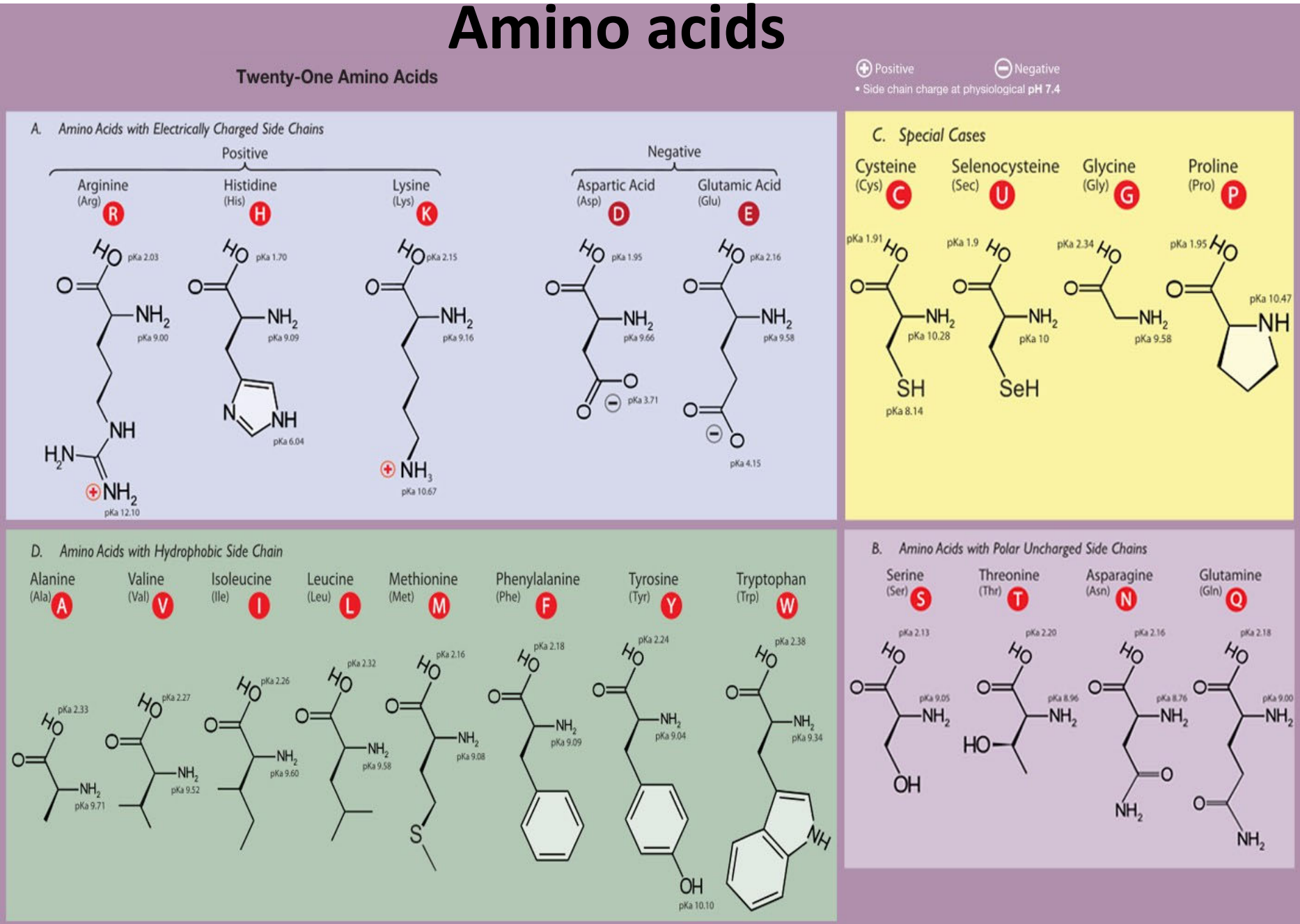
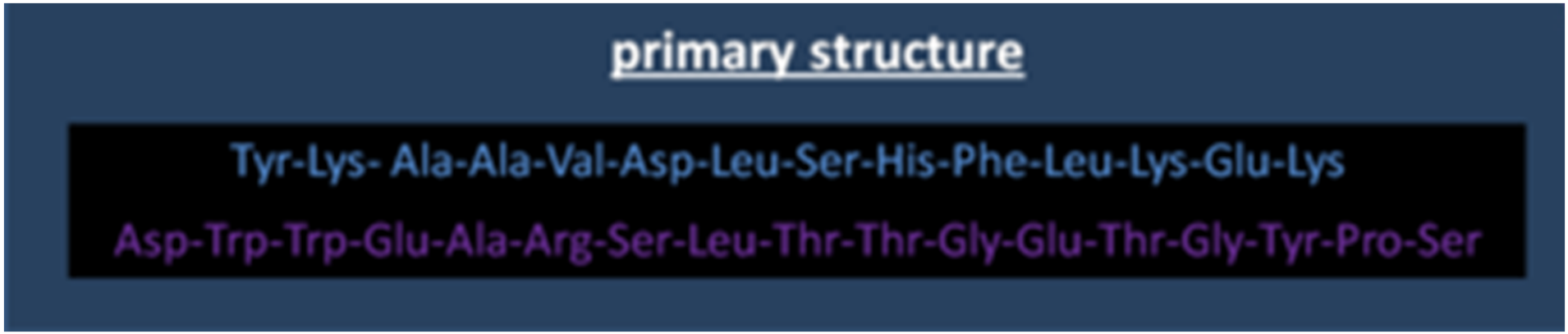
Lecture 6 Visualization of Protein Structurers

Amino acids


Twenty-One Amino Acids
Positive
Negative
• Side chain charge at physiological pH 7.4
A. Amino Acids with Electrically Charged Side Chains
Positive
Arginine (Arg) R
Histidine (His) H
Lysine (Lys) K
Negative
Aspartic Acid (Asp) D
Glutamic Acid (Glu) E
C. Special Cases
Cysteine (Cys) C
Selenocysteine (Sec) U
Glycine (Gly) G
Proline (Pro) P
D. Amino Acids with Hydrophobic Side Chain
Alanine (Ala) A
Valine (Val) V
Isoleucine (Ile) I
Leucine (Leu) L
Methionine (Met) M
Phenylalanine (Phe) F
Tyrosine (Tyr) Y
Tryptophan (Trp) W
B. Amino Acids with Polar Uncharged Side Chains
Serine (Ser) S
Threonine (Thr) T
Asparagine (Asn) N
Glutamine (Gln) Q

Level of protein structures

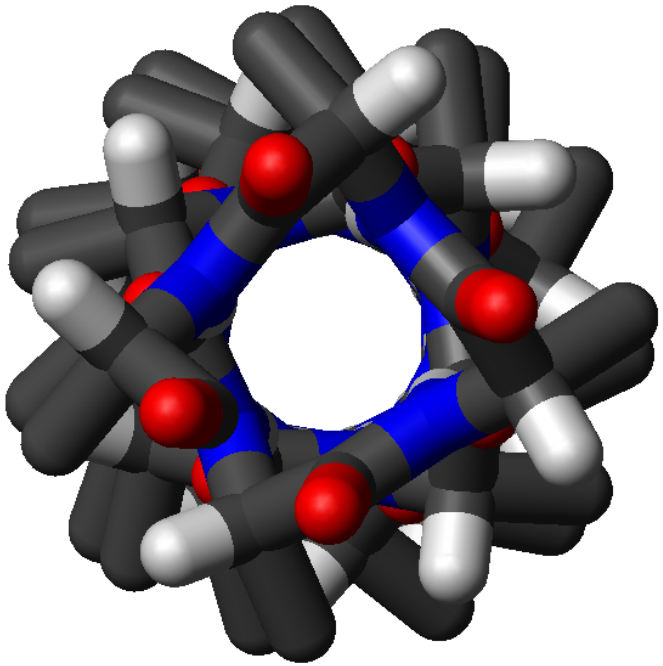
1st : The primary structure of a peptide or protein is the **linear sequence** of its amino acid structural units

