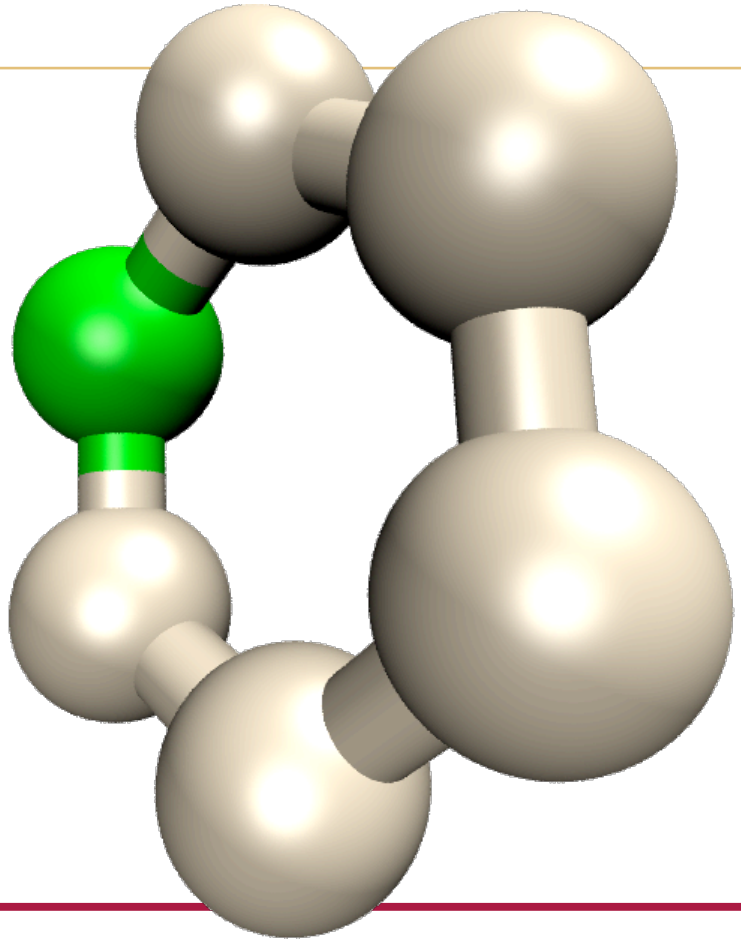
~~Using~~ OpenMM Zephyr



February 2009

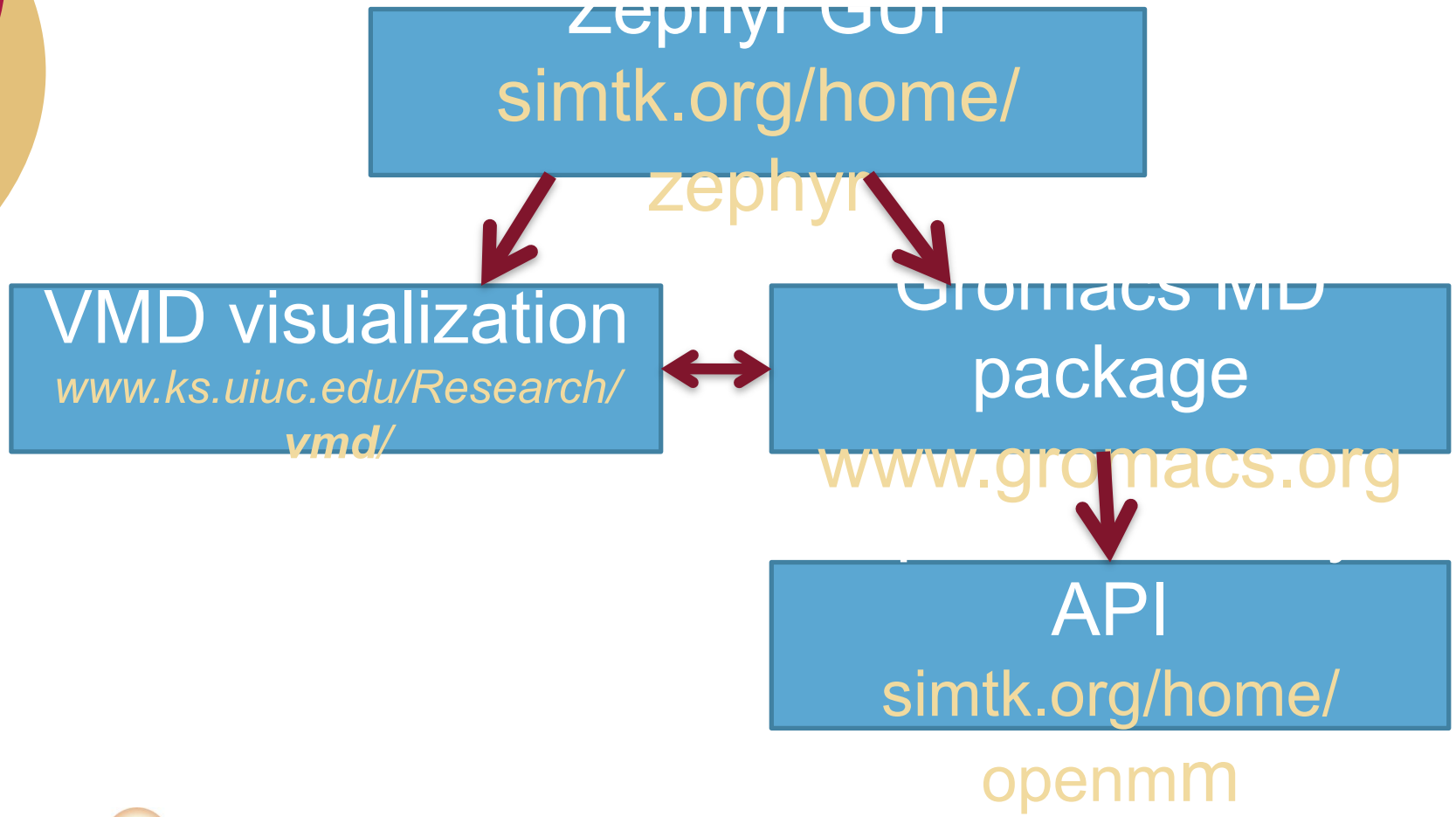
