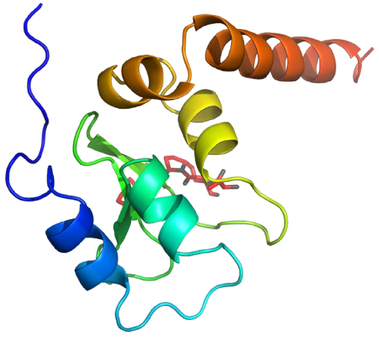
Computational Modeling of Proteins for Infectious Disease Researchers



Workshop

presented by NIH and the Centers for Research on Structural Biology of Infectious Diseases (CRSTAL-ID)

Reminder:
Please make your lunch orders before we get started.

Computational Modeling and analysis for Protein-Protein Interactions.

Bart Staker, PhD.
Seattle Children's Research Institute
Seattle Structural Genomics Center for Infectious Disease

Seattle Children's Research Institute

Day 2: [**Tuesday** July 21st, 2026]

- **Session 3: Protein:Protein Complexes**
- **8:00 AM – 8:30 AM**: Coffee and Networking
- **8:30 AM – 9:30 AM:** Application talk**.**
  - Protein complex prediction and validation
  - **9:30 AM – 9:45 AM**: Coffee Break
- **9:45 AM – 11:45 PM**: Hands-On Session:
  - Protein-Protein Interaction Analysis
  - Virtual Reality view of a complex
- **11:45 AM – 12:00 AM**: Wrap-up
- **12:00 PM - 1:00 PM**: Lunch Break
- Lunch on your own.

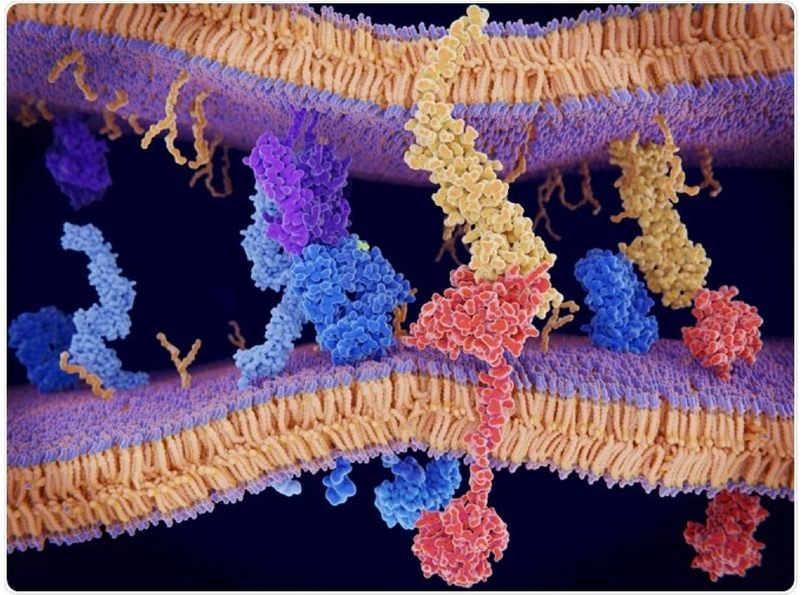
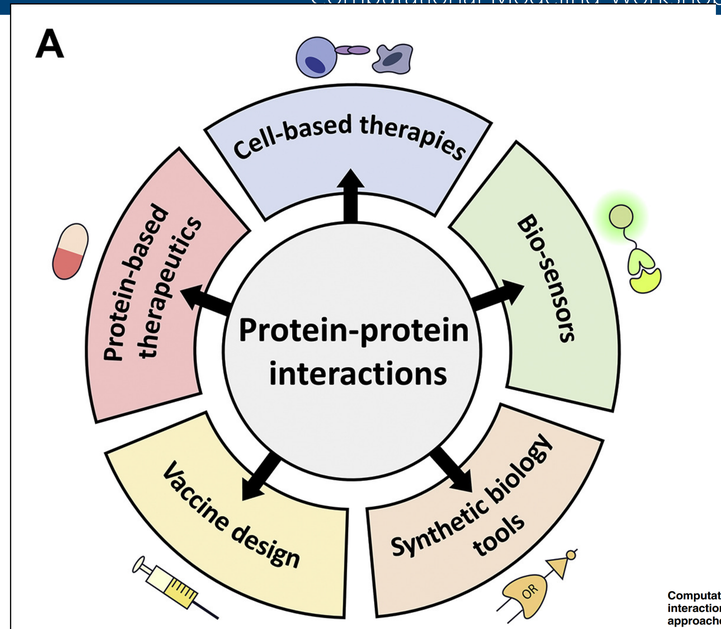
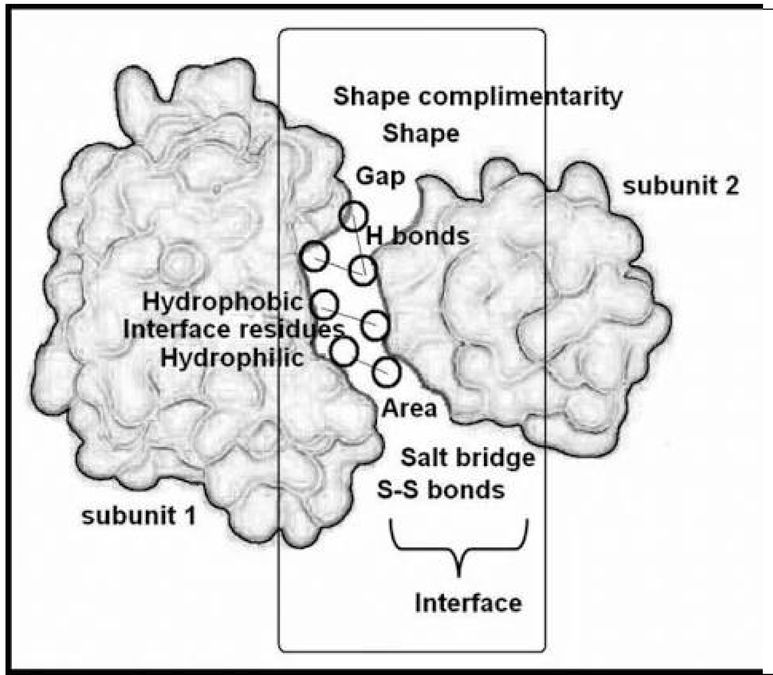


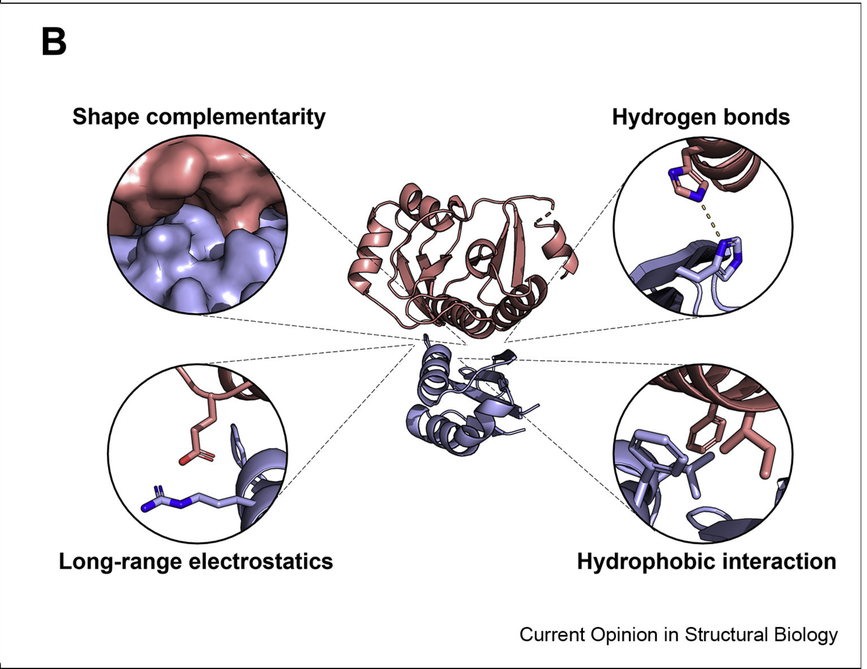
Image Credit: ©Juan Gaertner/Shutterstock.com

Why are protein protein interactions important?



