A glimpse of computational methods in biological physics:

Case study on three proteins and a cellular organelle

Klaus Schulten

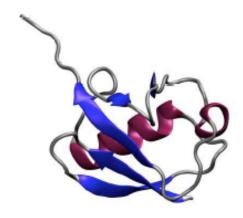
Theoretical and Computational Biophysics Group September 16, 2014

Administrative issues

- Assignment from this lecture: read the case study on ubiquitin and choose Three out of the six exercises to complete: http://www.ks.uiuc.edu/Training/CaseStudies/ index.html#ubqcs
- You will have to download and install the software VMD: http://www.ks.uiuc.edu/ Research/vmd/
- For any questions and assistance on the assignment, email Marco Tjioe at tjioe2@gmail.com
- Assignments should be given to Professor Selvin before class on October 1 (next Tuesday).

Case Study: Ubiquitin

Eduardo Cruz-Chu and JC Gumbart



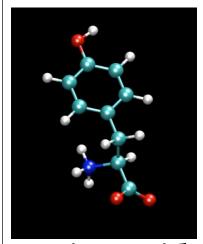
1 Introduction

Without a doubt, the most organized and coordinated machine known is the biological cell. Inside its micrometer-scale diameter, a wide variety of macromolecules (DNA, proteins, sugars, lipids, etc.) work together in a cooperative way, balancing energy and matter to keep the cell alive. Within the cell, proteins are the overachievers. They allow the movement of water and ions through the cell membrane, help ATP to store energy, assist DNA during replication, recognize foreign infections, and more. However, all of these functions don't work independently of each other. To maintain harmony and efficiency between various functions, most processes have to be turned on or off according to different cellular stages and changes within the environment.

To this end, together with the mechanisms to assemble functional proteins and to turn on their functions, there should be counterparts to suppress and disassemble proteins when they are no longer needed. The cellular machine depends on assembly and disassembly to regulate the effective concentration of proteins and their corresponding activities [1]. Furthermore, defective

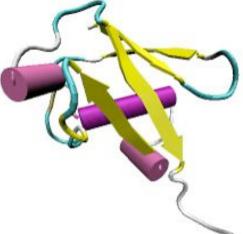
Introduction to Protein Structures and Molecular Graphics Tool

10 20 30 40 50 60 MQIFVKTLTG KTITLEVEPS DTIENVKAKI QDKEGIPPDQ QRLIFAGKQL EDGRTLSDYN 70 IQKESTLHLV LRLRGG

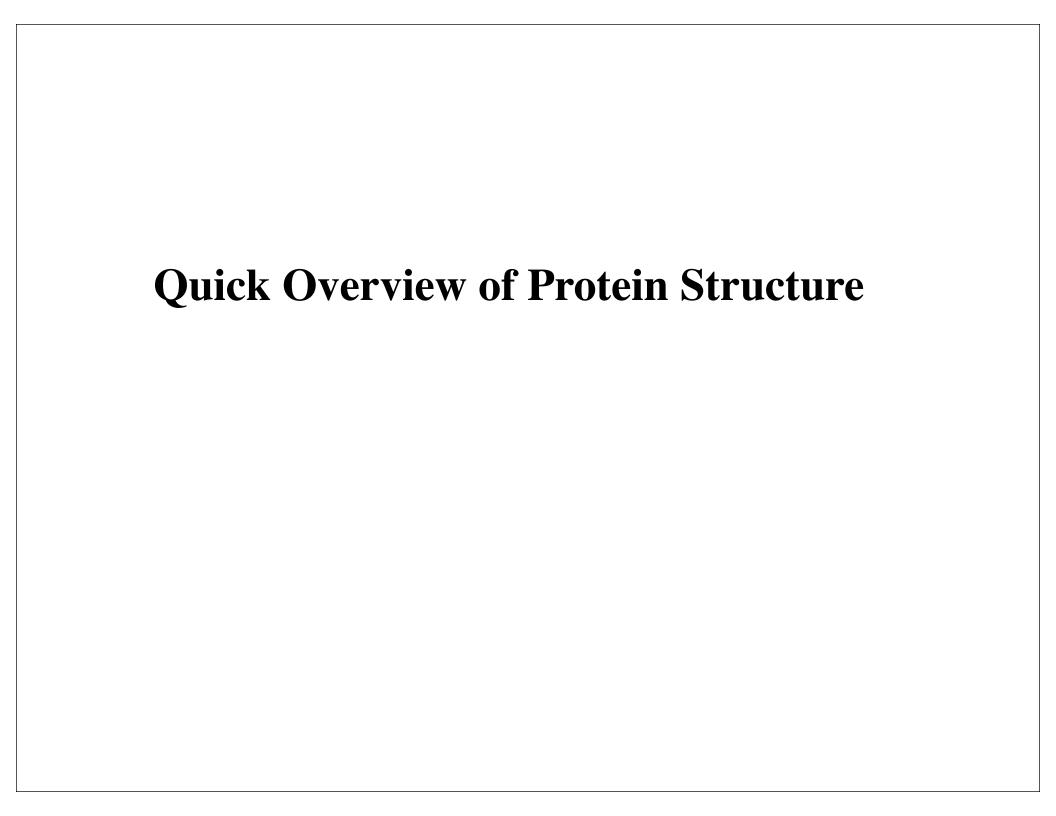


amino acid tyrosine, Y

Amino Acid Isoleucine	SLC	DNA codons					
		ATT,	ATC,	ATA			
Leucine	L	CTT,	CTC,	CTA,	CTG,	TTA,	TTG
Valine	v	GTT,	GTC,	GTA,	GTG		
Phenylalanine	F	TTT,	TTC				
Methionine	М	ATG					
Cysteine	С	TGT,	TGC				
Alanine	A	GCT,	GCC,	GCA,	GCG		
Glycine	G	GGT,	GGC,	GGA,	GGG		
Proline	P	CCT,	ccc,	CCA,	CCG		
Threonine	T	ACT,	ACC,	ACA,	ACG		
Serine	S	TCT,	TCC,	TCA,	TCG,	AGT,	AGC
Tyrosine	Y	TAT,	TAC	211			
Tryptophan	W	TGG					
Glutamine	Q	CAA,	CAG				
Asparagine	N	AAT,	AAC				
Histidine	H	CAT,	CAC				
Glutamic acid	E	GAA,	GAG				
Aspartic acid	D	GAT,	GAC				
Lysine	K	λλλ,	AAG				
Arginine	R	CGT,	CGC,	CGA,	CGG,	AGA,	AGG
Stop codons	Stop	TAA,	TAG,	TGA			



Ubiquitin

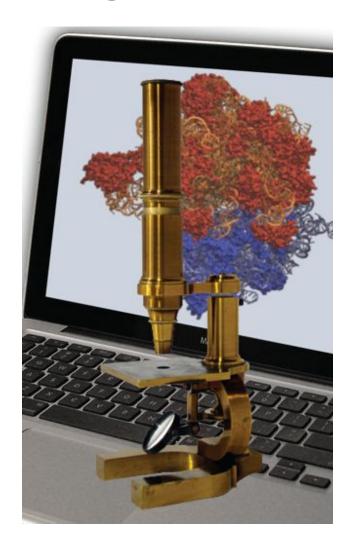


What Proteins are Made of: Primary Structure (Sequence) of Amino Acids

Proteins: polymeric molecules linking amino acids through peptide bonds

Peptide bond linking two amino acids

Looking at Proteins Through the Program VMD

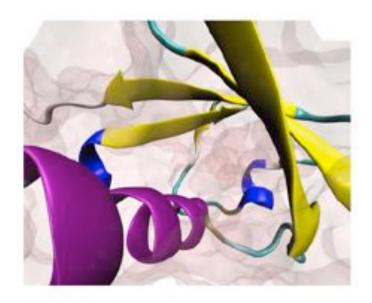


University of Illinois at Urbana-Champaign Beckman Institute for Advanced Science and Technology Theoretical and Computational Biophysics Group Computational Biophysics Workshop

Learn to use VMD from the "Using VMD" tutorial available at http://www.ks.uiuc.edu/Training/Tutorials/

