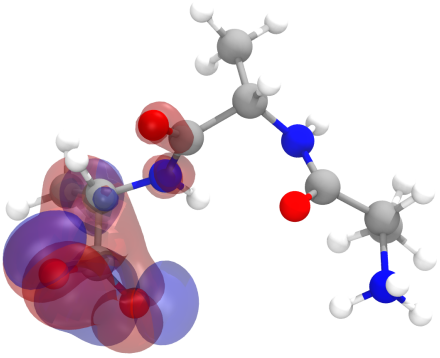
University of Illinois at Urbana-Champaign
NIH Resource for Macromolecular Modelling and Bioinformatics
Beckman Institute
Computational Biophysics Workshop

NAMD-QM/MM TUTORIAL

Unix/MacOSX Version



NAMD-QM/MM Suite Developers: Marcelo C. R. Melo, Rafael C. Bernardi

NAMD Tutorial Contributors: Marcelo C. R. Melo, Rafael C. Bernardi, João Ribeiro, Maximilian Scheurer, Till Rudack, James Phillips, John Stone, Zaida Luthey-Schulten

December 2017

A current version of this tutorial is available at
`http://www.ks.uiuc.edu/Training/Tutorials/`
Join the `tutorial-l@ks.uiuc.edu` mailing list for additional help.

Contents

Introduction

This tutorial provides a first introduction to NAMD-QM/MM interface and its basic capabilities. It can also be used as an introduction for the non-expert QM/MM user.

The tutorial assumes that you already have a working knowledge of VMD and that NAMD 2.12 or later has been installed correctly on your computer. For installation instructions, please refer to the NAMD User's Guide. For the accompanying VMD tutorial go to

`http://www.ks.uiuc.edu/Training/Tutorials/`

For a detailed description of NAMD the reader is referred to the NAMD User's guide located at

`http://www.ks.uiuc.edu/Research/namd/current/ug/`

The examples in the tutorial will focus on the study of a Calcium coordination site from a carbohydrate binding module, a protein that tightly binds sugars. Throughout the text, some material will be presented in separate "boxes". Some of these boxes include complementary information to the tutorial, such as details about QM/MM simulations, and tips or shortcuts for using NAMD. These boxes are not required for understanding the tutorial and may be skipped if you are short on time.

Boxes with an exclamation sign are especially important and should not be skipped.

Warning! The goal of this tutorial is to introduce QM/MM by performing some short molecular dynamics simulations. Therefore, the examples provided are optimized so simulations can be done in a reasonable period of time on a common computing facility. This means that some parameters and conditions under which simulations are done in this tutorial are not suitable for scientific studies. Whenever this happens it will be pointed out and alternatives or more appropriate parameters/conditions will be provided in case you want to improve the simulations and/or you have more computer power available.

System Requirements

The QM/MM interface of NAMD/VMD is available for MacOS X and Unix.

- **Mac**
  Operating System: Mac OS X 10.4.7 or later
  Memory: 1 GB RAM
  Processor: Intel Core 2 Duo 2.0 Ghz CPU (or comparable)
  Graphics card: NVIDIA GeForce 320M, or comparable

- **Linux**
  Operating System: Any reliable Linux distribution released within the last three years.
  Memory: 1 GB RAM
  Processor: Intel Core 2 Duo 2.0 Ghz CPU (or comparable)
  Graphics card: NVIDIA GeForce 320M, or comparable

The speed of a QM/MM calculation is heavily dependent on the level of the QM treatment, and the size of the QM region. Performing larger and more complex calculations than those presented in this tutorial will likely require a more powerful computer.

Required Programs

For more details check `http://www.ks.uiuc.edu/Research/qmmm/`

The following programs are required for this tutorial:

- **NAMD:** Available at `http://www.ks.uiuc.edu/Research/namd/` (for all platforms). QM/MM support is available in version 2.12 or newer, for both UNIX and MacOS. We recommend the use of the nightly build version of NAMD, as it will include recent bug-fixes.
- **VMD:** Available at `http://www.ks.uiuc.edu/Research/vmd/` (for all platforms) QM/MM support is available in version 1.94 or newer. A VMD version with the most up-to-date QM/MM implementations is available in: `http://www.ks.uiuc.edu/Research/qmmm/`
- **ORCA:** Available at `https://orcaforum.cec.mpg.de/`.
- **MOPAC:** Available at `http://openmopac.net/downloads.html`.

Installation Guide

Both NAMD and VMD are distributed pre-compiled. In both cases instalation can be performed in less than a minute. More information available at `http://www.ks.uiuc.edu/Development/`

- **NAMD:** A NAMD binary distribution need only be untarred or unzipped and can be run directly in the resulting directory. When building from source code, "make release" will generate a self-contained directory and .tar.gz or .zip archive that can be moved to the desired installation location. Windows and CUDA builds include Tcl .dll and CUDA .so files that must be in the dynamic library path.
- **VMD:** To install the pre-compiled *MacOS X* bundle version of VMD, open the VMD disk image and drag the VMD application into an appropriate directory. Once the VMD application has been placed appropriately it should be ready for immediate use as no other installation steps are required.