Computational design of novel protein–protein interactions – An overview on methodological approaches and applications
Anthony Marchand[a,b], Alexandra K. Van Hall-Beauvais[a,b] and Bruno E. Correia[a,b]


Shape complimentarity
Shape
Gap
subunit 2
H bonds
Hydrophobic
Interface residues
Hydrophilic
Area
Salt bridge
S-S bonds
subunit 1
Interface


B
Shape complementarity
Hydrogen bonds
Long-range electrostatics
Hydrophobic interaction
Current Opinion in Structural Biology

Learning Objectives

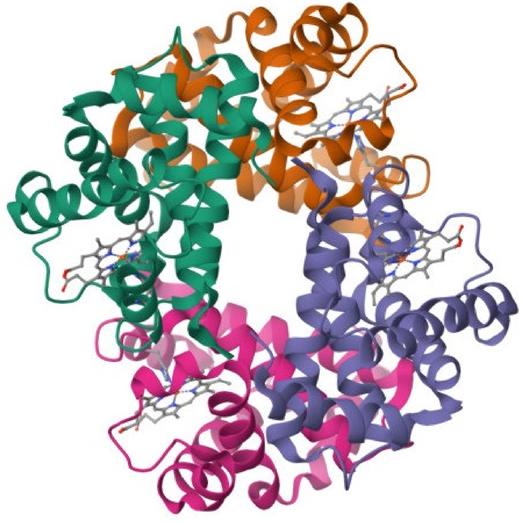
By the end of this session participants will be able to:

- Predict candidate protein-protein interactions using modern computational tools.
- Critically evaluate confidence in predicted interactions.
- Analyze interfaces and identify likely binding determinants.
- Prioritize experimental validation strategies.
- Understand current capabilities and limitations of AI-based structural biology.

The first protein-protein complex structure

was …

The second crystal structure ever solved

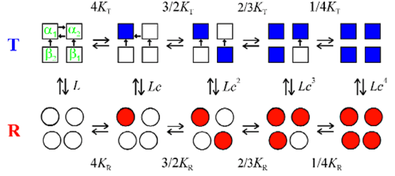


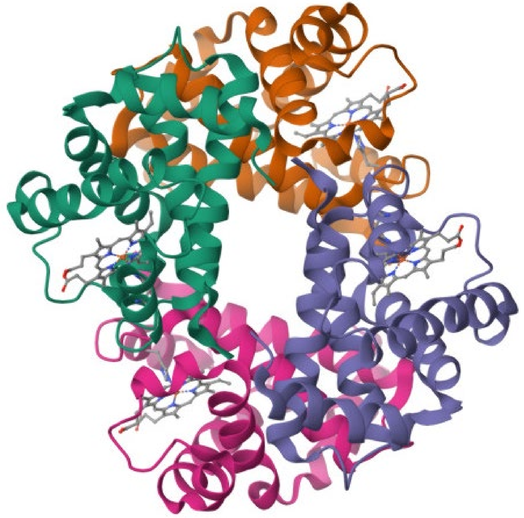
2DHB. Max Perutz, deposited in 1973
HORSE DEOXYHAEMOGLOBIN

The first protein-protein complex structure

was …

The second crystal structure ever solved



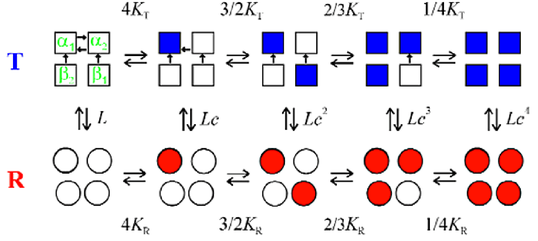


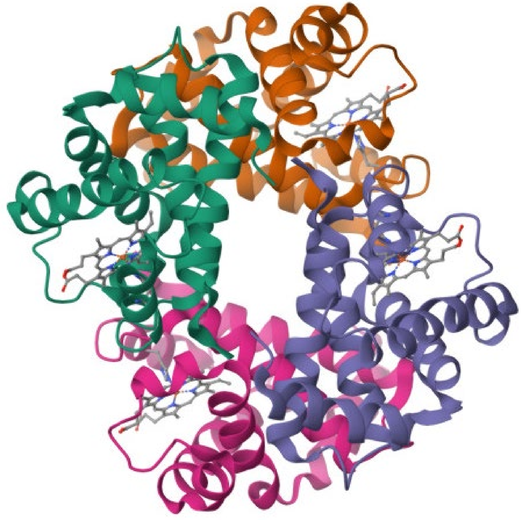
2DHB. Max Perutz, deposited in 1973
HORSE DEOXYHAEMOGLOBIN

The first protein-protein complex structure

was …

The second crystal structure ever solved





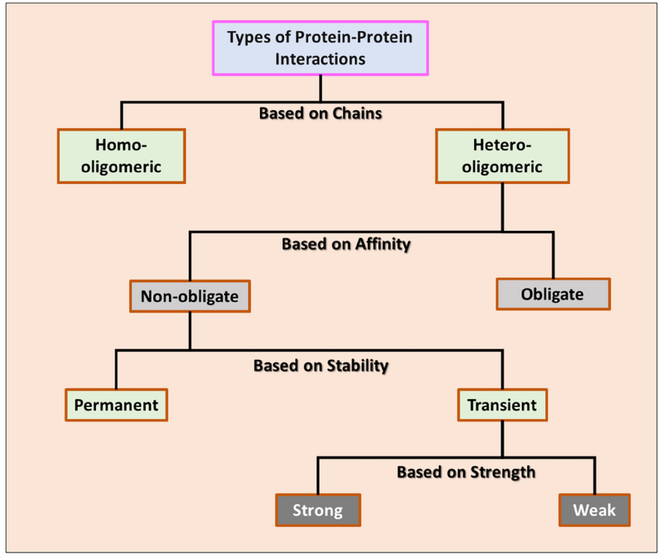
2DHB. Max Perutz, deposited in 1973
HORSE DEOXYHAEMOGLOBIN

shutterstock.com · 2690866113

Types of protein-protein complexes

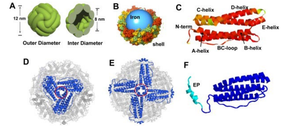
A topological review on protein–protein interactions: the development and promises in the era of omics

August 2024 · Journal of Proteins and Proteomics 15(3):523-544
DOI:10.1007/s42485-024-00160-w

