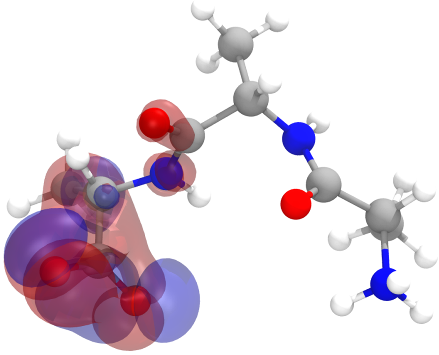
**University of Illinois at Urbana-Champaign**
**NIH Resource for Macromolecular Modelling and Bioinformatics**
**Beckman Institute**
**Computational Biophysics Workshop**

# NAMD-QM/MM Tutorial: Advanced Module Free Energy Profile of Reaction Mechanisms

Unix/MacOSX Version



**NAMD Tutorial Contributors: Marcelo C. R. Melo, Rafael C. Bernardi, João Ribeiro**

April 2018

A current version of this tutorial is available at
`http://www.ks.uiuc.edu/Research/QMMM/`

# Contents

# Introduction

This tutorial provides a detailed methodology of how to prepare, run and analyze complex QM/MM simulations. In real-life applications, it is usual to find molecular complexes that have not had their 3D structure experimentally determined. You will learn how to make use of a similar structure available in the PDB, mutate the ligand of interest, and set up Steered Molecular Dynamics (MD) simulations that use both classical and quantum mechanical calculations. From there, you will learn how to select representative stages of a chemical reaction, and prepare a String-Method optimization of the reaction coordinate. The optimized path will then be used to run an extended ABF (eABF) simulation that calculates the free energy change along the reaction.

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/`

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/VMD.

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

**Warning!** The goal of this tutorial is to provide an overall methodology and demonstrate it by performing some short QM/MM MD 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: 4 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 two years.
  Memory: 4 GB RAM
  Processor: Intel Core 2 Duo 2.0 Ghz CPU (or comparable)
  Graphics card: NVIDIA GeForce 320M, or comparable

# 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`.
- **Jupyter Notebook:** Available at `http://jupyter.org/install`. An updated list of packages used during the tutorial can be found in the tutorial notebook itself.

## 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 Preparing up your system

*In this section you will learn how to use VMD's Molefacture tool and TCL scripting to prepare problem-specific QM/MM simulations. You will learn how to make use of a similar structure available in the PDB, mutate the ligand of interest, as well as the protein structure, in order to perform simulations that use both classical and quantum mechanical calculations.*

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 Creating the molecular system

In real-life applications, it is usual to find yourself studying a molecular complexes that has not had its 3D structure experimentally determined. Either because a different ligand has been used in the crystallization process, or because the protein has amino acids with post-translational modifications, or even because mutations were applied to the protein to keep it from performing its function, maintaining the product bound to the active site but unmodified. In such cases, we generally start off by using a similar structure, and make the necessary modifications.

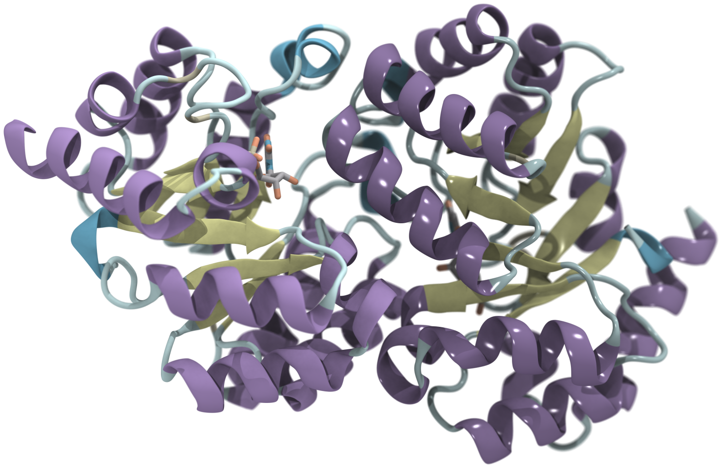


Figure 1: Orotidine 5'-monophosphate (OMP) decarboxylase enzyme (ODCase) as download from the PDB (PDB ID 1X1Z). The protein is shown in cartoon representation and the OMP ligand in licorice.

Our example is the Orotidine 5'-monophosphate (OMP) decarboxylase enzyme (ODCase). This is one of the most efficient enzymes ever studied, reducing the energy barrier of the decarboxylation of OMP by several orders of magnitude. The half-life of the substrate in neutral aqueous solution is about 78 million years, but when catalyzed by ODCase, this changes to 18 ms.

