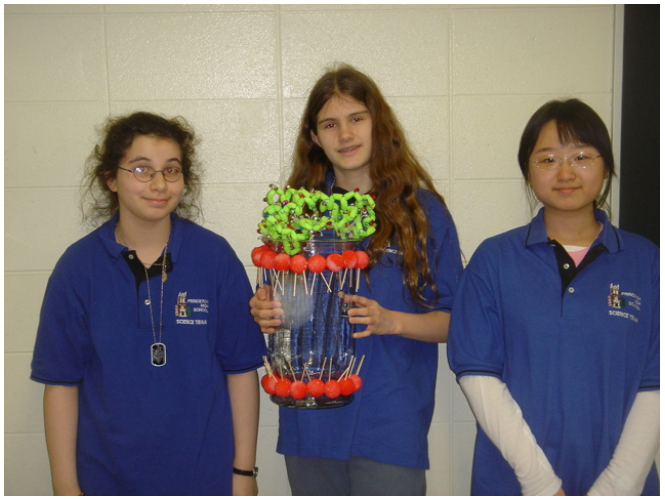
Protein Modeling (Coaches Clinic)



Shuchismita Dutta

October 27, 2007

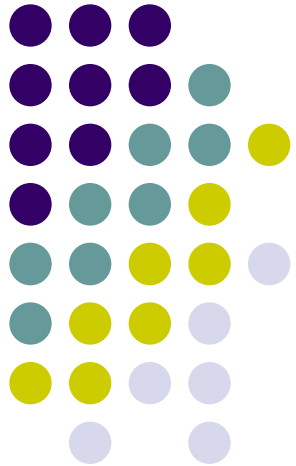
2007 State Champions:
Ola Hadaya, Sarah Goodman, and Yong Kim from Princeton High School

General Introduction

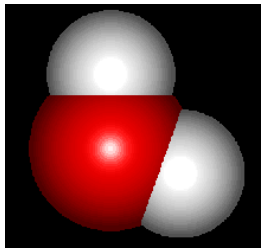
Water

Amino Acids

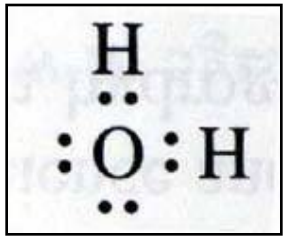
Protein folding

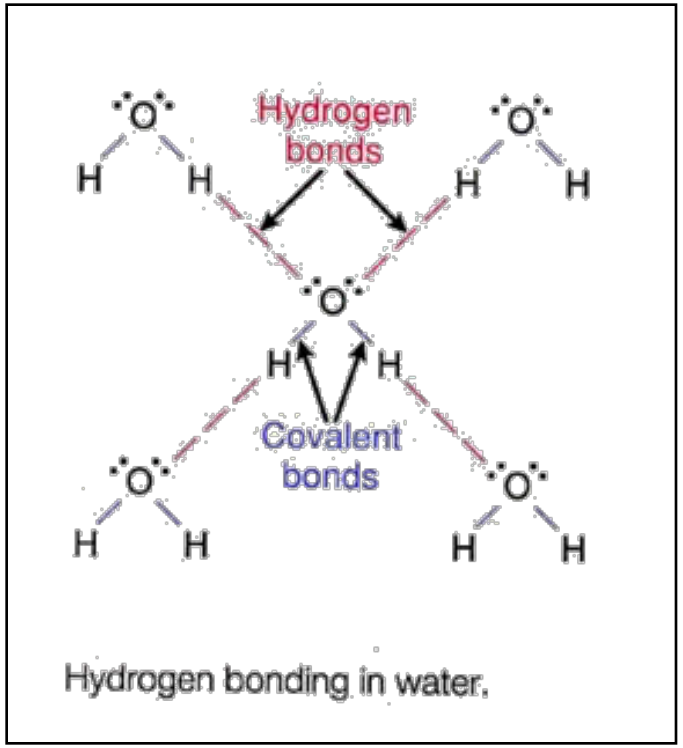


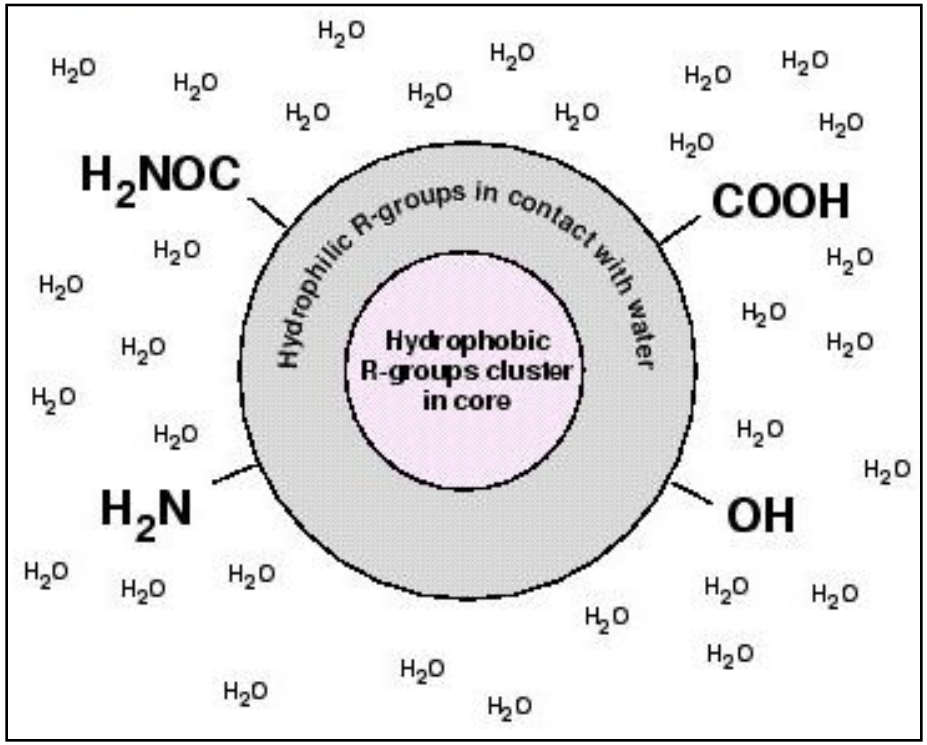
Water: love it or leave it



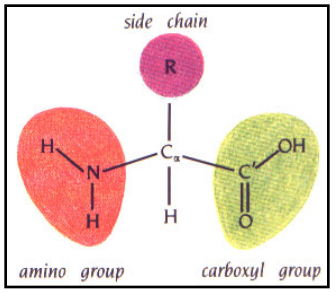
- Hydrogen bonds
- Hydrophobic and Hydrophilic structures