Christopher Bruns

simbios

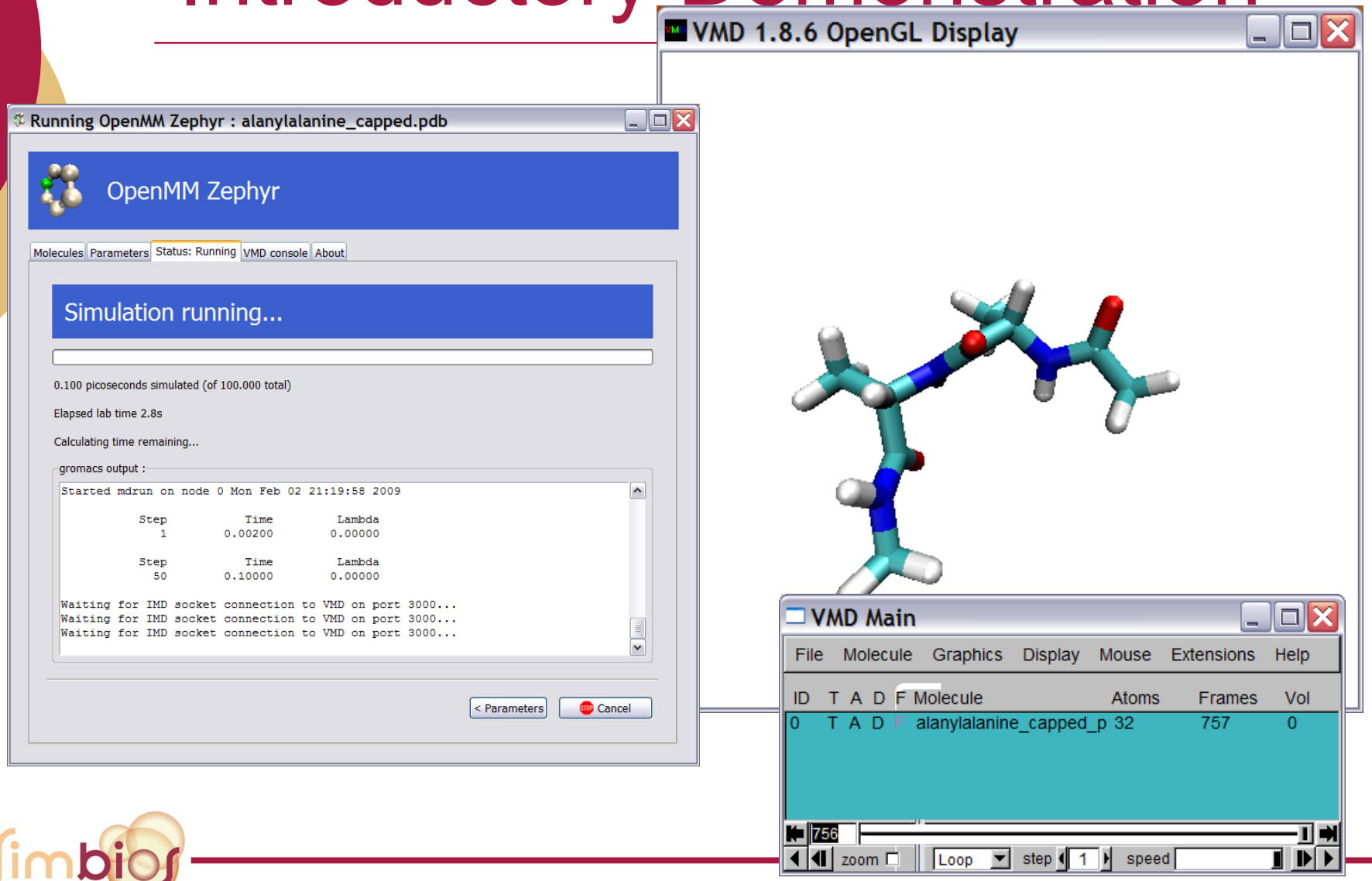
What is OpenMM Zephyr?