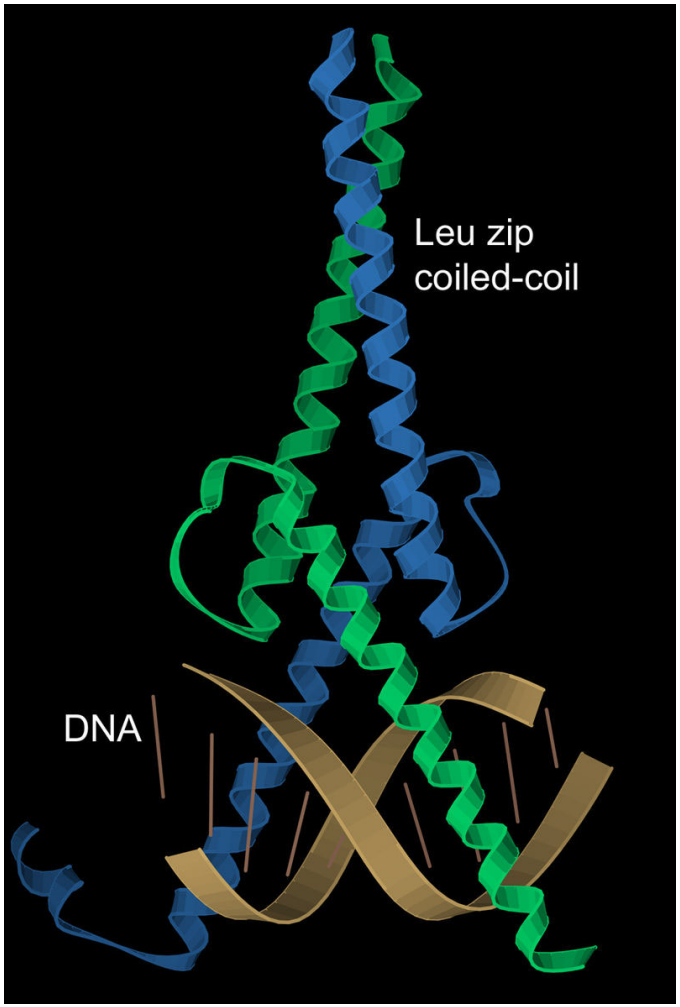


- 2^{nd}: Protein secondary structure is **the general three-dimensional form of local segments of peptides**.
- Secondary structure can be formally defined by the pattern of **hydrogen bonds** of amino acids, (**such as alpha helices and beta sheets**) that are observed in an atomic-resolution structure.

Alpha helix

The alpha helix (α-helix) is a common secondary structure of proteins and is a right hand-coiled or spiral conformation (helix) in which every backbone N-H group donates a hydrogen bond to the backbone C=O group of the amino acid four residues earlier (i+4 → i, hydrogen bonding).





Hydrogen bonding in alpha helix i+4 → i

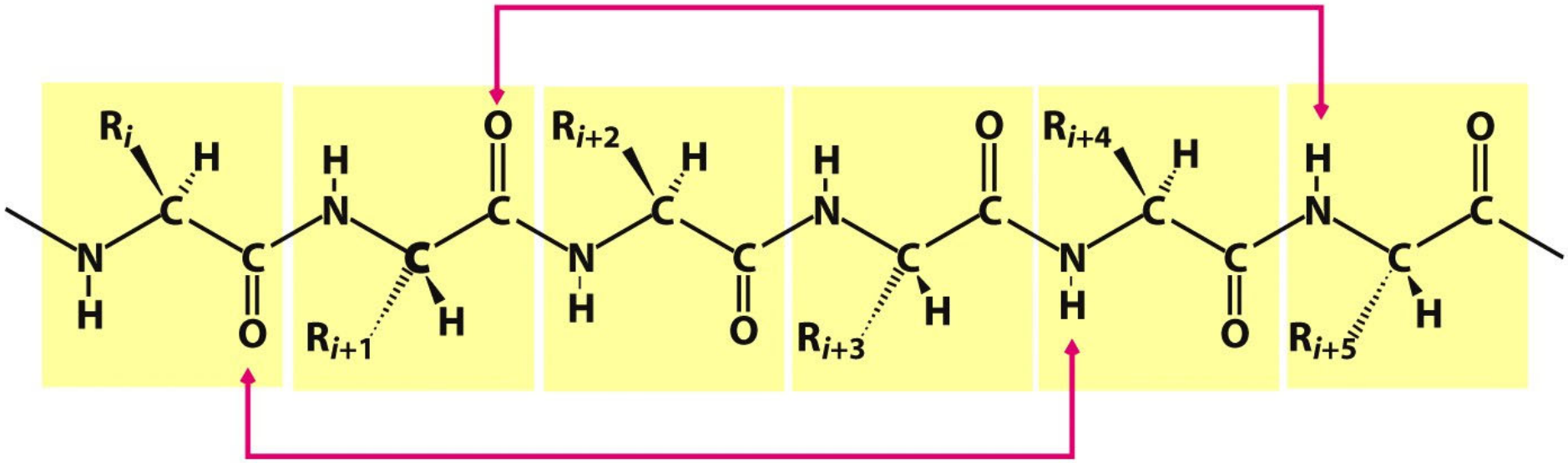
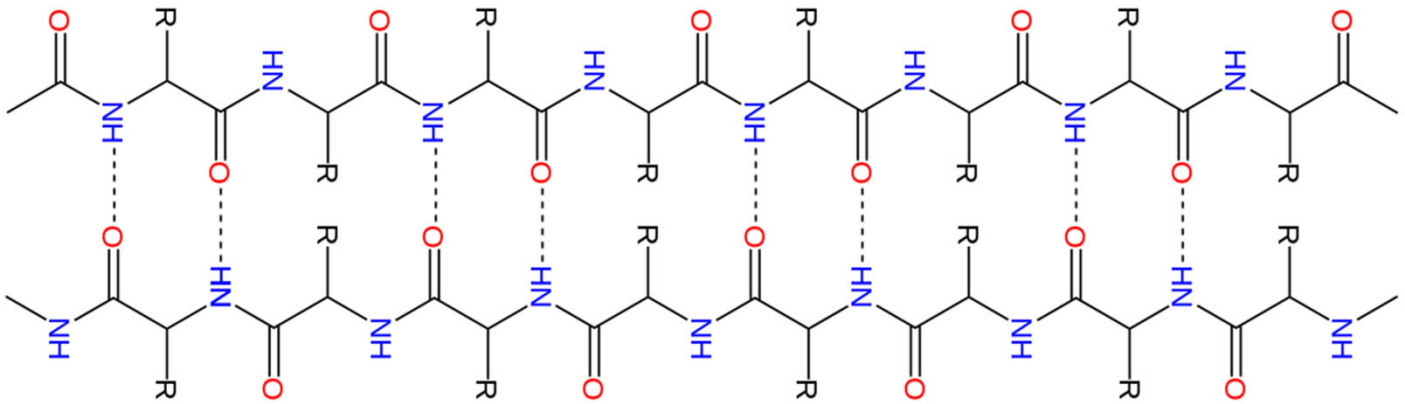


Figure 2.25
***Biochemistry*, Seventh Edition**

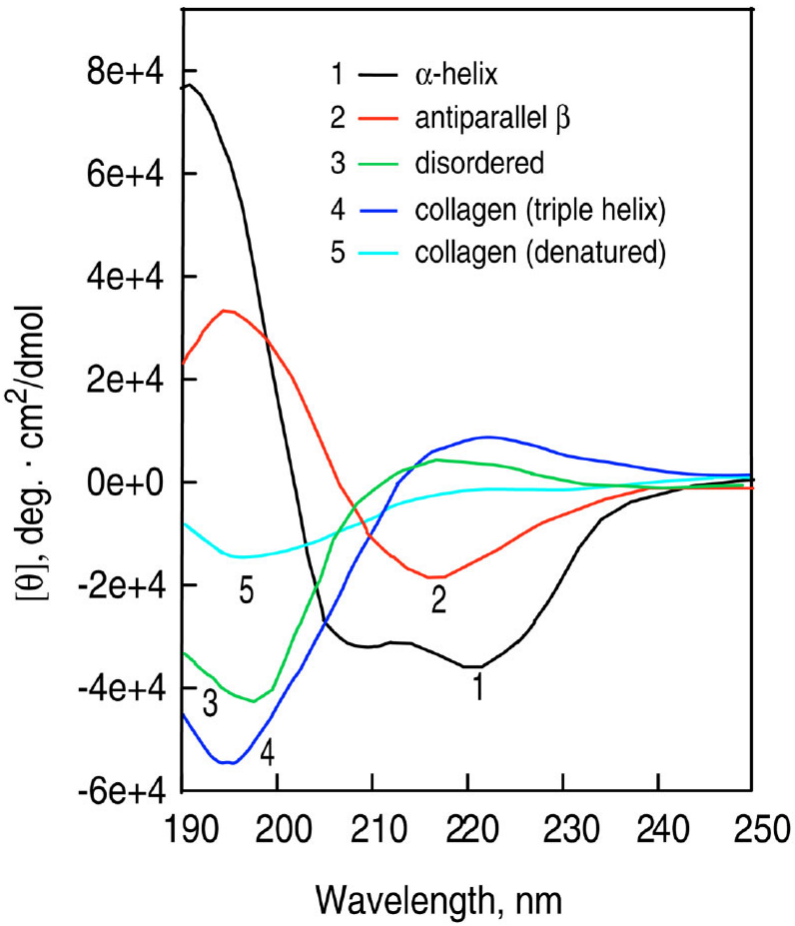
- The β sheet (also β-pleated sheet) is the second form of regular secondary structure in proteins. Beta sheets consist of beta strands connected laterally by at least two or three backbone hydrogen bonds, forming a generally twisted, pleated sheet. A beta strand (also β strand) is a stretch of polypeptide chain **typically 3 to 10 amino acids long** with backbone in an extended conformation.



