Coarse-grained representations in biology

Natural lump sizes:


atoms


residues


base pairs

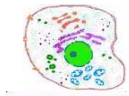

duplexes, alpha helices, beta strands

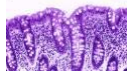

macromolecules

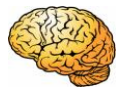

molecular assemblies


organelles

(picometer to micron scale)


cells


tissues


organs


limbs


organisms


societies


ecosystems

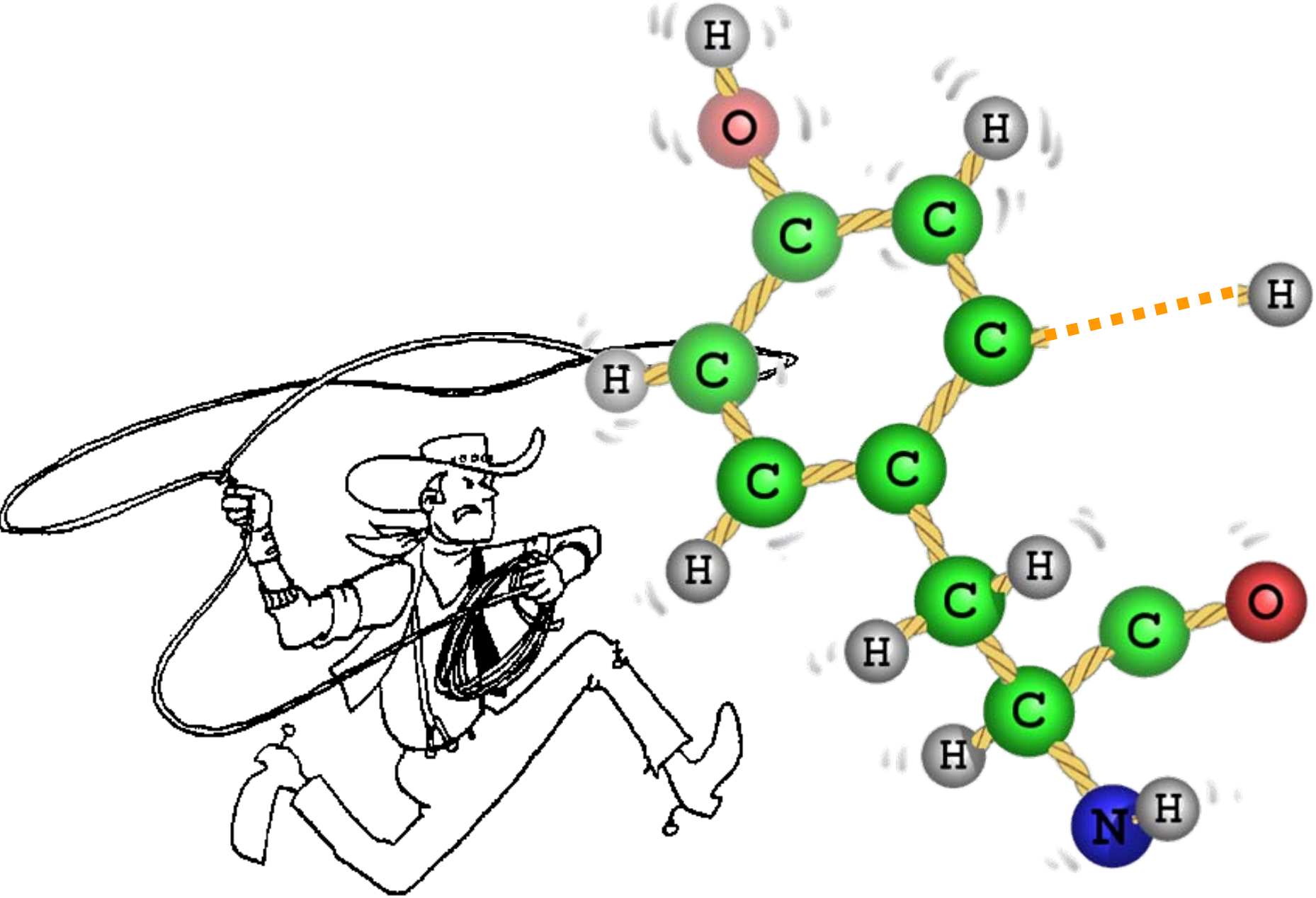

biospheres

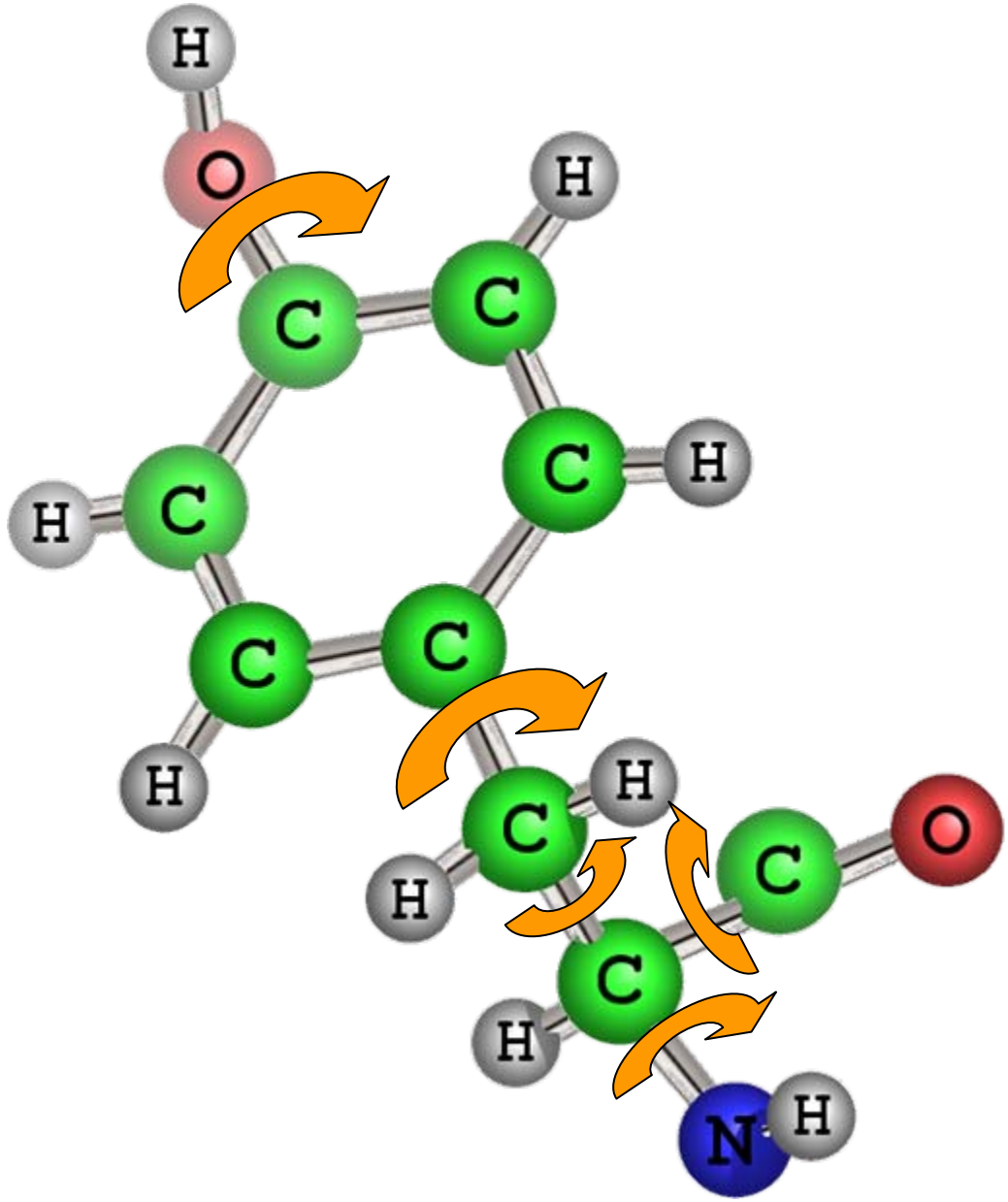

intergalactic federations

(micron to mega-light-year scale)

Coarse-grained simulation of atom-based molecular models

- Full atom force field
- coarse-grained motion permits longer integration time steps
- Increasing coarseness
  - full Cartesian
  - torsion only (internal coordinate)
  - full rigid