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

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



Introductory Demonstration



Zephyr Community Resources

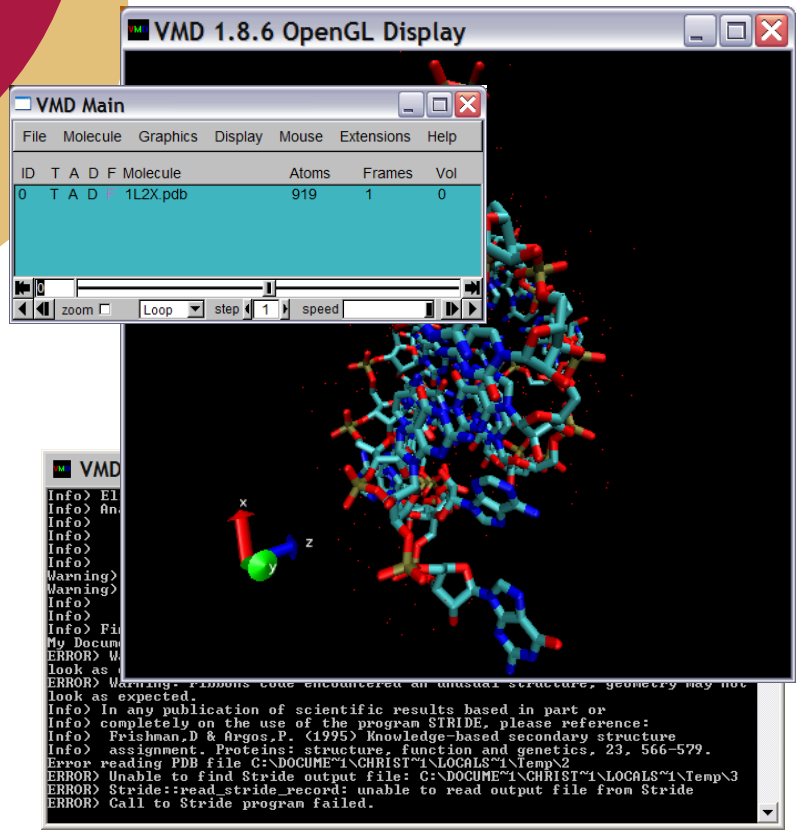
- Sign up for “OpenMM Zephyr News” mailing list
  - https://simtk.org/mail/?group_id=352
- OpenMM Zephyr forums
  - https://simtk.org/forum/?group_id=352

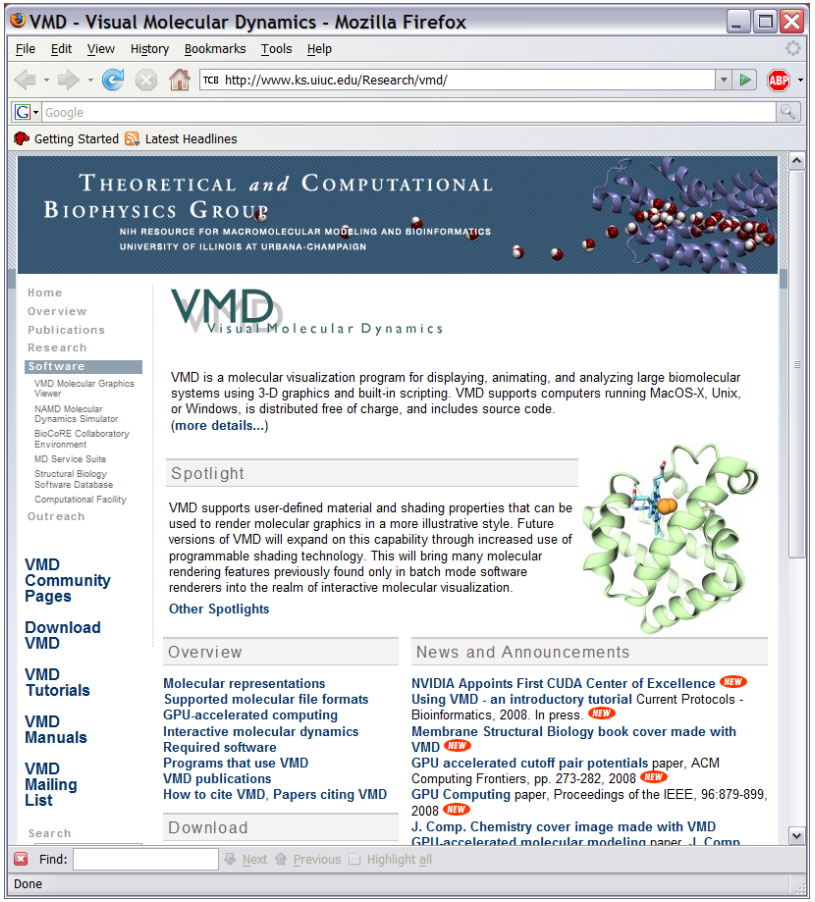
Current Zephyr Restrictions

- Windows XP or Vista
- 32 bit
- Implicit solvent only
- Amber96 force field only
- Standard protein/RNA/DNA molecules only
- For GPU accelerated dynamics:
  - must have supported GPU and drivers

