rho_i^0]$ and the valence z_i for each ion species i. Each ionic species i is represented by charged hard spheres of diameter d_i . JACFC offers a pretabulated crystal, effective and hydrated ionic diameter types and relative mobilities which are estimated using different experimental techniques [6]. To characterize the electrolyte solution, the user has to select the first ion species and its properties, then click the button “Add”, and subsequently repeat the procedure to add more ion species. The “Delete” and “Update” buttons allow users to remove and change the properties of a previously selected ion. Once the user selects all the ion species (e.g. valence and bulk concentration), the electroneutrality condition (e.g. $\sum_i z_i[\rho_i^0] = 0$) must be satisfied. Otherwise, the user will receive an error message to correct the input data. The status of this condition is displayed at the bottom of this section. Note that only ion species comprising the electrolyte solution must be defined in this section. Hydrogen and Hydroxide ions controlling the pH level of the electrolyte solution are assigned by JACFC automatically.

The *Solvent Model section* shown in Figure 7 allows users to characterize the structural and electric properties of the solvent. JACFC considers the solvent as neutral polar molecules. It requires the value of the uniform bulk dielectric permittivity constant ϵ to model the solvent electrical properties. Additionally, it requires the solvent molar bulk con-



(a) Electrical Signal propagation along F-actins

Figure 5: Model window

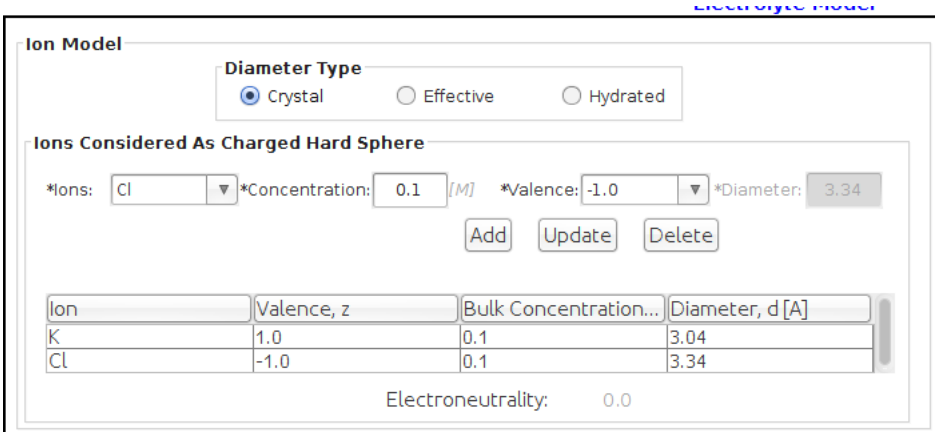


Figure 6: Ion Model.

centration $[\rho_w^0]$ and molecular diameter d_w to model the solvent entropy properties. The default values displayed in this section correspond to the experimental values for water.

The *Numerical Scheme and Runtime Option* section shown in Figure 5 allows the user to configure key parameters that play a fundamental role in the solver performance. They control the accuracy and computational cost. The *tolerance* number represents the numerical error required by the user to numerically obtain the normalized density profile solutions. A minimum of six digits of precision is highly recommended (default value). The radial grid *resolution* represents the regular separation distance h between two consecutive points in the domain discretized of the radial distance to solve CSDFT numerically. The value recommended for this parameter is 0.15Å which has been shown to work for most applications on computers without RAM memory restrictions. The domain of the solution ranges from the

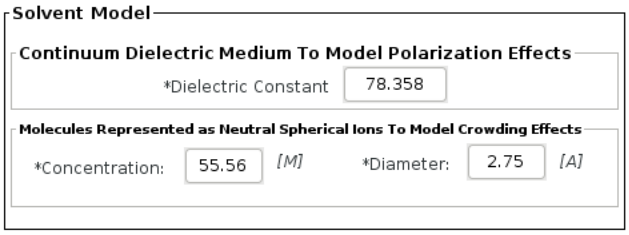


Figure 7: Solvent Model.

macroion surface (e.g. the radius *R)* to the cutoff *L*. The latter is determined automatically by CSDFTS to provide the correct (long range) asymptotic behavior of the mean electrostatic potential, and consequently, satisfy the electroneutrality condition of the system. The number of grid points is calculated as follows $N = (L - R)/h + 1$. The *Mixing Parameter* is a number between 0 and 1 which helps the solver to reach stability and convergence in the solution iteratively. The default value for this parameter is 0.35.

