

# Protein Modeling Division C 2012 San Diego Science Olympiad

## Nuts and Bolts of Protein Modeling

Dec. 10, 2011

San Diego Supercomputer Center/UCSD

2:00 - 4:00 pm

# Protein Modeling C Event

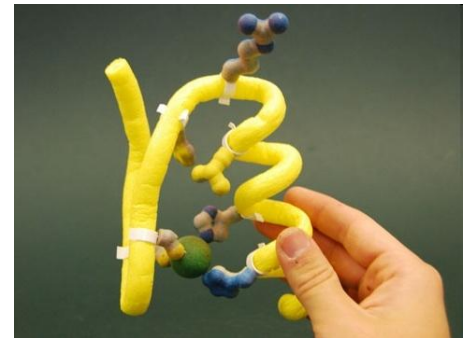
## 1. Pre-build models

- Build a large protein model
- Will **not** be impounded or scored at SD event!



## 2. On-site models

- Build a region of a protein
- 50% of score



## 3. Written test

- Multiple choice questions
- 1 written tie-breaker question
- 50% of score

### Multiple Choice Questions:

**A** 1. Which of the following amino acids is involved in disulfide bonds?

- A. Cysteine
- B. Histidine
- C. Tyrosine
- D. Methionine

**B** 2. Which of the following bonds is the strongest?

- A. Hydrogen bond
- B. Covalent bond
- C. Ionic bond
- D. Electrostatic bond

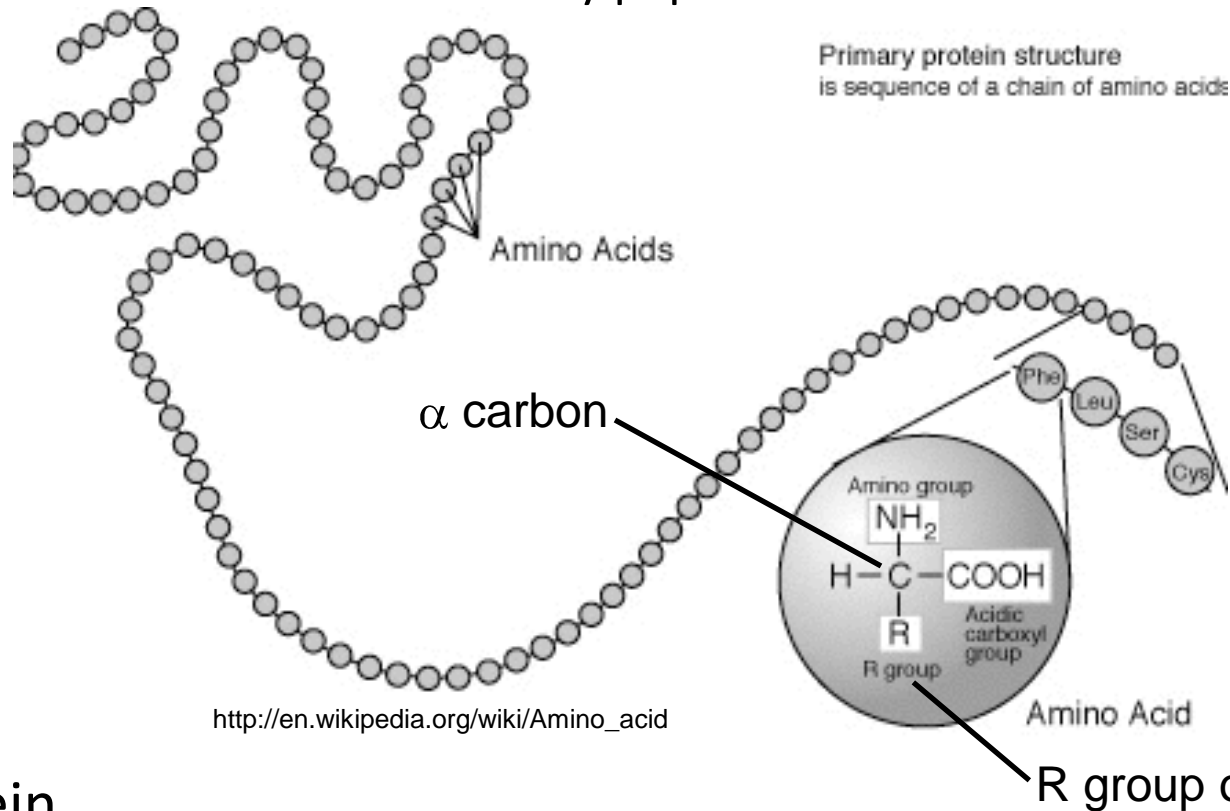
**D** 3. On your onsite model, you positioned amino acid Asp151. Which of the following statements is an accurate description of the significance of this amino acid?

- A. Asp151 is involved in binding Tamiflu, an antibiotic used to treat influenza.
- B. Asp151 plays a role in creating the tertiary structure necessary for the function of neuraminidase.
- C. Asp151 is one of the important amino acids necessary to dock the influenza virus on the host cell.
- D. Asp151 is an amino acid that is essential to the catalytic activity of the enzyme.

# What are Proteins?

- Polypeptides

- Polymers of amino acids (building blocks or monomers)
- Amino acids are connected by peptide bonds