Amino Acids



Basic

Acidic

Periodic Chart of Amino Acids

www.bachem.com

Hydrophobic

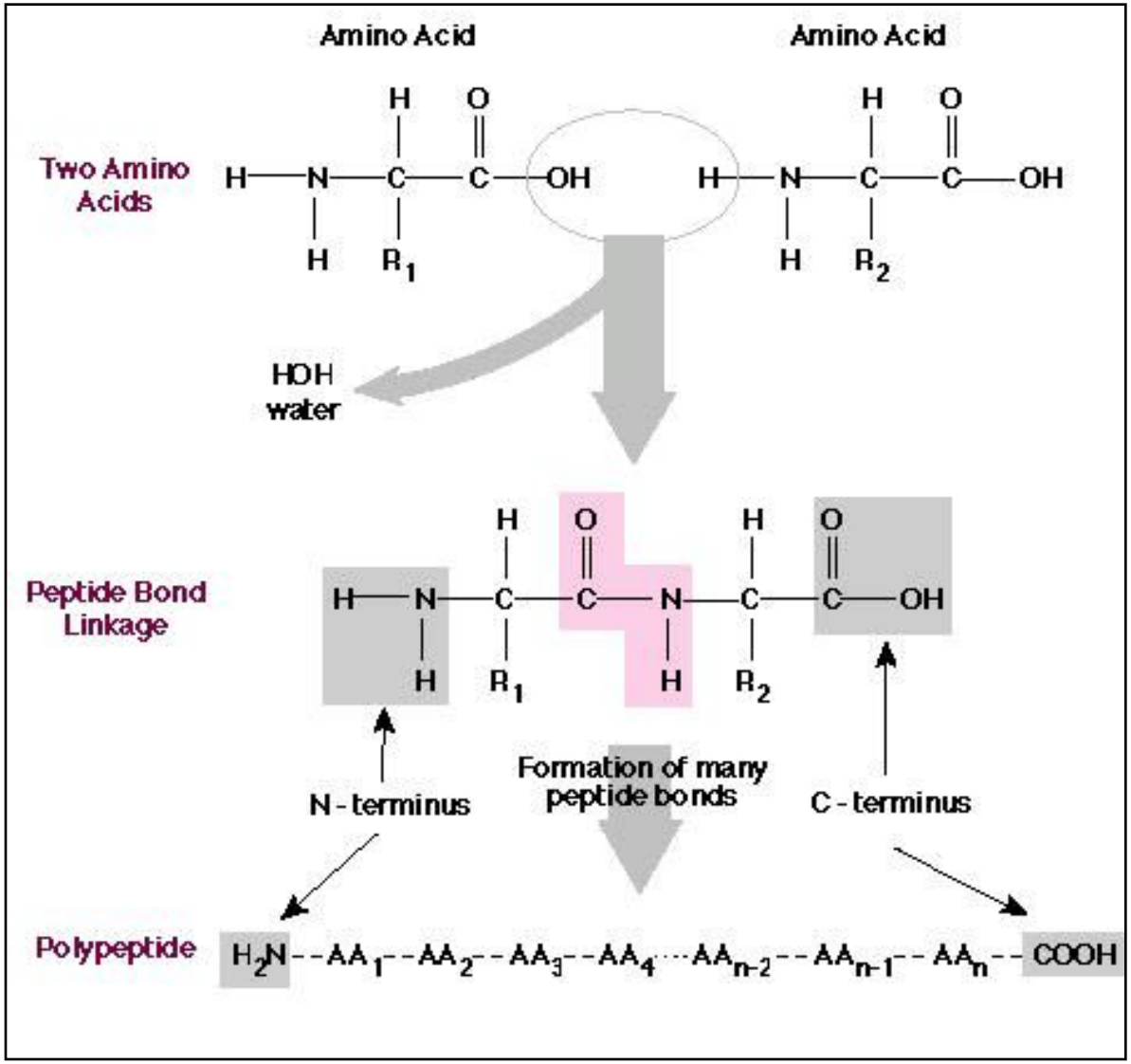
Hydrophilic

Basic
Nonpolar (hydrophobic)
Polar, uncharged
Acidic

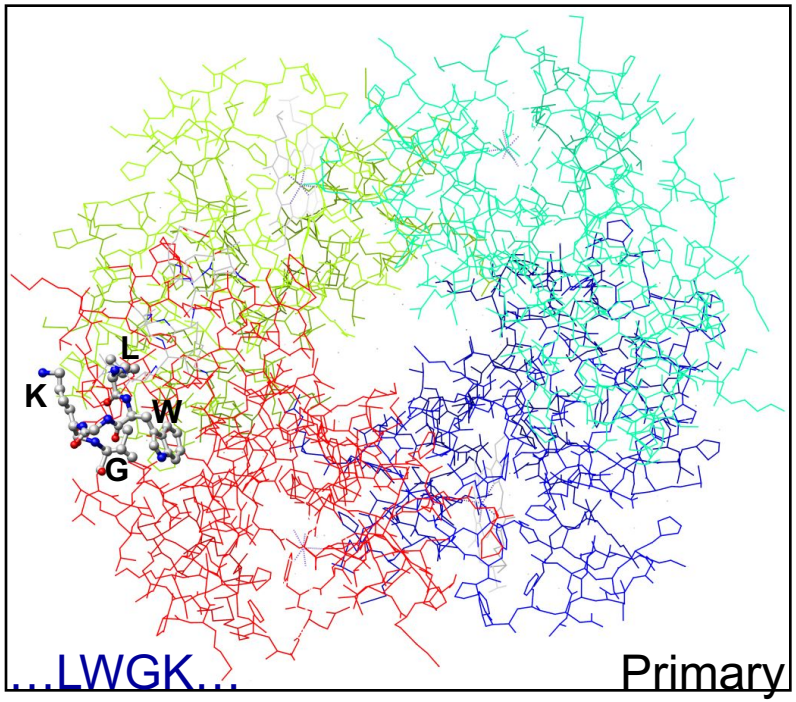
1-Letter Amino Acid Code
Molecular Weight
MW-H_2O
Molecular Formula
3-Letter Amino Acid Code
Chemical Structure
Chemical Name

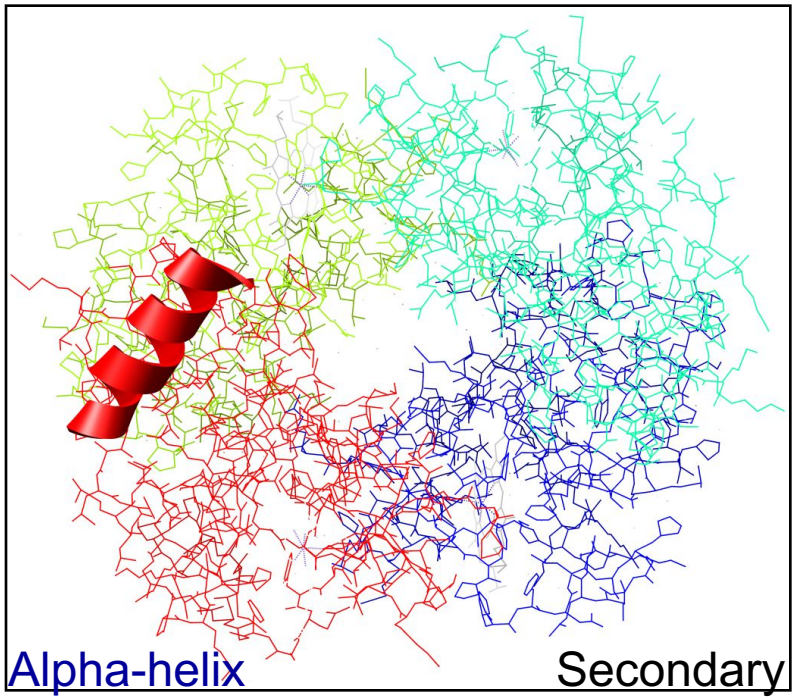
From www.bachem.com

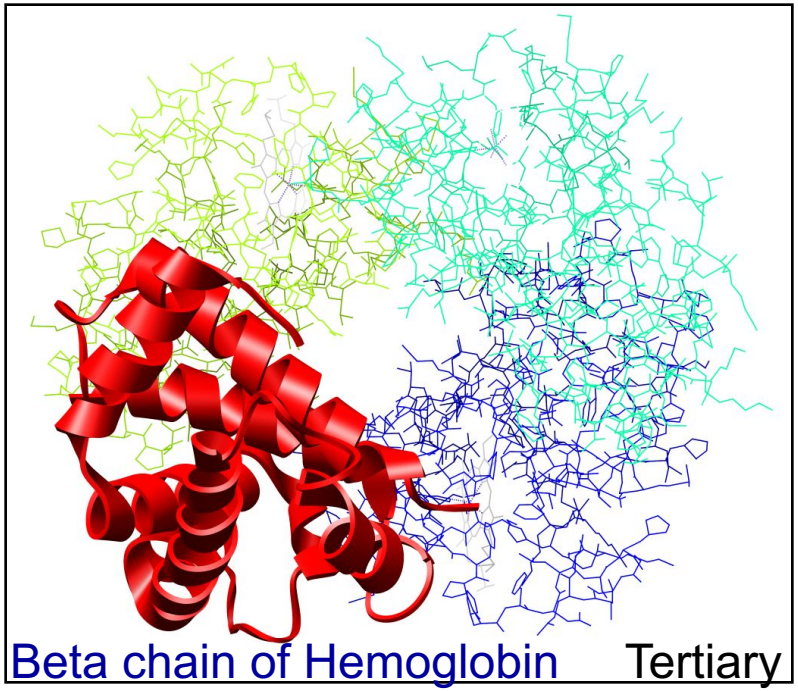
Proteins

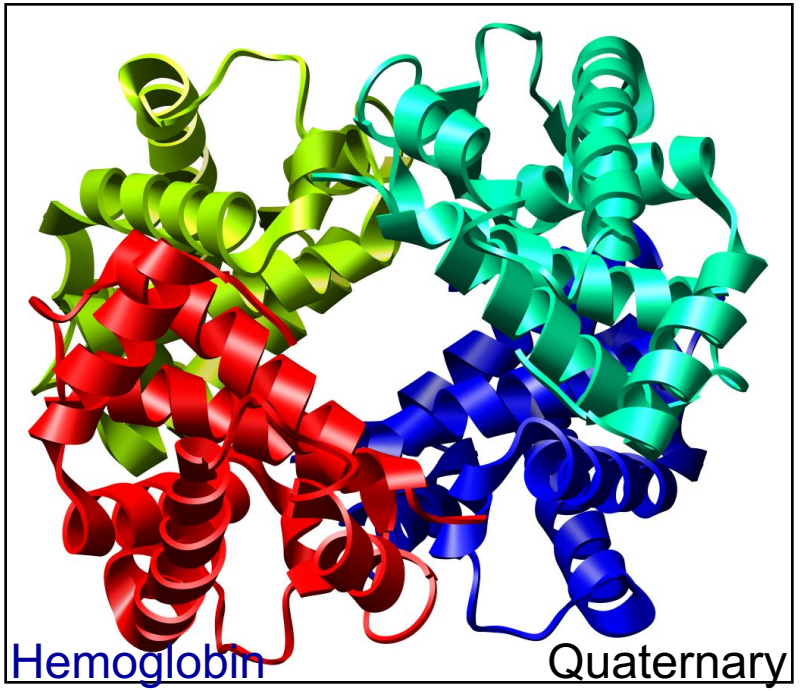

Amino Acid
Amino Acid
Two Amino Acids
HOH
water
Peptide Bond Linkage
N - terminus
Formation of many peptide bonds
C - terminus
Polypeptide
H2N--AA1--AA2--AA3--AA4···AAn-2--AAn-1--AAn--COOH

