

AMBER BINDING ENTHALPY CALCULATOR GUI starts with the command line call

python ENTHALPY.py

A simple graphical user interface for running small molecule ligand-protein binding enthalpy calculations from GPU accelerated AMBER molecular dynamics simulations, and viewing them in UCSF Chimera. If users do not have a license for GPU accelerated AMBER (i.e. pmemd.cuda), users can select the OpenMM molecular dynamics engine as an option for GPU accelerated molecular dynamics. The method sets up ensembles of replicate MD simulations in 4 states (protein-ligand complex, protein host alone, ligand guest alone and water alone). The method follows that of Fenley and Gilson (2014) as described here

<https://ambermd.org/tutorials/advanced/tutorial21/index.php>

and published here

<https://pubs.acs.org/doi/abs/10.1021/ct5004109>

Binding enthalpy describes the difference in heat due to the formation of the protein-ligand complex. If this formation gives off heat (i.e. is exothermic or energetically favored), then the binding enthalpy will be a negative number expressed in kcal/mol. In a system such as this, where covalent bonds are not broken, and assuming the change in entropy is negligible, the binding enthalpy will approximate the delta G (change in free energy) in the system during binding. For more information regarding the relationship between enthalpy, entropy and free energy please read this link.

<https://www2.chemistry.msu.edu/faculty/reusch/virttxtjml/energy1.htm>

This computer simulation approach is intended to mimic the results obtained through Isothermal Titration Calorimetry. For more information on this procedure

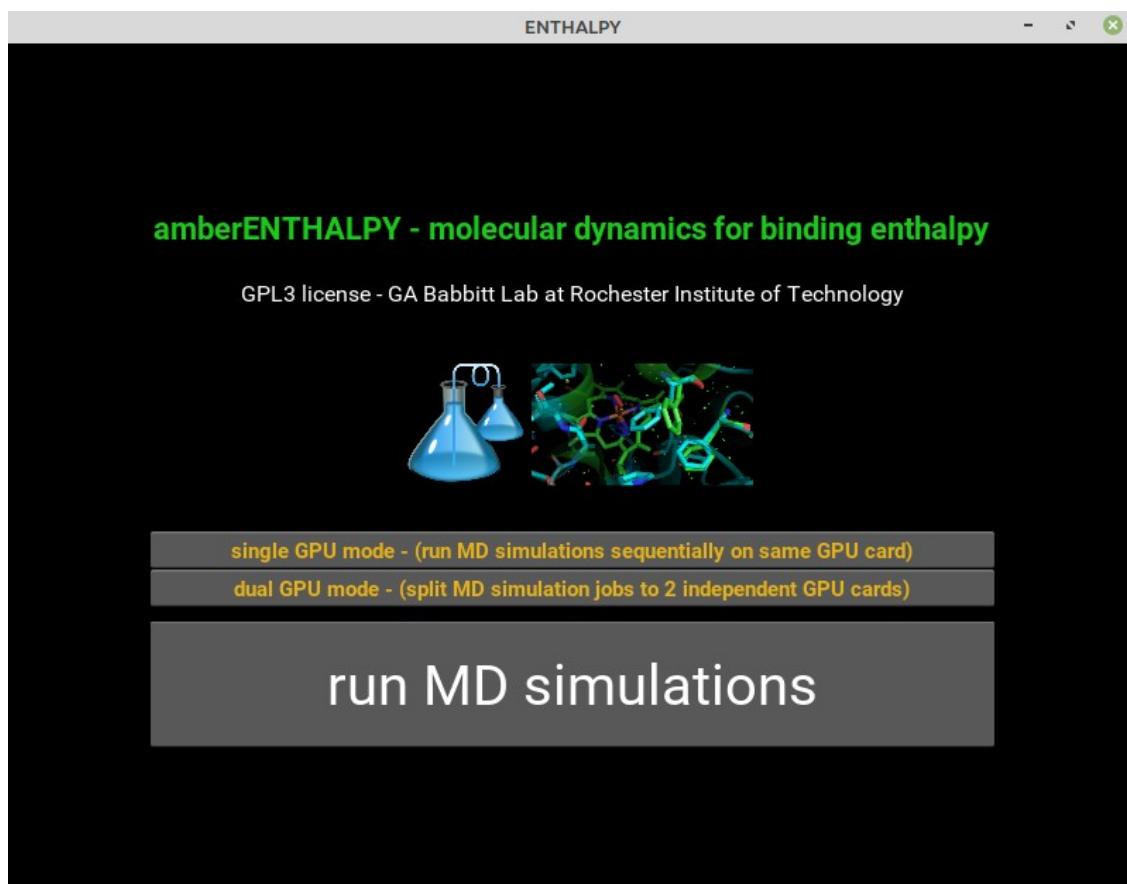
<https://www.jove.com/t/2796/isothermal-titration-calorimetry-for-measuring-macromolecule-ligand>

