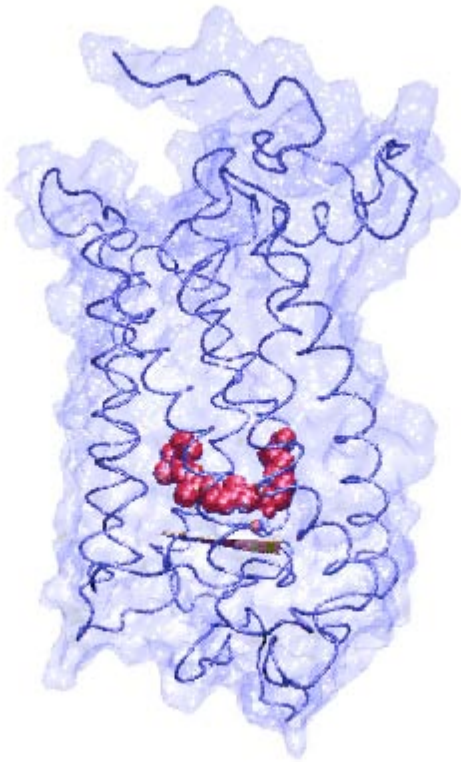
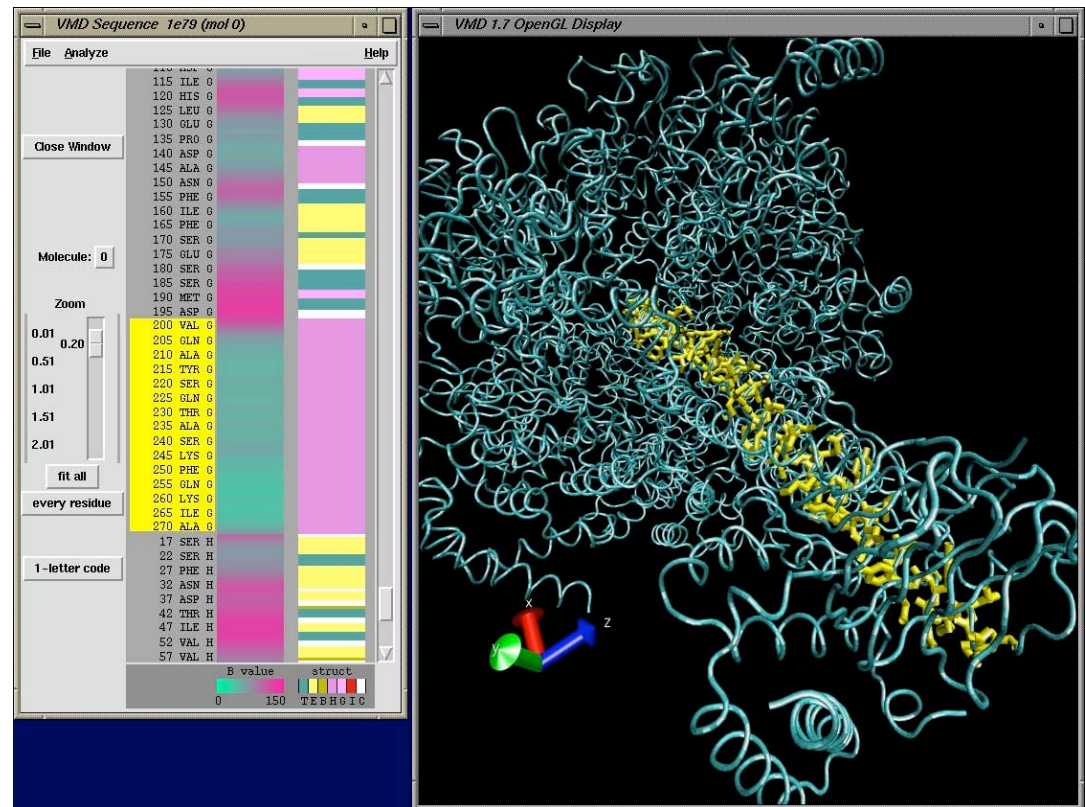


# Molecular Graphics Perspective of Protein Structure and Function



animation



sequence

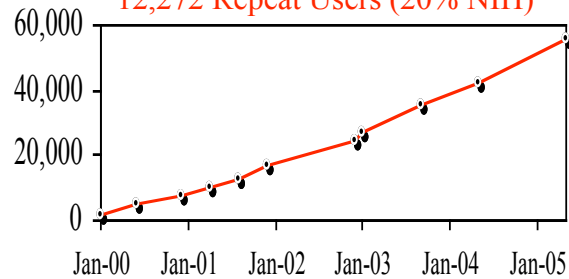
structure

# VMD Molecular Graphics

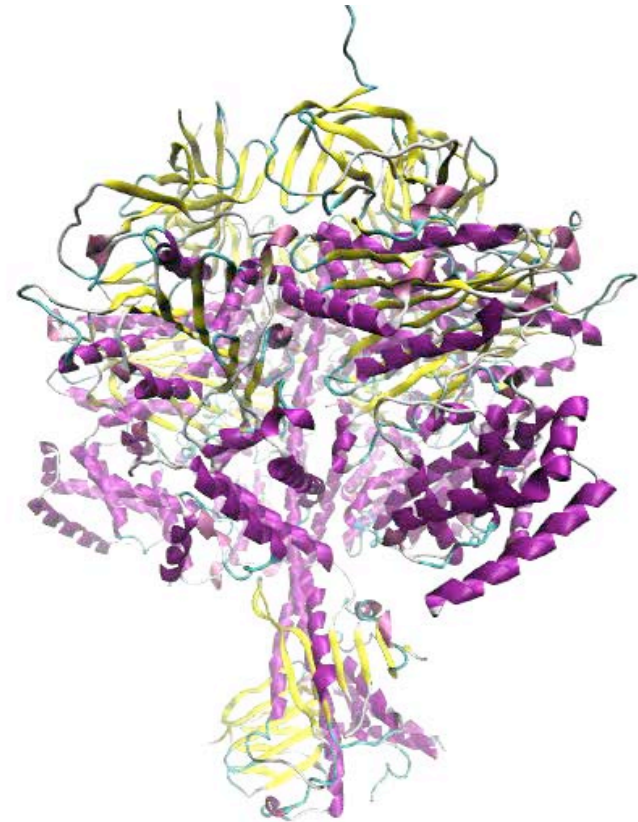
## VMD Registrants

55,422 Registrants (19% NIH)

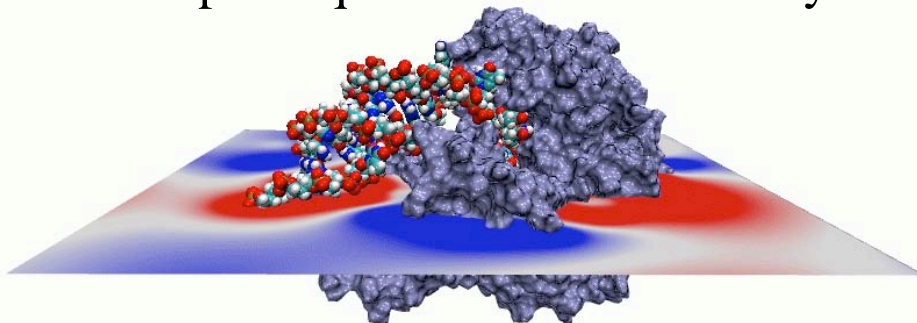
12,272 Repeat Users (20% NIH)



- > 50,000 registered users
- Platforms:
  - Unix (16 builds)
  - Windows
  - MacOS X
- Display of large biomolecules and simulation trajectories
- Multiple sequence - structure analysis