VMD Molecular Graphics Viewer

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





Downloading Zephyr

Install in a Folder with no space characters

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

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 → mdrun (minimize energy)

.gro + .mdp (simulation parameters) → grompp → .tpr + GBSA parameters .agb → mdrun (molecular dynamics)

mdrun → .pdb, .trr

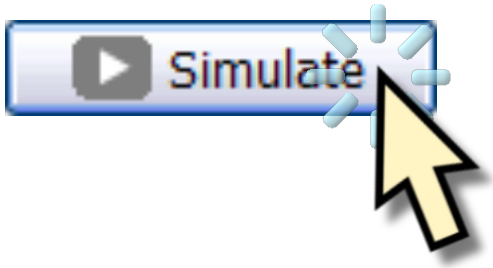
OUTPUT: structure and trajectory

OR

B: Click "Simulate" button in Zephyr

INPUT:
PDB structure file

.pdb

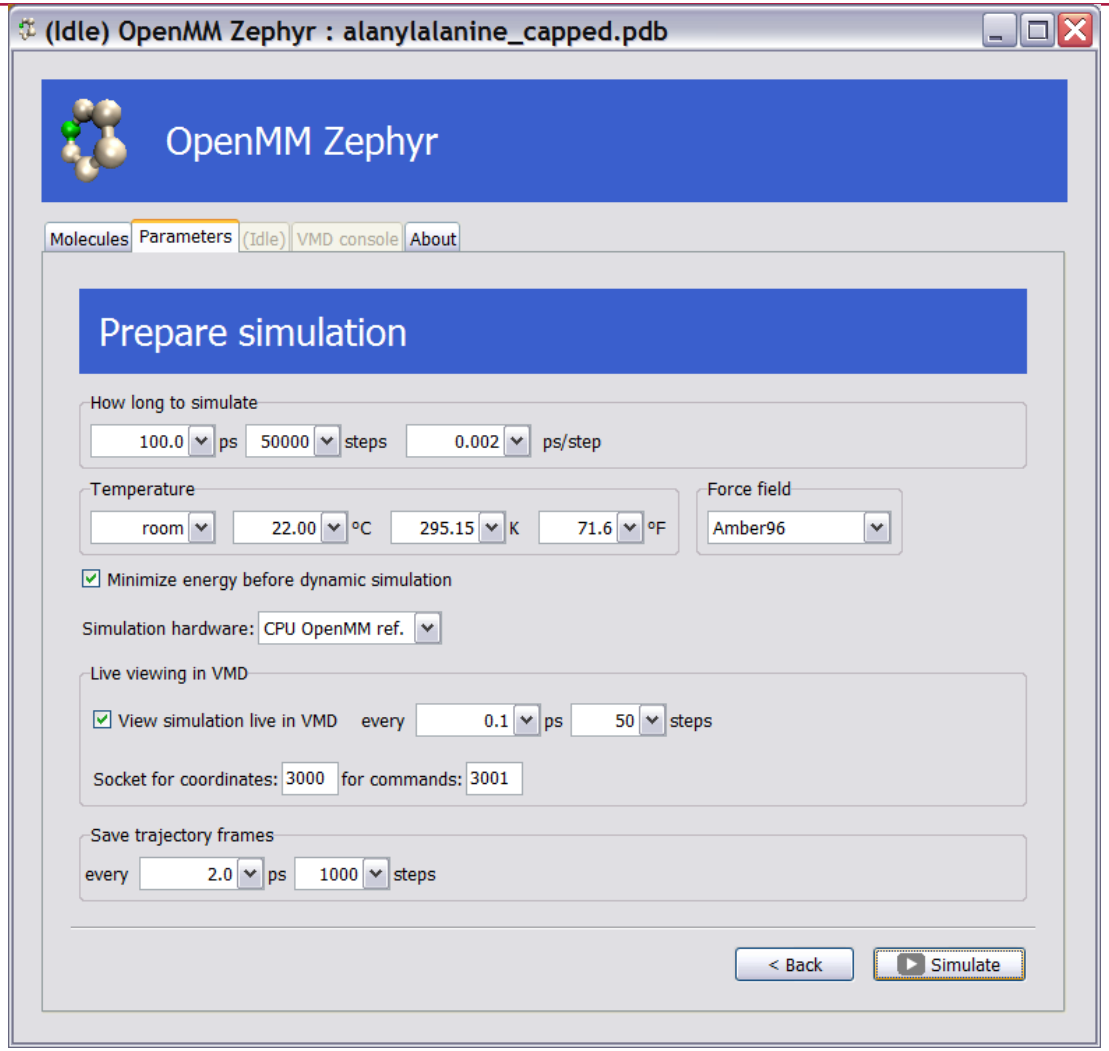


.pdb, .trr

OUTPUT: structure and trajectory

simbios

Simulation Parameters



Example: villin head piece

- based on PDB structure 1vii
- GPU acceleration
- VMD update frequency
- VMD representation