H
O
H
C
C
H
C
C
H
H
C
C
H
H
C
H
C
O
H
C
H
N
H

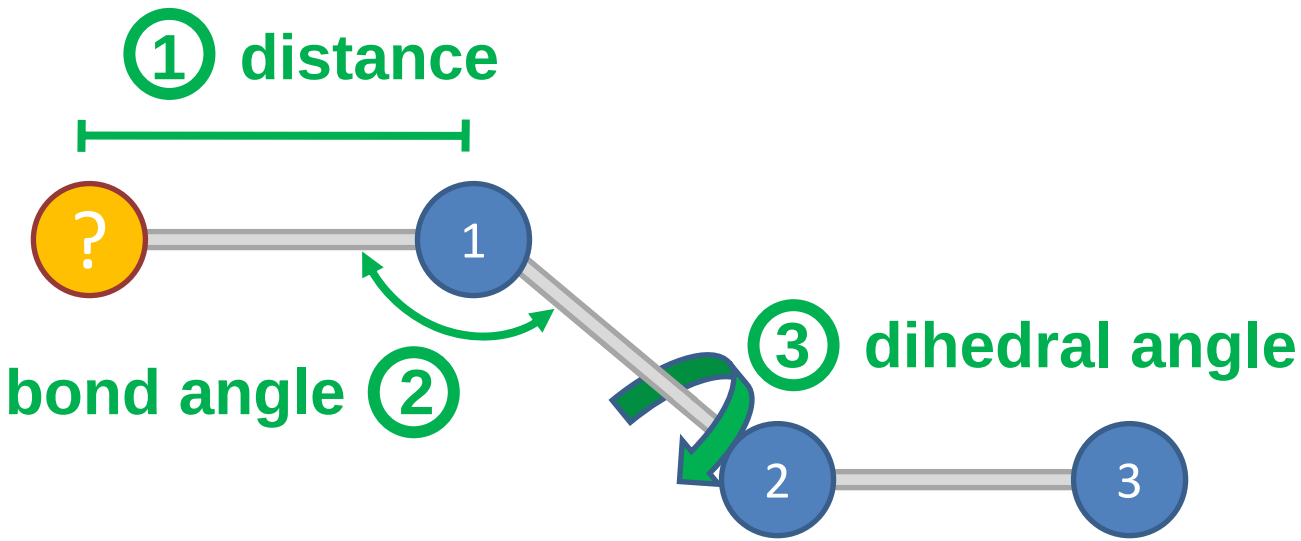

H
O
H
C
C
C
H
C
H
C
C
H
C
H
H
C
O
C
H
H
C
H
N
H

Molecule internal coordinates

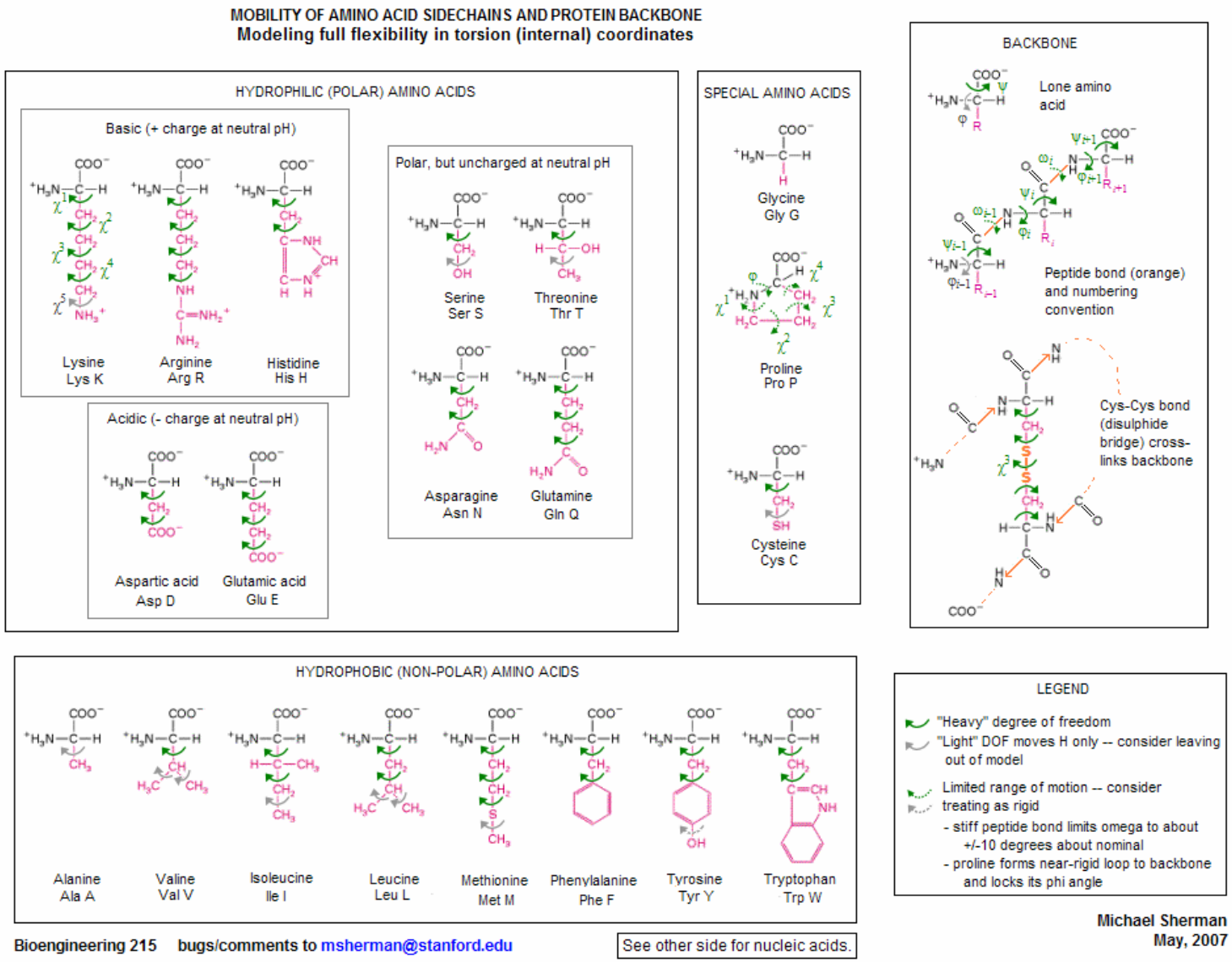
- Bond lengths, bond angles, dihedral angles