VMD view of F1-ATPase

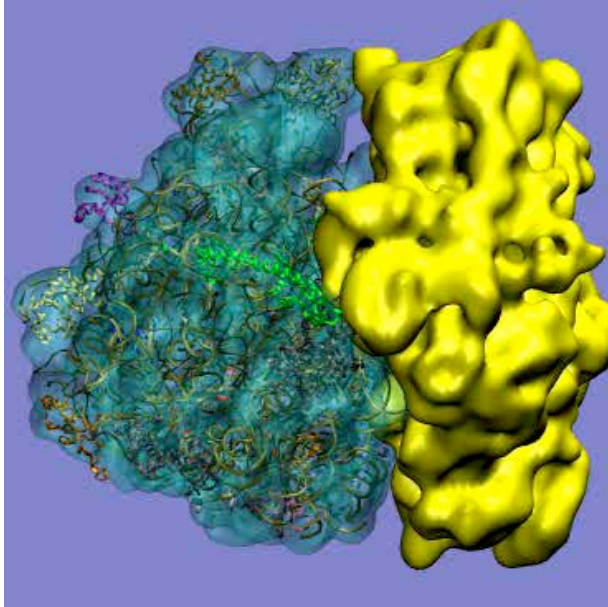


Electrostatic potential for an ATPase  
obtained with VMD's PME plugin

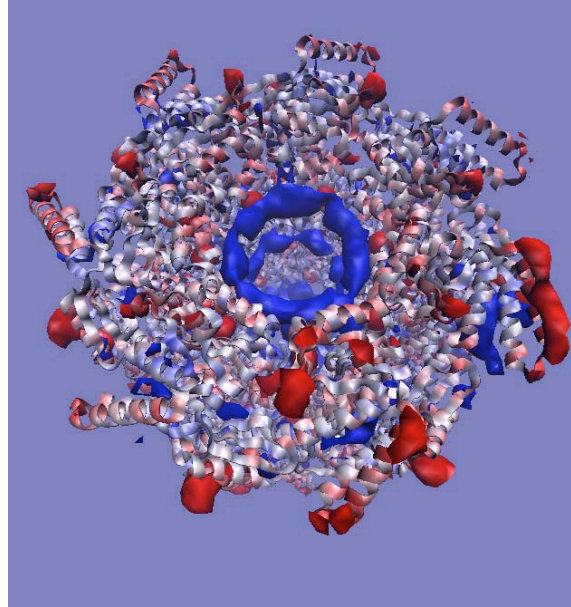


# VMD Permits Large Scale Visualization

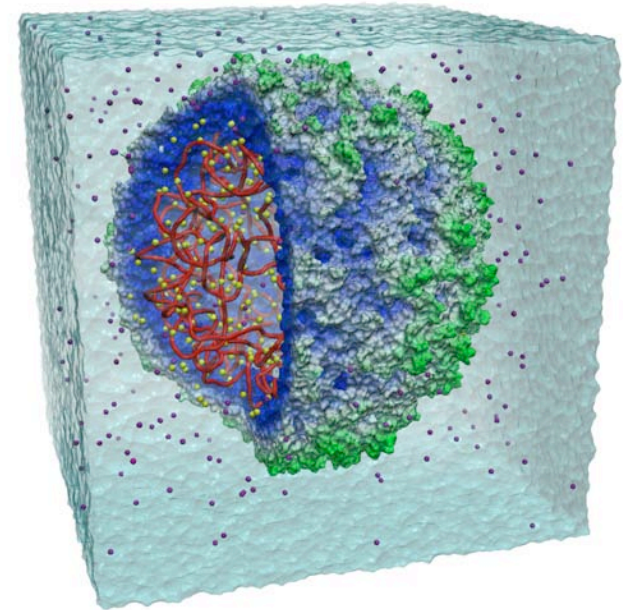
- Large structures: million atoms and up (ribosome in solvent: 2.6 million)
- Complex representations (many types for pdb structures)
- Long trajectories: thousands of time steps (terabytes; on parallel processors)
- Volumetric data (many types of EM, electrostatics, quantum chemistry data)
- 64 bit, I.e., multi-gigabyte data don't break 32-bit barriers
- Handles large data sets, e.g., GlpF: each 5 ns simulation of 100K atoms produces a 12GB trajectory
- Can load entire pdb (17 Gbytes) for pdb wide analysis



Ribosome



GroEL



STMV virus

# VMD Supports Database Access

Q: “How much does the length of arginine vary?”

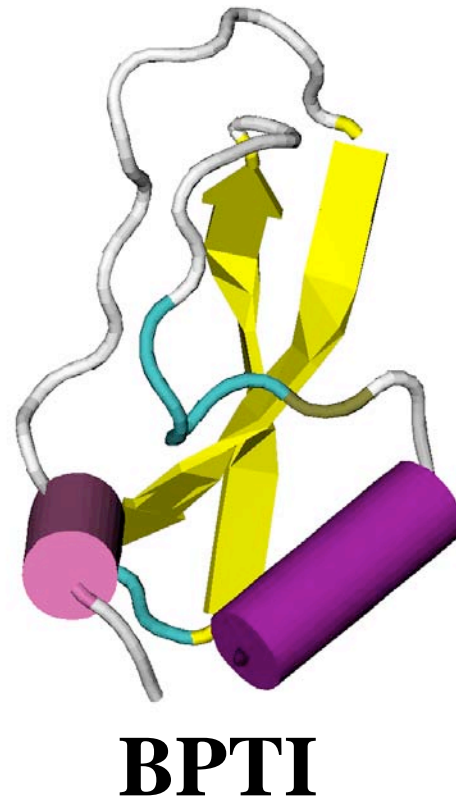
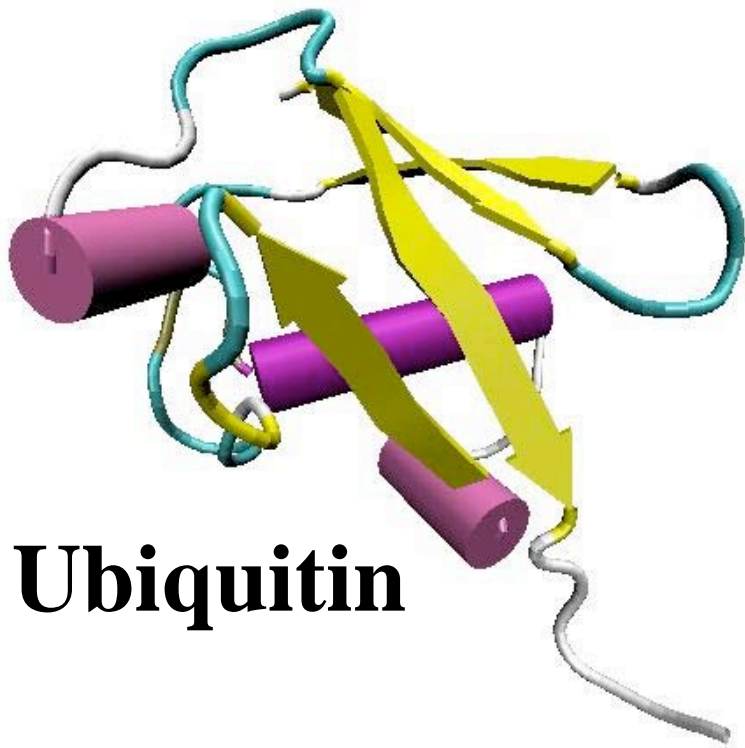
Answer:

- A VMD scan of all structures in the PDB revealed only small fluctuations (less than 10%) in the length of arginine residues
- Scan measured 591,338 arginine residues
- The “length” was defined by the CA-CZ distance
- Measured length was 5.5 +/- 0.49A
- Dominant factor in one-time scan is disk I/O
- Doing scan in memory runs 100x to 1000x faster
- Multiple instances of VMD can scan in parallel

# Focus on two proteins

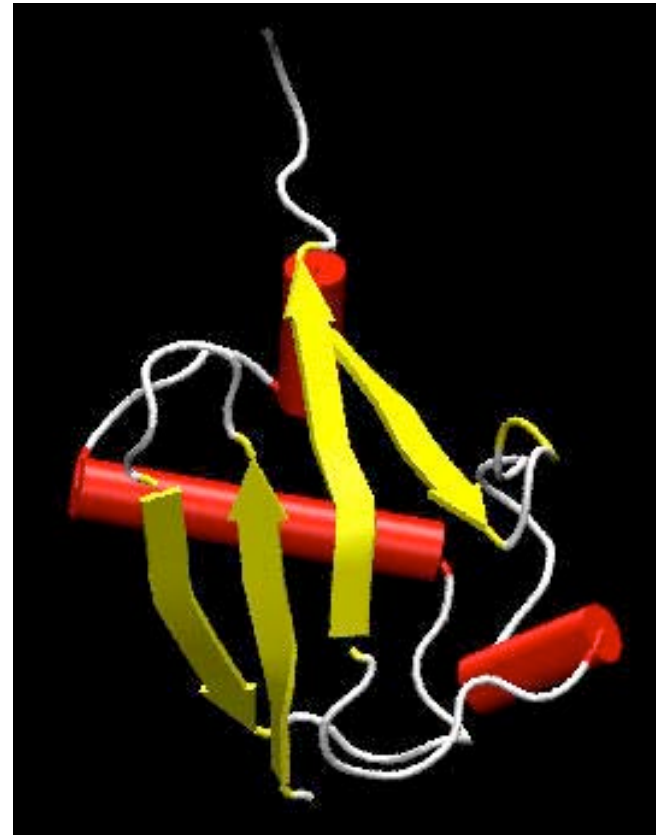
## Ubiquitin

## Bovine Pancreatic Trypsin Inhibitor (BPTI)



# Ubiquitin

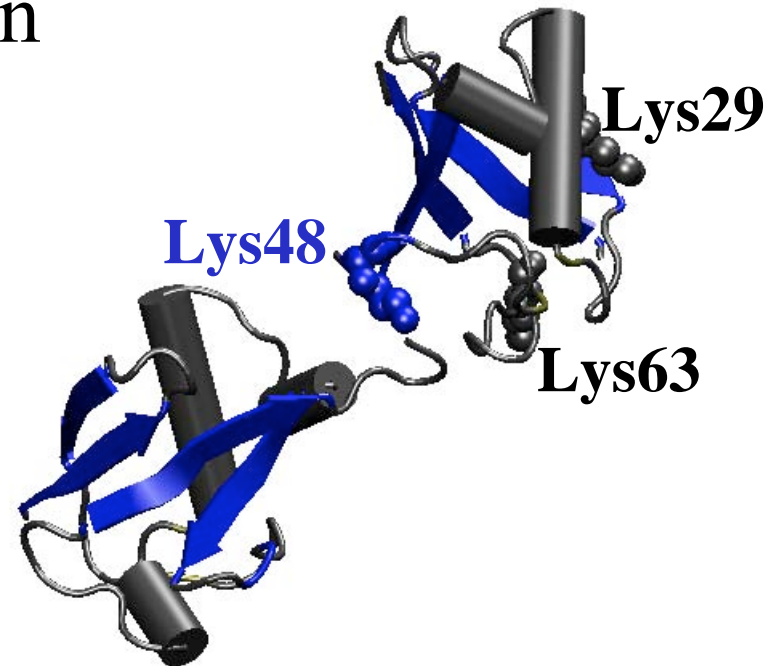
- 76 amino acids
- highly conserved
- covalently attaches to proteins and tags them for degradation
- other cell trafficking



- Glycine at C-terminal attaches to the lysine on the protein by an isopeptide bond.