IMPORTANT NOTE: this script will walk the user through tLeAP so as to balance the waters in the 4 simulations. It is important not to proceed with simulation unless the difference in waters = zero after running tLeAP. The number of water molecules in the system are balanced such that

protein-ligand COMPLEX + water ONLY = protein ONLY + ligand only

NOTE: The GUI downloads with a preloaded example files for geldanamycin bound to the Hsp90 chaperone protein (PDB: 1yet)

ALSO NOTE: use a slightly larger water box than you would in normal MD simulation. If any MD runs abort due to unstable box boundaries, then increase the water box size and restart from the beginning. The box size on the GUI indicates the buffer distance from the surface of the protein or ligand to the edge of the water box. Because the ligand and single starting water molecule used to build the empty water simulation are so much smaller than a typical protein, a buffer size of 1.5 times the box size is used to build the ligand ONLY and water ONLY systems.



a GUI based script PATHS.pl will prompt paths to working directories for UCSF Chimera, Amber forcefields, and Amber Home directory and write a .ctl file. Then the main menu will pop up. Requires Amber 16/18 license (or AmberTools with OpenMM) and some dependencies in python and perl

MD control settings

protein force field (e.g. leaprc.protein.ff14SB):

ligand force field (e.g. leaprc.gaff2):

pdb ID with ligand (e.g. 1yet_bound):

pdb ID without ligand (e.g. 1yet_unbound):

pdb ID for ligand (e.g. 1yet_ligandGELDANAMYCIN):

number of repeated MD sample runs:

number of protein chains (e.g. 3 = A/B/C):

METHOD OF SOLVATION

☒ implicit - Generalized Born

☒ explicit - Particle Mesh Ewald

MD SIMULATION ENGINE

☒ pmemd.cuda (amber) - licensed

☒ OpenMM - open source

Length of MD heating run (ps):

0 200 400 600 800 1000

Length of MD equilibration run (ns):

0.0 5.0 10.0 15.0 20.0

Length of each MD sample run (ns):

0.0 5.0 10.0 15.0 20.0

extra salt conc (M) (implicit only):

0.000 0.200 0.400 0.600

To install the software needed for this GUI, download the Amber and AmberTools tarball files and run

```
perl AMBER_installer.pl
```

Dr. Gregory A. Babbitt¹ and Dr. Ernest P. Fokoue²

¹Thomas H. Gosnell School of Life Sciences, Rochester Institute of Technology, Rochester NY, USA 14623

²School of Mathematical Sciences, Rochester Institute of Technology, Rochester NY, USA 14623

This binding enthalpy calculator GUI is part of the DROIDS software suite for comparative protein dynamics

The Babbitt and Fokoue Labs at RIT have developed DROIDS v3.0, a software package for comparative protein dynamics, which applies metrics of distributional divergence and statistical analysis to the root mean square fluctuations (rmsf) of protein backbone atoms and maps these results to both static and moving image of proteins. We have also developed maxDemon v1.0, a multi-method machine learning application that trains on the comparative protein dynamics, identifies functionally conserved dynamics, and deploys classifications of functional dynamic states to newly generated protein simulations. Nine different types of machine learners can be deployed on the dynamics of each amino acid, then the resulting classifications are rendered upon movie images of the novel MD runs. This results in movies of protein dynamics where the conserved functional states are identified in real time by color mapping, allowing users to see both when and where a novel MD simulation displays a specific functional state defined by the comparative training. DROIDS+maxDemon designed to compare impacts of genetic variants and drug binding variants on the functional aspects of protein dynamics.