Using VMD

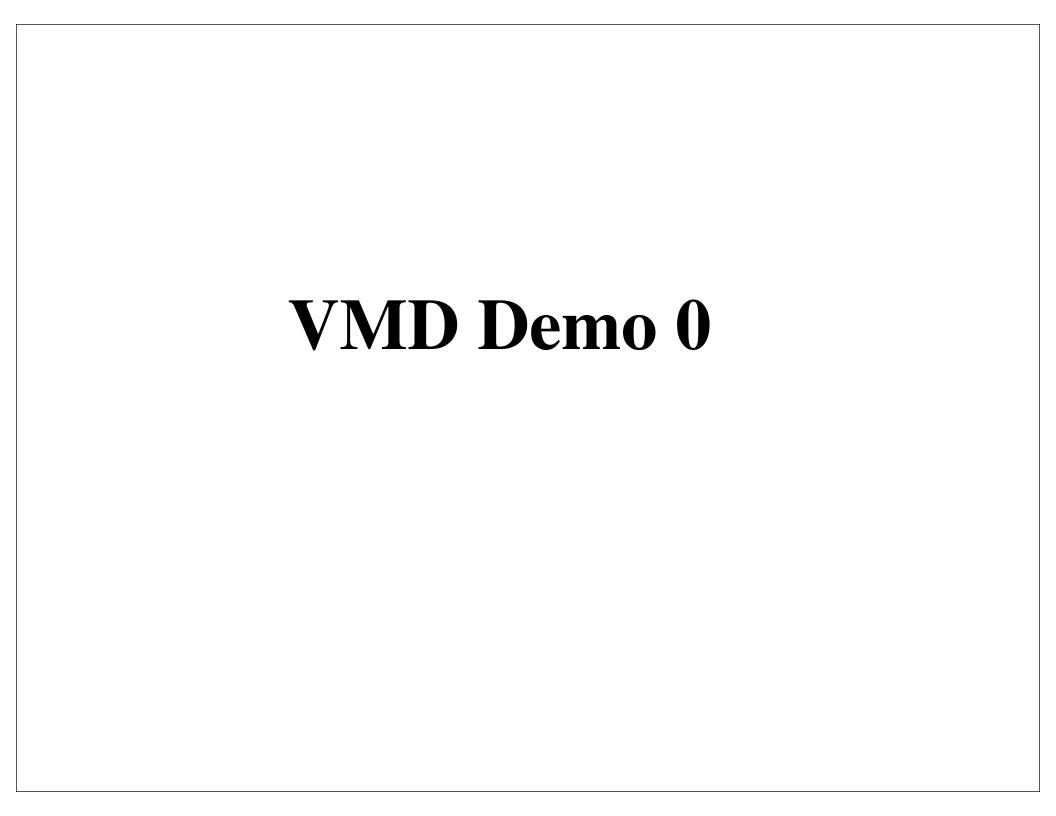
VMD for Mac OS X, Unix, and Windows is available for download at http://www.ks.uiuc.edu/Research/vmd/

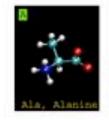


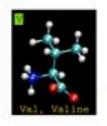
VMD Developer: John Stone

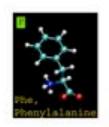
Tutorial Contributors:

Alek Aksimentiev, Anton Arkhipov, Robert Brunner, Jordi Cohen, Brijeet Dhaliwal, John Eargle, Jen Hsin, Fatemeh Khalili, Eric H. Lee, Zan Luthey-Schulten, Patrick O'Donoghue, Elijah Roberts, Anurag Sethi, Marcos Sotomayor, Emad Tajkhorshid, Leonardo Trabuco, Elizabeth Villa, Yi Wang, David Wells, Dan Wright, Ying Yin

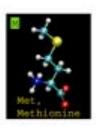






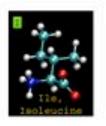


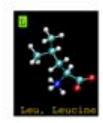




Protein Primary Structure

non-polar

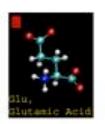




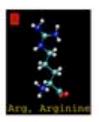
The twenty amino acids

charged









Protein Structure and

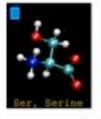
Function, Gregory Petsko and Dagmar Ringe, 2004

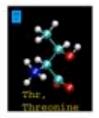
Molecular Biology of The Cell

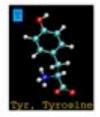
Alberts, Johnson, Lewis, Raff, Roberts, Walter, 2008, 3rd Ed.

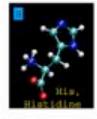
Introduction to Protein Structure, 2nd ed. Carl Branden & John Tooze, 1999

