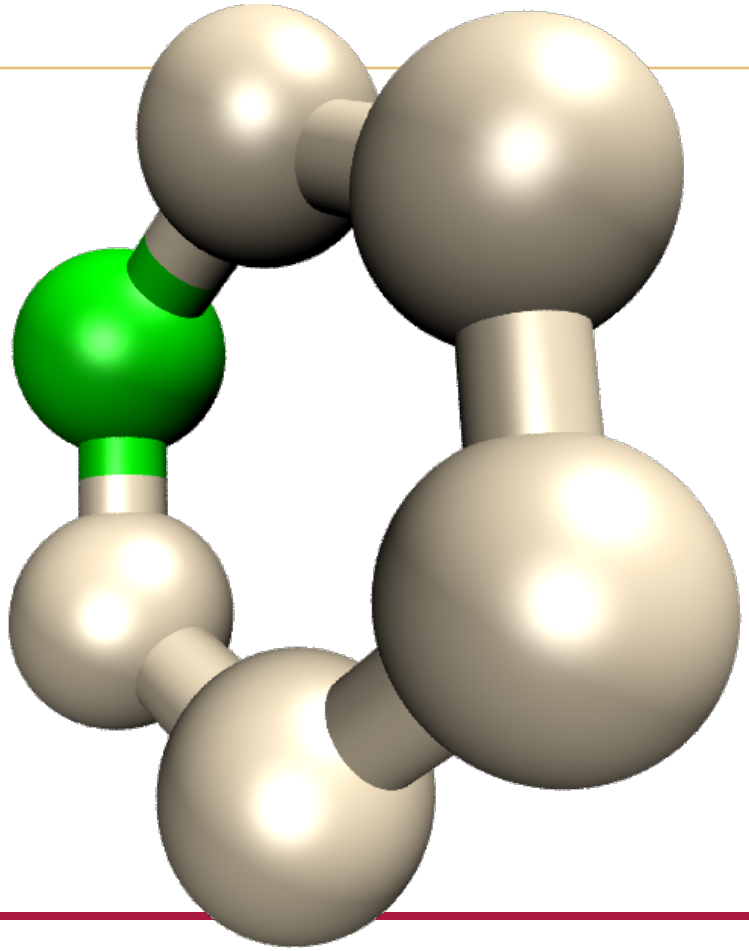
Molecular Simulation with OpenMM Zephyr



February 2010

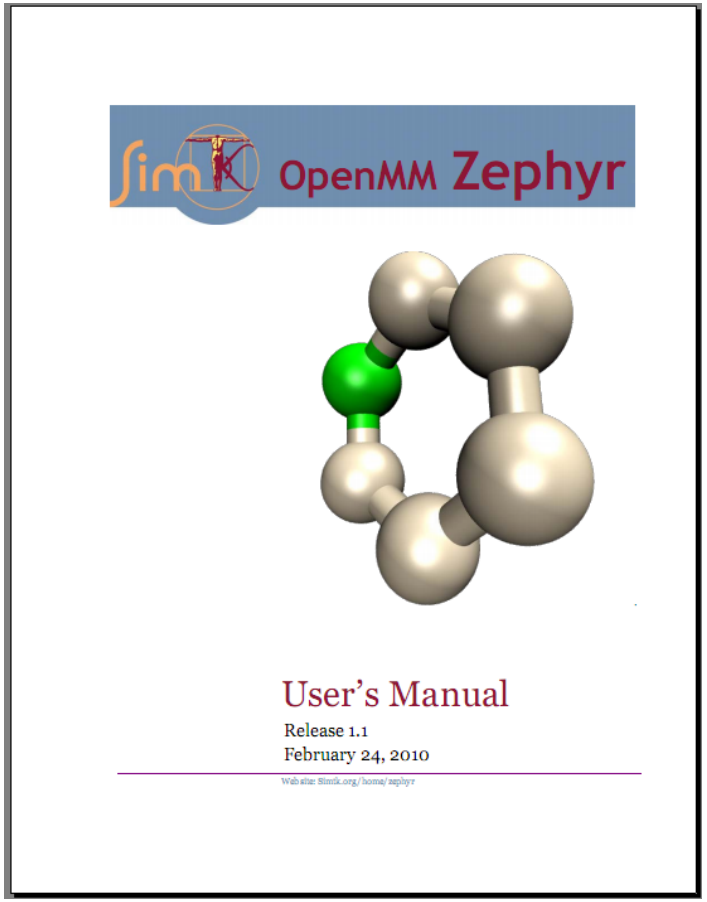
Christopher Bruns

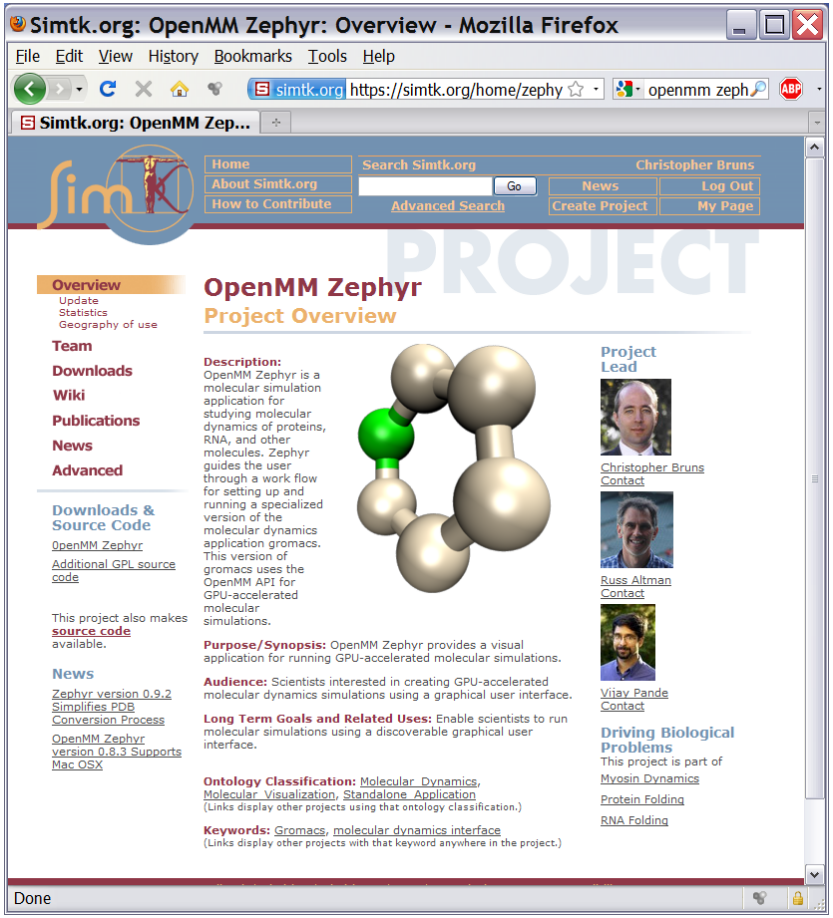
simbios

What is OpenMM Zephyr?

- Graphical user interface for running GPU accelerated molecular dynamics simulations
- Automates running of gromacs programs
  - http://www.gromacs.org/
