Protein Bioinformatics
260.655
Computer lab #1
Friday, April 11, 2008
Sean Prigge and Ingo Ruczinski

| Goals: | Approx. Time |
|---|---|
| [1] Use the Protein Data Bank PDB website. | 10 minutes |
| [2] Use the WebMol Viewer. | 20 minutes |
| [3] Use the Cn3D on the NCBI website. | 20 minutes |
| [4] Use the ExPASy site including DeepView, the Swiss-Pdb Viewer. | 15 minutes |

---

[1] Use the PDB website (http://www.rcsb.org/pdb/home/home.do)

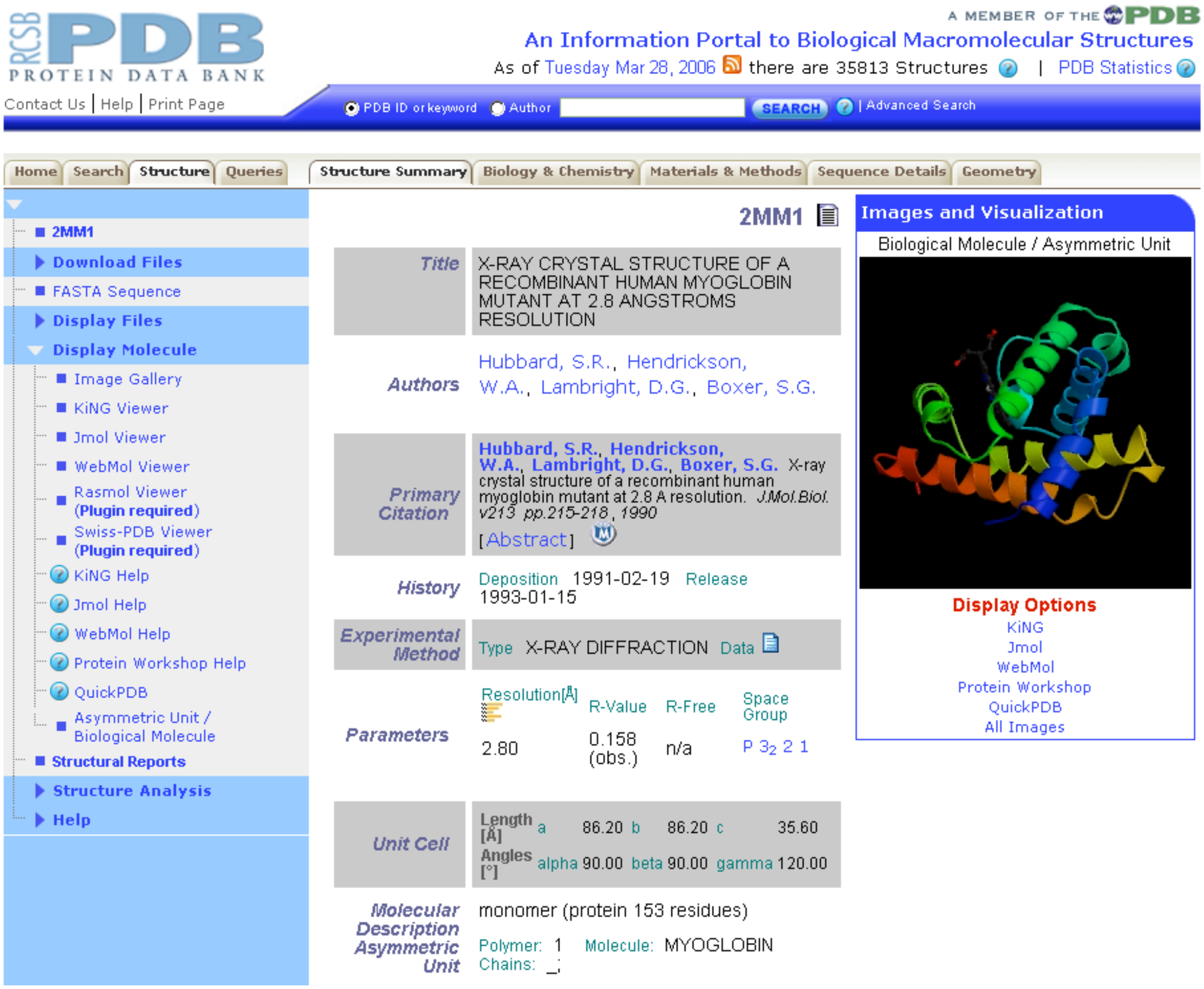
1a. Find a structures of human myoglobin:

Strategy 1. type ‘myoglobin’ in the search box
- how many structures are there?