  To install the pre-compiled *Unix* version of VMD, then only three steps remain to be done after you uncompress and untar the distribution.

  Edit the configure script. If necessary, change the following values:

  $install_bin_dir

  This is the location of the startup script 'vmd'. It should be located in the path of users interested in running VMD.

  $install_library_dir

  This is the location of all other VMD files. This includes the binary and helper scripts. It should not be in the path.

  Next generate the Makefile based on these configuration variables. This is done by running ./configure .

  After configuration is complete, cd to the src directory and type make install. This will put the code in the two directories listed above. After this, you just type vmd to begin, provided that vmd is in your path.
- **ORCA:** Instructions available at:

  `https://orcaforum.cec.mpg.de/`
- **MOPAC:** Instructions available at:

  `http://openmopac.net/manual/Installing_MOPAC.html`

Getting Started

- If you have downloaded this tutorial at home, you will also need to download the appropriate files, unzip them, place them in a directory of your choosing, and then navigate to that directory.

1 Basics of QM/MM

In this section you will learn how to use the NAMD-QM/MM interface to set up basic molecular dynamics (MD) simulations that use classical and quantum mechanical calculations. You will learn about typical input and output files, and basic options available to set up your simulation.

NOTE: You will be generating output data in this section by performing simulations and using other features of NAMD. If you are not able to produce the output, correct versions have been provided for each section and may be found in the `example-output` folder.

1.1 What is Needed

In order to run any QM/MM simulation, NAMD requires at least four things:

- a Protein Data Bank (pdb) file which stores atomic coordinates and/or velocities for the system. Pdb files may be generated by hand, but they are also available via the Internet for many proteins at `http://www.pdb.org`.
- a Protein Structure File (psf) which stores structural information of the protein, such as various types of bonding interactions.
- a force field parameter file. A force field is a mathematical expression of the potential which atoms in the system experience. CHARMM, X-PLOR, AMBER, and GROMACS are four types of force fields, and NAMD is able to use all of them. The parameter file defines bond strengths, equilibrium lengths, etc.
- a configuration file, in which the user specifies all the options that NAMD should adopt in running a simulation. The configuration file tells NAMD how the QM/MM simulation is to be run and how the QM code is to be executed.
- a "QM"-PDB file which indicates how atoms are partitioned in different QM regions. QM-PDB files may be generated by hand or by using VMD's QwikMD interface, as will be demonstrated in this tutorial.

1.2 Preparing Your System

In order to set up a system that will be studied with QM/MM simulations, we will use QwikMD to download a protein structure, prepare it for classical MD simulations, and then use the result to initiate a QM/MM simulation where only a small portion of the system is simulated using Quantum Mechanics.

