QGUI

user manual

A Graphical User Interface for Q

Geir Villy Isaksen
Tor Arne Heim Andberg
Davide Michetti
Bjørn Olav Brandsdal

UiT - The Arctic University of Norway

Contents

1 Introduction **2**

2 Installation **2**
- 2.1 System requirements . . . 2
- 2.2 Linux and Mac . . . 3
- 2.3 Windows . . . 4

3 Overview **5**

4 Settings **6**
- 4.1 Equilibration files setup . . . 7
- 4.2 Submission script and submission command . . . 7
- 4.3 Force Fields . . . 9

5 Prepare **10**
- 5.1 PDB files . . . 10
- 5.2 Missing parameters . . . 11
- 5.3 Topology . . . 13
- 5.4 Trjmask . . . 14

6 Molecular dynamics simulations **15**
- 6.1 MD analysis . . . 16
  - 6.1.1 Energies . . . 17
  - 6.1.2 Root Mean Square Fluctuation (RMSF), Root Mean Square Deviation (RMSD), Distance, Angle and Torsion 18

7 LIE calculations **20**
- 7.1 LIE tutorial . . . 23

8 FEP calculations **27**
- 8.1 FEP tutorial . . . 30

9 EVB simulations **30**
- 9.1 EVB - Tutorial . . . 31
  - 9.1.1 EVB - Setup . . . 31
  - 9.1.2 EVB - Analysis . . . 33
  - 9.1.3 EVB - Reference reaction . . . 34

9.1.4 EVB - Reaction energy 35
9.1.5 EVB - Arrhenius plot 37

1 Introduction

Qgui is a high-throughput interface for automated setup and analysis of free energy calculations and empirical valence bond (EVB) simulations. [1] The interface is developed to work with the molecular dynamics (MD) package Q (`http://xray.bmc.uu.se/~aqwww/q/`).[2] Qgui automates various tasks such as parameter assignment, topology generation, EVB reference reaction calibration, analysis of extensive free energy simulations etc.

2 Installation

Qgui is written in the python language (`https://www.python.org`) using object oriented design principles based on the Tkinter GUI library (`https://docs.python.org/2/library/tkinter.html`) In the following we will describe how to get Qgui up and running on your computer system.

2.1 System requirements

Required software and packages:

- The MD package Q (Qprep, Qdyn, Qfep and Qcalc)
- Python 2.x (preferably 2.7)
- NumPy (Base N-dimensional array package)
- Matplotlib (Comprehensive 2D Plotting)

If you do not already have **Q**, you can get it from:

`http://xray.bmc.uu.se/ aqwww/q/`

The easiest way to obtain Python 2.7 with the required packages used by Qgui is to install **Anaconda** which is the leading open data science platform powered by Python:

`https://www.continuum.io/downloads/`

[1] Isaksen GV et al. (2015) *J Mol Graph Mod* 10.1016/j.jmgm.2015.05.007
[2] Marelius J et al. (1998) *J Mol Graph Mod* Aug-Dec;16(4-6):213-25, 26

NOTE: if you install Anaconda, you do not need to worry about Numpy and Matplotlib since they are included in this package.

Optional software:

- PyMOL (visualization of structures during setup and analysis)
- Maestro free academic version (Assigning force field parameters)

Qgui communicates with PyMOL which can be used for visualizing MD trajectories and efficient definition of changing atoms and bonds in the FEP and EVB setups (see sections 8 and 9, respectively). **PyMOL** can be installed free of charge from:

`http://sourceforge.net/projects/pymol/`

The software can also communicate with the ffld_server utility that comes with Maestro (Schrödinger, LLC, New York) for automatic assignment of OPLS force field parameters. This is in general useful if you have molecules with missing parameters and for auto-assigning state dependent parameters in FEP and EVB. The **Maestro** academic version is free of charge and can be downloaded from:

`http://www.schrodinger.com/freemaestro/`

2.2 Linux and Mac

Extract the downloaded Qgui_vxyz.zip and go to the resulting `Qgui_mac_linux` directory. Here you will find the installation script `INSTALL.py`. Run the script by typing

```
python INSTALL.py
```

in your terminal window. The default directory for the installation is

- LINUX: `/home/$USER/QGUI`
- MAC: `/Users/$USER/QGUI`

This can easily be changed to any other directory by editing the `install_path` variable in the beginning of `INSTALL.py`. Make sure that you write a valid absolute path. Also, the default system variable file is

`.bash_profile`. This can also be edited in the installation file by changing the `bashfile` variable.

After successfully installing QGUI, you should restart or open a new terminal window. Now, the software can be launched from anywhere by typing `Qgui` in the terminal window.

2.3 Windows

We have never tried to use QGUI on Windows machines, but the principle should be similar as for Linux/OS as described above. We currently do not have any installation scripts prepared for windows, so the system variables must be set up manually.

3 Overview

From the Qgui main window (Fig 1) one can access all the various parts for setting up and analyzing MD/LIE/FEP/EVB simulations. The currently loaded structure (*.pdb) and topology (*.top) are also visible and updated in the main window.

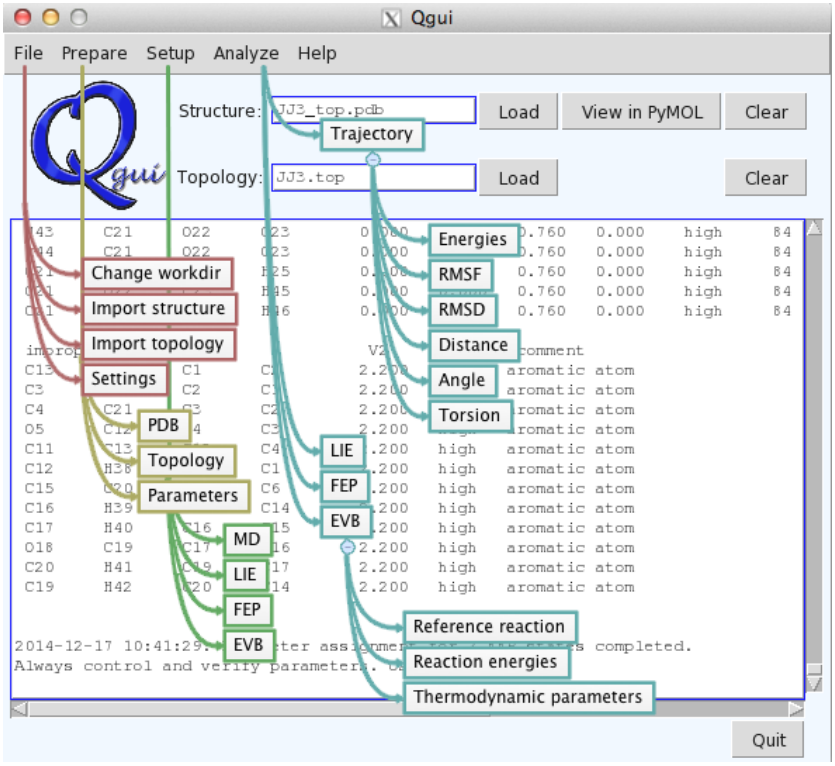

Figure 1: The main window of Qgui with illustrated overview of dropdown menus and their content.

The main window itself functions primarily as a logging and feedback window for the various parts included in the Qgui package. These are easily accessed through the five drop-down menus illustrated in Fig 1.

4 Settings

The Global settings can be accessed by clicking on **File** and then **Settings** (Fig 2) .

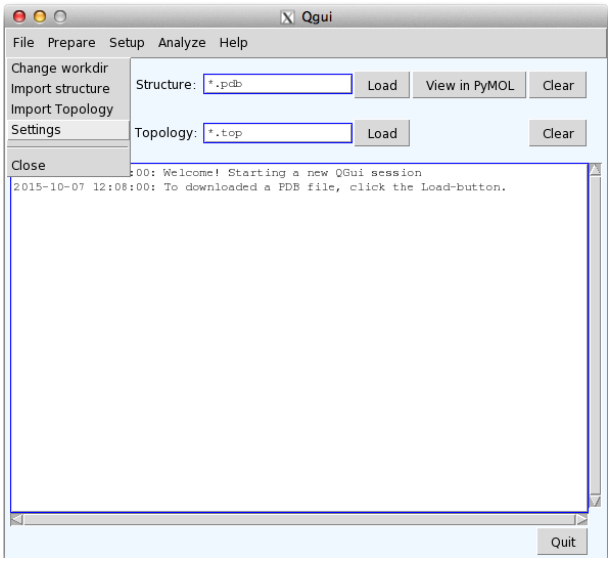

Figure 2: The main window of Qgui showing the location of **Settings** option.

In this window you can define (Fig 3):

- the parameter and library files required to generate topologies for the various simulations;
- the names of the installed **Q** executables (Qprep5, Qdyn5, Qfep5 and Qcalc5) included in the users path;
- the default equilibration procedure and corresponding settings to be run prior to all MD simulations;
- the **$SCHRODINGER** path (optional and only required for automatic parameter generation);

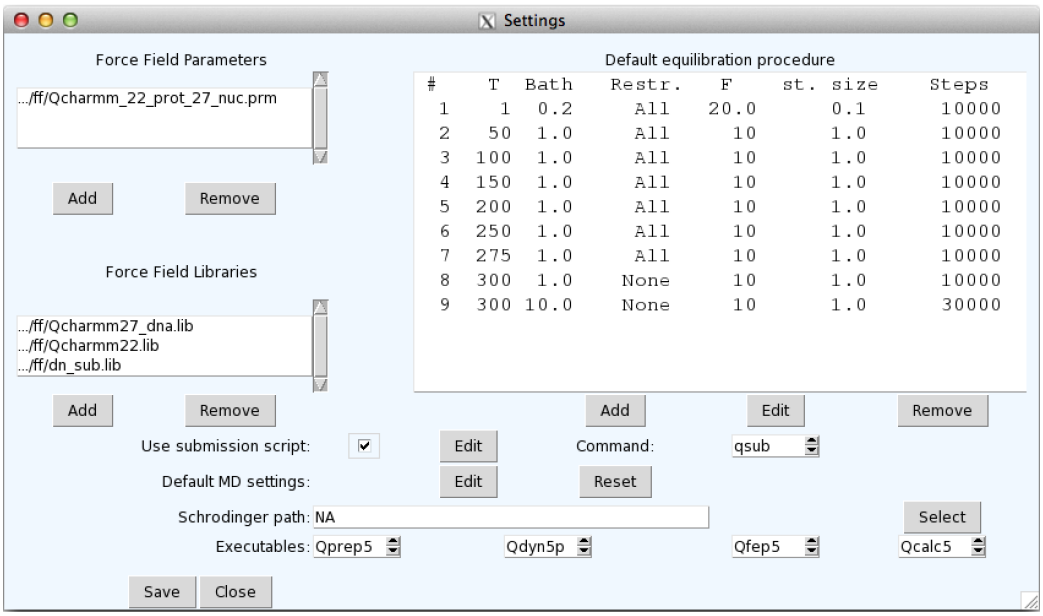

Figure 3: The Global settings window.