We note that the computing time is highly dependent on radial grid *resolution* for the electric potential. The user may use a lower *resolution* to speed up the calculations and reduce the allocated RAM memory at the risk of loosing accuracy in the calculations.

The *Cytoskeleton Filament Model section* is used to characterize the effective filament radius R and the characteristic longitudinal length $\ell = e/\lambda$ (electron charge e per axial charge density λ at pH 7). Typically, these parameters are obtained from experiments. Different experiments may generate different values, though. The user can provide this information using the *experimental data* model. The *Illustration* button provides visualization of the graphical representation (see Figure 8).

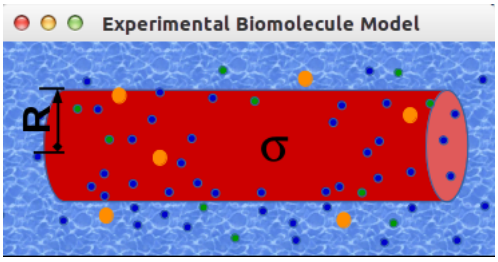


Figure 8: experimental Model.

Alternatively, the *Molecular structure model* option offers a novel calculation for these parameters from molecular structure models. The *Illustration* button provides visualization of the multi-scale approach (see Figure 11). The user has to upload the filament molecular structure (pdb file), select a force field, and define the pH level as shown in Figure 5. The web application has pre-loaded the most recent 13 biologically assembled monomers for actin filaments wild-type molecular structure models [10–14] (see Figure 9).

Additionally, the user is able to upload other molecular structures in pdb format including those accounting for mutations and isoforms. The uploaded molecular structure is used by JACFC to automatically run the pdb2pqr / propka application (http://sourceforge .net/projects/pdb2pqr/). This application assigns atomic charges and sizes, adds missing hydrogens, optimizes the hydrogen bonding network, and renormalizes atomic charges of the residues exposed to the surface due to pH effects (protonation/deprotonation process).

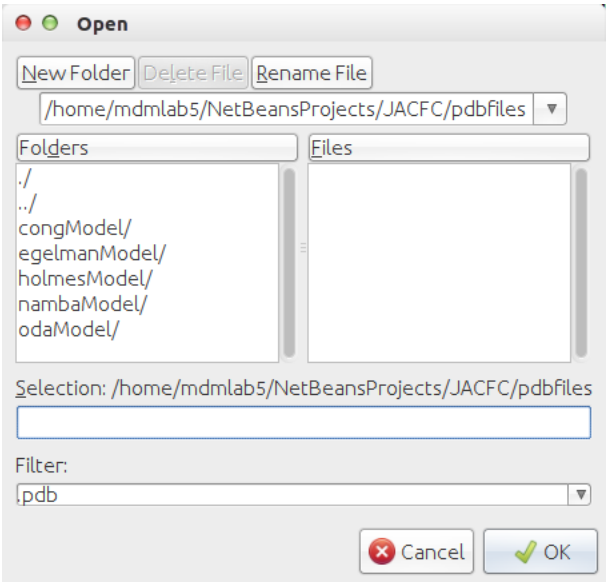


Figure 9: Preloaded Molecular structure Models.

JACFC also uses the molecular structure to estimate the filament volume. It automatically runs either provol (default) or 3v application which rolls a probe particle of radius 1.4Å and use a grid mesh of 0.5Å to generate the filament surface. Finally, the web application uses the information obtained from the total filament charge and volume to estimate the effective filament radius, as well as, the length ℓ. These steps can be monitored in the Progress bar (see Figure 10).