Strategy 2 type ‘human myoglobin’ or ‘sapiens myoglobin’
-do these searches yield the same results?

1b. Click on structure 2MM1

What information is displayed on the top page for this structure?
What is the resolution of the structure?
What is the R-factor?
Are there any other molecules bound to this protein?



**2MM1**

| | |
|---|---|
| Title | X-RAY CRYSTAL STRUCTURE OF A RECOMBINANT HUMAN MYOGLOBIN MUTANT AT 2.8 ANGSTROMS RESOLUTION |
| Authors | Hubbard, S.R., Hendrickson, W.A., Lambright, D.G., Boxer, S.G. |
| Primary Citation | **Hubbard, S.R., Hendrickson, W.A., Lambright, D.G., Boxer, S.G.** X-ray crystal structure of a recombinant human myoglobin mutant at 2.8 A resolution. *J.Mol.Biol. v213 pp.215-218 , 1990*<br>[Abstract] |
| History | Deposition 1991-02-19 Release 1993-01-15 |
| Experimental Method | Type X-RAY DIFFRACTION Data |

| Parameters | Resolution[Å] | R-Value | R-Free | Space Group |
|---|---|---|---|---|
| | 2.80 | 0.158 (obs.) | n/a | P $3_2$ 2 1 |

| Unit Cell | | | | | | |
|---|---|---|---|---|---|---|
| Length [Å] | a | 86.20 | b | 86.20 | c | 35.60 |
| Angles [°] | alpha | 90.00 | beta | 90.00 | gamma | 120.00 |

Molecular Description Asymmetric Unit: monomer (protein 153 residues)

Polymer: 1 Molecule: MYOGLOBIN Chains: _;

1c. Now look at the tools on the left hand side

How can you get the amino acid sequence of this protein?

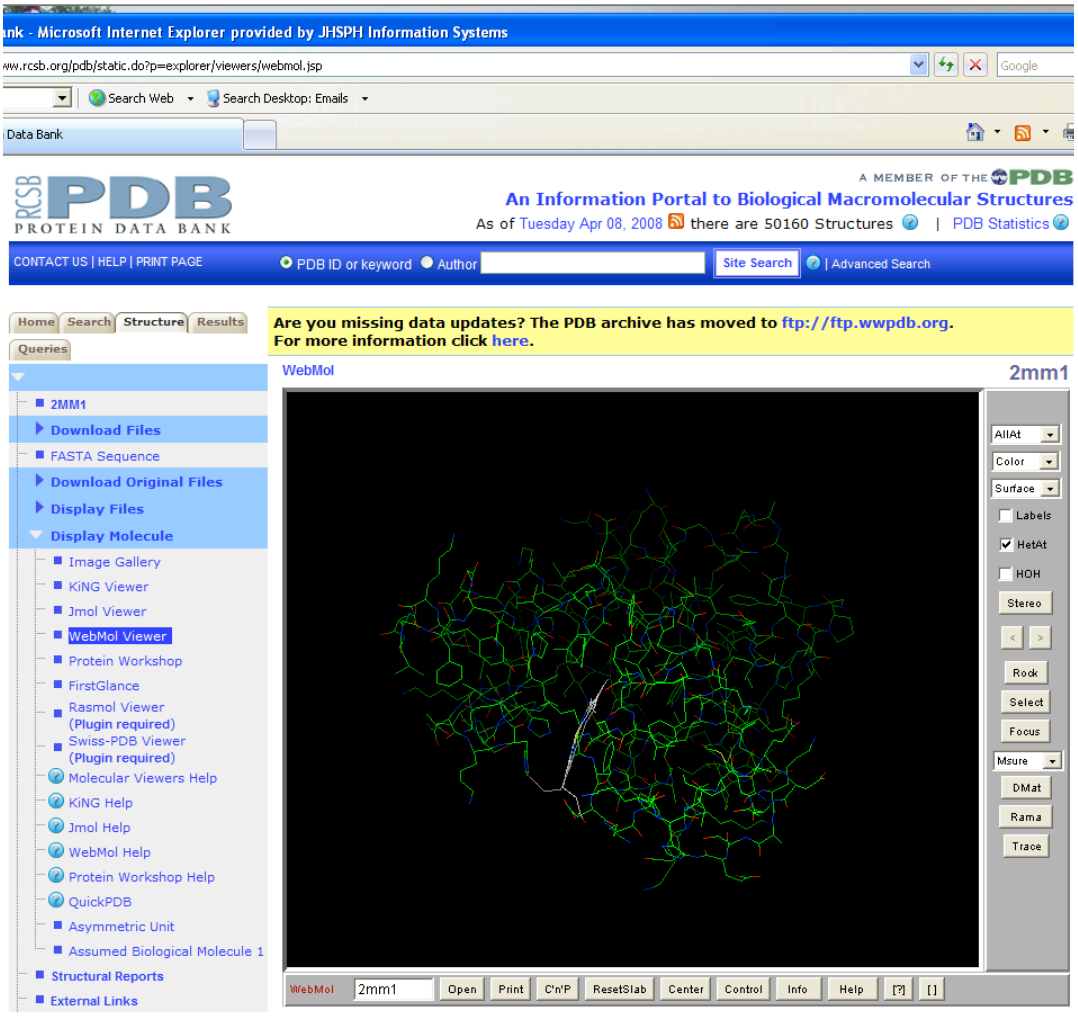
- Advanced: Use the sequence for an advanced sequence search.