- Beta-meander motif

How to determine the structure of peptides

- Circular Dichroism (CD)



α-helical proteins have negative bands at 222 nm and 208 nm and a positive band at 193 nm.

Proteins with well-defined **antiparallel β-pleated sheets** (β-helices) have negative bands at 218 nm and positive bands at 195 nm, while disordered proteins have very low ellipticity above 210 nm and negative bands near 195
Nm.

- 3rd structure: Tertiary structure refers to the three-dimensional structure of monomeric and multimeric protein molecules. The alpha-helixes and beta pleated-sheets are folded into a compact globular structure.

Monomeric lactate dehydrogenase

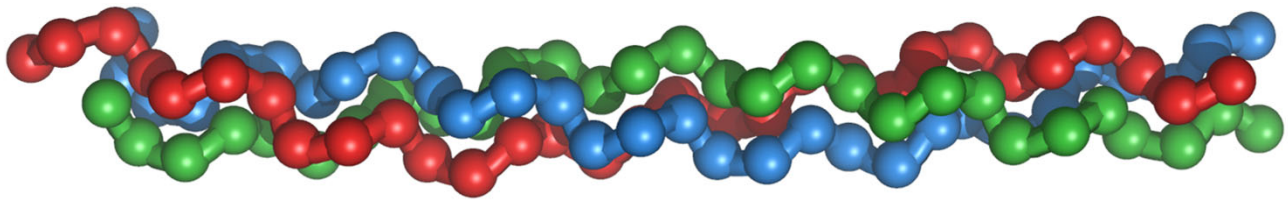
Identification of alpha-helix and beta-sheet

Collagens

- Collagen is the main structural protein in the extracellular space in the various connective tissues in animals. As the main component of connective tissue, it is the most abundant protein in mammals, making up from 25% to 35% of the whole-body protein content.

- Collagen, in the form of elongated fibrils, is mostly found in fibrous tissues such as tendons, ligaments and skin. It is also abundant in corneas, cartilage, bones, blood vessels, the gut, intervertebral discs and the dentin in teeth

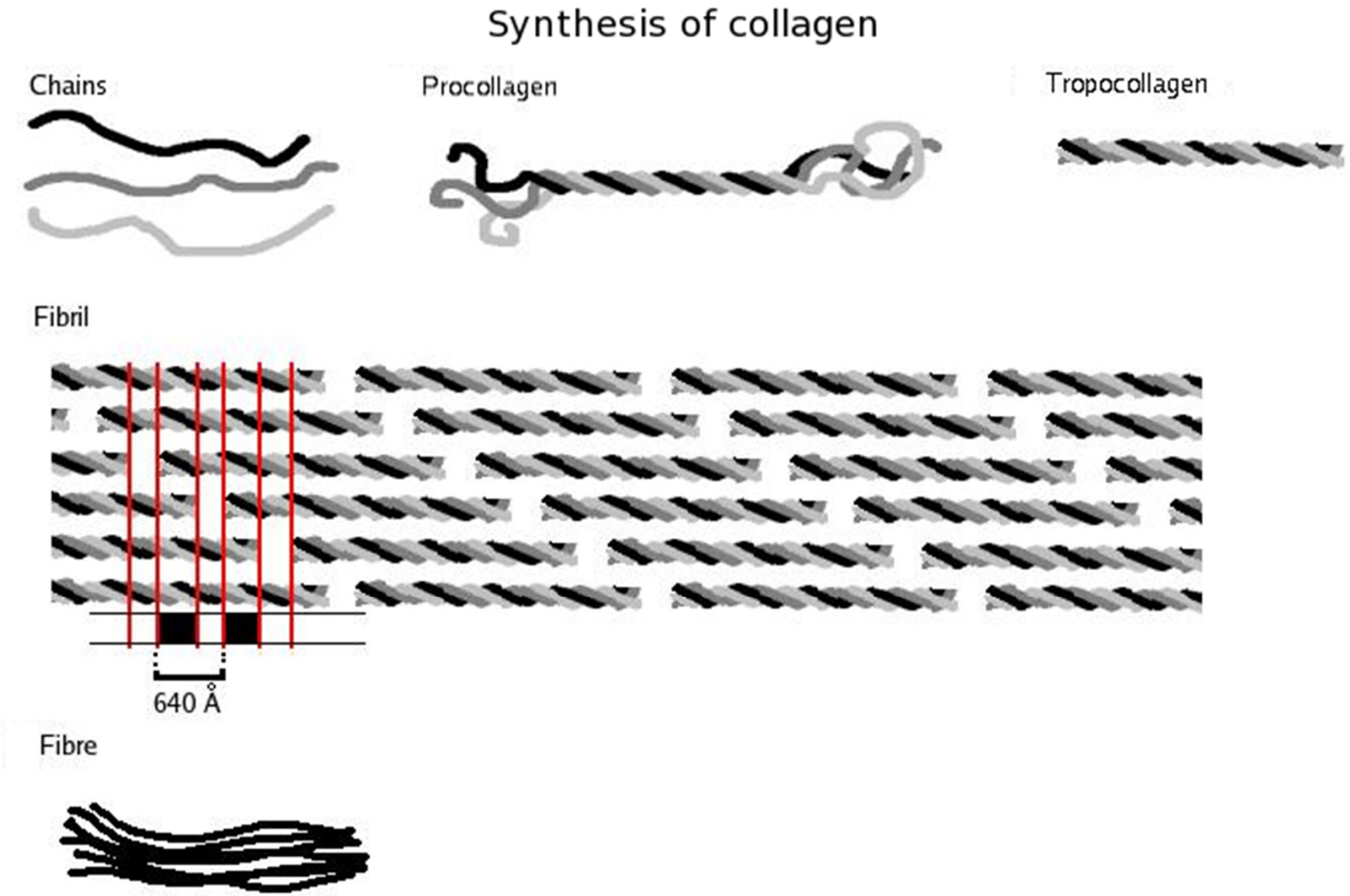
- A single collagen molecule (also known as tropocollagen) is used to make up larger collagen aggregates, such as fibrils. It is approximately 300 nm long and 1.5 nm in diameter, and it is made up of three polypeptide strands



What is Collagen?