- Discoverable interface
- Showcases OpenMM library

Zephyr Resources





https://simtk.org/home/zephyr

Getting Zephyr and VMD

Zephyr

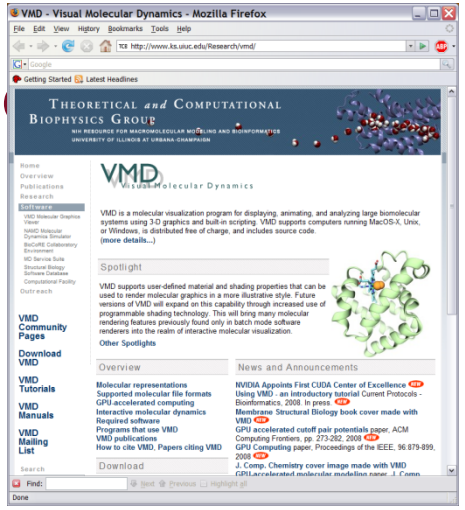
https://simtk.org/home/zephyr

Downloads => (installer)

Get User Guide also

VMD

http://www.ks.uiuc.edu/Research/vm



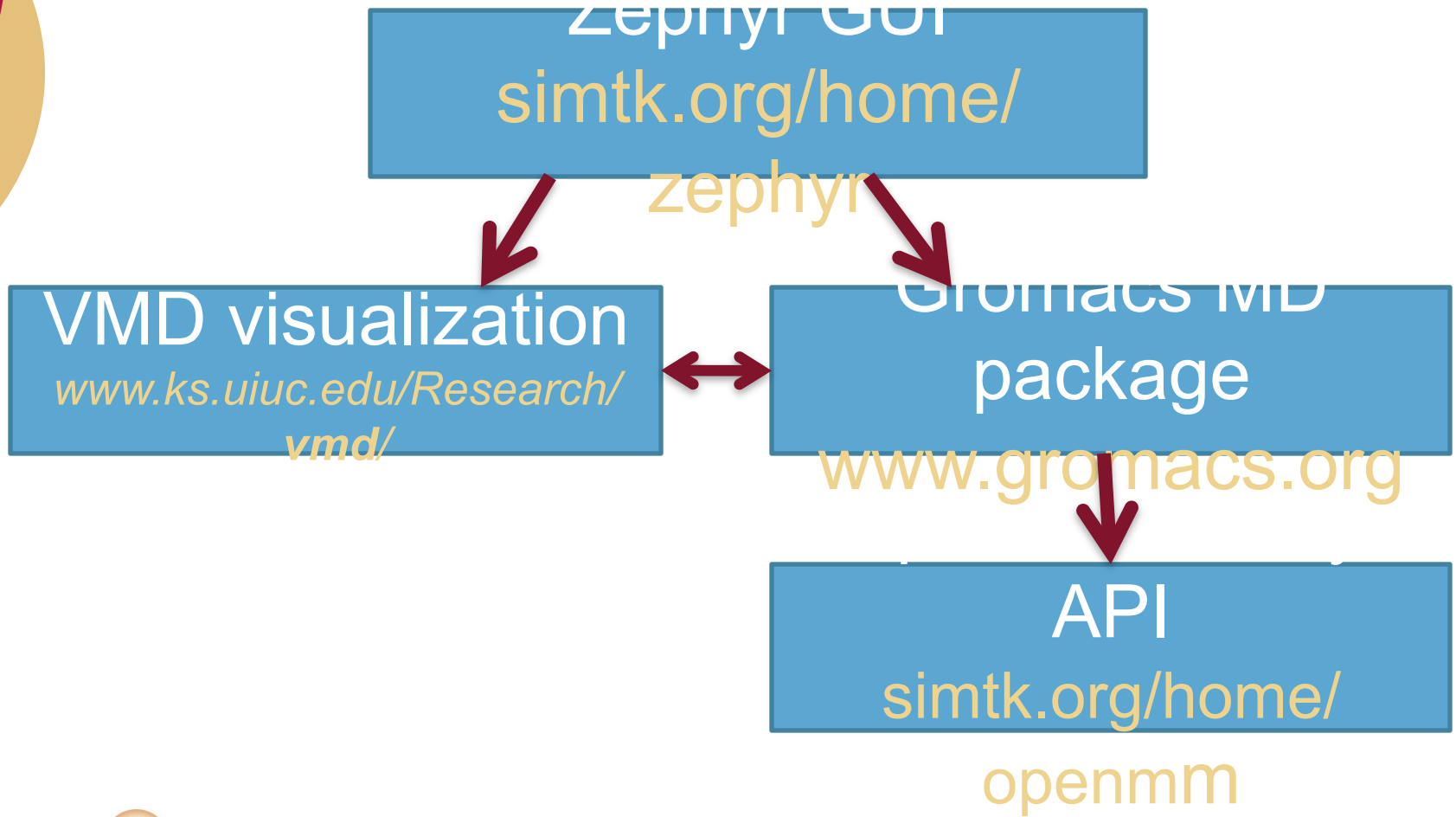
Install in a Folder with no space characters (Windows)

