ProDy

Protein Dynamics & Sequence Analysis

NMWIT Tutorial

Release

Ahmet Bakan, Cihan Kaya

May 17, 2024

CONTENTS

1 Normal Mode Wizard **1**
- 1.1 Required Programs 1
- 1.2 Getting Started 1

2 Using NMWiz **3**
- 2.1 Main Window 3
- 2.2 NMD Format 3
- 2.3 NMWiz Window 5

3 Principal Component Analysis **9**
- 3.1 Atom Selection 9
- 3.2 PCA Calculation 9
- 3.3 ANM Calculation 9
- 3.4 Comparison 10

4 Essential Dynamics Analysis **11**
- 4.1 EDA Calculation 11
- 4.2 ANM Calculation 11
- 4.3 Comparison 11
- 4.4 Suggestions 11

5 Extending a Model **13**
- 5.1 ANM Calculation 14
- 5.2 Visualization 14
- 5.3 Animation 14

6 NMWiz Commands **15**
- 6.1 Dataset handles 15
- 6.2 Combining modes 15
- 6.3 Inverting a mode 16
- 6.4 Changing coordinates 16

Bibliography **17**

CHAPTER

ONE

NORMAL MODE WIZARD

Normal Mode Wizard (NMWiz) is a VMD plugin for depiction, animation, and comparative analysis of normal modes. Normal modes may come from principal component of structural ensembles, essential dynamics analysis of simulation trajectories, or normal mode analysis of protein structures. In addition, NMWiz can be used to depict any vector that describes a molecular motion.

1.1 Required Programs

The latest versions of **ProDy_** (1.4 or later) and **VMD_** (1.9.2 or later) are required.

1.2 Getting Started

To follow this tutorial, you will need the following files:

```
 12M Feb 29 20:20 1dlo_anm_cross-correlations.txt
224K Feb 29 20:20 1dlo_anm.nmd
661K Feb 29 20:20 1dlo.pdb
3.3M Feb 29 20:20 p38_100frames.dcd
219K Feb 29 20:20 p38.pdb
```

CHAPTER

TWO

USING NMWIZ

This chapter provides an introduction to NMWiz and its capabilities.

2.1 Main Window

NMWiz is located in *VMD Main* → *Extensions* → *Analysis* menu and labeled as *Normal Mode Wizard*, as shown below:

When you first open NMWiz, you will see the *NMWiz 1.2 - Main* window shown below:

NMWiz 1.2 - Main window offers options to load precalculated data or to generate data for molecules loaded in VMD:

- *Load NMD File* lets you load precalculated data (`.nmd` files) for analysis.
- *From Molecule* lets you load precalculated data stored in molecular data formats recognized by VMD.
- *ProDy interface* facilitates ANM, GNM, and PCA/EDA calculations from VMD. This functionality is illustrated in the following chapters.
- *Structure Comparison* facilitates calculating, visualizing, and animating deformations between two frames of the same molecule or between two separate structures.

2.2 NMD Format

NMD files (extension `.nmd`) are plain text files that contain at least normal mode and coordinate data. In addition to PCA, EDA, NMA, ANM, or GNM data, arbitrary vectors can be stored in NMD files. The following data fields are recognized:

- **coordinates**: Coordinates must be provided in one line as a list of decimal numbers. The number of atoms in the system is deduced from size of this data line.
- **mode**: Normal mode array. Each normal mode array must be provided in one line as a list of decimal numbers. Mode array may be preceded by mode index and mode length (square root of variance or inverse frequency).
- **title**: A title for the dataset.
- **names**: Atom names. Default is “CA” for all atoms.
- **resnames**: Residue names. Default value is “GLY”.
- **chainids**: Chain identifiers. Default value is “A”.
- **resids**: Residue numbers. If this data line if not found, residue numbers are assigned starting from 1 and incremented by one for each atom.
- **betas**: Beta factors. Default value is 0 (zero). B-factors are used to color the protein representation.