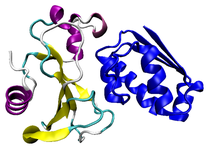


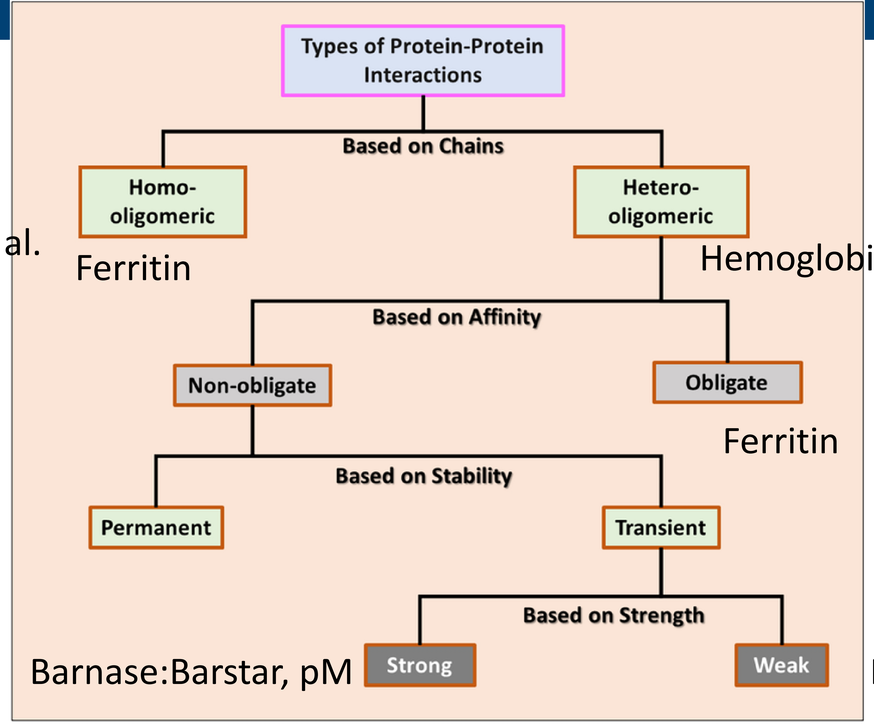
By dependence on the partner: obligate vs. non-obligate

- *Obligate* — the protomers are not stable on their own and are only found as part of the complex.
- *Non-obligate* — the components can exist independently as monomers or together as a complex in vivo.



Malate dehydrogenase et al.





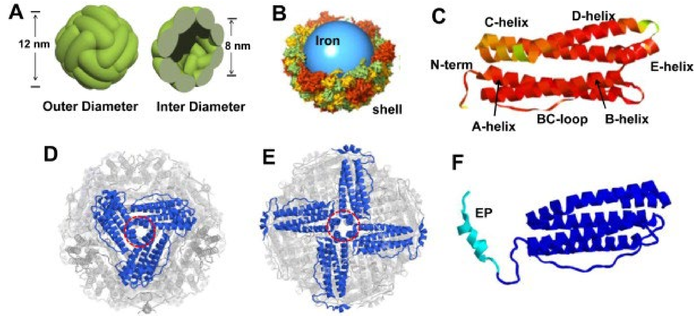
Ferritin

Hemoglobin,

Ferritin

Barnase:Barstar, pM

PD1:PDL1, uM



Experimental Methods to determine protein-protein interactions.

Major Classes of PPI Methods

Genetic / Reporter-based

Detect interactions through reporter gene activation

Biochemical / Affinity-based

Capture protein complexes using antibodies or affinity tags

Proximity Labeling

Label nearby proteins inside living cells

Biophysical

Measure binding affinity and kinetics quantitatively

Structural

Reveal how proteins interact within complexes

Discovery Methods

- **Yeast Two-Hybrid (Y2H)**
  - Screens for novel interacting proteins
  - Library screening and interaction mapping
  - Higher false-positive rate
- **Affinity Purification-Mass Spectrometry (AP-MS)**
  - Identifies components of protein complexes
  - Excellent for large interaction networks

Validation Methods

- **Co-Immunoprecipitation (Co-IP)**
  - Gold standard for confirming interactions
  - Tests interactions under near-native conditions
- **Pull-down Assays**
  - Tagged bait captures potential binding partners
  - Common first-pass biochemical assay

Methods for Living Cells and Quantitative Analysis

In Living Cells

- **BioID / APEX**
  - Covalently labels nearby proteins
  - Captures weak and transient interactions
  - Low background after stringent purification
- **FRET / BiFC**
  - Visualize interactions in living cells
  - Reveal when and where proteins interact

Quantitative Characterization

- **SPR**
  - Measures binding kinetics
  - Determines association, dissociation, and affinity
- **ITC**
  - Measures affinity, stoichiometry, enthalpy, and entropy
  - Label-free solution assay

Structural Analysis

- **Crosslinking Mass Spectrometry (XL-MS)**
  - Identifies interacting residues
  - Provides distance restraints for structural modeling
- **Structure Determination**
  - X-ray crystallography
  - Cryo Electron Microscopy
- **Computational Methods**
  - Prediction
  - Analysis

Choosing the Right Method

Goal	Recommended Methods
Discover new partners	Y2H, AP-MS
Confirm interaction	Co-IP, Pull-down
Detect weak/transient interactions	BioID, APEX
Observe interactions in living cells	FRET, BiFC
Measure affinity & kinetics	SPR, ITC
Determine interaction architecture	XL-MS

So...What Is the Purpose of Computational Modeling?

Why use computational models?

- Predict **how proteins interact** before doing experiments
- Generate **testable hypotheses**
- Prioritize the **most promising experiments**
- Visualize molecular interactions that are difficult to observe directly
- Interpret and extend experimental results

What can computational modeling do?

- Predict protein structures
- Predict protein-protein complexes
- Identify likely binding interfaces
- Estimate effects of mutations
- Suggest mechanisms of interaction
- Guide drug and protein design

What can't it do?

- ❌ Prove an interaction exists
- ❌ Replace experimental validation
- ❌ Always predict the biologically relevant complex

Computational methods for predicting PPIs.

Computational Methods for Predicting Protein-Protein Interactions

Sequence-Based Methods

- Detect interacting partners from sequence features
- Co-evolution and correlated mutations
- Machine learning using known interaction datasets

Structure-Based Methods

- **Protein docking** predicts how two proteins fit together
- Physics-based scoring of binding interfaces
- Template-based modeling from known complexes

AI-Based Methods

- **AlphaFold-Multimer**
  - Predicts structures of protein complexes
  - Uses deep learning trained on known protein structures
- Other AI tools (Boltz, Chai-1, RoseTTAFold All-Atom)
  - Alternative approaches for complex prediction

Network & Systems Biology

- Predict interactions from biological networks
- Gene co-expression
- Functional pathways
- Evolutionary conservation
- Protein interaction databases

Strengths

- Rapid and inexpensive
- Screens thousands of candidate interactions
- Guides experimental design
- Reveals structural hypotheses

Limitations

- Predictions require experimental validation
- Performance varies by protein type and available data
- Flexible and transient interactions remain challenging

Computational Tools for Predicting Protein-Protein Interactions

Purpose	Example Software	Primary Use
Interaction Databases	IntAct, BioGRID, STRING, MINT	Known and predicted PPIs
Structure Prediction	AlphaFold 3, AlphaFold-Multimer, Boltz-2, Chai-1, RoseTTAFold All-Atom	Predict structures of protein complexes
Protein Docking	HADDOCK, ClusPro, RosettaDock, ZDOCK	Predict how two proteins bind
Interface Analysis	ChimeraX, PDBePISA, PRODIGY	Analyze interfaces and estimate binding strength
Molecular Dynamics	GROMACS, OpenMM, AMBER, NAMD	Test stability of predicted complexes
Binding Free Energy	OpenMM/OpenFE, MM/PBSA, FoldX, Rosetta	Estimate effects of mutations and binding affinity
Visualization	ChimeraX, PyMOL, Mol*	Explore and present protein complexes

Computational Tools for Predicting Protein-Protein Interactions

Purpose	Example Software	Primary Use
Interaction Databases	IntAct, BioGRID, STRING, MINT	Known and predicted PPIs
Structure Prediction	AlphaFold 3, AlphaFold-Multimer, Boltz-2, Chai-1, RoseTTAFold All-Atom	Predict structures of protein complexes
Protein Docking	HADDOCK, ClusPro, RosettaDock, ZDOCK	Predict how two proteins bind
Interface Analysis	ChimeraX, PDBePISA, PRODIGY	Analyze interfaces and estimate binding strength
Molecular Dynamics	GROMACS, OpenMM, AMBER, NAMD	Test stability of predicted complexes
Binding Free Energy	OpenMM/OpenFE, MM/PBSA, FoldX, Rosetta	Estimate effects of mutations and binding affinity
Visualization	ChimeraX, PyMOL, Mol*	Explore and present protein complexes

