MPBEC's User Guide Release 1.0 *

Contents

[] * Contact information: Dr. Marcelo Marucho, Biophysics Computation Laboratory, Department of Physics and Astronomy, the University of Texas at San Antonio, TX 78249,USA. Email Address: marcelo.marucho@utsa.edu

I. OVERVIEW

In this user-guide we introduce MPBEC (Matlab Program for Biomolecular Electrostatic Calculations), a free, cross-platform, open-source software that provides an easy and efficient way to perform biomolecular electrostatic calculations. In many biologically relevant systems the solution of the (approximate) linear Poisson-Boltzmann (PB) equation is close to that one obtained from the (exact) nonlinear PB equation, even when the linear approximation does not hold. Additionally, the numerical calculation of the linear PB equation is much more efficient making it the most popular approach for biomolecular electrostatic calculations.

The linearized PB equation is given by the following expression

$$-\nabla.\left(\epsilon\left(\mathbf{r}\right)\nabla u\left(\mathbf{r}\right)\right)+\bar{\kappa}^{2}\left(\mathbf{r}\right)u\left(\mathbf{r}\right)=\frac{4\pi e_{c}^{2}}{K_{B}T}\sum_{i=1}^{N}z_{i}\delta\left(\mathbf{r}-\mathbf{r}_{i}\right) \tag{1}$$

where $u\left(\mathbf{r}\right)=e_{c}\Phi\left(\mathbf{r}\right)/K_{B}T$ is the dimensionless electric potential, $\Phi(\mathbf{r})$, the mean electrostatic potential, $\epsilon(r)$, the dielectric permittivity tensor, K_B, the Boltzmann constant, T, the absolute temperature, e_c, the elementary charge, $\bar{\kappa}$, the Debye-Hückel function (related to the ionic strength), and z is the charge valence.

MPBEC is a MATLAB script based on the Adaptative Poisson Boltzmann Solver (APBS)[1], one of the most popular approaches used to solve the PB equation. MPBEC takes advantage of optimized Matlab routines specially designed to solve large sparse linear system of equations. This feature enables the software to run on single processor computers at low-to-moderate computational cost depending on the computer performance, the biomolecule size, and the grid resolution, among other factors. MPBEC does not require user programming, text editing, or extensive statistical skills since it offers a graphical user interface (GUI) (see Figure 1) that allows users to take advantage of the visually guided setup of the required input data, eliminating errors and reducing the time spent on the setup of the input file. As a unique feature, the GUI provides helpful information about how to fill out the input data by moving the mouse pointer over the corresponding text or blank box (see Figure 2). The GUI tests all the input data before running MPBEC to make sure that the solver is properly configured for the required calculations. Consequently, the GUI does not run MPBEC if the pre-test algorithm detects any unusual/nonphysical value assigned to the required parameters as well as if any required input data is missing. Instead, the GUI sends the user warning messages to inform what is the issue(s) that must be

fixed (see Figure 3). Additionally, before any calculation MPBEC determines the available RAM memory of the user computer and estimates the RAM memory required to perform the required calculations to avoid memory allocation issues. The GUI provides default values of key input parameters and preselects some relevant algorithms to speed up the setup of the input data. These predefined parameters and algorithms may be easily changed at any time if needed. The GUI also integrates a web browser to assist the user in obtaining molecular structures (pdb files) from the protein data bank website. In such a case, the GUI uses pdb2pqr, an open source application (http://sourceforge.net/projects/pdb2pqr/), to automatically convert the molecular structures into the format required to solve the PB equation (pqr files) [1]. Moreover, the GUI facilitates the user the pre- and post- analysis of the biomolecular electrostatic calculations. It integrates Jmol, an open-source Java viewer for chemical structures in 3D (http://www.jmol.org/), to analyze molecular structures (see Figure 4) and display mean electrostatic potential solutions of the PB equation at the solvent accessible surface of the biomolecule (see Figure 5). As a result, the GUI enables the user to visualize the charge distribution, size, and shape of molecules in 3D, which is of fundamental importance in determining how molecules interact with one another. Additionally, the GUI utilizes Matlab plot tools to visualize the surface mean electrostatic potential along the plane $z = L_z/2$, where L_z is the box size along the z direction (see Figure 6). This visualization is useful for analyzing the behavior of the mean electrostatic potential at short and long distances from the biomolecule(s). It is also relevant for analyzing the effective range, intensity, and nature of the electrostatic interaction between the analyzed biomolecule(s) and surrounding ions. At the end of the calculations, MPBEC writes out a log file containing all information displayed on the Matlab console, including step by step intermediate calculations and computing times (see Figure 7). Additionally, all required output data files are properly saved and organized in a user-designated folder for post-analysis purposes.

In the next sections of this user-guide we explain how to run and use the GUI. We also provide several examples to illustrate the use of the GUI and the solver performance. In the Appendix we provide a description on how to run MPBEC without using the GUI.

II. HOW TO RUN THE GRAPHICAL USER INTERFACE

There are two ways to run the GUI. We recommend to use the provided batch files. For Windows users, just double click "mpbec-win.bat" to open the GUI automatically in Matlab console mode. On the other hand, Mac (Linux) users have to run the batch file "mpbec-mac.sh" ("mpbec-lin.sh") from a terminal. For instance, if you are a Linux user and your /path/to/the/file is "/home/smith/Documents/MPBEC/mpbec-lin.sh", you have to run "source /home/smith/Documents/MPBEC/mpbec-lin.sh".
Mac (Linux) users may run "find ∼ -name mpbec-mac.sh 2> /dev/null" ("find /home -name mpbec-lin.sh 2> /dev/null") to get the /path/to/the/file if needed.
Alternatively, the user can run Matlab as usual, change the current directory to
"~/MPBEC/program/src/Main"
using the Matlab search toolbar, and type "Welcome" in the Matlab window console. This GUI introduces MPBEC and provides the user the option to run the Beginner or Advanced level mode depending on the experience of the user performing biomolecular electrostatics calculations (See Figure 8). In the "Beginner" mode users are guided via three simple steps to fill out the minimum information required by the software to perform the calculations (See Figure 9). The user only has to upload the molecular structure files, choose, if needed, an energetic calculation option, provide a name for the input file, and run the calculation. MPBEC will automatically assigned the remaining parameters. On the other hand, the "Advanced" user mode offers full access to the configuration of MPBEC.

Once MPBEC has finished all the calculations, the user may change the configuration of the GUI and run MPBEC again. If the user closes the GUI before all the calculations are finished, the current directory may not be the one required to run the GUI again. Make sure that the current directory in the Matlab console is "~/MPBEC/program/src/Main" before running the GUI. If any unexpected message / warning is received in the Matlab window console and the user is unable to run the GUI, the user must exit Matlab and start over again..

III. HOW TO USE THE GRAPHICAL USER INTERFACE

It is highly recommended to provide the mandatory input data first, e.g. the *"molecular structure"* file(s), the "*output directory",* and the "*input file name*", and subsequently any other information if needed. Preselected options and predefined values presented in the main panel can be easily changed by the user at any time if needed. The user should be able to obtain the required PB solution without using optional panels. However, the user may need to provide additional information using optional panels depending on specific (advanced) calculations. MPBEC runs by pressing the *"Compute Solution"* button.

Below we explain how to change any preselected option and predefined value as well as how to setup optional panels.