- Mandatory for Compound construction
- Optional for Compound simulation

Internal coordinates

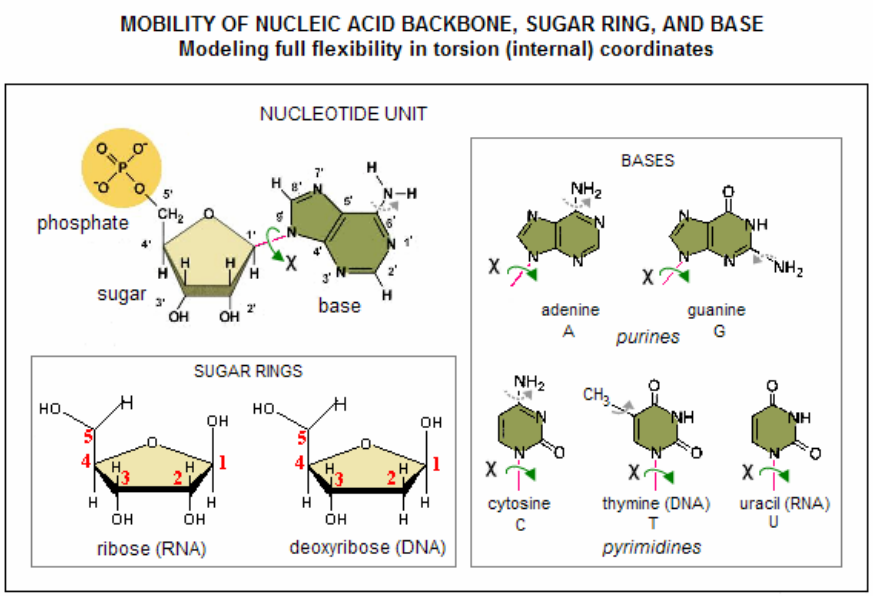
- Instead of explicit Cartesian X, Y, Z coordinates for an atom
- Given three other atom positions, one distance, one bond angle, and one dihedral angle