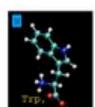
polar

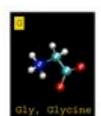




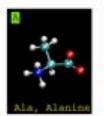


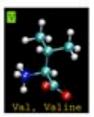


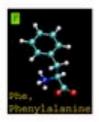


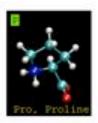


URL: http://lectures.molgen.mpg.de/ProteinStructure



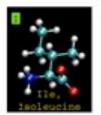


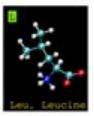


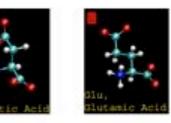


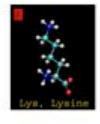


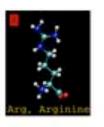
Alanine



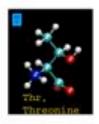


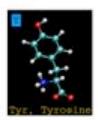


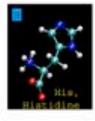


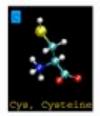


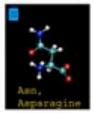








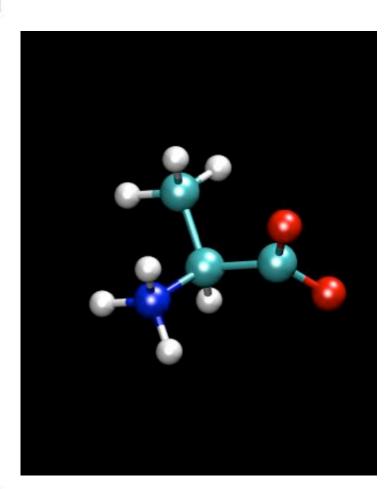


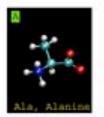




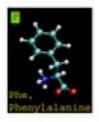


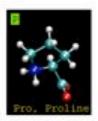


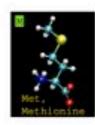




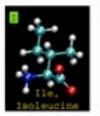




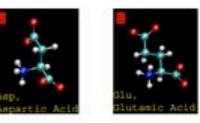




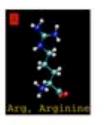
Valine

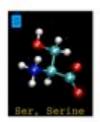








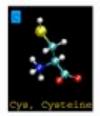


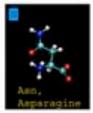




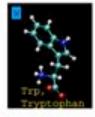




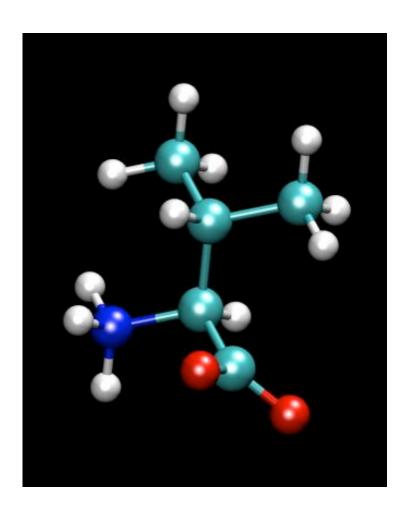


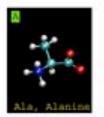


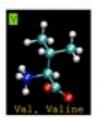




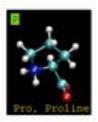


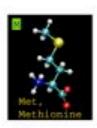




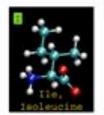


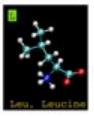


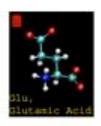


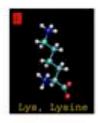


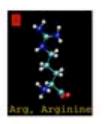
Phenylanaline

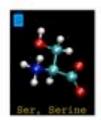


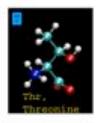






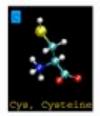


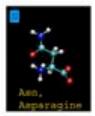




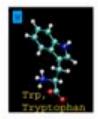


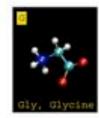


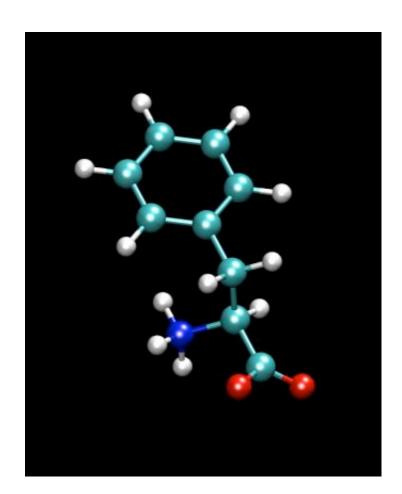


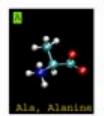


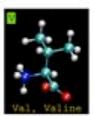




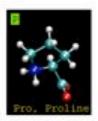








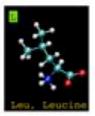


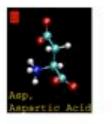


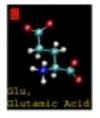


Proline

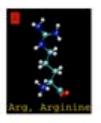


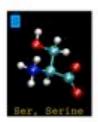






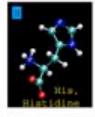


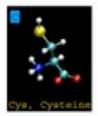


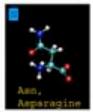




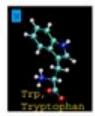




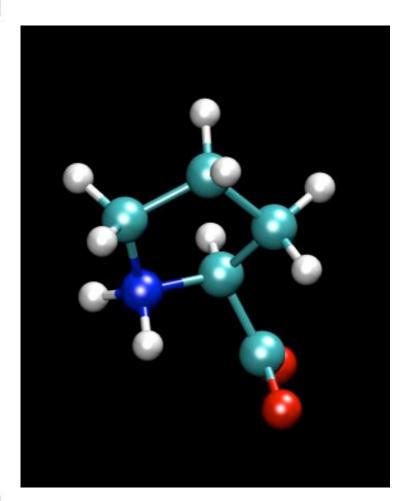


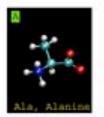


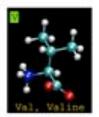


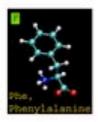


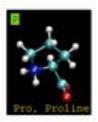


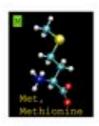








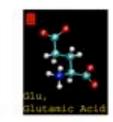


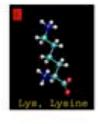


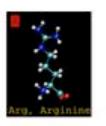
Methionine



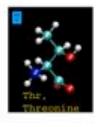






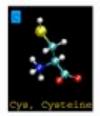






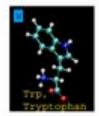


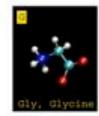


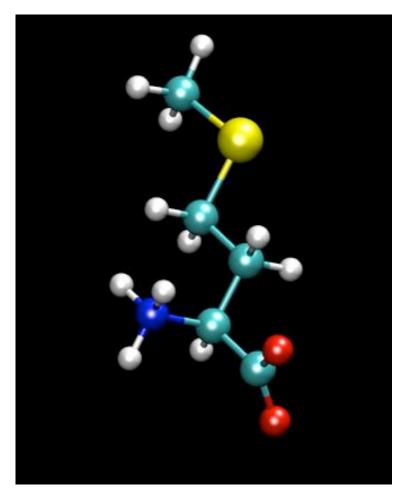


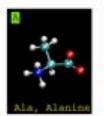




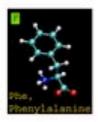


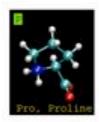








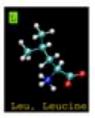


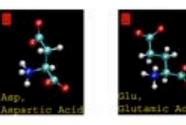


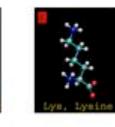


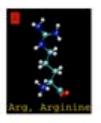
Isoleucine

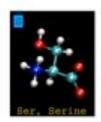