Tools We'll Use in This Workshop

⭐ **AlphaFold-Multimer**

- Predict protein complexes
- Generate structural hypotheses

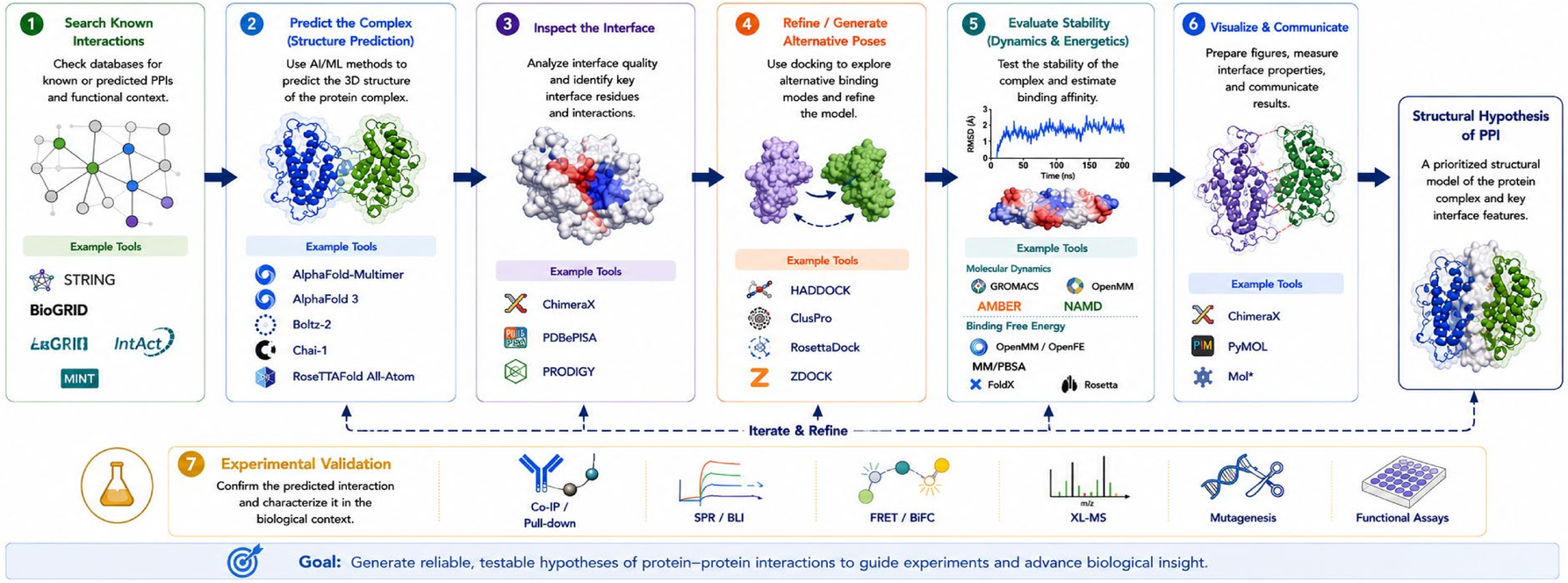
⭐ **ChimeraX**

- Visualize complexes
- Measure interfaces
- Identify hydrogen bonds and contacts
- Compare predicted vs. experimental structures

⭐ **PDBePISA** *(optional)*

- Calculate buried surface area
- Identify interface residues
- Predict biological assemblies

Computational Workflow for Predicting Protein–Protein Interactions (PPIs)

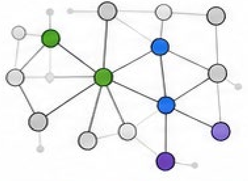


Question?

Does Bart's favorite protein, n-myristoyltransferase (NMT), have any binding partners?

1 Search Known Interactions

Check databases for known or predicted PPIs and functional context.



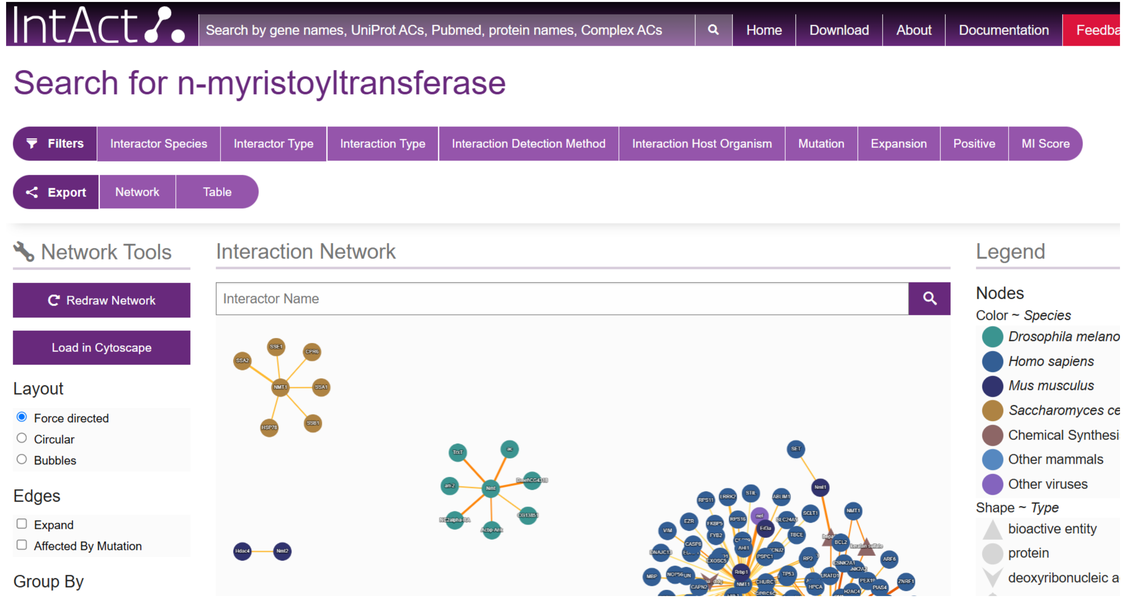
Example Tools

STRING

BioGRID

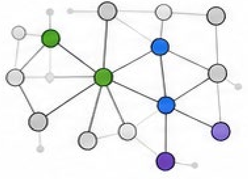
IntAct

MINT



1 Search Known Interactions

Check databases for known or predicted PPIs and functional context.