GUI description. There are several panels in the GUI that help and guide the user to configure and setup all the required parameters to solve the linear PB equation (see Figure 1).

The "*Charge map approx*" and "*Dielectric and kappa maps approx*" panels (also referred as chm and dkm, respectively) enable the user to select several options to calculate the charge density, dielectric, and kappa (ion-accessibility) maps. SPL2 and MOL, two of the most frequently used approximations, are preselected for the *"Charge map approx"* and "*Dielectric and kappa maps approx"* panels, respectively. Alternatively, the user can activate the "*Map files (optional)*" panel to upload map files in dx format. Clearly, the use of this optional panel automatically de-activates the "*charge map approx*" and *"Dielectric and Kappa map approx*" panels.

By default, the GUI is preconfigured to study biomolecules under normal physiological conditions. A monovalent NaCl salt solution at 0.15M concentration is predefined in the "*ionic species (optional)*" panel. Additionally, the solvent and solute dielectric coefficients are predefined to model water (ϵ_w=78.54) and proteins (ϵ_p =2) at room temperature (T=298.15K), respectively. These values can be changed as needed.

The "*molecular structure*" panel enables the user to upload molecular structures in either pdb or pqr format (see Figure 10). At least one file must be uploaded (molecule 1). For binding energy calculations, the user must also upload the molecular structures corresponding to the second biomolecule (molecule 2) and the complex system (molecule 1 plus molecule 2). It is not mandatory to upload the reference system file when only

one molecule is uploaded (molecule 1). The pqr format is preselected to upload molecular structures. If the user does not have pqr file(s), the user must provide the corresponding pdb file(s). In the case neither pqr nor pdb file(s) is available, the user may use the "*Need pdb files?*" button to get the molecular structures from the protein data bank website (internet access is required). MPBEC automatically converts pdb into pqr files using pdb2pqr and AMBER (http://ambermd.org/), one of the most popular force fields used in biomolecular electrostatic calculations. The force field is defined in the file "~/MPBEC/program/src/Main/pdb2pqrscript.st" and it can be changed at any time. The uploaded molecular structure of each of the molecules can be visualized by clicking on the corresponding visualization button ("*visualize mol 1?*", "*visualize mol 2?*", "*visualize complex?*").

The "*write out files*" panel, located at the middle of the main panel, is preselected to write out the mean electrostatic potential solution file (in dx format) and the surface plots for post-analysis purposes. The user is also able to generate map files by clicking on the corresponding blank boxes. The "*output directory*" box must be used to select a pre-existing directory to save the output files. By default, the software preselects the output directory "~MPBEC/program/src/examples/solutions". MPBEC creates one folder in the selected output directory to save data corresponding to the target (fine) grid calculation. If the user selects *"Optional Focus Calculation"*, MPBEC creates a second folder to save data corresponding to the focus (coarse) calculation. Each of these folders may contain several files to know: the logout file, the input files, the maps files, the energy file, the mean electrostatic potential file, the propka file and the surface plots, depending on the required calculations and selected options.

The "*grid information*" panel is located below the "*write out files*" panel. In the preselected "*automatic*" option, the GUI estimates the box size and the number of grid points based on the dimensions of the biological system. The GUI generates a box size 1.5 times larger than the length of the system along each direction. For average size systems smaller than 100Å, the *"automatic"* option uses a fine mesh size less than or equal to 0.5Å and number of grid points less than or equal to 300 along each direction. These conditions provide good mesh resolution while preventing excessive computational cost and massive use of RAM memory. Otherwise, the *"automatic"* option uses the focusing boundary condition. Alternatively, the user may run MPBEC using the *"manual"* option. In this case, the user

must provide the information on the box size and number of grid points along each of the Cartesian axis. For large biomolecules, the *"Optional focus calculation"* is more efficient than the corresponding target (fine) grid calculations. Another advantage of using focus calculation is to ensure that the a priori imposed Dirichlet boundary condition is satisfied.

The "*Linear Solver Setup*" panel enables the user to select the Matlab linear solver algorithm ("*Biconjugate gradient*", "*Gmres*", "*Minres*") and to define the error tolerance ("*Digits of precision*") required on the mean electrostatic potential solution. The Biconjugate gradient method and 6 digits of precision are preselected by default. The "*Boundary Condition*" panel has two options. While computationally more expensive, the preselected "*Dirichlet (mdh)*" option usually provides more accurate boundary conditions than "*Dirichlet (sdh)*".

The "*input file name*" panel is used to provide the input file name which must end with the extension ".inm". MPBEC uses the input file name to create output file names.

The "*Energy Calculations (optional)*" panel is located in the right-hand lower quarter of the main panel. By default, there is no preselected energy calculation. The user may select one or more energetic calculations depending on the number of uploaded biomolecules. For one biomolecule (e.g. molecule 1), the user may select three options: "*Electrost. Pot. Energy*" (Electrostatic potential energy), "*Solvation free Energy*" and/or "*pKa*". In the later case, the user must provide the "*pH*" level of the aqueous electrolyte solution. On the other hand, the "*Binding Energy*" calculation can only be selected in the case that the user uploads the files corresponding to two biomolecules (protein and ligand) and the complex system.

IV. SOLVER PARAMETERS AND ALGORITHM OPTIONS

A. Discretization of the physical parameters involved with the solution of the PB equation

The accuracy of the mean electrostatic potential solution depends on the algorithms and parameters used in the discretization of the charge density distribution and the kappa and dielectric functions appearing into the PB Eq. (1).

Charge density map **(chm)**. The position of the atoms obtained from the pdb or pqr file may not coincide with any of the grid points. The *chm* panel has the following options to specifies the algorithm used to spread the point charges (summation term in Eq. 1) on

the grid points.

SPLO, Trilinear interpolation (linear splines). This method locates the charge of each atom in the nearest neighbor grid point. The resulting charge density map may not be a smooth function, having sharp corners at the data points. SPLO has been shown to be an efficient algorithm to generate charge density maps but the accuracy of the PB solution may depend on the mesh, the box size, and the protein structure.

SPL2, Cubic B-spline interpolation (basis spline). This option results in the location of the charge in the nearest and next-nearest grid points. Whereas this calculation requires higher computational cost than SPLO, it provides softer distributions of charges which in turn generate mean electrostatic potential solutions less sensitive to the *"Grid information"* parameters. More information on interpolating splines can be found in reference [2].

***Dielectric and kappa maps* (dkm)**. This panel specifies the algorithm and parameters used to build up the dielectric (first term in Eq. 1) and the kappa maps (second term in Eq. 1).

MOL. The dielectric coefficients are assigned according to the definition of the solvent accessible surface. This surface is calculated using the Collony molecular surface approach [3]. A solvent sphere is used to probe the molecule shape and create a surface defined by the union of the solvent-sized spheres, equidistant to the surface atoms in the molecule. The radius of the solvent probe *"srad"* is preassigned with the value 1.4Å, which represents the radius of a water molecule (see Figure 1). The *"surf dens"* parameter is usually set equal to 10 and represents the number of points per $Å^2$ used in the calculation of the Collony molecular surface. The volume outside the solvent accessible surface is assigned with *"solvent dielectric coefficient"* (preassigned as 78.54) whereas the rest of the space is assigned with the *"solute dielectric coefficient"* (preassigned as 2). Setting *"srad"* equal to zero generates a typical van der Waals surface. The ion-accessibility map (kappa map) is built up on a modified van der Waals surface. The radius of each atom in the molecule is inflated with the radius of the ion species (See *"Ionic species"* panel in Figure 1). The space inside the modified surface is assigned with an ion-accessibility coefficient equal to zero whereas the volume outside of this surface is assigned with the corresponding kappa values. The MOL method is very efficient, mainly for large proteins. However, the abrupt transition of the dielectric and kappa maps at the solute-solvent interface may give rise to discontinuous variations of the electrostatic free energies.