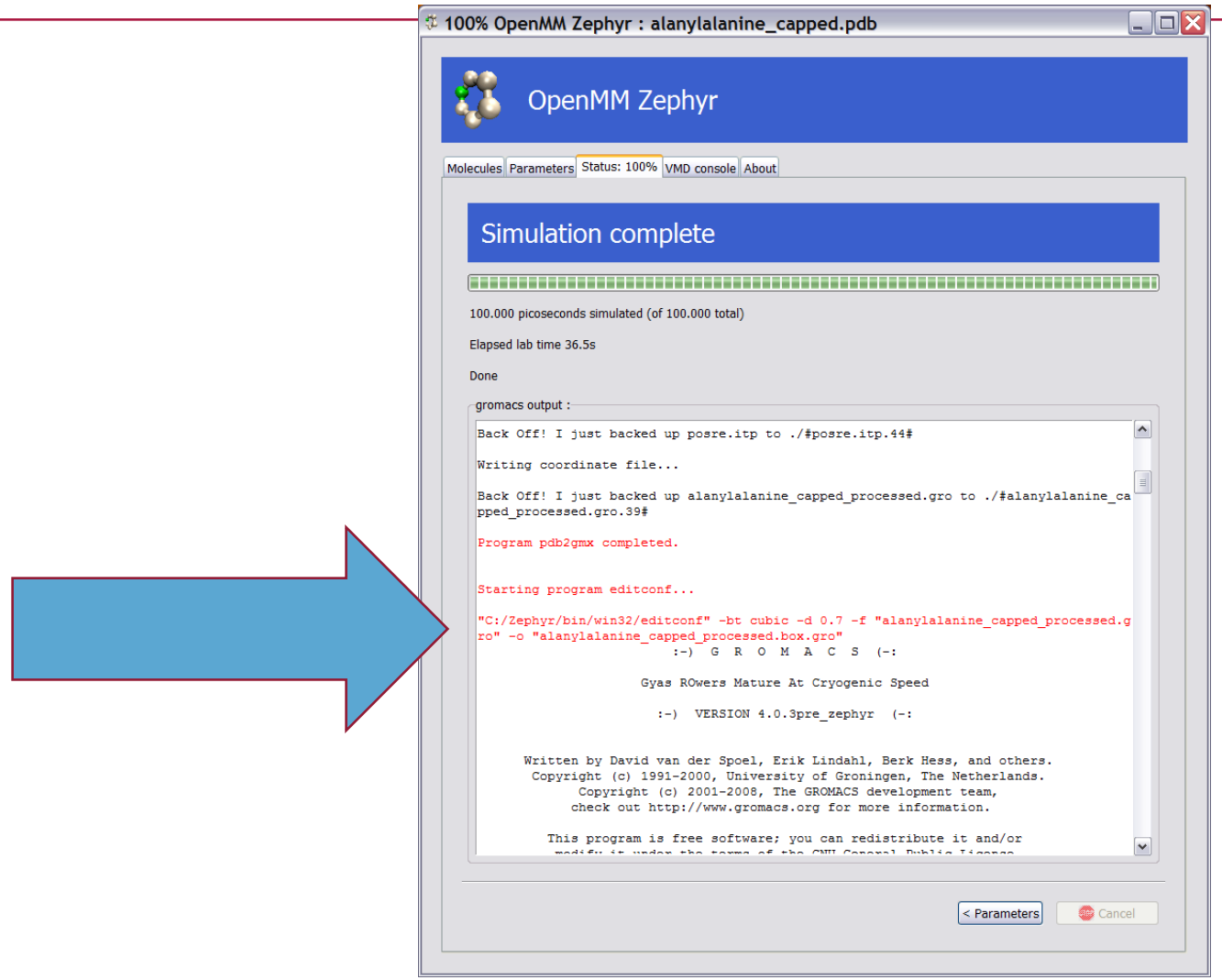
Example: RNA

- gcaa.pdb file of RNA duplex
- Investigate gromacs commands
- Save output from VMD