Protein Structure

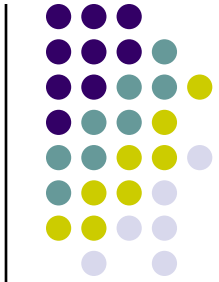






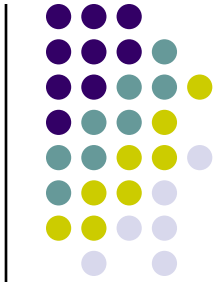


Some Rules to ‘Fold’ a Protein



- Covalent interactions
  - Protein Sequence
  - Di-sulfide bridges
- Non-covalent interactions
  - Hydrophobic interactions
  - Hydrogen bonds
  - Salt bridges (positive negative interactions)
  - Metal coordination

Protein Modeling: Toober and Thumb Tacks model

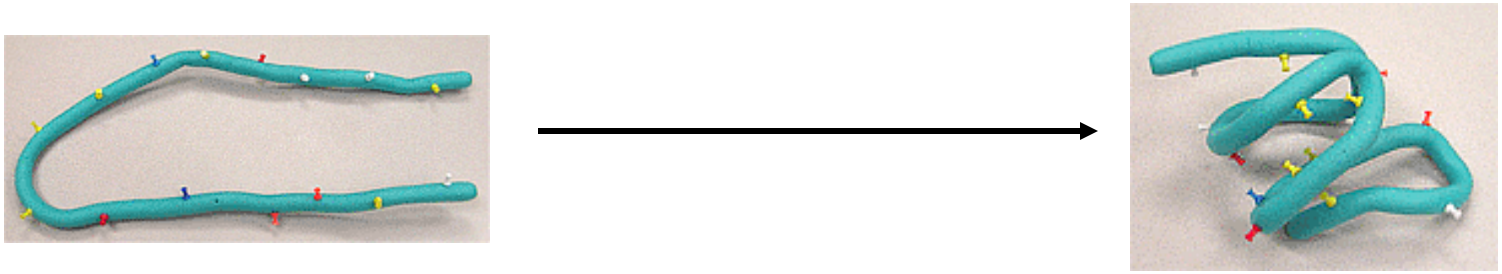


- 1Toober
- 10 thumb tacks
  - 1 Blue (Basic)
  - 1 Red (Acidic)
  - 4 Yellow (Hydrophobic)
  - 2 White (Hydrophilic)
  - 2 Green (Cysteine)

http://www.3dmoleculardesigns.com/15_Tacks.pdf

Interaction rules

- Hydrophobic (yellows should be away from water, and whites should be near water)
- Charge based (red and blue should pair up)
- Disulphide (the greens should pair up to form a bond)

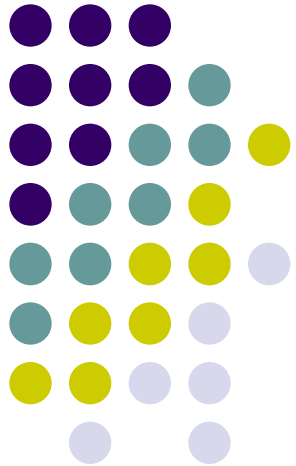


PDB & Protein Models

Protein Structures

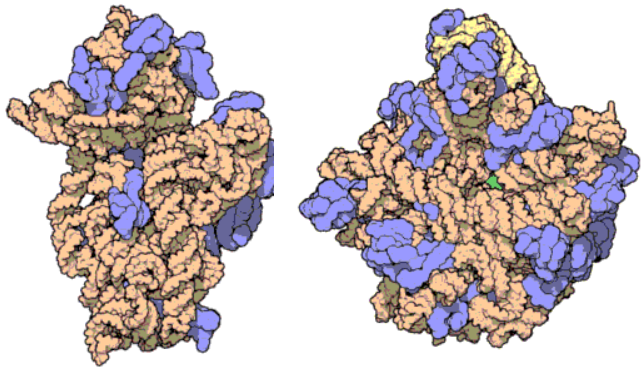
Protein Data Bank

Protein Modeling



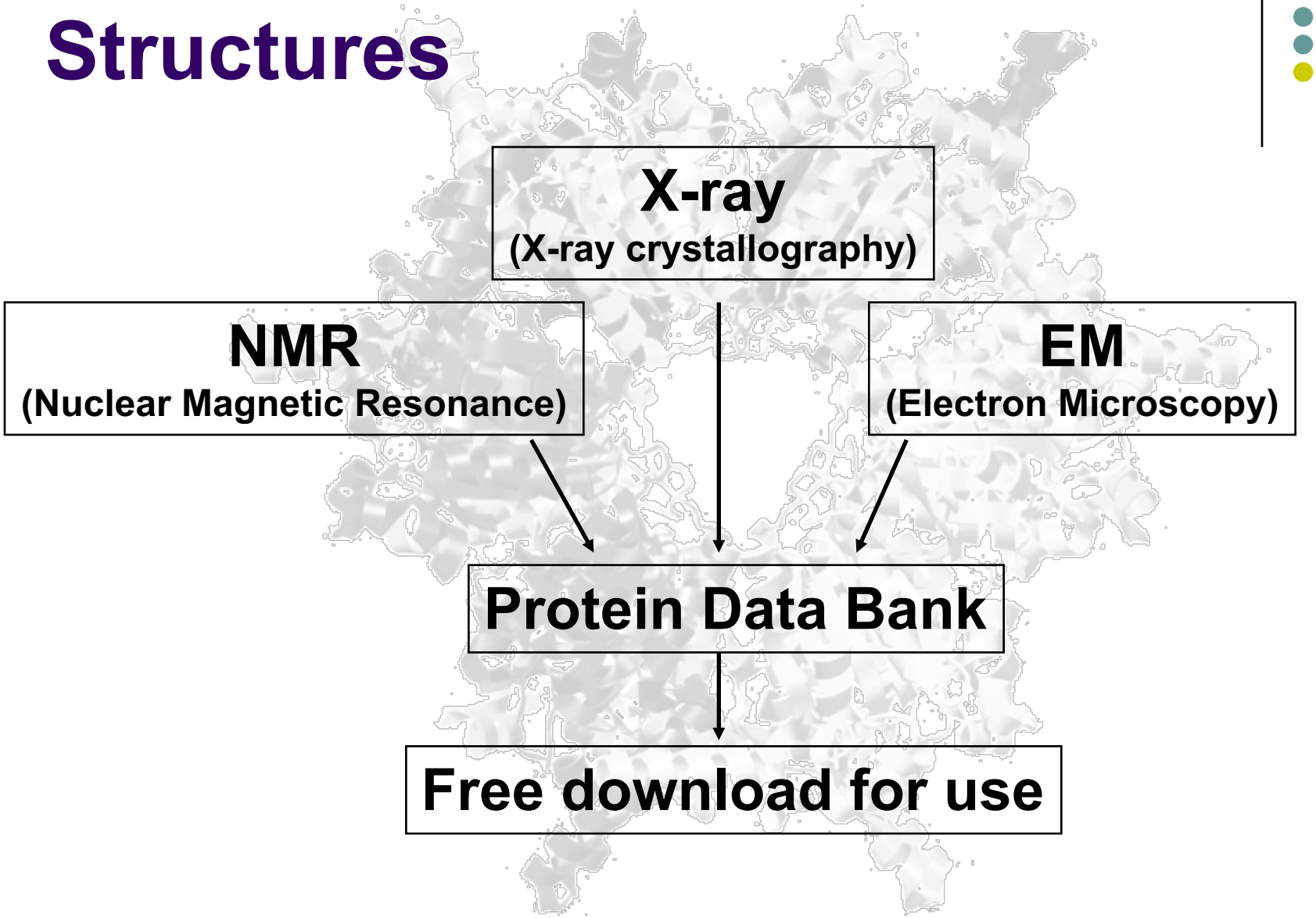
Why Structure?

shown: the ribosome



- Allows you to “visualize” the shape and details of the protein
- Offers clues about the role in the body
- May hold key to developing new medicines and diagnostic procedures for diseases like avian flu, HIV, West Nile Virus, parts of the protein associated with Alzheimer’s disease, Cancers, etc.