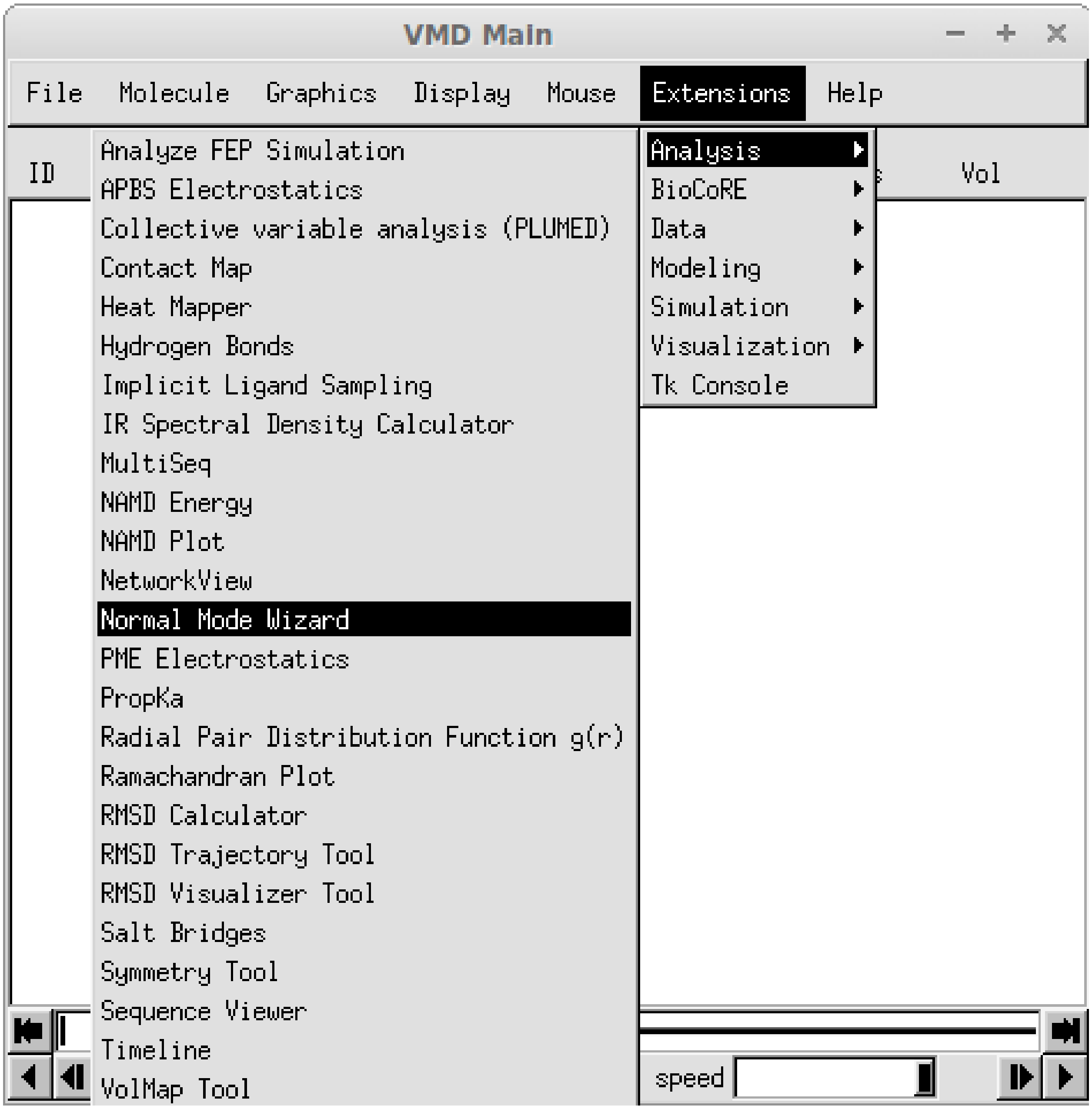

VMD Main
File
Molecule
Graphics
Display
Mouse
Extensions
Help
ID
Vol
Analyze FEP Simulation
APBS Electrostatics
Collective variable analysis (PLUMED)
Contact Map
Heat Mapper
Hydrogen Bonds
Implicit Ligand Sampling
IR Spectral Density Calculator
MultiSeq
NAMD Energy
NAMD Plot
NetworkView
Normal Mode Wizard
PME Electrostatics
PropKa
Radial Pair Distribution Function g(r)
Ramachandran Plot
RMSD Calculator
RMSD Trajectory Tool
RMSD Visualizer Tool
Salt Bridges
Symmetry Tool
Sequence Viewer
Timeline
VolMap Tool
Analysis
BioCoRE
Data
Modeling
Simulation
Visualization
Tk Console
speed

Fig. 2.1: Location of NMWiz plugin in VMD menus.

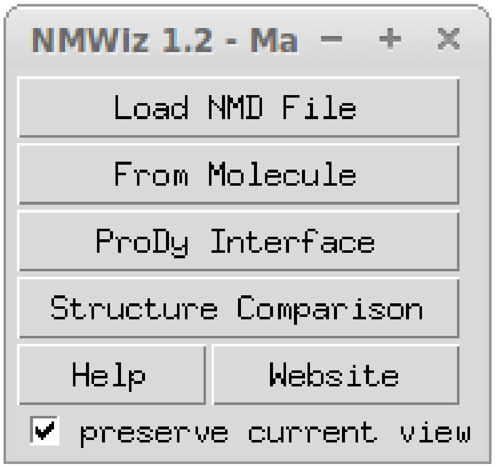


Fig. 2.2: NMWiz Main window.