- it can attach to other ubiquitin molecules and make a polyubiquitin chain.

There are 7 conserved lysine residues in ubiquitin.



Two ubiquitins attached together through LYS 48. LYS 63 and LYS 29 are also shown there.



# Ubiquitination Pathway



The Nobel Prize in Chemistry 2004

"for the discovery of ubiquitin-mediated protein degradation"

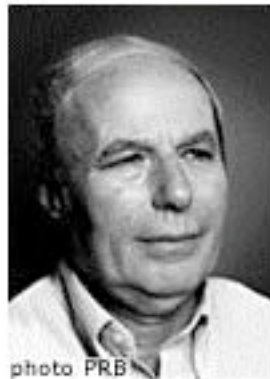


**Aaron Ciechanover**

🕒 1/3 of the prize  
Israel

Technion – Israel  
Institute of  
Technology  
Haifa, Israel

b. 1947



**Avram Hershko**

🕒 1/3 of the prize  
Israel

Technion – Israel  
Institute of  
Technology  
Haifa, Israel

b. 1937  
(In Karcag, Hungary)



**Irwin Rose**

🕒 1/3 of the prize  
USA

University of  
California  
Irvine, CA, USA

b. 1926

Ubiquitin-mediated protein degradation

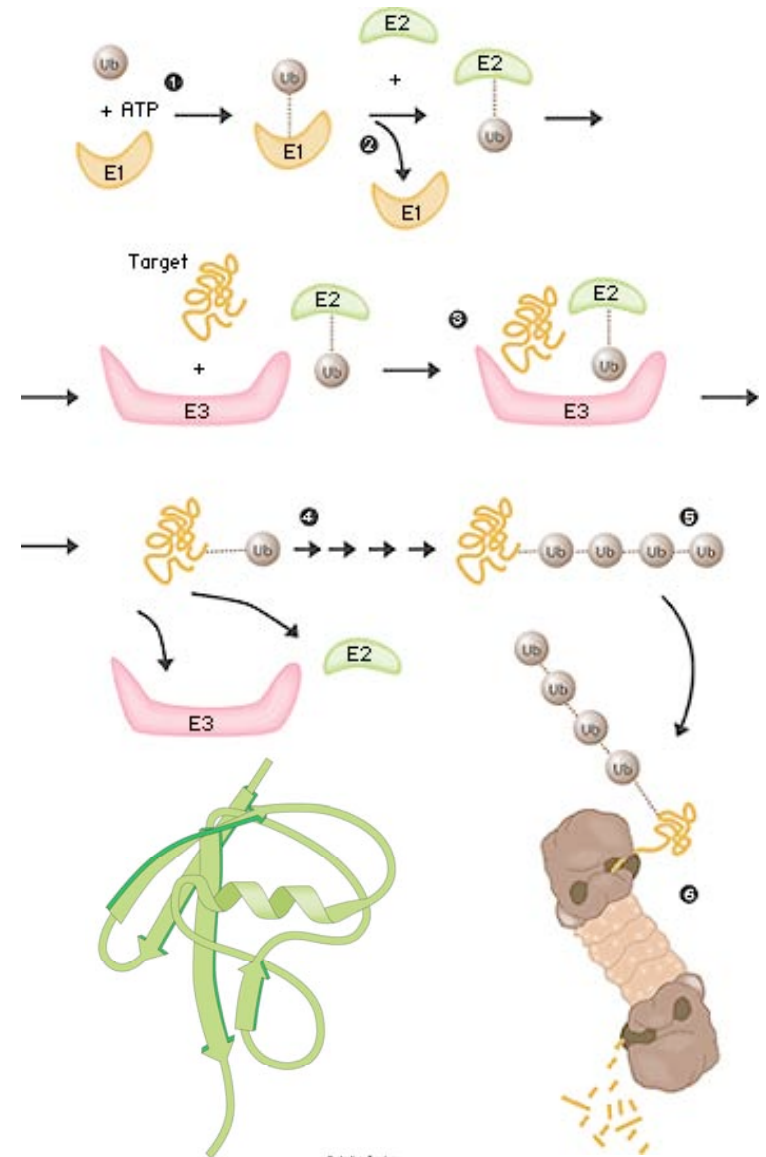
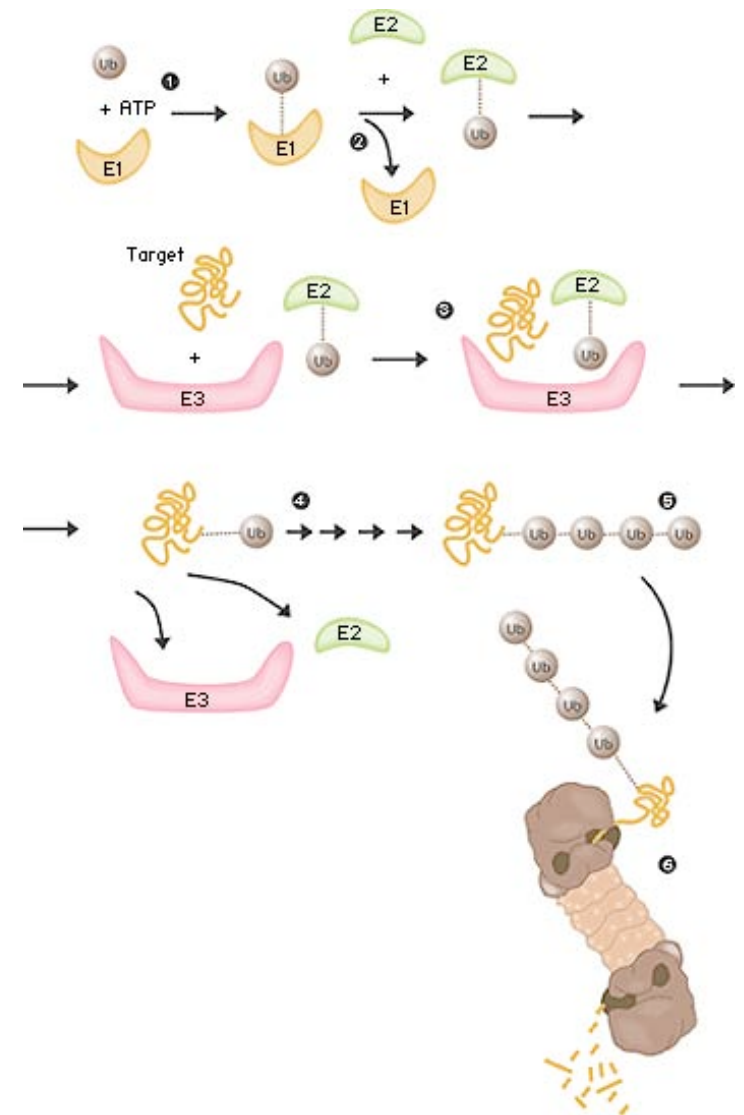


Illustration: Tgatom



# Ubiquitination Pathway

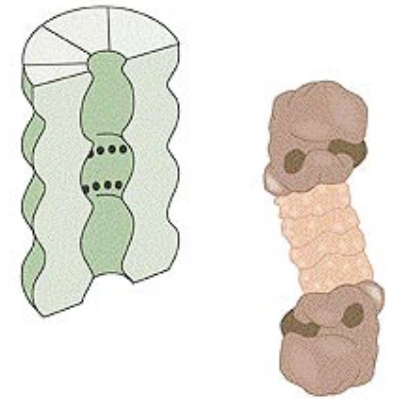
- Activation by E1 (ATP dependent process)  
(thiol-ester linkage between a specific cysteine residue of E1 and Glycine on ubiquitin)
- Transfer to a cysteine residue on E2  
(ubiquitin conjugation enzyme)
- Transfer of ubiquitin by E3 to the substrate lysine residue.
- E3 recognizes the ubiquitination signal of the protein.



# Ubiquitin Functions

- tagging misfolded proteins to be degraded in the proteasome (kiss of death).
- regulates key cellular processes such as cell division, gene expression, ...

A chain of at least four ubiquitins is needed to be recognized by the proteasome.



The cell's waste disposer, the proteasome. The black spots indicate active, protein-degrading surfaces.