Understanding Protein Structures



The Protein Data Bank

RCSB PDB PROTEIN DATA BANK

A MEMBER OF THE PDB

An Information Portal to Biological Macromolecular Structures

As of Tuesday Oct 17, 2006 there are 39464 Structures | PDB Statistics

Contact Us | Help | Print Page

PDB ID or keyword | Author | SEARCH | Advanced Search

Home | Search

- Home
- Tutorial About This Site
- Getting Started
- Download Files
- Deposit and Validate
- Structural Genomics
- Dictionaries & File Formats
- Software Tools
- Educational Resources
- BioSync
- General Information
- Acknowledgements
- Frequently Asked Questions
- Known Problems
- Report Bugs/Comments

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.

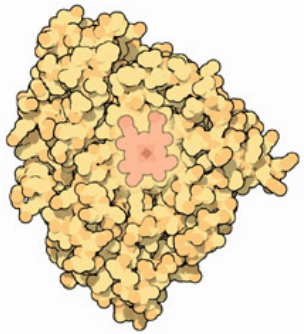
This site offers tools for browsing, searching, and reporting that utilize the data resulting from ongoing efforts to create a more consistent and comprehensive archive.

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: Cytochrome p450



If you have a headache and take a drug to block the pain, you'll notice that the effects of the drug wear off in a few hours. This happens because you have a powerful detoxification system that finds unusual chemicals, like drugs, and flushes them out of your body. This system fights all sorts of unpleasant chemicals that we eat and breathe, including drugs, poisonous compounds in plants, carcinogens formed during cooking, and environmental pollutants. The cytochrome p450 enzymes are our first line of defense in this chemical battle.

- More ...
- Previous Features

NEWS

- Complete News
- Newsletter
- Discussion Forum

17-October-2006

RCSB PDB Focus: Exploring Domains in Protein Structure

Domains can be thought of as the smallest structural units from which proteins are assembled that retain properties of the whole protein, such as a hydrophobic core.

The RCSB PDB offers various ways of exploring domains in protein structures.

- Full Story ...

10-October-2006

DOIs Available for Released Entries in the PDB Archive

Structures released by the wwPDB into the PDB Archive are now being assigned a Document Object Identifier DOI. The DOI System is used to identifying content objects (such as journal articles, books, and figures) in the digital environment.

- Full Story ...

The Structure Explorer Page

RCSB PDB PROTEIN DATA BANK

A MEMBER OF THE PDB

An Information Portal to Biological Macromolecular Structures

As of Tuesday Oct 17, 2006 there are 39464 Structures | PDB Statistics

Contact Us | Help | Print Page

PDB ID or keyword | Author | SEARCH | Advanced Search

Home | Search | Structure | Queries | Structure Summary | Biology & Chemistry | Materials & Methods | Sequence Details | Geometry

- 1ZAA
- NDB Atlas
- Download Files
- FASTA Sequence
- Display Files
- Display Molecule
- Structural Reports
- External Links
- Structure Analysis
- Help

1ZAA

Title	ZINC FINGER-DNA RECOGNITION: CRYSTAL STRUCTURE OF A ZIF268-DNA COMPLEX AT 2.1 ANGSTROMS
Authors	Pavletich, N.P., Pabo, C.O.
Primary Citation	**Pavletich, N.P., Pabo, C.O.** Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1 A. *Science v252 pp.809-817 , 1991* [Abstract]
History	Deposition 1992-09-17 Release 1993-10-31
Experimental Method	Type X-RAY DIFFRACTION Data N/A

Parameters

Resolution[Å]	R-Value	R-Free	Space Group
2.10	0.182 (obs.)	n/a	$C\,2\,2\,2_1$

Unit Cell

Length [Å]	a	45.40	b	56.20	c	130.80
Angles [°]	alpha	90.00	beta	90.00	gamma	90.00

Molecular Description Asymmetric Unit

Polymer: 1 Molecule: DNA (5'-D(*AP*GP*CP*GP*TP*GP*GP*GP*CP*GP*T)-3') Chains: A
Polymer: 2 Molecule: DNA (5'-D(*TP*AP*CP*GP*CP*CP*CP*AP*CP*GP*C)-3') Chains: B
Polymer: 3 Molecule: PROTEIN (ZIF268) Chains: C

Classification Transcription/dna

Source Polymer: 1 Scientific Name: Synthetic construct Polymer: 2 Scientific Name: Synthetic construct Polymer: 3 Scientific Name: **Mus musculus** Common Name: **Mouse** Mouse

Chemical Component

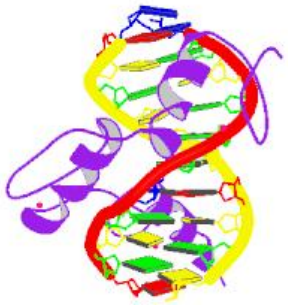
Identifier	Name	Formula	Drug Similarity	Ligand Structure	Ligand Interaction
ZN	ZINC ION	Zn^{2+}	[View]	[View]	[View]

SCOP Classification (version 1.69)

Domain Info	Class	Fold	Superfamily	Family	Domain	Species
d1zaac1	Small proteins	C2H2 and C2HC zinc fingers	C2H2 and C2HC zinc fingers	Classic zinc finger, C2H2	ZIF268	Mouse (Mus musculus)
d1zaac2	Small proteins	C2H2 and C2HC zinc fingers	C2H2 and C2HC zinc fingers	Classic zinc finger, C2H2	ZIF268	Mouse (Mus musculus)
d1zaac3	Small proteins	C2H2 and C2HC zinc fingers	C2H2 and C2HC zinc fingers	Classic zinc finger, C2H2	ZIF268	Mouse (Mus musculus)