Note: All data types must be listed in a single line. The size of data lines must match the number of atoms in the system, i.e. the size of coordinates line.

For example, `1dlo_anm.nmd` contains following lines:

```
nmwiz_load 1dlo_anm.nmd
name 1dlo
atomnames CA CA CA CA CA CA CA CA CA CA ...
resnames PRO ILE SER PRO ILE GLU THR VAL PRO VAL ...
resids 1 2 3 4 5 6 7 8 9 10 ...
chainids A A A A A A A A A A ...
bfactors 71.47 63.68 46.35 42.28 36.52 31.66 20.85 18.49 24.94 29.01 ...
coordinates 172.541 -33.818 72.141 172.591 -29.981 71.748 171.498 -29.271 ...
mode 1 5.08 0.010 -0.014 -0.030 0.011 -0.013 -0.029 0.006 -0.012 -0.027 ...
mode 2 3.99 -0.003 -0.014 0.007 0.001 -0.016 0.002 0.001 -0.019 0.000 ...
mode 3 2.82 0.010 0.027 0.000 0.006 0.028 0.003 0.005 0.028 0.004 0.002 ...
...
```

The first line in this file contains `nmwiz_load 1dlo_anm.nmd` command. With this, you can make **`vmd`** load this file first thing when it starts:

```
$ vmd -e 1dlo_anm.nmd
```

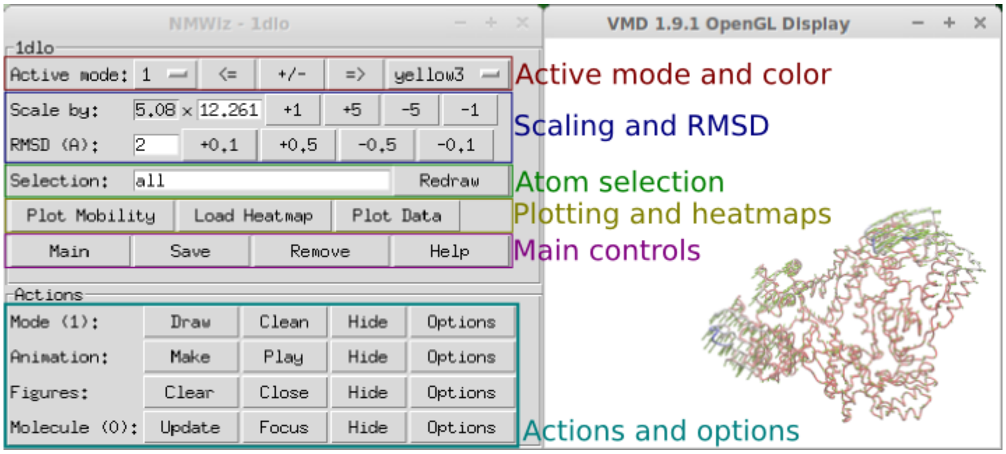
Note: Any data field name that is not recognized by NMWiz, such as `nmwiz_load 1dlo_anm.nmd` will be omitted.

2.3 NMWiz Window

Now, let's load `1dlo_anm.nmd` file to introduced the *NMWiz* window:

2.3.1 Changing and Editing Modes

The active mode is represented by arrows starting at the atom locations and pointing in the direction of motion. The top pane of the NMWiz window contains tools for changing some of the basic features of the displayed mode.



Active mode: 1

Scale by: 5.08 x 12.261

RMSD (A): 2

Selection: all

NMWiz - 1dlo
1dlo
<=
+/-
=>
yellow3
+1
+5
-5
-1
+0.1
+0.5
-0.5
-0.1
Redraw
Plot Mobility
Load Heatmap
Plot Data
Main
Save
Remove
Help
Actions
Mode (1): Draw Clean Hide Options
Animation: Make Play Hide Options
Figures: Clear Close Hide Options
Molecule (0): Update Focus Hide Options
VMD 1.9.1 OpenGL Display
Active mode and color
Scaling and RMSD
Atom selection
Plotting and heatmaps
Main controls
Actions and options

Fig. 2.3: NMWiz GUI

Active mode and color

The active mode can be changed using the drop down menu or incremented using the <= and => buttons. Changing the active mode will hide previously generated depictions, and the newly depicted mode will be assigned a distinct color.