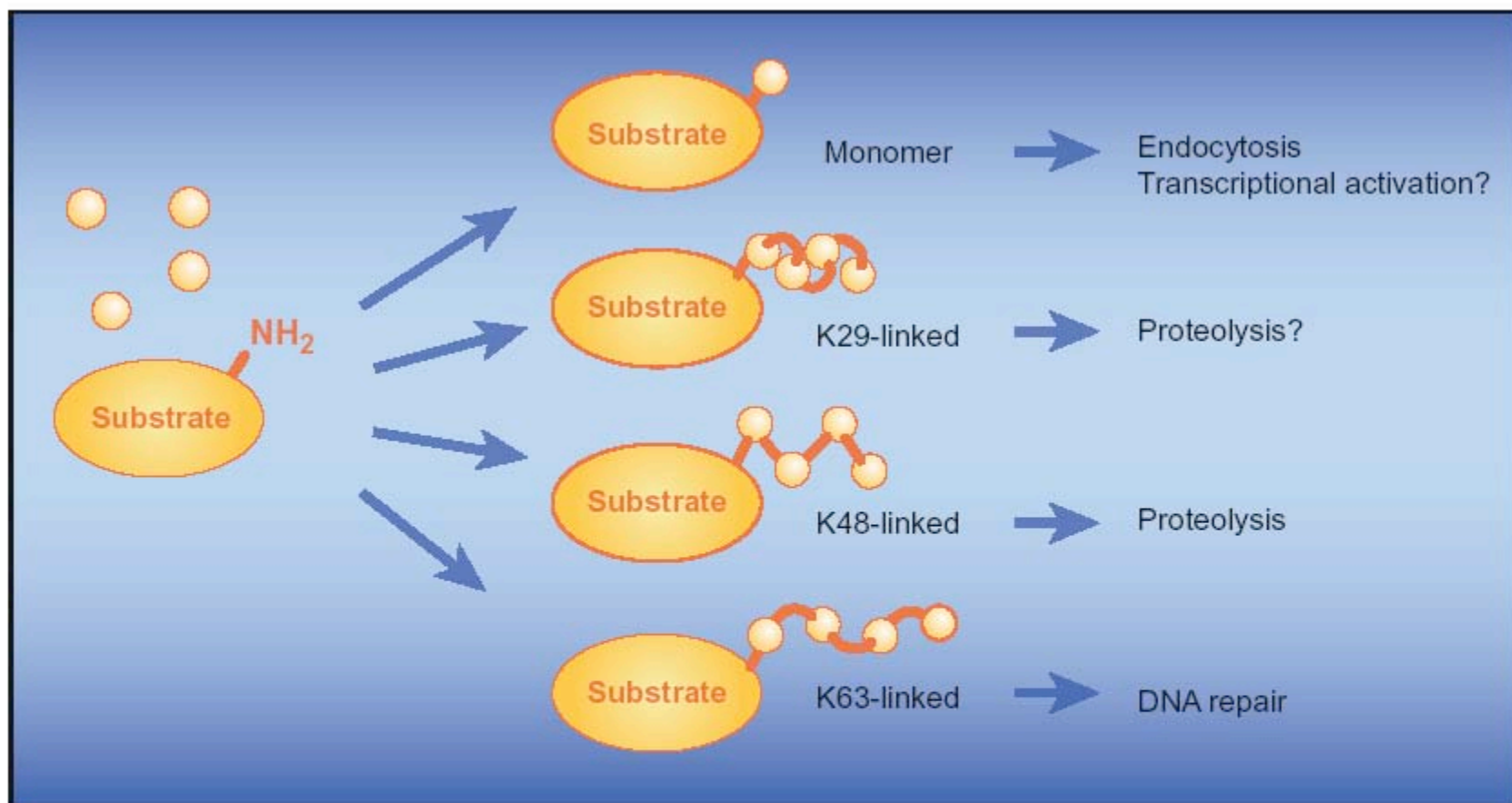
# Different types of ubiquitin signals arise from

- Length of the ubiquitin chain
- How ubiquitins are attached together
- Where the signals are read

## **Examples:**

- multi-ubiquitin chains, linked through Lysine 48, target protein for proteasome degradation
- K63 linkages direct DNA repair

# Mono-ubiquitylation versus multi-ubiquitylation



**Multifaceted.** Ubiquitin can attach to its various substrate proteins, either singly or in chains, and that in turn might determine what effect the ubiquitination has. (K29, K48, and K63 refer to the particular lysine amino acid used to link the ubiquitins to each other.)

Marx, J., Ubiquitin lives up its name, *Science* 297, 1792-1794 (2002)

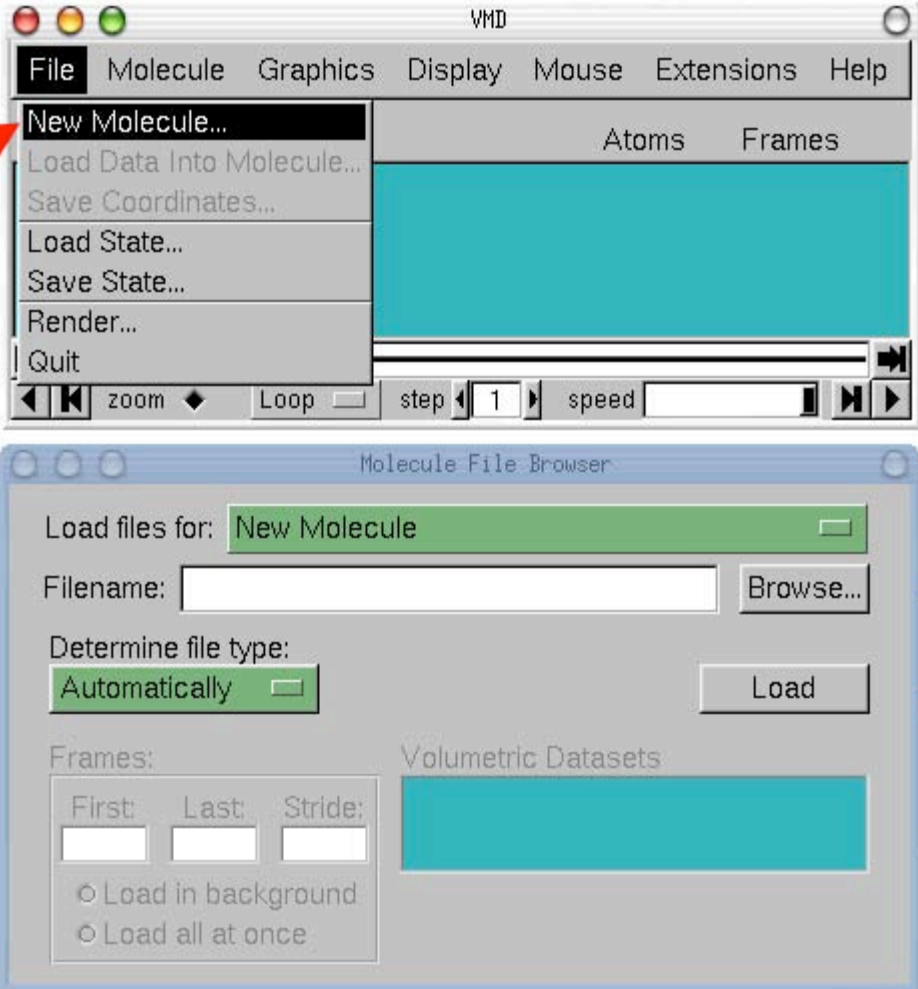


**Inspect ubiquitin with VMD**

# Basics of VMD

## Loading a Molecule

New Molecule (a)



The image shows two windows from the VMD software. The top window is the main VMD interface with the 'File' menu open. The 'New Molecule...' option is highlighted. The bottom window is the 'Molecule File Browser' dialog. It has a 'Load files for:' field set to 'New Molecule', a 'Filename:' field, a 'Browse...' button, a 'Determine file type:' field set to 'Automatically', and a 'Load' button. There are also 'Frames:' fields for 'First', 'Last', and 'Stride', and radio buttons for 'Load in background' and 'Load all at once'.

(b) Molecule file browser

(c) Browse

(d) Load

# Basics of VMD

## Rendering a Molecule

Current graphical representation

(a)

Draw style

(b)

Coloring

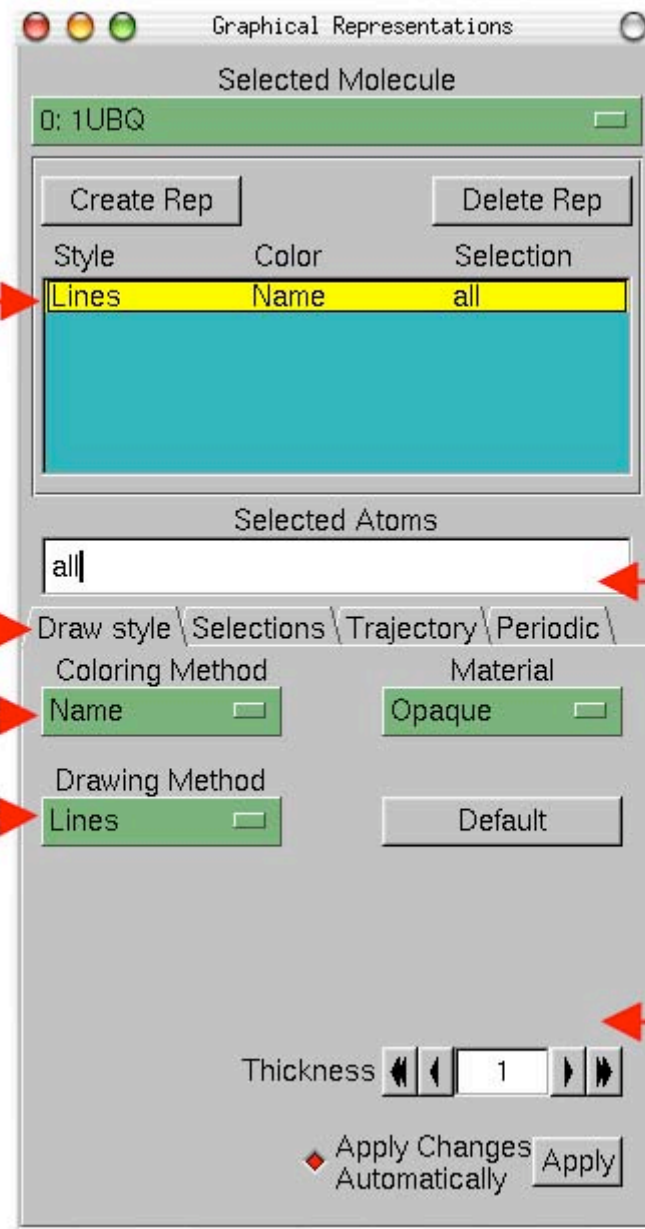
(c)