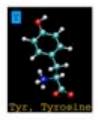




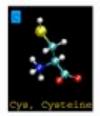


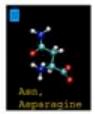




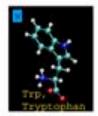




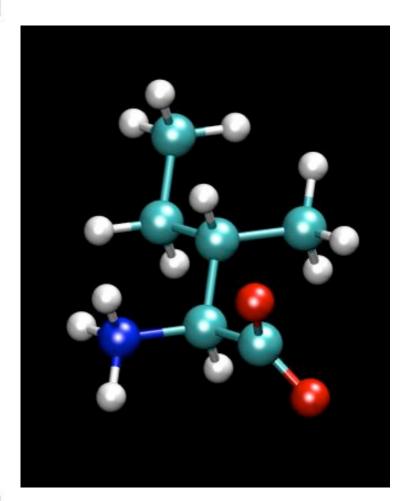


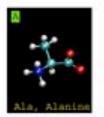




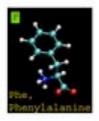


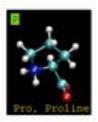






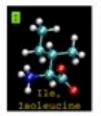




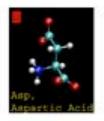


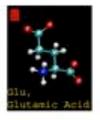


Leucine

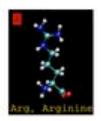


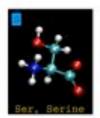








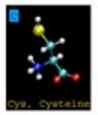






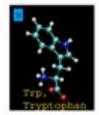


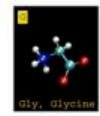


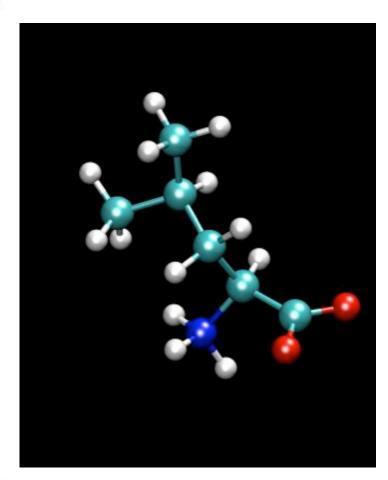


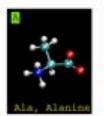




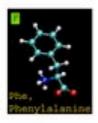


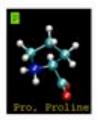


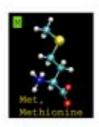




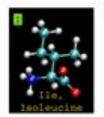


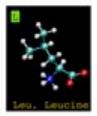


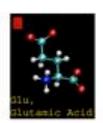




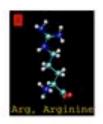
Aspartate

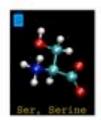


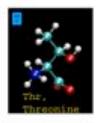




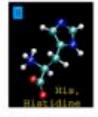


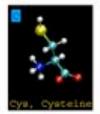


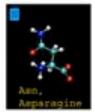




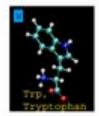


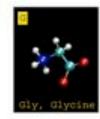


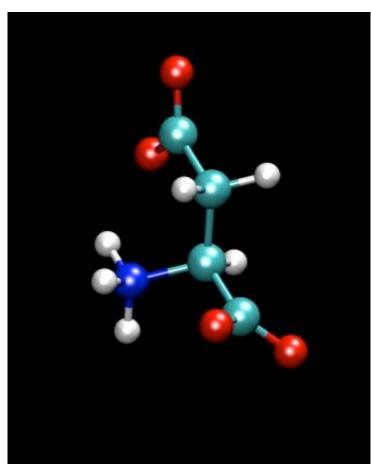


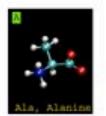




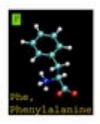


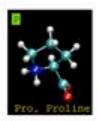






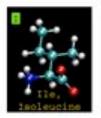


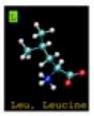


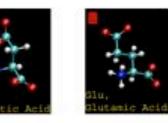




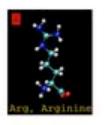
Glutamate

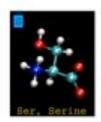








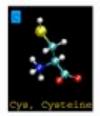


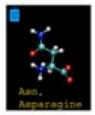




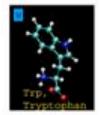




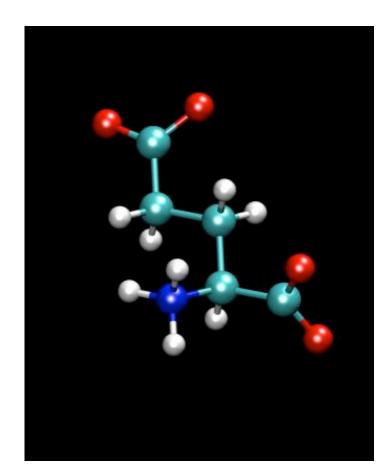


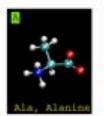




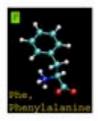




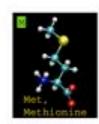




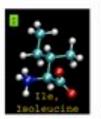


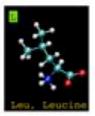


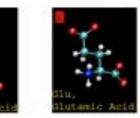




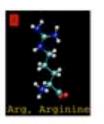
Lysine

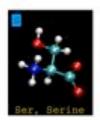










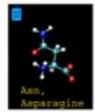




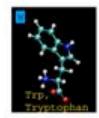


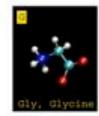


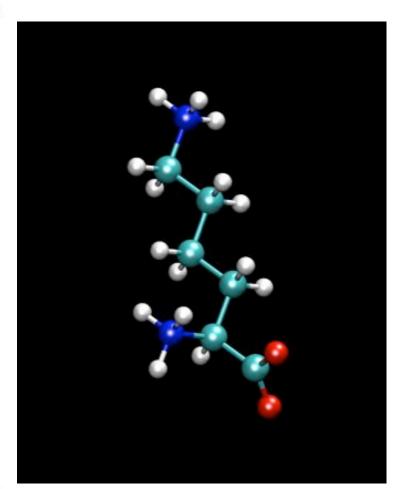


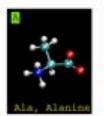


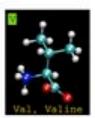


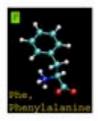


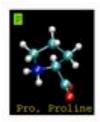






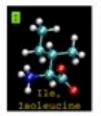




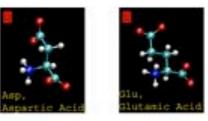


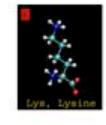


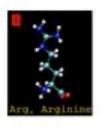
Arginine

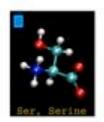


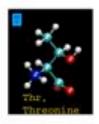






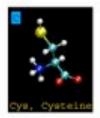


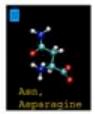




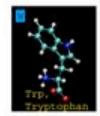


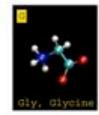


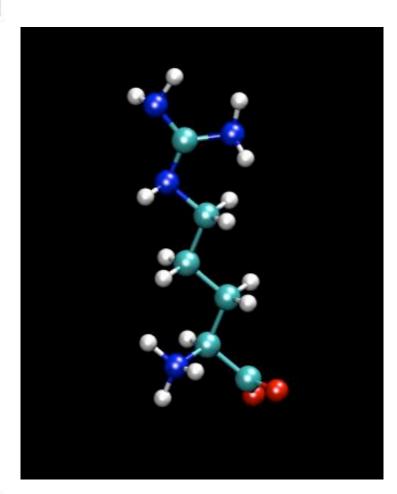


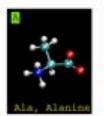


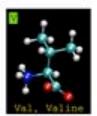


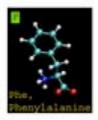


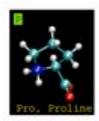


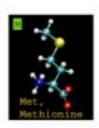




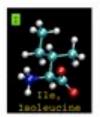


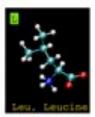


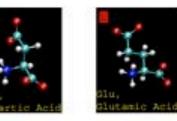


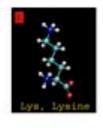