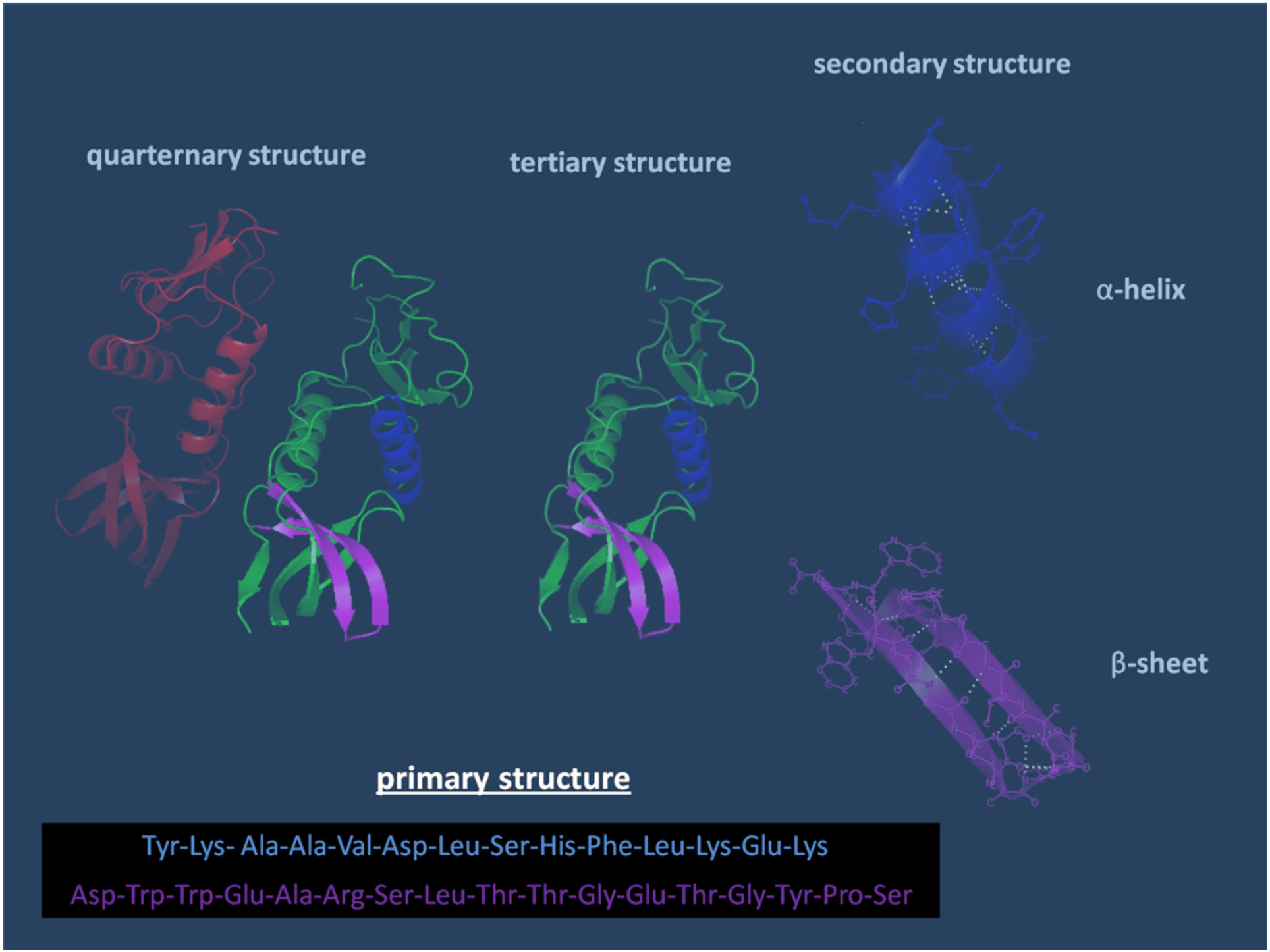
https://www.youtube.com/watch?v=WzoIjL426YY

Collagen chemistry



- 4^{th} structure: Quaternary structure is the three-dimensional structure of a multi-subunit protein and how the subunits fit together.

Tetrameric LDH


secondary structure
quarternary structure
tertiary structure
α-helix
β-sheet
primary structure
Tyr-Lys- Ala-Ala-Val-Asp-Leu-Ser-His-Phe-Leu-Lys-Glu-Lys
Asp-Trp-Trp-Glu-Ala-Arg-Ser-Leu-Thr-Thr-Gly-Glu-Thr-Gly-Tyr-Pro-Ser

AlphaFold: A breakthrough in predicting protein folding by A.I. and machine learning

Article | Open Access | Published: 15 July 2021

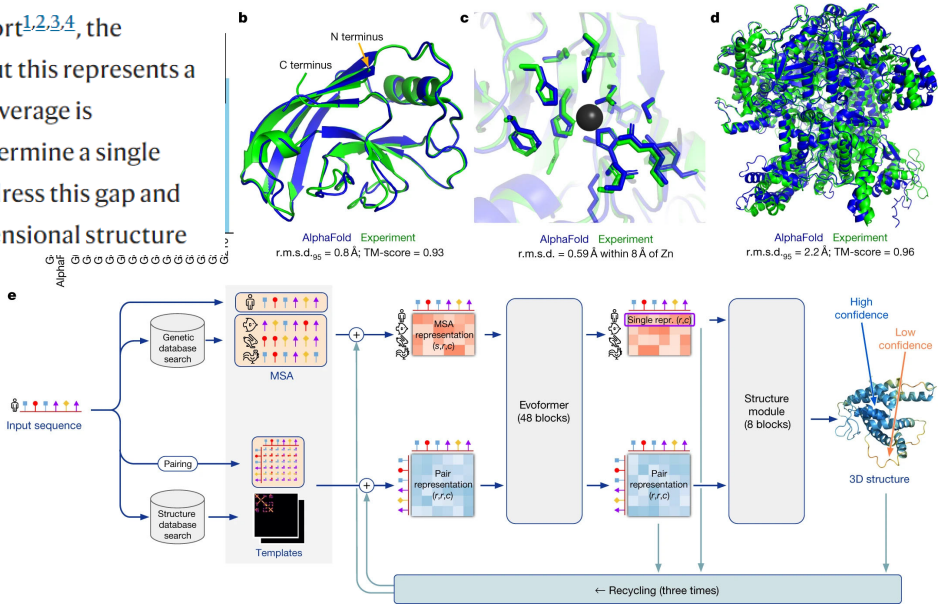
Highly accurate protein structure prediction with AlphaFold

John Jumper, Richard Evans, [...] Demis Hassabis

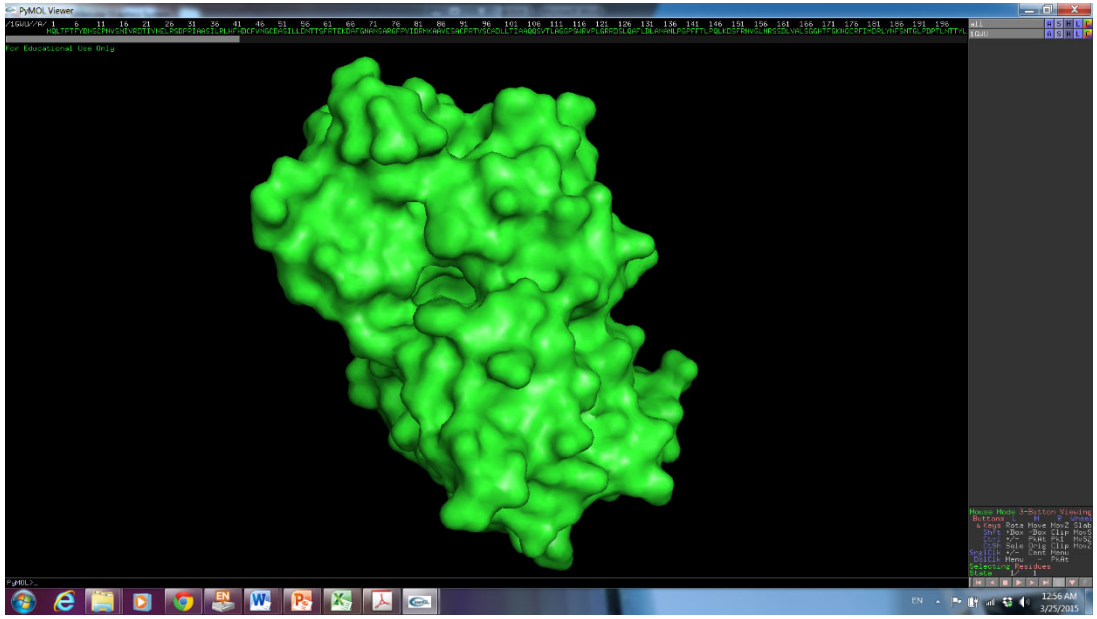
Nature **596**, 583–589 (2021) | Cite this article

Abstract

Proteins are essential to life, and understanding their structure can facilitate a mechanistic understanding of their function. Through an enormous experimental effort[1,2,3,4], the structures of around 100,000 unique proteins have been determined[5], but this represents a small fraction of the billions of known protein sequences[6,7]. Structural coverage is bottlenecked by the months to years of painstaking effort required to determine a single protein structure. Accurate computational approaches are needed to address this gap and to enable large-scale structural bioinformatics. Predicting the three-dimensional structure



Pymol



Registration For Educational-Use-Only PyMOL Builds

Schrödinger offers **Educational-use-only** PyMOL builds available at no cost to **teachers and high school and college students** (including online courses, homeschooling, etc.) for classroom instruction, homework assignments, and to provide a means for creating high quality figures. Please note that it is not provided for the purposes of academic research or publication.