Saving output

- trajectory and coordinate files in simulation folder
- save from VMD

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



Time Scales in Molecular Mechanics

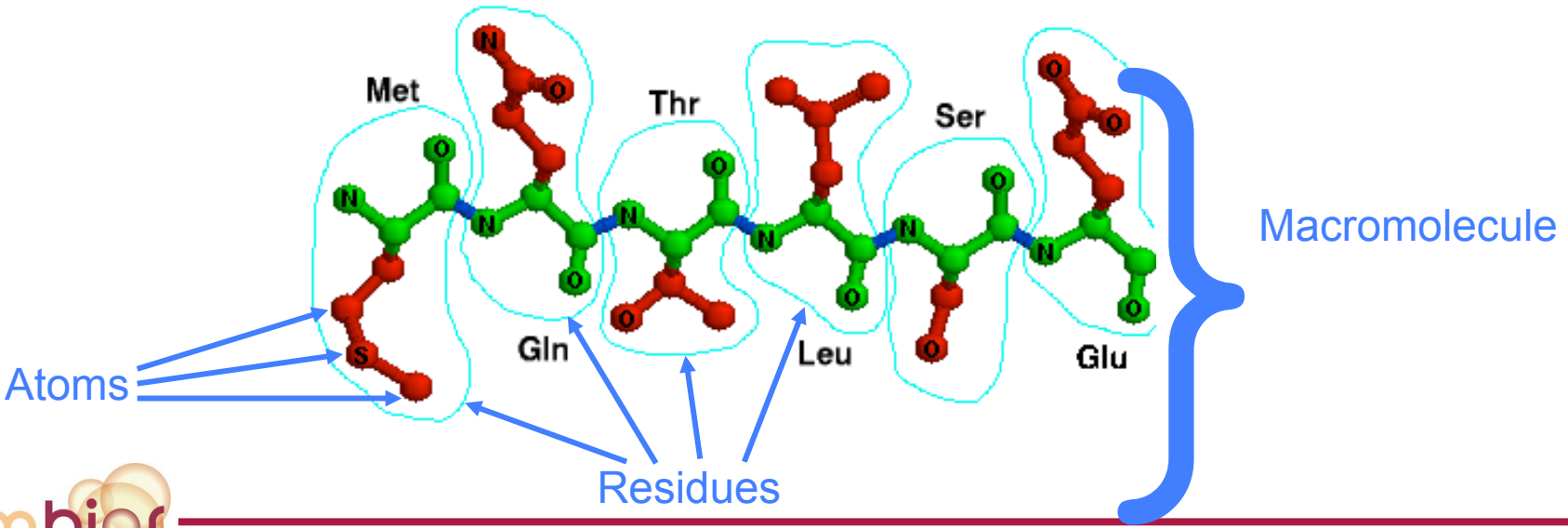
femtosecond	10^{-15} second	bond vibration
picosecond	10^{-12} second	side chain motion
nanosecond	10^{-9} second	protein tumbling
microsecond	10^{-6} second	helix/coil transition
millisecond	10^{-3} second	protein folding
second	10^{0} second	protein synthesis

Example: polyalanine

- viscosity

Macromolecules, residues, and atoms

- RNA, DNA, and protein are large molecules called *macromolecules*
- Macromolecules consist of chains of *residues*



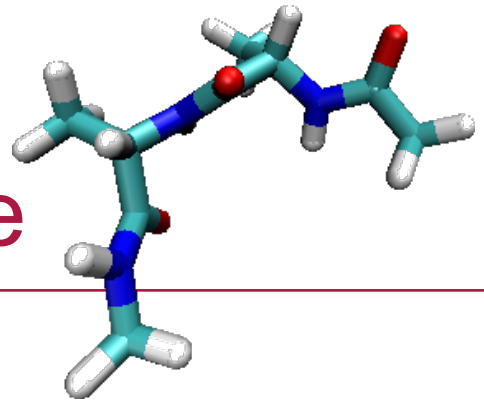
simbios

PDB File format

- Text format: humans can read it too
- Most of the information is atomic coordinates
- PDB files usually contain multiple molecules, including water molecules

```
...
ATOM   1751  N   GLY C 250      32.286   1.882  43.206  1.00 22.00
ATOM   1752  CA  GLY C 250      32.365   1.086  41.969  1.00 21.39
ATOM   1753  C   GLY C 250      31.538   1.735  40.864  1.00 20.79
ATOM   1754  O   GLY C 250      30.621   2.527  41.152  1.00 21.58
...
```