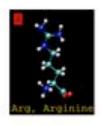
Serine

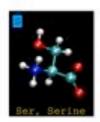








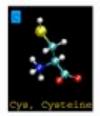


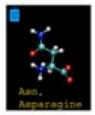




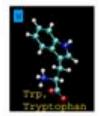


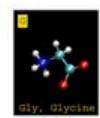


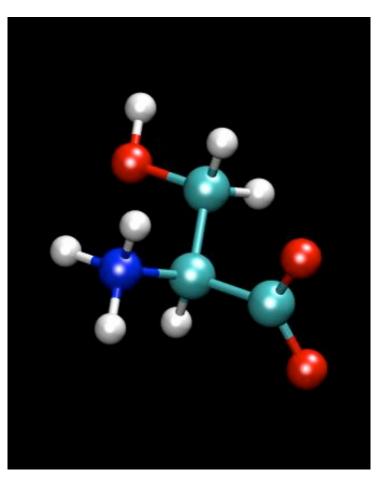


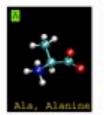


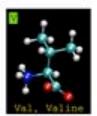


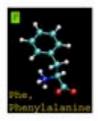


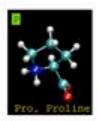


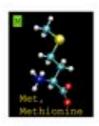




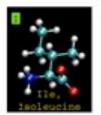


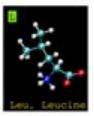


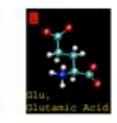




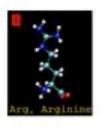
Threonine

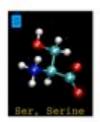


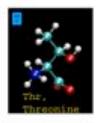




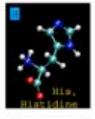




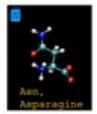




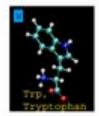




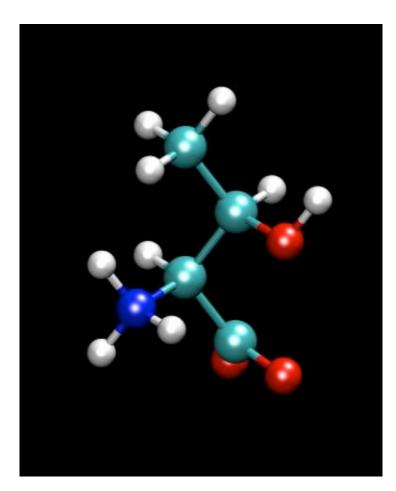


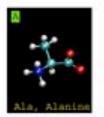


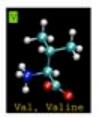


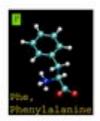


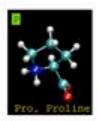


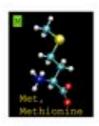




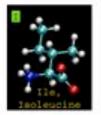


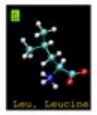


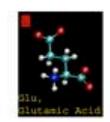




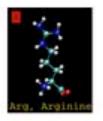
Tyrosine

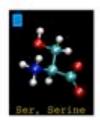






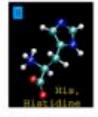


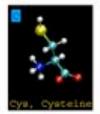


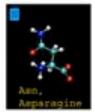








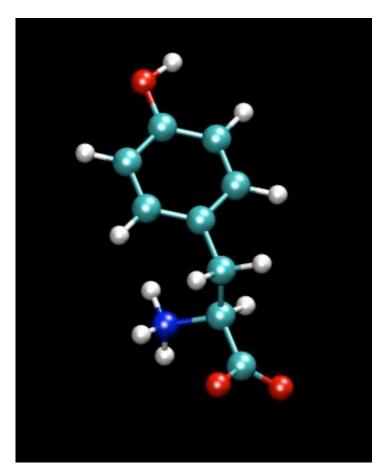


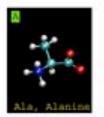


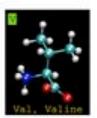


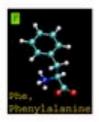


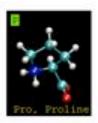


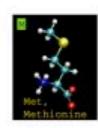




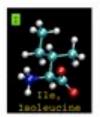


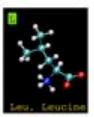


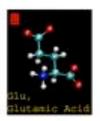




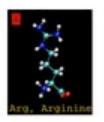
Histidine

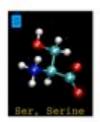








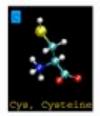


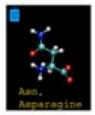




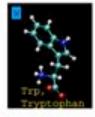


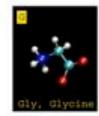


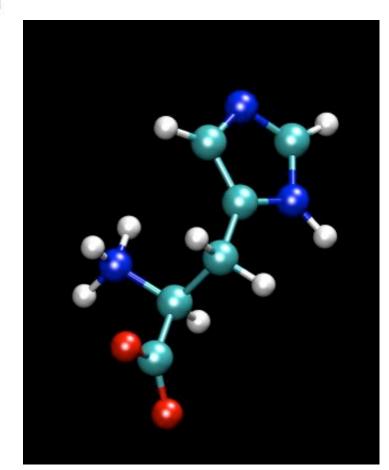


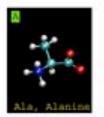


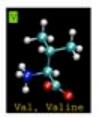


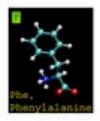


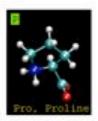


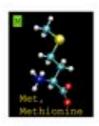




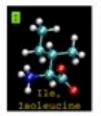


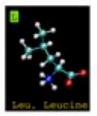


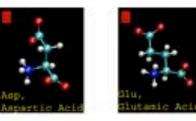


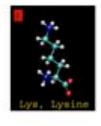


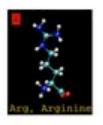
Cysteine

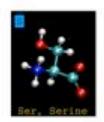




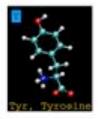




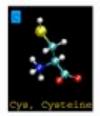


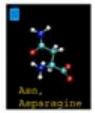




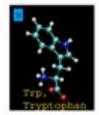


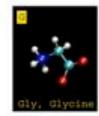


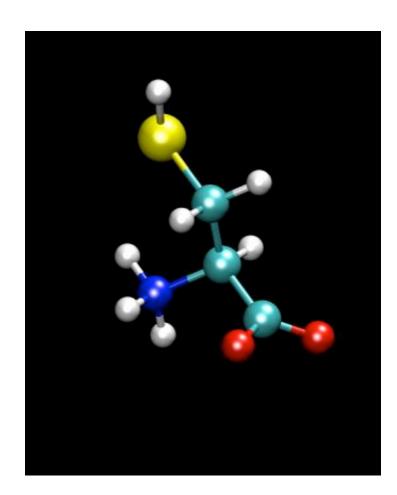


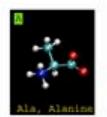


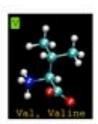




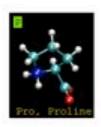


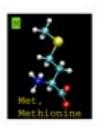




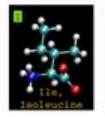


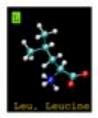


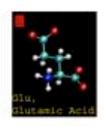




Asparagine

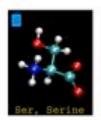








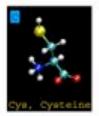


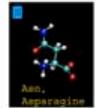


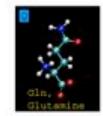


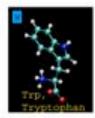