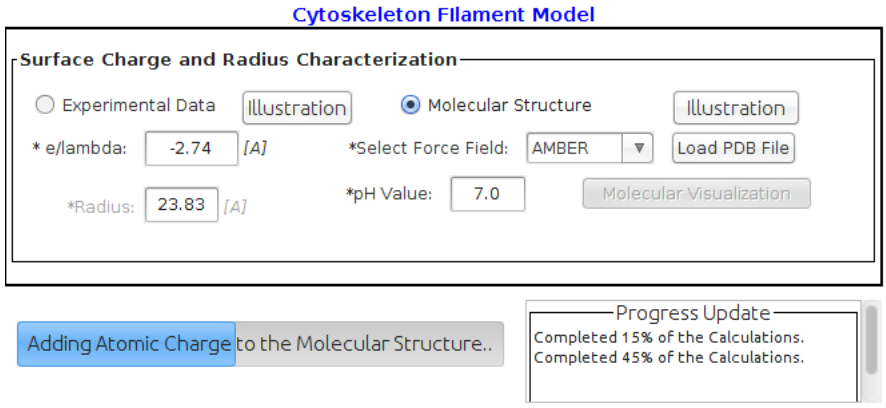


Figure 10: Progress bar tool

Once these calculations are done, the user can visualize the uploaded molecular structure using the *Molecular Structure Visualization* button (see Figure 12). It will open a new window using the 3D Jmol viewer (http://www.jmol.org/).

Note that the default values for the parameters ℓ and R assigned to the experimental data option are those predicted by the Cong molecular structure model for wild-type actin filaments [10] at pH 7 and using the Amber force field.

After all changes on preselected parameters and options are completed, all boxes are filled out and precalculations are over, the user is ready to press the button *"Run JACFC*.

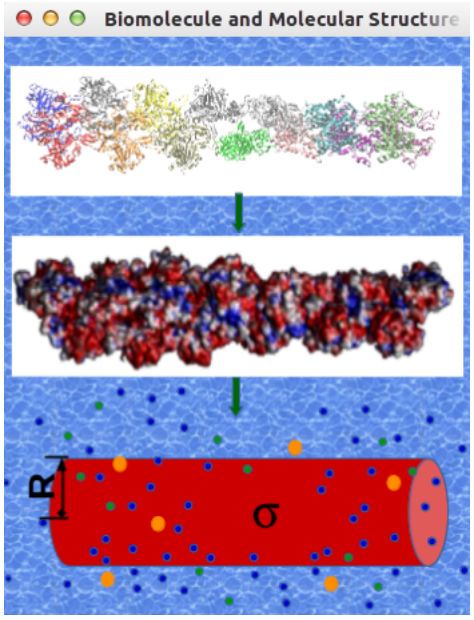


Figure 11: Molecular structure Model representation.

The web application will test all the input data, and send the user warning messages if unusual/nonphysical values are assigned to the required parameters or missing information is detected. Once the configuration pass all the tests, the application will run the simulation and show the performed calculations in the progress bar.

3. Results Visualization

The *results visualization* window will appear once the simulation is over. This screen allows the user to select the output file(s) and visualize the numerical solution(s) in two dimension plot(s) (see Figures 13). To change the plot properties the user must right click on it. The user can also visualize the data on spreadsheets as shown in Figure 13 c).

4. Output files

JACFC generates the following output files: "DensityProfile.dat (normalized ions and water density profile distributions"); "ElectrostaticPotential.dat" (mean electrostatic potential); "IntegrateCharge.dat" (the integrated charge), and "IonContributions.dat"(normalized electrostatic potential energy, particle crowding entropy energy, and ion-ion electrostatic correlation energy to the ionic potential of mean force). Explicit expressions for these properties are provided in Appendix A. The size of these files may vary depending on the number of ions species, number of grid points, and electrolyte model. These files are formatted in a multi-column arrangement where the row number corresponds to the number of grid points, the first column contains the discretized distance in unit of Å and the remaining column(s) provide the numerical solution(s) (see Table I). Other relevant data including SCD, ZP, and PS are saved in the log file "out.txt", whereas information on the input data and solver configuration are included in the file "inputfile.inp". If the *molecular structure* option is