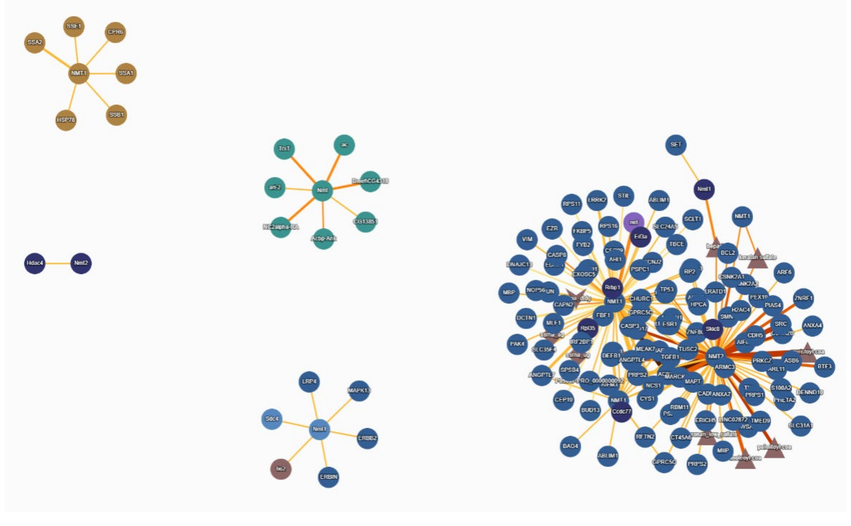
Example Tools

STRING

BioGRID

IntAct

MINT

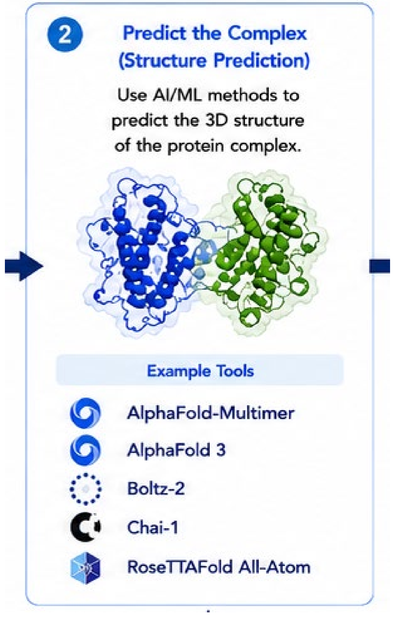


Showing 1 to 25 of 289 entries

1 2 3 4 5 … 12

Show 25 entries

Select	Molecule A	Molecule B	Identifier A	Identifier B	Type A	Type B	Species A	Species B	Host Organism	Positive interaction	Detection Method	Publication ID
☐	ACBD6	NMT2	UniProt Q9BR61	UniProt O60551	protein	protein	Homo sapiens	Homo sapiens	Homo sapiens HEK293 embryonic kidney cell	✓	anti tag coip	26621918
☐	ACBD6	NMT2	UniProt Q9BR61	UniProt O60551	protein	protein	Homo sapiens	Homo sapiens	Homo sapiens HEK293 embryonic kidney cell	✓	gal4 vp16 complement	26621918
☐	ACBD6	NMT2	UniProt Q9BR61	UniProt O60551	protein	protein	Homo sapiens	Homo sapiens	Homo sapiens HEK293 embryonic kidney cell	✓	pull down	26621918
☐	ACBD6	NMT2	UniProt Q9BR61	UniProt O60551	protein	protein	Homo sapiens	Homo sapiens	Homo sapiens HEK293T embryonic kidney cell	✓	anti tag coip	10.1016/j.cell.2021 33961781
☐	ACBD6	NMT2	UniProt Q9BR61	UniProt O60551	protein	protein	Homo sapiens	Homo sapiens	Homo sapiens U2OS osteosarcoma cell	✓	anti tag coip	40205054



ColabFold v1.6.2: AlphaFold2 using MMseqs2

Easy to use protein structure and complex prediction using AlphaFold2 and Alphafold2-multimer. Sequence alignments/templates are generated through MMseqs2 and HHsearch. For more details, see bottom of the notebook, checkout the ColabFold GitHub and Nature Protocols.

Old versions: v1.4, v1.5.1, v1.5.2, v1.5.3-patch

Mirdita M, Schütze K, Moriwaki Y, Heo L, Ovchinnikov S, Steinegger M. ColabFold: Making protein folding accessible to all. *Nature Methods*, 2022

Input protein sequence(s), then hit `Runtime` -> `Run all`

query_sequence

MADESETAVKPPAPPLPQMMEGNGNGHEHCSDCENEEDNSYNRGGLSPANDTGAKKKKKKQKKKKEKGSETDSAQDQPVKMNSLPAERIQEIQKAIELFSVGQGPAKTMEEASKRSY

- Use `:` to specify inter-protein chainbreaks for **modeling complexes** (supports homo- and hetro-oligomers). For example **PI...SK:PI...SK** for a homodimer

Sequence1:Sequence2

Sequence1:Sequence2

Input protein sequence(s), then hit `Runtime` -> `Run all`

query_sequence

MADESETAVKPPAPPLPQMMEGNGNGHEHCSDCENEEDNSYNRGGLSPANDTGAKKKKKKQKKKKEKGSETDSAQDQPVKMNSLPAERIQEIQKAIELFSVGQGPAKTMEEASKRSY

- Use `:` to specify inter-protein chainbreaks for **modeling complexes** (supports homo- and hetro-oligomers). For example **PI...SK:PI...SK** for a homodimer

jobname

nmt1 acbd6

num_relax

0

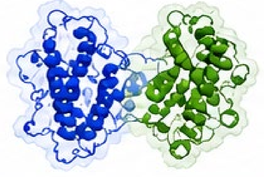
- specify how many of the top ranked structures to relax using amber

template_mode

none

2 Predict the Complex (Structure Prediction)

Use AI/ML methods to predict the 3D structure of the protein complex.



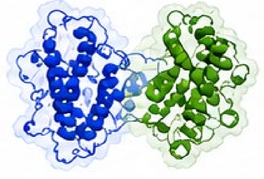
Example Tools

- AlphaFold-Multimer
- AlphaFold 3
- Boltz-2
- Chai-1
- RoseTTAFold All-Atom

```
Downloading alphafold2_multimer_v3 weights to .: 100%|██████████| 3.82G/3.82G [00:25<00:00, 160MB/s]
2026-07-15 17:39:03,038 Running on GPU
2026-07-15 17:39:03,512 Found 5 citations for tools or databases
2026-07-15 17:39:03,513 Query 1/1: nmt1acbd6_505d5_0 (length 778)
PENDING:   0%|          | 0/300 [elapsed: 00:00 remaining: ?]2026-07-15 17:39:03,816 Sleeping for 5s. Reason: PENDING
RUNNING:   2%|▏         | 5/300 [elapsed: 00:05 remaining: 05:28]2026-07-15 17:39:09,087 Sleeping for 9s. Reason: RUNNING
RUNNING:   5%|▍         | 14/300 [elapsed: 00:14 remaining: 05:01]2026-07-15 17:39:18,353 Sleeping for 7s. Reason: RUNNING
RUNNING:   7%|▋         | 21/300 [elapsed: 00:22 remaining: 04:51]2026-07-15 17:39:25,619 Sleeping for 10s. Reason: RUNNING
RUNNING:  10%|▉         | 31/300 [elapsed: 00:32 remaining: 04:39]2026-07-15 17:39:35,915 Sleeping for 8s. Reason: RUNNING
RUNNING:  13%|█▏        | 39/300 [elapsed: 00:40 remaining: 04:30]2026-07-15 17:39:44,186 Sleeping for 7s. Reason: RUNNING
RUNNING:  15%|█▍        | 46/300 [elapsed: 00:48 remaining: 04:33]2026-07-15 17:39:52,358 Sleeping for 5s. Reason: RUNNING
RUNNING:  17%|█▋        | 51/300 [elapsed: 00:54 remaining: 04:26]2026-07-15 17:39:57,626 Sleeping for 10s. Reason: RUNNING
RUNNING:  20%|█▉        | 61/300 [elapsed: 01:04 remaining: 04:11]2026-07-15 17:40:07,910 Sleeping for 8s. Reason: RUNNING
RUNNING:  23%|██▏       | 69/300 [elapsed: 01:12 remaining: 04:02]2026-07-15 17:40:16,218 Sleeping for 8s. Reason: RUNNING
RUNNING:  26%|██▍       | 77/300 [elapsed: 01:20 remaining: 03:52]2026-07-15 17:40:24,495 Sleeping for 5s. Reason: RUNNING
RUNNING:  27%|██▋       | 82/300 [elapsed: 01:26 remaining: 03:48]2026-07-15 17:40:29,768 Sleeping for 9s. Reason: RUNNING
RUNNING:  30%|██▉       | 91/300 [elapsed: 01:35 remaining: 03:37]2026-07-15 17:40:39,061 Sleeping for 9s. Reason: RUNNING
RUNNING:  33%|███▎      | 100/300 [elapsed: 01:44 remaining: 03:27]2026-07-15 17:40:48,323 Sleeping for 10s. Reason: RUNNING
RUNNING:  37%|███▋      | 110/300 [elapsed: 01:55 remaining: 03:16]2026-07-15 17:40:58,589 Sleeping for 10s. Reason: RUNNING
RUNNING:  40%|███▉      | 120/300 [elapsed: 02:05 remaining: 03:05]2026-07-15 17:41:08,886 Sleeping for 9s. Reason: RUNNING
RUNNING:  43%|████▎     | 129/300 [elapsed: 02:14 remaining: 02:56]2026-07-15 17:41:18,153 Sleeping for 8s. Reason: RUNNING
RUNNING:  46%|████▌     | 137/300 [elapsed: 02:22 remaining: 02:48]2026-07-15 17:41:26,422 Sleeping for 5s. Reason: RUNNING
RUNNING:  47%|████▋     | 142/300 [elapsed: 02:28 remaining: 02:43]2026-07-15 17:41:31,696 Sleeping for 7s. Reason: RUNNING
```