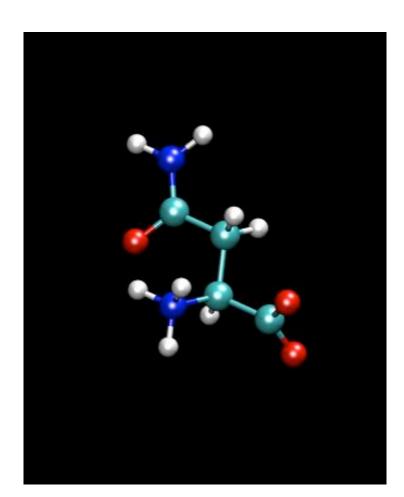


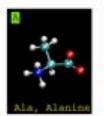






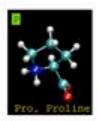


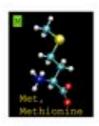




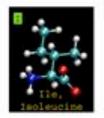


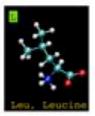






Glutamine

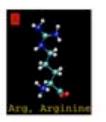


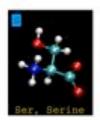


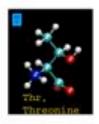


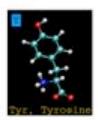


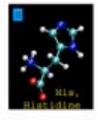


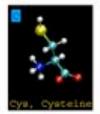


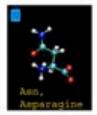




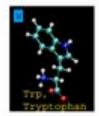


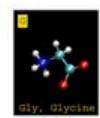


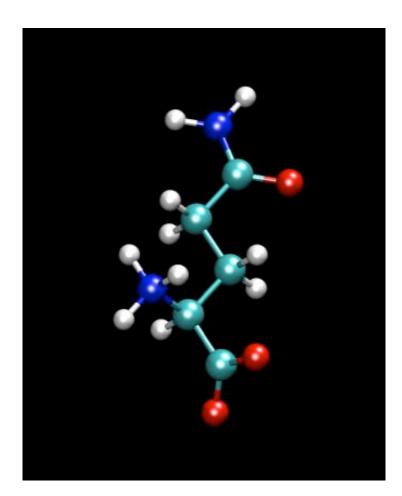


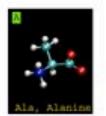


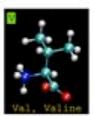




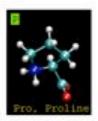


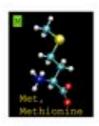




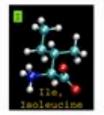


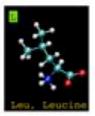


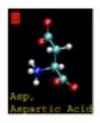


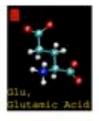


Tryptophane

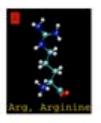


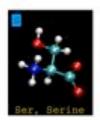


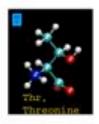




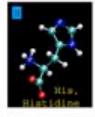








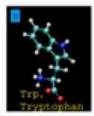




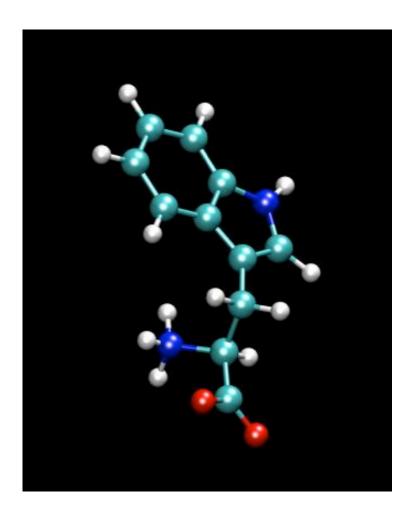


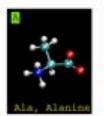




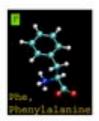


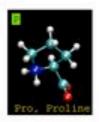


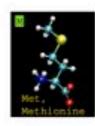




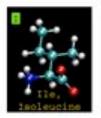


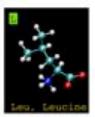


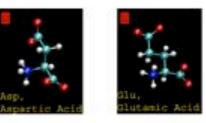




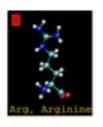
Glycine

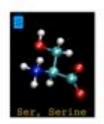








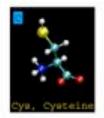


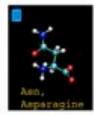




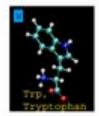


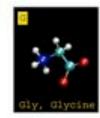


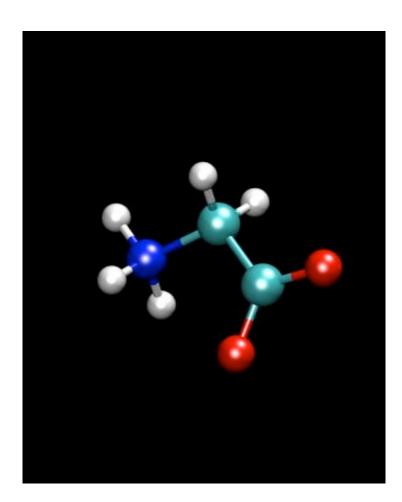




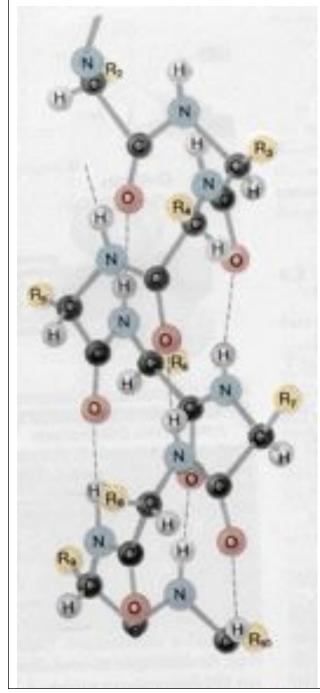






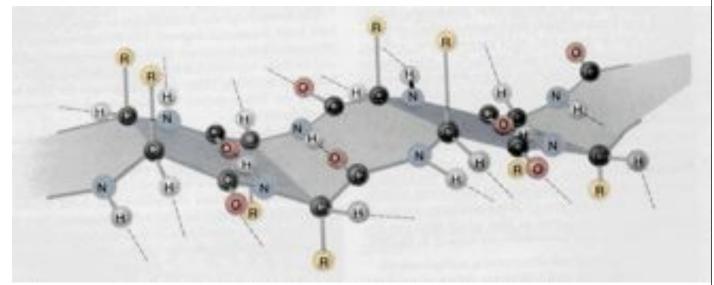


Protein Secondary Structure



An antiparallel beta sheet

Beta sheets are created, when atoms of beta strands are hydrogen bound. Beta-sheets may consist of parallel strands, antiparallel strands or out of a mixture of parallel and antiparallel strands.



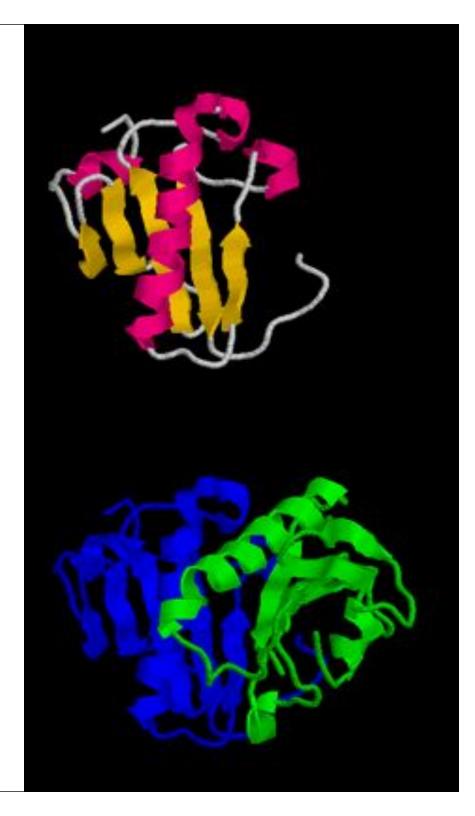
An alpha helix

The backbone is formed as a helix. An ideal alpha helix consists of 3.6 residues per complete turn. The side chains stick out. There are hydrogen bonds between the carboxy group of amino acid and the amino group of another amino acid n+4. The mean phi angle is -62 degrees and the mean psi angle is -41 degrees

Tertiary and Quarternary Structures of Proteins

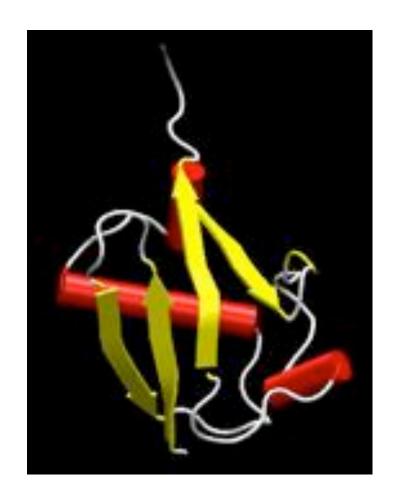
Tertiary structure describes the packing of alpha-helices, beta-sheets and random coils with respect to each other on the level of one whole polypeptide chain.

Quaternary structure only exists, if there is more than one polypeptide chain present in a complex protein. Then quaternary structure describes the spatial organization of the chains

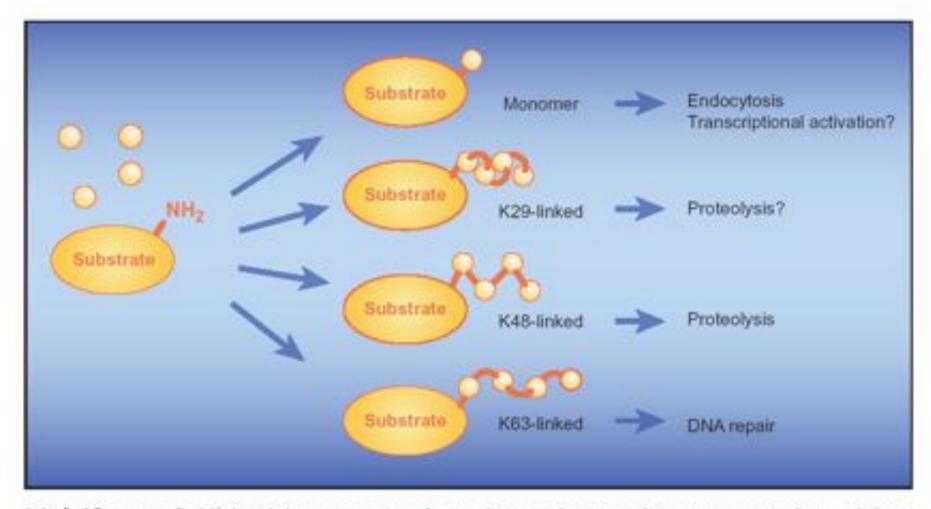


Focus on one protein Ubiquitin

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



Mono-ubiquitylation versus multi-ubiquitylation

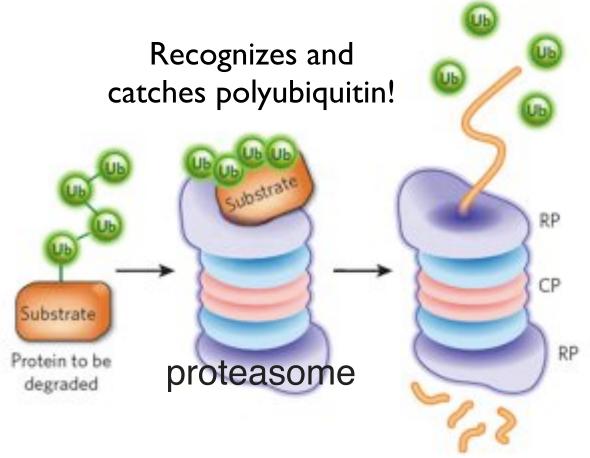


Multifaceted. Ubiquitin can attach to its various substrate proteins, either singly or in chains, and that in turn might determine what effect the ubiquitination has. (K29, K48, and K63 refer to the particular lysine amino acid used to link the ubiquitins to each other.)