Images and Visualization

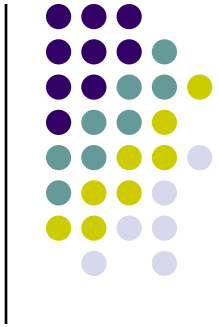
Biological Molecule / Asymmetric Unit



Display Options

KiNG
Jmol
WebMol
Protein Workshop
QuickPDB
All Images

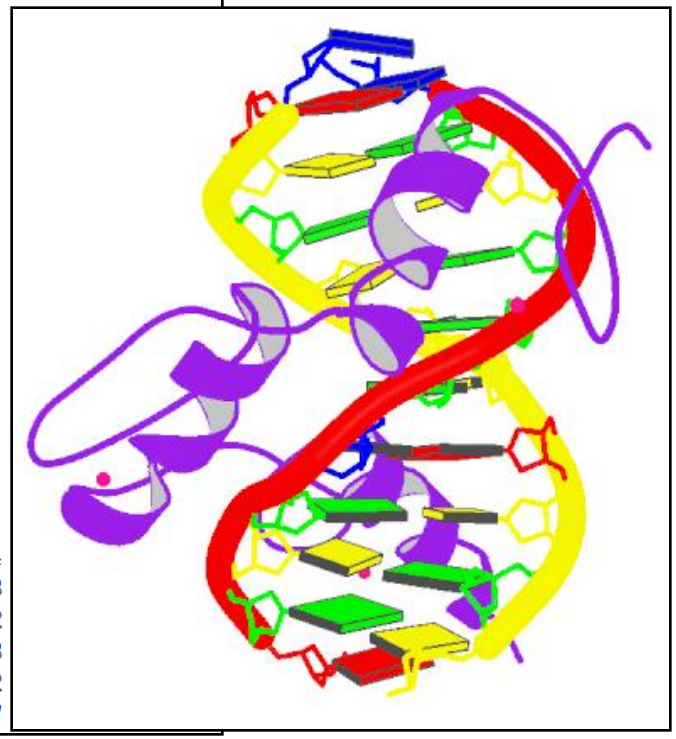
The PDB file



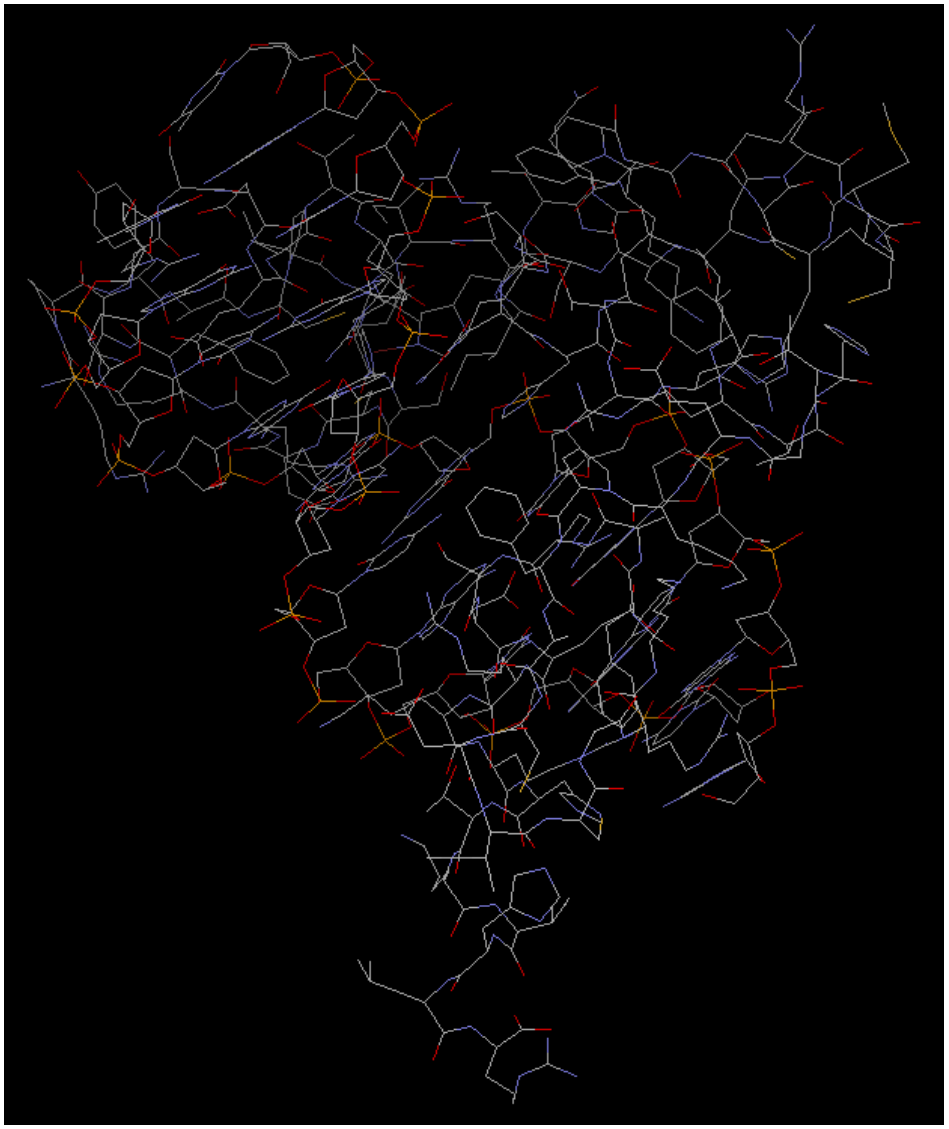
```
HEADER    TRANSCRIPTION/DNA                         17-SEP-92   1ZAA
TITLE     ZINC FINGER-DNA RECOGNITION: CRYSTAL STRUCTURE OF A ZIF268-
TITLE    2 DNA COMPLEX AT 2.1 ANGSTROMS
COMPND    MOL_ID: 1;
COMPND   2 MOLECULE: DNA (5'-D(*AP*GP*CP*GP*TP*GP*GP*GP*CP*GP*T)-3');
COMPND   3 CHAIN: A;
COMPND   4 ENGINEERED: YES;
COMPND   5 MOL_ID: 2;
COMPND   6 MOLECULE: DNA (5'-D(*TP*AP*CP*GP*CP*CP*CP*AP*CP*GP*C)-3');
COMPND   7 CHAIN: B;
COMPND   8 ENGINEERED: YES;
COMPND   9 MOL_ID: 3;
COMPND  10 MOLECULE: PROTEIN (ZIF268);
COMPND  11 CHAIN: C
SOURCE    MOL_ID: 1;
SOURCE   2 SYNTHETIC: YES;
SOURCE   3 MOL_ID: 2;
SOURCE   4 SYNTHETIC: YES;
SOURCE   5 MOL_ID: 3;
SOURCE   6 ORGANISM_SCIENTIFIC: MUS MUSCULUS;
SOURCE   7 ORGANISM_COMMON: MOUSE;
SOURCE   8 GENUS: MUS;
SOURCE   9 SPECIES: MUSCULUS
KEYWDS    PROTEIN-DNA COMPLEX, DOUBLE HELIX
EXPDTA    X-RAY DIFFRACTION
AUTHOR    N.P.PAVLETICH,C.O.PABO
REVDAT   3   01-APR-03 1ZAA    1       JRNL
REVDAT   2   08-MAR-95 1ZAA    1       SHEET
REVDAT   1   31-OCT-93 1ZAA    0
JRNL        AUTH   N.P.PAVLETICH,C.O.PABO
JRNL        TITL   ZINC FINGER-DNA RECOGNITION: CRYSTAL STRUCTURE OF
JRNL        TITL 2 A ZIF268-DNA COMPLEX AT 2.1 A.
JRNL        REF    SCIENCE                       V. 252   809 1991
JRNL        REFN   ASTM SCIEAS  US ISSN 0036-8075
REMARK   1
REMARK   2
REMARK   2 RESOLUTION. 2.10 ANGSTROMS.
REMARK   3
REMARK   3 REFINEMENT.
REMARK   3   PROGRAM     : X-PLOR, TNT
REMARK   3   AUTHORS     : TRONRUD,TEN EYCK,MATTHEWS
REMARK   3
```

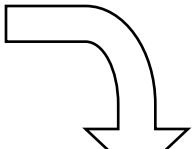
The PDB file – con't.

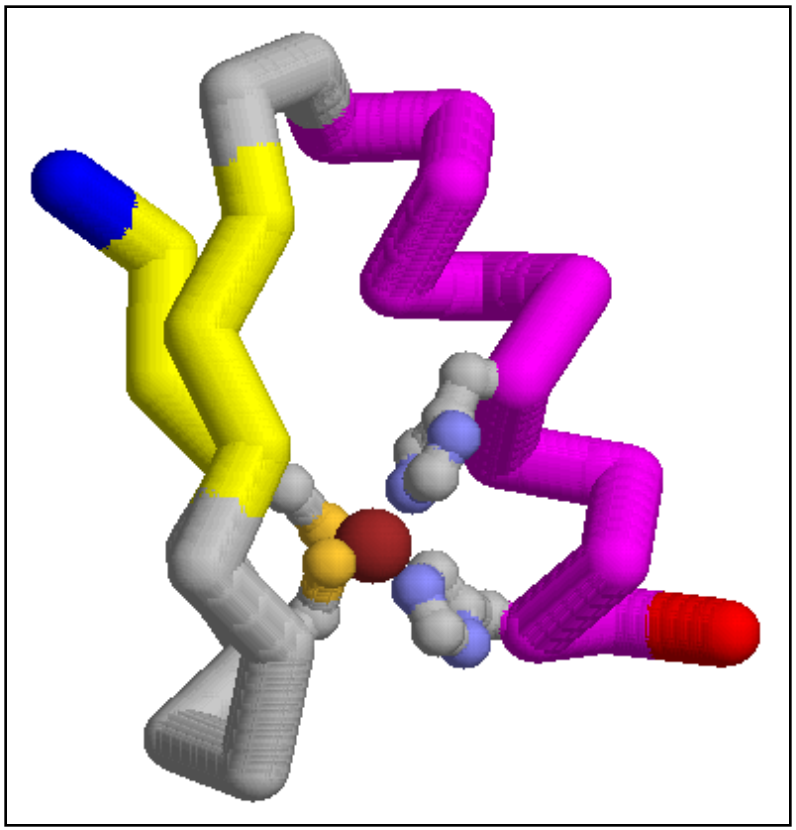
```
DBREF  1ZAA C    1    87  UNP    P08046   EGR1_MOUSE     332    418
SEQRES   1 A   11   DA  DG  DC  DG  DT  DG  DG  DG  DC  DG  DT
SEQRES   1 B   11   DT  DA  DC  DG  DC  DC  DC  DA  DC  DG  DC
SEQRES   1 C   87  HIS GLU ARG PRO TYR ALA CYS PRO VAL GLU SER CYS ASP
SEQRES   2 C   87  ARG ARG PHE SER ARG SER ASP GLU LEU THR ARG HIS ILE
SEQRES   3 C   87  ARG ILE HIS THR GLY GLN LYS PRO PHE GLN CYS ARG ILE
SEQRES   4 C   87  CYS MET ARG ASN PHE SER ARG SER ASP HIS LEU THR THR
SEQRES   5 C   87  HIS ILE ARG THR HIS THR GLY GLU LYS PRO PHE ALA CYS
SEQRES   6 C   87  ASP ILE CYS GLY ARG LYS PHE ALA ARG SER ASP GLU ARG
SEQRES   7 C   87  LYS ARG HIS THR LYS ILE HIS LEU ARG
HET     ZN  D 201       1
HET     ZN    202       1
HET     ZN  D 203       1
HETNAM      ZN ZINC ION
FORMUL   4   ZN    3(ZN 2+)
FORMUL   7  HOH   *129(H2 O)
HELIX    1  A1 SER C   19  THR C   30  1
HELIX    2  A2 SER C   47  THR C   58  1
HELIX    3  A3 SER C   75  HIS C   85  1
SHEET    1 AB1 2 TYR C    5  CYS C    7  0
SHEET    2 AB1 2 ARG C   14  PHE C   16  1
SHEET    1 AB2 2 PHE C   35  CYS C   37  0
SHEET    2 AB2 2 ARG C   42  PHE C   44 -1
SHEET    1 AB3 2 PHE C   63  CYS C   65  0
SHEET    2 AB3 2 ARG C   70  PHE C   72 -1
CRYST1   45.400   56.200  130.800  90.00  90.00  90.00 C 2 2 21
ORIGX1      1.000000  0.000000  0.000000        0.00000
ORIGX2      0.000000  1.000000  0.000000        0.00000
ORIGX3      0.000000  0.000000  1.000000        0.00000
SCALE1      0.022026  0.000000  0.000000        0.00000
SCALE2      0.000000  0.017794  0.000000        0.00000
SCALE3      0.000000  0.000000  0.007645        0.00000
ATOM      1  O5'  DA A   1      12.718  25.492  40.619  1.00 20.74
ATOM      2  C5'  DA A   1      14.060  25.855  40.385  1.00 17.93
ATOM      3  C4'  DA A   1      15.009  24.791  40.891  1.00 21.22
ATOM      4  O4'  DA A   1      15.274  24.953  42.304  1.00 21.73
ATOM      5  C3'  DA A   1      14.597  23.333  40.746  1.00 23.22
ATOM      6  O3'  DA A   1      15.741  22.606  40.427  1.00 27.17
```