Look at ‘Download Original Files’ and ‘Display Files’

- This is how you can save a full PDB record to your computer for later use

Now look at the ‘Display Molecule’ tab and click on ‘WebMol Viewer’

[2] Experiment with the WebMol viewer



An Information Portal to Biological Macromolecular Structures

As of Tuesday Apr 08, 2008 there are 50160 Structures | PDB Statistics

**Are you missing data updates? The PDB archive has moved to ftp://ftp.wwpdb.org. For more information click here.**

Use the top right tool to show mainchain or backbone

Use the 'color' tool to color by: secondary struct, B-factor, or buried/Exposed atoms

Use the 'surface' tool. What is the difference between SASA and VDW?

- solvent available surface area, Van der Waals

What do the 'HetAt' and 'HOH' tools do?

-HeteroAtom, Water

Use the 'Select 'and 'Focus' tools to explore the heme group. What atom is in the center of the heme?

Use the 'Msure' (measure) tool to look at distances and angles in the heme.

what is the distance in angstroms between the N-terminal and C-terminal atoms?

What is the 'Dmat' tool and what does it tell us?

What is the 'Rama' tool? Click on alpha-helical and non-helical residues.

Advanced: Experiment with the tools on the bottom of the screen.

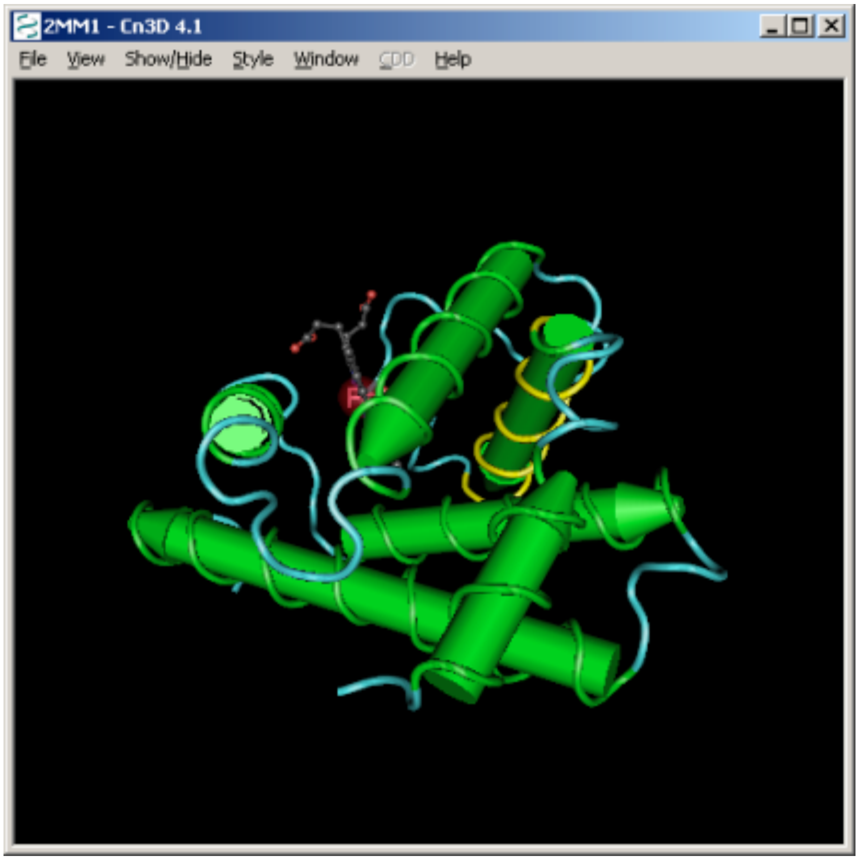
[3] Use the NCBI website. (http://www.ncbi.nlm.nih.gov/)

3a. Find myoglobin. Strategy #1: enter ‘myoglobin’ as a search term at the front page of NCBI.

3b. NCBI strategy #2. Use the taxonomy browser; select ‘Homo sapiens’; click on the ‘protein’ link. Now add the word ‘myoglobin’ in the search box.

3c. NCBI strategy #3: Do the same as in 3b, but select ‘Structure links’ in the ‘Display’ box. This should yield 6 hits including 2MM1.

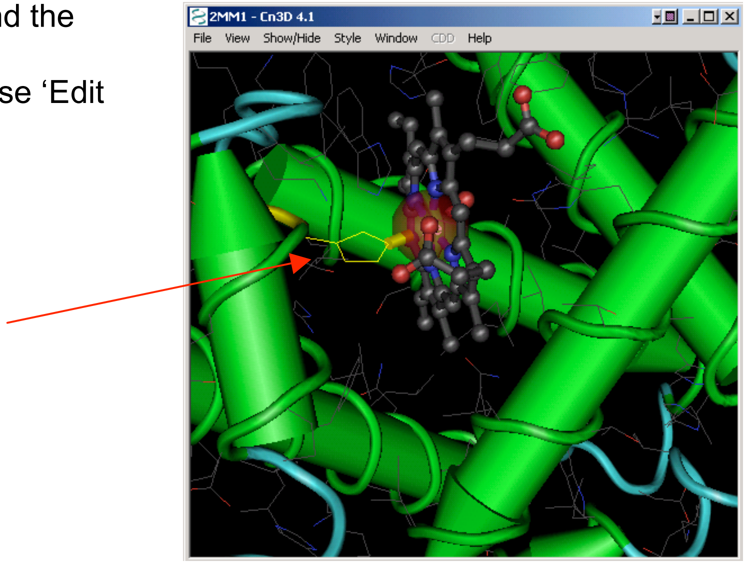
3d. Click on 2MM1 and view structure 2MM1 with Cn3D. A screen capture is shown here:



3e. Highlight an alpha helix in the primary amino acid sequence (shown below) and view the corresponding structure.

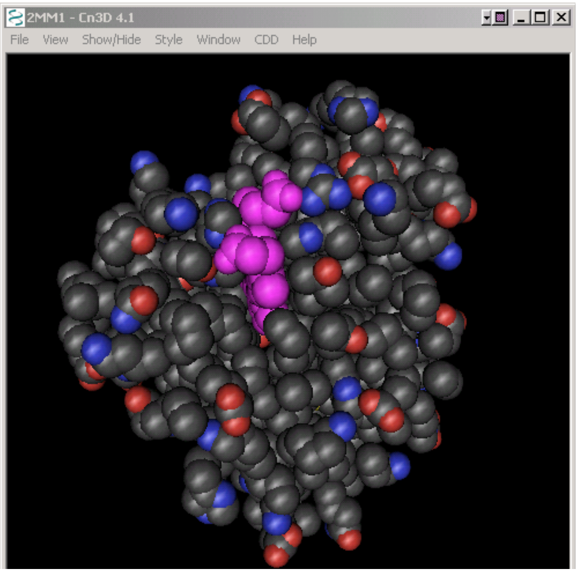
2MM1 g l s d g e w q l v l n v w g k v e a d i p g h g q e v l i r l f k g h p e t l e k f d r f k h l k s e d e m k a s e d l k k h g a t v l t a l g g i l k k k g h h e a e