- Install in “C:\Zephyr”
- NOT “Program Files”
- NOT “My Documents”
- Issue with this version of gromacs
- No spaces in the entire path

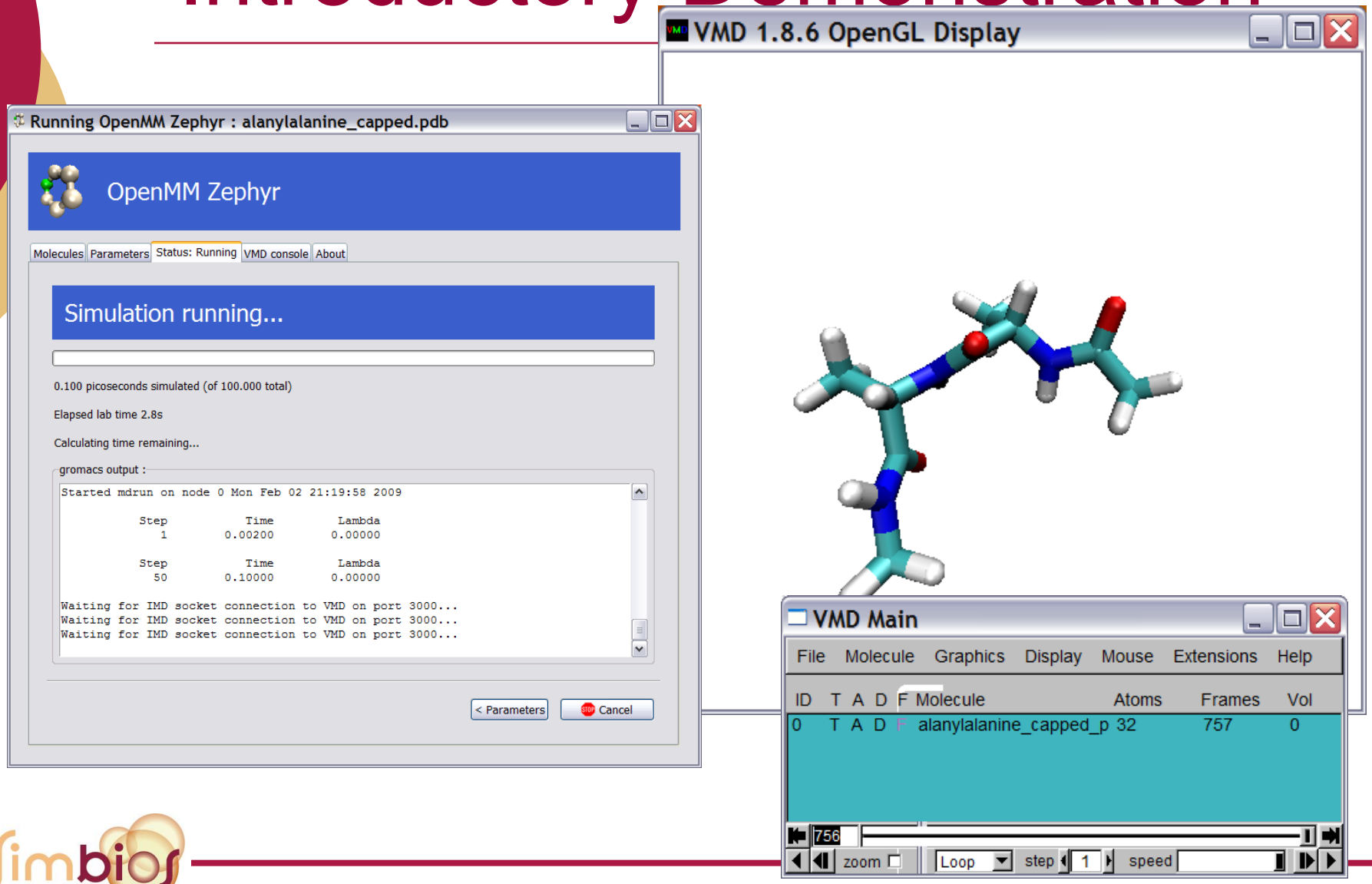
Overview of exercises

- Be sure to get handout with just the exercise, to help you work at your own pace.
  - Dinucleotide (adenylyladenylate) 2 residues
    - change viscosity
    - replay simulation in VMD
  - Tetraloop (GCAA) 12 residues
    - attempt GPU acceleration
    - write trajectory in PDB format using VMD
  - Unfold an alpha helix
    - modify temperature
  - double helix (255d)
    - From PDB web site

