

NOTE: This PDF file is for reference purposes only. This lab should be accessed directly from the web at [https:// sassie-web.chem.utk.edu/docs/simulate/prody/prody.html](https://sassie-web.chem.utk.edu/docs/simulate/prody/prody.html).

[Return to Main Documents Page](#)

## ProDy

Generates structures by performing molecular normal mode analysis of input structure using the open-source molecular dynamics package [ProDy](#).

---

### Accessibility

The ProDy module is accessible from the [Beta](#) section of the main menu.

The ProDy module is accessible from the [Simulate](#) section of the main menu.

---

### Basic Usage

The purpose of the module is perform generate normal mode trajectories of an input protein structure.

---

### Notes

- The input structure must be a complete structure without missing residues. Atom and residue naming should be compatible with those defined in the CHARMM force field for subsequent usage with certain other SASSIE modules. See [Notes on Starting Structures and Force Fields](#) and [PDB Scan](#) for further details.
  - For a system with N atoms, the 3N - 6 vibrational normal modes can be calculated. The first 6 represent translational and rotational degrees of freedom and are not included.
  - The output file format for the trajectories along each normal mode ([ProDy\\_sampleModes](#)) is DCD since in most cases many structures are generated. One can use [Extract Utilities](#) to convert DCD files to multi-frame PDB files. The trajectories are merged starting with 1st (non-trivial) normal mode and ending on last normal mode. Individual trajectories along each mode also appear in output. The DCD file for the *merged* trajectories is in XPLO format, and cannot, in general, be read by ProDy (as is) for any subsequent (advanced input) usage.
  - The program uses the anisotropic network model (ANM) in ProDy to calculate normal modes. Default selections for the node atoms are CA atom for proteins; C2 atom for nucleic acids; C2 atom for sugars; Fe atom for hemes; C1 atom for lipids; ions; and O atom for water. User is referred to [ProDy\\_AtomFlags](#) for a list of acceptable ProDy residue names that fall under the above mentioned categories. All residues in the input PDB file should be named according to the ProDy convention, or removed from the input PDB file. The program will convert nucleotide names to Prody acceptable format for the user. Other residues must be named manually, and, if necessary (see first note above), named to be compatible with the CHARMM force field. For example, the ProDy program will execute with either STE or STEA lipid residue names, but only STEA is CHARMM compatible.
  - The output file format for the ensemble of randomly sampled conformations from the normal modes ([ProDy\\_sampleModes](#)) is provided in both DCD and PDB format.
  - NOT YET IMPLEMENTED: An Advanced Input option to enter a prody application command ([ProDy\\_Applications](#)). The command must begin with either *prody* or *evol*, followed by the desired subcommand (eg. *prody pca*; *evol fetch*), followed by necessary options.
  - NOTE YET IMPLEMENTED: Advanced Input options to reject structures based on Rg value, position of atoms in the Z-direction and via atomic constraints provided as a list in a text file as described in [Constraints](#). These options are not mutually exclusive and can be used in the same run as needed.
  - It is likely necessary and advisable to apply [Energy Minimization](#) to structures within the DCD files representing trajectories along the modes, or the ensemble of randomly sampled modes, to remove instances of physically unrealistic bond stretches that appear following the ProDy ANM calculation. As many as 5000 to 10000 steps of minimization may be needed for this.
  - A typical workflow involves generating trajectories along the requested number of normal modes of the input structure using this module, then energy minimizing the ensemble using [Energy Minimization](#), then calculating scattering from the trajectory(s) (or randomly sampled ensemble) using modules in [Calculate](#), and finally, comparing results to experimental data using modules in [Analyze](#).
- 

### Screen Shots and Description of Input Fields

---

This example generates a set of 5 normal modes using the ProDy ANM module for the lysozyme protein.