3f  View the side chains and find the distal histidine (binds to the iron of the heme group).  Use ‘Edit global style’ under ‘style’.

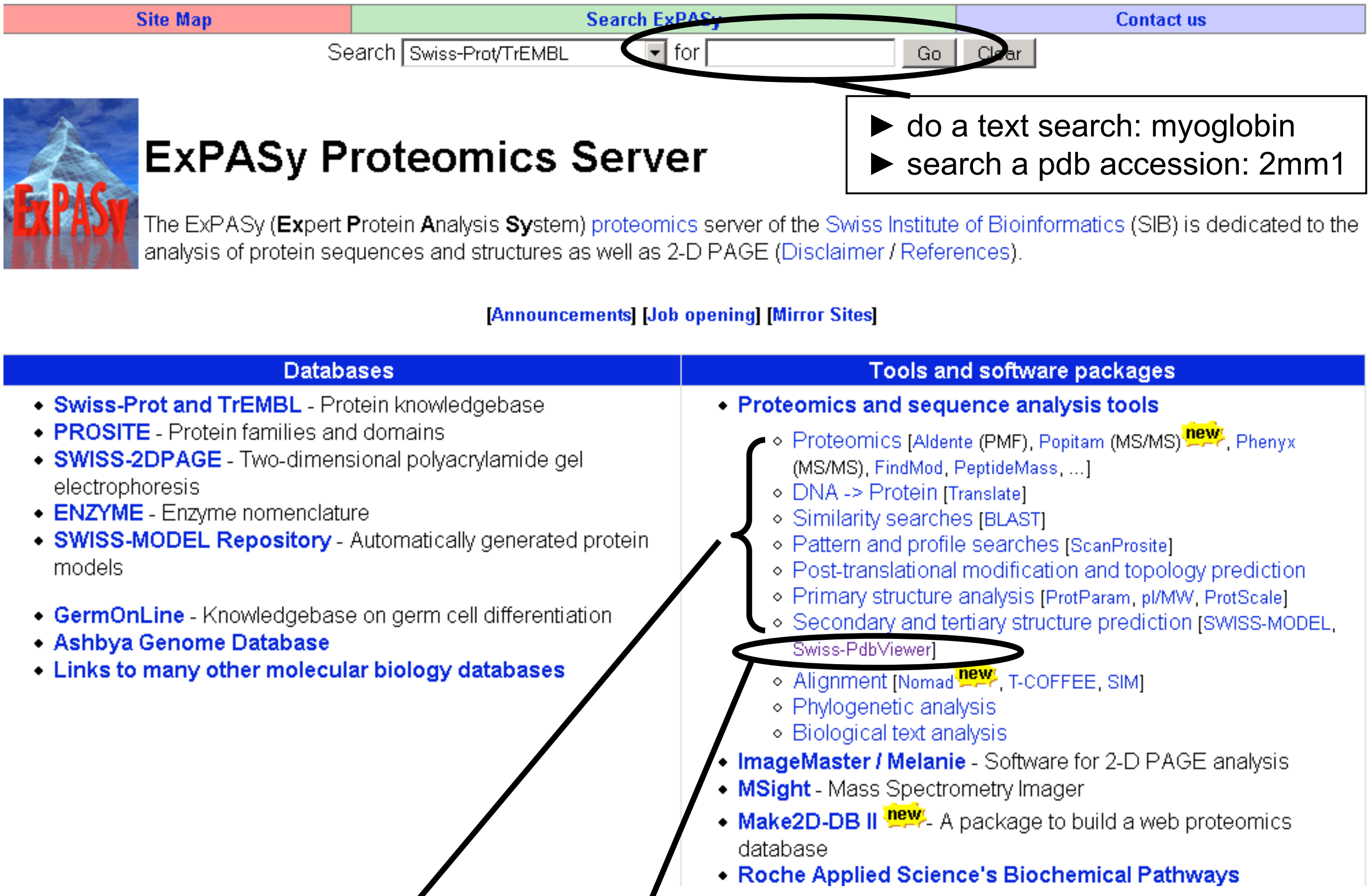


3g.  Find the conserved proline in the sequence ‘HPE’ (after the second helix). Highlight this proline in the sequence window and find it in the myoglobin structure.  Think of a structural reason why this proline is conserved in myoglobin and hemoglobin sequences.

3h.  If there is time, try to reproduce this space-filling model using the ‘Edit global style’ menu under ‘style’.  Hint: to color the heme purple, select ‘object’ under ‘color scheme’ for heterogens.



[4] Use the ExPASy website



## ExPASy Proteomics Server

The ExPASy (**Ex**pert **P**rotein **A**nalysis **Sy**stem) proteomics server of the Swiss Institute of Bioinformatics (SIB) is dedicated to the analysis of protein sequences and structures as well as 2-D PAGE (Disclaimer / References).

**[Announcements] [Job opening] [Mirror Sites]**

**Databases**

- **Swiss-Prot and TrEMBL** - Protein knowledgebase
- **PROSITE** - Protein families and domains
- **SWISS-2DPAGE** - Two-dimensional polyacrylamide gel electrophoresis
- **ENZYME** - Enzyme nomenclature
- **SWISS-MODEL Repository** - Automatically generated protein models

- **GermOnLine** - Knowledgebase on germ cell differentiation