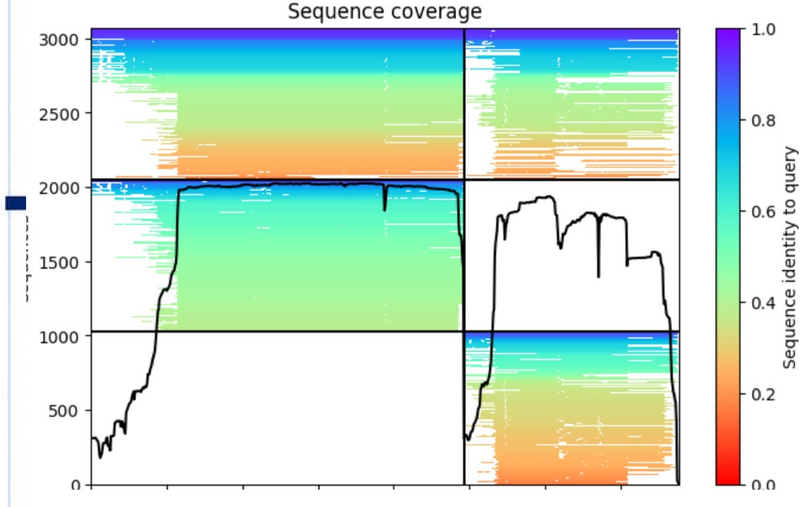
2 Predict the Complex (Structure Prediction)

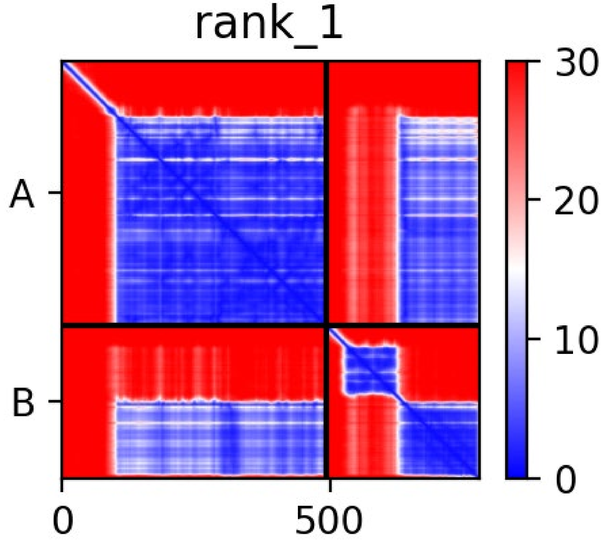
Use AI/ML methods to predict the 3D structure of the protein complex.

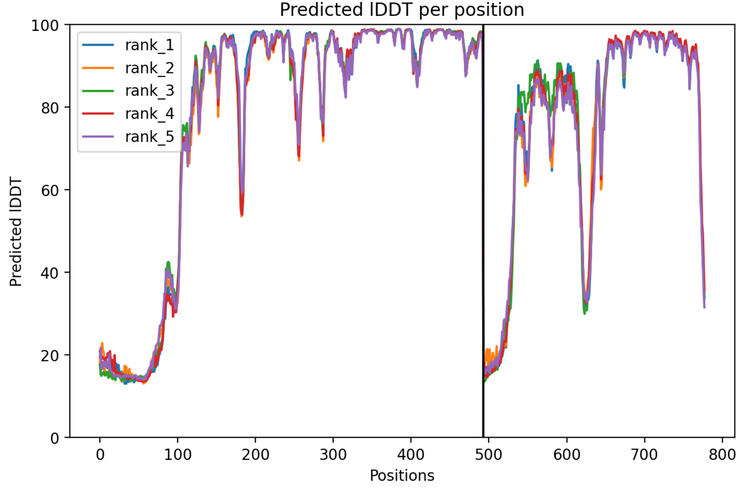


Example Tools

- AlphaFold-Multimer
- AlphaFold 3
- Boltz-2
- Chai-1
- RoseTTAFold All-Atom

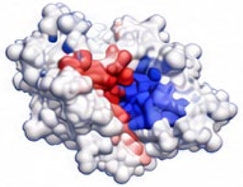






3 Inspect the Interface

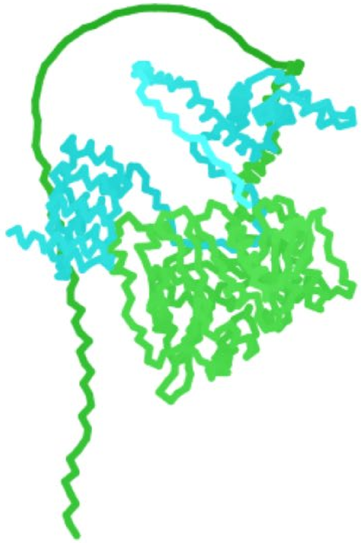
Analyze interface quality and identify key interface residues and interactions.



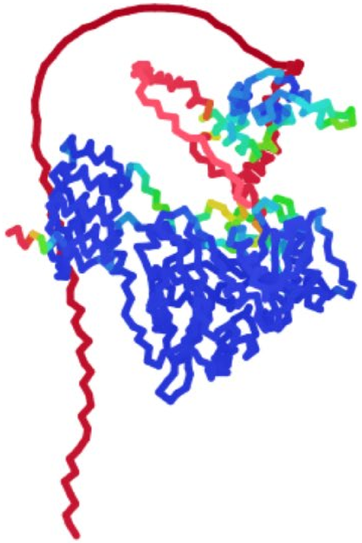
Example Tools

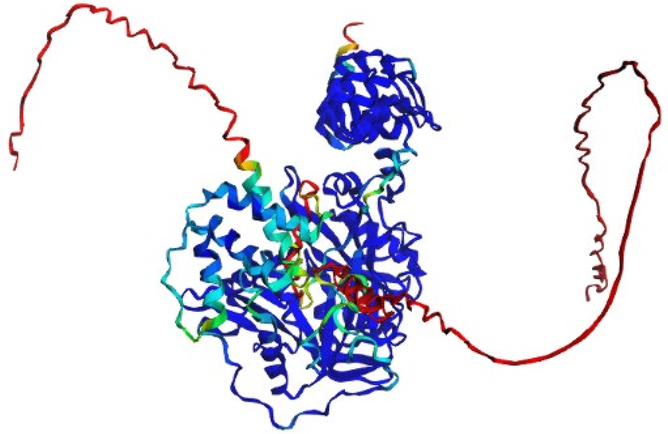
- ChimeraX
- PDBePISA
- PRODIGY

colored by chain



colored by pLDDT



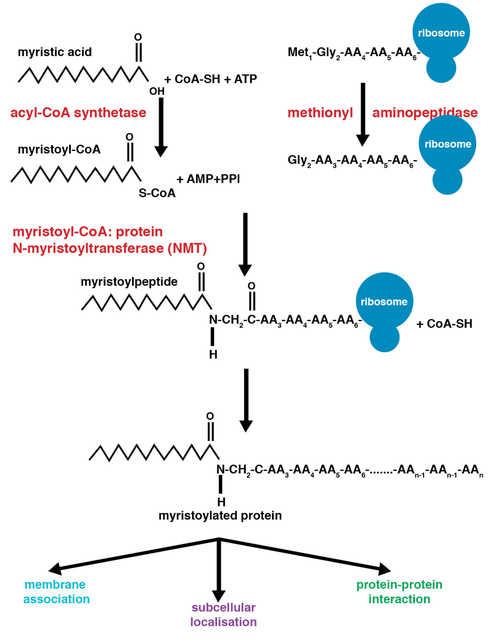


So, what is going on here?