SMOL. This option includes a harmonic interpolation that smooths the surface on the dielectric and kappa maps generated by MOL, which reduces the sensitivity of the PB solution to *"Grid information"* parameters at the expense of increasing the computational cost [4].

SPL2. More suitable approaches (compared to MOL and SMOL) are required in those calculations demanding the evaluation of the slope of the mean electrostatic potential (e.g., electric force). In this case, the dielectric and kappa maps are built up on a cubic-spline surface where an intermediate dielectric region is introduced at the solute-solvent boundary to smooth the transition. The parameter *"Swin"* specifies the rate of change (e.g. the window width) in the dielectric interface [5]. The preassigned value for this parameter is 0.3Å.

B. Energy Calculations

The free energy calculations are usually divided into two contributions: a nonpolar and a polar contribution. MPBEC calculates only the polar contribution to the free energy. Although the polar free energy calculations alone are not always sufficient to explain biological phenomena, they are the foundation for more complex, electrostatic calculations. A detailed description of polar free energy calculations can be found in reference [6].

The ***"Energy Calculations (optional)"*** panel provides the user four free energy calculations (kJ/mol) which may be chosen depending on the number of uploaded molecular structures.

Electros. Pot. Energy. MPBEC calculates the total electrostatic potential free energy stored in a single molecule. This energy represents the work required to assemble the molecule. It is calculated as

$$G_{elec} = \frac{1}{2}\sum q_i \Phi(\mathbf{r_i}) \tag{2}$$

where $\Phi(\mathbf{r_i})$ is the mean electrostatic potential acting on the atom i located at the position $\mathbf{r_i}$ and carrying on charge q_i.

Solvation free energy. MPBEC calculates the polar contribution to the free energy released during the solvation of the biomolecule from vacuum to the solvent of interest as follows

$$\triangle G = G_{elec}^{Solv} - G_{elec}^{Ref} \tag{3}$$

where G_{elec}^{Ref} (reference energy) is the total electrostatic potential free energy of the molecule in the reference state (solute uniform dielectric medium), whereas G_{elec}^{Solv} (solvated energy) is that one in the solvated state (inhomogeneous dielectric medium formed by a cavity of solute dielectric constant surrounded by a solvent dielectric constant). In this calculation, the values of the solute and solvent dielectric constants are preassigned as 2 and 78.54, respectively.

Binding energy. MPBEC calculates the polar contribution to the binding free energy between two molecules (protein and ligand) defined by

$$\triangle G_{binding} = G_{elec}^{complex} - G_{elec}^{protein} - G_{elec}^{ligand} \tag{4}$$

Although this approach does not account for conformational changes of the complex after formation, it is a good approximation to study interactions between biomolecules.

pKa. Protein active sites are able to protonate/deprotonate residues depending on the number of hydrogen ions (pH level) present in the aqueous electrolyte solution. This titration process may affect the function, stability, and molecular recognition of proteins. MPBEC uses PROPKA (www.propka.org) to calculate the protonation (positioning of hydrogen ions) to a given molecular structure (in pdb format) at the specified pH level [7]. PROPKA generates a new pdb file containing the re-normalized protein charge distribution, which is subsequently used to solve the linear PB equation.

C. Focus Calculations

In focusing calculations, the Matlab script solves the PB equation twice automatically. The first time it calculates the solution in the coarse (large) grid box and, subsequently, it uses this solution on the faces of the fine (small) grid box as the boundary condition to calculate the solution in the fine grid. Note that the Coarse box size must be greater than the fine box size.

D. Linear equation system solver

The *Linear Solver Setup* panel provides three options ("*Biconjgrad*", "*Gmres*", and "*Minres*") to select the algorithm that MPBEC uses to numerically solve $Au = b$, namely the

discretized version of the linear PB Eq. 1. The square matrix A contains the information on the dielectric and kappa maps, the column vector b contains the information on the charge map and boundary condition, and u represents the unknown discretized dimensionless mean electrostatic potential vector.

MPBEC uses sparse matrix calculations and iterative linear solvers to solve the discretized linear PB equation. The three solvers used by MPBEC are:

Gmres. MPBEC calls the Matlab function *[u,flag]=gmres(A,b,restart,tol,maxit,L,U,uo)* to solve the linear equations system using generalized minimal residual method. The solver is automatically restarted after *10* iterations.

Minres. MPBEC calls the Matlab function *[u,flag]=minres(A,b,tol,maxit,L,U,uo)* to solve the linear equations system using minimal residual method.

Biconjgrad. MPBEC calls the Matlab function *bicgstab(A,b,u,tol,maxit,L,U,uo)* to solve the linear equations system using biconjugate gradients stabilized method.

Notice that none of the aforementioned three options is the best for all biomolecules and aqueous electrolyte solutions (e.g. for all matrices A and vectors b). There are always examples in which one option outperforms the other. Therefore, multiple solvers should be tried in practice to see which one is the best for a given problem. In general, the "*Biconjgrad*" option performs well for all the proteins provided in the example section. If this method does not converge, MPBEC automatically uses *"Gmres"*, and subsequently *"Minres"*, if needed. More information on iterative methods for solving linear equations can be found in reference [8].

V. COMPUTER REQUIREMENTS

Java and Matlab must be installed on the computer. MPBEC was successfully tested on PC Windows 7 and Windows 8.1, Linux Fedora 16 and Fedora 19, and OSX Yosemite running Matlab R2010a and newer versions. We recommend to run MPBEC on computers having at least 2 GB of free RAM memory and 2.5 GHz CPU speed or higher. Some examples of estimated RAM memory allocations required by the automatic grid configuration to run MPBEC are provided in Table VII. As it happens with other applications, it is recommended to keep only one copy of the MPBEC software in your computer to avoid potential Matlab path search issues.

VI. EXAMPLES

In this section we present the results obtained for several biological systems to illustrate the use of different solver algorithms and parameters as well as the solver performance. We also compare MPBEC against APBS (version 1.4) results for testing purposes. In all these systems we use the following parameter values: solvent dielectric coefficient= 78.54, temperature= 298.15K, boundary condition in the fine grid= mdh, solver method= biconjgrad, and digits of precision= 6. We also use the *"manual"* option to setup the grid information unless otherwise is stated. The molecular structure input files used in this section can be found in the directory ~/MPBEC/program/src/examples.

Table I shows the polar solvation free energy results obtained for the Born ion in a box of 32Åx32Åx32Å for different number of grid points and, consequently, different grid spacing. We use the *"Solvation free energy"* option in the *"Energy calculations"* panel, and set the ionic concentrations equal to 0, chm= SPL0, dkm= MOL, and solute dielectric coefficient= 1. Table II shows the polar solvation free energy results obtained for the same biological system using 193 grid points along each direction, and different charge (SPL0, SPL2) and dielectric (MOL, SMOL, SPL2) approximations. Tables III and IV show the polar solvation free energy results obtained for Methanol and Methoxide (ionized methanol), respectively, in a box of 32Åx32Åx32Å for different charge approximations (SPL0, SLP2). These molecules are in the presence of a single monovalent salt (NaCl) with ionic radius= 2Å and 0M, 0.05M, and 0.5M salt concentrations. We use 193 grid points along each direction and set dkm= MOL, and solute dielectric coefficient= 2. In these examples we observe that the solvated and reference state energies are sensitive to the different parameters and approximations whereas the polar solvation free energy remains approximately constant.