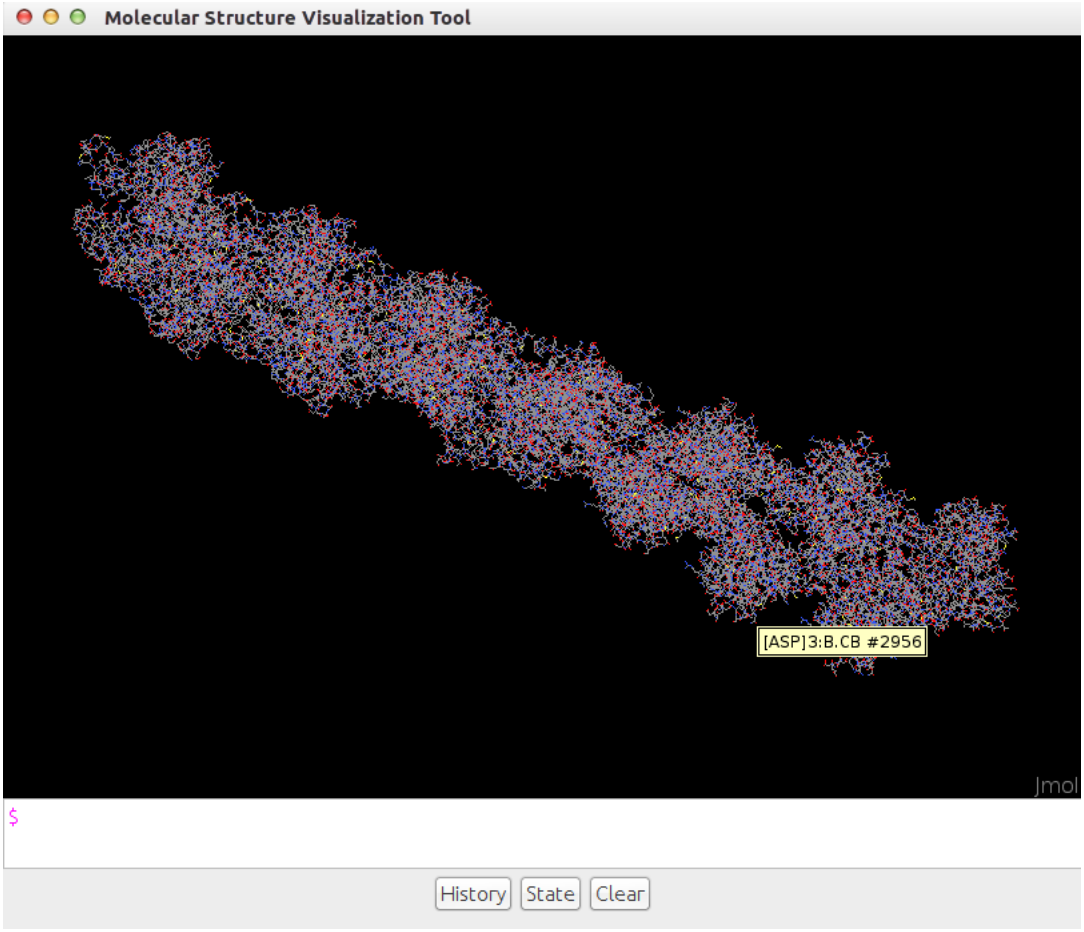


Figure 12: Molecular structure visualization.

selected, JACFC generates the following additional files: “input.pdb” (copy of the molecular structure uploaded by the user); “input.pqr” (pdb2pqr output data); “input.propka” (propka output data) ; “out.pdb” (molecular structure information with heterogem atoms removed); “out.xyzr” (atomic positions in xyzr format); “vol.txt” (3v and provol output data including information on the macroion surface and volume as well as the number of residues and atoms); “lengthandcharge.txt” (total macroion charge and dimensions), and “radiusandcharge.txt” (parameters characterizing the macroion radius and charge density).

Once the simulation is over, the output files can be downloaded in a zip file by pressing the *Download Results* button (See Figure 2).