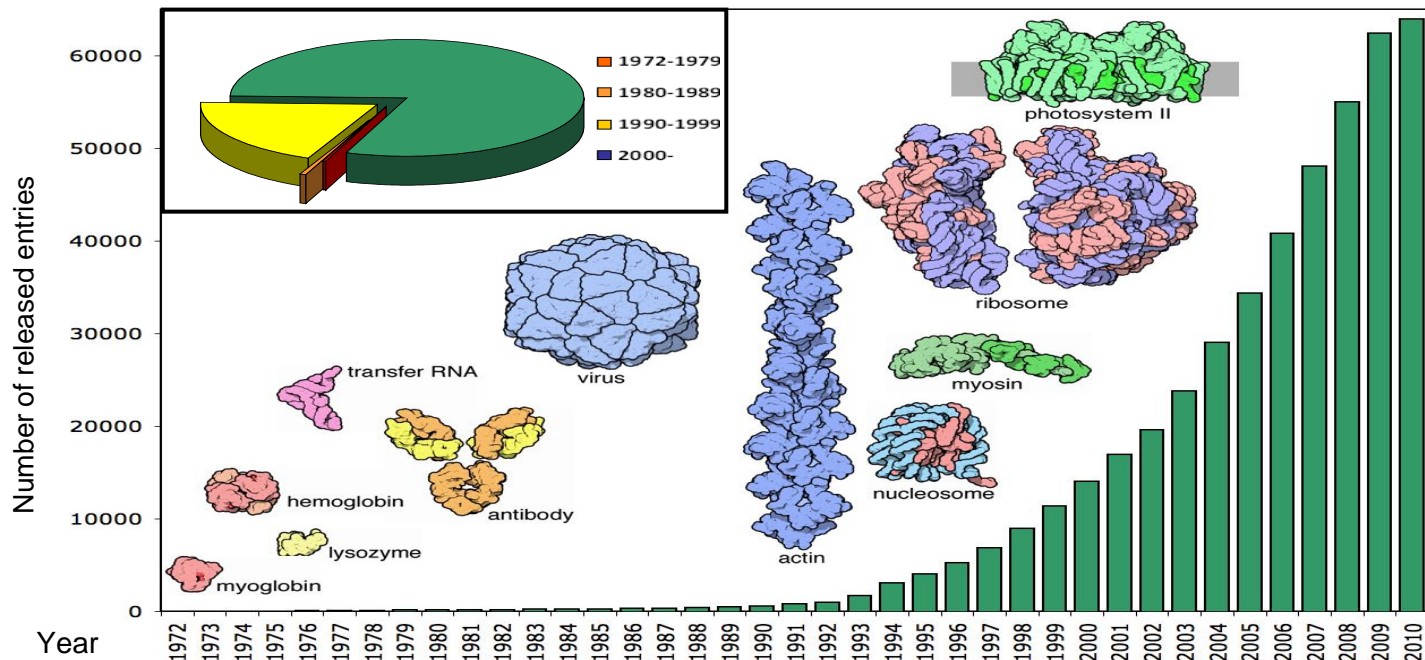


- Protein

- Consists of one or more polypeptides (chains)

# Where can I get Protein Structures?

- RCSB Protein Data Bank (PDB) (<http://www.rcsb.org>)
- Experimentally-determined structures of proteins, nucleic acids, and complex assemblies
- Building blocks of all forms of life
- Hosted at SDSC
  - RCSB PDB: collaboration between Rutgers University and UCSD

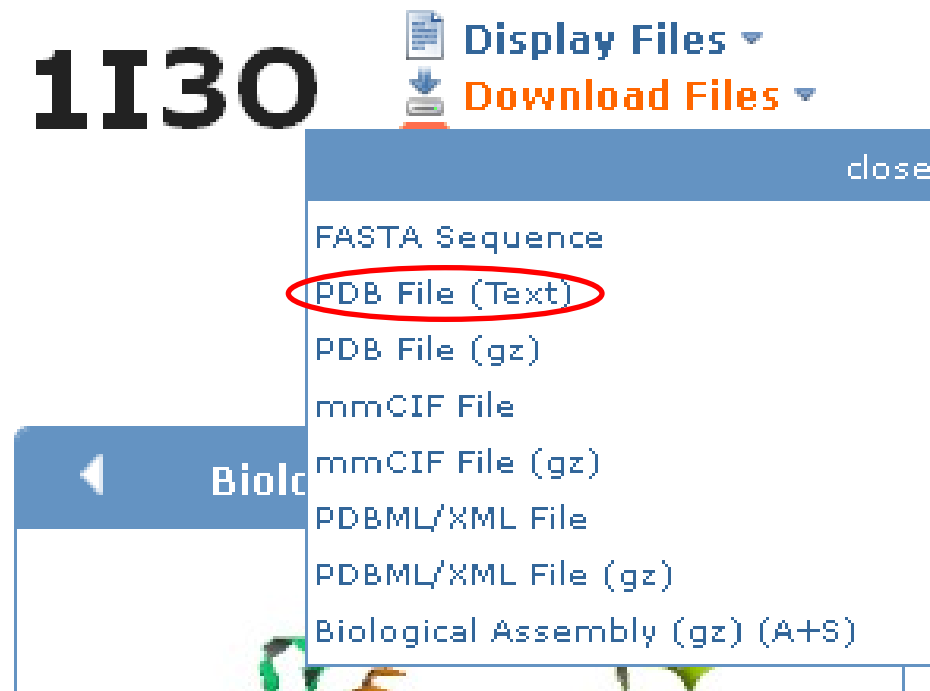


# How do I access Protein Structures?

- Go to <http://www.rcsb.org>
- The pre-build model is a caspase-3 structure
- Can you find caspase-3 structures?
- Can you find the Caspase Molecule of the Month article?
- You can find specific protein structures by their 4-character code
  - 1I3O (letter “o”, not zero!) for the pre-build model

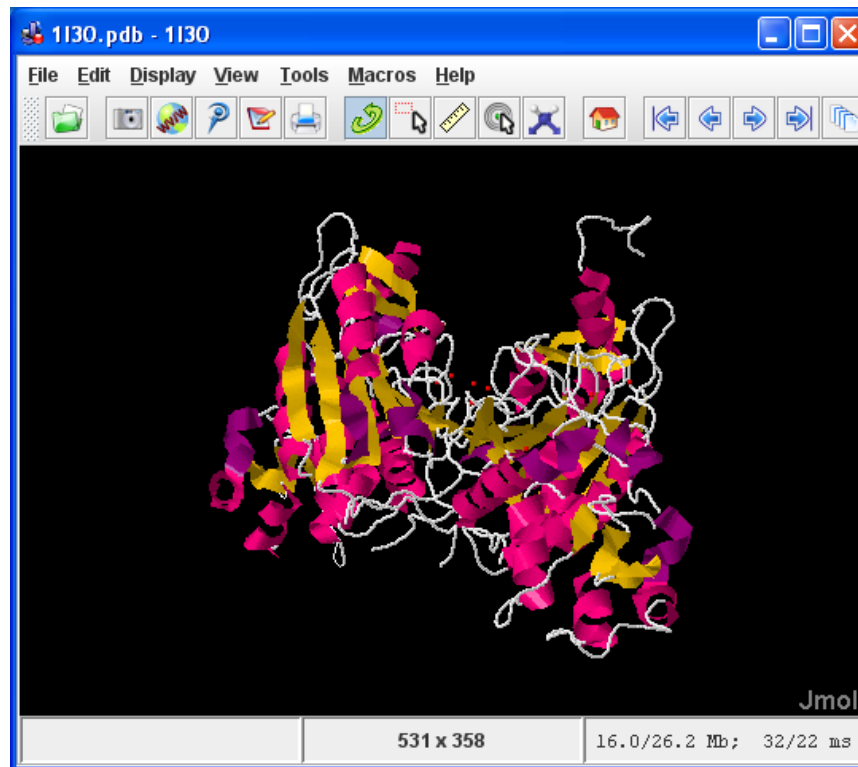
# How do I download a PDB structure?