---

**Prody**

run name: run\_0

reference pdb:  lysozyme.pdb or  Local: lysozyme.pdb

number of normal modes: 5

number of conformations per normal mode: 50

number of frames to traverse for each mode: 10

average sampled RMSD: 1.0

average sampled RMSD for normal mode trajectory: 1.5

Advanced Input

Check Box for Advanced Input:

FEEDBACK  
DOCS

- **run name** User defined name of folder that will contain the results.
- **reference pdb** PDB file with naming information and coordinates of the structure.
- **number of normal modes** Number of normal modes to calculate.
- **number of conformations from sampled normal modes** Number of conformations to generate by random sampling of normal modes ([ProDy\\_sampleModes](#)).
- **number of frames to traverse for each mode** Number of steps to traverse each mode in both directions ([ProDy\\_sampleModes](#)).
- **average sampled RMSD** Average RMSD of randomly sampled conformations with respect to initial conformation ([ProDy\\_sampleModes](#)).
- **maximum RMSD for normal mode trajectory** Maximum RMSD of conformations for trajectory from traversed mode with respect to initial conformation ([ProDy\\_sampleModes](#)).

### Example Output

```

=====
DATA FROM RUN: Thu Sep 1 14:34:03 2016

Configurations and statistics saved in run_0/prody/ directory

DCD data were written to run_0/prody/run_0.dcd

=====

progress:
percent done: 100.0

```

Results are written to a new directory within the given "run name" as noted in the output.

Several files are generated and saved to the "run name" prody directory: the output DCD file containing merged trajectories for modes 1 to 5; separate DCD files for individual trajectories along each mode; a DCD and PDB file with an ensemble of randomly-sampled conformations from each of the normal modes; a file containing the Kirchhoff matrix; a file containing the hessian matrix; a file with ANM results in NMD format for subsequent visualization; two additional files with ANM results in NMD format with normal modes extended to back-bone and all atoms, respectively; two text files with eigenvalues and eigenvectors, respectively; a text file containing square fluctuations along each mode; a text file with cross-correlations; a text file with the covariance matrix constructed using the 5 modes; and a text file with the  $\beta$ -value (temperature factor) of the course-grain atom on each residue.

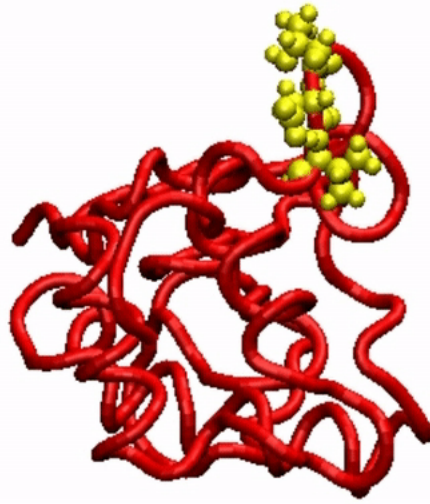
---

## Visualization

The first image below shows an animation of the first (lowest frequency) normal mode. The image was generated by uploading the file with normal modes extended to all atoms using the VMD Normal Mode Wizard. A short step-by-step guide to creating this image can be found on the [VMD visualization/animation of normal modes](#) page. A more detailed tutorial on the use of VMD's Normal Mode Wizard can be found on the [ProDy\\_NMWiz tutorial](#) site.