- **Ashbya Genome Database**
- **Links to many other molecular biology databases**

**Tools and software packages**

- **Proteomics and sequence analysis tools**
    - Proteomics [Aldente (PMF), Popitam (MS/MS) new, Phenyx (MS/MS), FindMod, PeptideMass, ...]
    - DNA -> Protein [Translate]
    - Similarity searches [BLAST]
    - Pattern and profile searches [ScanProsite]
    - Post-translational modification and topology prediction
    - Primary structure analysis [ProtParam, pI/MW, ProtScale]
    - Secondary and tertiary structure prediction [SWISS-MODEL, Swiss-PdbViewer]
    - Alignment [Nomad new, T-COFFEE, SIM]
    - Phylogenetic analysis
    - Biological text analysis
- **ImageMaster / Melanie** - Software for 2-D PAGE analysis
- **MSight** - Mass Spectrometry Imager
- **Make2D-DB II** new - A package to build a web proteomics database
- **Roche Applied Science's Biochemical Pathways**

► do a text search: myoglobin
► search a pdb accession: 2mm1

► download Swiss-Pdb Viewer here

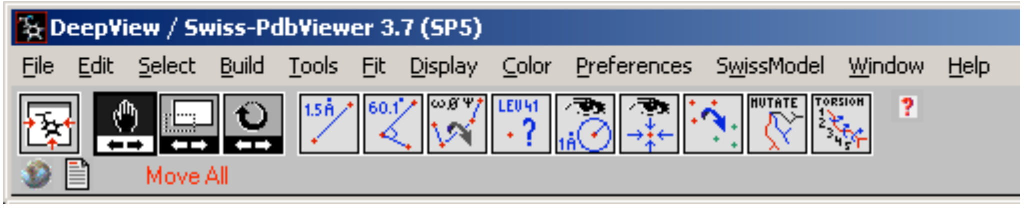
► explore a variety of these web-based tools for primary, secondary, and tertiary structure analysis

[4] Use DeepView, the Swiss-Pdb Viewer

[4a] Download DeepView from ExPASy (onto PC, Mac, Linux)

[4b] Save 2MM1.pdf onto your hard drive; open in DeepView.

[4c] Use each of the tools below.