1 Open VMD and then open QwikMD by clicking `Extensions -> Simulation -> QwikMD` menu item in the VMD main window.

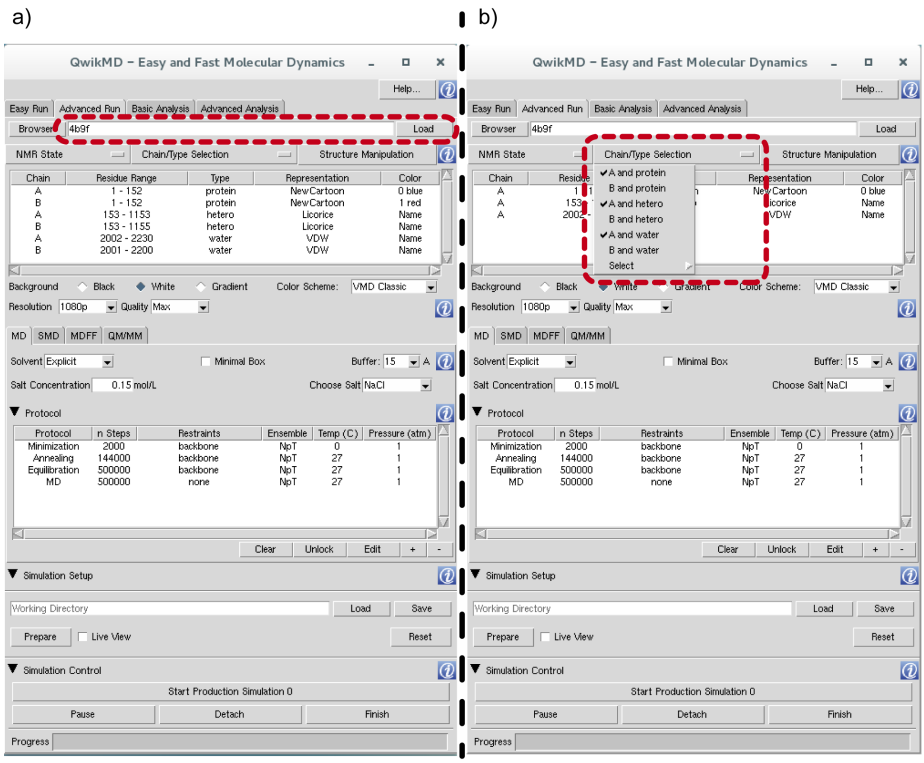


Figure 1: **a)** QwikMD main window showing the system automatically downloaded after entering the PDB code in the top navigation bar, and clicking "Load". **b)** Highlight for the selection drop-down menu.

2 Type 4b9f in the navigation bar and click "Load". This will automatically download the PDB file form the PDB databank.

You should see the window populated with information from the downloaded structure (Figure 1a) and the structure will be displayed in VMD's main window (Figure 2). We will only use one of the protein chains displayed in the X-ray structure, so our system will only contain one protein chain and one Calcium-binding site.

3 Click on "Chain/Type Selection" and un-select the protein, "hetero" and water selections for the "B" chain (Figure 1b).

We will use the "Structure Manipulation" window to remove sulfates used for the crystallization buffer. This window can also be used to make mutations and changes to your structure before the beginning of the simulation.

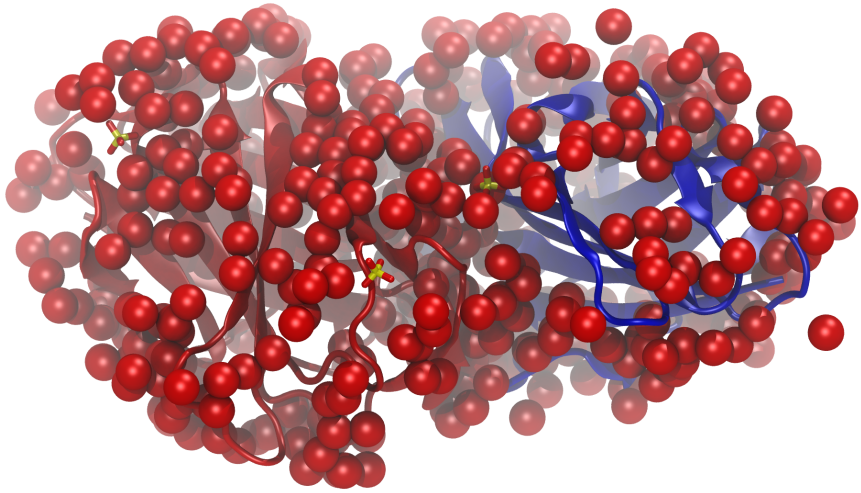


Figure 2: VMD main window showing the loaded system.

4 Click on "Structure Manipulation" and scroll down to residue 1153 named "SO4". Select the residue and click on the "Delete" button. Now click on the "Apply" button. (Figure 3a).

5 Still on the "Structure Manipulation", select the residue 153 and click on the "Rename" button. Now click on the "Res NAME" field of the line and select "Calcium" from the drop-down list. Now click on the "Apply" button (Figure 3b).

QwikMD automatically re-names several atoms and residues to match the naming system of parameters sets like that of the CHARMM force field. In this case, the Calcium ion was already re-named from the PDB naming system to the CHARMM naming system, but this served as an example for future systems that may not have known names among CHARMM parameter sets.

6 Run the structural verification by clicking the "Check" button in the "Structure Manipulation" window. The red box next to "Topologies & Parameters" should become green. This indicates QwikMD has all the parameters required to simulate the molecules in the system. Close the window.

Now that we have the desired protein and ion complex, we need to prepare the system for an MD simulation, where we will equilibrate the structure before beginning a QM/MM simulations. The protein-ion complex will be placed in a water box with counter ions.