- Ignored known structural information
- Ignored known biology
- Generated garbage

This protein is probably myristoylated by NMT?
***check sequence for N-terminal MGXXX**

```
7OWQ_1      ------------------------------------------------------------  0
P30419      MADESETAVKPPAPPLPQMMEGNGNGHEHCSDCENEEDNSYNRGGLSPANDTGAKKKKKK  60

7OWQ_1      --------------------------------GGSEFSVGQGPAKTMEEASKRSYQFW  26
P30419      QKKKKEKGSETDSAQDQPVKMNSLPAERIQEIQKAIELFSVGQGPAKTMEEASKRSYQFW  120
                                            . . **********************

7OWQ_1      DTQPVPKLGEVVNTHGPVEPDKDNIRQEPYTLPQGFTWDALDLGDRGVLKELYTLLNENY  86
P30419      DTQPVPKLGEVVNTHGPVEPDKDNIRQEPYTLPQGFTWDALDLGDRGVLKELYTLLNENY  180
            ************************************************************

7OWQ_1      VEDDDNMFRFDYSPEFLLWALRPPGWLPQWHCGVRVVSSRKLVGFISAIPANIHIYDTEK  146
P30419      VEDDDNMFRFDYSPEFLLWALRPPGWLPQWHCGVRVVSSRKLVGFISAIPANIHIYDTEK  240
            ************************************************************

7OWQ_1      KMVEINFLCVHKKLRSKRVAPVLIREITRRVHLEGIFQAVYTAGVVLPKPVGTCRYWHRS  206
P30419      KMVEINFLCVHKKLRSKRVAPVLIREITRRVHLEGIFQAVYTAGVVLPKPVGTCRYWHRS  300
            ************************************************************

7OWQ_1      LNPRKLIEVKFSHLSRNMTMQRTMKLYRLPETPKTAGLRPMETKDIPVVHQLLTRYLKQF  266
P30419      LNPRKLIEVKFSHLSRNMTMQRTMKLYRLPETPKTAGLRPMETKDIPVVHQLLTRYLKQF  360
            ************************************************************
```



Before the break

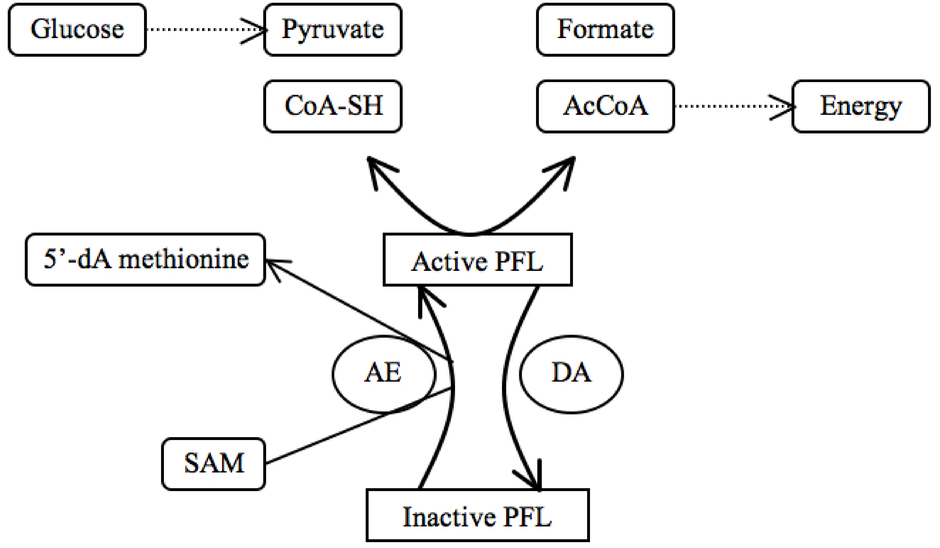
Let's start an alphafold job.

Question:
What is the likely multimeric state of Formate acetyltransferase (PFL) from Amygdalobacter nucleatus?

Formate acetyltransferase (PFL) from Amygdalobacter nucleatus

Amygdalobacter nucleatus:
A novel anaerobic bacterium discovered in 2023 directly from human vaginal microbiome samples using advanced anaerobic cultivation methods. For years, its genetic footprint was known as 'BVAB2' linked to reproductive health issues.

Formate acetyltransferase (PFL) is essential for anaerobic metabolism, makes a good drug target.