- On the structure summary page for 1I3O select Download Files->PDB File (Text)
- Save this file (1I3O.pdb) to your desktop



# Visualize Protein with Jmol

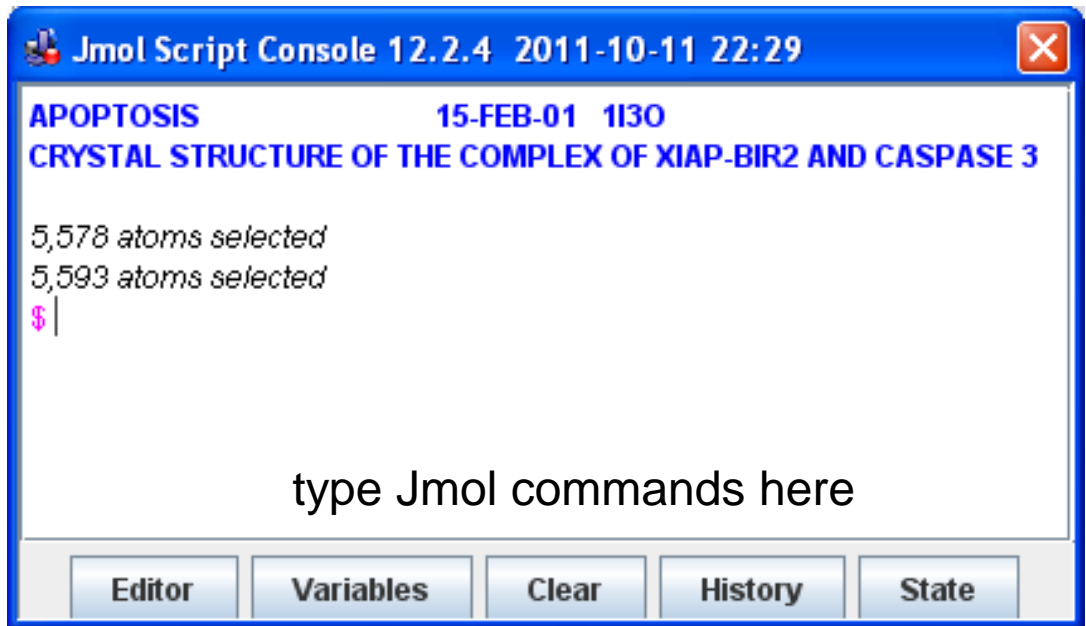
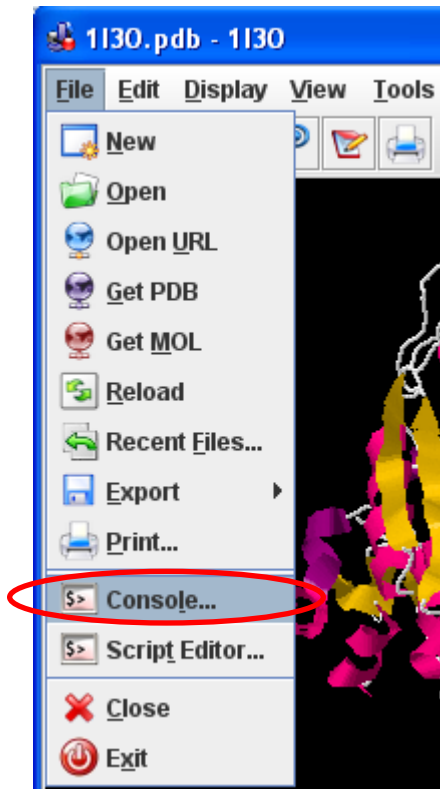
- Click on the Jmol shortcut on your desktop
- Open file 1I30.pdb from your desktop



- Alternatively, you get it directly from the PDB by File->Get PDB if you have internet access

# Jmol Console

- Science Olympiad will use Jmol commands instead of the Jmol menus



# Changing Graphical Representation of Macromolecules



cartoon

backbone

trace

cpk

lines connect the  $\alpha$  carbon atoms

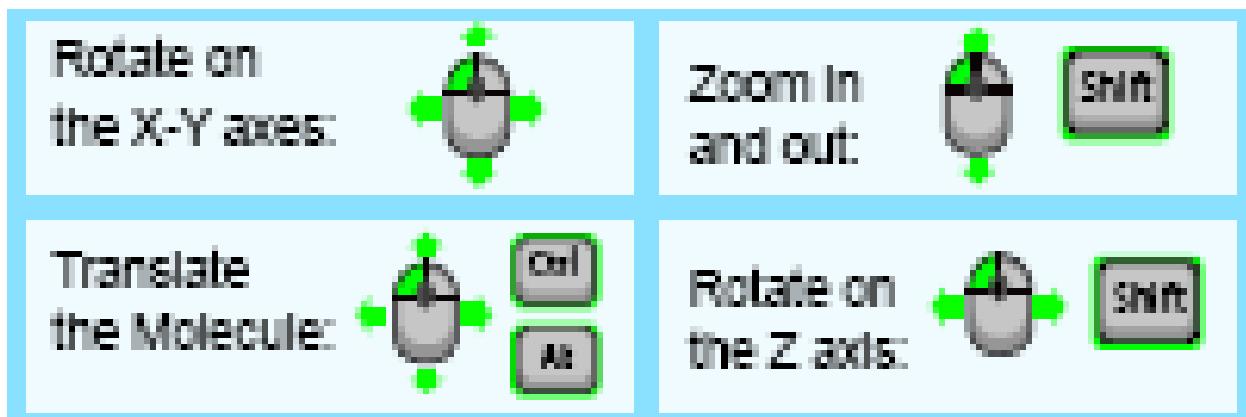
all atoms as “balls”

Type the following Jmol commands into the console:

1. cartoon off
2. cpk off
3. backbone on
4. backbone 100
5. backbone off
6. trace 100
7. trace 200
8. trace off
9. cpk
10. cpk off
11. cartoon on