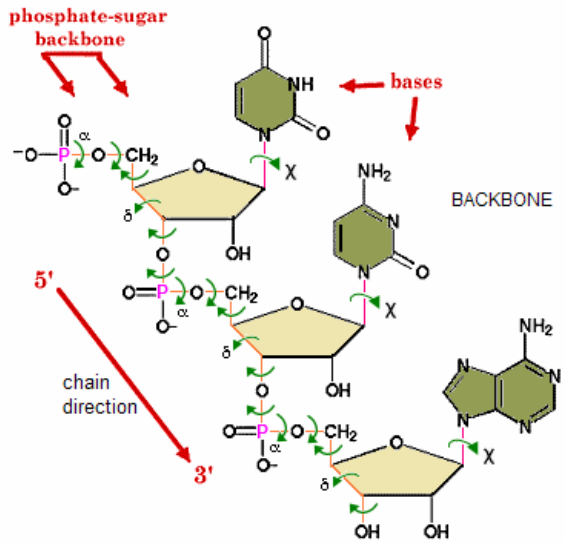


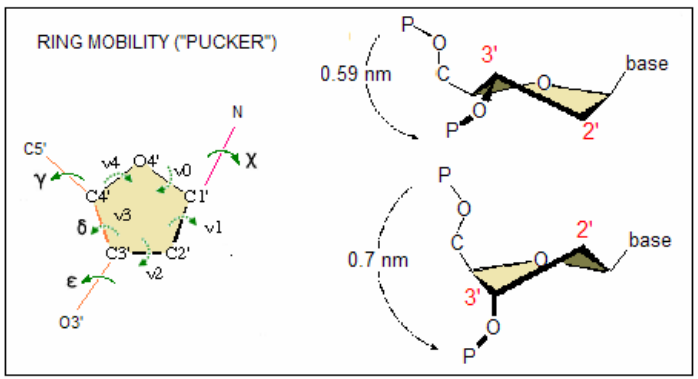
Protein Internal Coordinates

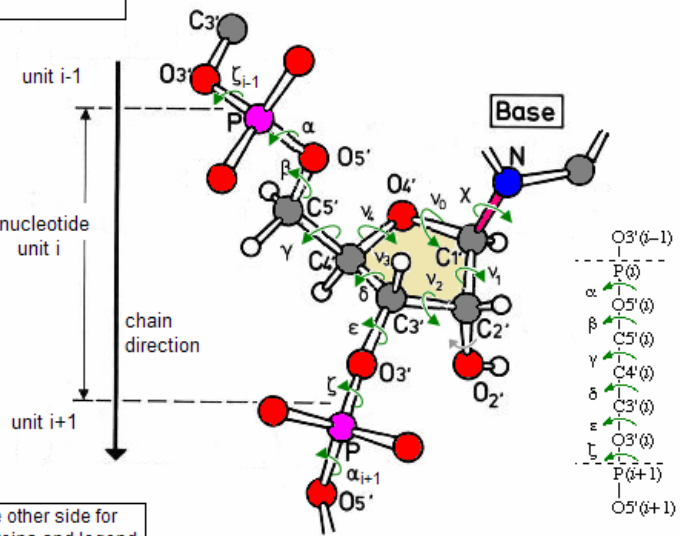


Nucleic Acid Internal Coordinates









Michael Sherman
May, 2007
Bioengineering 215 bugs/comments to msherman@stanford.edu

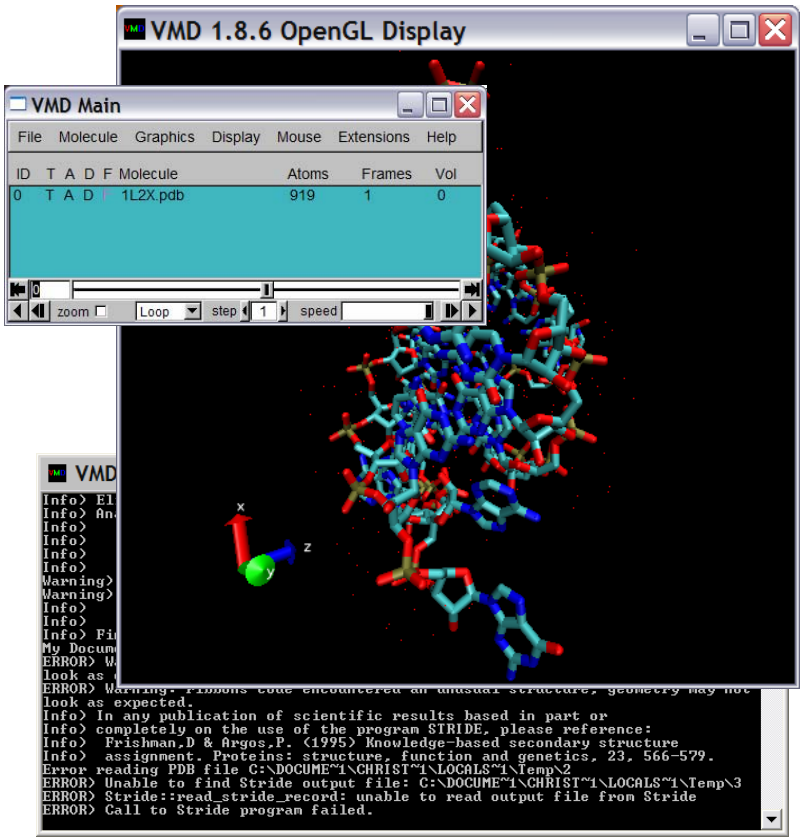
See other side for proteins and legend.

