AFM Tutorial:

Fitting a Model for Neat Cyclopentane Using AFM Tools and afmtogmx.

Raymond Weldon and Feng Wang

University of Arkansas

Fayetteville, AR 72701

This tutorial describes the key steps in fitting a model for neat cyclopentane without using partial charges. As reported in Sci. Rep. 999, 1234 (2025), for alkanes, it is sufficient to fit force field models without using partial charges, as the binding energy is dominated by dispersion. While it is possible to fit the model completely using AFM tools, the afmtogmx utility developed by Raymond Weldon greatly simplifies the creation of Gromacs input files from the CRYOFF .off file.

1. Prerequisites.

The tutorial assumes a basic understanding of the AFM workflow and the format of the `.ref` and `.ff` input files. The functionality of selected AFM tools used in this tutorial is described in the AFM Tools manual.

It also assumes proficiency in Linux and bash.

The tutorial assumes the cryoff program has been installed as `/usr/local/bin/cry.431.x`.

As the Intel compilers were used to create this copy of cry.431.x, the environment for the Intel compiler runtime must be set. Assuming the intel oneapi compilers are installed under /usr/local/intel/openapi, this can be accomplished with:

source /usr/local/intel/oneapi/setvars.sh intel64

Please set the intel environment before activating conda, (see below) as the intel python may interfere with the conda python.

2. Installing of the tools and set up of the execution environments.

a. Install conda

In this tutorial, we install conda in the /scr/fengwang/Raymond/anaconda3 directory. Details on installing anaconda and links for the conda installer can be found at https://docs.conda.io/projects/conda/en/latest/user-guide/install/linux.html.

wget https://repo.anaconda.com/archive/Anaconda3-2025.06-0-Linux-x86_64.sh

allow anaconda3 to be downloaded.

After it is downloaded, run

bash Anaconda3-2025.06-0-Linux-x86_64.sh -p /scr/fengwang/raymond/anaconda3

In this example, the directory conda installed is */scr/fengwang/raymond/anaconda3*, you may want to change this according to your system.

I choose not to let anaconda to modify my .bashrc file. In this case, one has to manually to create the conda hook.

eval "$(/scr/fengwang/raymond/anaconda3/bin/conda shell.bash hook)"

b. Install afmtogmx

cd ..

cd 02_Add_Ray_Package/

conda env create -f environment.yml

conda activate afmpython

cp -a afmtogmx /scr/fengwang/raymond/anaconda3/envs/afmpython/lib/python3.11/site-packages

c. set up AFM tools.

The AFM tools does not need to be installed. We will simply run it from the

01_Install_Software/03_AFMTools/AFMTools directory provided by the tutorial distribution package.

To do this, one has to set up the environment variables by

source 01_Install_Software/03_AFMTools/AFMTools/setenv.afm

Note that to use the execution environment later after installation, the commands in bold will have to be executed again.

3. Fitting of Dispersion.

As MP2 is used to fit the neat cyclopentane model, the best method is to fit dispersion computed with SAPT. The `total.ref` file in the `02_Fitting/dispersion` directory provides the SAPT E2 dispersion energies in the `fitE` line, followed by the coordinates of the cyclopentane dimer.

The corresponding dispersion.ff file instructs the `cryoff` program to fit energy (`fitE`) and use ridge regression. The λ factor for ridge regression was determined by a fit without the `ridge` keyword using the default SVD algorithm. `Cryoff` will print out all the singular values of the fit. The ridge λ is set to be the square of 50% of the smallest singular value, which is 4.1889E-13 in this case.

It is worth pointing out that the EXClusion section sets all intramolecular pairwise interactions to be excluded, as SAPT only produces intermolecular energies. As shown in the [POW] section, the fit is performed with a combination of $1/r^6$ and $1/r^8$ terms.

The details for creating the .ff and .ref files will not be provided in this tutorial.

The fit can be performed with the `runme.sh` file.

The .off file provides important information such as the RMSE of the fit and a linear regression of the fitted and reference energies. The final parameters are printed at the end of the file.

