

EXPERIMENT 5: STRUCTURE COMPARISON

Dr. Zhiyi Wei
SUSTC

Structural superposition/alignment



RMSD (Root-Mean-Square Deviation)

- The way to evaluate the structural similarity by measuring the average distance between the atoms of superimposed proteins

$$\text{RMSD} = \sqrt{\frac{1}{N} \sum_{i=1}^N \delta_i^2}$$

where δ is the distance between N pairs of equivalent atoms (usually the backbone atoms)

- Indication of RMSD
 - $<1\text{\AA}$, near identical
 - $1\text{-}2\text{\AA}$, highly similar
 - $2\text{-}4\text{\AA}$, similar
 - $>4\text{\AA}$, usually no similarity

Methods

- **Least-square fitting (LSQ)**
- DALI
 - Distance alignment matrix
 - Break the input structures into hexa-peptide fragments
 - Calculate a distance matrix by evaluating the contact patterns between successive fragments
- **Combinatorial extension (CE)**
 - Similar to DALI
- **Sequential Structure Alignment Program (SSAP)**
 - Double dynamic programming
 - Produce a structural alignment based on atom-to-atom vectors in structure space

Simple alignment in PyMOL

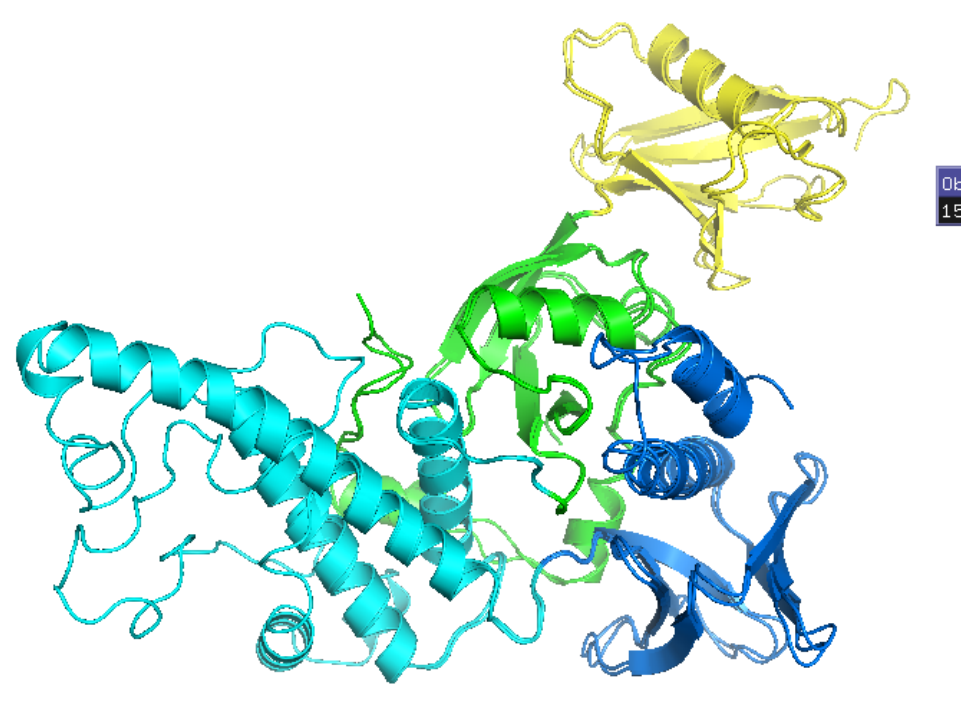
PyMOL Viewer

/refine/A/A/18 21 26 31 36 41 46 51 56 61 66 71 76 81 86 91 96 101 106

WALSVHVTDLNRDVTLRVTGEVHIGGVMLKLVKLDVKKDWSHDALWWEKKRTWLLKTHWTLDKCGIQADAKLQFTPPQHKLRLQLPNMKYVK

/15-acetate/A/A/12 16 21 26 31 36 41 46 51 56 61 66 71 76 81 86 91 96 101

MGSDGTWELSVHVTDLNRDVTLRVTGEVHIGGVMLKLVKLDVKKDWSHDALWWEKKRTWLLKTHWTLDKCGIQADAKLQFTPPQHKLRLQLP



all

refine

3ivf-coo

<F0>

<F1>

<F2>

<F3>

<F01>

<tF3>

<tF2>

<tF1>

Action:

zoom

orient

center

origin

drag matrix

reset matrix

drag coordinates

clean

preset

find

align

generate

assign sec. struc.

rename object

duplicate object

delete object

hydrogens

remove waters

state

masking

sequence

movement

compute

Object:

15-acetate

Align:

to molecule

to selection

enabled to this

all to this

states (* /ca)

states

matrix from

matrix to

matrix reset

Mouse Mode 3-Button Viewing

Buttons L M R Wheel

& Keys Rota Move MovZ Slab

Shft +Box -Box Clip MovS

Ctrl +/- PKAt PK1 MvSZ

CtSh Sele Orig Clip MovZ

SnglClk +/- Cent Menu

Db1Clk Menu - PKAt

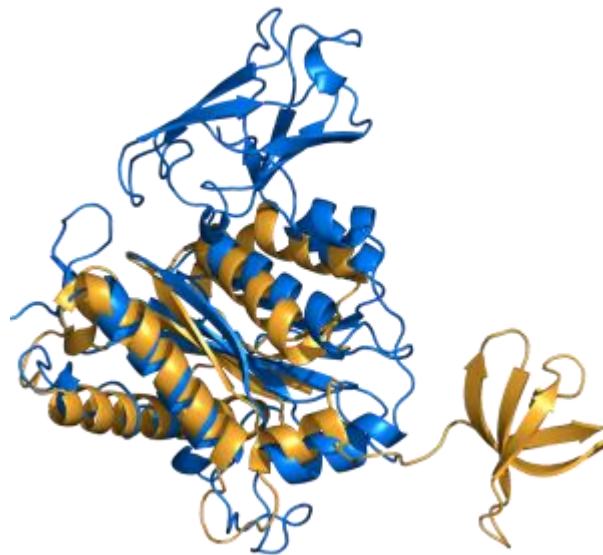
Selecting Residues

State 1/ 1

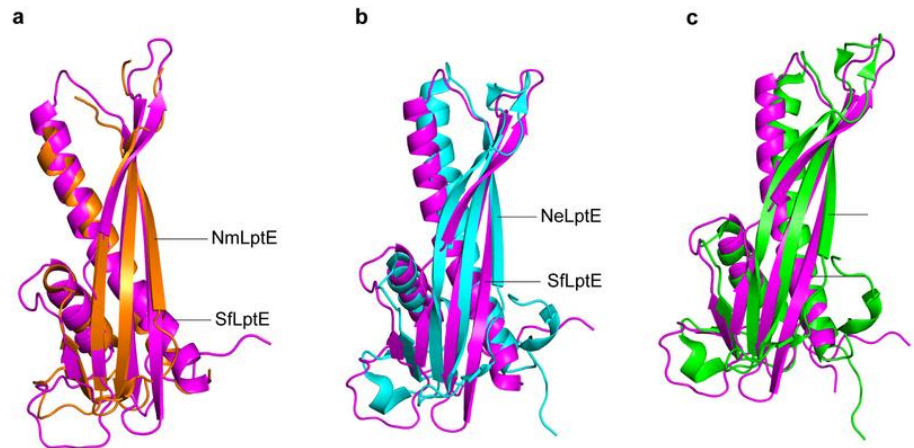
PyMOL>_

Cealign plugin in PyMOL

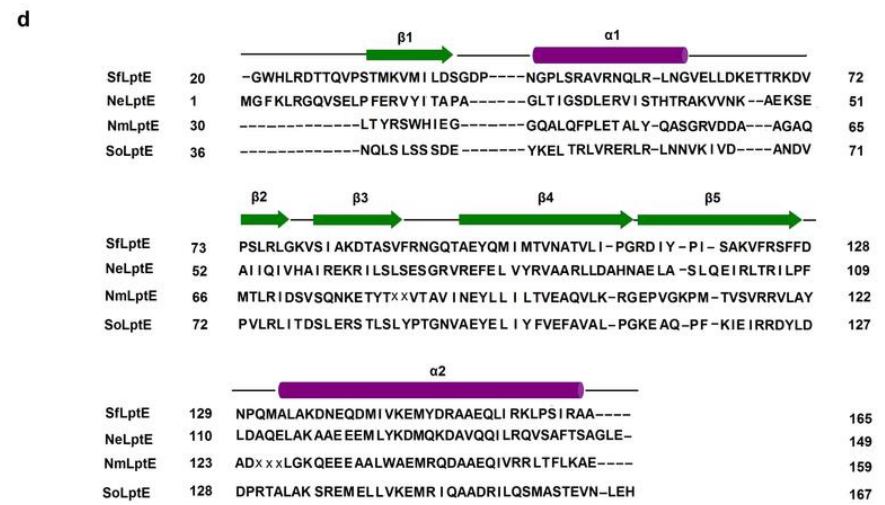
- Used for accurate structural alignment
- Better results for structures with low sequence similarity



Structure-based sequence alignment



Much more accurate!



Tasks

1. Open human hemoglobin structure (PDB id: 4HHB) in PyMOL
 - PDB Loader in Plugins
2. Create new objects containing either alpha-subunit or beta-subunit
 - You can use different sequence mode to make selection
3. Align beta-subunit to alpha-subunit
 - RMSD value and indication?
4. Analyze the pairwise sequence alignment ^{structure}
 - Which region cannot be aligned? Your explanation?
5. Load human neuroglobin (PDB id: 4MPM) and a globin-like protein (PDB id: 4BJA), and align them to alpha-globin, respectively
 - Evaluate your results
6. Use Cealign to realign the globin-like protein to alpha-globin
 - Compared with the simple align, which one give a better alignment?
7. Compare the structure-based alignment and simple sequence alignment of the three globins (this part you can finish at home)
 - Which one give you a better alignment? Why?

Lab report format

- Title
- Your name and student No.
- Introduction
- Methods
- Results
- Conclusions