Example: alanylalanine

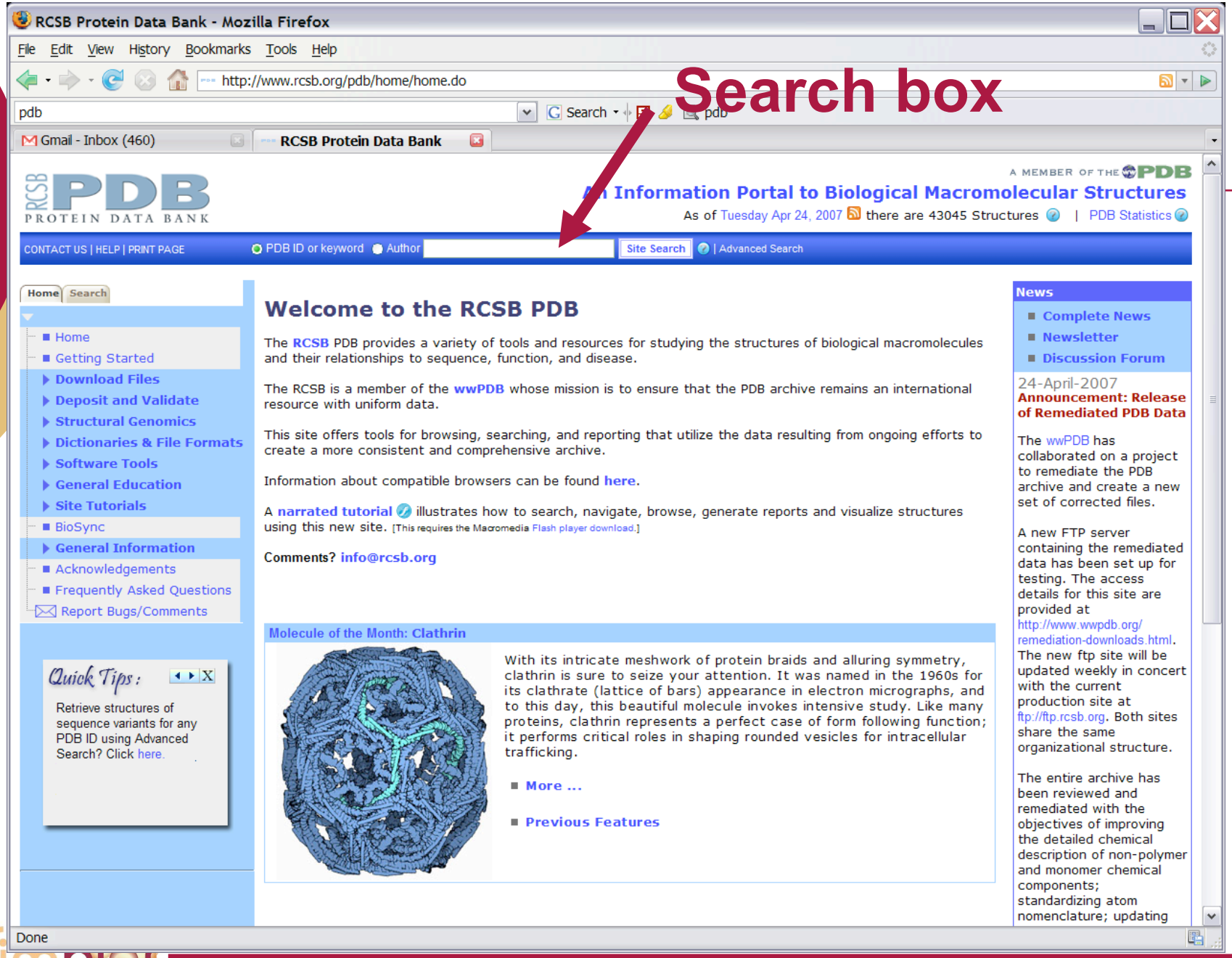


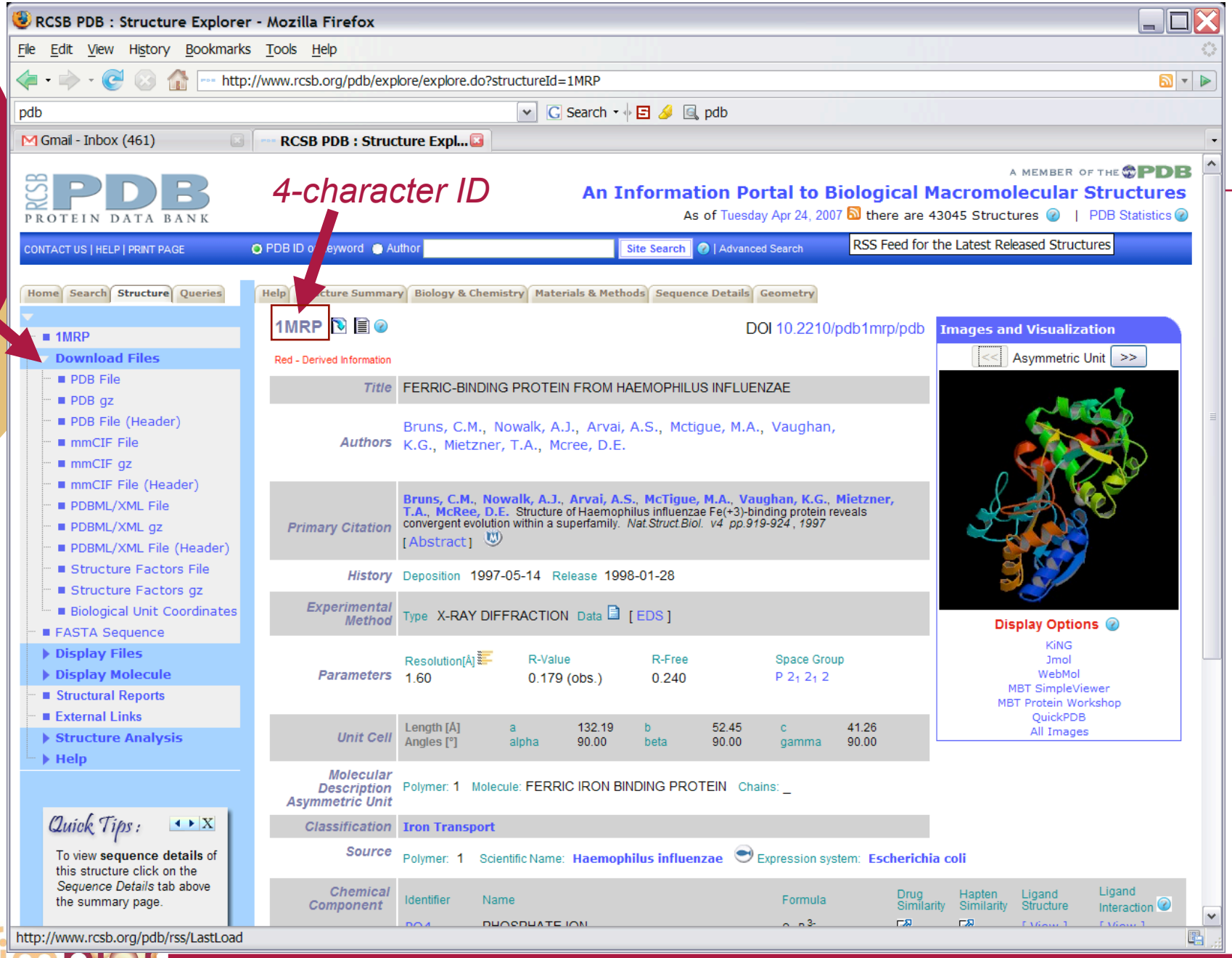
atom name → | residue name → | residue number → | x | y | z

```
ATOM      1  CH3 ACE A   1      22.028   8.496  12.678  1.00 21.24           C
ATOM      2  C   ACE A   1      20.512   8.634  12.719  1.00 24.01           C
ATOM      3  O   ACE A   1      19.827   7.940  13.456  1.00 25.32           O
ATOM      4  N   ALA A   2      19.934   9.503  11.885  1.00 22.94           N
ATOM      5  CA  ALA A   2      18.479   9.595  11.851  1.00 22.62           C
ATOM      6  C   ALA A   2      17.892   8.295  11.241  1.00 25.61           C
ATOM      7  O   ALA A   2      17.014   7.634  11.796  1.00 25.04           O
ATOM      8  CB  ALA A   2      17.994  10.774  11.004  1.00 22.91           C
ATOM      9  N   ALA A   3      18.390   7.947  10.077  1.00 20.56           N
ATOM     10  CA  ALA A   3      17.869   6.776   9.374  1.00 21.22           C
ATOM     11  C   ALA A   3      17.951   5.462  10.174  1.00 28.38           C
ATOM     12  O   ALA A   3      17.008   4.685  10.245  1.00 27.17           O
ATOM     13  CB  ALA A   3      18.564   6.635   8.058  1.00 21.05           C
ATOM     14  N   NME A   4      19.103   5.218  10.736  1.00 23.99           N
ATOM     16  CH3 NME A   4      19.295   4.009  11.474  1.00 20.74           C
```

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
pdb
Search
pdb
Gmail - Inbox (461)
RCSB PDB : Structure Expl...
4-character ID
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
Author
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