Force Field Potential:

Intra-Potential:

Inter-Potential:

C~C: POW	-885.37631	-6.0000000	Min:	6.514857	Max:	18.551504
C~C: POW	-14109.008	-8.0000000	Min:	6.514857	Max:	18.551504

Showing the C6 and C8 parameters being 885.37631 and -14109.008 in default AFM unit of kcal/mol for energy and Å for distance.

4. AFM Iterations

The 298K directory provides an example of creating a reference file based on one MD sampling run.

From the MD simulation in this directory, 120 frames were extracted and placed in the 01_grofiles directory. The files with the coordinates of atoms in the QM region, in AFM tools pxyz format, are in the 02_pxyz_files directory. ORCA calculations of forces are in the 03_force_calculations directory, with the final reference file assembled in 04_ref_files.

The .ref files can be concatenated to create the final .ref file for the fit.

The shell or Python scripts used to create these files are as follows:

01_sample_gro.sh

02_gen_pxyz.py

03_gen_force_input.sh

04_gen_ref.py

a. The 01_sample_gro.sh

This script writes a series of .gro files with the molecules at the box boundary made whole.

An atom (TAR) is inserted near the center of the box, around which a random selection of the first molecule will be made.

b. 02_gen_pxyz.py

This script generates the pxyz file containing 6 cyclopentane molecules, as described in the paper. This is accomplished by assigning each atom a mark. Generally, the fitting zone of the QM region will have a higher mark, the MM region a lower mark, and atoms to be discarded will have the lowest mark. In this particular exercise, there is no MM region, and all QM atoms are being fit. Note that the `mark_byrange` script and a random number generator can also be used to pick the first molecule.

Inserting a TAR atom at the center of the box avoids the need to center the box around the randomly selected molecule. The TAR atom was marked as -1 and removed.

c. 03_gen_force_input.sh

This script update a template ORCA input file ('template') with conformations from each pxyz file to create orca input files for the 120 ORCA calculations.

d. 04_gen_ref.py

This script assembles the .ref file for each conformation and compute net force and net torque.

5. Intermolecular Fit

The input files for the intermolecular fit can be found in the 02_Fitting/inter directory.

The total.ref file contains 1200 frames, which corresponds to the global fit.

The inter.ff file has blank dispersion parameters. While these parameters can be updated with the off2ff script in AFM tools, it is straightforward to edit this file manually, as it only contains two parameters.

Change the SRD lines to:

[SRD] 2

C C FIX -885.37631 -6 2.045

C C FIX -14109.008 -8 2.045

The fit can be performed by running the commands in the runme.sh file.

After the fit is completed, the inter.off file contains the intermolecular parameters and related quality of fit metrics. To start the intramolecular fit, the intra.ff file must be created using intermolecular parameters from the intermolecular fit. The easiest way to accomplish this is to use the off2ff tool, which takes directives from the proto.copy file, parameters from the inter.off file, and creates the intra.off file using the template_intra.ff file.

This can be accomplished by running:

source postfit

Successful extraction of intermolecular parameters from the inter.off file with the postfit script will produce an intra.ff file, which can be copied to the 02_Fitting/intra directory for the intramolecular fit.

6. Intramolecular Fit

The total.ref file for the intramolecular fit is identical to that of the intermolecular fit.

However, the intra.ff file defines intramolecular terms and requests an intramolecular fit with the 'intra' keyword. The intramolecular fit fits atomic forces rather than total forces and total torques.

The intramolecular fit is accomplished with the runme script in the 02_Fitting/intra directory.

7. Creating the Input File for Gromacs

With the intra.off file, the gen_md_input.py script in the 02_Fitting/tabpot_topology directory will create the Gromacs input file topol.top and the associated tabulated potential files for running Gromacs with cyclopentane.

The gen_md_input.py script uses the afmtogmx module created by Raymond Weldon, the help pages for afmtogmx can be accessed with the python help function.

python

import afmtogmx

help(afmtogmx)

help(afmtogmx.function)