Mode scale and direction

Mathematically, each mode's sign is arbitrary. NMWiz depicts normal modes as they are read from the input file, but the +/- button changes the mode's sign, inverting its direction. The size of the arrows can be adjusted using the *Scale by* or *RMSD* controls. * **Scale by**: The first number is the inverse of the mode frequency, and the second is a real-valued scaling factor that may be incremented or decremented using the adjacent buttons. The +/- button controls the sign of the scale factor. By default, NMWiz scales all modes to 2 Angstroms RMSD. * **RMSD**: Arrow length can be adjusted to show a specified displacement along the active mode. The sum of squares of the arrow lengths will be the square of the specified RMSD.

Atom selections

Arrows are drawn for selected atoms. You can enter any valid atom selection string to display arrows for a specific part of the molecule.

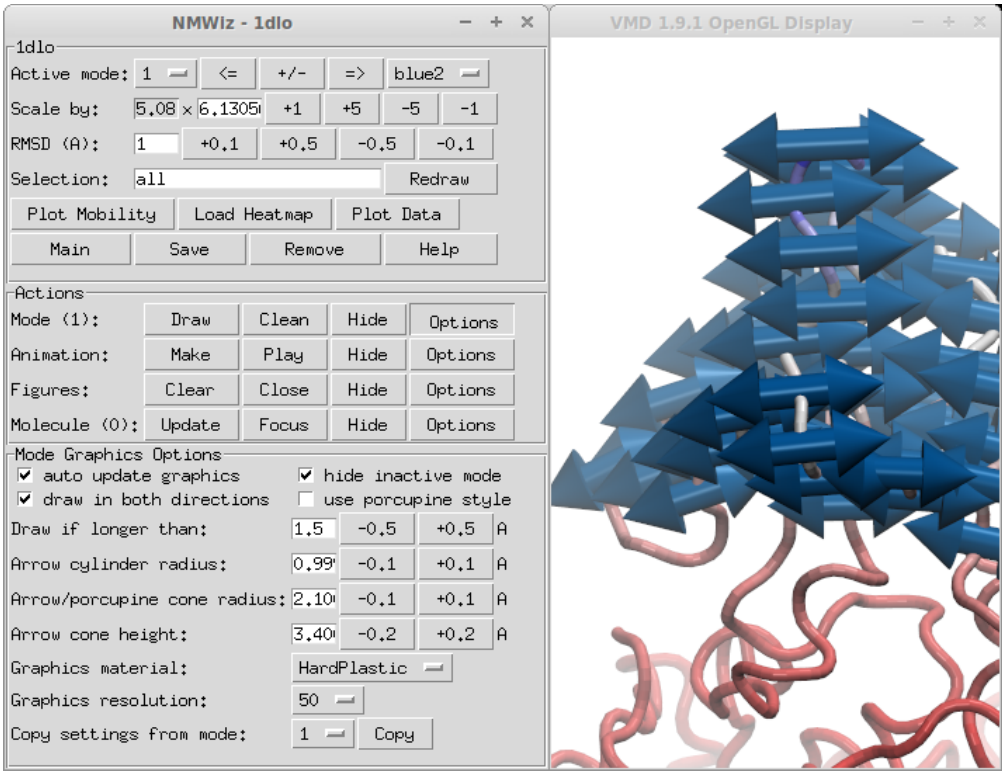
Mode options

Mode Graphics Options panel allows for customizing the look of arrows. You can change the size, length, material, and resolution of arrows, and also select to use porcupine quills instead of arrows.

Note: Mode representation is not redrawn after manual changes in variables. When this is the case click one of *Draw* or *Redraw* buttons.

2.3.2 Representations

You can change molecule representations through NMWiz. The *Molecule* number in the *Actions* panel shows the index of the molecule that the changes will affect. If you would like NMWiz not to interfere with your customized visualization of this molecule, you can turn this off in the molecule *Options* panel. Alternatively, you can *Hide* this molecule, and make custom representations for a molecule that you load into VMD manually.