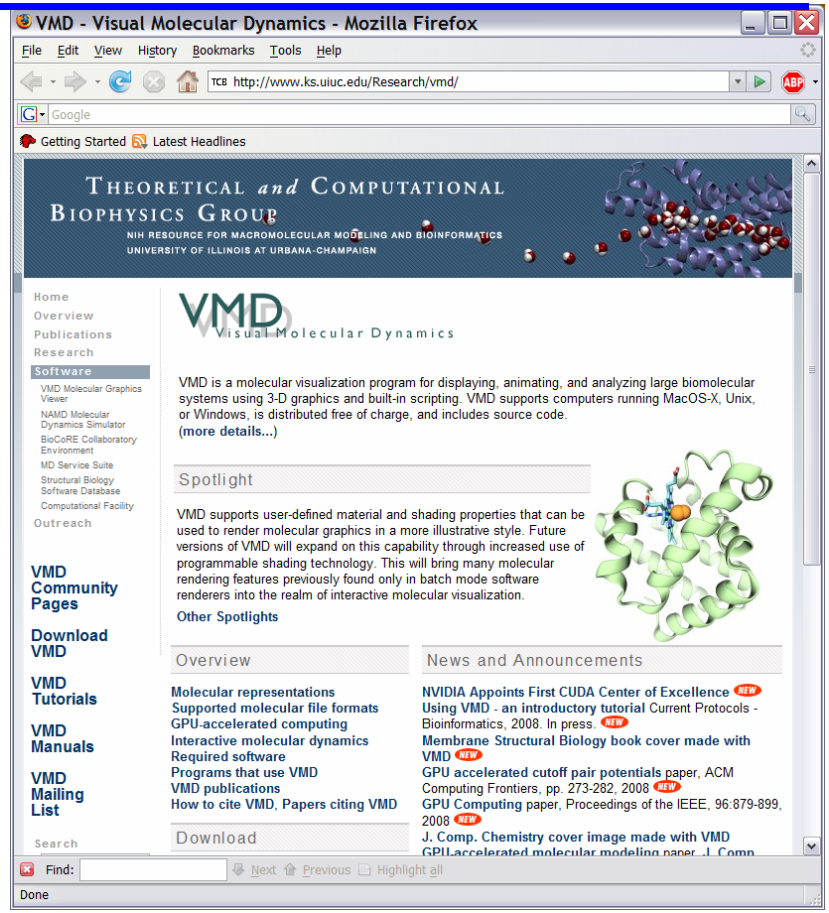
Mobility Exercises

- Simulate rigid protein
- Create second protein
- Convert to full Cartesian model
- Which way is faster?
- Simulate multi-grain AdenylateMobilitiesVTK example

VMD Molecular Graphics Viewer

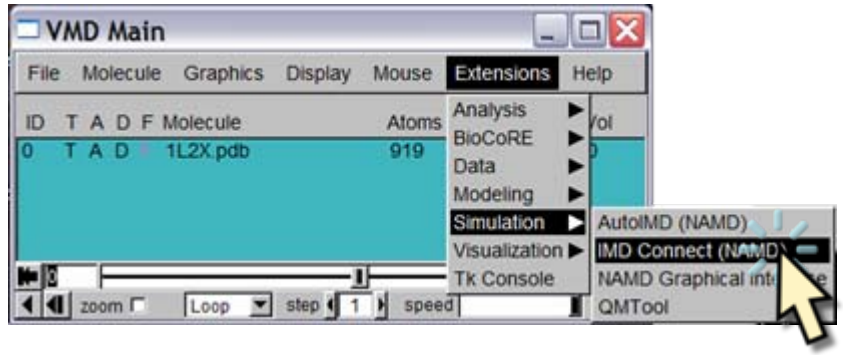
http://www.ks.uiuc.edu/Research/vmd/

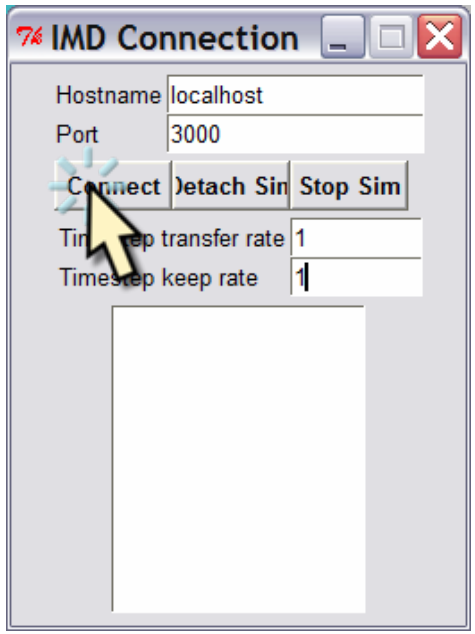




Viewing SimTK simulation in VMD

1. Start SimTK simulation using PeriodicVMDReporter
2. Load molecule in VMD
   1. VMD Main->File->New Molecule…
   2. Molecule File Browser->Browse
   3. Molecule File Browser->Load
3. Connect VMD to SimTK
   1. VMD Main->Extensions->Simulation->IMD Connect (NAMD)
   2. In IMD Connection window
      1. Hostname: localhost
      2. Port: 3000 (match argument to PeriodicVMDReporter)
      3. Timestep transfer rate: 1 (to see every reported step) PRESS ENTER
      4. Timestep keep rate: 1 (to remember every step for replay value) PRESS ENTER
      5. Press “Connect” button