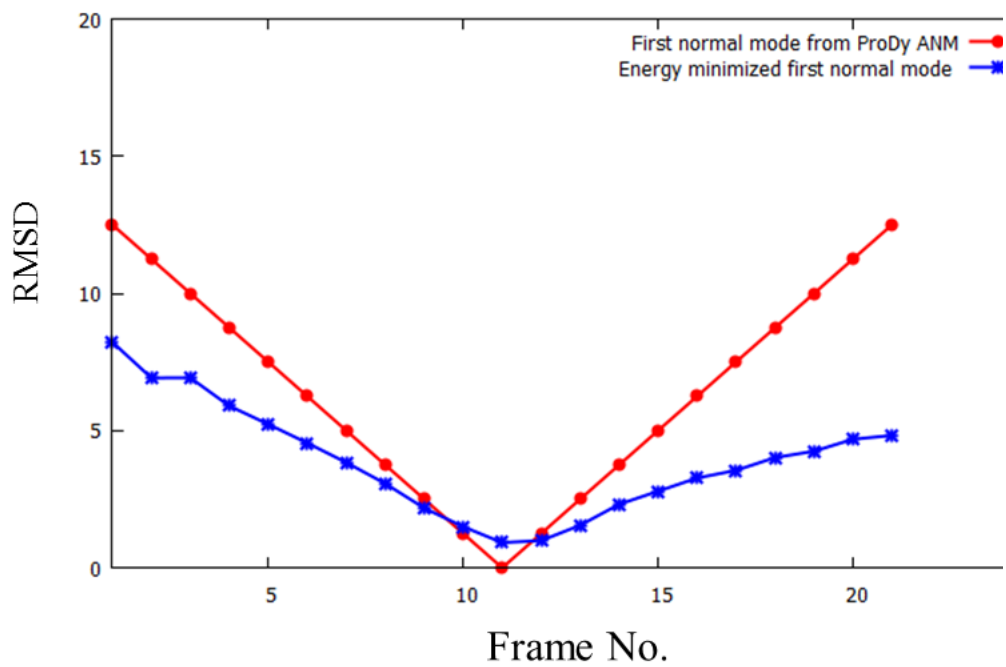


The three residues highlighted in yellow emphasize stretching of bonds to lengths that are not physically realistic. The image below shows the analogous animation after the original trajectory for the first mode is subjected to 1000 steps of [Energy Minimization](#).



This is further demonstrated in the third image below, which shows a plot of the root mean square deviation (RMSD), relative to input structure, versus frame number (for residue 118-the middle residue of the highlighted three) for the trajectory along the first mode from ProDy, before and after [Energy Minimization](#).

RMSD relative to input structure for lysozyme residue 118  
using frames along first normal mode



This, shows that energy minimization can remove the physically unrealistic bond stretches, restoring those structures that deviate far from the initial input structure.

### Energy Minimization Input Fields

Below is an example of the inputs needed to perform a 1000 step [Energy Minimization](#) of the trajectory for the first normal mode, `traverse_0.dcd`.

### Energy Minimization

run name	run_0	
reference pdb	<input type="button" value="Browse..."/> lysozyme.pdb	or <input type="button" value="Browse server"/> Local: lysozyme.pdb
input filename (dcd or pdb)	<input type="button" value="Browse..."/> No file selected.	or <input type="button" value="Browse server"/> Server:
PSF file name	<input type="button" value="Browse..."/> lysozyme.psf	or <input type="button" value="Browse server"/> Local: lysozyme.psf
output file name (dcd)	min_traverse_0.dcd	
number of processors	2	
keep run output files	no	
run type	minimization	
	number of minimization steps	1000

---

Advanced Input

Check Box for Advanced Input

### Example Output

Upon completion, the output screen should look something like the one below.

```

=====
DATA FROM RUN: Thu Sep  1 19:53:03 2016
Total number of frames = 21
Minimized structures saved to : ./run_0/energy_minimization/
=====

progress:
percent done: 100.0

```

### Visualization

The trajectories for traverse\_0.dcd and min\_traverse\_0.dcd can be directly compared in VMD to see the effect of the 1000 step minimization.