Our main goal is to try to visualize the impact of one of the longest time scale processes in the universe (i.e molecular evolution over 100s millions of years) on one of the shortest time scale processes (i.e. molecular motion over femtoseconds). To achieve this goal we use state-of-the-art biophysical simulations and graphics to design a gaming PC into a 'computational microscope' that is capable seeing how mutations and other molecular events like binding, bending and bonding affect the functioning of proteins and nucleic acids. DROIDS-1.0 (Detecting Relative Outlier Impacts in molecular Dynamic Simulation) is a GUI-based pipeline that works with AMBER16, Chimera 1.11 and CPPTRAJ to analyze and visualize comparative protein dynamics on GPU accelerated Linux graphics workstations. DROIDS employs a statistical method (multiple test corrected KS tests on all backbone atoms of each amino acid) to detect significant changes in molecular dynamics simulated on two homologous PDB structures. Quantitative differences in atom fluctuation are displayed graphically and mapped onto movie images of the protein dynamics at the level of individual residues. P values indicating significant changes are also able to be similarly mapped. DROIDS is useful for examining how mutations or binding interactions affect protein dynamics. DROIDS was produced by student effort at the Rochester Institute of Technology under the direction of Dr. Gregory A. Babbitt as a collaborative project between the Gosnell School of Life Sciences and the Biomedical Engineering Dept. Visit our lab website (<https://people.rit.edu/gabsbi/>) and download DROIDS 1.0 from Github at <https://github.com/gbabbitt/DROIDS-1.0>. We will be posting video results periodically on our youtube channel at <https://www.youtube.com/channel/UCJTbQGq01pBCMDQikn566Kw>

please cite

Gregory A. Babbitt, Jamie S. Mortensen, Erin E. Coppola, Lily E. Adams, Justin K. Liao,
DROIDS 1.20: A GUI-Based Pipeline for GPU-Accelerated Comparative Protein Dynamics,
Biophysical Journal,
Volume 114, Issue 5,
2018,
Pages 1009-1017,
ISSN 0006-3495,
<https://doi.org/10.1016/j.bpj.2018.01.020>.
(<http://www.sciencedirect.com/science/article/pii/S0006349518301462>)
Abstract: Traditional informatics in comparative genomics work only with static

representations of biomolecules (i.e., sequence and structure), thereby ignoring the molecular dynamics (MD) of proteins that define function in the cell. A comparative approach applied to MD would connect this very short timescale process, defined in femtoseconds, to one of the longest in the universe: molecular evolution measured in millions of years. Here, we leverage advances in graphics-processing-unit-accelerated MD simulation software to develop a comparative method of MD analysis and visualization that can be applied to any two homologous Protein Data Bank structures. Our open-source pipeline, DROIDS (Detecting Relative Outlier Impacts in Dynamic Simulations), works in conjunction with existing molecular modeling software to convert any Linux gaming personal computer into a “comparative computational microscope” for observing the biophysical effects of mutations and other chemical changes in proteins. DROIDS implements structural alignment and Benjamini-Hochberg-corrected Kolmogorov-Smirnov statistics to compare nanosecond-scale atom bond fluctuations on the protein backbone, color mapping the significant differences identified in protein MD with single-amino-acid resolution. DROIDS is simple to use, incorporating graphical user interface control for Amber16 MD simulations, cpptraj analysis, and the final statistical and visual representations in R graphics and UCSF Chimera. We demonstrate that DROIDS can be utilized to visually investigate molecular evolution and disease-related functional changes in MD due to genetic mutation and epigenetic modification. DROIDS can also be used to potentially investigate binding interactions of pharmaceuticals, toxins, or other biomolecules in a functional evolutionary context as well.