Tip: to change background color to white type: background white

# How to use the Mouse?

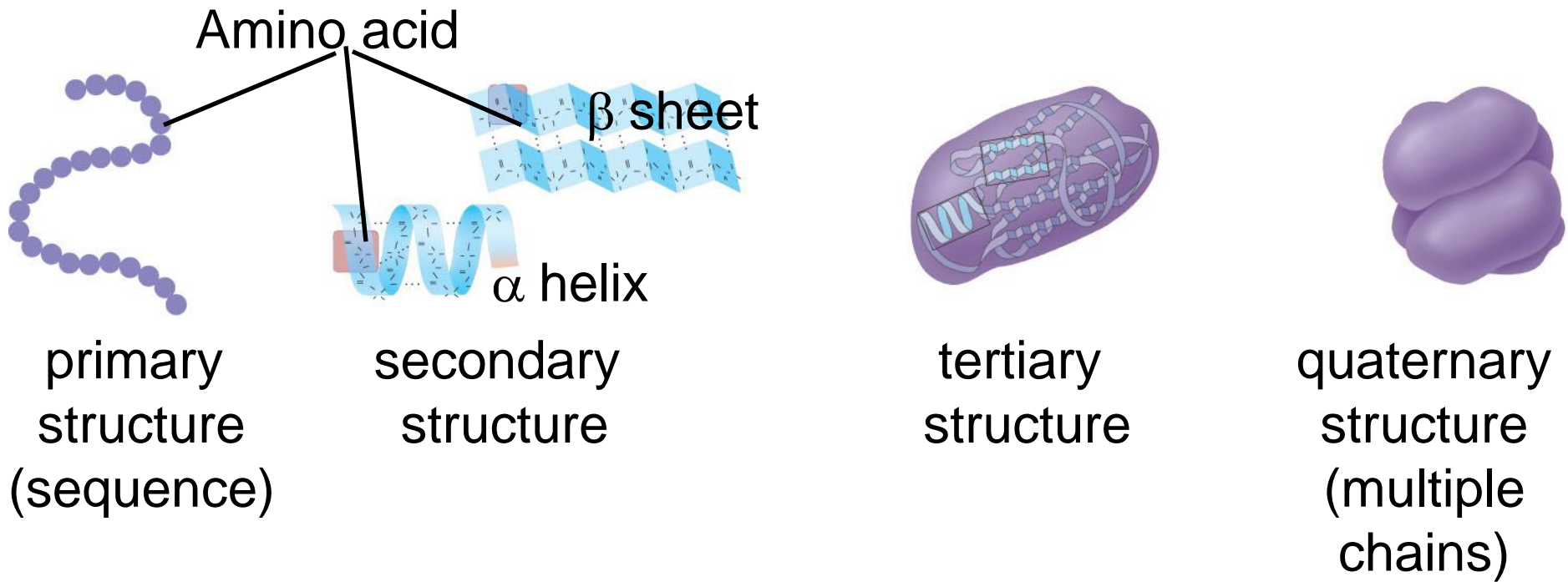


<http://cbm.msoe.edu/includes/jmol/SOJmols/JmolQuickReferenceSheet.pdf>

Another good Jmol tutorial:

[http://www.callutheran.edu/Academic\\_Programs/Departments/BioDev/omm/scripting/molmast.htm](http://www.callutheran.edu/Academic_Programs/Departments/BioDev/omm/scripting/molmast.htm)

# 4 Levels of Protein Structure



# Find a simple Beta Sheet

- Run the following advanced search at <http://www.rcsb.org>

The screenshot shows the 'Advanced Search Interface' on the RCSB PDB website. The search criteria are set to 'Secondary Structure Content'. The search is based on the percent content or number of alpha helices and beta sheets. The search results show 17 PDB Entries.

Advanced Search Interface

Secondary Structure Content

Search by percent content or number of alpha helices and beta sheets

	Percent		Number		
	min	max	min	max	
Alpha Helices:	<input type="text"/>	- <input type="text"/>	0 <input type="text"/>	- 0 <input type="text"/>	Result Count 17 PDB Entries
Beta Sheets:	50 <input type="text"/>	- 100 <input type="text"/>	2 <input type="text"/>	- 2 <input type="text"/>	

Add Search Criteria +

Remove Similar Sequences at 90% Identity  
Match all of the above conditions.

Clear All Parameters Submit Query

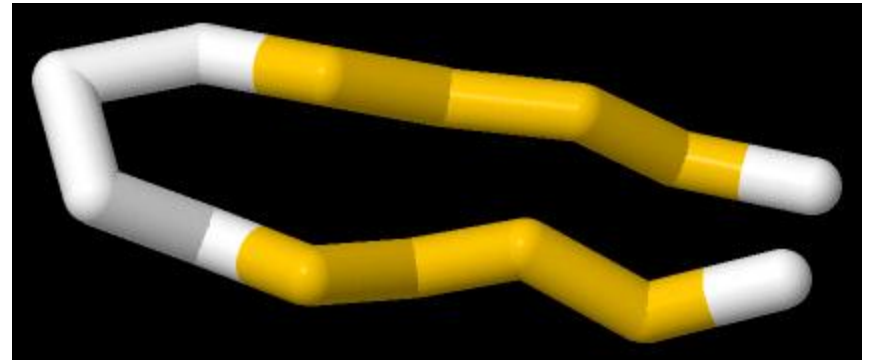
- Review the list of results
- Click on PDB entry 1NIZ and download the PDB file to your desktop

# Visualize 1NIZ in Jmol

Open 1NIZ.pdb

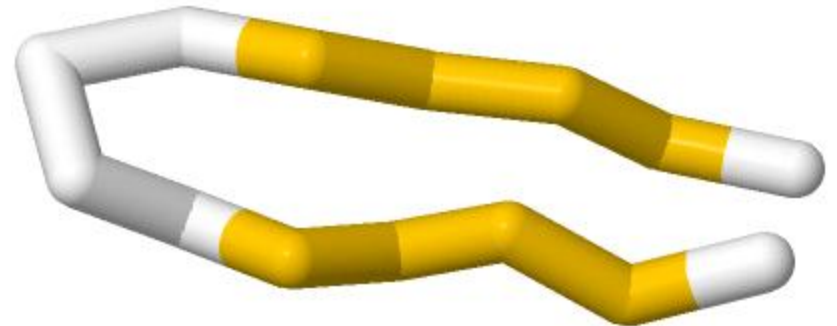
Display the backbone

```
cartoon off  
backbone 200
```



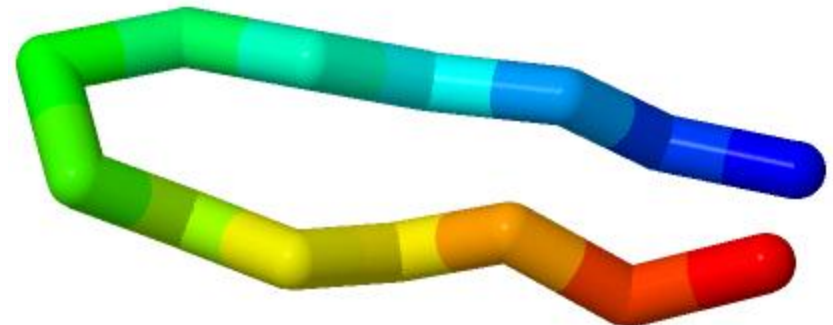
Change background to white

```
background white
```

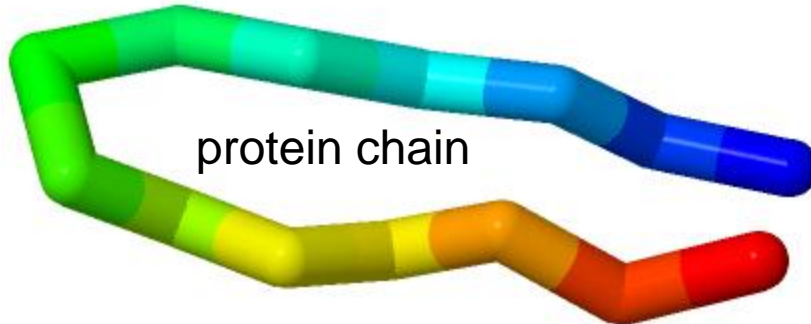


Show monomers

```
color monomer
```



# Primary Structure



Begin of protein chain:  
amino-terminus or  
N-terminus (blue)

End of protein chain:  
carboxyl-terminus or  
C-terminus (red)

Move mouse over an atom  
for more information



Legend:

LYS: lysine, first residue of the protein chain  
5: residue number 5  
A: chain A  
CA: C- $\alpha$ , the alpha carbon of LYS  
#2: atom number 2