Tables V and VI illustrate the use of the *"Binding energy"* and *"Optional Focus calculation"* options for two complex systems, namely hca (protein)-acetazolamide (ligand) and 1d7h (protein)-DMSO (ligand). We use the *"Optional Focus calculation"* with 129 grid points along each direction, a focus (coarse) box of 96Åx96Åx96Å and a target (fine) box of 57Åx57Åx57Å. We also set dkm= MOL, chm= SPL0, and solute dielectric coefficient= 2. While the hca-acetazolamide complex is not immersed in a electrolyte, the 1d7h-DMSO complex is in presence of a monovalent salt with concentration of 0.01M and ionic radius of 2Å .

Figure 11 illustrates the use of the *"pKa"* option. It presents the surface potential of the protein 2KAC at three pH levels (12, 7, and 3). In this case, we select the *"Automatic"* option in the *"grid information"* panel and use the preassigned parameters. The propka file saved in the *"output directory"* provides useful information about the charge renormalization at several pH levels. In particular, the total protein charge increases from -15e to 0.7e when the pH level decreases from 7 to 3, whereas it decreases from -15e to -20e when the pH level increases from 7 to 12. Additionally, the isoelectric point predicted for this protein is 3.33.

Table VII presents the polar solvation free energy results obtained for several proteins using the preselected parameters and the "*automatic*" grid information option. We also include the registered computing time obtained for proteins of different sizes using Matlab R2012b in a single CPU Intel Xeon X5670 Westmere 2.93 GHz under Linux Fedora 19.

APPENDIX: HOW TO RUN MPBEC WITHOUT USING THE GUI

Running MPBEC without using the GUI is not recommended unless MPBEC is implemented into other Matlab programs or the user has Matlab programming and text editing skills as well as a strong background in biophysics. The user must first include the Matlab search path corresponding to the folder ~/MPBEC and subfolders from the Matlab console. It must also generate the input file named "inputfile.inm" using any text editor (see input file structure section for instructions). This file must be saved in the directory ~/MPBEC/program/src/Main. Then, the user has to make sure that the current folder in Matlab search toolbar is ~/MPBEC/program/src/Main before running MPBEC. The user also must check that the biological system satisfies the electroneutrality condition, among other pretests that the GUI does automatically. Finally, the user has to type MPBEC('inputfile.inm') on the Matlab commander window and press Enter to run the MPBEC one argument MATLAB function. MPBEC can be also called from other Matlab programs as usually done for any other Matlab function.

1. Input File structure (fine grid calculations)

The "inputfile.inm" file parsing is strict. Input data must contain the value of these parameters in the exact following column order (see example in Figure 12):

dime

glen

bulk

bc

solversetup

dielx_str

diely_str

dielz_str

kappa_str

charge_str

pqr_str

pqr_cent_str

energy

in_name_str

name_str

wout

Description of each arrow

dime This line is used to specify the number of grid points. The values for this keyword are: ***nx ny nz*** which represent the (integer) number of grid points in the x-, y-, and z-directions, respectively.

glen This line is used to specify the mesh domain lengths. The values for this keyword are: ***xlen ylen zlen*** which represent (floating point) grid lengths L_x, L_y and L_z (in Å) along the x-, y-, and z-directions, respectively.

bulk This line is used to specify bulk properties and parameters involved in the calculation of the dielectric and kappa maps. The values for this keyword are: ***Ionicstrength maxionR dielw dielp T srad surfdens swin.*** ***Ionicstrength*** defines the value of the ionic strength $I = \frac{1}{2}\sum c_i z_i^2$ where the sum is over all different ionic species "i", and z_i and c_i are the ionic valence and concentration (in moles), respectively. This quantity is used in the calculation of the Dirichlet boundary condition and the kappa map. **maxionR** is the maximum ionic radius (in Å) used in MOL and SMOL approaches to calculate the kappa

map. **dielw** is the value of the solvent dielectric coefficient involved in the calculation of the dielectric map. It is usually set equal to 78.54 for water in the solvated state and equal to the solute dielectric coefficient in the reference state. **dielp** is a number, usually between 2 and 10, which defines the value of the solute dielectric coefficient involved in the calculation of the dielectric map. **T** is the temperature of the system in Kelvin degrees. **srad** defines the molecular solvent radius in Å used in MOL and SMOL approaches to calculate the Collony molecular surface. It is usually set equal to 1.4 to model water. **surfdens** is usually set equal to 10 and represents the number of points per $Å^2$ used in MOL and SMOL approaches to calculate the Collony molecular surface. **swin** defines the value of the window width parameter (in Å) used in SPL2 approach to calculate the dielectric and kappa maps, usually set equal to 0.3.

bc This line is used to specify the boundary condition used in the calculation of the mean electrostatic potential. The words for this keyword are either **mdh** (if the user requires multiple Dirichlet boundary condition), **sdh** (if the user requires single Dirichlet boundary condition), or **focusname.inm** (if the user requires focus boundary condition).

solversetup This line is used to specify the linear solver setup. The words for this keyword are: **digpres selectsolver. digpres** is the number of digits of precision required on the solution of the linearized PB equation. This number is usually set equal to 6. **selectsolver** specifies the linear solver method. The words for this keyword are either **biconjgrad** (it uses Biconjugate Gradient solver stabilized by LU decomposition), **gmres** (it uses Generalized minimal residual solver stabilized by LU decomposition), or **minres** (it uses the minimal residual solver).

dielx_str This line is used to specify the file name of the shifted x-component of the dielectric coefficients in dx format (for instance **dielx.dx**). If this file is not uploaded, the user must set this keyword equal to **nouploadedfiles.**

diely_str Same description provided for **dielx_str** but in this case the file name refers to the shifted y-component of the dielectric.

dielz_str Same description provided for **dielx_str** but in this case the file name refers to the shifted z-component of the dielectric.

kappa_str This line is used to specify the file name of the ionic accessibility coefficients kappa in dx format (for instance **kappa.dx**). If this file is not uploaded, the user must set this keyword equal to **nouploadedfiles.**

charge_str This line is used to specify the file name in dx format of the charge density coefficients in units of $1/\text{Å}^3$ (for instance **charge.dx**). If this file is not uploaded, the user must set this keyword equal to **nouploadedfiles.**

pqr_str This line is used to specify the pqr file name of the molecular structure (for instance **moleculetarget.pqr**). This file determines the target molecule to solve the PB equation.

pqr_cent_str This line is used to specify the pqr file name of the molecule which contains the molecular structure that defines the center of grid. If there is only one molecule in the system, **pqr_cent_str** and **pqr_str** must contain the same keyword. In systems having two molecules **pqr_cent_str** specifies the pqr file name corresponding to the reference molecule. In such case the center of grid is defined by the coordinates of that molecule for all the molecules comprised by the complex system. An example of this file name is **moleculereference.pqr.**

dkm This line is used to specify the method used in the calculation of the dielectric and kappa maps. The words for this keyword are either **MOL** (if the user requires the Collony molecular surface approach), **SMOL** (if the user requires the harmonic smoothing approach of the Collony molecular surface), or **SPL2** (if the user requires cubic spline approach).