There is no available structure for the complex we wish to study, so we will build one starting from the X-Ray structure of a mutated enzyme complexed with 6-hydroxyuridine-5'-phosphate (BMP), available in the PDB with ID 1X1Z (Fig. 1).

If you open this PDB code in VMD, you will notice the enzyme was crystallized as a dimer, and we know from the literature that the dimer form is relevant for its biological activity, but for the purposes of this tutorial, only one protein chain will be used, since the monomer is still enzymatically active (Fig. 2).

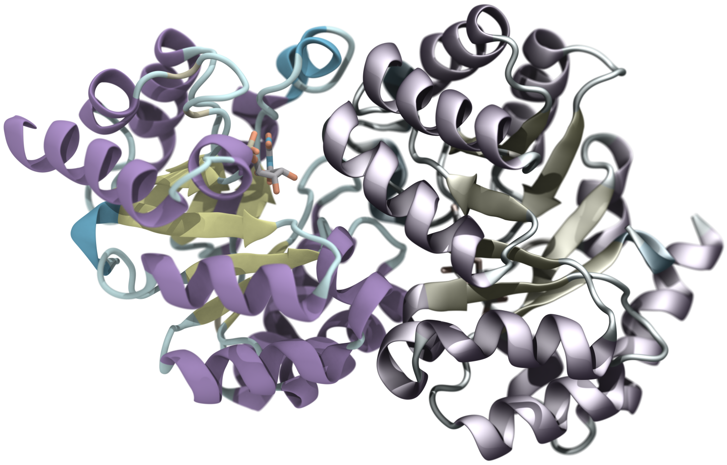


Figure 2: ODCase dimer, shown in different colors for clarity.

We are starting from scratch, so open a terminal window and navigate to the tutorial folder inside your home folder, and run the TCL script that isolates the monomer:

**1** cd `~/QMMM_String_eABF_Tutorial/setup`

In order to run a script using VMD's text mode, we use the *-dispdev text* option, as follows:

**2** vmd -dispdev text -e prepare.tcl

The content of the TCL script can be found bellow.

```
# Loads molecule by downloading directly from the PDB website
mol new 1X1Z

# Creates a selection of atoms from the downloaded structure.
#  We select only the protein from the chain A of the dimer.
set sel [atomselect 0 ”(protein and chain A)”]

# Write the extracted protein in a PDB file
$sel writepdb enzyme.pdb
```

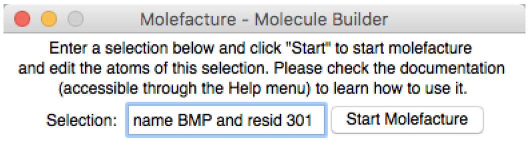
Listing 1: Contents of prepare.tcl

A file called “enzyme.pdb” will be created that only contains the protein chain from one of the monomers from the PDB structure. In order to get the correct ligand in place, we need to manually extract it and modify it using VMD. The particular extension we will use is called “Molefacture”, and it provides a very powerful tool to edit molecules atom-by-atom, even though it does not currently provide the clearest user interface.

**3** Open VMD and load the same PDB code used in the TCL script: 1X1Z

**4** In VMD, open Extensions → Modeling → Molefacture

**5** In the initial Molefacture window, use the following selection: resname BMP and resid 301 (Fig. 3). Click on **Start Molecule**.



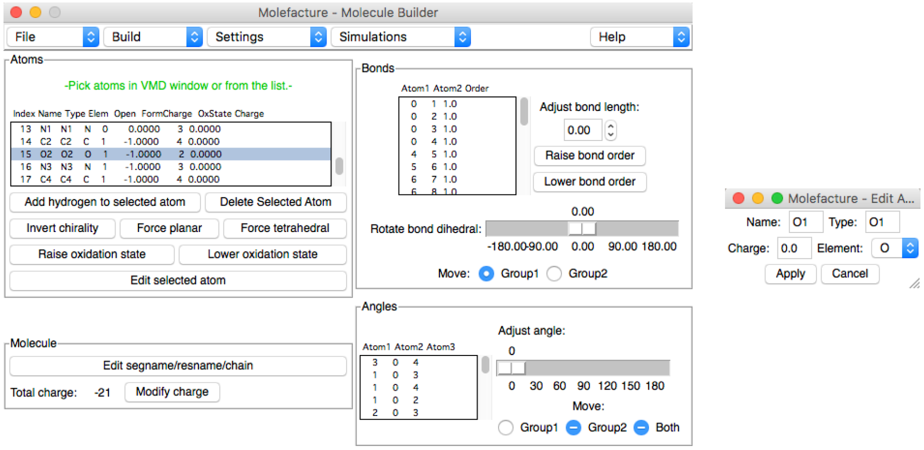
Enter a selection below and click "Start" to start molefacture and edit the atoms of this selection. Please check the documentation (accessible through the Help menu) to learn how to use it.