4.1 Equilibration files setup

In the Default equilibration procedure you can define the number of files/steps needed for all calculations (Fig 3). In particular for each file you can change the Temperature, bath coupling, restraints (None, All system, Solute or Solvent) with the related force constant (F column, Kcal/mol), step size (femtosecond) and number of steps.

4.2 Submission script and submission command

Simulations are most often run on computer clusters that typically involves a queuing system with a special command to submit jobs to the queue. These features can easily be prepared in the Qgui settings (Fig 3).

The submission command is the keyword used to submit jobs and is dependent on the computer cluster settings. If there are no special commands for running jobs on the system you are running jobs on, just set the command to `./` or bash.

The submission script allows the user to define routine tasks that are executed before and after the actual Q jobs are run/submitted. Once the "`Use submission script`" is selected in the settings window (Fig 3) the submission script will automatically be launched when pressing "`run`" in MD/LIE/FEP/EVB. This procedure is generally recommended compared to running jobs interactively.

The submission script file can be edited directly by pressing the "edit" button in the settings window (Fig 3). The only required line in this file is the "`#Qdyn I/O`", which is the point where the actual commands to run the relevant Q jobs will be written. An example of a submission script is shown in the box below.

```
#!/bin/bash

#PBS -lnodes=1:ppn=20

#PBS -lwalltime=08:00:00
# Expecting to run for 8 hours.

#PBS -lpmem=400MB
#PBS -A 1234xx
#PBS -m ae
##PBS -q express
workdir=/global/work/$USER/$PBS_JOBID
mkdir -p $workdir

executable=/home/$USER/bin/Qdyn5p
inputfiles=$PBS_O_WORKDIR/*.inp
outputfiles=$workdir/*.log
trj=$workdir/*.dcd
ene=$workdir/*.en
fep=$PBS_O_WORKDIR/*.fep
top=$PBS_O_WORKDIR/*.top

cd $PBS_O_WORKDIR
cp $executable $inputfiles $fep $top $workdir
cd $workdir

#Qdyn I/O

cp $outputfiles $re $ene $trj $PBS_O_WORKDIR

cd
rm -rf $workdir
```

4.3 Force Fields

The force field (FF) parameter (.prm) and library (.lib) files are the same as those originally defined for Q. However, there is one new feature available for

the library files that allows Qgui to toggle between for example neutral and charged versions of an amino acid in the Topology prepare tool (see section 5.3). The toggle definitions need to be in the top of the lib file and can in principle be defined to switch between any similar amino acid or molecule, as long as they both are defined in the lib file. This section is recognized by Qgui when the *[toggle_residues] is read:

```
*[toggle_residues] ! This is for topology prepare in Qgui
        1          2    3      4    5            6
*   QGUI_TOGGLE  ARG  ARN   -HH12           radius_to CZ
*   QGUI_TOGGLE  ARN  ARG   +HH12  NH1      radius_to CZ
*   QGUI_TOGGLE  HIP  HID   -HE2            radius_to CE1
*   ...
```

Here, every `QGUI_TOGGLE` (column 1) initiates the definition of a new toggle residue where columns 2 and 3 are the original and the new residue names, respectively. Column 4 defines what atom to be removed (-) or added (+). Whenever an atom is to be added, the next column will define what atom to bond to (column 5 above). There can be as many adds/deletes as the user wants here as long as the pattern above is followed. The last column (column 6 in the above example) defines what atom to measure the simulations sphere distance to (`radius_to`).

5 Prepare

5.1 PDB files

The prepare PDB tool allows the user edit the PDB-file and delete certain atoms or residues (figure 4).

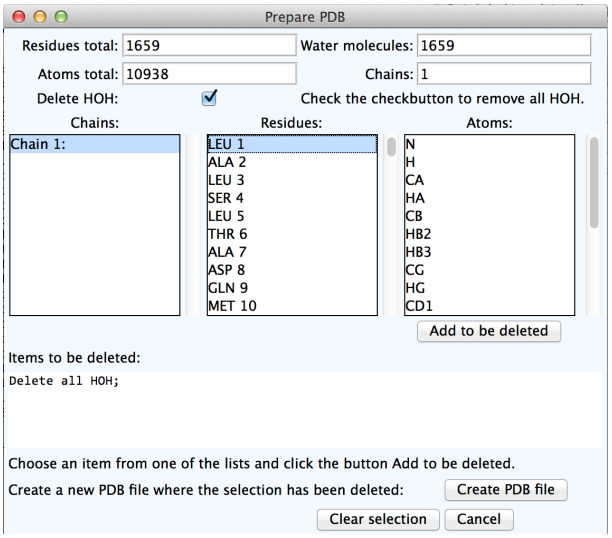

Figure 4: The prepare PDB window.

Selecting a chain will list all residues belonging to the chain and selecting a residue will list all atoms in the residue. Both residues and single atoms can be selected for deletion. To delete all water molecules at once you can check the delete HOH box. To commit selected changes click the "Create PDB" button and a new PDB file will be created.

5.2 Missing parameters

Qgui feature automatic parameter generation via the Schrodinger ffld server using the OPLS force-field. The parameter generator tool is located under Prepare → Parameters. Remember to always check automatically generated parameters manually.

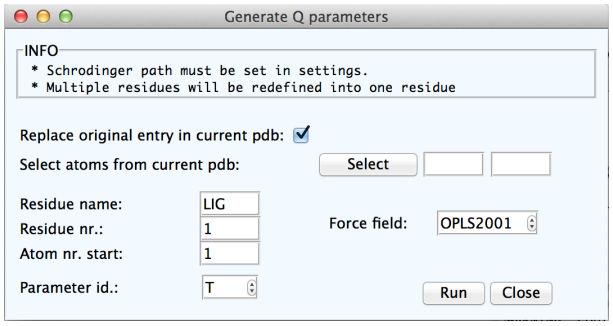

Figure 5: The parameter generator window.

Once you select the tool you will see the paramter generator window (Fig 5). Here you can enter data manually or select the atoms to parameterize directly from the pdb-file using the "Select" button (Fig 6).

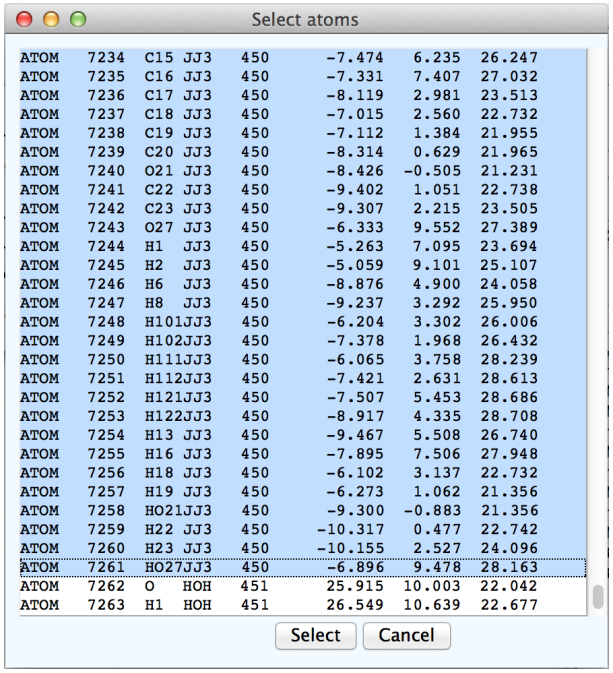

Figure 6: The select atoms to parameterize window.

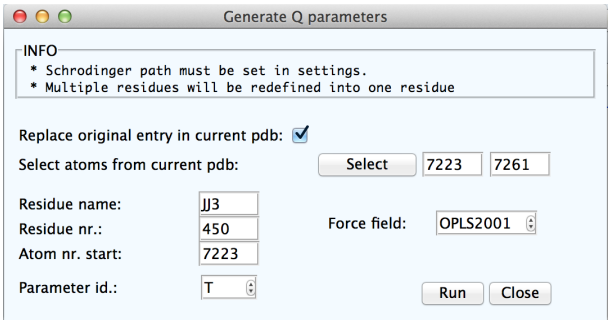

Figure 7: Completed parameter generator setup.

After selecting the atoms you will see a window similar to figure 7. Finally you have the option to select either OPLS2001 or OPLS2005 parameters, then choose "Run" to generate the parameters. Remember to add the newly generated files in File → Settings.

Replace original entry in current pdb will change the residue name of the selected atoms to the specified residue name. For example if a ligand consists of two amino acids this will merge them to a single residue with a single name and rename the atoms.

Parameter id, by default T, sets the first letter of the atom names in the automatically generated parameters. Thus a CA atom would be named TCA and so on. This is to prevent errors from overlapping atom names in the force fields.

Sometimes pdb-files do not include sufficient information for the correct interpretation of bond orders and hybridization in certain molecules. To read an .mae-file place an .mae file with the same name as the ligand in the work folder. In the example case the ligand was named JJ3, if a file named JJ3.mae was present Qgui would read this file instead and any information in the pdb-file would be disregarded for the purpose of generating the parameters.

5.3 Topology

To create a topology you must first load a pdb file and then select the topology generation tool, Prepare → Topology. Make sure you have force fields properly set up, see section 4.3. If parameters are missing check section 5.2 for details on how to create parameters.