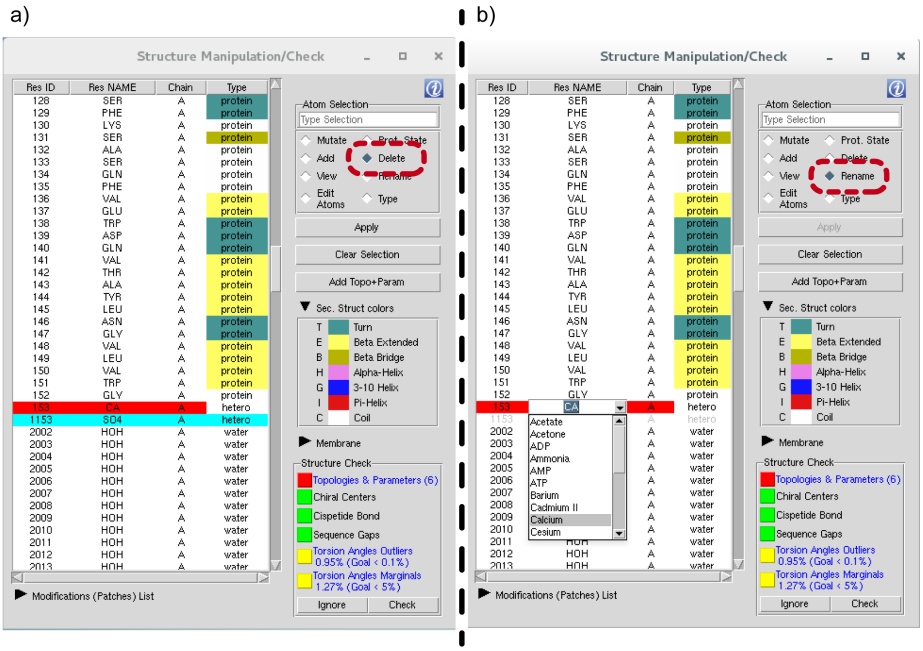


Figure 3: **a)** QwikMD "Structure Manipulation" window indicating the deletion of a residue from the system. **b)** Highlight for the "Rename"option and drop-down menu for residue names.

7 In the main QwikMD window, click on the "Advanced run" tab, and then on the "MD" tab. Then check the box for "Minimal Box". and select a 12 Åbuffer of solvent molecules around our protein (Figure 4a).

8 Finally, in the "Protocol" section, select the "MD" line and click the "-" (minus) button to remove this simulations step. Prepare your simulation under the "Simulation Setup" section and run it. The results will be used to initialize the QM/MM simulation.

We run a classical simulation so that we can equilibrate the molecule we just downloaded from the PDB, as well as the solvent atoms created to represent the solution where the protein is found.

1.3 Preparing Your QM/MM Simulation

After your classical simulations have ended, re-start VMD (if VMD was never closed, click on the "Reset" button under the "Simulation Setup" section).

1 Go to QwikMD's "Simulation Setup" section and load your previous MD simulation (Figure 5a).

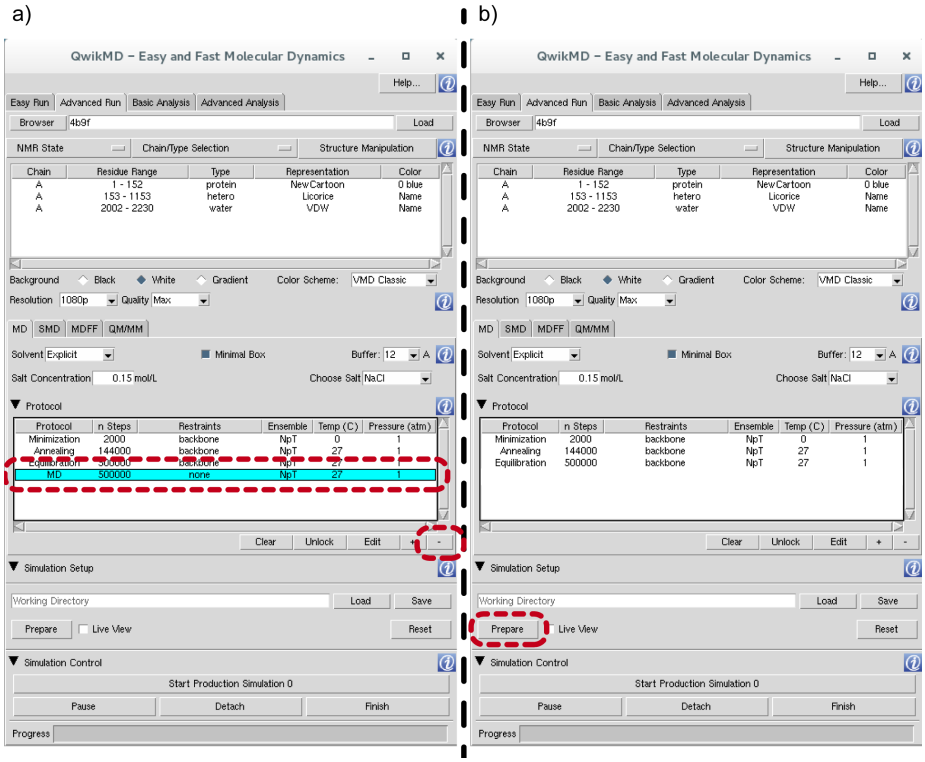


Figure 4: **a)** QwikMD main window indicating the selection of the "MD" simulation step. **b)** Highlight for the "Prepare" option with all simulation parameters set.