-> FAQ (Frequently Asked Questions)

The Educational-use-only PyMOL builds are provided "AS IS" with no obligation to grant download access, fix bugs, furnish updates, provide documentation, or meet any other need related to the educational-use PyMOL builds.

If you intend to use PyMOL products for academic research or publication, please purchase an Academic PyMOL subscription, which includes access to technical support, screencasts, and additional resources. See http://pymol.org/academic.

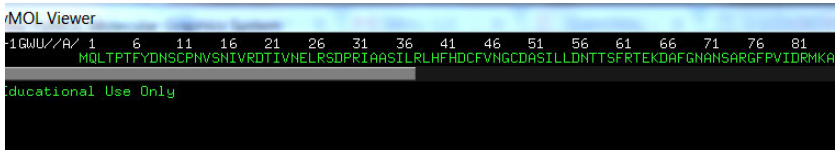
I am a:
Your First Name:
Your Last Name:
Your Email Address:
Your Telephone Number:
Institution:
Comments (optional):
Continue

https://pymol.org/2/
Student and teachers can download a free education version by registration on Pymol
https://pymol.org/edu/?q=educational
Please follow the registration requirement to receive a free download.

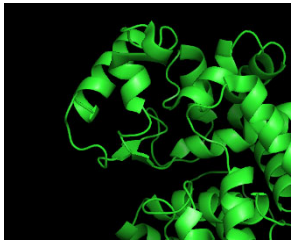
Pymol Methods:

Pymol can open the crystal data from Protein Data Bank http://www.rcsb.org/pdb/home/home.do

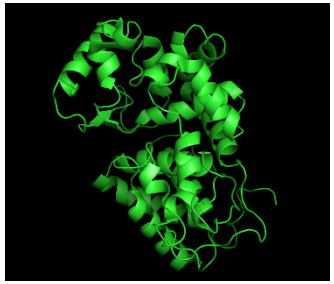
In this lab, several protein structures are provided on sakai, including horseradish peroxidase, glucose 6-phosphate dehydrogenase and lactate dehydrogenase.



M-1: Pymol can show polypeptide sequences (1st structure):

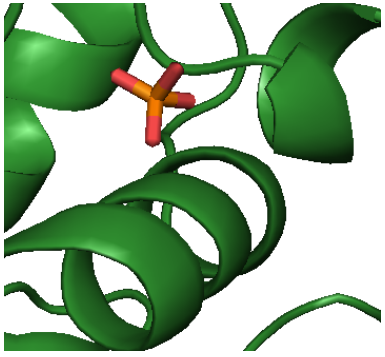


M-2: Show the folding of polypeptides (cartoon mode) of random loop, alpha helix and beta sheet (2nd structure).



M-3: Show the folding of a monomeric protein (3rd structure)

M-4: Show 4th structure of interaction between multimeric proteins:

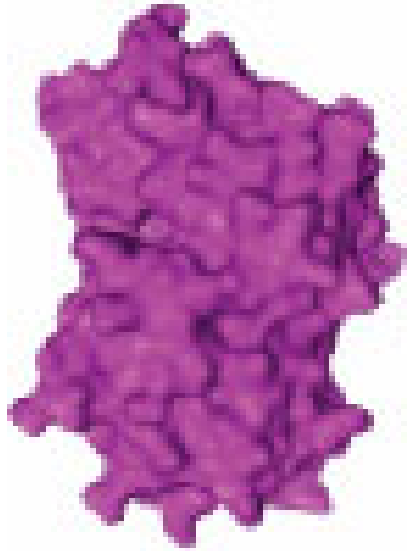


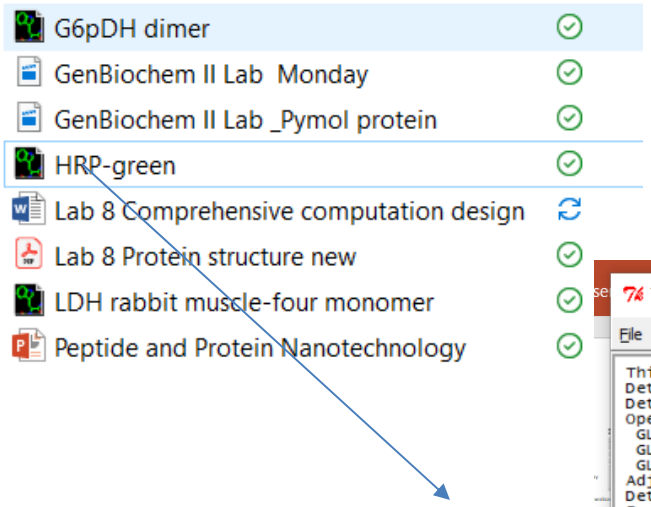
M-5: Show protein-ligand interaction:

M6: Measurement of Protein Sizes (under wizard):

Horseradish peroxidase (HRP)

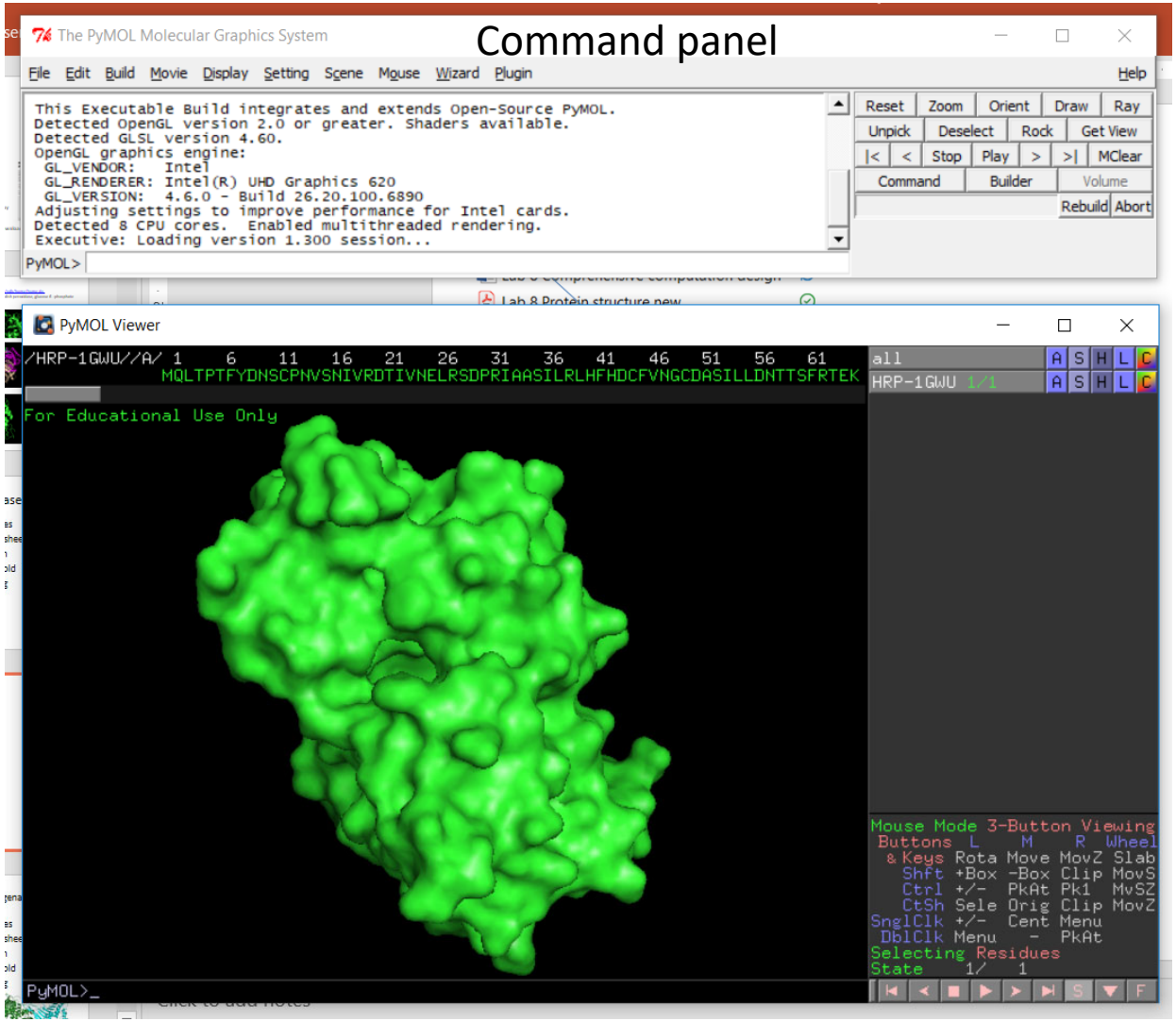
- List 4-level of protein structures
- How long is the polypeptide chain (aa):
- Identify alpha helix and beta-sheet
- Identify multimeric interaction
- Identify protein surface/ backbone
- Identify ligand-protein binding
- Protein size