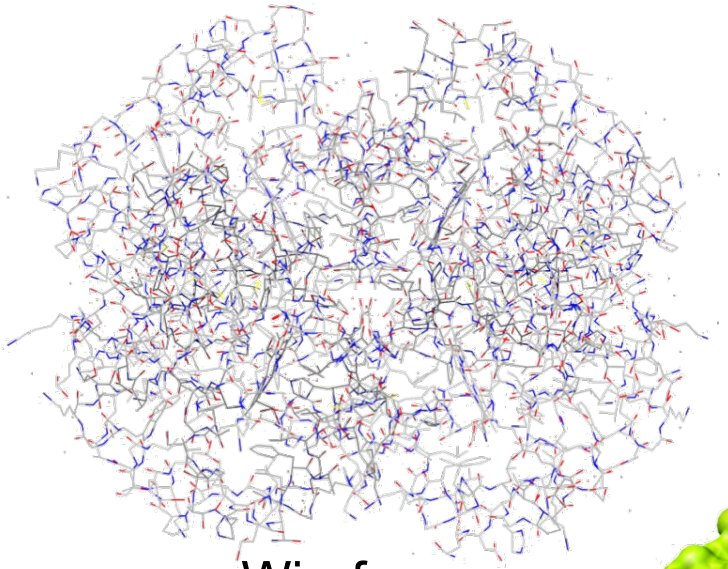
Visualization (RasMol or Jmol)