Zephyr is a GUI layer on OpenMM, Gromacs, and VMD

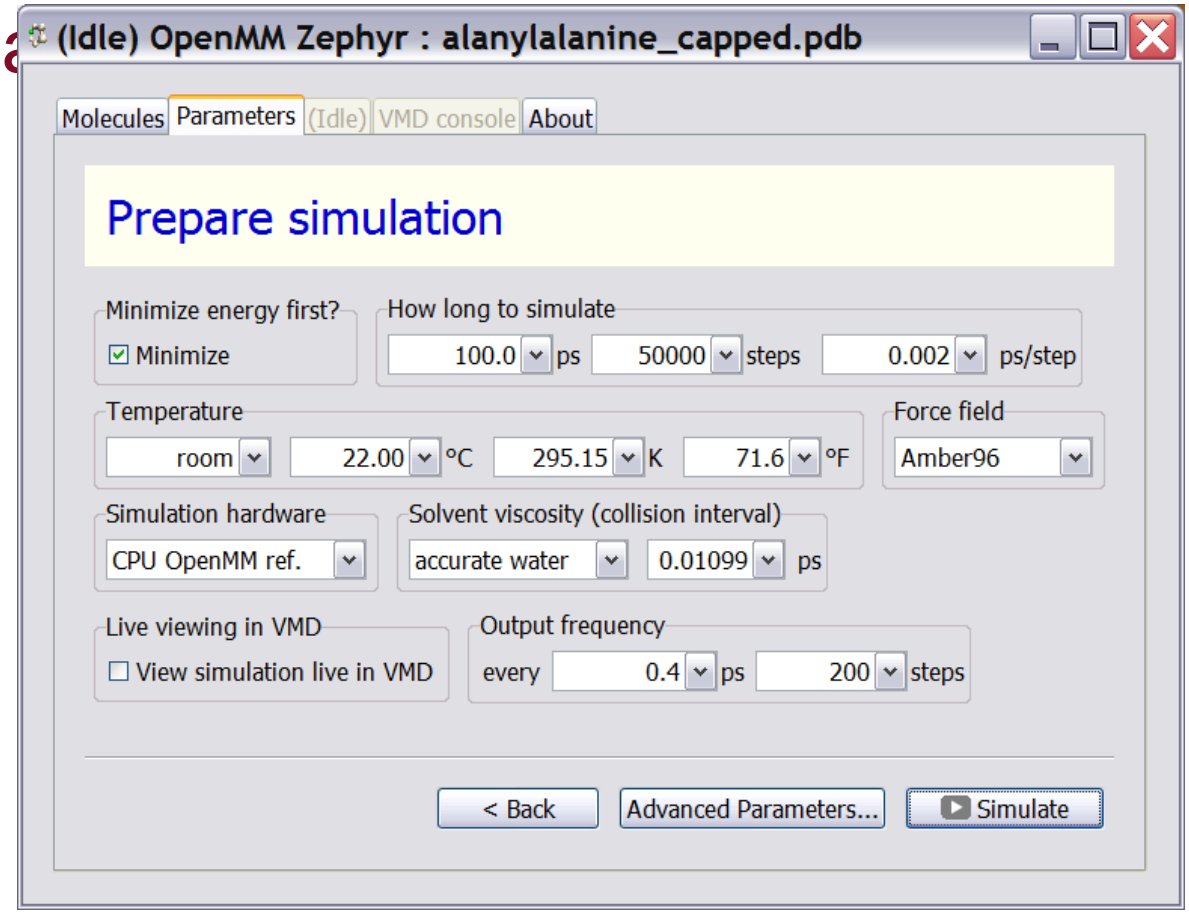


Introductory Demonstration



Simulation parameters

Implicit solvent only – unlike standard groma



Exercises: RNA Dinucleotide

1. Launch Zephyr and simulate adenylyladenylate.pdb for 40 picoseconds. Does it run to completion?
2. Raise solvent collision interval from 0.01099 to 1.0 ps. Simulate again. Does the simulation look any different?
3. Replay the trajectory in VMD. Vary the speed of the playback.
4. Restore the solvent collision interval to 0.01099 ps.

RNA Dinucleotide

- Increased collision interval means lower viscosity.
- Lower viscosity permits faster exploration of conformational space.
- VMD provides a rich environment for exploring molecular simulations.

Exercises - VMD

- Load a Zephyr trajectory from your simulations directory into VMD
  - VMDMain->File->New Molecule
    1. select Determine file type: Gromacs GRO
    2. click Browse-><whatever>.em.gro
    3. click Load
    4. select Determine file type: Gromacs TRR trajectory
    5. click Browse-><whatever>.md.trr
    6. click Load (again)
- Save trajectory in PDB format from VMD
  - Select trajectory in VMD Main window
  - File->Save Coordinates->pdb, all->Save "my_trajectory.pdb"

Zephyr Design Principles

1. Discoverability
   - Not a black box
   - Learn molecular dynamics by investigating simulation interface
2. Convention
   - Harvest best practices of experts for default work flow
3. Feedback
   - Reveal when things go wrong
   - Reveal when things go right

Either:

A: Run these gromacs programs

INPUT:
PDB structure file

.pdb → pdb2gm ← force field parameters

pdb2gm → .top, .gro

.top, .gro → editconf

editconf → .top

.top → grompp → .tpr

minimize energy

.tpr → mdrun

mdrun → .gro

.gro + .mdp (simulation parameters) → grompp → .tpr

molecular dynamics

.tpr + .agb (GBSA parameters) → mdrun → .pdb, .trr

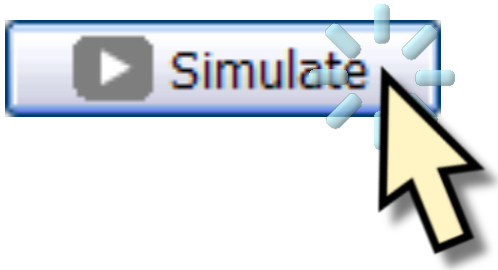
OUTPUT: structure and trajectory

OR

B: Click "Simulate" button in Zephyr

INPUT:
PDB structure file

.pdb



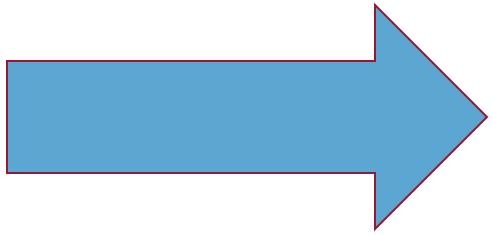
.pdb, .trr

OUTPUT: structure and trajectory

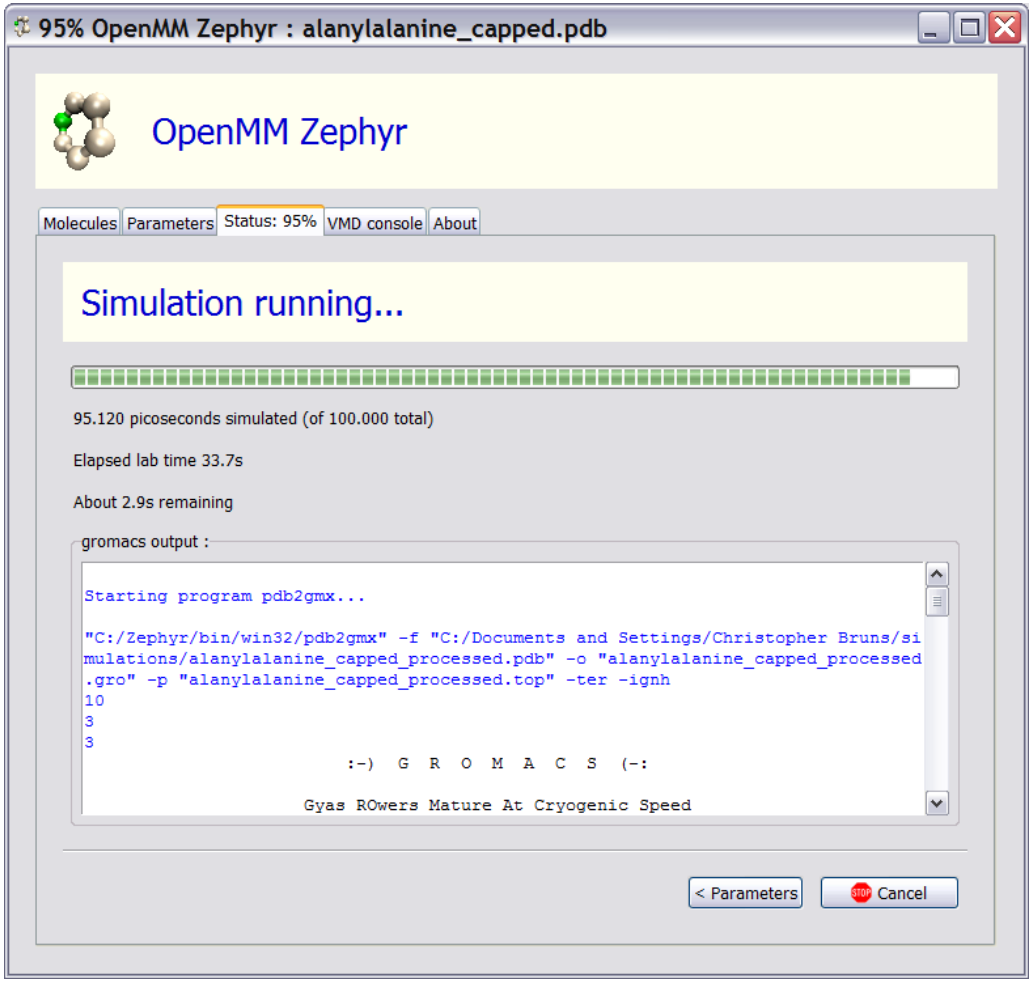
Exercises: Tetraloop hairpin

1. Begin simulating gcaa.pdb using default parameters. How long would it take to complete? Press “Cancel” to halt the simulation. (On Mac you may need to close VMD and/or Zephyr to halt)
2. Set “Simulation Hardware” to “GPU Nvidia”. Simulate again. What happens?
3. If GPU acceleration is unavailable:
    1. Restore hardware to “CPU OpenMM ref.”
    2. Reload adenylyladenylate.pdb

Read the blue lines to learn to run Gromacs-OpenMM on the command line



see also zephyr.log file



Zephyr Capabilities

- OS: Windows (XP or Vista), Mac, Linux
- 32 bit
- Implicit solvent only
- Amber96 force field only
- Standard protein/RNA/DNA molecules only
  - No ATP, heme, flavins, etc.
- For GPU accelerated dynamics:
  - you must have supported GPU and drivers

Time Scales in Molecular Mechanics

femtosecond	10^{-15} second	bond vibration
picosecond	10^{-12} second	side chain motion
nanosecond	10^{-9} second	protein/RNA tumbling
microsecond	10^{-6} second	helix/coil transition
	10^{-4} second	RNA duplex formation
millisecond	10^{-3} second	protein/RNA folding
second	10^{0} second	protein synthesis

Exercise: Unfolding an alpha helix

1. Simulate polyala10helix.pdb for 40 ps at 1000C.
2. Select in VMD Main window. Then Set VMD->Graphics->Representation to “NewCartoon”
3. Simulate same structure again for 60 ps at 300K. View as New Cartoon again. What is different?

Unfolding an alpha helix

- You will be unable to fold a protein in the lab time available, but you might be able to unfold something.