chm This line is used to specify the method in the calculation of the charge density map. The words for this keyword are either **SPL0** (if the user requires linear interpolating approach) or **SPL2** (if the user requires cubic interpolating approach).

energy This line is used to specify the calculation of the total electrostatic potential free energy given by Eq.(2). The words for this keyword are either **calceneryes** (if the user requires energy calculations), or **calcenerno** (if the user does not require energy calculations).

in_name_str This line is used to specify the full path to the input file directory containing all the files to be uploaded. An example of the input file directory is

C:\smith\Documents\MPBEC\program\src\examples\pqr

for Windows users and

/home/smith/Documents/MPBEC/program/src/examples/pqr

for Linux and Mac users.

name_str This line is used to specify the full path to the output file directory that contains the resulting output files. For instance, this directory may be

\MPBEC\program\src\examples\solutions

for Windows users and

/MPBEC/program/src/examples/solutions

for Linux and Mac users.

wout This line is used to specify the required output files. The values for this keyword are: **n1 n2 n3 n4 n5. n1** set equal to 1.0 writes out the charge map (dx file) whereas **n1** set equal to 0.0 does not. Similarly, **n2, n3, n4,** and **n5** are those parameters corresponding to the kappa map (dx file), dielectric maps (dx files), mean electrostatic potential solution (dx file), and electric potential surface plots, respectively.

2. Input File structure (coarse grid calculations using focusing)

For focusing calculations, the user has to generate two input files. One of them, the "inputfile.inm" file, specifies the parameters required to perform the calculation on the fine grid. Note that the keyword **focusname.inm** must be assigned in the **bc** line. The second input file, named "focusname.inm", provides the parameters corresponding to the focusing calculation using the same file structure described in the previous section for the "inputfile.inm" file. The files "inputfile.inm" and "focusname.inm" must be saved in the directory ~/MPBEC/program/src/Main. To run MPBEC on the Matlab commander window, the user has to type MPBEC('inputfile.inm') and press enter.

---

[1] S. Unni, Y. Huang, R. Hanson, M. Tobias, S. Krishnan, W.W. Li, J.E. Nielsen, N.A. Baker, Web servers and services for electrostatics calculations with APBS and PDB2PQR, J. Comput. Chem. 32 (2011) 1488–1491.

[2] G. Hämmerlin, K.H. Hoffmann, Numerical Mathematics, Springer, New York, 1991.

[3] M.L. Connolly, Solvent-accessible surfaces of proteins and nucleic acids, Science 221 (1983) 709-713.

[4] R.E. Bruccoleri, J. Novotny, M.E. Davis, K.A. Sharp, Finite difference Poisson-Boltzmann electrostatic calculations: Increased accuracy achieved by harmonic dielectric smoothing and charge antialiasing, J. Comput. Chem. 18 (1997) 268-276.

[5] W. Im, D. Beglov, B. Roux, Continuum solvation model: Computation of electrostatic forces

from numerical solutions to the Poisson-Boltzmann equation, Comp. Phys. Commun. 111 (1998) 59-75.

[6] F. Dong, B. Olsen, N.A. Baker, Computational methods for biomolecular electrostatics, Methods Cell. Biol. 84 (2008) 843-870.

[7] H. Li, A.D. Robertson,J.H. Jensen, Very Fast Empirical Prediction and Rationalization of Protein pKa Values, Proteins 61 (2005) 704-721.

[8] A. Greenbaum, Iterative Methods for Solving Linear Systems (Frontiers in Applied Mathematics), SIAM, 1987.