### Files Used and Created in Example

- input files
  - [lysozyme.pdb](#)
  - [lysozyme.psf](#)
- output files from Prody
  - [run\\_0.dcd](#)
  - [traverse\\_0.dcd](#) First normal mode
  - [traverse\\_1.dcd](#) Second normal mode

[traverse\\_2.dcd](#) Third normal mode  
[traverse\\_3.dcd](#) Fourth normal mode  
[traverse\\_4.dcd](#) Fifth normal mode  
[lysozyme\\_samp.pdb](#) Randomly-sampled conformations from each of the normal modes  
[lysozyme\\_samp.dcd](#) (pdb and dcd formats)  
[lysozyme\\_anm\\_kirchhoff.txt](#)  
[lysozyme\\_anm\\_hessian.txt](#)  
[lysozyme\\_anm.nmd](#)  
[lysozyme\\_anm\\_extended\\_bb.nmd](#)  
[lysozyme\\_anm\\_extended\\_all.nmd](#)  
[lysozyme\\_anm\\_evalues.txt](#)  
[lysozyme\\_anm\\_evectors.txt](#)  
[lysozyme\\_anm\\_sqflucts.txt](#)  
[lysozyme\\_anm\\_cross-correlations.txt](#)  
[lysozyme\\_anm\\_cross-correlations.hm](#)  
[lysozyme\\_anm\\_covariance.txt](#)  
[lysozyme\\_anm\\_beta.txt](#)

- output files from Energy Minimization

[min\\_traverse\\_0.dcd](#)  
[min\\_traverse\\_0.dcd.pdb](#)

## Advanced Input Options (NOT YET IMPLEMENTED!)

Advanced options are not yet implemented. ProDy will not run if the Check Box for Advanced Input is checked.

The screenshot shows the ProDy web interface with the following fields and options:

- run name:** run\_0
- reference pdb:** Browse... No file selected. or Browse server
- number of normal modes:** 15
- number of conformations per normal mode:** 50
- number of frames to traverse for each mode:** 10
- average sampled RMSD:** 1.0
- average sampled RMSD for normal mode trajectory:** 1.5
- Advanced Input:**
  - Check Box for Advanced Input:**
  - enter manual ProDy command:** [text input field]
  - low Rg cutoff:** 0
  - high Rg cutoff:** 300
  - check box to use atomic constraints:**

Buttons: Submit, Reset to default values

The input variables are listed below.

- **enter manual ProDy command** ProDy (Evol) application command, subcommand, -options. ([ProDy Applications](#))
- **low Rg cutoff** Structures with Rg values less than this value are discarded.
- **high Rg cutoff** Structures with Rg values greater than this value are discarded.
- **check box to use atomic constraints** Check box to implement the ability to discard structures that do not satisfy the atomic / geometric constraints provided in the user defined constraint file.
  - **constraint file name** Choose a text file with constraint definitions. See [Using Atomic Constraints](#) for guidance as to how to create such a file with desired constraints.

## Reference(s) and Citations

1. [ProDy: Protein Dynamics Inferred from Theory and Experiments](#) A. Bakan, L.M. Meireles, I. Bahar, Bioinformatics 27, 1575-1577 (2011). [BIBTeX](#), [EndNote](#), [Plain Text](#)
  2. [Dynamics of proteins predicted by molecular dynamics simulations and analytical approaches: Application to a-amylase inhibitor](#) P. Doruker, A. R. Atilgan, I. Bahar, Proteins 40, 512-524 (2000). [BIBTeX](#), [EndNote](#), [Plain Text](#)
  3. [Anisotropy of fluctuation dynamics of proteins with an elastic network model](#) A. R. Atilgan, S. R. Durrell, R. L. Jernigan, M. C. Demirel, O. Keskin, I. Bahar I. Biophys. J. 80, 505-515 (2001) [BIBTeX](#), [Endnote](#), [Plain Text](#)
  4. [SASSIE: A program to study intrinsically disordered biological molecules and macromolecular ensembles using experimental scattering restraints](#) J. E. Curtis, S. Raghunandan, H. Nanda, S. Krueger, Comp. Phys. Comm. 183, 382-389 (2012). [BIBTeX](#), [EndNote](#), [Plain Text](#)
- 

[Return to Simulate](#)

[Return to Main Documents Page](#)

[Go to top](#)

Supported via CCP-SAS a joint EPSRC (EP/K039121/1) and NSF (CHE-1265821) grant