B. Examples

We consider an actin filament immersed in 0.1M KCl (e.g. 0.1M K^+ and 0.1 M Cl^-, crystal ion type). We use the Cong molecular structure 3B5U.pdb (375 residues), Amber force field, and *pH* 7. We use predefined values for the remaining parameters. The Figures 14 a), b), c), and d) correspond to the normalized ions and water density profile distributions (DensityProfile.dat); the mean electrostatic potential (ElectrostaticPotential.dat); the integrated charge (“IntegrateCharge.dat”), and the normalized electrostatic potential energy, particle crowding entropy energy, and ion-ion electrostatic correlation energy to the ionic

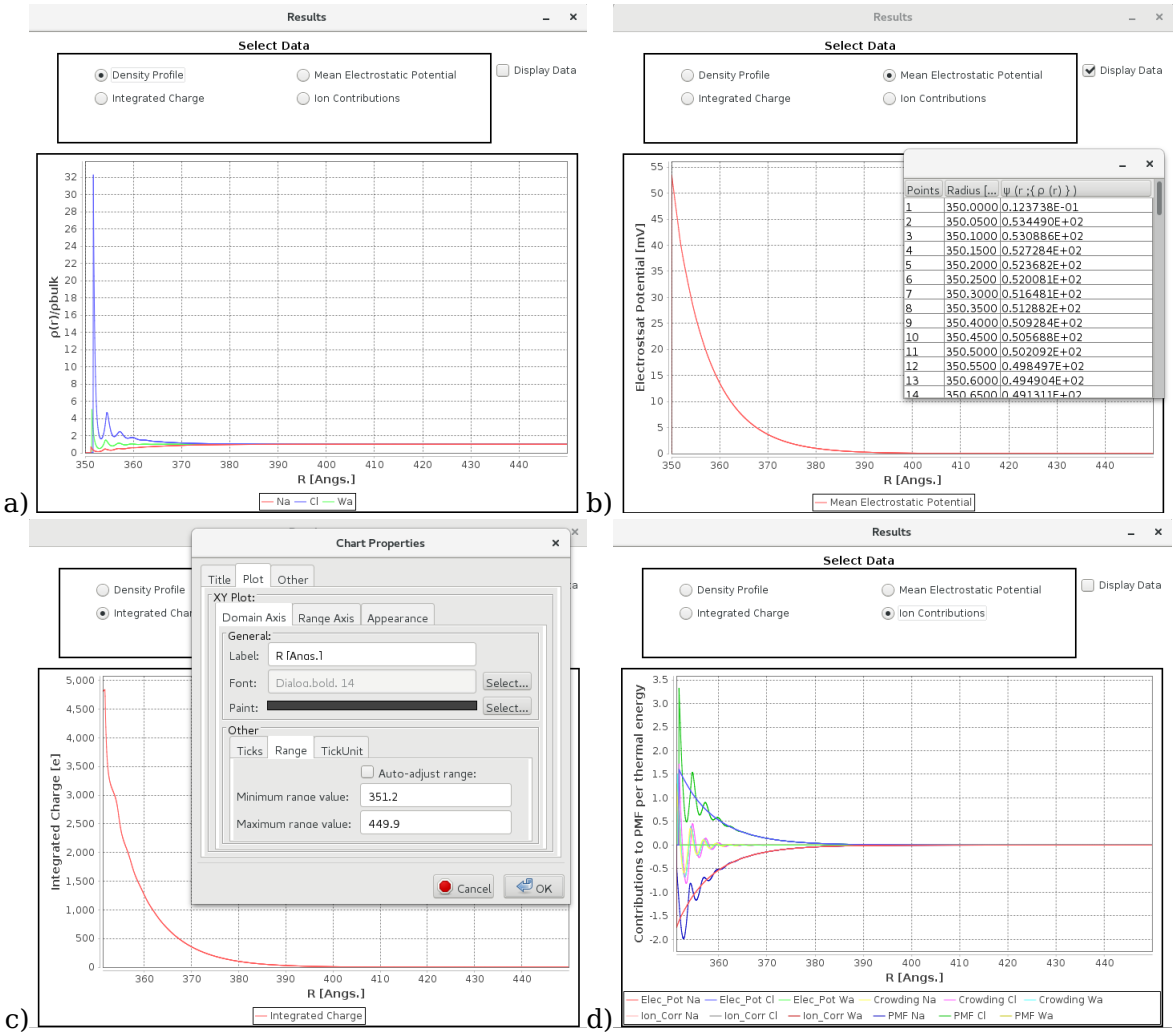


Figure 13: Results Visualization

potential of mean force (IonContributions.dat), respectivelly. These plots provide valuable information on the EDL properties of the filaments (analysis of similar plots can be found in these articles [7–10]).