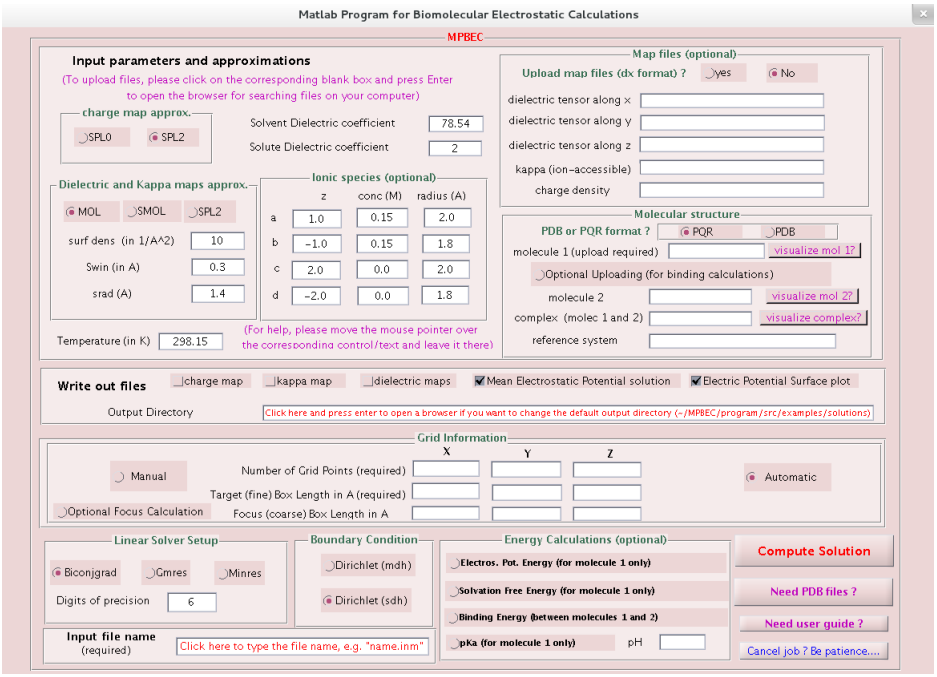

Figure 1: Main Screen

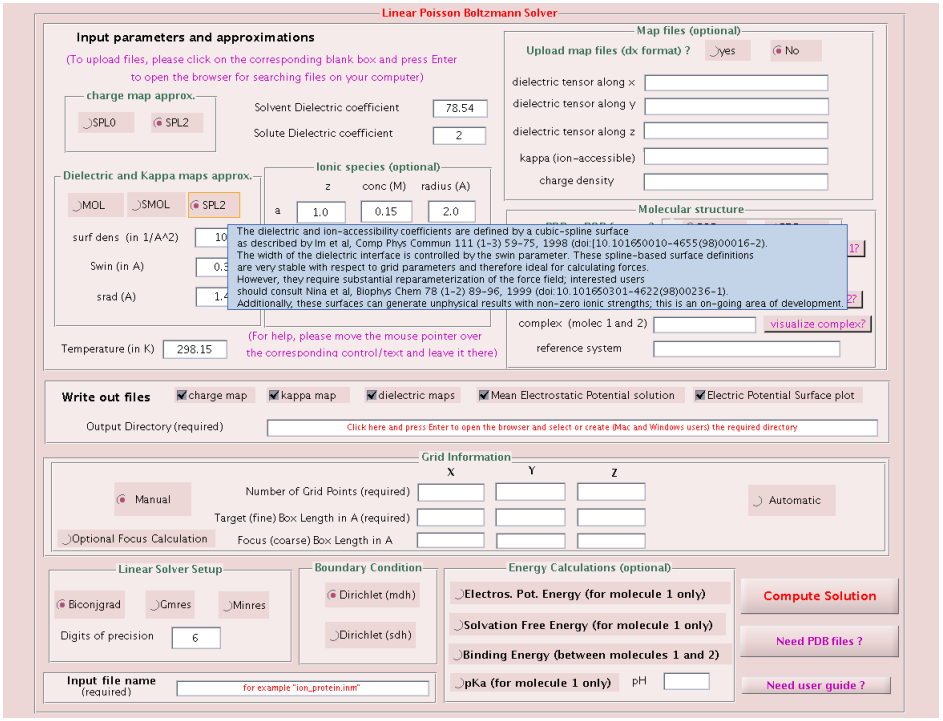

Figure 2: Getting help messages

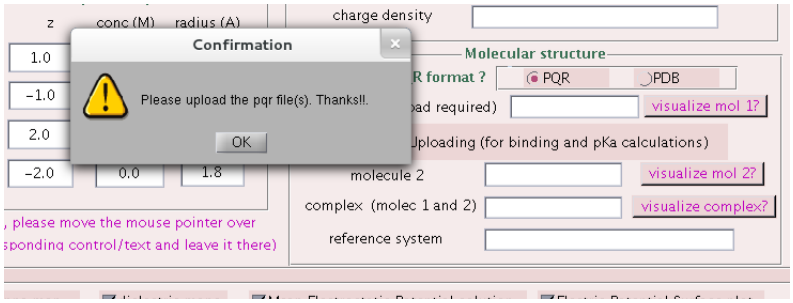

Figure 3: Getting warning messages generated by the pre-test algorithm when an error in the input data is detected

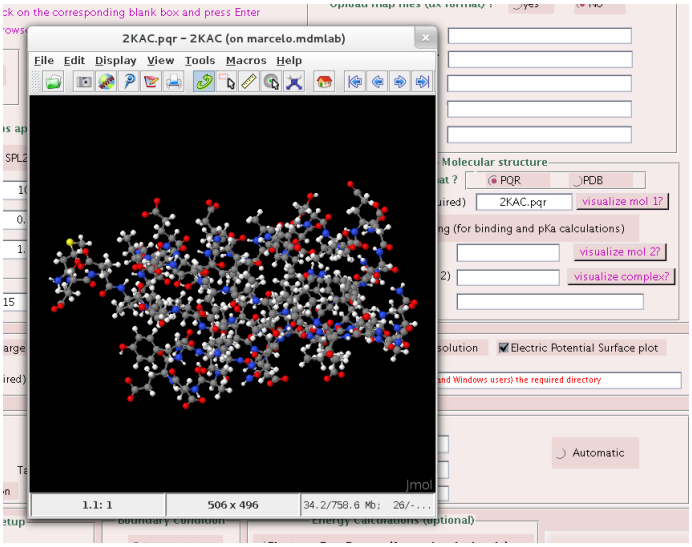

Figure 4: Using the integrated Jmol application to visualize molecular structures

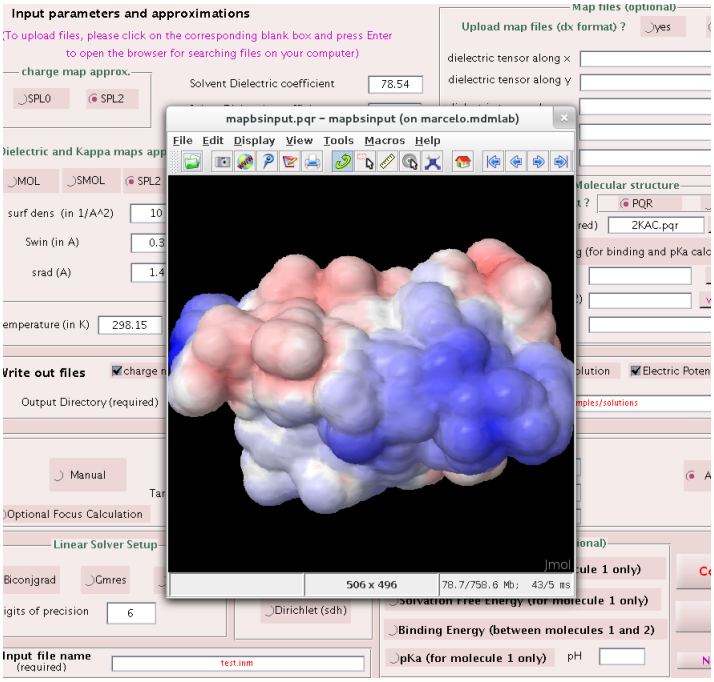

Figure 5: Using the integrated Jmol application to visualize the solution of the mean electrostatic potential on the biomolecule surface.

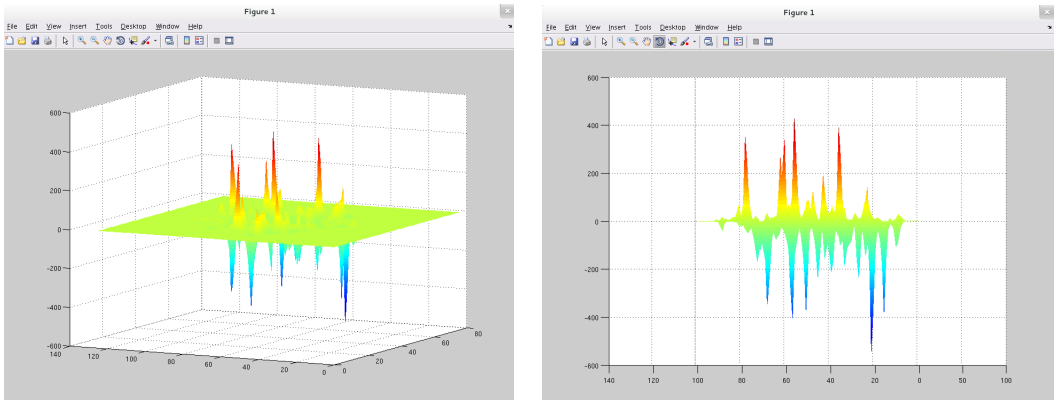

Figure 6: Using Matlab tools to visualize the asymptotic behavior of the solution of the mean electrostatic potential.

```
Processing charge data......
Done!

Generating the dielectric and ionic maps using SPL2 ....

molecular solvent radius (A) = 1.4,    Surface density (1/A^2) = 10    and Swin (A) = 0.3
Ionic Strength = 0.15,        Maximum Ionic Radius (A) = 2.5

APPROXIMATE SOLUTION FOR THE ELECTROSTATIC POTENTIAL

Constructing the sparse matrix A....
Done!....

Solving the linearized PB equation using the
Biconjugated gradient method stabilized by LU matrices

Performing the LU decomposition....
Done!....

Solving the linear matrix equation....
Done!


Residual error=  9.5468e-07,     Iteration number=  47

END OF THE CALCULATIONS

Elapsed time is 23.592578 seconds.

writing out the electrostatic potential solution
the file target_grid__MATLAB_pot.dx was generated

writing out the charge map
the file target_grid_MATLAB_rho.dx was generated

writing out the dielx map
the file target_grid_dielx.dx was generated

writing out the diely map
the file target_grid_diely.dx was generated

writing out the dielz map
the file target_grid_dielz.dx was generated

writing out the ionic accessible map
the file target_grid_kappa.dx was generated

Generating plots!....Please, close Jmol application to continue with the calculations
```