Drawing method

(d)

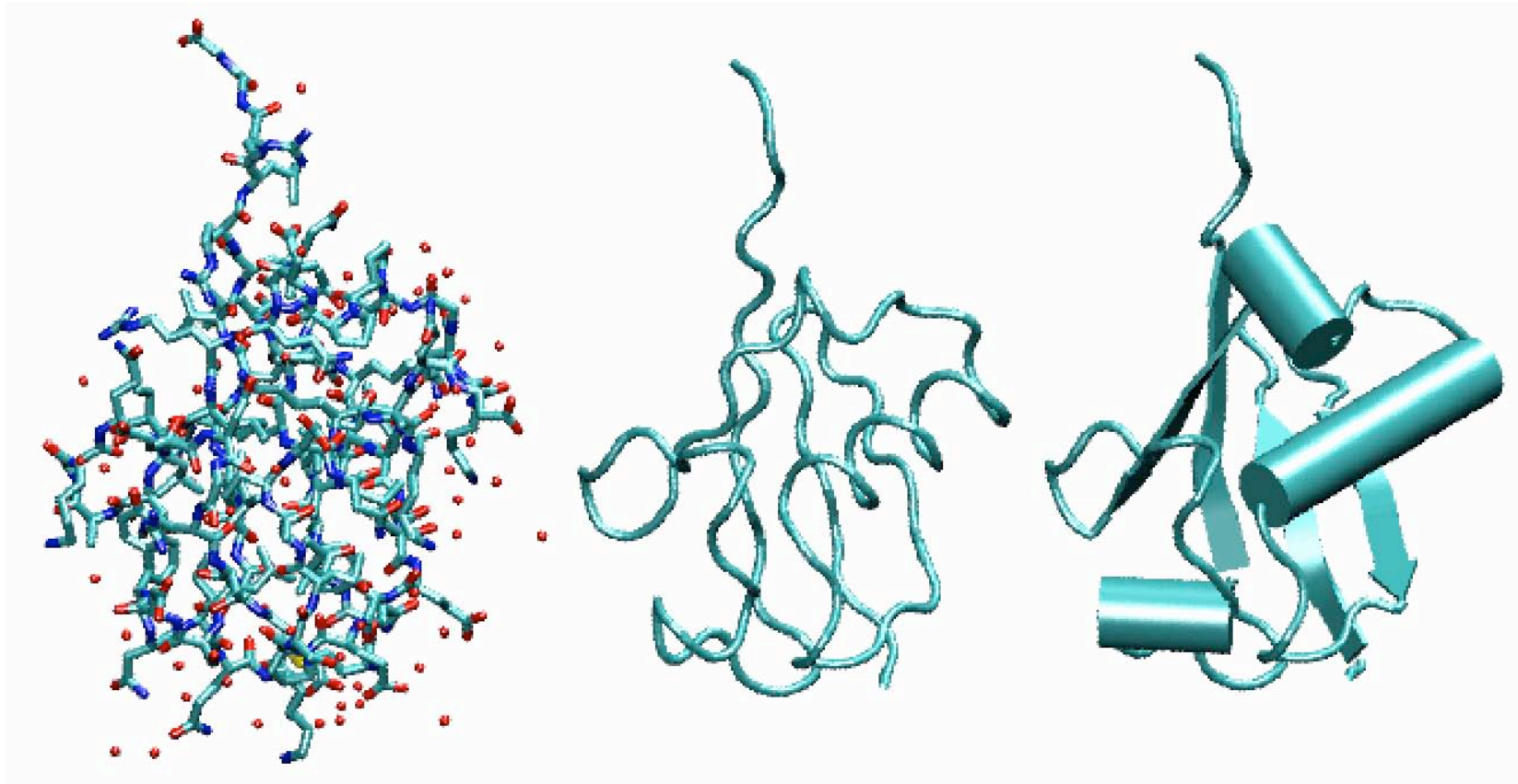
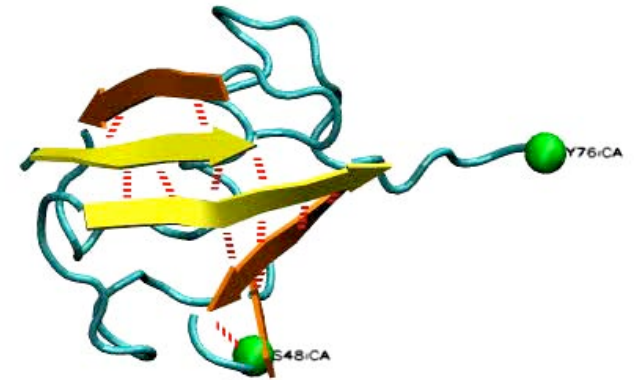
(f) Selected Atoms

(e) Resolution, Thickness



# Basics of VMD

Change rendering style



CPK

tube

cartoon



# Basics of VMD

Create Representation (a)



Create Rep



Delete Representation (b)

Delete Rep

Current Representation (d)

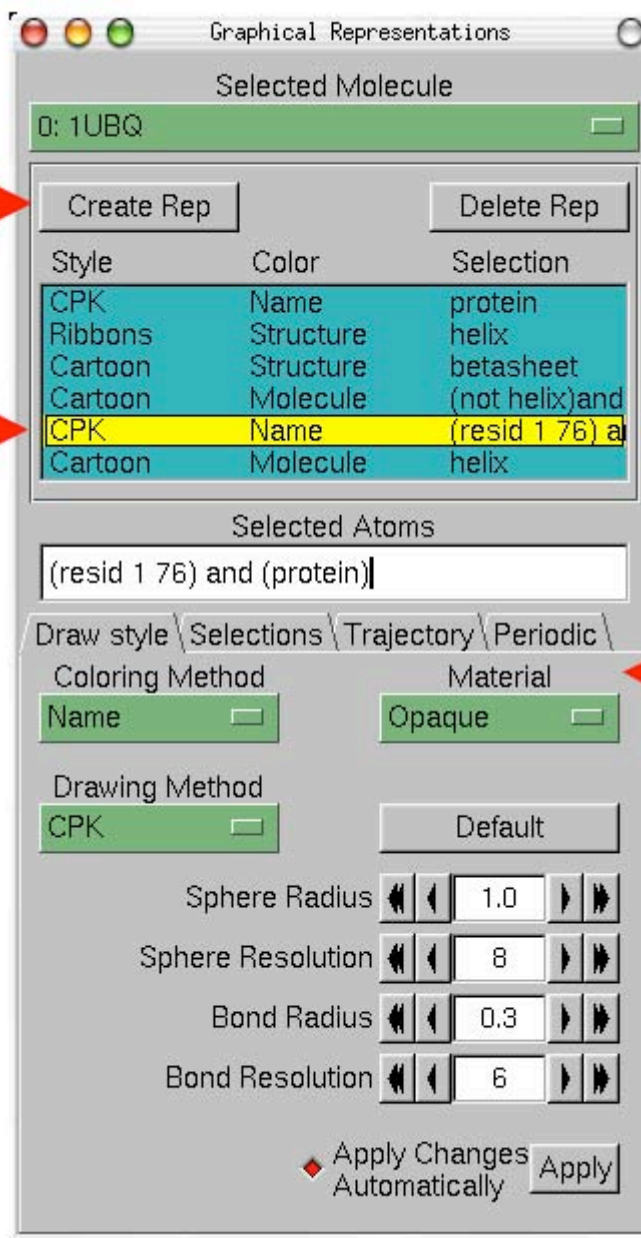


CPK Name (resid 1 76) a

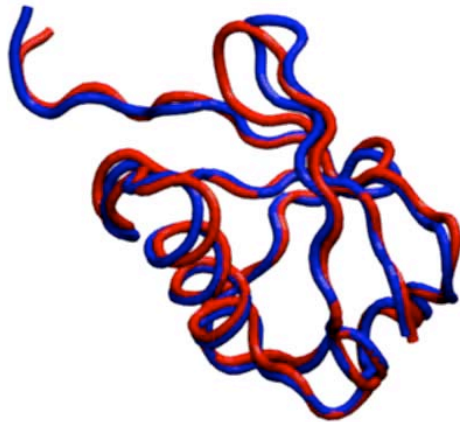
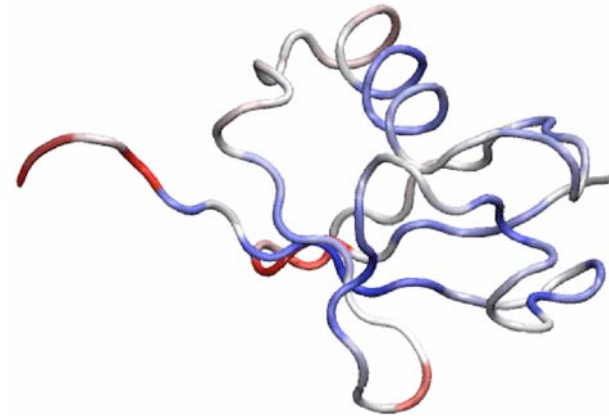
Multiple representations



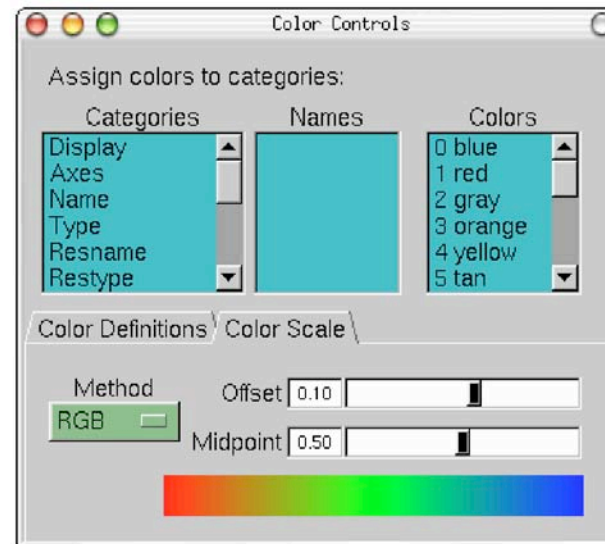
Material (c)



# VMD Scripting



Left: Initial and final states of ubiquitin after spatial alignment  
Right (top): Color coding of deviation between initial and final



The Color Controls window showing the Color Scale tab.