Simulation sphere: Sets the radius for the simulation sphere. Atoms outside the sphere will typically be restrained and chargeable amino acids

are usually neutral.
PyMOL: Opens PyMOL.
Simulation center: Defines the coordinates of the simulation center.
Change: User selects atom to be simulation center.
Center: Changes the simulation center to the systems center of mass.
Solvate: User chooses solvation model, either TIP3P, SPC or None.
Create S-S bond: Qgui will automatically identify disulfide bridges based on distance.
Total charge: Displays the total charge of the system.
Toggle all charges: Turns all chargeable amino acids within the simulation sphere to be charged/uncharged.
Toggle state: Changes state of a single residue, for example HID-HIE-HIP.
Neutralize system with NaCl: Adds sodium or chloride ions to neutralize the system.
Edit: Opens a window for editing the pdb-file. Note that changes made in this window is not stored in the actual pdb-file, but rather is stored temporarily and used when the topology is created.
Check: Runs a check on the integrity of the topology, listing errors and missing data.
Topology name: The name of the topology file.
Topology PDB: The name of the pdb-file matching the topology.
Write: Writes the Q input file for creating the topology.
Run: Creates the topology.

5.4 Trjmask

In certain simulations the trajectory does not include every atom in the topology. The Trjmask tool creates a pdb-file with the same atom mask, and precisely matches the atoms in the trajectory.

To create the Trjmask pdb select one topology file and one trajectory file (fig. 8), and click create. A file dialog then appears, choose the name you want for the pdb-file and save. This closes the Trjmask tool and automatically loads the topology and pdb-file in the main Qgui window.

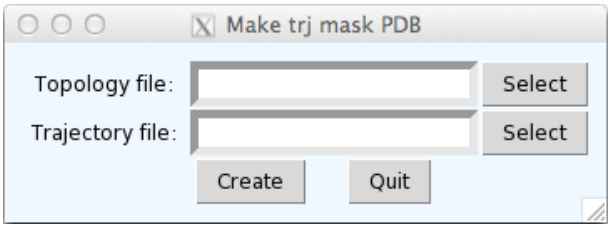

Figure 8: Trjmask window.

6 Molecular dynamics simulations

In order to setup a Molecular Dynamics simulation you should first load/prepare a topology to the corresponding pdb file as described in section 5.1. The MD preparation window can be accessed from the Setup drop-down menu (Fig 9).

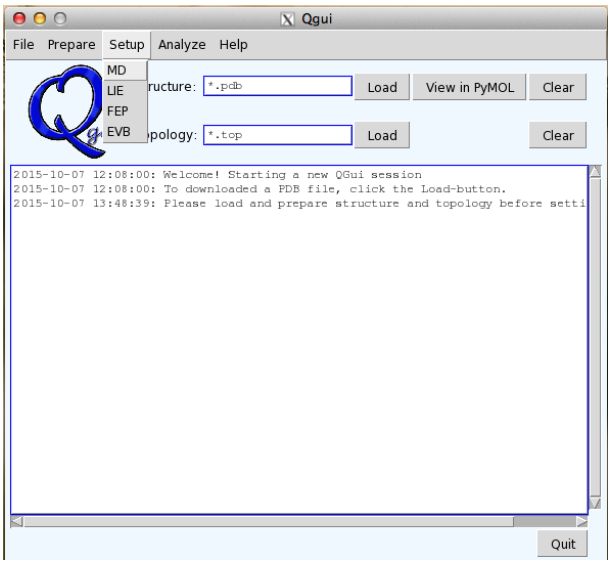

Figure 9: The main window showing the location of the MD setup.

As we can see in Fig 10, in the red square first you need to decide in how many .dcd files you want to split the productive simulation into. For long and heavy trajectories it is advisable to split the output in more than one file. By changing the value in **simulation time** you decide the length

in **nanosecond** of each file. In this section you can also define temperature, bath coupling, shake algorithm and lrf Taylor expansion. In the blue square the inner shall radius is automatically taken from the topology file and the default force constant is usually kept to 10 (kcal/mol). In the green square you can setup the frequency of output creation for trajectory, energy calculation, energy file and for updating the non-bonded list for interaction cutoffs. You can also decide how many atoms will be written in the trajectory file. In the violet square you can setup the different kind of restraints that **Q** offer: distance, atom, position and wall. Finally by pressing write the input file for the MD simulation will be written and you can automatically launch the job by pressing run.

Up to date only simulations in the water droplet model can be setup and run.

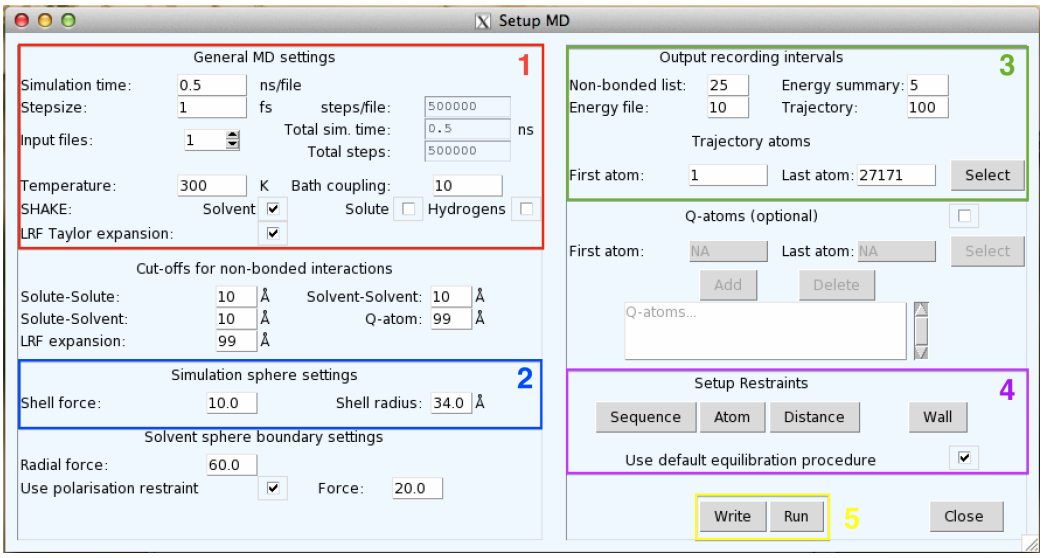

Figure 10: MD setup window;

6.1 MD analysis

In the analysis section there are six possible calculation options (Fig 11):

- Energies,
- Root Mean Square Fluctuation (RMSF),

- Root Mean Square Deviation (RMSD),
- Distance,
- Angle,
- Torsion,

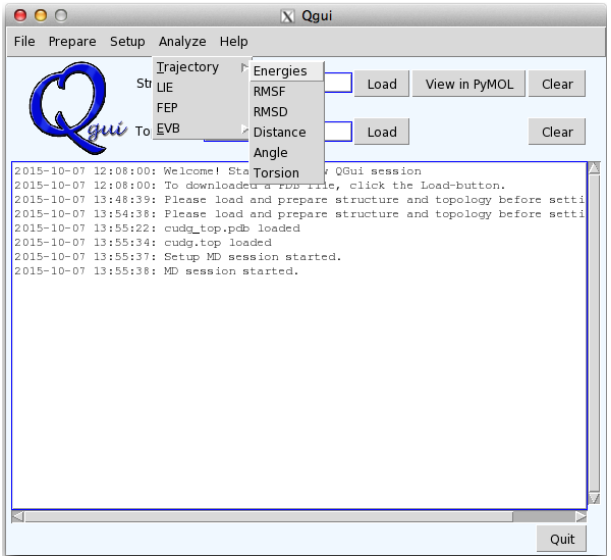

Figure 11: The main window showing the location of the MD analysis.

6.1.1 Energies

In the energies window in the **Title section** you give the title to the energy plots and you can add the trajectory files by clicking "+" in the Runs section. You can calculate the following type of energies (Fig 12):

- Energy of interaction Solute-Solute,
- Energy of interaction Solute-Solvent,
- Energy of interaction Solvent-Solvent,
- Restraints,
- Local Reaction Field,

- Total energy,

By choosing plot you can create a time-dependent plotting of the different energies.

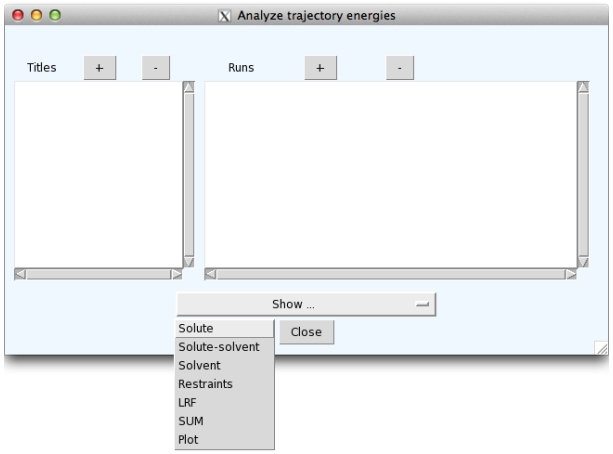

Figure 12: The MD energies analysis.

6.1.2 Root Mean Square Fluctuation (RMSF), Root Mean Square Deviation (RMSD), Distance, Angle and Torsion

For all these analysis the input (Fig 13) and output (Fig 14) windows look the same. In the first window you can select the atom that you are interested into and there are also buttons for quick selection of: backbone (bb), sidechains (sc) and solute atoms.