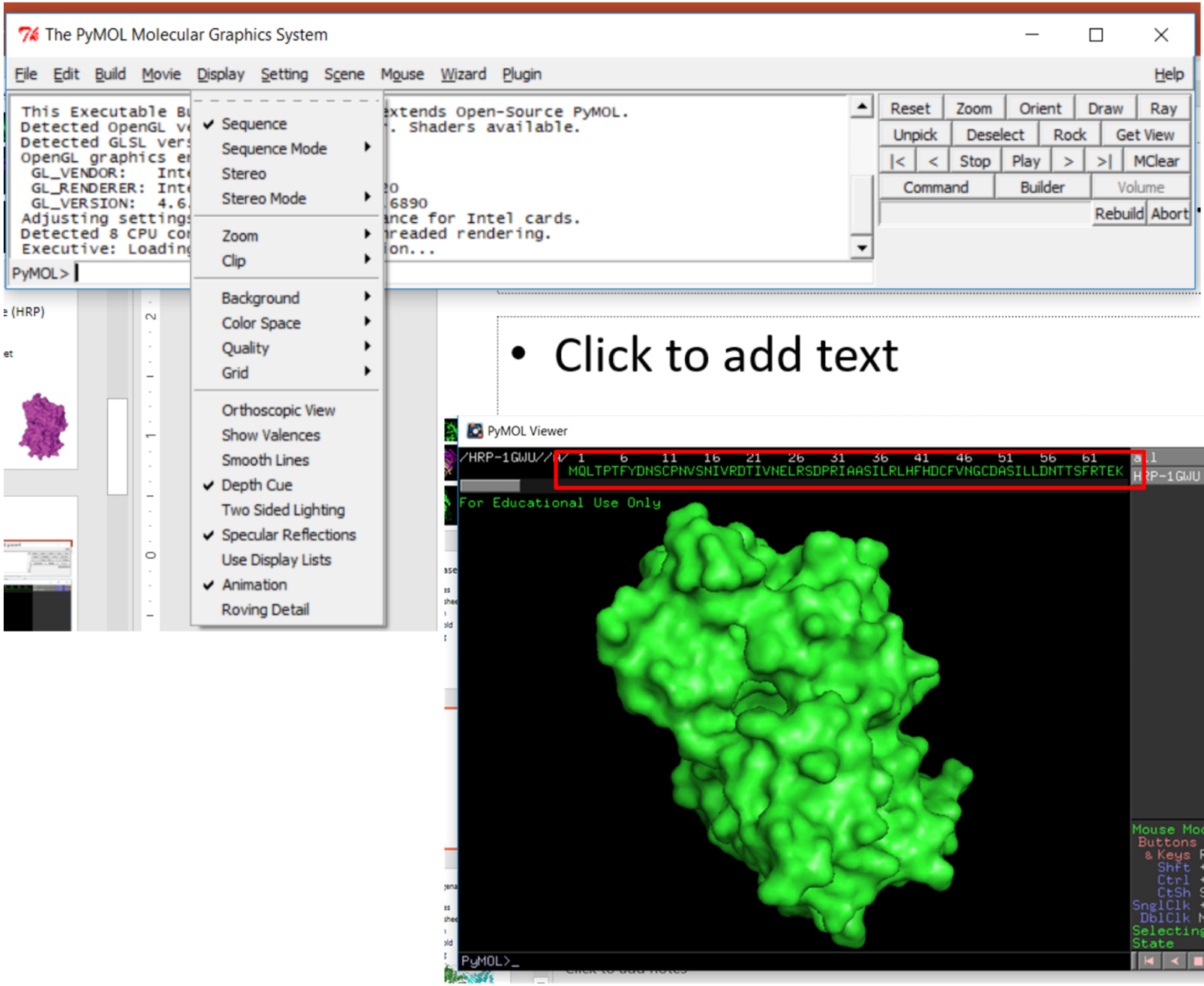




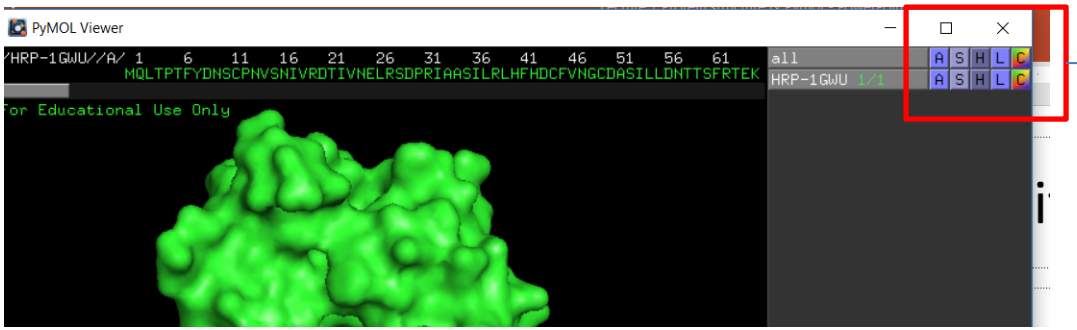
Command panel

View panel



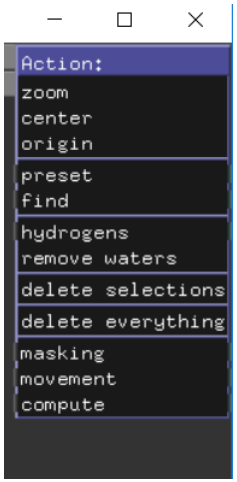

The PyMOL Molecular Graphics System
File Edit Build Movie Display Setting Scene Mouse Wizard Plugin Help
Sequence
Sequence Mode
Stereo
Stereo Mode
Zoom
Clip
Background
Color Space
Quality
Grid
Orthoscopic View
Show Valences
Smooth Lines
Depth Cue
Two Sided Lighting
Specular Reflections
Use Display Lists
Animation
Roving Detail
Reset Zoom Orient Draw Ray
Unpick Deselect Rock Get View
Stop Play MClear
Command Builder Volume
Rebuild Abort
PyMOL>
• Click to add text
PyMOL Viewer
MQLTPTFYDNSCPNVSNIVRDTIVNELRSDPRIAASILRLHFHDCFVNGCDASILLDNTTSFRTEK
For Educational Use Only
PyMOL>_