NMWiz - 1dlo

1dlo

Active mode: 1 blue2

Scale by: 5.08 x 6.1305

RMSD (A): 1

Selection: all

Mode Graphics Options

☑ auto update graphics
☑ hide inactive mode
☑ draw in both directions
☐ use porcupine style

Draw if longer than: 1.5 A

Arrow cylinder radius: 0.99 A

Arrow/porcupine cone radius: 2.10 A

Arrow cone height: 3.40 A

Graphics material: HardPlastic

Graphics resolution: 50

Copy settings from mode: 1

<=
+/-
=>
+1
+5
-5
-1
+0.1
+0.5
-0.5
-0.1
Redraw
Plot Mobility
Load Heatmap
Plot Data
Main
Save
Remove
Help
Actions
Mode (1):
Draw
Clean
Hide
Options
Animation:
Make
Play
Hide
Options
Figures:
Clear
Close
Hide
Options
Molecule (0):
Update
Focus
Hide
Options
-0.5
+0.5
-0.1
+0.1
-0.1
+0.1
-0.2
+0.2
Copy
VMD 1.9.1 OpenGL Display

Fig. 2.4: Mode Graphics Options panel

2.3.3 Animations

Clicking *Animation* → *Make* will start an animation along the active mode. The animated motion changes the structure along the mode in both directions with an amplitude given by the RMSD displayed in top panel.

You can adjust animation settings, such as number of frames, using the *Animation* → *Options* panel.

2.3.4 Plots and Heatmaps

You can generate plots and heatmaps that are linked to the protein display. Click *Plot Mobility*, for example. You will see a plot of the residue mean squared fluctuations for the active mode. Clicking the points on the plot will toggle labels for the corresponding atoms in the VMD molecular display.

2.3.5 Main Controls

Use the following controls for:

- *Main* - Open the NMWiz main window, useful when it's closed
- *Save* - Save current dataset in NMWiz format
- *Remove* - Remove current dataset and free memory
- *Help* - Get some help

2.3.6 Efficiency Tips

Displaying large number of primitive graphics for large molecules may make your system unresponsive. You can try some of the following tricks to speed up drawing arrows and manipulating the view.

Show Fewer Arrows

- Mode visualizations for large molecules may become cluttered with arrows. You can show a subset of arrows evenly distributed over the molecular system by making a selection using `%` and entering it into the *Selection* box:
  - `index % 2 == 0` - shows every other arrow
  - `resid % 4 == 0` - shows arrow for every fourth residue
- To only draw arrows that are longer than a user-specified threshold length, use the options in *Mode Graphics Options*, shown in Figure 2.4.

Depending on the molecule one or both of these tricks can provide a more responsive experience and a better (less crowded) depiction of motions.

CHAPTER

THREE

PRINCIPAL COMPONENT ANALYSIS

We will use ProDy Interface of NMWiz plugin to perform a comparative analysis of ubiquitin dynamics predicted using theory using anisotropic network model (ANM) and inferred from experimental structures using principal component analysis (PCA).

Let's fetch a PDB files that contains a solution (NMR) structure using prody fetch[1] command:

```
$ prody fetch 2k39
```

Alternatively, you may use a browser to download these files. **:pdb:'2k39'** file can be loaded into VMD at startup as follows:

```
$ vmd 2k39.pdb
```

This file contains 116 solution structure models of the protein Ubiquitin, so it will be suitable for comparison of theoretical predictions with experimental observations.

NMWiz - Main window can be accessed via the *Extensions* → *Analysis* menu of *VMD Main* window. Click on *ProDy Interface* and a window listing molecules and ProDy jobs will pop up.

3.1 Atom Selection

In both ANM and PCA calculations, we will omit the flexible C-terminal tail of ubiquitin, since incorporating this flexible tail skews the calculations and results highlight local tail motions instead of collective global motions that we are interested in.

We enter the selection string `resid <= 70 and protein and name CA` and click *Select*. 70 Cα atoms will be selected for calculations.

3.2 PCA Calculation

Now, we will use all ubiquitin models for the PCA calculation. Select *PCA Calculation* in the *ProDy Job Settings* panel. Keeping *PCA (EDA) Settings* unchanged, click *Submit Job*. Results will be loaded automatically when prody pca[2] command finishes the calculations.

3.3 ANM Calculation

First, we perform ANM calculations, which is the job type selected by default. We will use 78th NMR model for ANM calculations, because this model is the closest to the average of the ubiquitin NMR ensemble. Since VMD frame (model) numbers start from 0, we will change the *Frame number* to read 77.