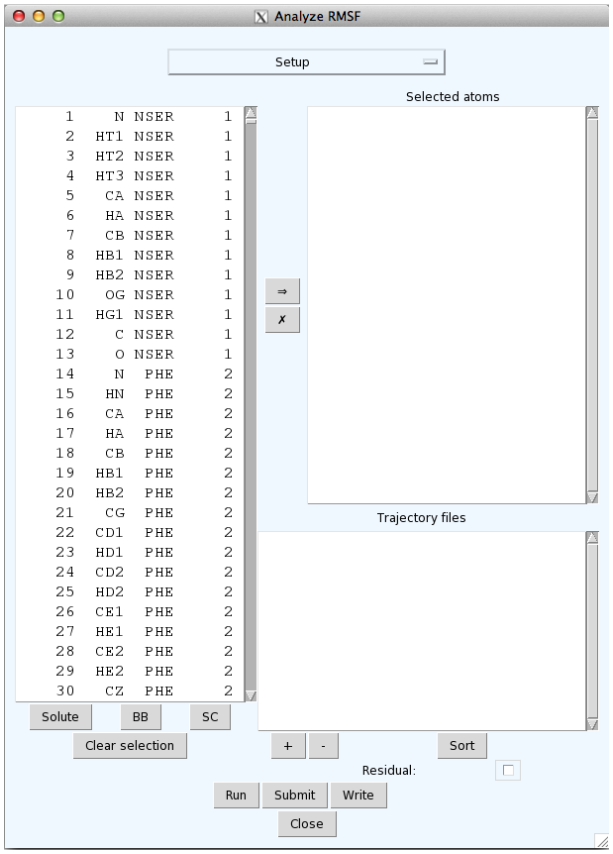

Figure 13: RMSF and RMSD main window.

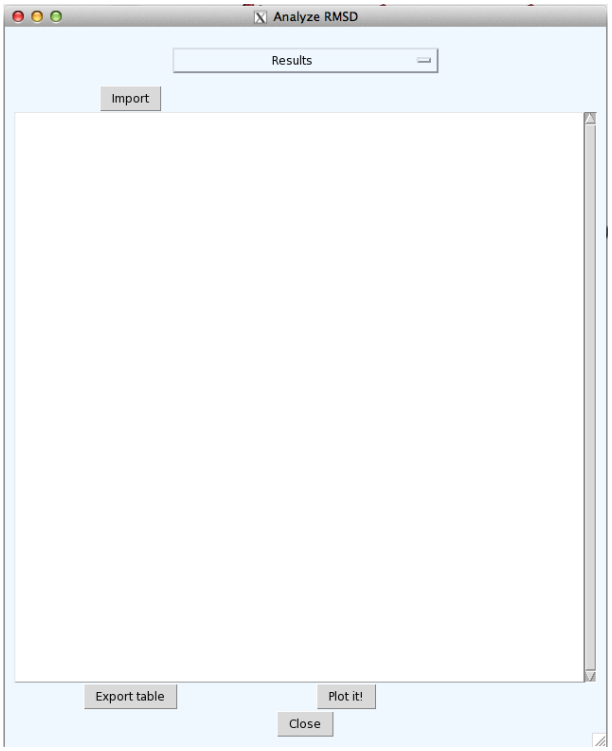

Figure 14: RMSF and RMSD output window.

7 LIE calculations

The linear interaction energy (LIE) method calculates the binding free energy of a ligand bound to a macro molecule according to the LIE equation:

$$\Delta G_{bind} = \alpha \Delta \langle U_{l-s}^{vdw} \rangle + \beta \Delta \langle V_{l-s}^{el} \rangle + \gamma \tag{1}$$

The energy contribution is divided into one van der Waals term scaled by an empirical parameter α and one electrostatic term scaled by a theoretical parameter β. To reproduce absolute binding energies a third paramater, γ, is needed (ADD REF).

Setup LIE tool is located at Setup $\rightarrow$ LIE (Fig 15). A LIE simulation consist of two separate MD simulations, one of the ligand in complex with the macro molecule and one of the ligand free in the solvent. Both can be set up using the Setup LIE tool.

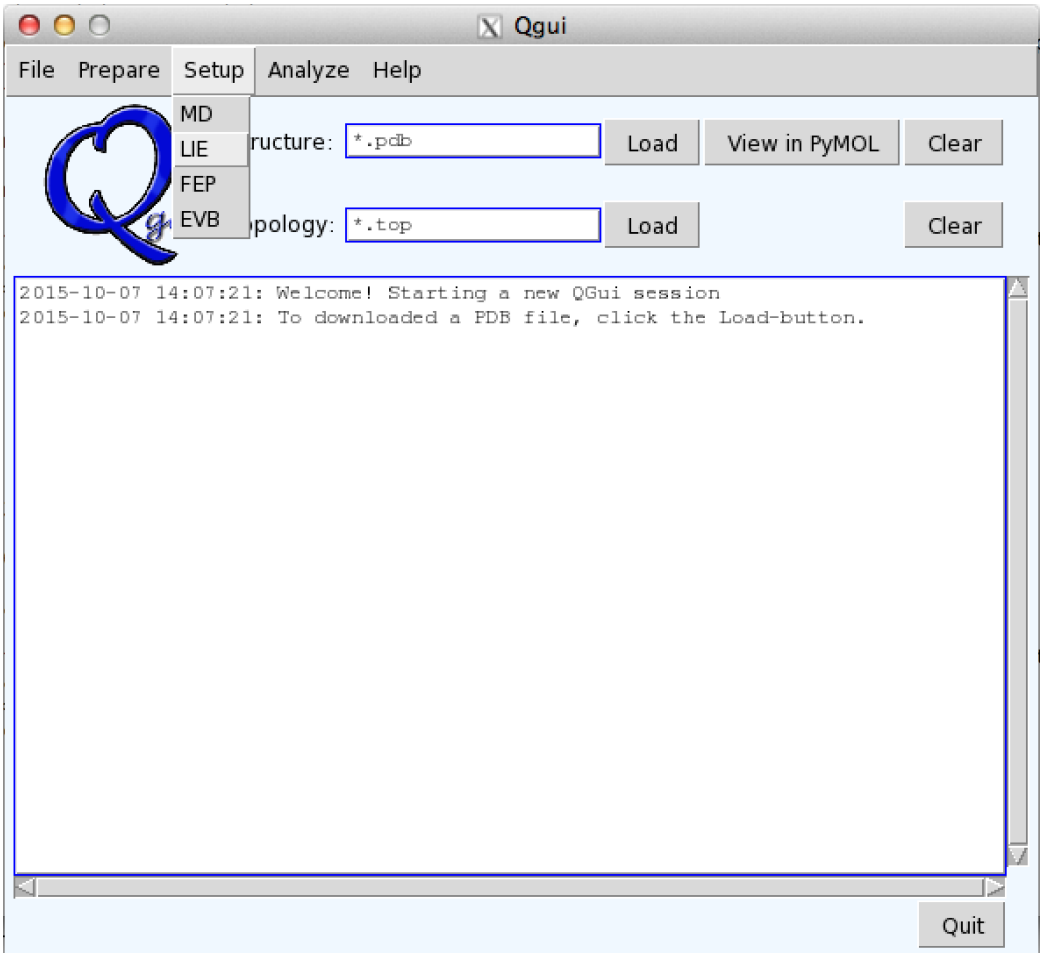

Figure 15: The main window showing the location of the Setup LIE tool

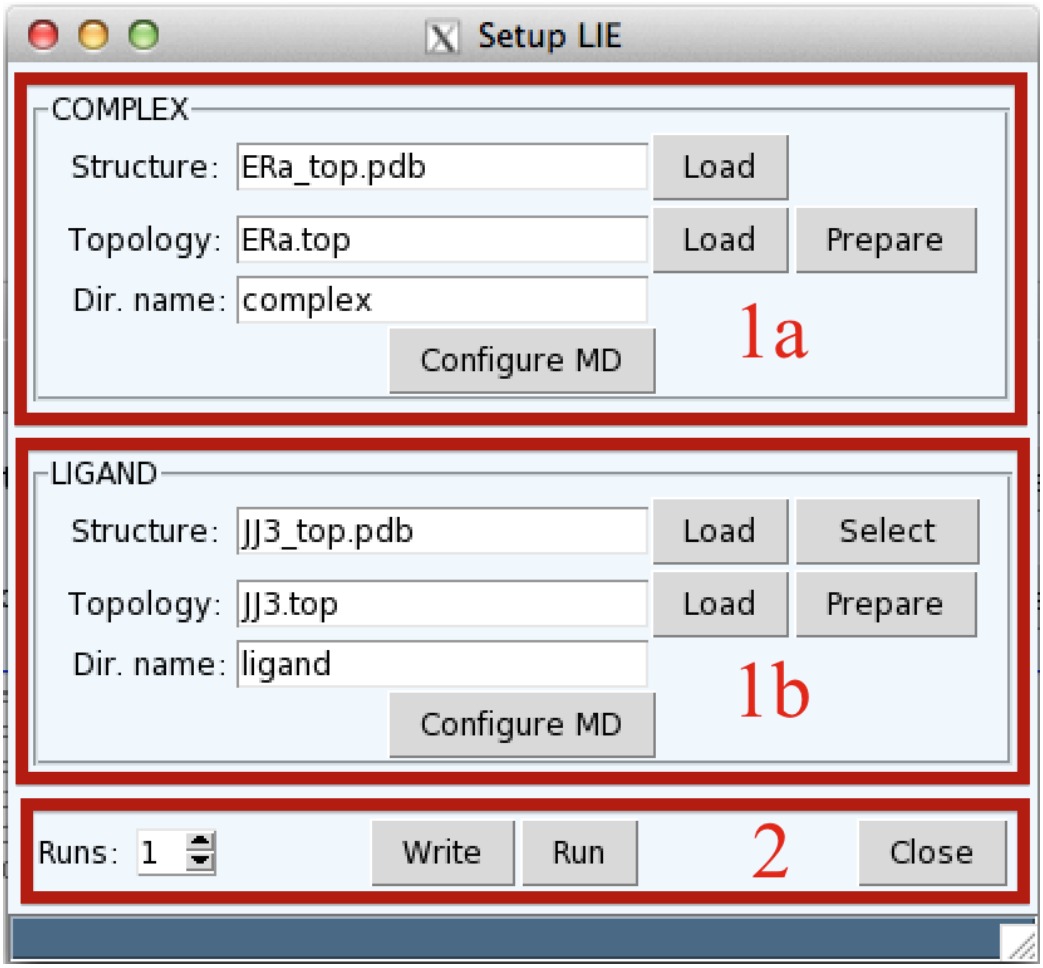

Figure 16: Setup LIE menu. 1a: Set up for complex simulation. 1b: Set up for ligand simulation. 2: Number of runs box. Write settings, run simulations and close buttons.