Pushing the envelope: Temperature

1. Increase temperature to 6000 kelvin, the temperature of the sun. Simulate. What happens; and why?
2. Cancel and restart with simulation step size set to 0.0005 ps. Is this what would really happen on the sun? (On Mac you may need to close VMD and/or Zephyr to halt)
3. Restore temperature and step size to 300K and 0.002ps

Temperature

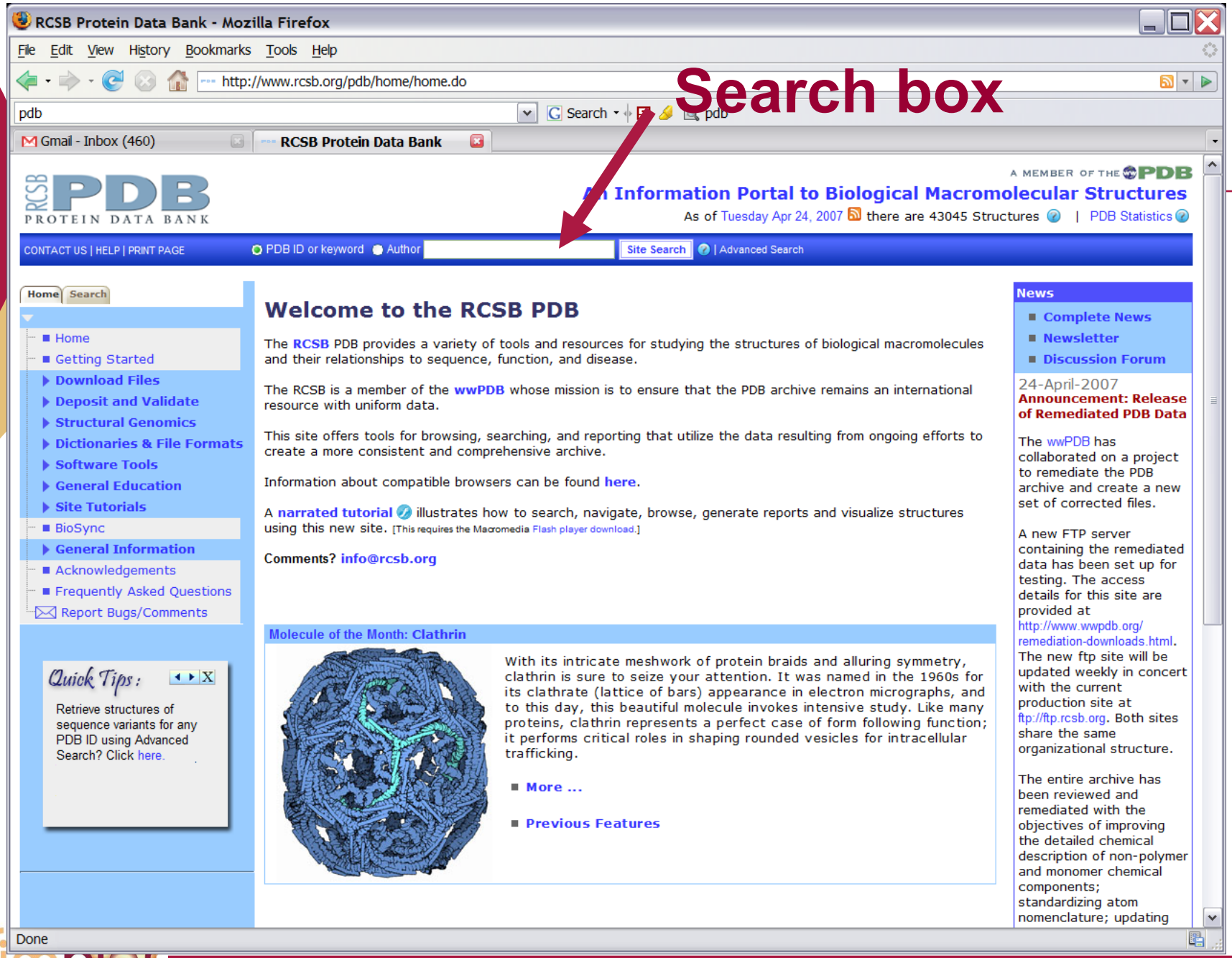
- Temperature of 6000 degrees requires a smaller step size for stable integration. “Blowing up” reveals numerical sensitivity of the integration algorithm, not physics.
- Molecular dynamics does not handle breaking or forming covalent bonds. i.e. no chemistry. Chemistry might really happen at sun temperature.