In this tutorial, you should load only the data from the equilibration trajectory, and mark that only the last step of the simulation should be loaded. We are using the last step form the classical equilibration step to select our QM region and initiate the QM/MM simulation.

2 In the "QM Options" section, make sure the "QM Software" indicates MOPAC. If it does not, select it form the drop-down menu (Figure 5b).

3 In the "QM Regions" section, click the "+" (plus) button to define a new QM region in your system. A new line should appear in this section (Figure 5b).

4 Click on the "n Atoms" column (see Figure 5b) to open the "QM Region Selection" window (Figure 7a) .

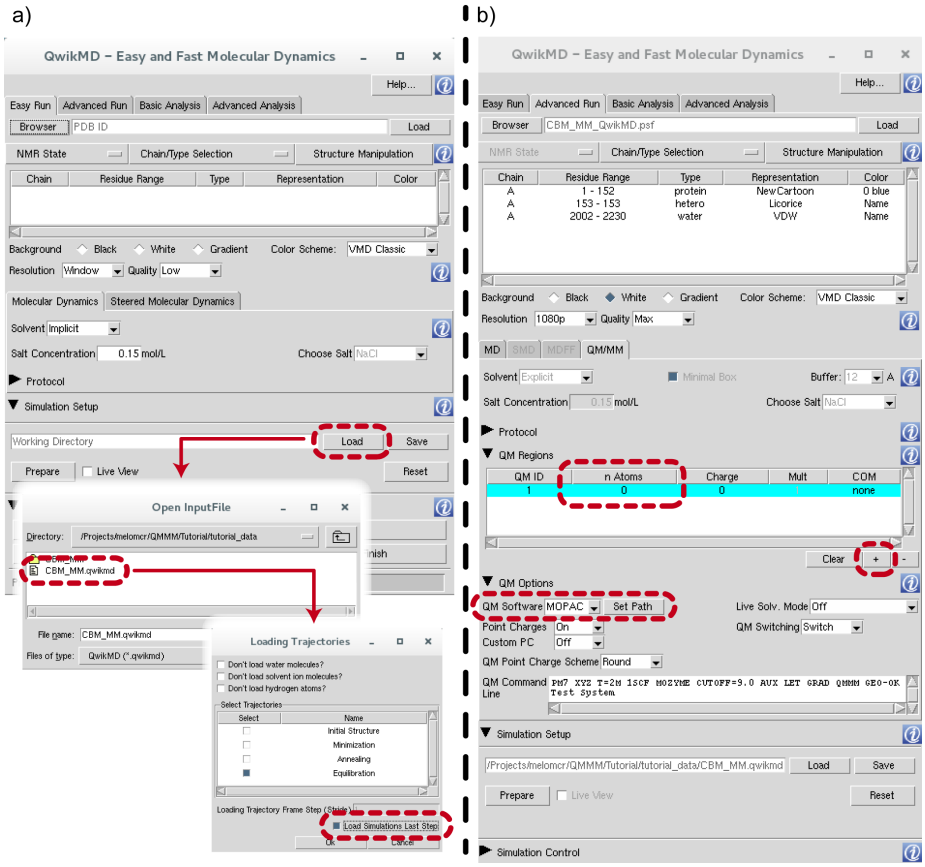
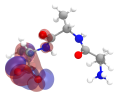


Figure 5: **a)** QwikMD windows indicating the steps for loading a previously ran MD simulation. Loading only the last simulation step will save time and memory space. **b)** QwikMD main window indicating the loaded data from previous simulations, the "QM Options" section with a highlight for the drop-down menu where MOPAC should be selected, a highlight for the button used to add QM regions to your system, and a highlight for the column indicating the size of each QM region, .



Semi-Empirical QM Calculations. Since semi-empirical methods only explicitly represent a fraction of all the electrons in a system, heavily charged molecules tend to present a big problem for the distribution of charges, occasionally generating abnormal charge distributions. For this reason, QwikMD will check the total charge of the QM region and, in case a simulation is being ran with MOPAC, will help you fine-tune your selection by indicating charged residues near your QM region. After adding and/or removing residues, the total charge for each QM region should be between $+1$ and -1.