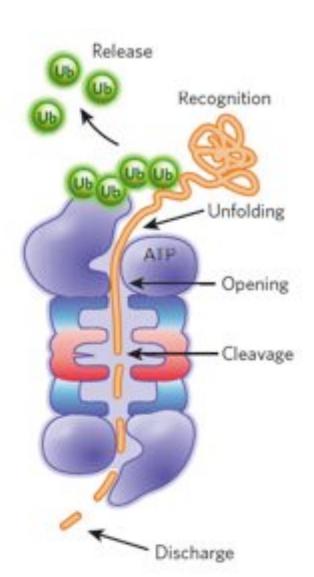
Marx, J., Ubiquitin lives up its name, Science 297, 1792-1794 (2002)

Ubiquitin's role in protein degradation

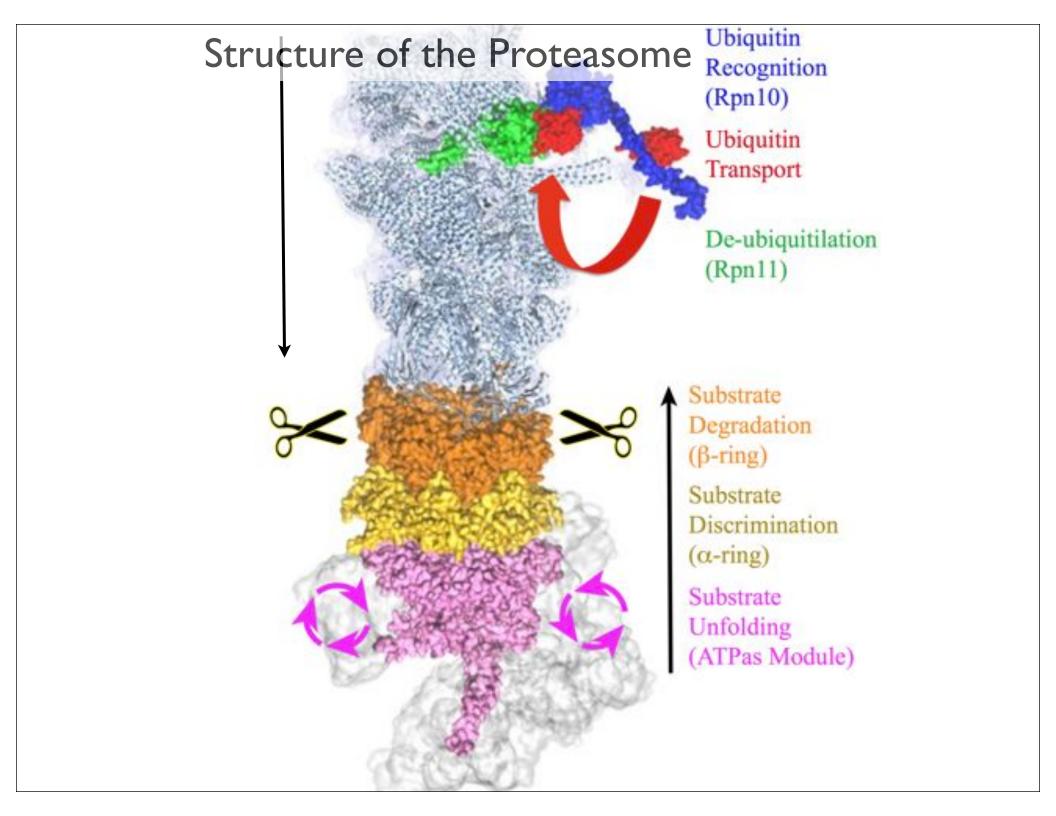
The substrate-polyubiquitin complex is then sent to proteasome



The substrate is sent through the proteasome barrel, where it is chopped up and recycled

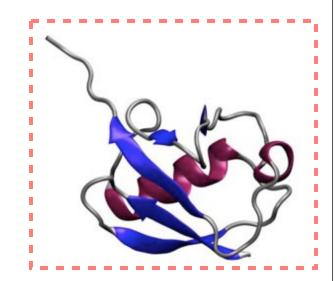


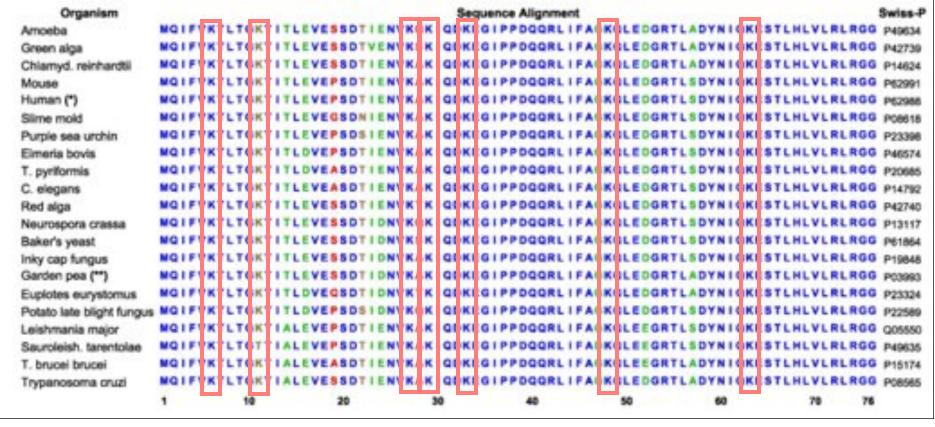
M. Hochstrasser, Nature, 458. (2009)

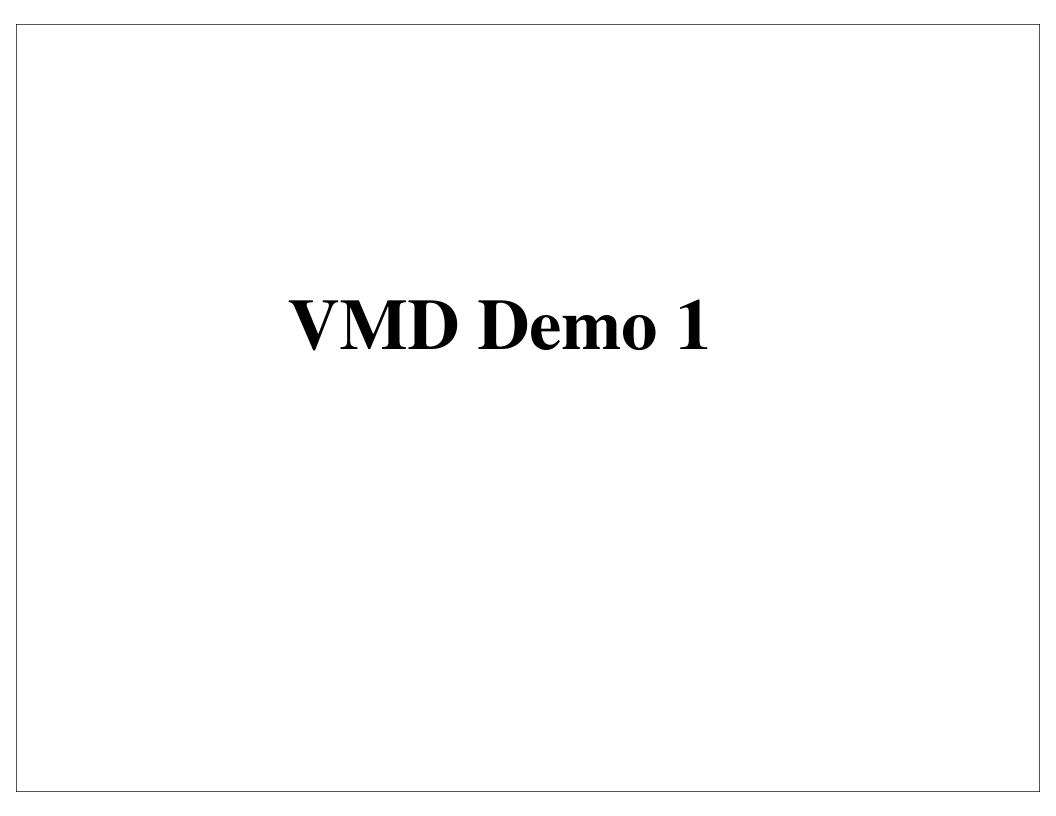


Highly conserved ubiquitin chain

- The sequence of ubiquitin is highly conserved, in particular the seven lysine residues
- A lysine residue in a ubiquitin can be linked to the Cterminus of another ubiquitin
- By using different lysine for such linkage, ubiquitin is used for different cellular purposes







Some readings

http://en.wikipedia.org/wiki/Ubiquitin

- J. Marx, "Ubiquitin lives up to its name." Science, 297. (2002)
- M. Hochstrasser, "Origin and function of ubiquitin-like proteins." *Nature*, 458. (2009)
- A. Varshavsky, "The early history of the ubiquitin field." *Protein Science*, 15. (2006)
- C. M. Pickart, "Back to the future with ubiquitin." *Cell*, 116. (2004)
- M. Carrion-Vazquez et al., "The mechanical stability of ubiquitin is linkage dependent." *Nature Structure Biology*, 10. (2003)