# VMD Sequence Window

(a)

VMD Sequence 1UBQ (mol 0)

File Help

Close Window

Molecule: 0

Zoom

0.01 0.51 1.01 1.00 1.51 2.01

fit all

every residue

1-letter code

23 ILE X  
24 GLU X  
25 ASN X  
26 VAL X  
27 LYS X  
28 ALA X  
29 LYS X  
30 ILE X  
31 GLN X  
32 ASP X  
33 LYS X  
34 GLU X  
35 GLY X  
36 ILE X  
37 PRO X  
38 PRO X  
39 ASP X  
40 GLN X  
41 GLN X  
42 ARG X  
43 LEU X  
44 ILE X  
45 PHE X  
46 ALA X  
47 GLY X  
48 LYS X  
49 GLN X  
50 LEU X  
51 GLU X  
52 ASP X  
53 GLY X  
54 ARG X  
55 THR X  
56 LEU X  
57 SER X  
58 ASP X  
59 TYR X  
60 ASN X  
61 ILE X  
62 GLN X  
63 LYS X  
64 GLU X

B value 0 150

struct T E B H G I C

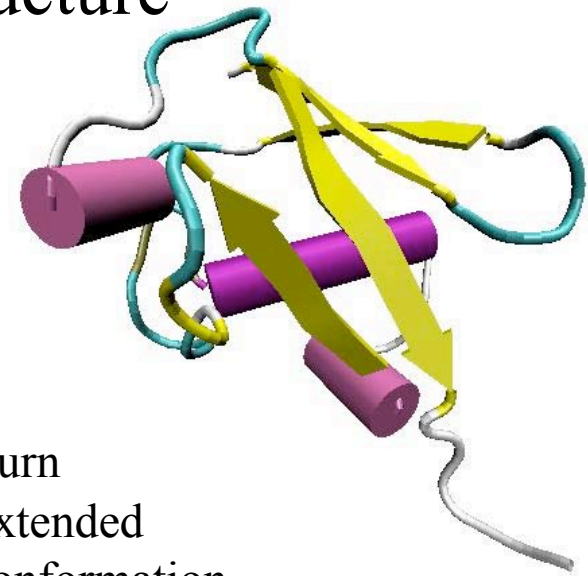
(d)

List of  
the residues

Zoom

(b) Beta Value

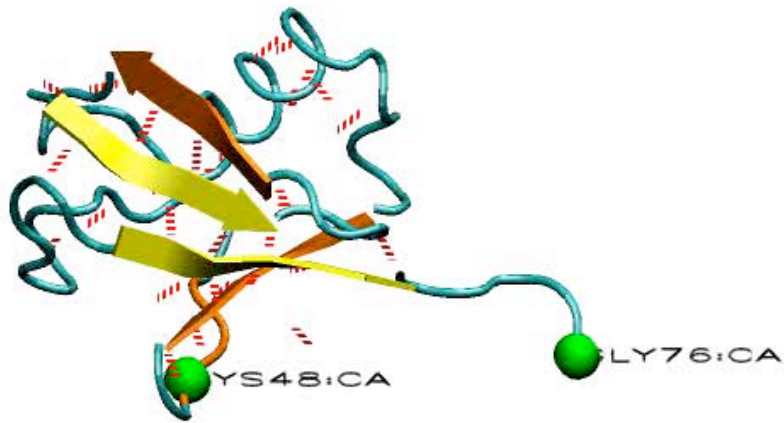
(c) Structure



T: Turn  
E: Extended  
conformation  
H: Helix  
B: Isolated Bridge  
G: 3-10 helix  
I: Phi helix

# VMD Macros to Color Beta Strands

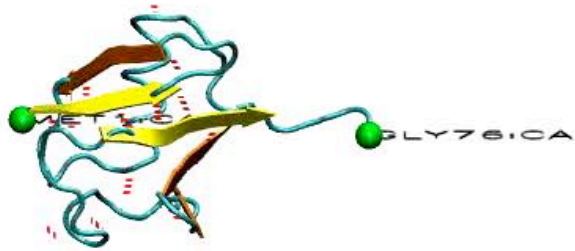
Use VMD scripting features to color beta strands separately; show hydrogen bonds to monitor the mechanical stability of ubiquitin



**Ubiquitin stretched between the C terminus and K48 does not fully extend!**



# Discovering the Mechanical Properties of Ubiquitin

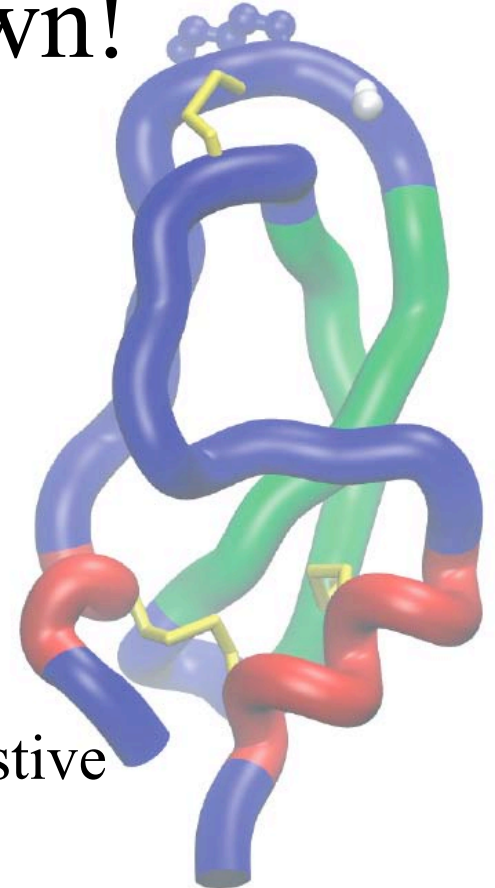


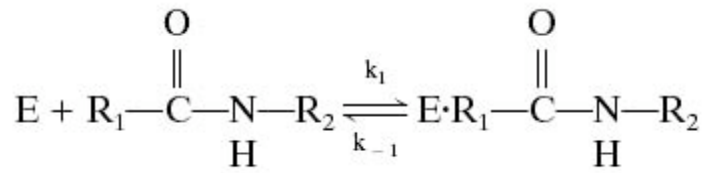
**Ubiquitin stretched between the C and the N termini extends fully!**

# Discover BPTI on your own!

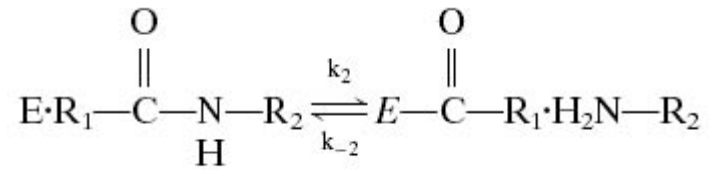
*bovine pancreatic trypsin inhibitor*

- small (58 amino acids)
- rigid
- binds as an **inhibitor** to Trypsin  
(a serine proteolytic enzyme, that appears in digestive system of mammals.)
- blocks its active site.

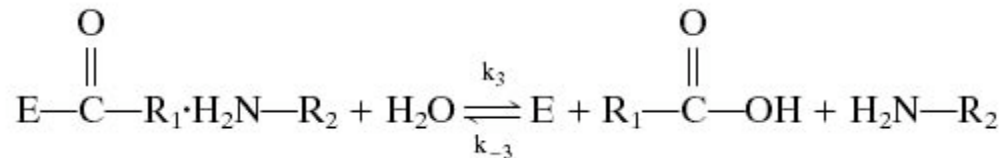




(Michaelis complex)



(Acyl-enzyme with leaving peptide  
noncovalently bound)



(Free enzyme and products)

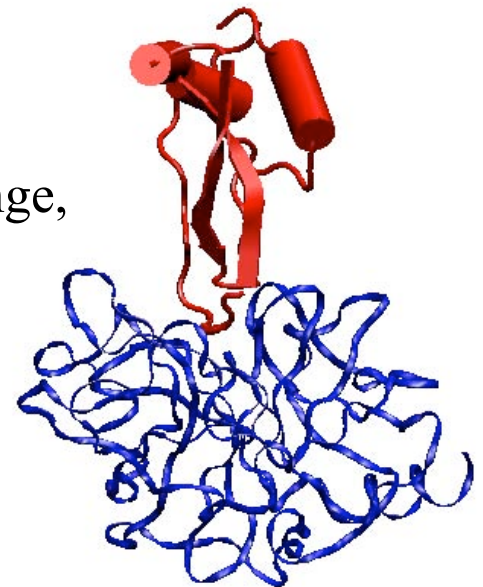
Mechanism of cleavage of peptides with serine proteases.