---

[1] http://prody.csb.pitt.edu/manual/apps/prody/fetch.html#prody-fetch

[2] http://prody.csb.pitt.edu/manual/apps/prody/pca.html#prody-pca

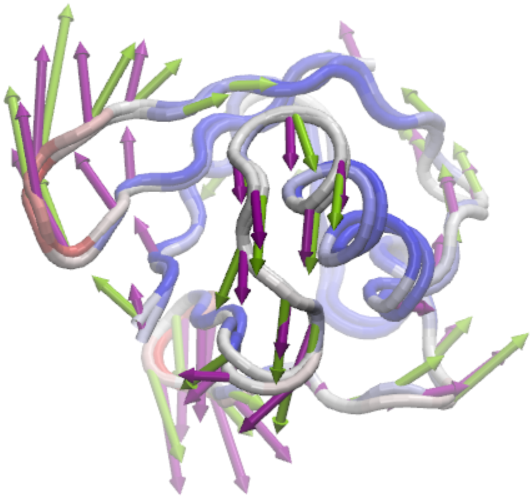


Fig. 3.1: PC 2 (purple) vs. ANM mode 2 (lime green)

We keep the rest of the parameters unchanged, and click *Submit Job*. This action will run prody anm[3] command in the background and load the results when calculations are finished.

3.4 Comparison

For each dataset you load into or generate via NMWiz, a GUI will pop up with independent controls for normal mode display, animation, and plotting. Select PC 2 and ANM mode 2 and try to get the view in the image in VMD display.

Clicking on *Options* buttons will expand the GUI showing more controls for fine tuning the display. Exploring these options is left to you. You can adjust the quality of the arrows and protein to get a publication quality image. In addition, you can make mobility plots and generate animations for the selected mode.

---

[3] http://prody.csb.pitt.edu/manual/apps/prody/anm.html#prody-anm

CHAPTER

FOUR

ESSENTIAL DYNAMICS ANALYSIS

In this part, we will perform essential dynamics analysis calculations for a p38 MAP kinase trajectory and compare results with normal modes from ANM calculations. You will need following files:

- NMWiz Tutorial files (ZIP)
- NMWiz Tutorial files (TGZ)

This files contain frame from a 40 ns long simulation of p38. Detailed analysis of this trajectory can be found in *[AB11]*.

Load the contents of this file into VMD as follows:

```
$ vmd p38.pdb p38_100frames.dcd
```

Click on *ProDy Interface* for performing ANM and EDA jobs.

4.1 EDA Calculation

Select *PCA Calculation* in the *ProDy Job Settings* panel. Set *First frame* 1 to exclude the X-ray coordinates from the calculation. You may also check *aligned* to make ProDy skip alignment step in PCA/EDA calculations when you already have done the alignment in VMD. In this case, the frames are already aligned. Click *Submit Job* and results will be loaded automatically when prody pca[4] command finishes the calculations.

4.2 ANM Calculation

We will perform ANM calculations for all Cα atoms and keep the rest of the parameters unchanged. To properly align the structures used for EDA and ANM, set *First frame* to 100 for the ANM calculation. Click *Submit Job* and results obtained from prody anm[5] command will load automatically.

4.3 Comparison

For each dataset you load into or generate via NMWiz, a GUI will pop up with independent controls for normal mode display, animation, and plotting. Select PC 1 and ANM mode 1 and try to get the view in the image in VMD display.

4.4 Suggestions

NMWiz writes a DCD or PDB file for PCA/EDA calculations. For large systems and long trajectories you may try one or more of the following for speedier calculations:

---

[4] http://prody.csb.pitt.edu/manual/apps/prody/pca.html#prody-pca

[5] http://prody.csb.pitt.edu/manual/apps/prody/anm.html#prody-anm

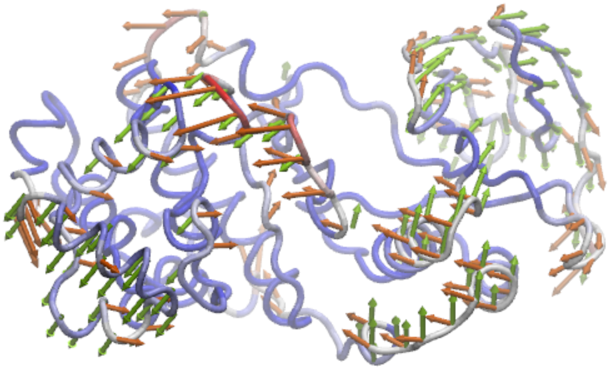


Fig. 4.1: EDA 1 (orange) vs. ANM mode 1 (lime green)

- select a subset of atoms, e.g. Cα atoms
- select a subset of frames, e.g. set *Skip frame* a value greater than 0
- use *DCD file* for faster IO operations and less disk usage
- alternatively, if you have trajectories in DCD format, use prody pca[6] directly to obtain results in NMD Format[7]

---

[6] http://prody.csb.pitt.edu/manual/apps/prody/pca.html#prody-pca

[7] http://prody.csb.pitt.edu/manual/reference/dynamics/nmdfile.html#nmd-format

CHAPTER

FIVE

EXTENDING A MODEL

In the previous calculations, we used only Cα atoms for calculating normal modes. Publication-quality graphics that use VMD's cartoon or ribbon representations require a least the full protein backbone. In this part, we will use NMWiz's *Extend model to* capabilities to extend NMA results to all backbone atoms of the protein.