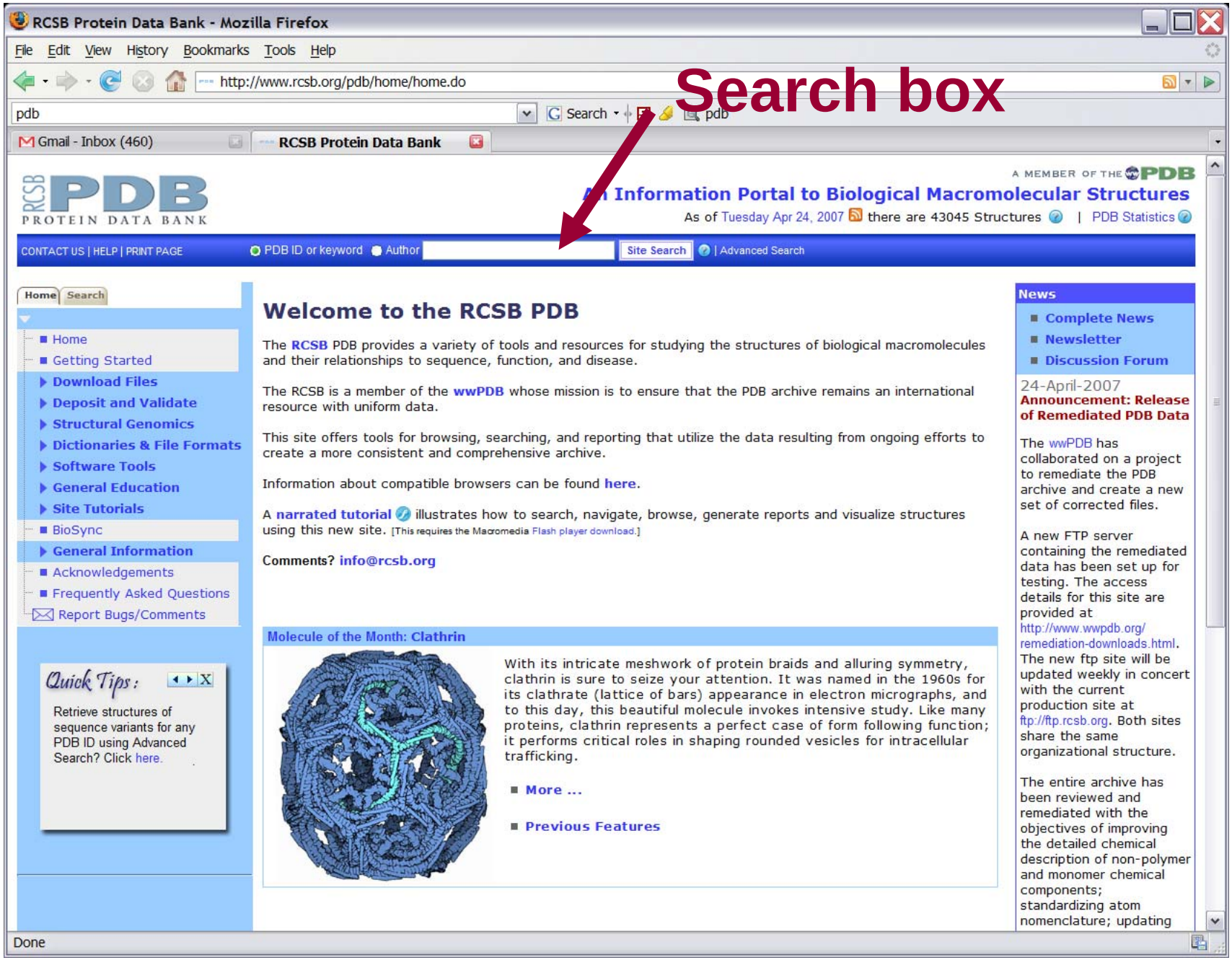


RNA Mobility Exercises

- Simulate multi-grain AdenylateMobilities example program and connect to VMD

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...

RCSB PDB PROTEIN DATA BANK

4-character ID

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 | Keyword | 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]	R-Value	R-Free	Space Group
	1.60	0.179 (obs.)	0.240	P 2_1 2_1 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
PO4	PHOSPHATE ION	O_4P^{3-}			[View]	[View]