The molecular structure generates a total filament charge of -154e and total length of 422Å (lengthandcharge.txt). It estimates a total volume = 665165Å^3 (vol.txt). CSDFT used a cutoff of 275.34Å and grid point number 1652 (see output.txt) for the calculations. However, it was used the plot properties in the Figures 14 to visualize the solutions up to a distance of 75Å . The solver took 870 iterations to reach the desired tolerance of 6 digits of precision. The approach predicts an effective filament radius $= 23.38$Å and characteristic longitudinal length $\ell = -2.74$ (radiusandcharge.txt). The corresponding surface charge density is $\sigma = \lambda/(2\pi R) = e/(2\pi R\ell) = -0.00415\frac{C}{m^2}$. It also predicts a Zeta Potential: -37.90mV and Surface Potential: -45.07mV. This result indicates a highly repulsive electric interaction between actin filaments preventing aggregation.

File name	Units	Description
DensityProfile.dat		The first column is for distances. The second column corresponds to the first ion species selected in ion model section, and so on. If the CSDFT is selected, the last column corresponds to water
ElectrostaticPotential.dat	KT/e	The first column is for distances. the second column contains the values of the mean
IntegratedCharge.dat	e	The first column is for distances. the second column contains the values of the integrated charge
IonConstributions.dat	J/KT	The first column is for distances. Then, there is one set of columns per contribution. The number of columns per contribution is equal to the number of ion species. The order of the set of columns from left to right reads : the electrostatic potential energy, excluded volume energy, the ion-ion correlation energy, and the potential of mean force (sum of the previous contributions). The order of the columns in each set of contributions is the following: first column corresponds to the first selected ion species, and so on.

Table I: Output Files description.

ACKNOWLEDGMENTS

This work is supported by NIH Grant SC1GM127187-03.

C. Appendix A: CSDFT summary

In this approach, we consider a rigid charged filament of effective radius R and uniform surface charge density σ surrounded by an electrolyte solution comprised of m ionic species. We use the solvent particle model to characterize the electrolyte. Each ionic species i is represented by bulk Molar concentration $[\rho_i^0]$, a charged hard sphere of diameter d_i, and total charge $q_i = ez_i$, where e is the electron charge and z_i is the corresponding ionic valence. Additionally, the solvent molecules are represented as a neutral ion species whereas the solvent electrostatics is considered implicitly by using the continuum dielectric environment with a dielectric constant ϵ. The macroion-liquid interaction induces inhomogeneous ion profiles $[\rho_i(r)]$ which are calculated using CSDFT as follows [8, 10]:

$$[\rho_i(r)] = \begin{cases} \left[\rho_i^0\right] exp\{\Delta E_i(r, \{[\rho_j]\}), r > R + d_i/2 \\ 0, r \leq R + d_i/2 \end{cases} \quad (1)$$

where $\Delta E_i(r, \{[\rho_j]) \equiv -\beta q_i \psi(r, \{[\rho_j]\}) + \Delta c_i^{(1)hs}(r; \{[\rho_j]\}) + \Delta c_i^{(1)res}(r; \{[\rho_j]\})$ stands for the ionic PMF per unit of thermal energy KT, $\beta = 1/kT$, k is the Boltzmann constant, T the temperature, and $c_i^{(1)hs}(r; \{[\rho_j]\})$ and $c_i^{(1)res}(r; \{[\rho_j]\})$ are the hard sphere (particle crowding) and residual electrostatic ion-ion correlation functions, respectively. $\psi(r, \{[\rho_j]\})$ represents the MEP of the system