*What is the residue number of the last residue?* 18

*What type of amino acid is the last residue?* THR: threonine

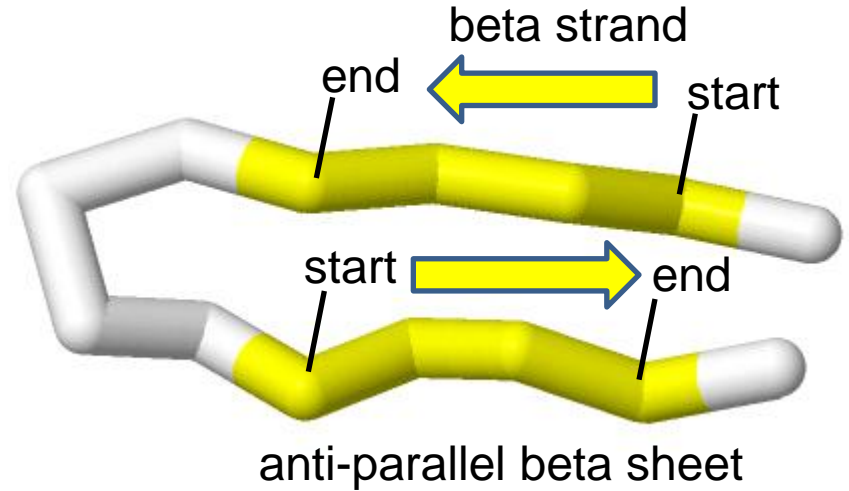
*How many residues does this protein have?* 14

# Secondary Structure

Color residues of beta sheet

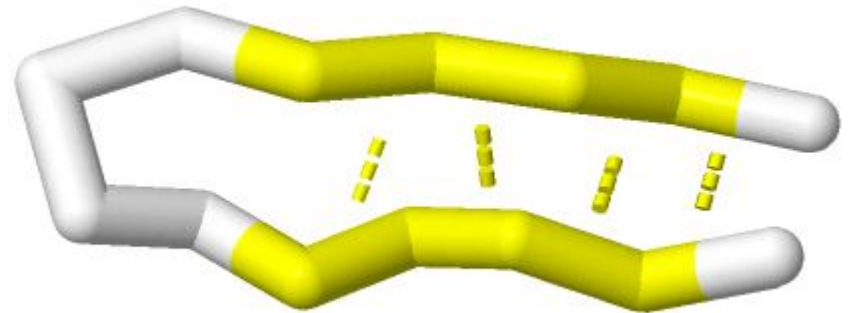
select all  
color white

select sheet  
color yellow



Show hydrogen bonds

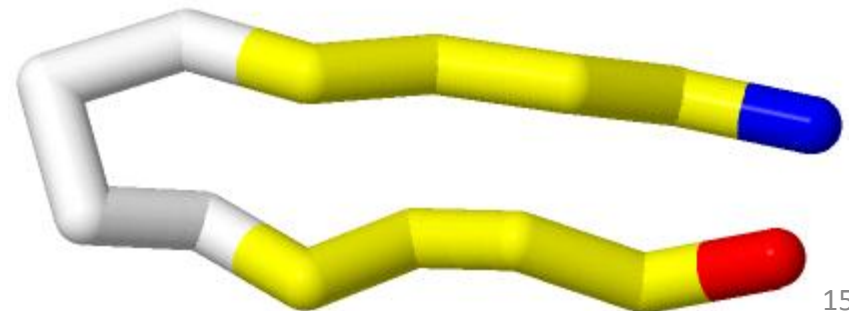
calculate hbonds  
hbonds 50



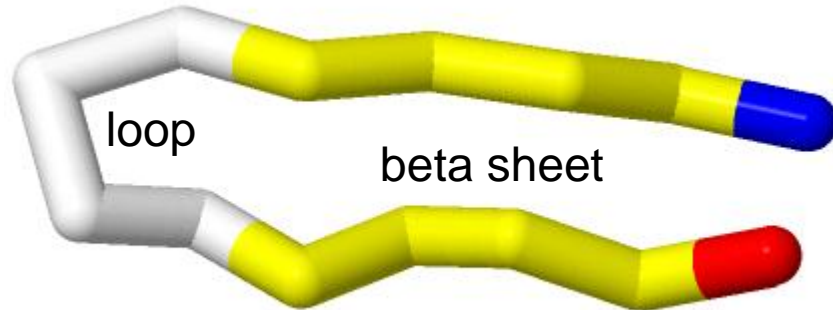
Color N- and C-terminus

select 5:A  
color blue

select 18:A  
color red



# Tertiary Structure



folded protein chain: tertiary structure

Build a wire model of this protein at a scale of 2 cm per residue

- Cut wire
- Mark each residue with sharpie
- Mark beginnings and ends of beta strands with electrical tape
- Bend the beta strands
- Fold the protein into its final shape (tertiary structure)
- Add blue and red wire connectors to the N- and C-terminus

# Find a simple Alpha Helix

- Run the following advanced search at <http://www.rcsb.org>

The screenshot shows the 'Advanced Search Interface' for the RCSB PDB. The search criteria are set to 'Secondary Structure Content'. The search is based on the percent content or number of alpha helices and beta sheets. The search parameters are:

	Percent		Number	
	min	max	min	max
Alpha Helices:	50	100	1	1
Beta Sheets:			0	0

The result count is 305 PDB Entries. There is a button 'Add Search Criteria' and a checkbox 'Remove Similar Sequences at 90% Identity'. The match criteria are set to 'all' of the above conditions. There are buttons for 'Clear All Parameters' and 'Submit Query'.

- Review the list of results
- We will use PDB entry 2L5M

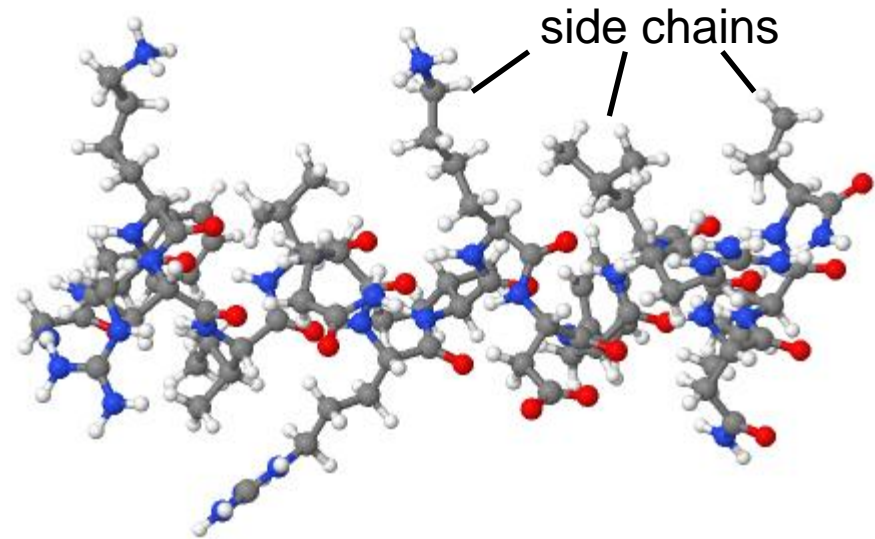
# Visualize 2L5M in Jmol

Get 2L5M directly from PDB

File->Get PDB 2L5M

cpk atom colors:

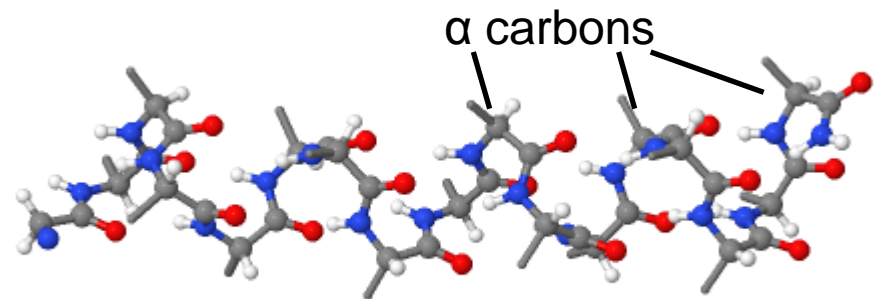
- Grey – carbon
- White – hydrogen
- Blue – nitrogen
- Red – oxygen
- Yellow - sulfur



protein in "ball and stick"  
(cpk and wireframe)

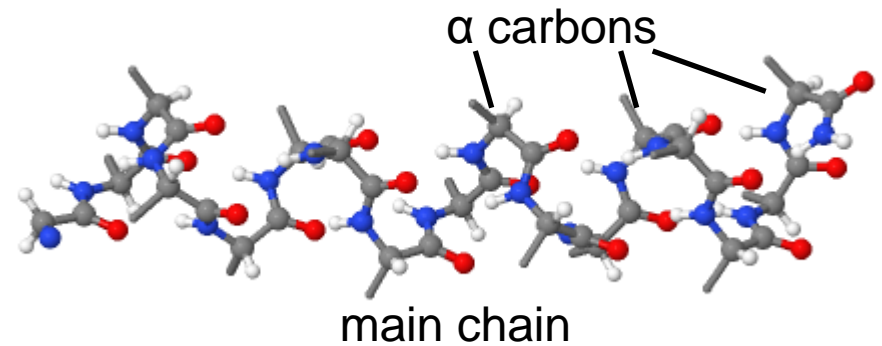
Show main chain only

select sidechain  
wireframe off  
cpk off



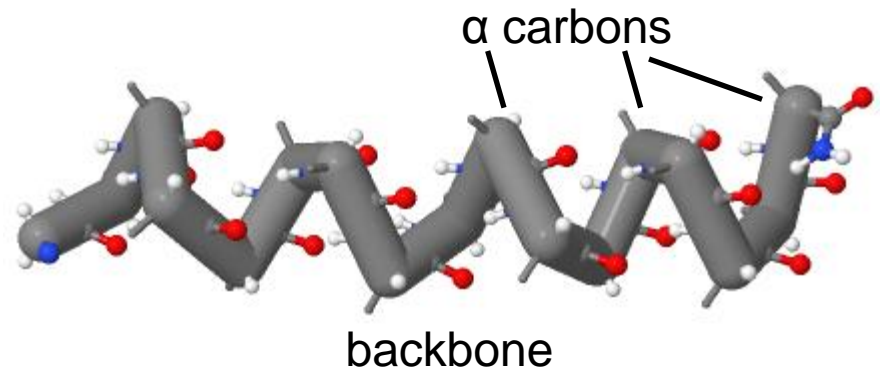
main chain

# Simplify Display of Protein



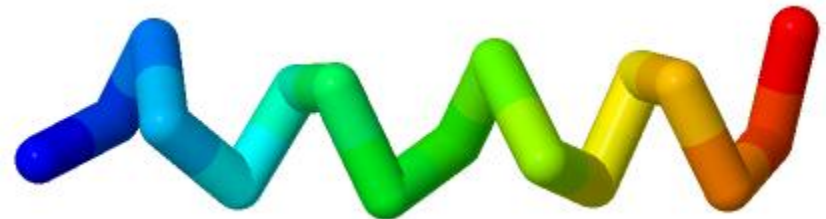
Show backbone

select all  
backbone 200



Show backbone with monomers

wireframe off  
cpk off  
color monomer



*How many residues are in this helix?*<sup>19</sup>

# Secondary Structure – Alpha Helix



Length of helix: 17 residues

Color helix

```
select helix  
color red
```



Show hydrogen bonds

```
calculate hbonds  
hbonds 50
```



Hydrogen bonds between residues  $i$  and  $i + 4$

# Helix Length

*What are the atom numbers of the N- and C-terminal alpha carbons?*

Tip: move mouse of the atom

Measure distance between N- and C-terminus

monitor 2 299



length = 2.47 nm =  $2.47 \times 10^{-9}$  m =  $24.7 \times 10^{-10}$  m = 24.7 Å (ångström)

# Helix Turns, Pitch, Rise

*How many turns are in this helix?*

4 ½ turns

---

*How many residues are there per turn?*

$(17-1) / 4.5 = 3.6$  residues / turn

---

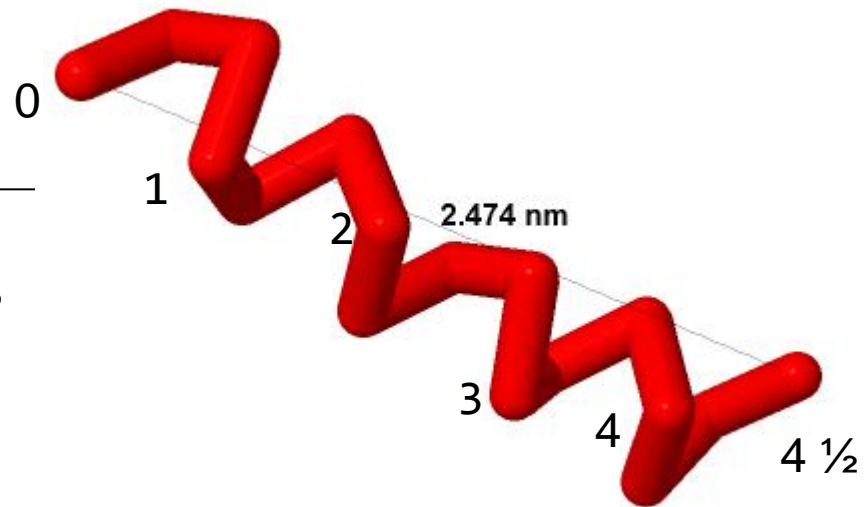
*What is the pitch of this helix (distance between two turns)?*