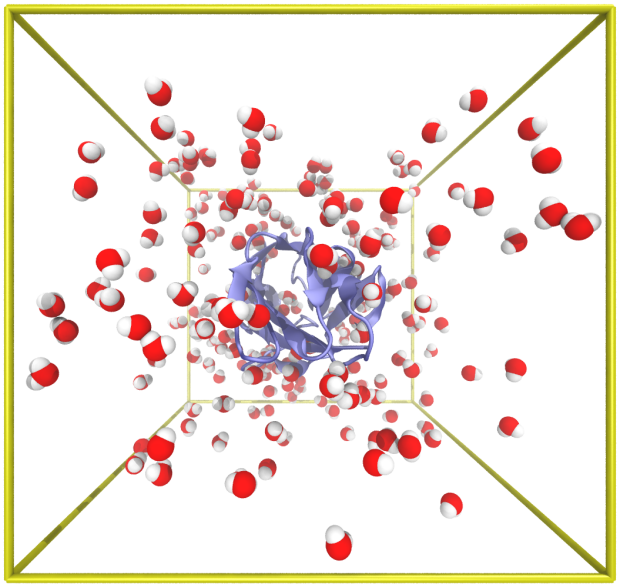


Figure 6: VMD main window showing the loaded MD system.

Total charge in semi-empirical calculations Semi-empirical calculations can display problems when distributing electrons across a QM region that is heavily charged. For this reason, when MOPAC is selected as the QM software, QwikMD will not allow the user to perform a simulation if a QM region which has a charge larger than ±1. If a selection is made that brakes this rule, QwikMD will warn you and indicate neighboring residues that could be added to (or removed from) the selection in order to neutralize some of the charge. As an initial test, we will select a small region that will have a total charge of +2 (assuming you loaded the pre-ran MD results provided with this tutorial). This will show you how QwikMD helps you adjust the QM region selection as to neutralize (or reduce) the total charge.

5 In the "QM Region Selection" window (Figure 7b), change the radius for solvent within the QM region to 5 Å. Now scroll down the residue list in the same window until you find the Calcium residue (Res ID: 153; Res NAME: CAL) and click on that line. The calcium ion and two water molecules should be selected. Click "Apply" so QwikMD can process your selection.

The new selection will be displayed by QwikMD with 7 atoms (the calcium ion and two water molecules) and the total charge shown on the window will be updated to +2. Since this is outside the ±1 range, a representation similar to

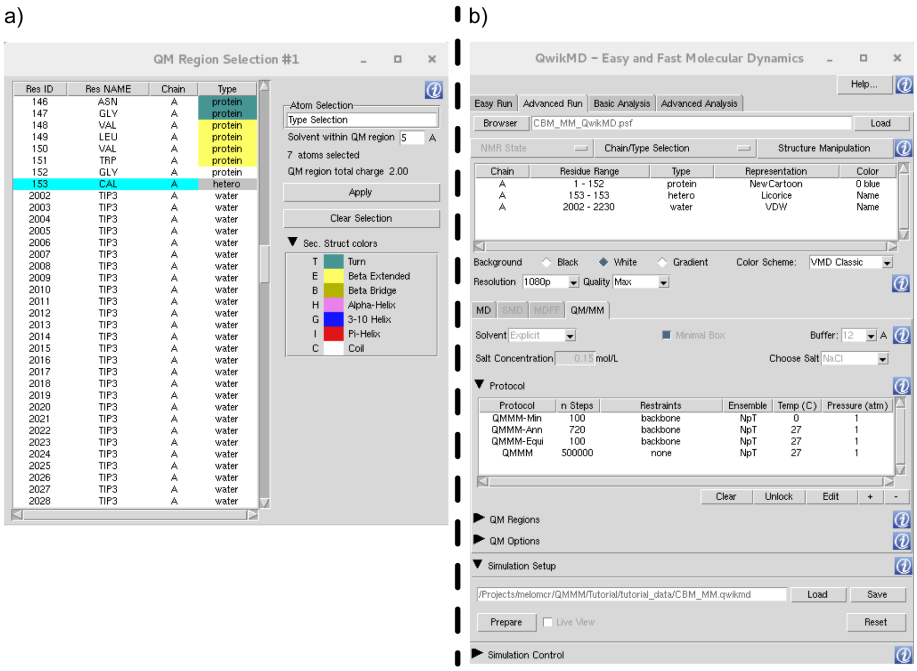


Figure 7: **a)** QwikMD "QM Region Selection" window, indicating the selection of the Calcium ion, and the radius for inclusion of solvent molecules in the QM region. **b)** QwikMD main window indicating the different simulation steps pre-prepared for QM/MM simulations.

Figure 8 will be created, indicating charged residues near the QM region (red for negatively charged and blue for positively charged).

6 Click on the transparent red residues and note the charge of the QM region changing to reflect the new selection.

NOTE: Every time you add or remove a residue from the QM region selection, QwikMD will update the selection of water molecules to reflect the new QM region selection.

Effective QM Region Selection In order to make a more effective selection for a QM/MM simulation of the ion binding pocket, we will re-define our QM region and select residues and water molecules around the Calcium ion.

7 In the "QM Region Selection" window, click on the Clear Selection button (if you closed the window, just go to QwikMD's main window, go to "QM Regions", and click on the QM region line under the "n Atoms" column).