Quick Tips :
To view sequence details of this structure click on the Sequence Details tab above the summary page.
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
http://www.rcsb.org/pdb/rss/LastLoad

Preparing PDB Files in a text editor

```
ATOM      1  N   MET A  41       1.177 -10.035  -3.493  1.00  2.04           N
ATOM      2  CA  MET A  41       0.292  -8.839  -3.377  1.00  1.55           C
ATOM      3  C   MET A  41      -0.488  -8.912  -2.063  1.00  1.22           C
ATOM      4  O   MET A  41      -1.039  -9.937  -1.709  1.00  1.32           O
ATOM      5  CB  MET A  41      -0.674  -8.793  -4.565  1.00  1.98           C
ATOM      6  CG  MET A  41      -0.091  -7.889  -5.657  1.00  2.27           C
```

```
ATOM      1  N   NMETA  41       1.177 -10.035  -3.493  1.00  2.04           N
ATOM      2  CA  NMETA  41       0.292  -8.839  -3.377  1.00  1.55           C
ATOM      3  C   NMETA  41      -0.488  -8.912  -2.063  1.00  1.22           C
ATOM      4  O   NMETA  41      -1.039  -9.937  -1.709  1.00  1.32           O
ATOM      5  CB  NMETA  41      -0.674  -8.793  -4.565  1.00  1.98           C
ATOM      6  CG  NMETA  41      -0.091  -7.889  -5.657  1.00  2.27           C
```

atom names

residue names

Some PDB Renaming Rules

- http://chemistry.csulb.edu/ffamber/
- Remove water molecules
- Residue names
  - LYS => LYP, CYS => CYN
  - HIS => HIE, HID, or HIP
  - first residue e.g. ALA => NALA
  - final residue e.g. ASN => CASN
- Atom names
  - final residue O => OC1, OXT => OC2

Submit feature requests and bug reports