Selection: name BMP and resid 301

Figure 3: Molefacture opening window, with selection of ligand residue.

The main Molefacture window will open showing four enclosed boxes: Atoms, Bonds, Angles and Molecule (4).

You will see the ligand molecule in the main VMD window (Fig. 5), where each atom has its open valences highlighted in colored “sticks”, along with their net charges. You will notice the structure does not have any hydrogen atoms, giving the molecule a net charge of -21. We will use Molefacture to add the missing hydrogens and transform the oxygen “O1” into a carbon, and then add oxygen atoms to it to create the CO2 group that will later be “decarboxylated”.



```
Index Name Type Elem  Open  FormCharge  OxState Charge
  13   N1   N1    N    0     0.0000       3  0.0000
  14   C2   C2    C    1    -1.0000       4  0.0000
  15   O2   O2    O    1    -1.0000       2  0.0000
  16   N3   N3    N    1    -1.0000       3  0.0000
  17   C4   C4    C    1    -1.0000       4  0.0000
```

Total charge: -21

```
Atom1 Atom2 Order
  0     1   1.0
  0     2   1.0
  0     3   1.0
  0     4   1.0
  4     5   1.0
  5     6   1.0
  6     7   1.0
  6     8   1.0
```

Adjust bond length: 0.00

Rotate bond dihedral: 0.00 (-180.00 -90.00 0.00 90.00 180.00)

Move: ● Group1 ○ Group2

```
Atom1 Atom2 Atom3
  3     0     4
  1     0     3
  1     0     4
  1     0     2
  2     0     3
```

Adjust angle: 0 (0 30 60 90 120 150 180)

Move: ○ Group1 ⊖ Group2 ⊖ Both

Name: O1 Type: O1

Charge: 0.0 Element: O

Figure 4: Molefacture main window (left), and atom editing window (right).

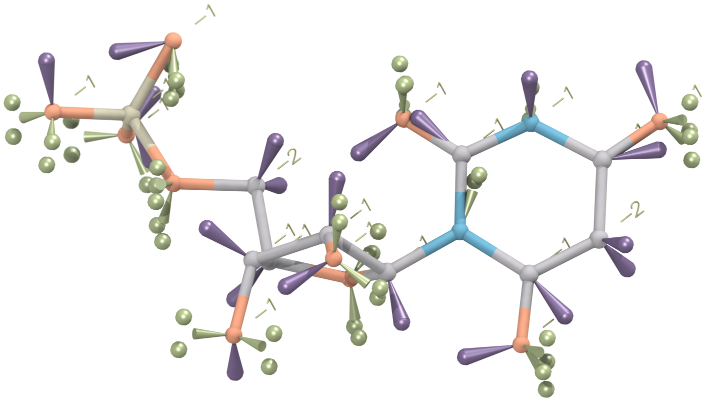


Figure 5: Ligand shown in main VMD window. Color scheme may vary depending on your presets.

**VMD Shortcuts.** Knowing a few keyboard shortcuts can really improve your VMD experience. While VMD's "OpenGL Display" window is selected, clicking the $p$ key activates the selection mode of your mouse (to "pick" atoms, get it?), so every time you click on an atom it is selected for the current activity you are performing, such as editing a molecule, atom-by-atom, in Molefacture. Clicking the $r$ key activate the "rotation" mode of your mouse, allowing you to spin the molecule in any direction. The $c$ key will make the mouse click select an atom to be the "center" of the system's rotations, and the $t$ key activates the "translation" mode, to move the system without spinning it.

**6** Select the “O1” oxygen atom. This can be done through VMD’s graphical interface by clicking on the atom, or by selecting it from the atom list in the main Molefacture window, inside the “Atoms” box.

**7** Click on **“Edit selected atom”** in Molefacture’s main window. A new (tiny) window will pop up for atom editing (Fig. 4).

**8** In the “Edit Atom” window, change the *name* of the atom to “C7”, the *type* to “C” and the *element* to “C”. Click on **Apply**.

**9** In Molefacture’s main window, make sure to have the (new) “C7” atom selected, and then click on “**Add hydrogen to selected atom**” two times.

**10** Select one of the hydrogen atoms you just added to “C7”, and then click on **“Edit selected atom”**.

**11** In the “Edit Atom” window, change the *name* of the atom to “OA”, the *type* to “O” and the *element* to “O”. Click on **Apply**.

**12** For the second hydrogen atom, change the *name* of the atom to **“OB”**, the *type* to “O” and the *element* to “O”. Click on **Apply**.