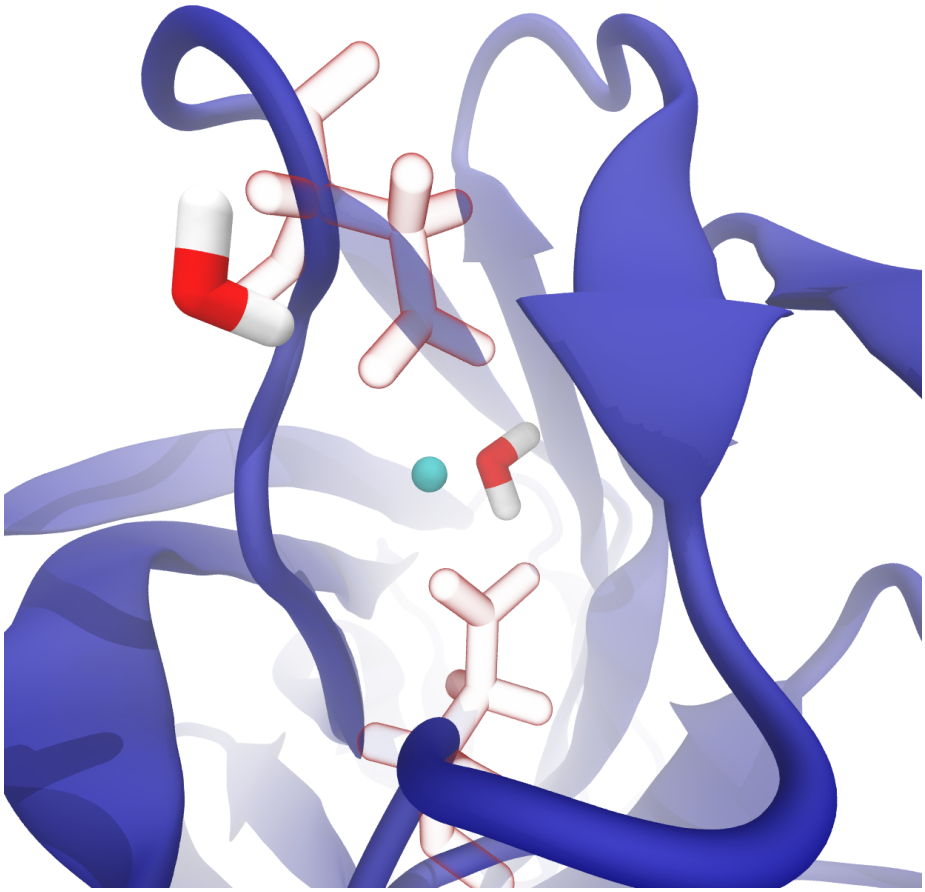


Figure 8: VMD main window showing the selected QM region in solid colors, and nearby charged residues in transparent red representation for negatively charged and transparent blue for positively charged.

8 Change the solvent selection radius to 0 (zero).

9 Place the cursor in the Atom Selection text field and type the following selection: same residue as (within 5 of (resname CAL)).

The parenthesis are not necessary in this selection, but helps us parse the statement. This selection is made using VMD's *atomselect* language, which allows easy and flexible selection of regions within your system using logical statements and direct access to properties of atoms and residues. To understand this selection, we will read it starting from its inner most statement: First, we select every residue with name CAL, which only occurs once in this particular system, and is the Calcium ion. We then extend the selection to include every

atom within 5 Åof the residue, and then expand it again to include all atoms of residues that were partially included in the 5 Åradius around the Calcium ion (Figure 9). Once again, assuming you loaded the pre-ran MD results provided with this tutorial, a total of 148 atoms will be selected by this "Atom Selection" statement.