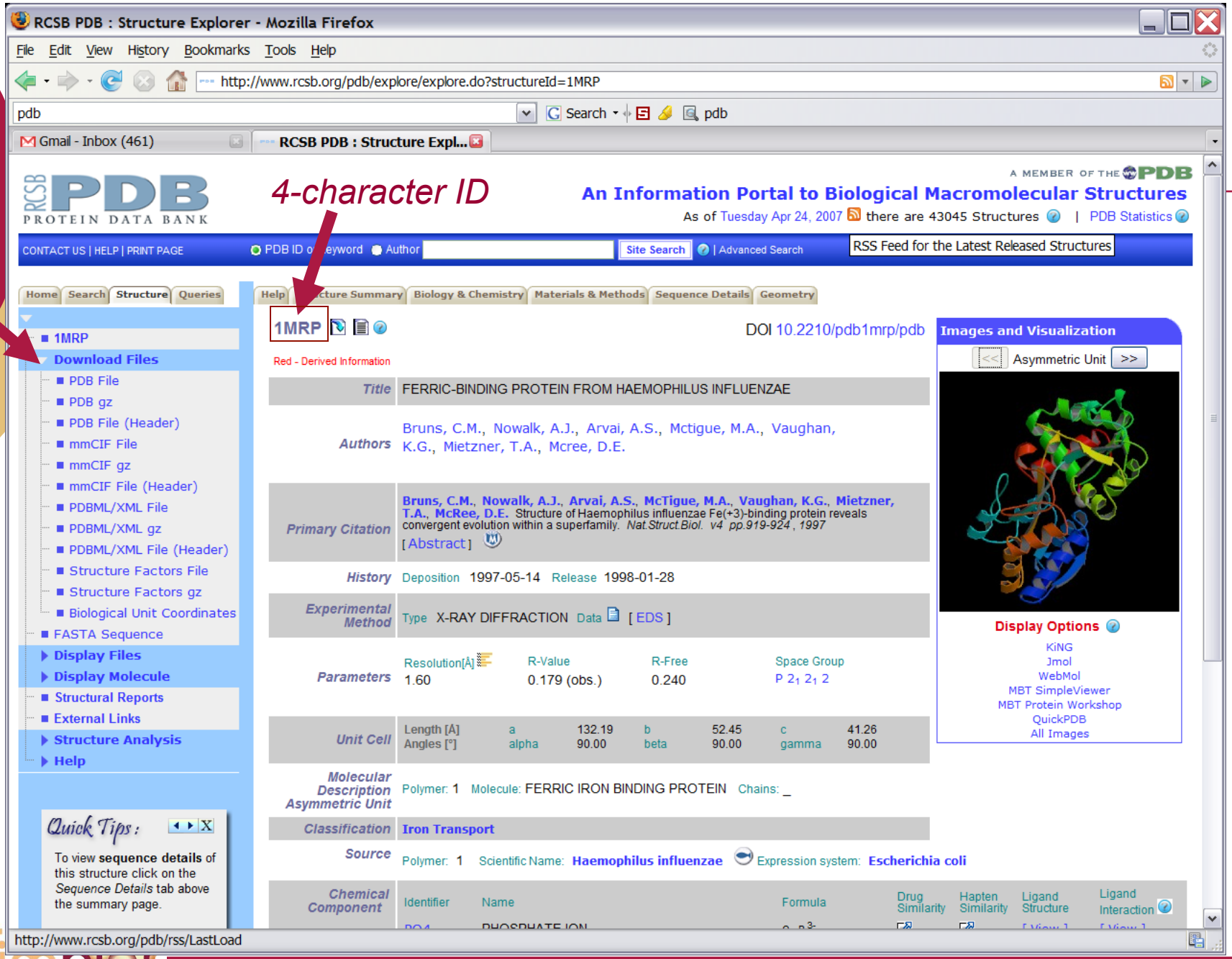
Tips for fast simulation

- Increase step interval for VMD and trajectory output. (Is it slower or faster?)
- Use GPU acceleration
- Increase solvent collision interval to 1 (lower viscosity)

PDB: The Protein Data Bank

- http://www.rcsb.org/pdb/
- Repository of DNA, RNA, and protein atomic structures
- Contains experimental results, not perfect models
- Entries identified by 4 character ID, e.g. “1MRP”, “1GRZ”
- Use “advanced search” at PDB web site to find structures.