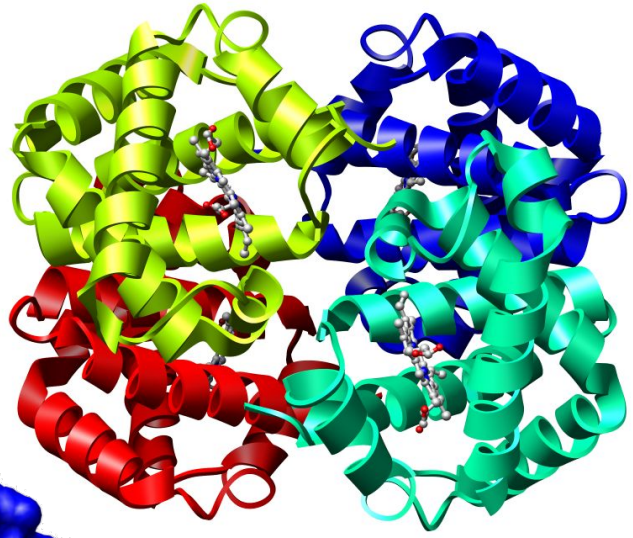




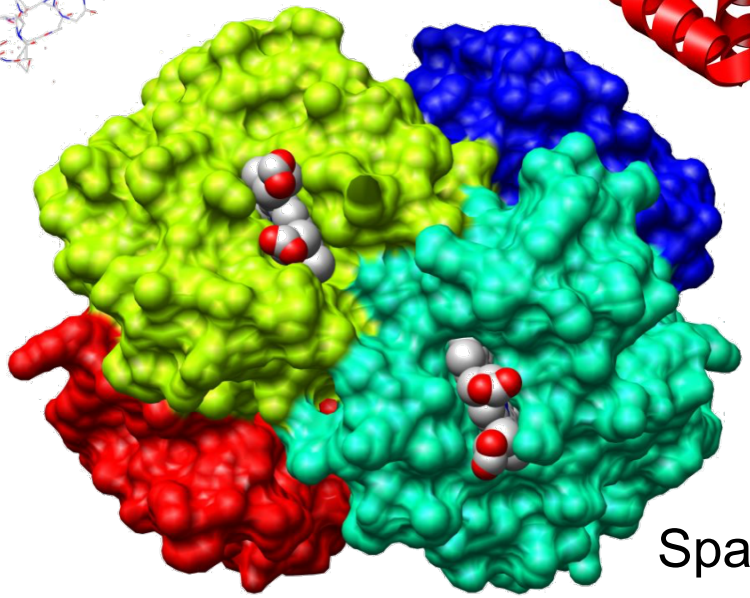
Structure Representation



Wireframe

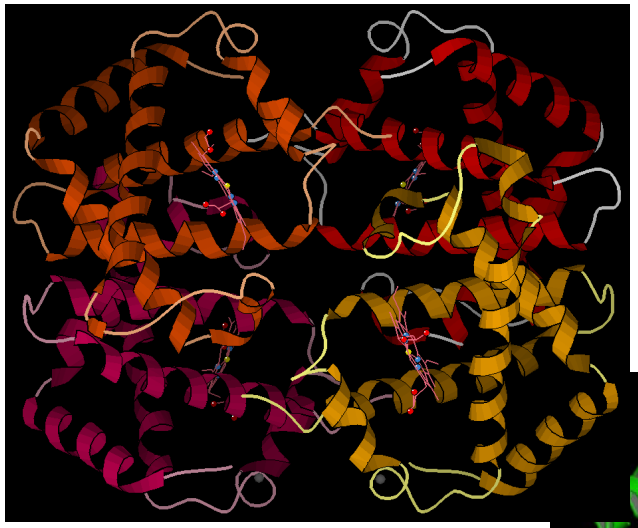


Ribbons

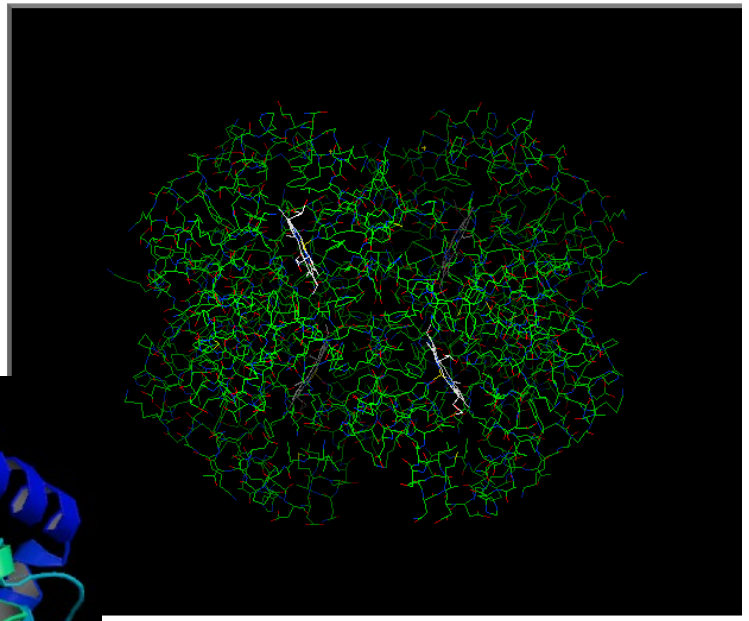


Spacefill

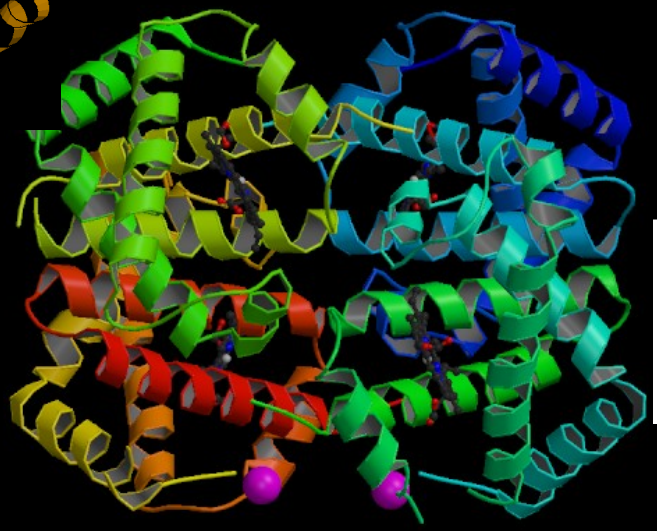
Graphics



KiNG

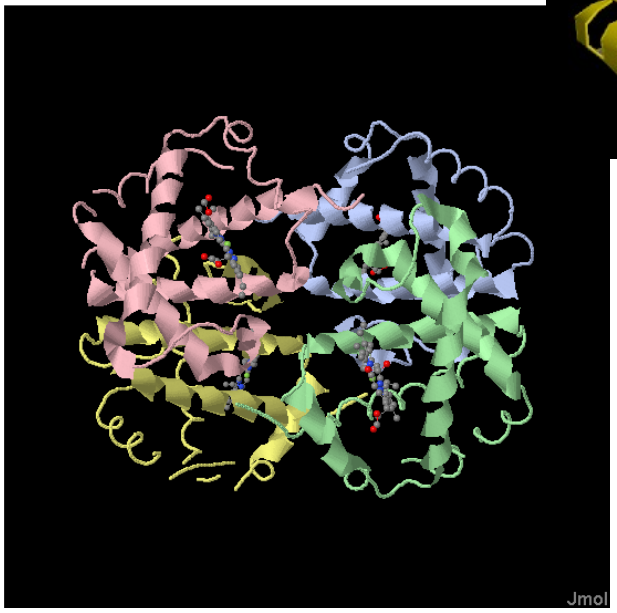


WebMol

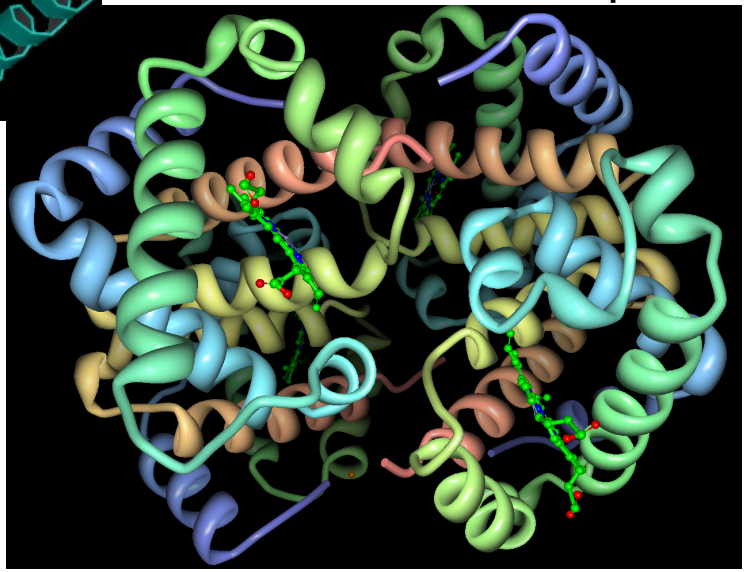


Default image

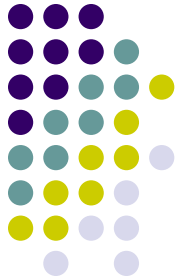
Jmol



Protein Explorer



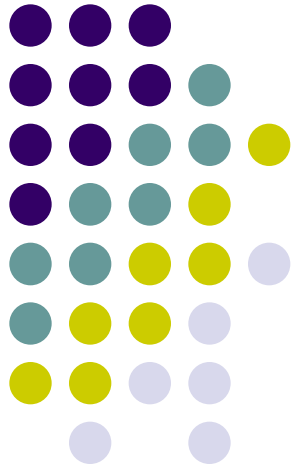
Zinc Finger Toober Model



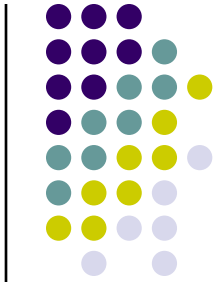
- Download the 1ZAA pdb file (www.pdb.org)
- Create image in Jmol, identify key features
- Fold a Mini-Toober model
  - Material modifications
    - Blue thumb tack (N-terminus)
    - Red thumb tack (C-terminus)
    - Colored Pipe-cleaners to represent Cys, His, Arg18, Phe16 and Leu22

NJSO 2008

Basic information



Information



- Regional and State level contests
  - Protein Calmodulin (PDB ID 1CLL)
- Rules
  - Pre-build: Bring in toober model and short abstract
  - On-site build: Build a designated part of 1CLL using a mini-Toober and Jmol
  - On-site exam: Answer questions about structure, function, importance and history of modeled protein. Materials will be provided.

Molecule of the Month

RCSB Protein Data Bank

RCSB PDB PROTEIN DATA BANK

A MEMBER OF THE PDB

An Information Portal to Biological Macromolecular Structures

As of Tuesday Oct 23, 2007 there are 46818 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
  - RCSB PDB Newsletters
  - Educational Resources
  - Molecule of the Month
  - Education Corner
  - Biological Unit Tutorial
  - Birch Aquarium
- Site Tutorials
- BioSync
- General Information
  - About the PDB
  - Citing the PDB
  - News & Publications
  - Related External Links
- Acknowledgements
- Frequently Asked Questions
- Report Bugs/Comments

Quick Tips:

Try the *Web Services API* for software developers using C/C++, Java, Python and Perl. Click here.

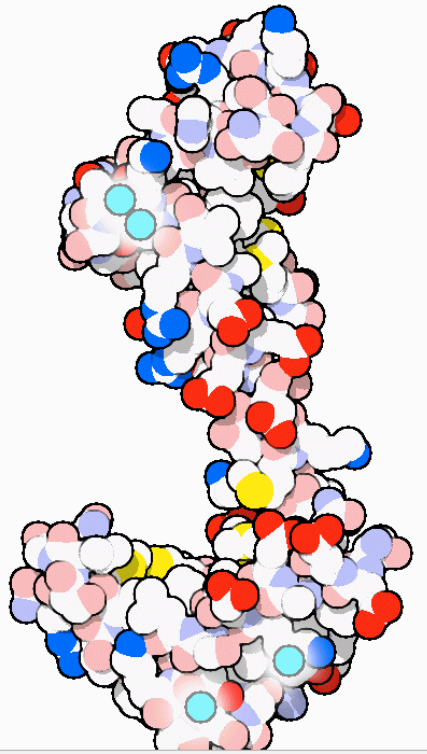
Are you missing data updates? The PDB archive has moved to ftp://ftp.wwpdb.org.
For more information click here.

Calmodulin

August 2003 Molecule of the Month
by Shuchismita Dutta and David S. Goodsell
Previous Features

Structures and Signals

Calcium is the most plentiful mineral element found in your body, with phosphorous coming in second. This probably doesn't come as a surprise, since your bones are strengthened and supported by about two kilograms of calcium and phosphorous. Your body also uses a small amount of calcium, in the form of calcium ions, to perform more active duties. Calcium ions play essential roles in cell signaling, helping to control processes such as muscle contraction, nerve signaling, fertilization and cell division. Through the action of calcium pumps and several kinds of calcium binding proteins, cells keep their internal calcium levels 1000-10,000 times lower than the calcium levels in the blood. Thus when calcium is released into cells, it can interact with calcium sensing proteins and trigger different biological effects, causing a muscle to contract, releasing insulin from the pancreas, or blocking the entry of additional sperm cells once an egg has been fertilized.