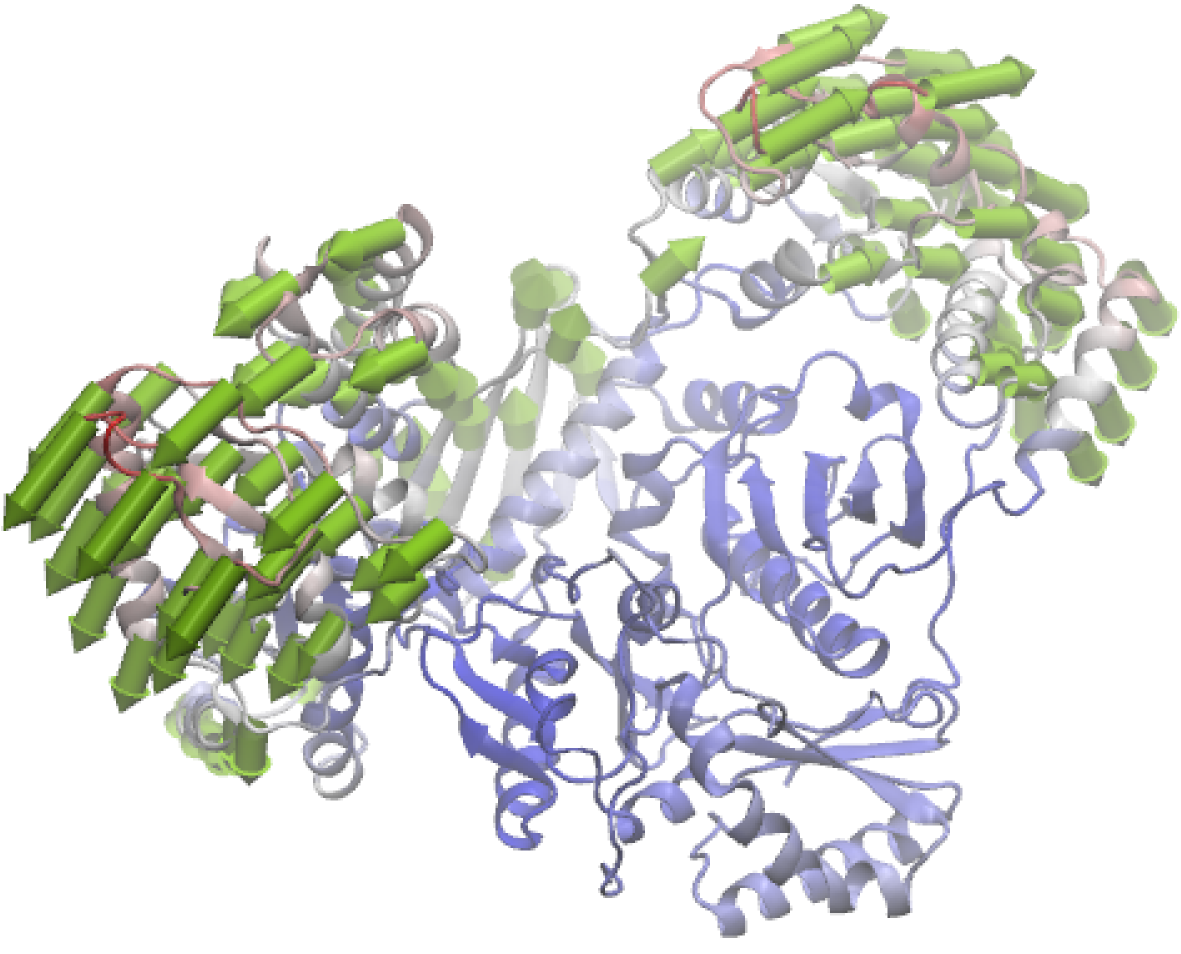


Fig. 5.1: ANM mode 1 for HIV Reverse Transcriptase

5.1 ANM Calculation

Let's fetch an X-ray structure of the protein HIV reverse transcriptase (RT) and load into VMD:

```
$ prody fetch 1dlo
$ vmd 1dlo.pdb
```

In the *ProDy Interface*, we select *ANM Calculation*, check *backbone* option, and click *Submit Job*. Model will be calculated for 971 selected Cα atoms, but the normal modes will be extended to all backbone atoms.