Radisky E. and Koshland D. Jr., Proc. Natl. Acad. Sci., USA, 99, 10316-10321

**Trypsin:** A proteolytic enzyme that hydrolyzes peptide bonds on the carboxyl side of **Arg** or **Lys**.

## BPTI: A “standard mechanism” inhibitor

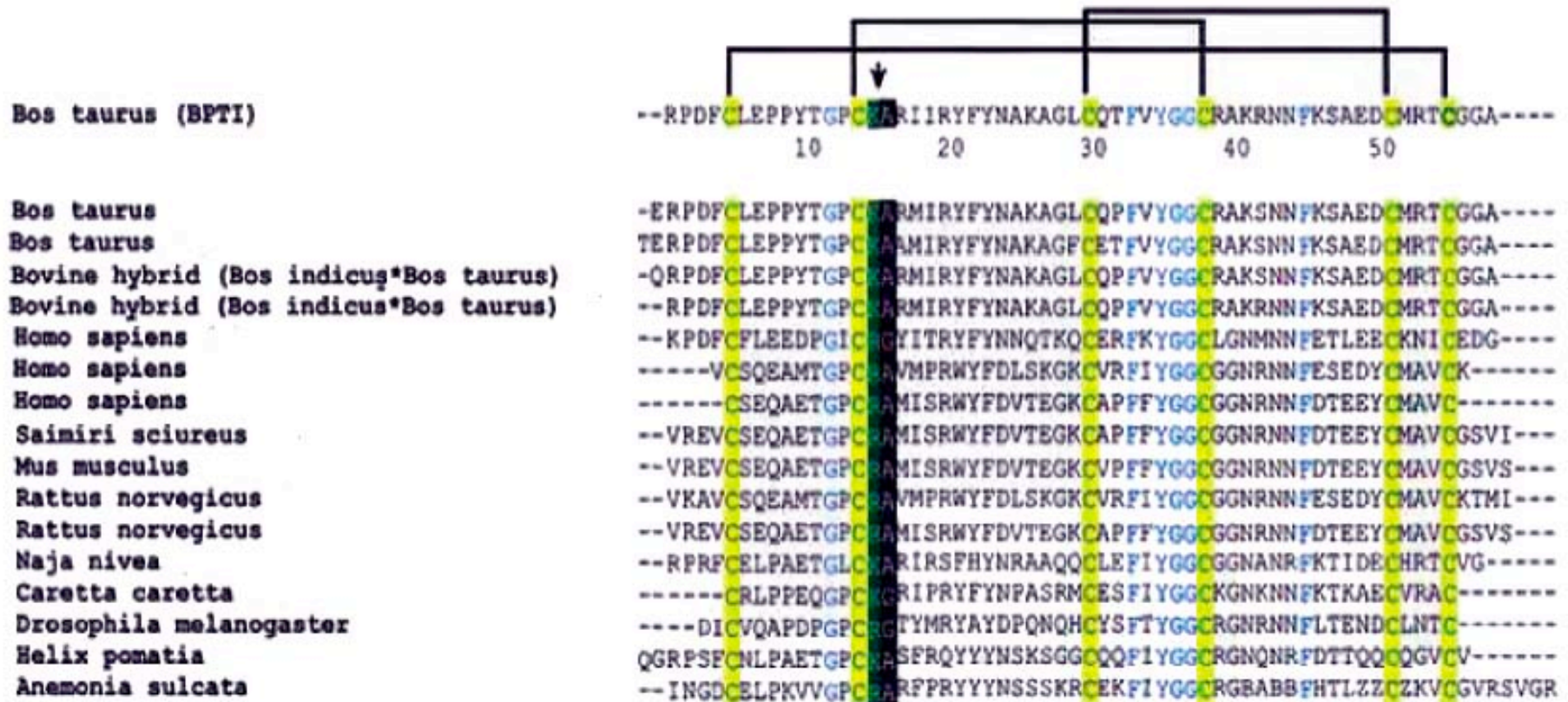
- Binds to Trypsin as a substrate.  
forms an acyl-enzyme intermediate rapidly.
- Very little **structural changes** in trypsin or BPTI.  
several H-bonds between backbone of the two proteins change,  
little reduction in conformational entropy → binds tightly
- Remains uncleaved.  
hydrolysis is  $10^{11}$  times slower than for other substrates  
  
Structures of the **protease binding region**, in the proteins of all 18 families of standard mechanism inhibitors are similar.



# Why does Trypsin cleave BPTI so slowly?

- Disruption of the non-covalent bonds in the **tightly bonded** enzyme-inhibitor complex increases the energy of transition states for bond cleavage.
- Water molecules do not have access to the active site, because of the **tight binding** of Trypsin and BPTI.
- After the cleavage of the active-site peptide bond, the newly formed termini **are held in close proximity**, favoring reformation of the peptide bond.
- The **rigidity** of BPTI may also contribute by not allowing necessary atomic motions.

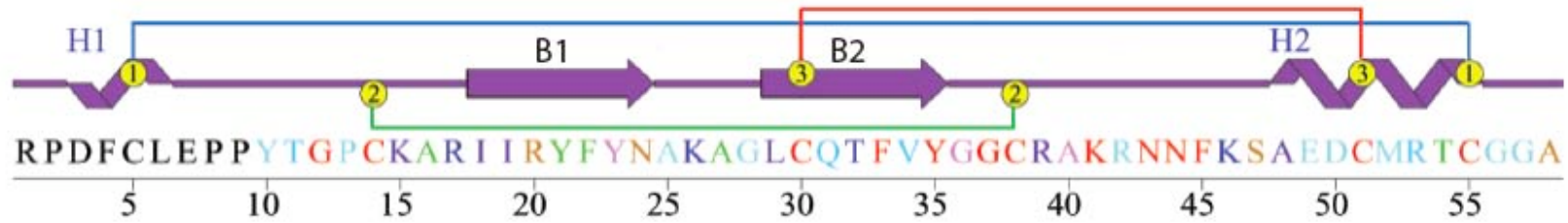
# Amino acid sequence alignment of BPTI-like proteins



Reactive conserved P and P' residues are highlighted in black and indicated by the arrow. Six conserved cysteine residues are highlighted in yellow. Three disulfide bonds formed by the cysteines are indicated by black lines. Other residues that are conserved in all proteins are labeled in blue.

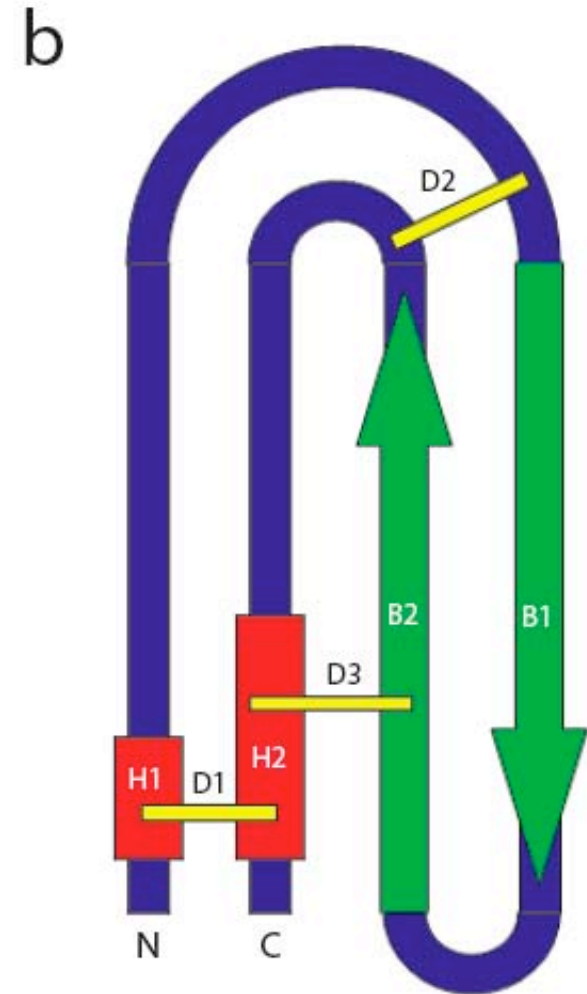
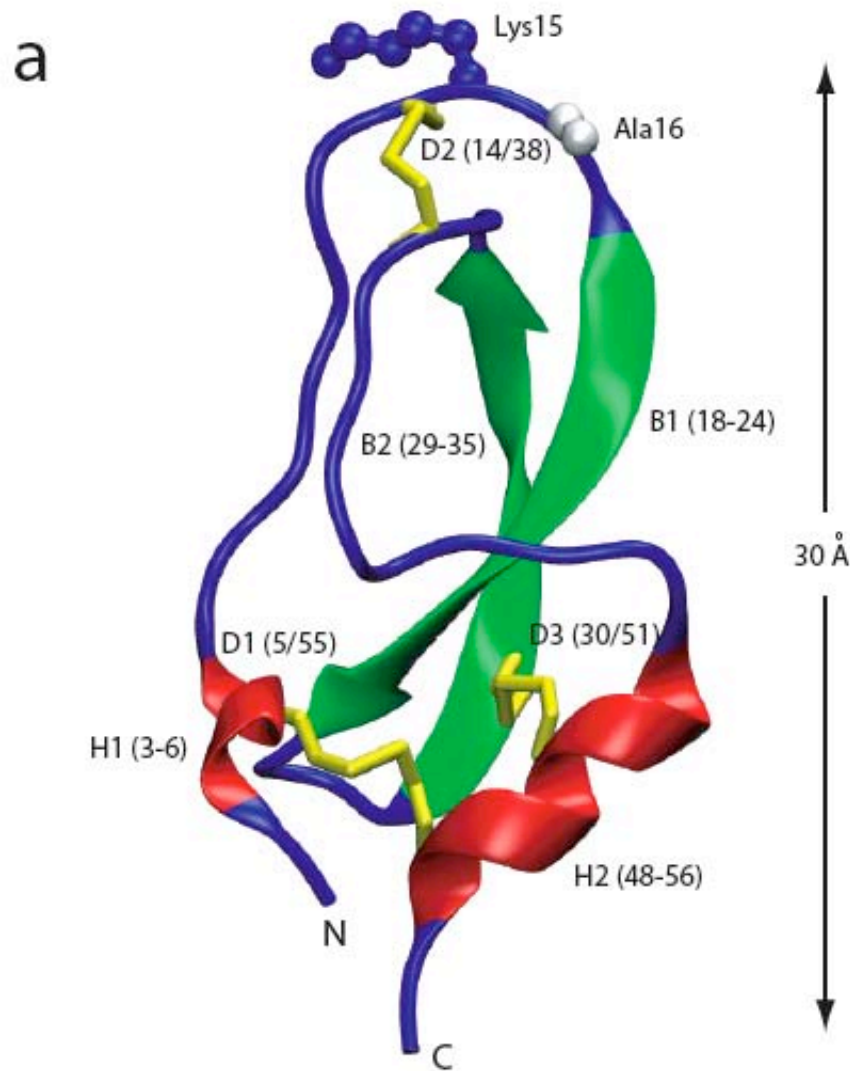