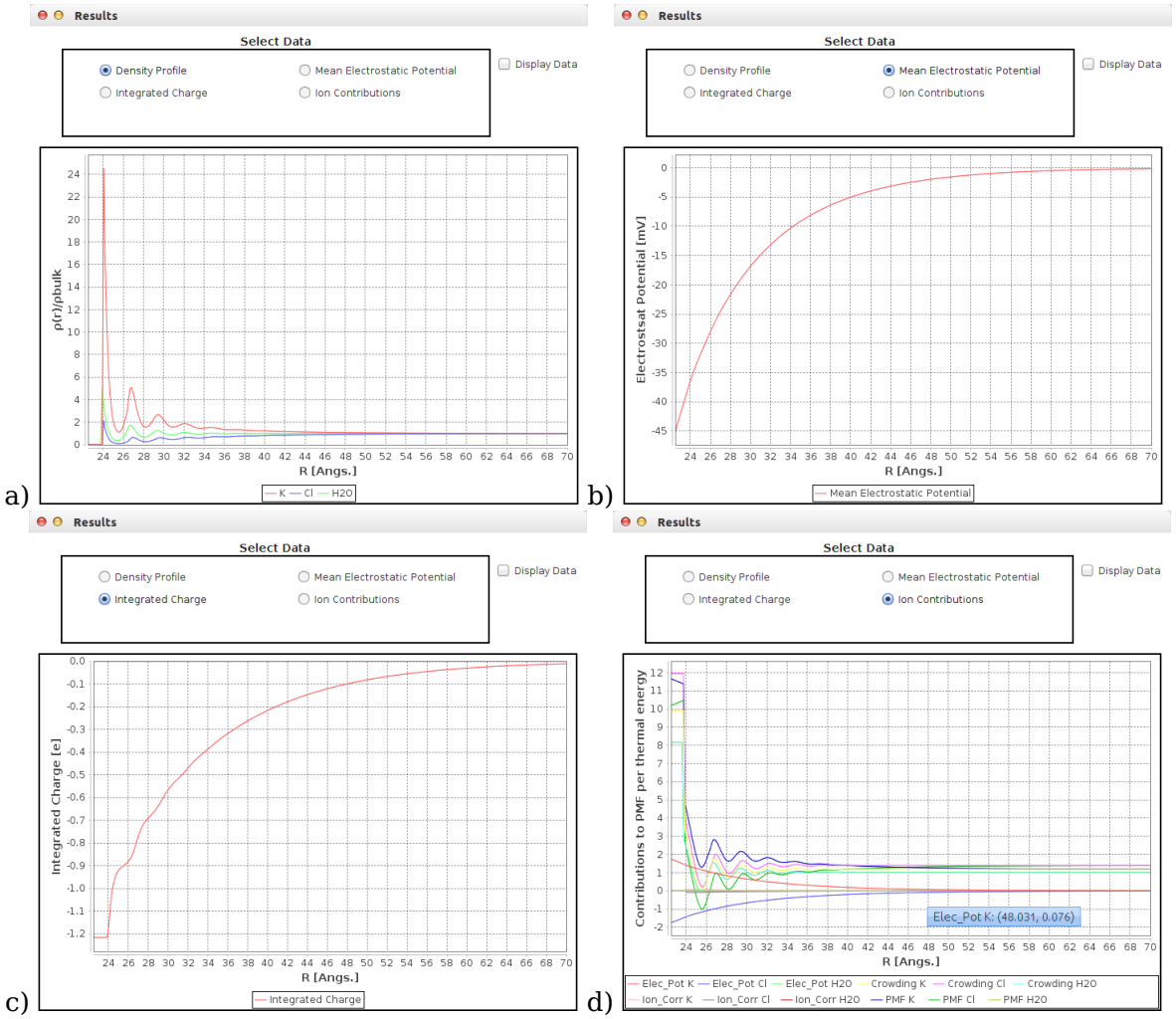


Figure 14: EDL properties

$$\psi(r,\{[\rho_j]\}) = \frac{e}{\epsilon}\int_r^{\infty}\frac{dr'}{r'}P(r',\{[\rho_j]\},n) \tag{2}$$

and

$$P(r',n) = \frac{1}{r'^{n-1}}\left\{\frac{R^n\sigma}{e} + \int_R^{r'} dr' r'^n \sum_i z_i\rho_i(r')\right\} \quad \left\{\begin{array}{l}\text{n=1 for cylindrical macroions}\\ \text{n=2 for spherical macroions}\end{array}\right. \tag{3}$$