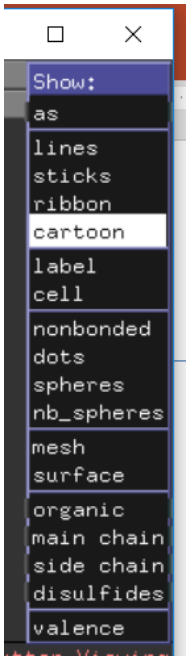
View panel



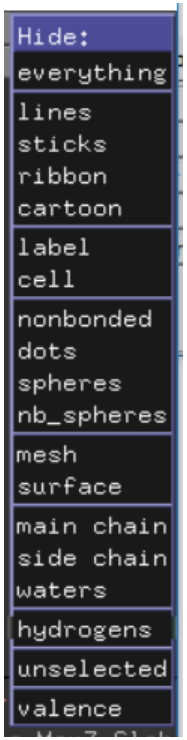
View option

A: Action; S: Show; H: Hide; C: Color

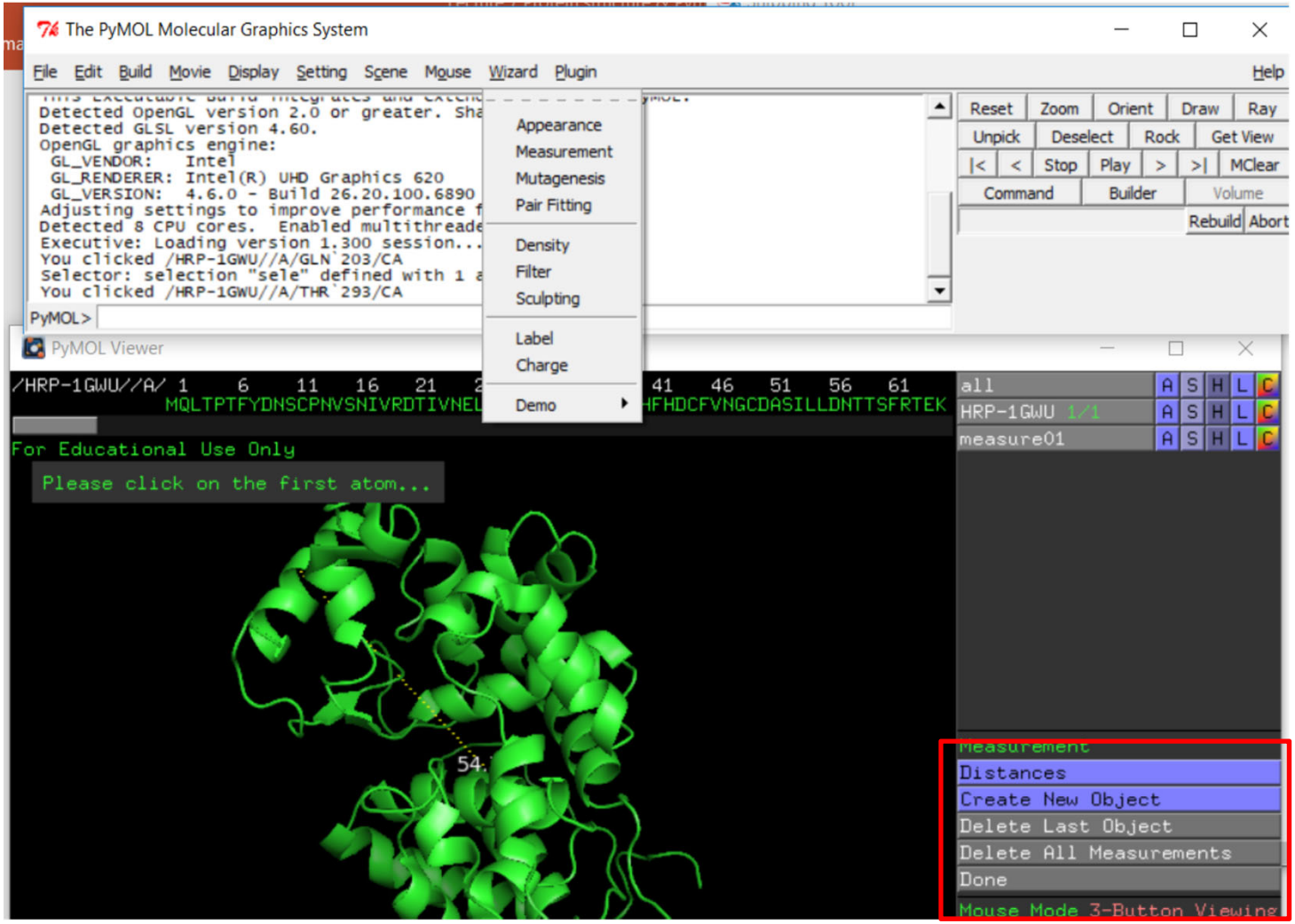




Most useful:
Sticks: Chemical bonds
Cartoon: 2nd structures, alpha helix or beta sheet
Surface: Protein surface

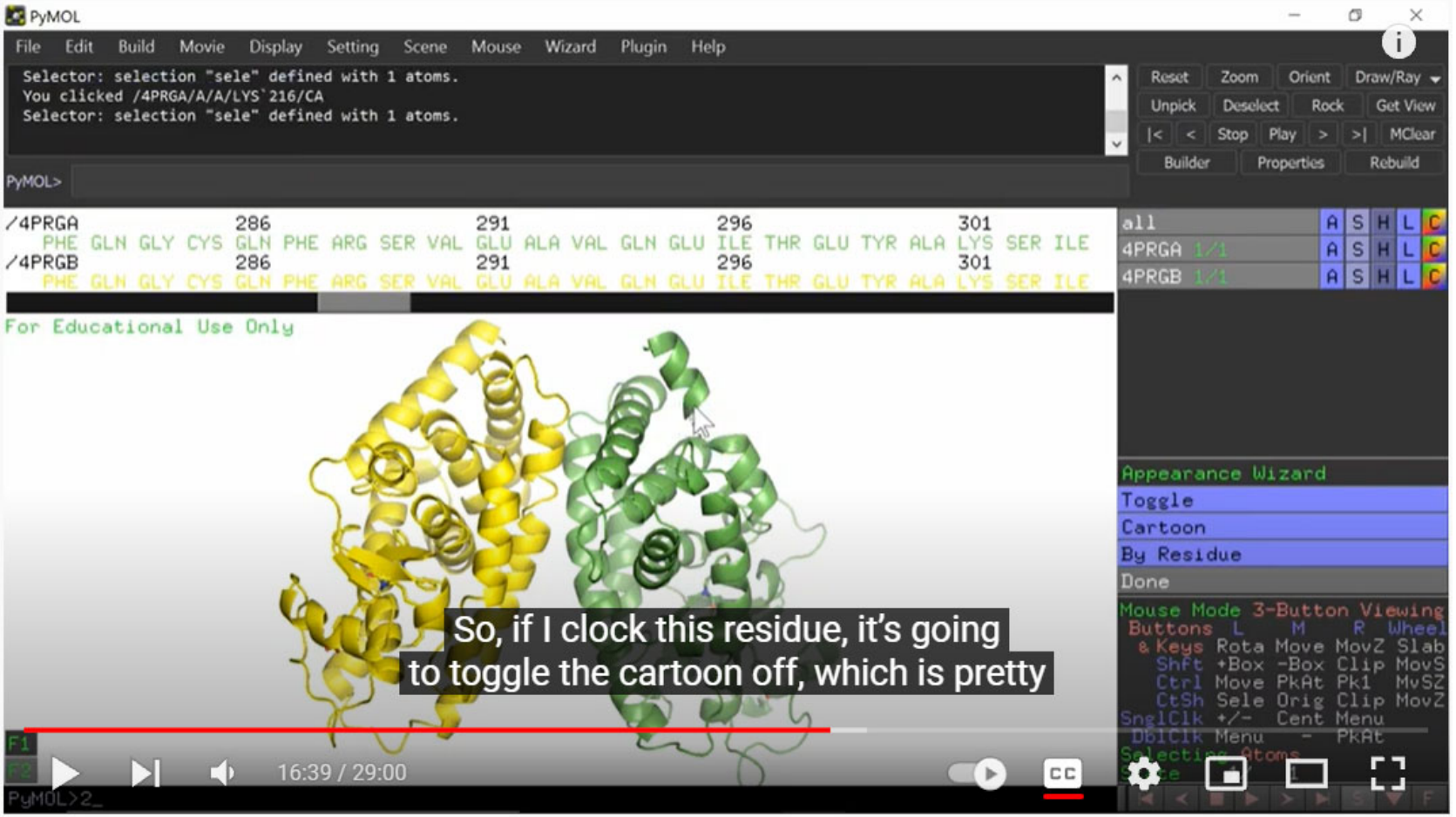


Protein size measurement



Under wizard -→ measurement, choose two position on protein to measure the distance. The unit is Angstrom (0.1 nm)

If there are too many overlaid measurements, you can choose to delete them in the right bottom corner