# BPTI secondary structure

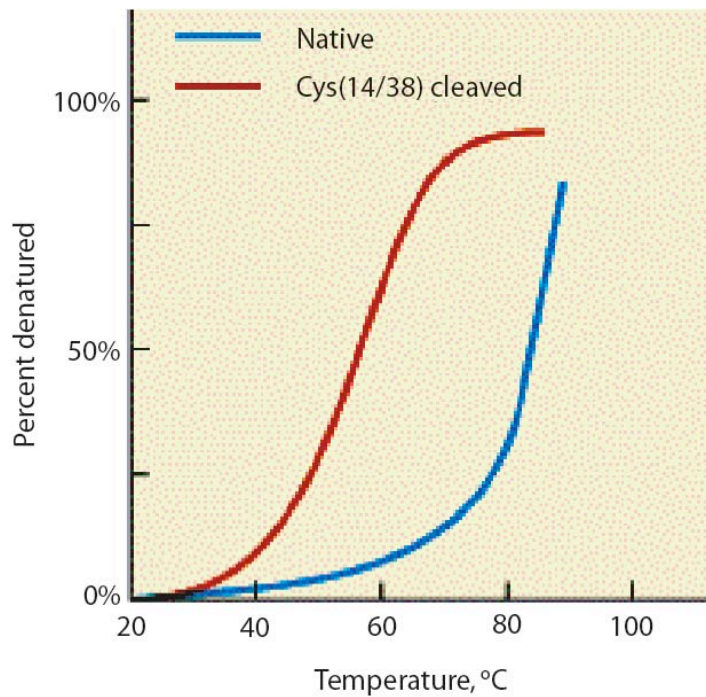


BPTI secondary structure: Conservation is indicated by color using rainbow scale coloring (Blue to red= low to high)

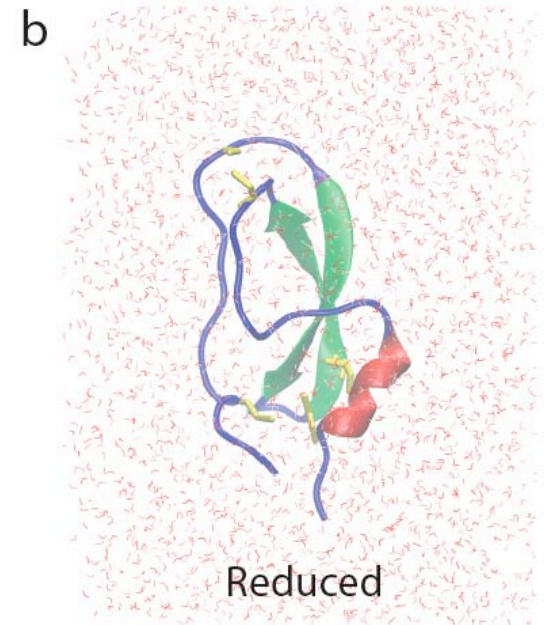
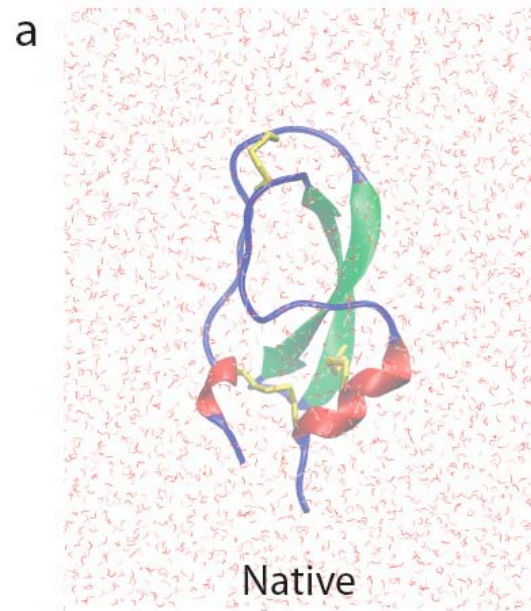
# BPTI Tertiary Structure



# Stability of native and reduced BPTI



Experiment



Can be tested through simulation

# BPTI case study

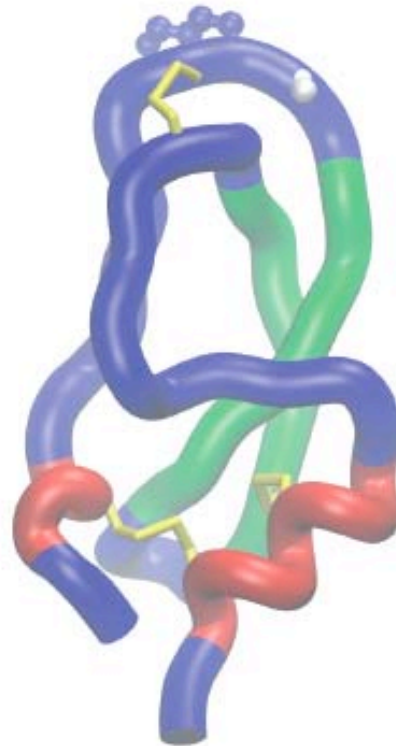
Chalermpol Kanchanawarin

Department of Physics and Beckman Institute,

University of Illinois at Urbana-Champaign

Urbana, IL 61801, USA

Date: Tuesday 11<sup>th</sup> January 2005



# **Inspect BPTI with VMD**

# VMD supports multiple sequence analysis

The screenshot displays the VMD 1.8.3a2 OpenGL Display interface. The main window shows a 3D protein structure with various domains colored in blue, green, yellow, and orange. An 'Extensions' menu is open, listing several plugins: sequence, autoimd, apbsrun, imd, contactmap, pdbtool, ramaplot, rmsd, solvate, timeline, multiseq, tkcon, and vmdmovie. A 'treeWindow' is open, showing a phylogenetic tree with five entries: d1efwa3.ent (Thermus thermophilus B), d1c0aa3.ent (Escherichia coli B), d1n9wb1.ent (d1n9wb1.ent), d1asza2.ent (Saccharomyces cerevisiae E), and d1b8aa2.ent (Pyrococcus kodakaraensis A). A scale bar indicates a distance of 0.56. Below the tree is a 'Sequence Display' window showing a multiple sequence alignment of the protein sequences. The alignment is displayed on a cyan background with yellow highlights on specific residues.

**Extensions**

- sequence
- autoimd
- apbsrun
- imd
- contactmap
- pdbtool
- ramaplot
- rmsd
- solvate
- timeline
- multiseq
- tkcon
- vmdmovie

**treeWindow**

Tree

- d1efwa3.ent Thermus thermophilus B
- d1c0aa3.ent Escherichia coli B
- d1n9wb1.ent d1n9wb1.ent
- d1asza2.ent Saccharomyces cerevisiae E
- d1b8aa2.ent Pyrococcus kodakaraensis A

0.56

**Sequence Display**

```
d1b8aa2.ent  IDTEGERLLGKYM--MENENAPLYFLYQYPS-----EAKPFYIMKYDN-----K--PEICRAFDLEYRGV
d1asza2.ent  LSTENEKFLGKLV--RDKYDTDFYILDKFP-----EIRPFYTMPDPA-----N--PKYSNSYDFMRGE
d1n9wb1.ent  LSEEAERLLGEYA--KERWGSDFVTRYP-----SVRPFYTYP--EE-----DGTRSFDLLFRGL
d1c0aa3.ent  ---GSD-KP-DLRDE---SKWAPLWVIDFPMFE--DDGEGGLTAMHHPPTSPPK--DMTAAELKAAPENAVANAYDMVINGY
d1efwa3.ent  ---GSD-KP-DL-RR---EGFRFLWVDFPFLLEWDEEEEAWTYMHHPPTSPPHED--LPLLEKDPGRVRALAYDLVINGV
```

Version 1.8.3  
(Zaida Luthey-Schulten  
+ coworkers)