pitch = length / number of turns =  $24.7 \text{ \AA} / 4.5 = 5.5 \text{ \AA}$

---

*What is the rise per residue?*

rise =  $24.7 \text{ \AA} / (17-1) = 1.5 \text{ \AA}$



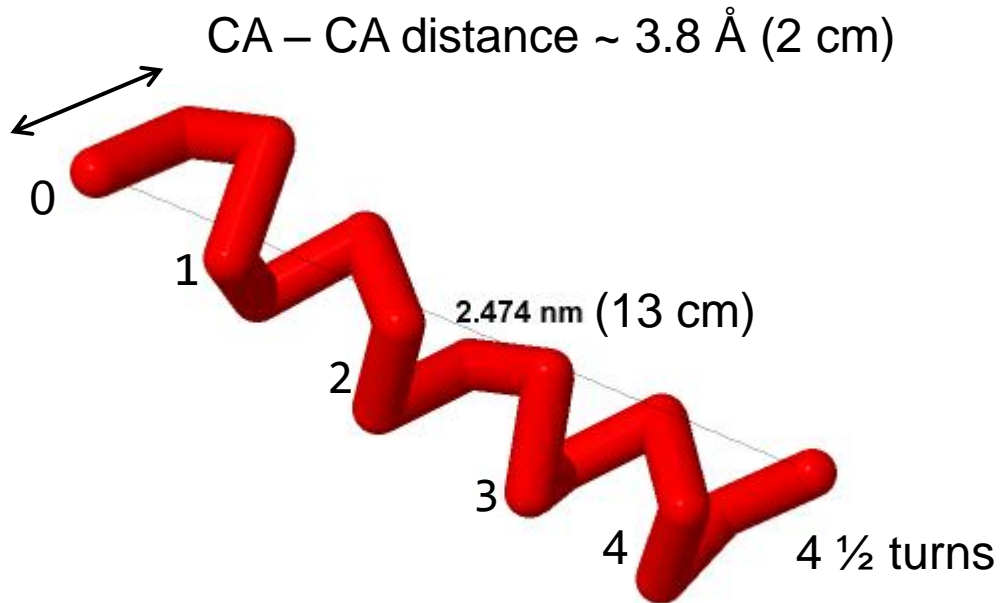
# Alpha Helix Properties

- Hydrogen bonds between residues  $i$  and  $i + 4$
- 3.6 residues per turn
- 1.5 Å per residue (rise)
- 5.4 Å per turn (pitch) =  $1.5 \text{ Å} * 3.6$
- Right-handed

1 Å (ångström) =  $10^{-10} \text{ m}$  =  $0.1 \times 10^{-9} \text{ m}$  = 0.1 nm



# Build an Alpha Helix



How long should this helix be in our model?

$$3.8 \text{ \AA} \rightarrow 2 \text{ cm}$$

$$24.7 \text{ \AA} \rightarrow x \text{ cm}$$

$$x = 13 \text{ cm (length of folded helix)}$$

Build a wire model of this protein at a scale of 2 cm per residue

- Cut wire
- 4  $\frac{1}{2}$  right-hand turns around wooden dowel (2 cm diameter)
- Adjust length to 13 cm

# Build Zinc Finger Model

- Use the sample competition environment
  - Google for: zinc finger sample
  - <http://cbm.msoe.edu/includes/jmol/SOJmols/ZincFingerSample.html>
  - This environment will be used during the competition

# Setup Zinc Finger Model

1. Display zinc ion as cpk

```
select Zn; cpk 300
```

2. Display all secondary structure elements

```
select all; color structure
```

3. Display and color important side chains

Find the two CYS and two HIS residues that bind to zinc.

Then display and color the side chains, for example [CYS]7:C  
(CYS, residue 7, chain C)

```
select 7:C and (sidechain or alpha)  
wireframe 75  
cpk 100  
color cpk
```

4. Color the N- and C-terminus blue and red

# Building Exercise: Zinc Finger Model

- Group A (6 students)
  - Use mini-toober set
  - These materials will be used at the competition and are provided by the organizers
- Group B
  - Build backbone model of Zinc finger using wire, electrical tape, and wire connectors
- Then group A and B switch activities

# Pre-build Competition Environment



Important resources

Parts of pre-build set  
\$20 + S&H  
or use  
alternative materials

A screenshot of a web browser window displaying the "2012 Science Olympiad Protein Modeling Event" page. The page has a red header and contains the following sections:

- Pre-Build Competition**: Caspase-3 based on 1i3o.pdb.
- Directions:** The Pre-Build Model should represent chains A and B of caspase-3 based on the PDB file 1i3o.pdb. This section of the protein has two chains (A and B). Chain A is 143 amino acids long and will require a toober that is 286cm long. Chain B is 101 amino acids long and will require a toober that is 202cm long. (1 amino acid = 2cm)
- Resources:** [Pre-Build Amino Acid Numbering Map](#), [Jmol Quick Reference Sheet](#)
- Parts List:**
  - 1 Mini-Toober (286cm long)
  - 10 crosslinkers
  - 2 red end caps
  - 1 Mini-Toober (202cm long)
  - 2 blue end caps

The right side of the browser window shows a Jmol application with a 3D wireframe model of the protein. Below the model is a command window with the text "Script completed" and "messagecallback = 'showmsg'".

Jmol application  
pre-loaded with  
Caspase-3  
(PDB file 1i3o.pdb)

Jmol command  
window

<http://cbm.msoe.edu/includes/jmol/SOJmols/2012PreBuild.html>

With this environment, you can only work on the pre-build model. The on-site competition will use this environment, so you should become familiar with it.

SD Science Olympiad will not impound or score this model, but it will be scored at the state or national events.

# Explore Pre-build Competition Environment

- Uses PDB entry 1I3O.pdb (letter “o”, not zero!)
- Structure summary page:  
<http://www.rcsb.org/pdb/explore/explore.do?structureId=1i3o>
- 1I3O has 6 chains (A, B, ..., F)
  - quaternary structure
- The Pre-Build Model should represent chains A and B of caspase-3
- Need to restrict model to chains A and B