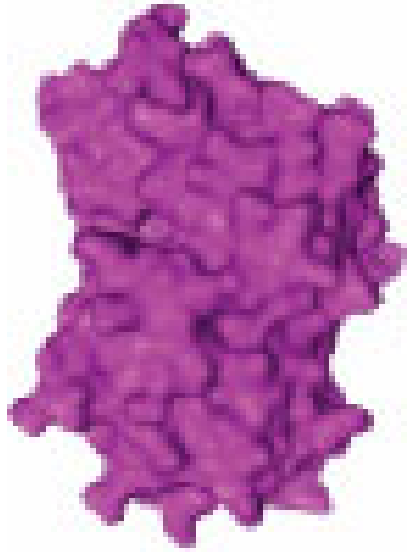
Molecular Visualization Using PyMOL: 1 Introduction/User Interface | Bevan Brown Lab

16,004 views • Jul 2, 2019 275 2 SHARE SAVE ...

https://www.youtube.com/watch?v=aDmOe1ZgTz0

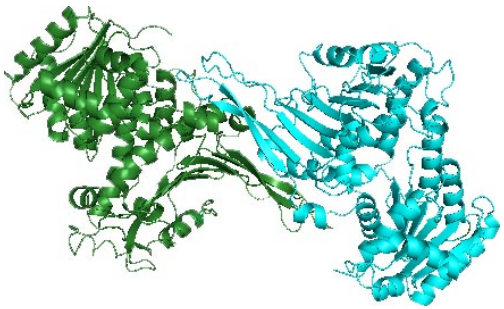
Project 1: Horseradish peroxidase (HRP)

- How many polypeptide chains?
- How long is the polypeptide chain (aa):
- Identify alpha helix and beta-sheet
- Identify protein surface/ backbone
- Identify ligand-protein binding
- Protein size



Project 2: Glucose-6 phosphate dehydrogenase (G6PDH)

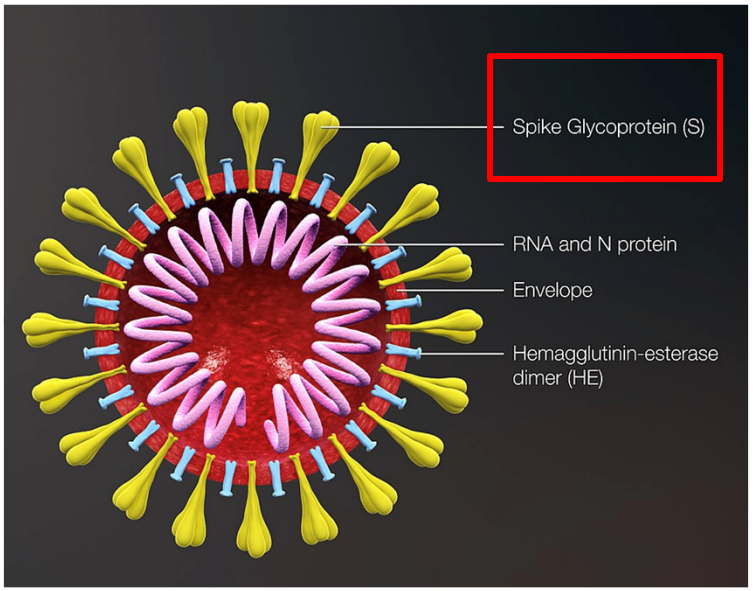
- List 4-level of protein structures
- How long is the polypeptide chain (aa):
- Identify alpha helix and beta-sheet
- Identify multimeric interaction (cartoon, identify interface)
- Identify protein surface/ backbone
- Protein size

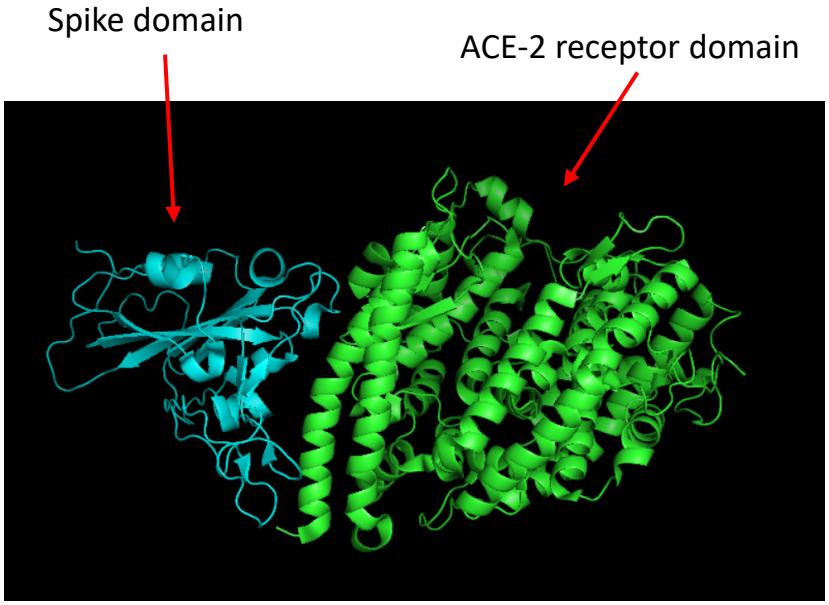


Project 3: Lactate Dehydrogenase (LDH)

- How many polypeptide chains?
- How long is the polypeptide chain (aa):
- Alpha helix/ beta-sheet
- Protein backbone
- Protein surface
- Multimeric interaction (cartoon, identify interface)
- Identify ligand (NAD)-protein binding
- Protein size

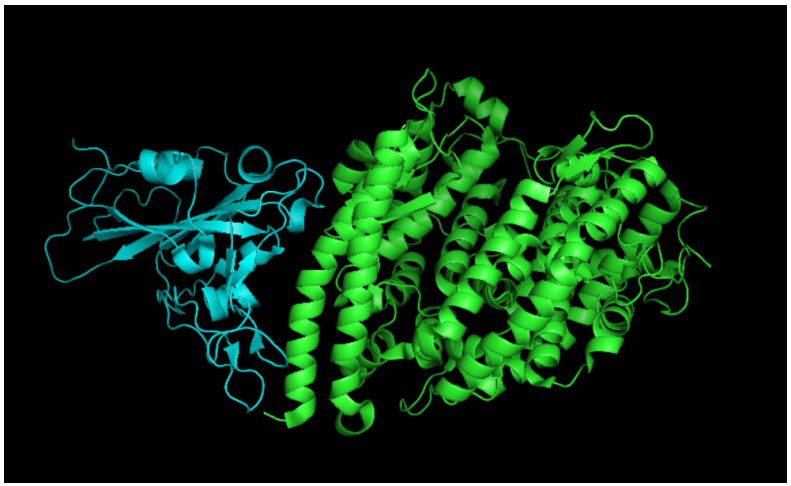
Extra Project 4: SARS-COV-2: Spike protein for targeting on human ACE-2 receptor, trigger infection





https://asm.org/Articles/2020/January/2019-Novel-Coronavirus-2019-nCoV-Update-Uncoating

Extra Project 4:



http://www.rcsb.org/structure/6LZG

- Identify spike domain and ACE-2 receptor in the complex
- How many AA in Spike domain?
- How many AA in ACE-2 receptor domain?
- Measure the size of this complex
- Highlight binding interface of “483 – EGFN….GVGY-504” on spike protein

A Data Report is required for the Project 1-3

TASK-1: Horseradish peroxidase (HRP), use cartoon of figures to illustrate following structures

- How many polypeptide chains in HRP?
- How long is the polypeptide chain (aa):
- Identify alpha helix and beta-sheet
- Identify protein surface/ backbone
- Identify ligand-protein binding (label heme cofactor)
- Protein size (in A or nanometer)

Task -2: Gluose 6-phsophate dehydrogenase (G6PDH):

- List 4-level of protein structures, give an example of each structure
- How long is the polypeptide chain (aa):
- Identify alpha helix and beta-sheet
- Identify multimeric interaction (cartoon, identify interface)
- Identify protein surface/ backbone
- Protein size

Task -3: Lactate dehydrogenase (LDH):

- How many polypeptide chains in LDH?
- How long is the polypeptide chain (aa):
- Identify alpha helix and beta-sheet (Show cartoon structure)
- Show protein surface?
- Protein size (in diameters, use measurement tool)
- Label LDH – NAI (NAD analogue) ligand interaction

Summarize your results for Task 1 -3, submit a word document of Task results. You must use figures from Pymol to demonstrate each task.