Sensing Calcium

As its name suggests, calmodulin is a CALcium MODULated proteIN. It is abundant in the cytoplasm of all higher cells and has been highly conserved through evolution. Calmodulin acts as an intermediary protein that senses calcium levels and relays signals to various calcium-sensitive enzymes, ion channels and other proteins. Calmodulin is a small dumbbell-shaped protein composed of two globular domains connected together by a flexible linker. Each end binds to two calcium ions. PDB entry **3cln**, shown here, has all four sites filled with calcium ions and the linker has formed a long alpha helix separating the two calcium-binding domains.

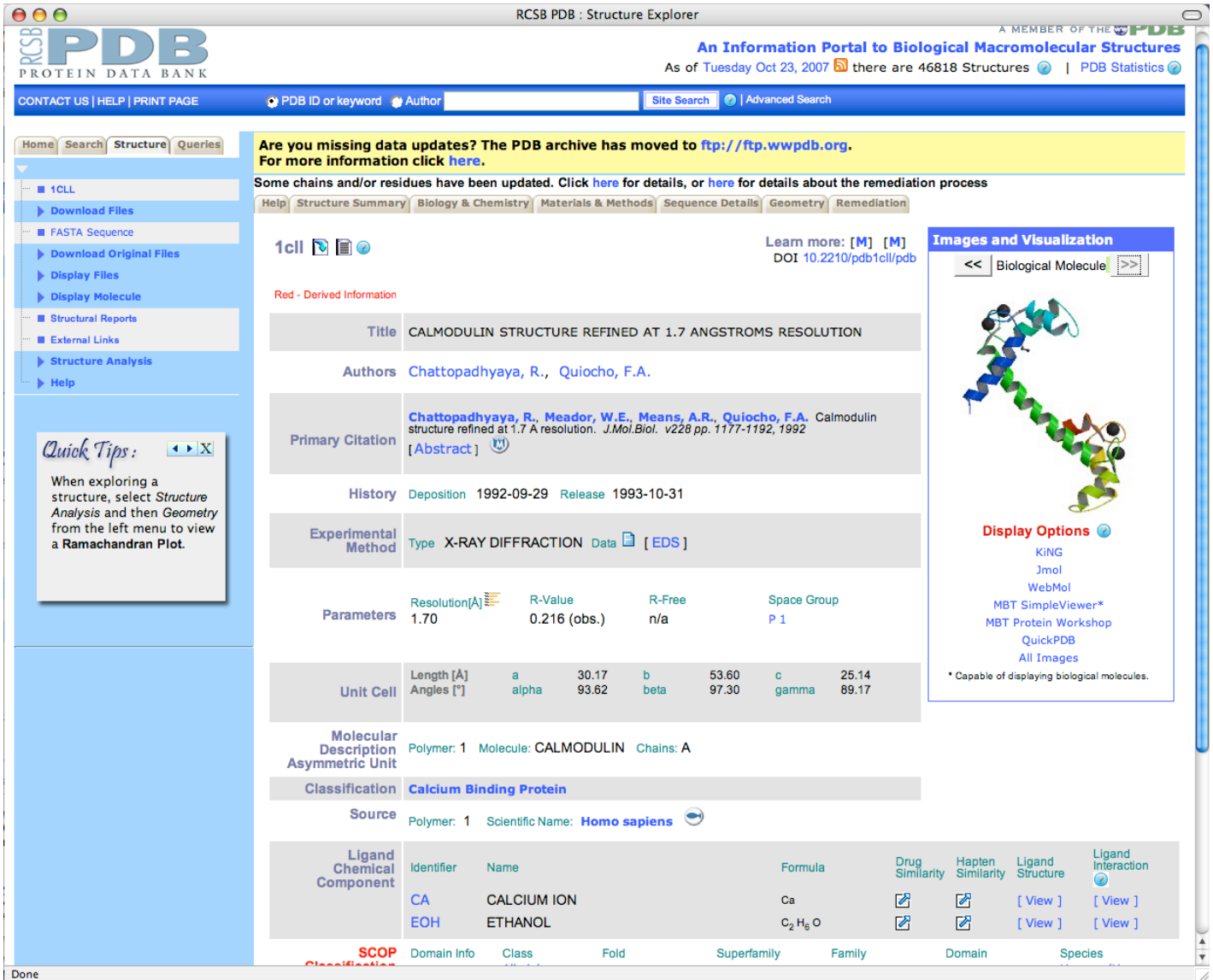
Calmodulin Look-alikes

Many different proteins are sensitive to calcium levels inside (and outside) cells. In the late 1960's, before the discovery of calmodulin, troponin C (see, for instance, PDB entry **1tcf**) was the first protein shown to be sensitive to calcium. Troponin C senses rising calcium levels and triggers muscle contraction. The structures of troponin C and calmodulin are remarkably similar, the major difference being the length of the linker connecting the two calcium-binding globular domains. The calcium-binding region of the protein, shown in detail in a later section, is almost identical. This motif has since been found in dozens of other calcium-sensitive proteins.

Next: A Functionally Versatile Machine

Done

Structure Explorer Page

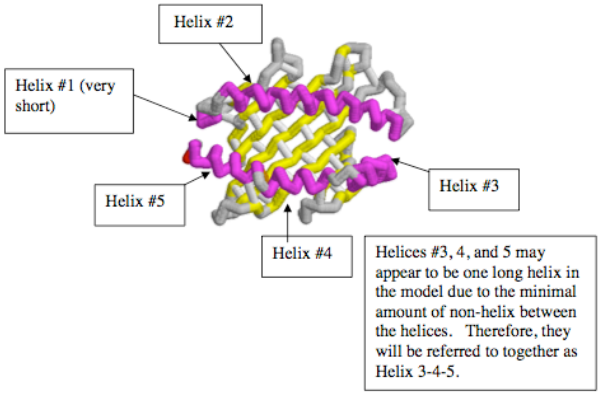


Jmol

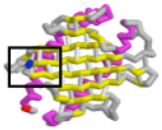

RCSB Protein Data Bank
RCSB PDB PROTEIN DATA BANK
A MEMBER OF THE PDB
An Information Portal to Biological Macromolecular Structures
As of Tuesday Oct 23, 2007 there are 46818 Structures | PDB Statistics
CONTACT US | HELP | PRINT PAGE
PDB ID or keyword
Author
Site Search
Advanced Search
Home
Search
Structure
Queries
1CLL
Download Files
FASTA Sequence
Download Original Files
Display Files
Display Molecule
Image Gallery
KiNG Viewer
Jmol Viewer
WebMol Viewer
Protein Workshop
Rasmol Viewer (Plugin required)
Swiss-PDB Viewer (Plugin required)
Molecular Viewers Help
KiNG Help
Jmol Help
WebMol Help
Protein Workshop Help
QuickPDB
Asymmetric Unit
Assumed Biological Molecule 1
Structural Reports
External Links
Structure Analysis
Help
Quick Tips :
When exploring a structure, select Structure
Are you missing data updates? The PDB archive has moved to ftp://ftp.wwpdb.org. For more information click here.
Jmol
1cll
Jmol
For help select one of the options below:
Help interacting with Jmol
Simple Interaction Guide (requires flash)
Advanced Jmol Help
Jmol script completed

How we judge

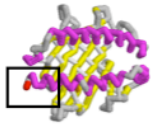
Alpha Helices
This is how the alpha helices will be referred to in the rubric and the guide to the rubric.



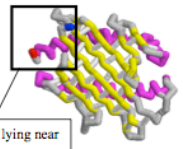
1. Blue Cap on N-terminal Amino Acid (Gly1) (1 pt)
   - To receive one point, the blue cap needs to be located at the N-terminus of the protein, which is the middle strand of the beta sheet. Please see the figure to the right for the correct positioning of the end caps.



2. Red cap on C-terminal Amino Acid (Lys176) (1 pt)
   - To receive one point, the red cap needs to be located at the C-terminus of the protein, which is the end of the long helix. Please see the figure to the right for correct positioning of the end caps.



3. N and C termini are on the same side of the model, next to each other (1 pt)
   - To receive one point, the blue cap needs to be located near the red cap. Please see the figure to the right for positioning of the caps.



2

Rubrics available from education.pdb.org/olympiad

Help

- Details and links at http://education.rcsb.org/olympiad/
- If you have questions or to borrow the “Introduction to Protein Structure” collection suitcase please write to buildmodel@rcsb.rutgers.edu