The Setup LIE window consist of two separate simulation setups for complex and ligand (fig 16), labelled 1a and 1b in the figure. Both setups are similar requiring a pdb file, a topology file, directory name and has a configure MD button.
Structure → Load: Loads pdb file.
Structure → Select (ligand only): The user can choose atoms from the complex pdb file and define them as the ligand. Qgui will then

automatically create a new pdb file from the selected atoms and load it.
Topology → Load: Loads topology (.top) file.
Topology → Prepare: Creates a topology file, see section 5.3 for details on how to create a topology.
Dir. name: Choose the names for the complex and ligand simulation directories. The default names are complex and ligand.
Configure MD: LIE simulations require the user to define Q-atoms (fig 19). This tells the MD program for which atoms to sample the ligand surrounding energies that are necessary for calculating the LIE binding energy. See section 6 for details.
Tip: For ligand simulations it is useful to add a restraint on the ligand to prevent it from drifting too close to the simulation boundary.
Runs: Select the number of parallel MD simulations.
Write: Creates the folders and writes the MD settings files.
Run: Submits the MD simulation jobs.
Close: Closes the Setup LIE window.

For a step by step guide on how to set up a LIE simulation please look at the LIE tutorial (section 7.1).

7.1 LIE tutorial

To set up the simulation the user starts by loading the pdb and topology file for the complex (Fig 17). If the topology has not been made yet it is possible to create it now using the prepare button. For information on how to create a topology see section 5.3.

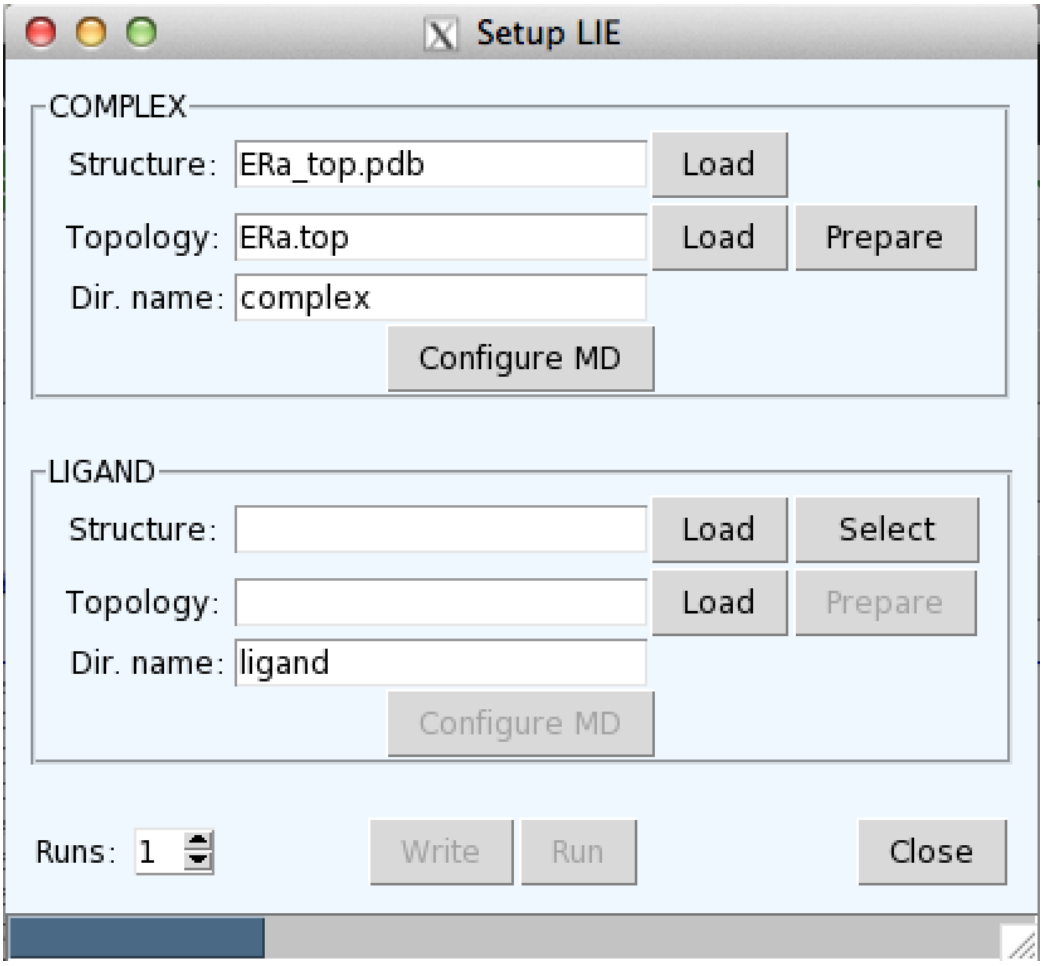

Figure 17: The setup window for LIE simulations.

The ligand simulation requires the pdb data of the ligand. This can be loaded in one of two ways, you can either load a pdb file or you can select it directly from the complex pdb file using the Select button (Fig 18). The topology for the ligand can be loaded or generated as with the complex simulation.

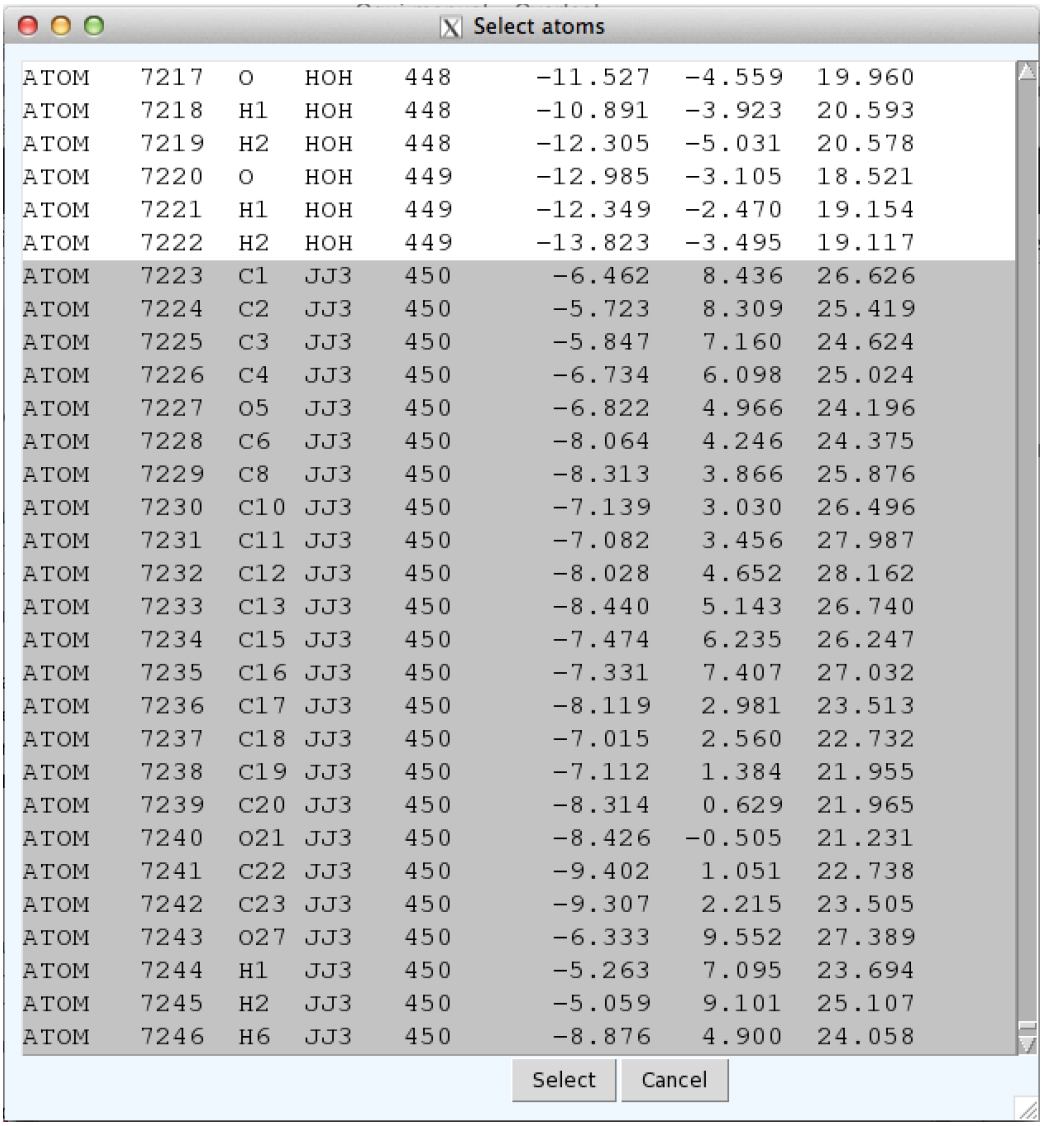

Figure 18: The select atoms window, where you can select the atoms belonging to the ligand and Qgui will automatically create the pdb file.

The user must also select the Configure MD tool for both complex and ligand simulations. For general MD settings see section 6. LIE simulations additionally require the user to define the ligand as Q-atoms (figure 19). The Q-atoms can be chosen by atom number or selected using the Select

tool, then click add to confirm the selection. For the ligand MD setup it is necessary to add a restraint to the ligand to ensure it will remain in the centre of the simulation.