Gregory A. Babbitt, Ernest P. Fokoue, Joshua R. Evans, Kyle I. Diller, Lily E. Adams,

DROIDS 3.0—Detecting Genetic and Drug Class Variant Impact on Conserved Protein Binding Dynamics,

Biophysical Journal,

Volume 118, Issue 3,

2020,

Pages 541-551,

ISSN 0006-3495,

<https://doi.org/10.1016/j.bpj.2019.12.008>.

(<http://www.sciencedirect.com/science/article/pii/S0006349519343905>)

Abstract: The application of statistical methods to comparatively framed questions about the molecular dynamics (MD) of proteins can potentially enable investigations of biomolecular function beyond the current sequence and structural methods in bioinformatics. However, the chaotic behavior in single MD trajectories requires statistical inference that is derived from large ensembles of simulations representing the comparative functional states of a protein under investigation. Meaningful interpretation of such complex forms of big data poses serious challenges to users of MD. Here, we announce Detecting Relative Outlier Impacts from Molecular Dynamic Simulation (DROIDS) 3.0, a method and software package for comparative protein dynamics that includes maxDemon 1.0, a multimethod machine learning application that trains on large ensemble comparisons of concerted protein motions in opposing functional states generated by DROIDS and deploys learned classifications of these states onto newly generated MD simulations. Local canonical correlations in learning patterns generated from independent, yet identically prepared, MD validation runs are used to identify regions of functionally conserved protein dynamics. The subsequent impacts of genetic and/or drug class variants on conserved dynamics can also be analyzed by deploying the classifiers on variant MD simulations and quantifying how often these altered protein systems display opposing functional states. Here, we present several case studies of complex changes in functional protein dynamics caused by temperature, genetic mutation, and binding interactions with nucleic acids and small molecules. We demonstrate that our machine learning algorithm can properly identify regions of functionally conserved dynamics in ubiquitin and TATA-binding protein (TBP). We quantify the impact of genetic variation in TBP and drug class variation targeting

the ATP-binding region of Hsp90 on conserved dynamics. We identify regions of conserved dynamics in Hsp90 that connect the ATP binding pocket to other functional regions. We also demonstrate that dynamic impacts of various Hsp90 inhibitors rank accordingly with how closely they mimic natural ATP binding.

Before you start you should collect .pdb files you want to compare and move them into the ambermdGUI folder. Naming convention should be PDB_ID.pdb. Be sure to check that they are 'sensibly' prepared. Edit in UCSF Chimera if necessary.

Remove mirrored structures or unusual ligands used in crystal prep. Atypical Amber preparations (e.g. beyond adding H, removing crystallographic waters and missing atoms using teleap) can be done at the command line using Antechamber for further ligand library prep

NOTE: this program assumes .pdb files are ready for run through pdb4amber, antechamber and teleap

Dependencies - python-kivy, perl, perl module (Descriptive), perl-tk, python, python-tk, python-gi, R-base, R-dev, R package(ggplot2), UCSF Chimera 1.11, evince(pdf viewer)

Amber18 (licensed from Univ of Ca; visit ambermd.org), Ambertools18 (tested on Linux Mint 18 and 19 Cinnamon 64-bit Kernel 4.4.0-53-generic)

maxDemon R packages - see installer script and/or user manual for complete list of machine learning packages to be installed

BabbittLab - Rochester Inst. Technol. Rochester NY

DROIDS 3.0 Copyright 2019 G.A. Babbitt.

DROIDS 3.0 is free software: you can redistribute it and/or modify it under the terms of the GNU General Public License as published by the Free Software Foundation, either version 3 of the License, or (at your option) any later version.

DROIDS 3.0 is distributed in the hope that it will be useful, but WITHOUT ANY WARRANTY; without even the implied warranty of MERCHANTABILITY or FITNESS FOR A PARTICULAR PURPOSE. See the GNU General Public License for more details.

You should have received a copy of the GNU General Public License along with DROIDS 3.0. If not, see <http://www.gnu.org/licenses/>.

Visit us on GitHub and at <https://people.rit.edu/gabsbi/>