# Tips for Setting up the Pre-Build Model

Restrict display to chains A and B

```
restrict :a,:b
```

Display chain A in yellow

```
select :a; color yellow
```

Display chain B in cyan

```
select :b; color cyan
```

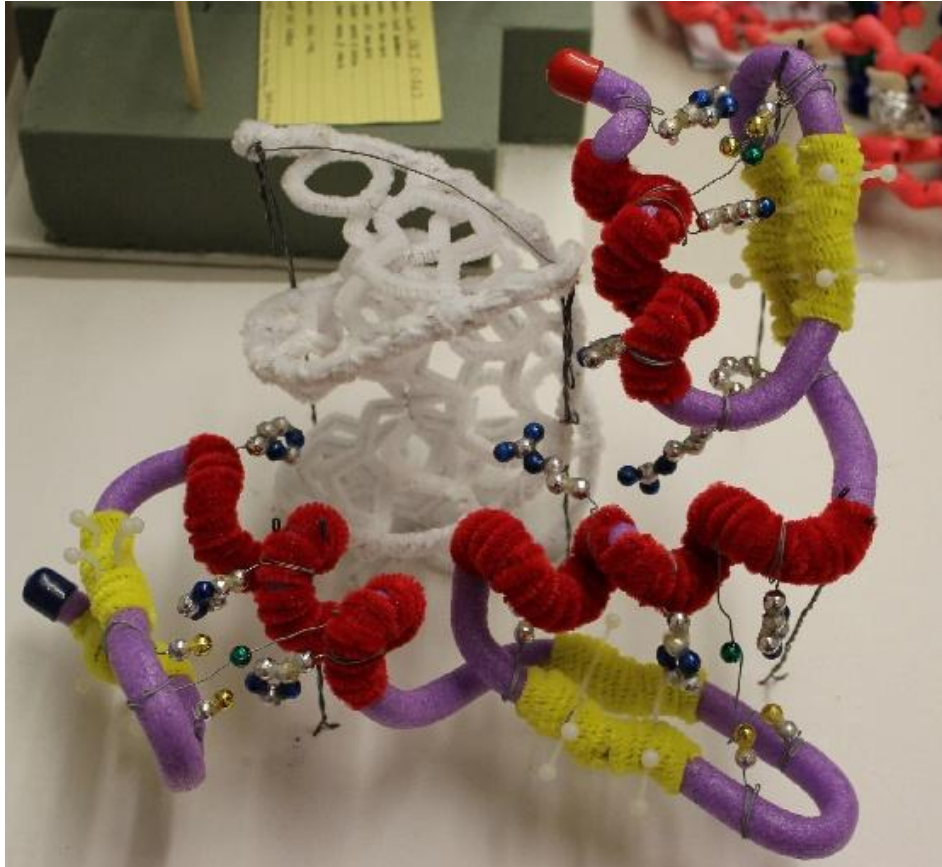
Display all secondary structure

```
select all; color structure
```

# Cautions about 2012 Event

- Pre-Build Competition Environment
  - Review “Pre-Build Amino Acid Numbering Map”
    - Numbering is non-consecutive!
  - Jmol Quick Reference Sheet has errors!
- 2012 On-Site Description
  - <http://cbm.msoe.edu/includes/pdf/SO/2011-2012OnsiteBuildInformation.pdf>
  - Description of written test is incorrect !
  - See 2012 Protein Modeling rules for written test description

# Example of a Pre-Build Model



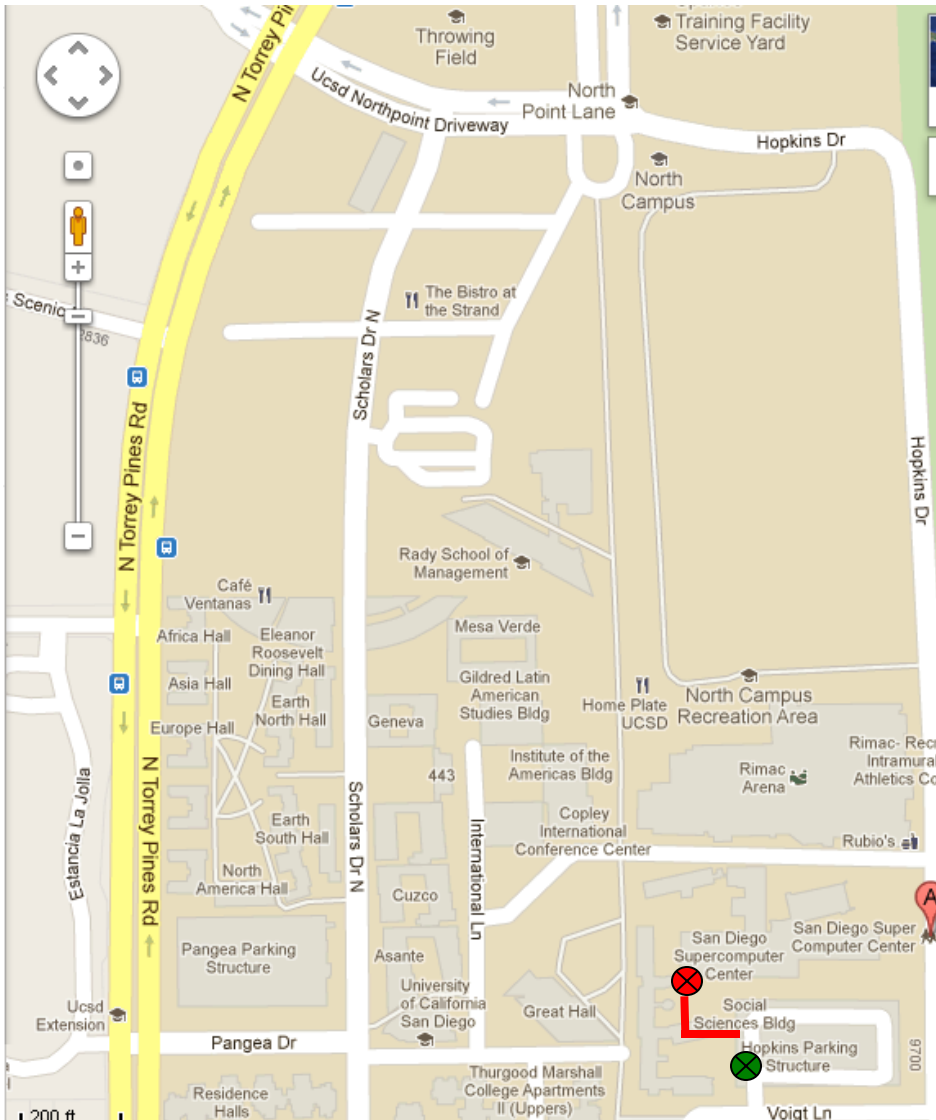
<http://cbm.msoe.edu/stupro/so/Photos.html>

- Mini-toober model with creative “decorations” that illustrate function of protein
  - Secondary structure
  - Important side chains
  - Additional molecules of importance (here DNA)
  - (cross-linkers or wires to stabilize model)
- Arts & Crafts stores (i.e., Michaels) are a good places to look for “decorations”
- How are pre-build models scored at national tournaments
  - <http://cbm.msoe.edu/stupro/so/SONTEExamRubric.html>

# Protein Modeling C Workshops

- Mastering the Protein Modeling C Event -  
Sat. Jan. 7, 2012  
San Diego Supercomputer Center/UCSD, Room 279\*,  
2:00 – 4:00 pm
  - Introduction to the on-site model
  - Advanced hands-on protein visualization and model building at computer lab
  - In-depth review of the biological concepts related to this event

# Workshop Location: San Diego Supercomputer Center, UCSD



- From Torrey Pines Road turn east onto UCSD Northpoint Driveway
- Continue on Hopkins Dr
- Turn right on Voigt Ln and drive up the steep hill. Park in the Hopkins Parking Structure (free on weekends)
- Go to the top of the parking structure (level 7) and cross the pedestrian bridge to the Social Sciences patio.
- Turn right to exit the patio northward along the north-south walkway of the Social Sciences building. The walkway intersects the sidewalk that leads to SDSC.
- Cross this sidewalk and go up the steps. The training room 279 is straight ahead.