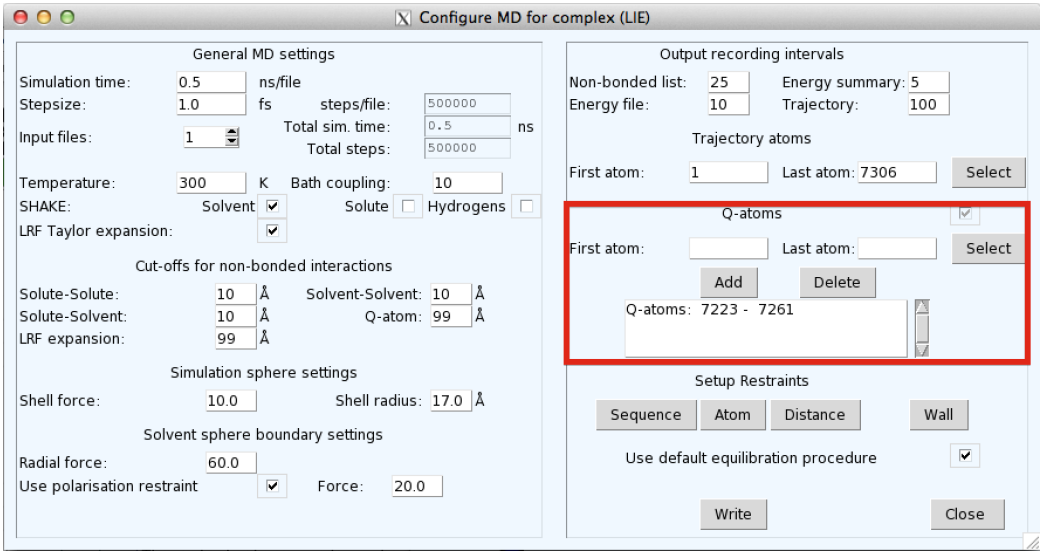

Figure 19: MD settings window for LIE simulations with select Q-atoms section highlighted.

Once everything is set up the user can choose the number of parallel runs for the simulation using the Runs tab. Finally the user can write the settings files with Write and submit the simulations using the Run button (figure 20).

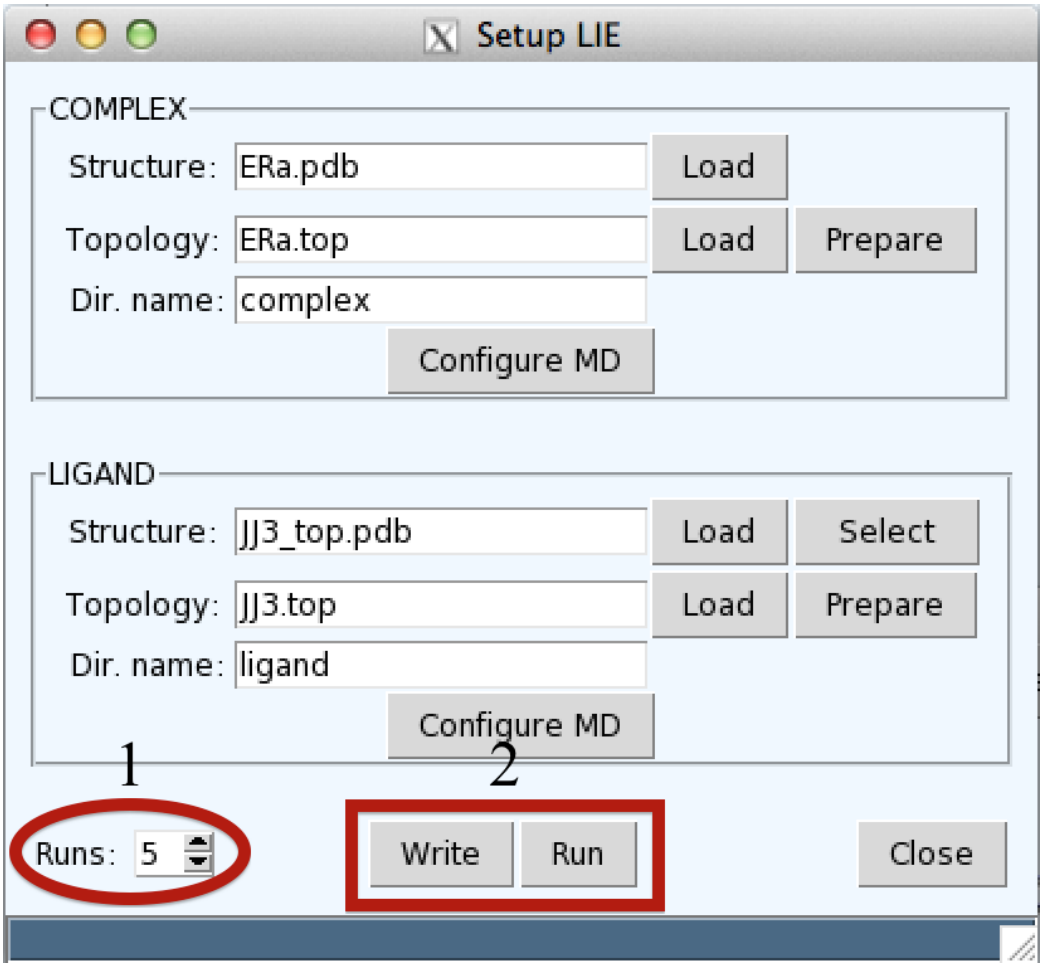

Figure 20: Completed LIE setup. 1: Choose number of parallel runs. 2: Write settings and Run simulations.

8 FEP calculations

The free energy perturbation (FEP) method calculates the relative free energy between two states (different ions, functional groups, etc.) by running an MD simulation over n discrete steps using a linear combination of the two states. ADD REF

Energy is calculated according to Zwansigs formula:

$$\Delta G_m = \beta^{-1} \sum_{m=0}^{n-1} ln\langle exp[-\beta(\epsilon_{m+1} - \epsilon_m)]\rangle \quad (2)$$

To set up a FEP calculation the user must first load a pdb and topology file of the system (see section 5.3 on details on how to create a topology). Then select the FEP tool, Setup → FEP.

The user then defines Q-atoms, these will be the atoms that can change during the FEP and for which energy will be sampled. Qgui currently support up to 4 state FEP, and the number of states can be changed in the setup section of the tool. (figure 21)

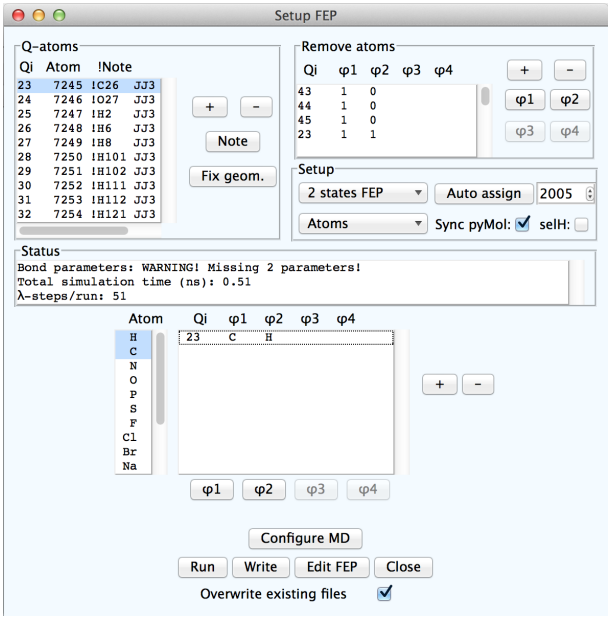

Figure 21: The FEP setup tool

PyMOL is integrated with QGUI and checking the sync PyMOL box will open a PyMOL window where the changes between the states can be observed directly (figure 22). It is also possible to select atoms directly from the PyMOL window by clicking them.

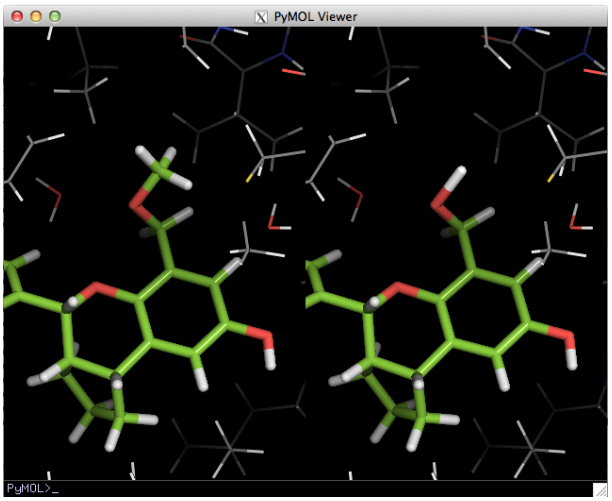

Figure 22: The sync PyMOL window of a two state FEP. State 1 left, state 2 right.

A number of different parameters can be changed during the FEP (figure 23). For a detailed guide check out the FEP tutorial (section: 8.1).

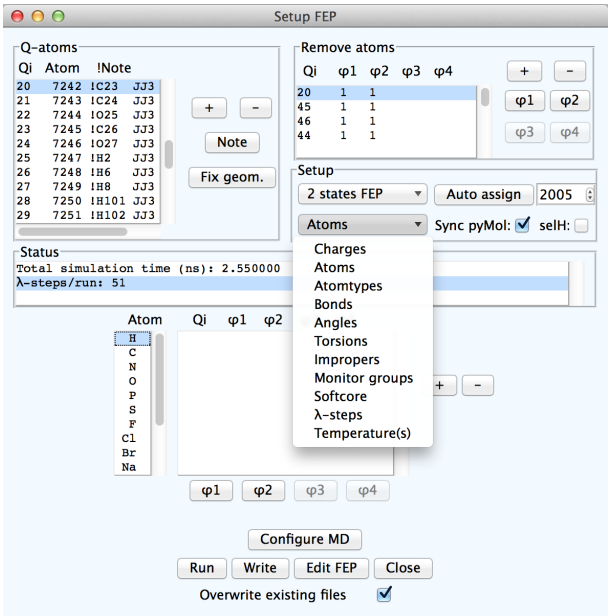

Figure 23: The FEP settings drop down menu

Qgui can also generate parameters for the new FEP states by clicking the "Auto assign" button. The user should carefully inspect the automatically generated parameters and add or change missing or wrong parameters manually. As with the parameter generator (section 5.2), this is also based on the Schrodinger ffld server and use the OPLS force field.

Before running the job the user needs to configure the MD settings, see section 6 for details. The FEP settings can be edited manually using the "Edit FEP" button. This is intended as the last step, as changes made here will be overwritten if the user commits further changes in the menus.