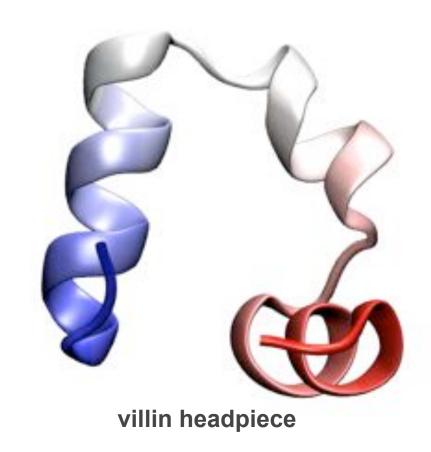
http://nobelprize.org/nobel_prizes/chemistry/laureates/2004/

If you need to get a quick lesson on VMD, here is a short tutorial:

http://www.ks.uiuc.edu/~jhsin/papers/HSIN2008.pdf

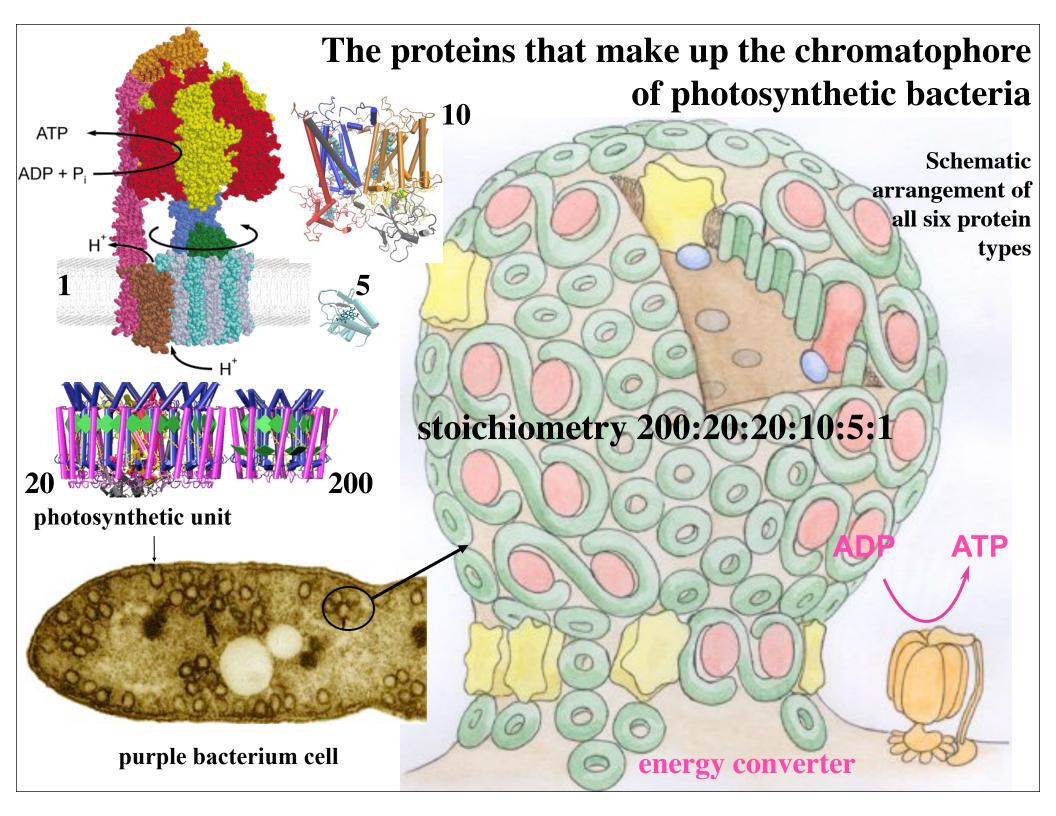
Protein Folding

- Folding of the Protein called Villin Headpiece
- First protein folded in computer simulation
- Visualization of the "trajectory" of the folding protein reveales how this protein finds it native conformation from an initially stretched-out conformation



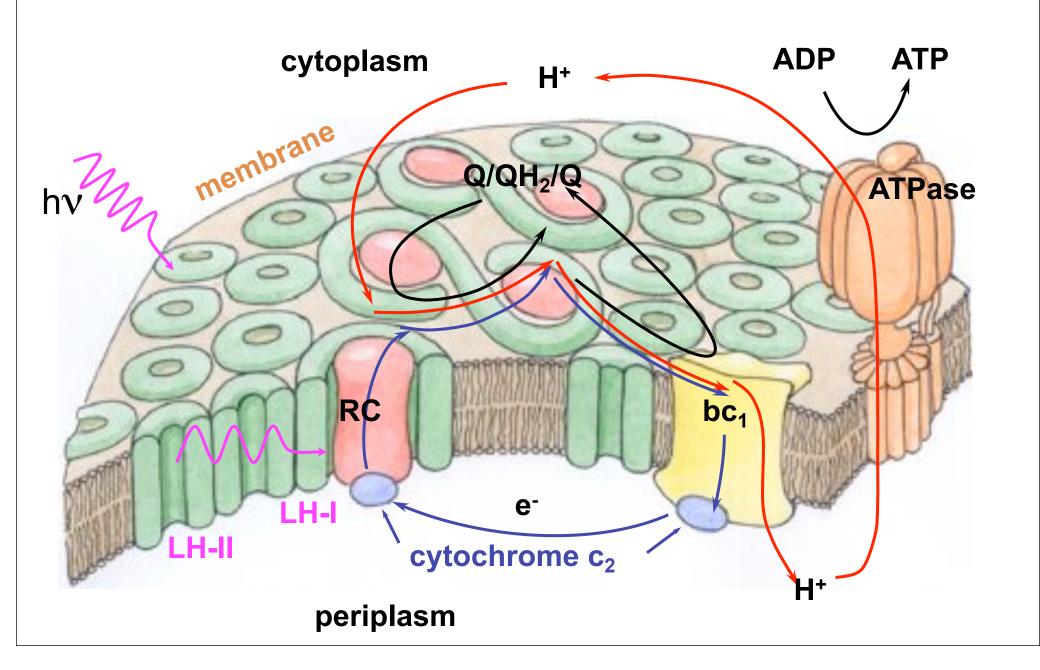
Observe folding process in unprecedented detail

VMD Demo 2

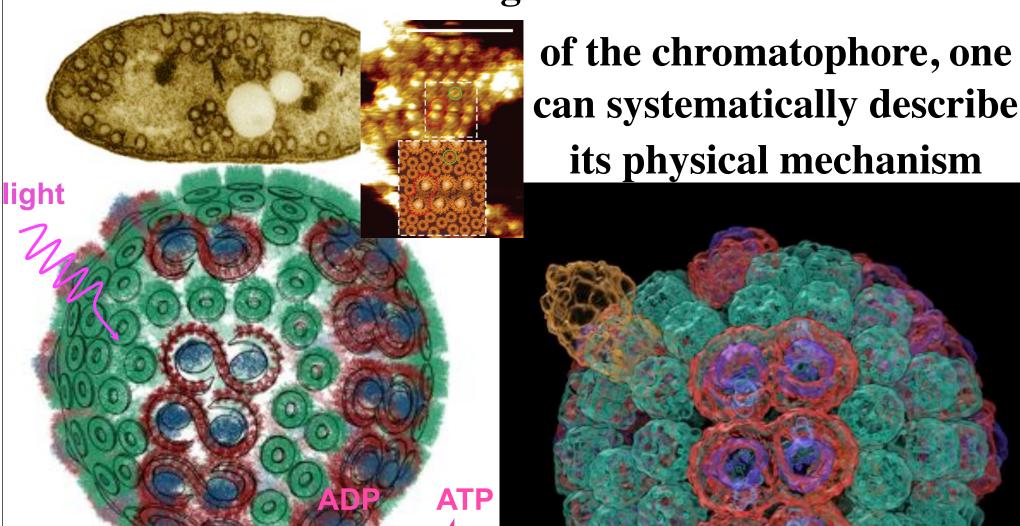


Chromatophore of Purple Bacteria

(section of the chromatophore membrane)



Knowing the Atomic Level Structure



M. Sener, J. Olsen, N. Hunter, and K. Schulten. *PNAS*, **104**: 15723-15728, 2007