Figure 7: Matlab screen example when the code is running

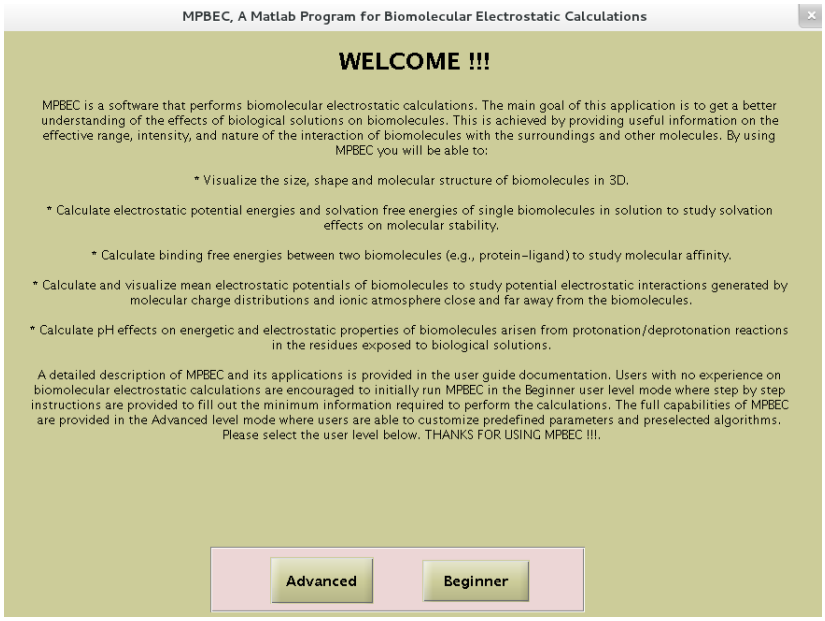

Figure 8: "Welcome" GUI

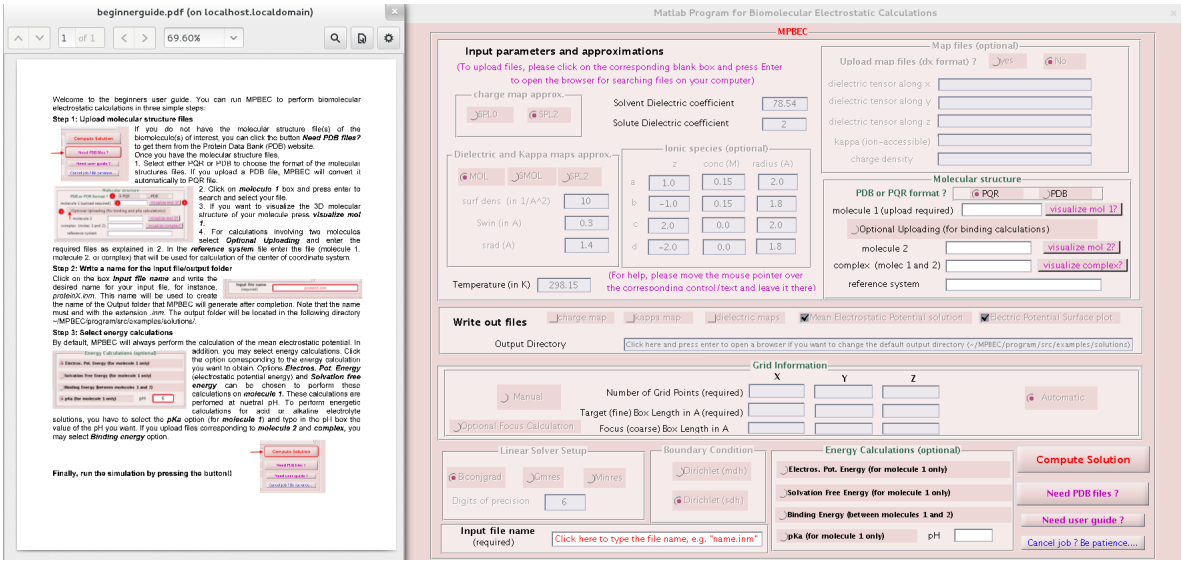

Figure 9: Beginner GUI and step by step guide document

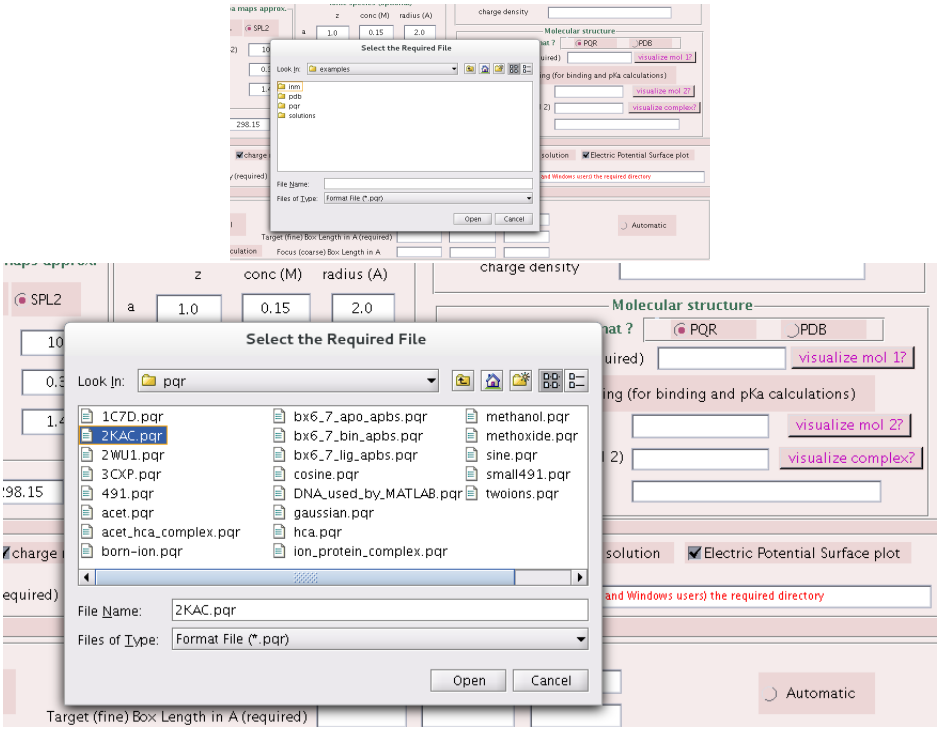

Figure 10: Input files uploading

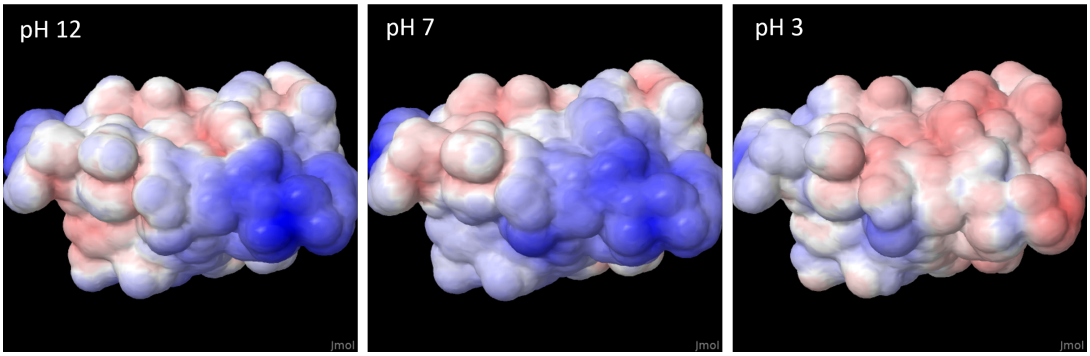

Figure 11: Mean electrostatic surface potential for different pH values

inputfile_example.inm

```
126 77 77
62.620500 38.037000 31.545000
0.150000 2.000000 78.540000 2.000000 298.150000 1.400000 10.000000 0.300000
mdh
6 biconjgrad
nouploadedfiles
nouploadedfiles
nouploadedfiles
nouploadedfiles
nouploadedfiles
2KAC.pqr
2KAC.pqr
MOL
SPL2
calcenerno
/home/username/MPBEC/program/src/examples/pqr/
/home/username/MPBEC/program/src/examples/solutions
0.000000 0.000000 0.000000 1.000000 1.000000
```

Figure 12: Input file example to run MPBEC without using the GUI.