8.1 FEP tutorial

9 EVB simulations

For EVB simulations you will need the same files necessary for FEP: topology, fep file and the input files for MD. To setup the FEP file, MD input files and to run the jobs you need to go in the dropdown option **Setup** and then **EVB** (Fig 24).

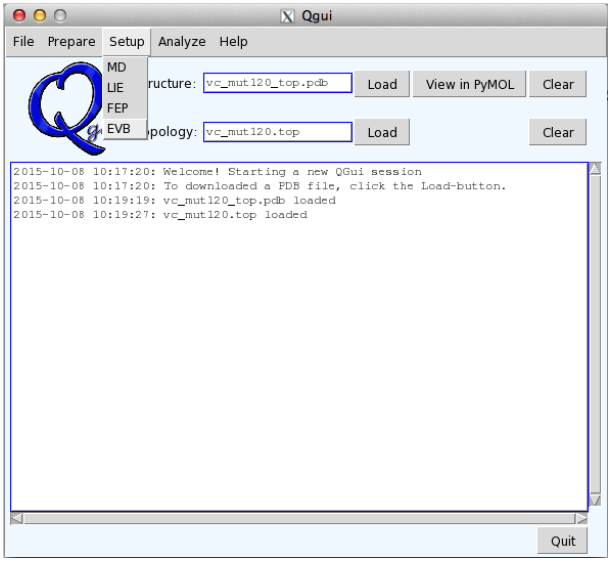

Figure 24: The main window showing the location of the EVB setup.

9.1 EVB - Tutorial

9.1.1 EVB - Setup

After loading the PDB file and topology of your structure, the first step in the EVB setup is to prepare the FEP file or load one if you have already (25). You have to start by choosing the Q atoms and by defining the bonds that are gonna be created/broken in the red square (MISSING PYMOL PART!!!). In order to define the bonds you have to select the involved atoms in the Q-atoms section and then press plus in the bonds section. You then have to define the state in which the bond is present, by selecting the connected atoms and then pressing $\varphi_(i)$.

In the status panel (Fig 25) are listed the missing parameters for the creation of the FEP file.

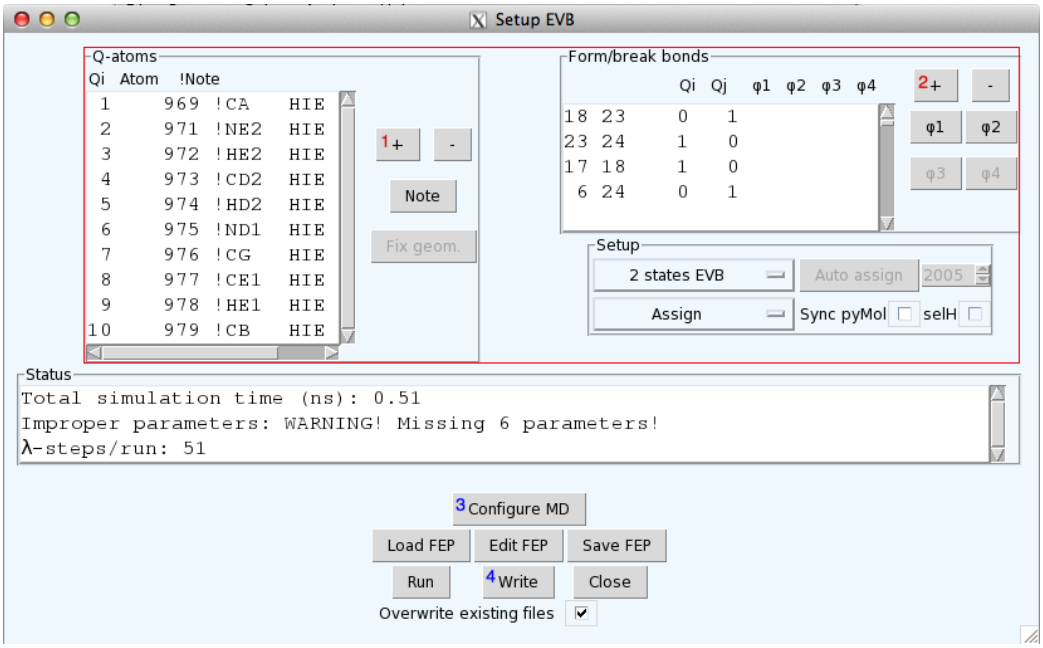

Figure 25: Main window for evb setup.

Once defined the Q-atom and the bonds broken/formed, you have to specify the Q-atoms atomtype change between the state (Fig 26). Afterwards you will see that all the bond, angle, torsion and improper parameters are

listed in the FEP file. In the red square are present the atomtypes related to your topology. So you will probably have to import from the parameter file (defined in the settings panel, see 4) all the extra atomtypes you need, by pressing **import**.

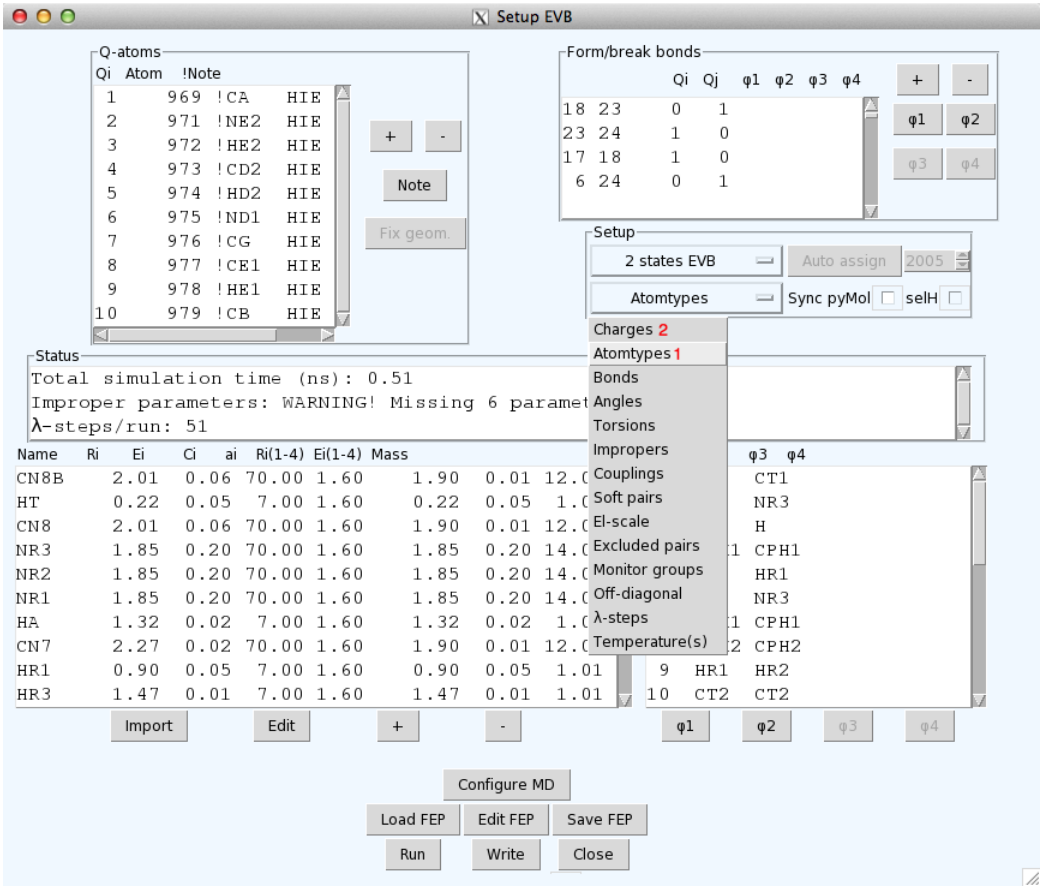

Figure 26: Subwindow for atomtype change definition in FEP file.

After preparing/loading the FEP file you can proceed in the last step of the EVB setup: MD input files preparation. In this section you would normally keep the default values for the MD setup, while the only details that you need to add, if required, are the restraints for your system (see 6).

9.1.2 EVB - Analysis

With the EVB wizard you can (Fig 27):

1. Calibrate your evb parameters to a reference reaction in water;
2. Run a normal FEP at fixed temperature, to get the ΔG^o of reaction and the $\Delta G^{\ddagger}$ of activation;
3. Run an Arrhenius simulation to get also the $\Delta H^{\ddagger}$ and $\Delta S^{\ddagger}$;

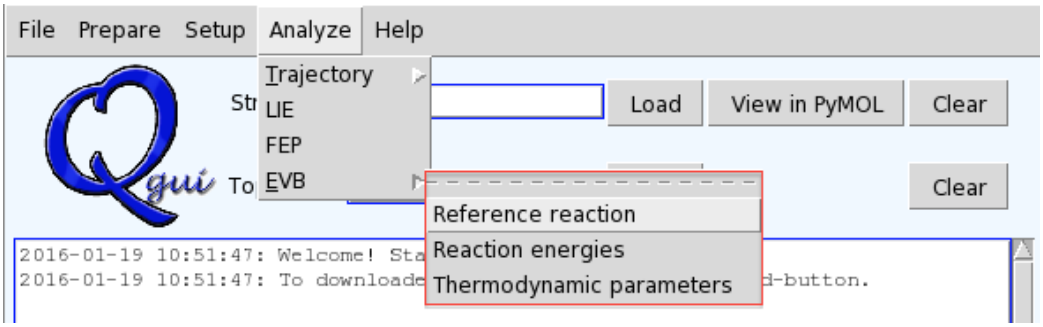

Figure 27: Evb analysis tab.