the integrated charge (see Figure 4). Expression (2) is the formal solution of the PB equation for an homogeneous anisotropic dielectric media ϵ

$$\begin{array}{c}\nabla^2\psi(r,\{[\rho_j]\}) = -\frac{1}{\epsilon}\sum_{i=1}^{m} z_i\left[\rho_i(r)\right]\} \\ \epsilon\partial\psi(r,\{[\rho_j]\})/\partial r|_{r=s} = -\sigma, \qquad \psi(r,\{[\rho_j]\})|_{r\to\infty}\to 0,\end{array} \tag{4}$$

with the surface charge layer position defined as $s \equiv R+ < \{d_i\} >$ and $< \{d_i\} > \equiv N_A l_B^3 \sum_i |z_i|\,[\rho_i^0]d_i/(2m)$. In the latter definition N_A and l_B stand for the Avogadro number

and the Bjerrum length, respectively.

---

[1] M. Marucho, A Java Application to Characterize Biomolecules and Nanomaterials in Electrolyte Aqueous Solutions, Comput Phys Commun. 2019 Sep;242:104-119.

[2] Vergara-Perez, S. and Marucho, M. (2016). MPBEC, a Matlab Program for Biomolecular Electrostatic Calculations, Comput. Phys. Commun. 198 : 179-194

[3] Christian Hunley, Diego Uribe, and Marcelo Marucho, A Multi-scale approach to describe electrical impulses propagating along Actin filaments in both intracellular and in-vitro conditions, RCS Adv. 2018, 1207.

[4] Ovanesyan, Z.; Medasani, B.; Fenley, M. O.; Guerrero-García, G. I.; Olvera de la Cruz, M. and Marucho, M. (2014). Excluded volume and ion-ion correlation effects on the ionic atmosphere around B-DNA: Theory, simulations, and experiments, J Chem Phys 141 : 225103 (PMCID: PMC4265039).

[5] https://www.webswing.org

[6] R. D. Shannon (1976). "Revised effective ionic radii and systematic studies of interatomic distances in halides and chalcogenides". Acta Crystallogr A. 32: 751–767.

[7] Medasani, B.; Ovanesyan, Z.; Thomas, D. G.; Sushko, M. L. and Marucho, M. (2014). Ionic asymmetry and solvent excluded volume effects on spherical electric double layers: A density functional approach, J Chem Phys 140 : 204510 (PMCID: PMC4039739).

[8] Ovanesyan, Z.; Aljzmi, A.; Almusaynid, M.; Khan, A.; Valderrama, E.; Nash, K. L. and Marucho, M. (2016). Ion-ion correlation, solvent excluded volume and pH effects on physicochemical properties of spherical oxide nanoparticles, J Colloid Interface Sci 462 : 325-333

[9] Hunley, C. and Marucho, M. (2017). Electrical double layer properties of spherical oxide nanoparticles, Phys. Chem. Chem. Phys. 19 : 5396-5404

[10] Yao Cong et all, Crystallographic Conformers of Actin in a Biologically Active Bundle of Filaments, J. Mol. Biol. (2008) 375, 331–336.

[11] Vitold E. Galkin, Albina Orlova, Matthijn R. Vos, Gunnar F. Schro¨der, and Edward H. Egelman, Near-Atomic Resolution for One State of F-Actin, Structure 23, 173–182, January 6, 2015.

[12] K. Holmes et all, Electron cryo-microscopy shows how strong binding of myosin to actin releases nucleotide, NATURE | VOL 425 | 25 SEPTEMBER 2003.

[13] Takashi Fujii , Atsuko H. Iwane , Toshio Yanagida & Keiichi Namba, Direct visualization of secondary structures of F-actin by electron cryomicroscopy, 7 2 4 | N AT U R E | VO L 4 6 7 | 7 O C T O B E R 2 0 1 0.

[14] Toshiro Oda et all, The nature of the globular- to fibrous-actin transition, NATURE | Vol 457 | 22 January 2009.