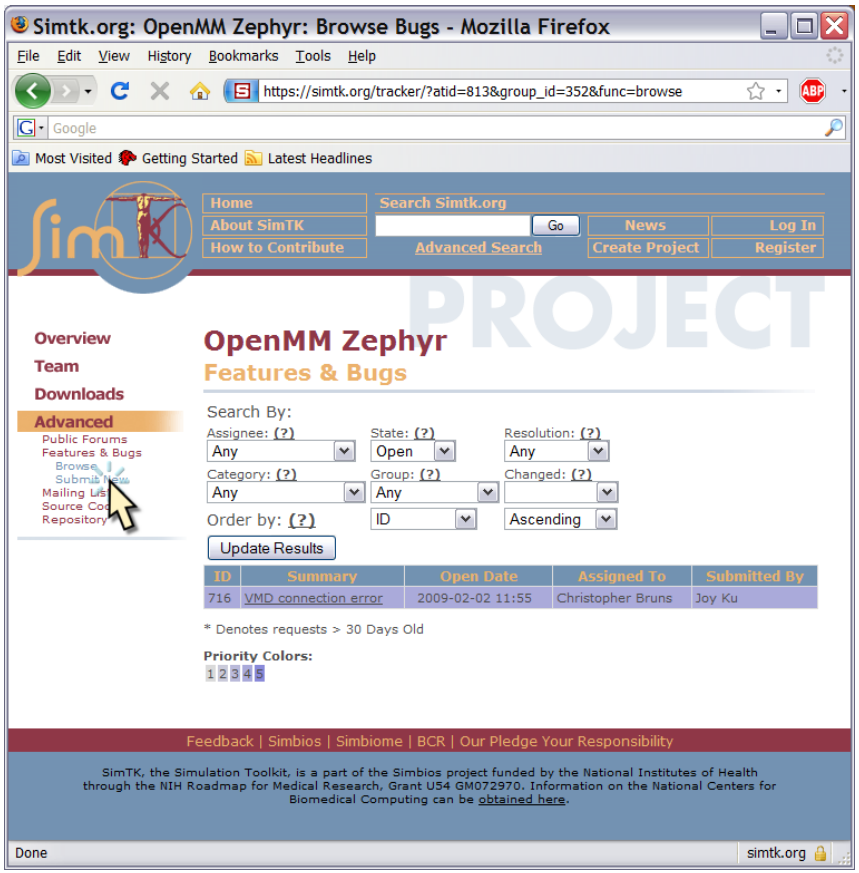

RCSB Protein Data Bank - Mozilla Firefox
File Edit View History Bookmarks Tools Help
http://www.rcsb.org/pdb/home/home.do
Search box
pdb
Search
Gmail - Inbox (460)
RCSB Protein Data Bank
RCSB PDB PROTEIN DATA BANK
A MEMBER OF THE PDB
An Information Portal to Biological Macromolecular Structures
As of Tuesday Apr 24, 2007 there are 43045 Structures | PDB Statistics
CONTACT US | HELP | PRINT PAGE
PDB ID or keyword Author
Site Search | Advanced Search
Home Search
Home
Getting Started
Download Files
Deposit and Validate
Structural Genomics
Dictionaries & File Formats
Software Tools
General Education
Site Tutorials
BioSync
General Information
Acknowledgements
Frequently Asked Questions
Report Bugs/Comments
Quick Tips :
Retrieve structures of sequence variants for any PDB ID using Advanced Search? Click here.
Welcome to the RCSB PDB
The RCSB PDB provides a variety of tools and resources for studying the structures of biological macromolecules and their relationships to sequence, function, and disease.
The RCSB is a member of the wwPDB whose mission is to ensure that the PDB archive remains an international resource with uniform data.
This site offers tools for browsing, searching, and reporting that utilize the data resulting from ongoing efforts to create a more consistent and comprehensive archive.
Information about compatible browsers can be found here.
A narrated tutorial illustrates how to search, navigate, browse, generate reports and visualize structures using this new site. [This requires the Macromedia Flash player download.]
Comments? info@rcsb.org
Molecule of the Month: Clathrin
With its intricate meshwork of protein braids and alluring symmetry, clathrin is sure to seize your attention. It was named in the 1960s for its clathrate (lattice of bars) appearance in electron micrographs, and to this day, this beautiful molecule invokes intensive study. Like many proteins, clathrin represents a perfect case of form following function; it performs critical roles in shaping rounded vesicles for intracellular trafficking.
More ...
Previous Features
News
Complete News
Newsletter
Discussion Forum
24-April-2007
Announcement: Release of Remediated PDB Data
The wwPDB has collaborated on a project to remediate the PDB archive and create a new set of corrected files.
A new FTP server containing the remediated data has been set up for testing. The access details for this site are provided at http://www.wwpdb.org/remediation-downloads.html. The new ftp site will be updated weekly in concert with the current production site at ftp://ftp.rcsb.org. Both sites share the same organizational structure.
The entire archive has been reviewed and remediated with the objectives of improving the detailed chemical description of non-polymer and monomer chemical components; standardizing atom nomenclature; updating
Done


RCSB PDB : Structure Explorer - Mozilla Firefox
File Edit View History Bookmarks Tools Help
http://www.rcsb.org/pdb/explore/explore.do?structureId=1MRP
RCSB PDB PROTEIN DATA BANK
4-character ID
An Information Portal to Biological Macromolecular Structures
As of Tuesday Apr 24, 2007 there are 43045 Structures | PDB Statistics
CONTACT US | HELP | PRINT PAGE
Site Search | Advanced Search
RSS Feed for the Latest Released Structures
Home Search Structure Queries
Help Structure Summary Biology & Chemistry Materials & Methods Sequence Details Geometry
1MRP
Download Files
PDB File
PDB gz
PDB File (Header)
mmCIF File
mmCIF gz
mmCIF File (Header)
PDBML/XML File
PDBML/XML gz
PDBML/XML File (Header)
Structure Factors File
Structure Factors gz
Biological Unit Coordinates
FASTA Sequence
Display Files
Display Molecule
Structural Reports
External Links
Structure Analysis
Help
1MRP
DOI 10.2210/pdb1mrp/pdb
Red - Derived Information
Title FERRIC-BINDING PROTEIN FROM HAEMOPHILUS INFLUENZAE
Authors Bruns, C.M., Nowalk, A.J., Arvai, A.S., Mctigue, M.A., Vaughan, K.G., Mietzner, T.A., Mcree, D.E.
Primary Citation Bruns, C.M., Nowalk, A.J., Arvai, A.S., McTigue, M.A., Vaughan, K.G., Mietzner, T.A., McRee, D.E. Structure of Haemophilus influenzae Fe(+3)-binding protein reveals convergent evolution within a superfamily. Nat.Struct.Biol. v4 pp.919-924 , 1997
[Abstract]
History Deposition 1997-05-14 Release 1998-01-28
Experimental Method Type X-RAY DIFFRACTION Data [EDS]
Parameters Resolution[Å] 1.60 R-Value 0.179 (obs.) R-Free 0.240 Space Group P 2₁ 2₁ 2
Unit Cell Length [Å] a 132.19 b 52.45 c 41.26 Angles [°] alpha 90.00 beta 90.00 gamma 90.00
Molecular Description Asymmetric Unit Polymer: 1 Molecule: FERRIC IRON BINDING PROTEIN Chains: _
Classification Iron Transport
Source Polymer: 1 Scientific Name: Haemophilus influenzae Expression system: Escherichia coli
Chemical Component Identifier Name Formula Drug Similarity Hapten Similarity Ligand Structure Ligand Interaction
Images and Visualization
Asymmetric Unit
Display Options
KiNG
Jmol
WebMol
MBT SimpleViewer
MBT Protein Workshop
QuickPDB
All Images
Quick Tips:
To view sequence details of this structure click on the Sequence Details tab above the summary page.
http://www.rcsb.org/pdb/rss/LastLoad

Exercises

- Download structure 255d from the Protein Data Bank. Try simulating in Zephyr.
  - Be aware that most other PDB structures won't work without modification.

Submit feature requests and bug reports



Where to go from here

- Running OpenMM gromacs from the command line
  - see chapter 5 of the Zephyr Users Guide
- Pyopenmm: Python bindings for the OpenMM API
  - https://simtk.org/home/pyopenmm

Zephyr Community Resources

- Sign up for “OpenMM Zephyr News” mailing list
  - https://simtk.org/mail/?group_id=352
- OpenMM Zephyr forums
  - https://simtk.org/forum/?group_id=352