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

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
…
```

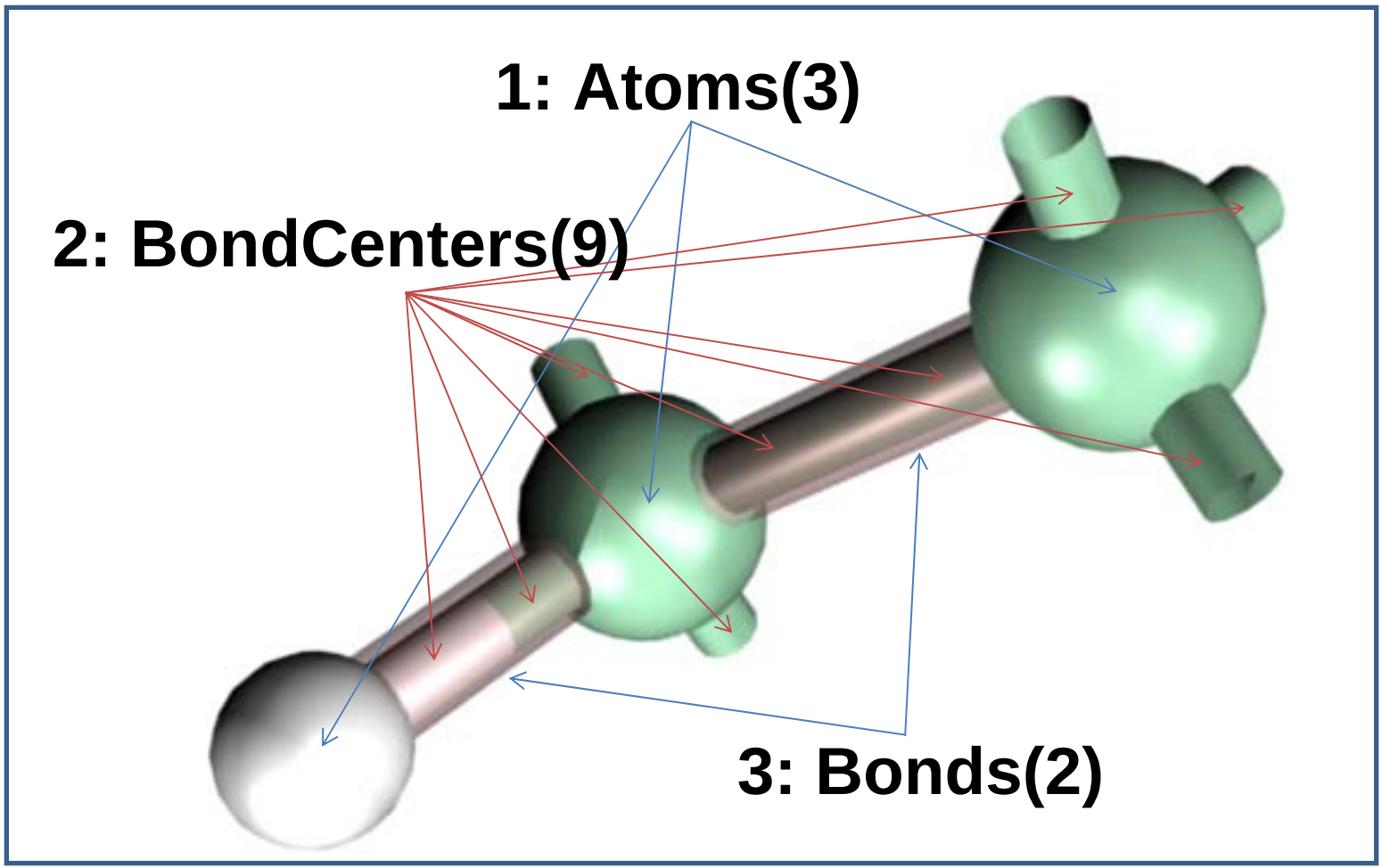
LoadPdb Exercises

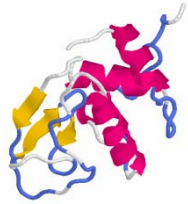
- Download 1AKG from Protein Data Bank http://www.rcsb.org/pdb
- Simulate 1AKG using Example
- Simulate 1AKG using stream constructor

Custom Molecule Construction

- BondCenters, Atoms, Bonds, Compounds
- Inboard bond center
- The first few atoms
- Ring closing bonds

Molecule Construction nomenclature





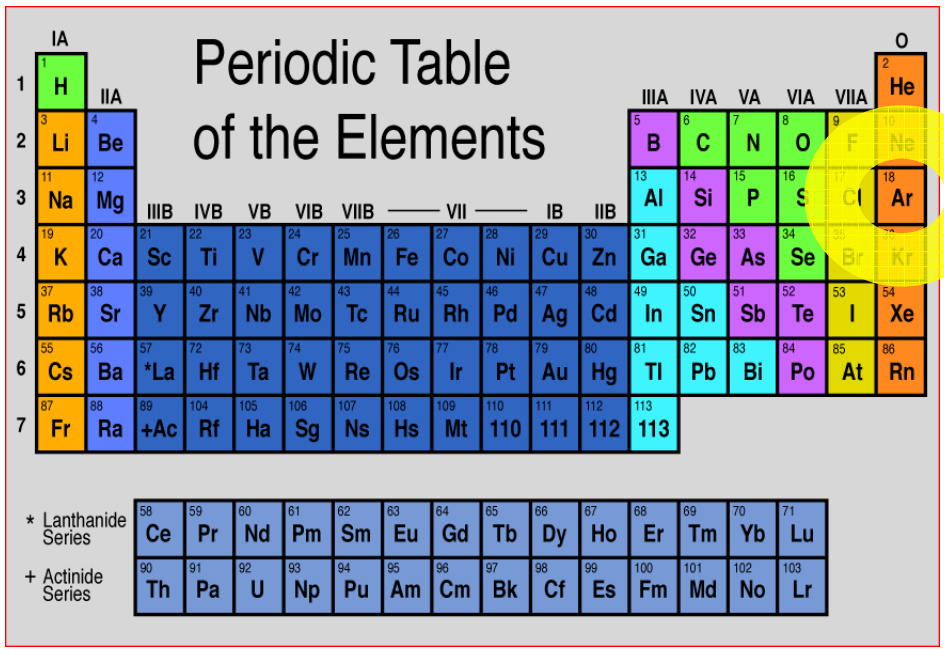
	Type	Description	Protein example	RNA example	Argon example
structure	Compound	Molecule or part of molecule	hen egg-white lysozyme	yeast phenylalanine transfer RNA	argon
structure	Atom	one particular atom in a molecule	serine 3 nitrogen	guanosine 3, carbon C3’	argon
	Biotype	links Atoms to ChargedAtom Types	serine N	guanylate C3’	argon
force field	ChargedAtom Type	Atom type with a particular charge	serine N: charge -0.42	guanylate C3’: charge +0.2022	argon: charge zero
force field	AtomClass	Atom type	sp^2 amide nitrogen (N)	tetrahedral carbon (CT)	argon

Biotype