Grid points x=y=z		MPBEC (kJ/mol)	APBS (kJ/mol)
97	Solvated energy	6387.46	6386.11
	Reference energy	6618.70	6618.78
	Solvation energy	-231.24	-232.67
129	Solvated energy	8595.29	8595.04
	Reference energy	8824.93	8825.04
	Solvation energy	-229.64	-230.00
161	Solvated energy	10801.5	10801.5
	Reference energy	11031.2	11031.3
	Solvation energy	-229.65	-229.78
193	Solvated energy	13007.7	13007.7
	Reference energy	13237.4	13237.6
	Solvation energy	-229.67	-229.87
257	Solvated energy	17420.6	17420.7
	Reference energy	17649.9	17650.1
	Solvation energy	-229.22	-229.42
353	Solvated energy	24039.7	24039.7
	Reference energy	24268.6	24268.9
	Solvation energy	-228.89	-229.16

Table I: Born ion model. Energy calculations for different number of grid points.

cgm	dkm		MPBEC (kJ/mol)	APBS (kJ/mol)
spl2	mol	Solvated energy	5352.48	5352.35
		Reference energy	5582.15	5582.22
		Solvation energy	-229.67	-229.87
	smol	Solvated energy	5353.38	5353.13
		Reference energy	5582.15	5582.22
		Solvation energy	-228.77	-229.09
	spl2	Solvated energy	5333.17	5333.23
		Reference energy	5582.15	5582.22
		Solvation energy	-248.98	-248.98
spl0	mol	Solvated energy	13007.7	13007.7
		Reference energy	13237.4	13237.6
		Solvation energy	-229.67	-229.87
	smol	Solvated energy	13008.6	12988.6
		Reference energy	13237.4	13237.6
		Solvation energy	-228.77	-229.09
	spl2	Solvated energy	12988.4	12988.6
		Reference energy	13237.4	13237.6
		Solvation energy	-248.98	-248.98

Table II: Born ion model. Energy calculations for different options of charge map (chm) and dielectric and kappa maps (dkm).

Ionic strength	chm		MPBEC (kJ/mol)	APBS (kJ/mol)
0 M	spl0	Solvated energy	2681.14	2681.17
		Reference energy	2706.61	2706.64
		Solvation energy	-25.47	-25.47
	spl2	Solvated energy	1648.74	1648.76
		Reference energy	1674.05	1674.07
		Solvation energy	-25.31	-25.31
0.05 M	spl0	Solvated energy	2681.14	2681.17
		Reference energy	2706.37	2706.29
		Solvation energy	-25.23	-25.12
	spl2	Solvated energy	1648.74	1648.76
		Reference energy	1673.81	1673.72
		Solvation energy	-25.07	-24.96
0.5 M	spl0	Solvated energy	2681.14	2681.16
		Reference energy	2705.93	2705.83
		Solvation energy	-24.80	-24.67
	spl2	Solvated energy	1648.74	1648.75
		Reference energy	1673.37	1673.25
		Solvation energy	-24.63	-24.50

Table III: Methanol model. Energy calculations for different charge map (chm) options and electrolyte concentrations.

Ionic strength	chm		MPBEC (kJ/mol)	APBS (kJ/mol)
0 M	spl0	Solvated energy	5633.92	5633.99
		Reference energy	5834.53	5834.60
		Solvation energy	-200.61	-200.61
	spl2	Solvated energy	2485.68	2485.71
		Reference energy	2686.32	2686.35
		Solvation energy	-200.64	-200.64
0.05M	spl0	Solvated energy	5633.56	5633.48
		Reference energy	5784.86	5777.48
		Solvation energy	-151.30	-144.00
	spl2	Solvated energy	2485.32	2485.20
		Reference energy	2636.65	2629.23
		Solvation energy	-151.33	-144.03
0.5 M	spl0	Solvated energy	5633.42	5632.90
		Reference energy	5762.95	5758.42
		Solvation energy	-129.53	-125.53
	spl2	Solvated energy	2485.18	2484.62
		Reference energy	2614.75	2610.18
		Solvation energy	-129.56	-125.56

Table IV: Methoxide (Ionized Methanol) model. Energy calculations for different charge map (chm) options and electrolyte concentrations.

Molecule		Grid	MPBEC (kJ/mol)	APBS (kJ/mol)
Hca	Electrostatic potential energy	Coarse	127649.66	127651.22
		Fine	298150.07	298150.51
Acetazolamine		Coarse	1147.24	1147.25
		Fine	2660.16	2660.19
Hca-Acetazolamine		Coarse	128787.93	128789.56
		Fine	300749.95	300754.49
	Binding energy		-60.29	-60.22

Table V: Hca-Acetazolamine complex model. Binding energy calculation using focusing boundary condition

Molecule		Grid	MPBEC (kJ/mol)	APBS (kJ/mol)
1D7H	Electrostatic potential energy	Coarse	70246.2034	70246.8009
		Fine	166946.98	166948.75
DMSO		Coarse	145.8681	145.8687
		Fine	310.5369	310.5395
1D7H-DMSO		Coarse	70394.73	70395.33
		Fine	167258.31	167260.10
	Binding energy		0.7968	0.7997

Table VI: 1d7h - DMSO complex model. Binding energy calculation using focusing boundary conditions.

PDB ID	Number of atoms	Estimated requiered memory (Mbs)	total computing time (sec)	Grid points	Grid length ($\AA$)	MPBEC			APBS
						$\frac{Solvated}{energy}$	$\frac{Reference}{energy}$	$\frac{Solvation}{energy}$	$\frac{Solvation}{energy}$
2MQW	2666	2039	249	166 179 179	83 89 66	184304	189127	-4823.35	-4724.75
2MRI	2878	1452	215	139 165 165	69 82 70	173961	177418	-3457.49	-3386.81
2MY8	1529	829	102	99 148 148	49 74 52	101809	104199	-2390.09	-2370.99
3CXP	3246	1272	206	145 149 149	72 74 75	163049	169131	-6082.29	-6074.04
4OZ5	2644	947	155	132 137 137	66 68 60	178314	182662	-4347.57	-4266.00
2KAC	931	287	50	126 77 77	63 38 32	67613.7	70213.2	-2599.51	-2483.57

Table VII: Protein solvation free energy (kJ/mol) calculations using the preselected parameters and the automatic option to configure the grid information.