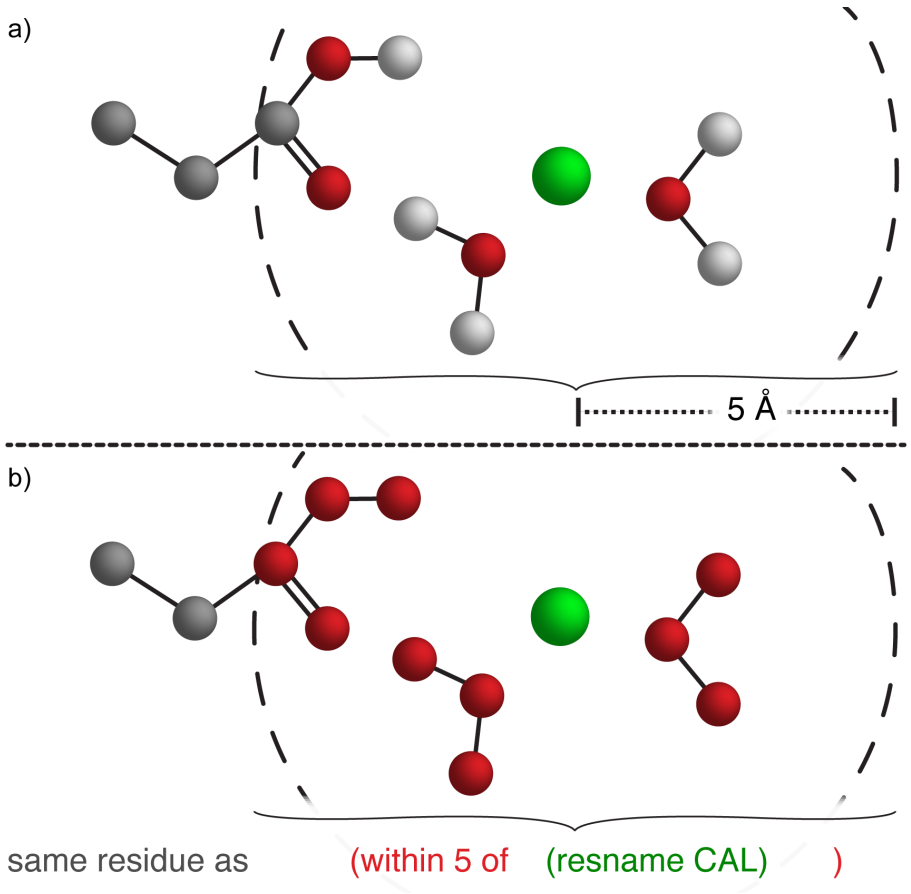


Figure 9: **a)** Simplified representation of the example used in the tutorial. At the center, the green sphere represents the Calcium ion, two water molecules are surrounding it, and a portion of an Aspartate residue. The dashed circle represents a 5 Å radius around the Calcium ion. **b)** The same system is represented in colors matching the *atomselect* statement.

We can now proceed to prepare the simulation by clearing the path shown in the Simulation Setup section, and clicking on "Prepare". You will give a path to the location where all input and configuration files will be saved, and when

asked which starting step you would like to use, select Equilibration.

NOTE: Reduced output files are provided with this tutorial, since a 500 thousand step simulation accumulates too much information for it to be feasible to download and analyze quickly.

2 Advanced

Speeding-up every-step I/O

If you are using a Linux distribution, it is possible to create a folder for temporary files mounted directly on RAM, usually by creating a folder under /dev/shm. You can change where NAMD stores temporary files from QM executions by clicking on a protocol, such as the QMMM-Min line, under the Protocol section, and clicking on the Edit button. This will show you the complete NAMD configuration file, where you can change any detail for the selected simulation step. The path given to the qmBaseDir keyword is the one used to run QM calculations, and can be substituted by the folder you created. Changes must be made to each protocol individually.

Dynamic Bonds in VMD

To improve the visualization of chemical reactions we added new features to two of the most used Representations of VMD, namely Licorice and CPK. Using VMD's "Tk Console" (VMD Main > Extentions > Tk Console), a new representation with dynamic bonds can be created or modified:

creation commands:

- mol representation CPK 1.1 0.3 8.0 6.0 1.5
- mol material Opaque
- mol selection all
- mol addrep 0

modify an existing rep (rep index 0 for molecule 0):

- mol modstyle 0 0 CPK 1.1 0.3 8.0 6.0 1.5

In both cases the last digit after CPK (1.5) represents the maximum bond distance.

3 Acknowledgment

The developement of NAMD and the tutorials are funded by the National Institute of Health (P41- RR005969 - Resource for Macromolecular Modeling and Bioinformatics). Proper citation is a primary way in which we demonstrate the value of our software to the scientic community, and is essential to continued NIH funding for NAMD. The authors request that all published work which utilizes NAMD/VMD include the primary NAMD/VMD citations:

Scalable molecular dynamics with NAMD. James C. Phillips, Rosemary Braun, Wei Wang, James Gumbart, Emad Tajkhorshid, Elizabeth Villa, Christophe Chipot, Robert D. Skeel, Laxmikant Kale, and Klaus Schulten. Journal of Computational Chemistry, **26**:1781-1802, 2005

VMD - Visual Molecular Dynamics William Humphrey, Andrew Dalke, and Klaus Schulten. Journal of Molecular Graphics, **14**:33-38, 1996

NAMD goes quantum: A new integrative suite for QM/MM simulations. Marcelo C. R. Melo, Rafael C. Bernardi, Till Rudack, Maximilian Scheurer, Christoph Riplinger, James C. Phillips, Julio Maia, Gerd B. Rocha, João V. Ribeiro, John E. Stone, Frank Neese, Klaus Schulten, Zaida Luthey-Schulten. *Submitted* (2017)