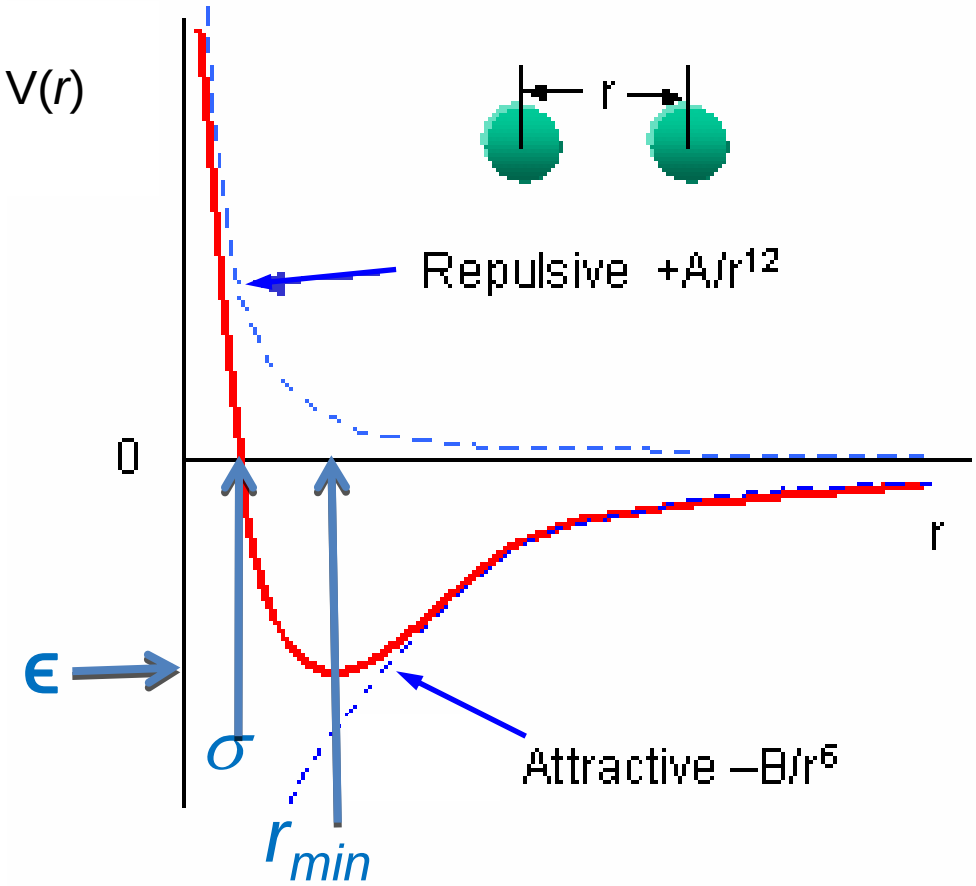
- Bridge between structural model and force field implementation
- Concept borrowed from TINKER molecular dynamics package

Argon

- Inert “noble” gas
- Has no chemistry
- van der Waals forces only



Lennard-Jones potential



$$r_{\min} = 2^{1/6}\sigma$$

Four ways to express length parameter:

	σ (r_0)	r_{min}
radius = diameter/2		
diameter		

$$V(r) = 4\epsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right]$$

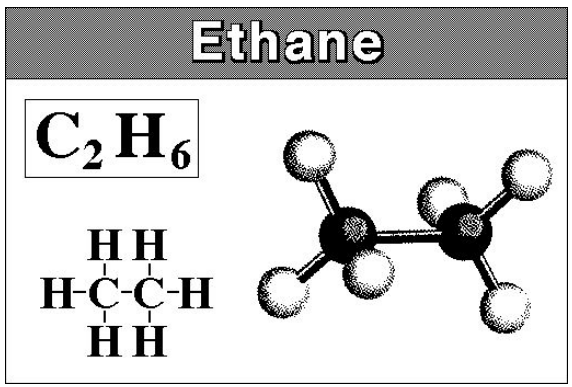
$$V(r) = \epsilon\left[\left(\frac{r_{min}}{r}\right)^{12} - 2\left(\frac{r_{min}}{r}\right)^{6}\right]$$

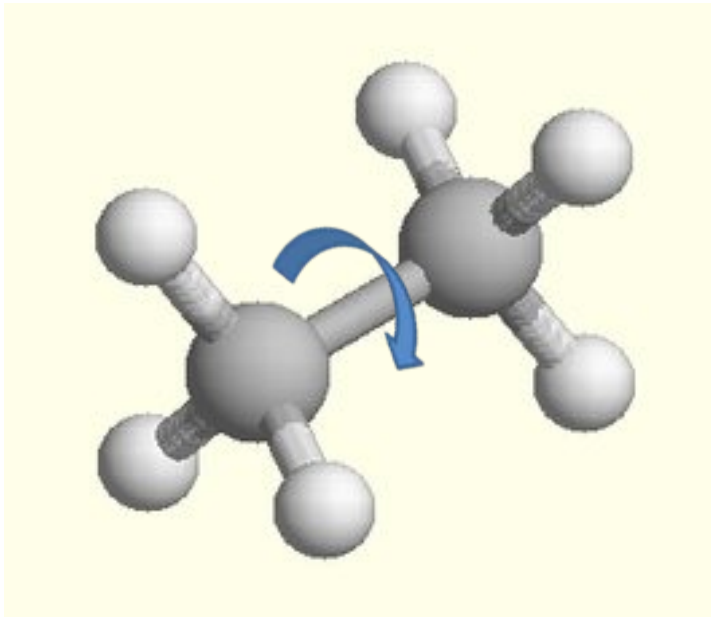
Argon Exercises

- Compile and run TwoArgons example program
- Add a third argon atom.
- What happens when atoms start too close together? Too far apart?
- Increase the attractive force between atoms ten-fold
- Why does this example lack temperature and enemy minimization stanzas?

Ethane

- Two carbons, six hydrogens
- One dihedral degree of freedom
- 3 kcal/mol energy barrier to rotation





Ethane Exercises

- Compile and run TwoEthanes example program
- Add a third ethane molecule

Submitting Bug Reports