9.1.3 EVB - Reference reaction

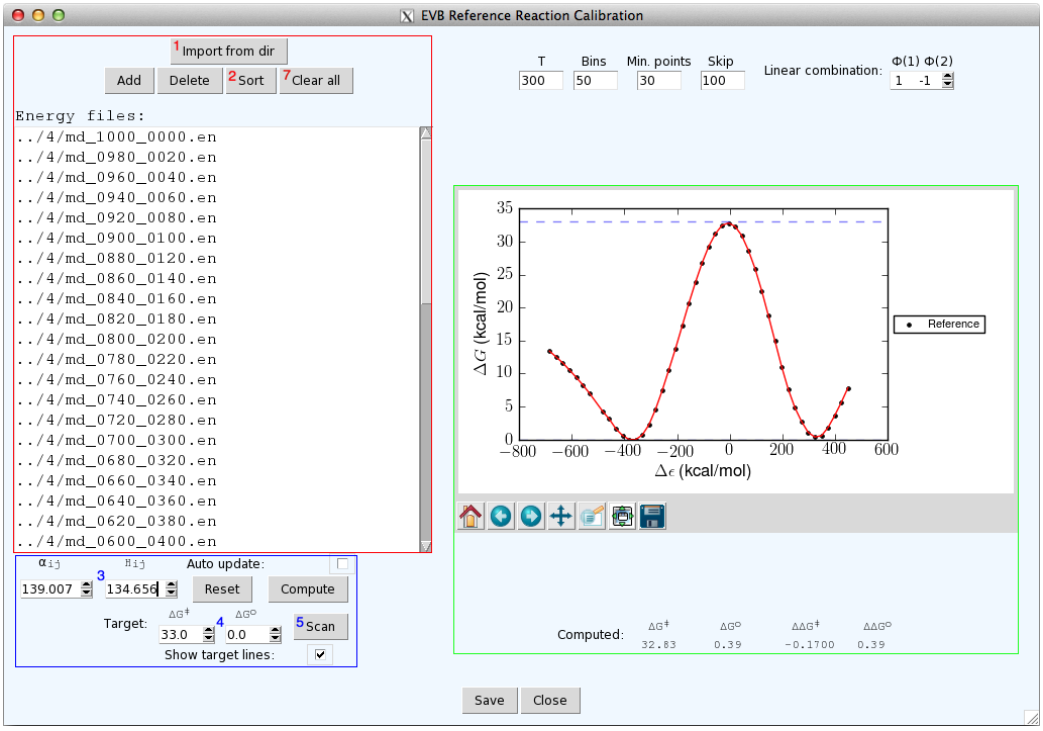

Figure 28: Reference reaction window for calibration of the EVB parameters.

9.1.4 EVB - Reaction energy

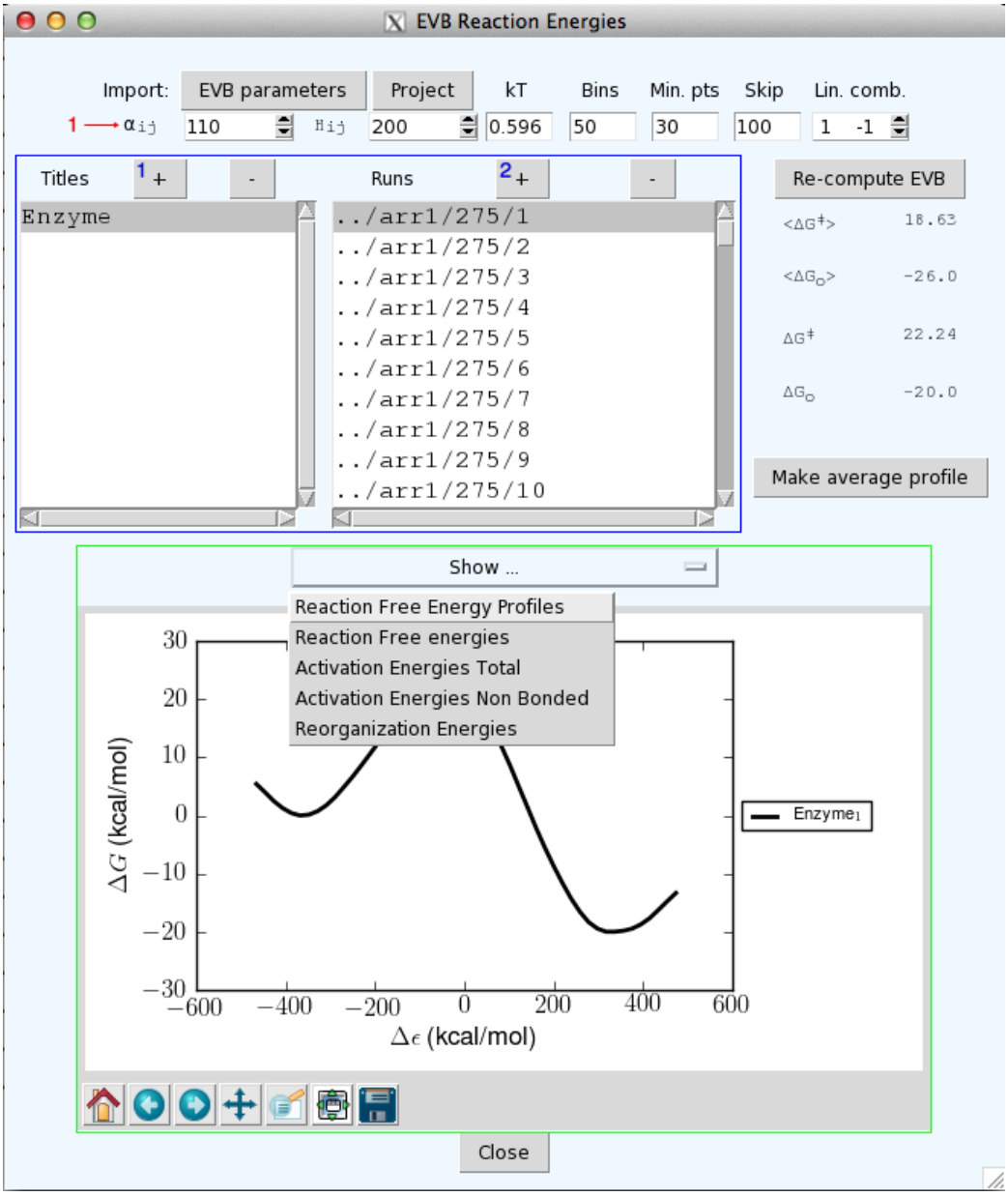

Figure 29: Reaction energy window to get the ΔG^o of reaction and the $\Delta G^{\ddagger}$ of activation.

9.1.5 EVB - Arrhenius plot

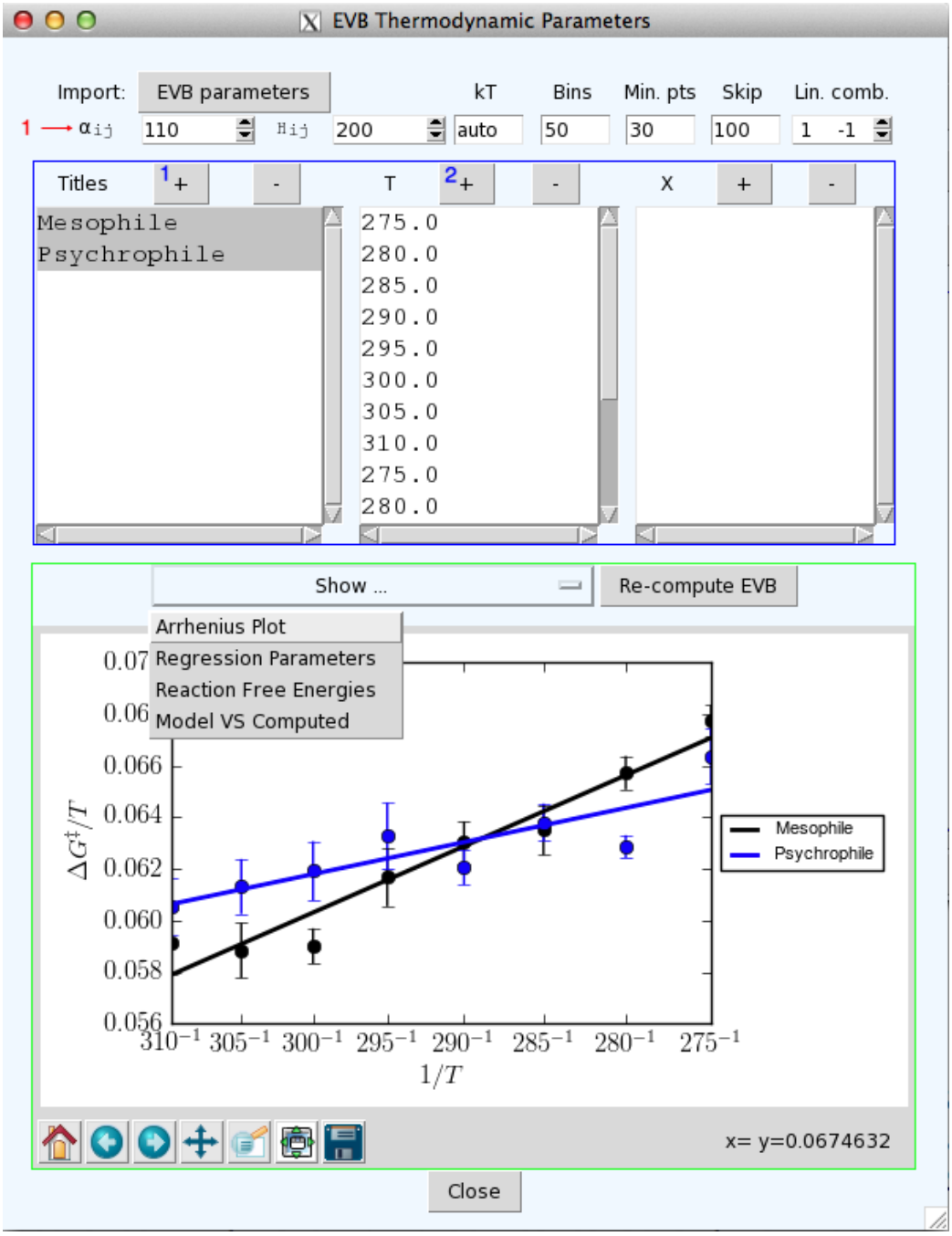

Figure 30: Arrhenius plot window to get the $\Delta H^{\ddagger}$ and $\Delta S^{\ddagger}$ of activation.