Now that the carboxylic group has been added to the molecule, we will add the missing hydrogen atoms along the rest of the molecule.

**13** Use Molefacture’s main window or VMD’s OpenGL interface to select the “C5” atom, and then click on “**Add hydrogen to selected atom**”.

**14** Repeat the same procedure for atoms (pay attention to the “prime” in the atoms’ name): N3, C1’, C2’, C3’, C4’, C5’ (add two hydrogens to C5’), O2’, O3’, OP2 and OP3.

**15** Use Molefacture’s main window or VMD’s OpenGL interface to select the “N3” atom, and then click on “**Force planar**”, to force a planar conformation for the nitrogen and its bonds.

After making the modifications and adding all hydrogen atoms, the structure should look like the one in Figure 6. The total charge of the ligand should read “-9”, which is reflective of the fact that we have not assigned any double bonds to the molecule. We do not need to assign double bonds, but you can do that in the “Bonds” section of Molefacture’s window, and get the total charge down to “-1”.

**16** In Molefacture’s main window, click in File → Write psf and pdb files. Use the prefix for output files “ligand”, and save it to the *setup* directory.

**17** When asked about the hydrogen atoms’ names, allow Molefacture to rename the atoms.

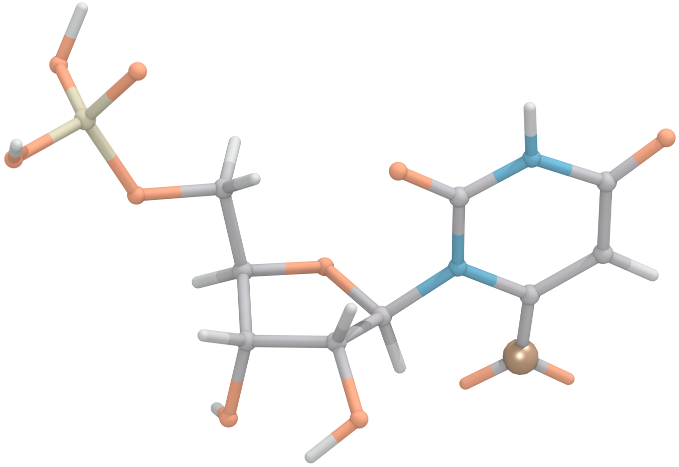


Figure 6: Ligand shown in main VMD window after all modifications. (Color scheme may vary depending on your presets)

### 1.2 Creating parameters for the ligand

In order to simulate the complete system in NAMD (protein and ligand in water), we need parameters for the ligand molecule we just created. One option is to use the CGenFF web service to automatically approximate CHAMRM36 parameters based on other molecules with similar chemical groups:

**1** Go to CgenFF web server at: https://cgenff.paramchem.org/

**2** If you do not have an account, create one now. It's free!

**3** Click in *Uploadmolecule* and select the *ligand.pdb* file. Make sure to check the box for "Guess bond orders from connectivity". Click on **Upload molecule**.

**4** Download the parameter file *ligand.str* and save it in the *toppar* folder.

## 2 Free Energy Calculations

From now on, we will use a jupyter notebook to combine the ligand and protein structure, and prepare the simulations. The chemical reaction in this tutorial can be studied using QM/MM MD simulations, but waiting for the system to randomly get to the perfect conformation, aligning the relevant atoms in just the right way, can take an unfeasible amount of time. That is why we must induce the reaction mechanism to happen using Steered Molecular Dynamics (SMD).

## 2.1 Preparing your Steered Molecular Dynamics (SMD) simulation

The tutorial jupyter notebook combines all procedures, code and comments related to the SMD. To open a jupyter notebook, open a command line window and navigate to the tutorial directory:

**1** cd `~/QMMM_String_eABF_Tutorial`

**2** Now run the following command: *jupyter notebook*

**3** Click on the notebook *Tutorial QMMM String eABF.ipynb* (this will open a new window)

**4** To make the tutorial easier to read and understand, some python code is hidden, and in order to do that, *you must execute the very first block of text* in the tutorial by clicking on it, and typing *Ctrl + Enter*

## 2.2 Notebook navigation and execution

Jupyter notebooks are composed of cells that can contain text and images, or source code that will be executed. You can select any cell at any point and execute it either by clicking the "Run" button, or by typing $Ctrl + Enter$. Typing $Shift$+*Enter* will execute the cell and automatically select the following cell.

# 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 scientific 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. Nature Methods (2018)

**Using collective variables to drive molecular dynamics simulations.** Giacomo Fiorin and Michael L. Klein and Jrme Hnin. Molecular Physics, **22**:3345-3362, 2013