1 2 3 4 5 6 7 8 9 10 11 12 13

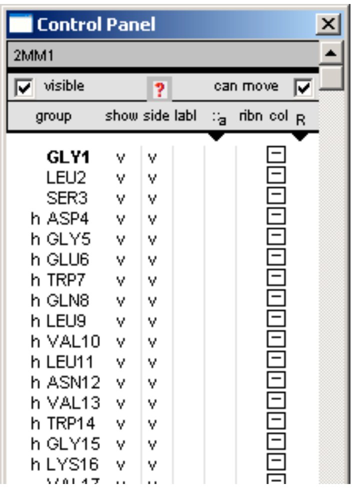
1 center

2 move (translate)
3 zoom in/out
4 rotate

5 click, select two atoms, determine distance in angstroms
6 measure bond angles (pick center atom, then two more atoms)
7 measure dihedral angles (ω, ϕ, φ)(omega, phi, psi) from a selected atom.
8 identify an atom (and the group to which it belongs).
Type: CA, CB, O
Group: LYS116, etc. x,y,z atom coordinates
9 display groups within a particular distance (e.g. 10A) from a selected atom. Note the selection on the control and graphics panels.
10 center display on a selected atom
11-13 advanced tools

[4d] Use the pull-down menu options such as the following:
-- Select → Group property → non-polar (and visualize this)
-- Select → Secondary structure → helices

► Go to window → control panel.

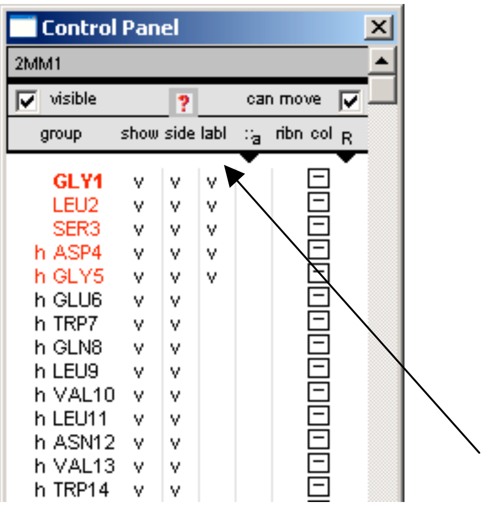


Control Panel

2MM1

☑ visible ? can move ☑

| group | show | side | labl | ::a | ribn col R |
|---|---|---|---|---|---|
| **GLY1** | v | v | | | ⊟ |
| LEU2 | v | v | | | ⊟ |
| SER3 | v | v | | | ⊟ |
| h ASP4 | v | v | | | ⊟ |
| h GLY5 | v | v | | | ⊟ |
| h GLU6 | v | v | | | ⊟ |
| h TRP7 | v | v | | | ⊟ |
| h GLN8 | v | v | | | ⊟ |
| h LEU9 | v | v | | | ⊟ |
| h VAL10 | v | v | | | ⊟ |
| h LEU11 | v | v | | | ⊟ |
| h ASN12 | v | v | | | ⊟ |
| h VAL13 | v | v | | | ⊟ |
| h TRP14 | v | v | | | ⊟ |
| h GLY15 | v | v | | | ⊟ |
| h LYS16 | v | v | | | ⊟ |



Control Panel

2MM1

☑ visible ? can move ☑

| group | show | side | labl | ::a | ribn col R |
|---|---|---|---|---|---|
| **GLY1** | v | v | v | | ⊟ |
| LEU2 | v | v | v | | ⊟ |
| SER3 | v | v | v | | ⊟ |
| h ASP4 | v | v | v | | ⊟ |
| h GLY5 | v | v | v | | ⊟ |
| h GLU6 | v | v | | | ⊟ |
| h TRP7 | v | v | | | ⊟ |
| h GLN8 | v | v | | | ⊟ |
| h LEU9 | v | v | | | ⊟ |
| h VAL10 | v | v | | | ⊟ |
| h LEU11 | v | v | | | ⊟ |
| h ASN12 | v | v | | | ⊟ |
| h VAL13 | v | v | | | ⊟ |
| h TRP14 | v | v | | | ⊟ |

►Shift/click to select the first five amino acid residues of myoglobin. They should appear red.

► Click “labl” (i.e. label)(see arrow, above right). Those five residues now have a “v”.

►Inspect the display panel; those five residues are labeled.