5.2 Visualization

When the results are loaded, you will see four arrows per residue (or node). Change the *Selection* string to read `name CA` and click *Redraw*. This will draw only one arrow per residue.

RT is a large structure and updating the display with every little change you make might be time consuming. You can uncheck *auto update graphics* option in *Mode Graphics Options* panel to speed up rendering.

To get the view displayed in the figure, you will need to hide arrows that are shorter than a given length using *Draw if longer than* option in the *Model->Options* menu. Draw an arrow for every forth residue using the selection `name CA and resid % 4 == 0`. The protein representation is *NewCartoon*.

5.3 Animation

You can generate a trajectory along the selected mode by clicking *Make* in *Animation* row. For large proteins, keeping the *Graphics resolution* low (10) will make the animation run smoother.

CHAPTER

SIX

NMWIZ COMMANDS

NMWiz commands allow altering molecule coordinate and normal modes, as well as adding new modes. Commands can also be used to generate combinations of modes and add them to the dataset (see the example at the end). These commands would be entered in the Tk Console[8] of **VMD_**.

Following are the level commands:

- `nmwiz list` - lists all normal mode dataset handles
- `nmwiz load` - loads a dataset in NMD format
- `nmwiz main` - shows NMWiz main window

6.1 Dataset handles

A handle of a dataset as follows:

```
set nmdhandle [nmwiz load p38_MAPK_1p38_anm.nmd]
```

Or you can use the list command to get a handle to the dataset at given index:

```
set nmdhandle [lindex [nmwiz list] 0]
```

This handle can provides the following commands:

- `$nmdhandle numatoms` - get number of atoms
- `$nmdhandle getcoords` - get molecule coordinates
- `$nmdhandle setcoords array` - set molecule coordinates
- `$nmdhandle nummodes` - get number of modes
- `$nmdhandle getmode index` - get mode array
- `$nmdhandle setmode index array` - set mode array
- `$nmdhandle getlen index` - get length of mode
- `$nmdhandle setlen index length` - set length of mode
- `$nmdhandle addmode array` - add new mode

6.2 Combining modes

Two mode arrays can be combined linearly to generate a new mode as follows:

---

[8] http://www.ks.uiuc.edu/Training/TutorialsOverview/vmd/tutorial-html/node4.html

```
set nmdhandle [lindex [nmwiz list] end]
set m1 [$nmdhandle getmode 1]
set l1 [$nmdhandle getlen 1]
set m2 [$nmdhandle getmode 2]
set l2 [$nmdhandle getlen 2]
$nmdhandle addmode [vecadd [vecscale $l1 $m1] [vecscale $l2 $m2]]
```

Note: Note that we used VMD vector subroutines for array operations. See VMD user guide for list of all vector and matrix subroutines which may be useful for manipulating normal mode arrays.

New mode will be available immediately in the NMWiz GUI and can be used in depictions and for animations.

6.3 Inverting a mode

Inverting a mode array and adding to the dataset is can be performed as follows:

```
set nmdhandle [lindex [nmwiz list] end]
set m1 [$nmdhandle getmode 1]
$nmdhandle setmode 1 [vecinvert $m1]
```

6.4 Changing coordinates

It is also possible to change the coordinates of the molecule. New or updated normal mode depictions will originate from the new coordinates. Below, we place all atoms of the molecule to the origin:

```
set nmdhandle [lindex [nmwiz list] end]
set zeros [vecscale 0 [$nmdhandle getcoords]]
$nmdhandle setcoords $zeros
```

In fact, NMWiz logo shown below was generated using this dataset with above change.

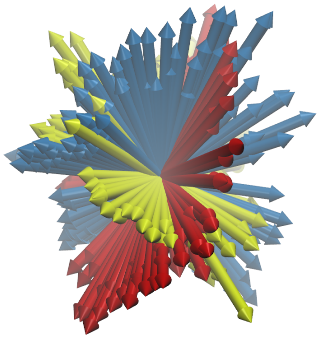


Fig. 6.1: Slowest three ANM modes for p38 generated by setting molecule coordinates to zero.

Acknowledgments

Continued development of Protein Dynamics Software *ProDy* and associated programs is partially supported by the NIH[9]-funded Biomedical Technology and Research Center (BTRC) on *High Performance Computing for Multiscale Modeling of Biological Systems* (MMBios[10]) (P41 GM103712).

---

[9] http://www.nih.gov/

[10] http://mmbios.org/

[AB11] Bakan A, Bahar I. Computational generation of inhibitor-bound conformers of p38 MAP kinase and comparison with experiments. *Pacific Symposium on Biocomputing* **2011** 16 181-192.