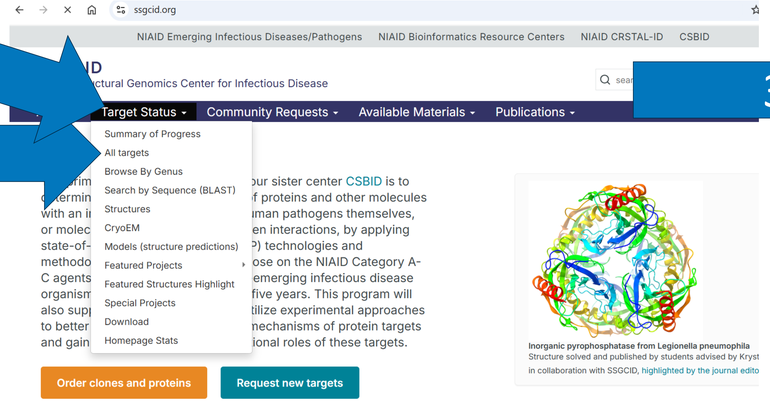


Go to www.ssgcid.org

Click on Target Status: All targets

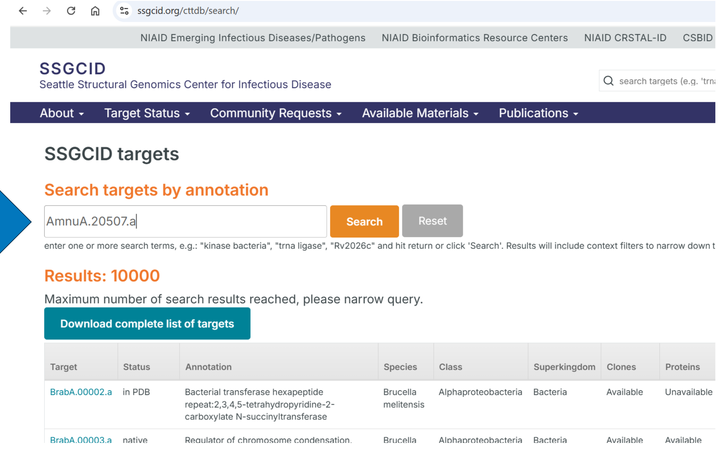
1

2



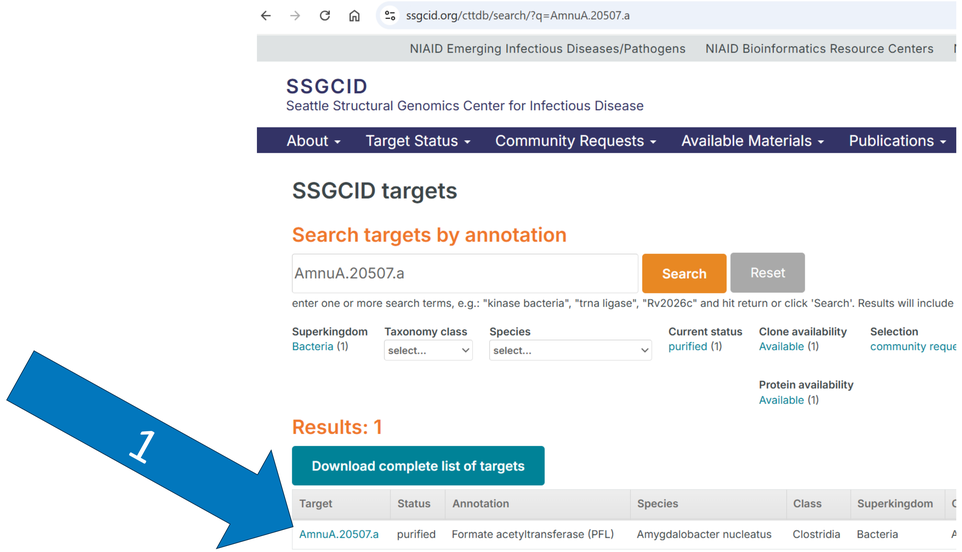
3

Type in AmnuA.20507.a into the search bar

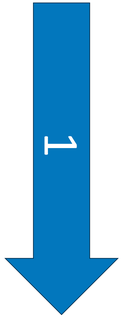


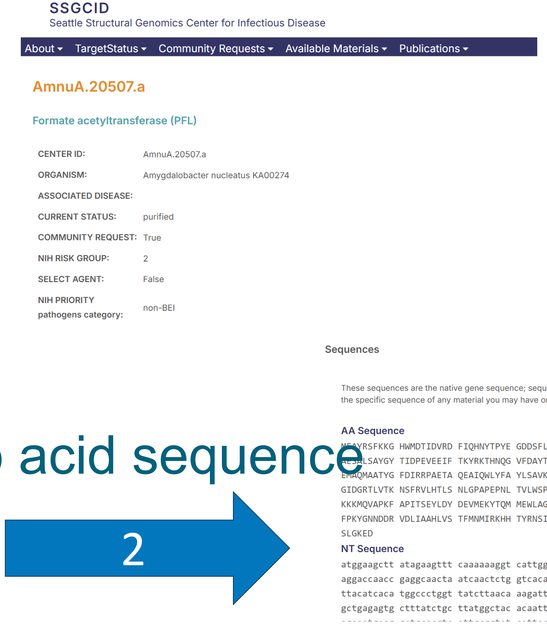
4

Click on AmnuA.20507.a



Keep scrolling



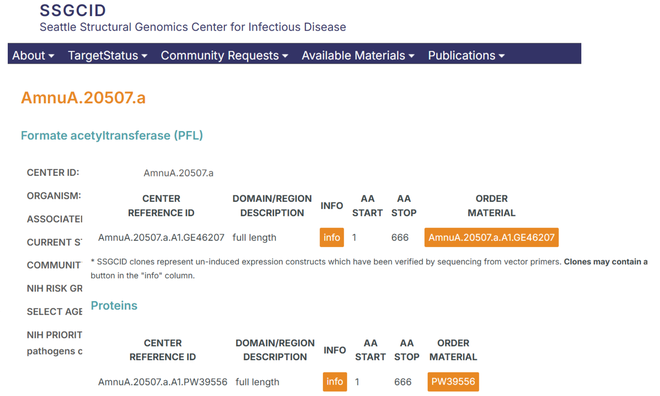


Copy the amino acid sequence

OR, get the UNIPROT Sequence and copy it

Keep scrolling

1



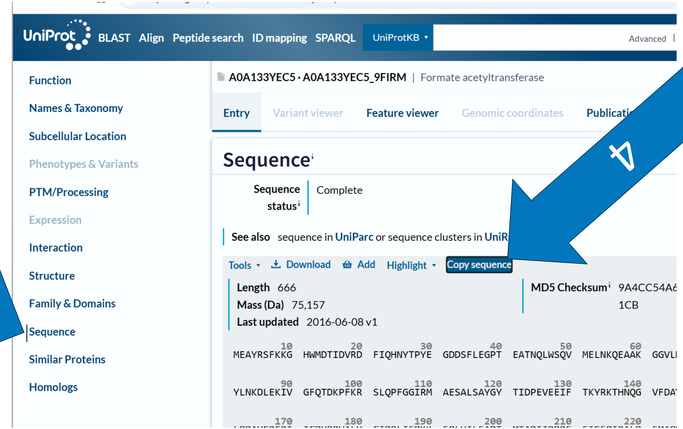
Click on Uniprot link

2

Use uniport ‘Copy Sequence’ to get rid of spaces

3

4



1

Go to alphafold server

https://alphafoldserver.com/

3

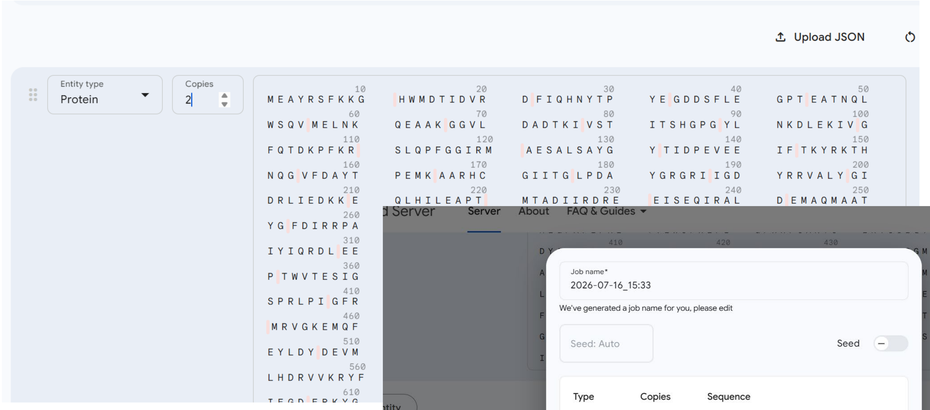
Paste sequence in

4

Choose 2,3,4,5,6 copies

5

Submit job

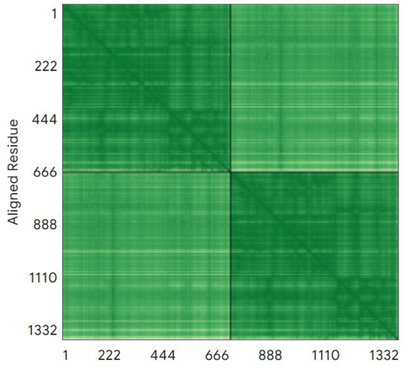


6

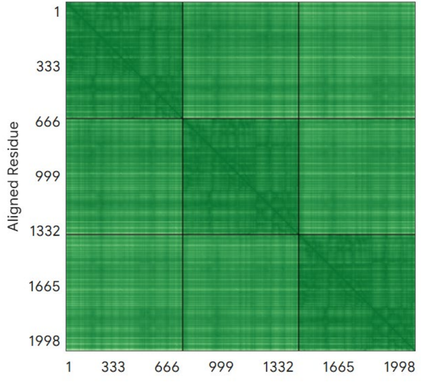
Get coffee

Coffee break

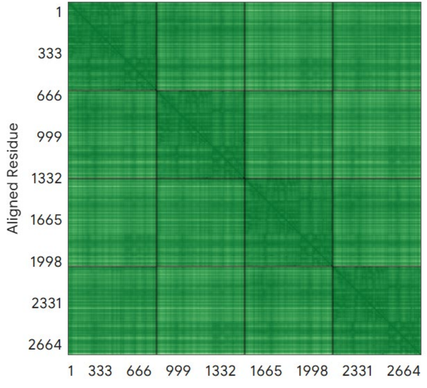
Let’s look at results



2

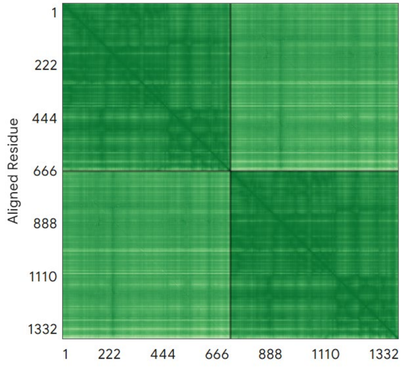


3

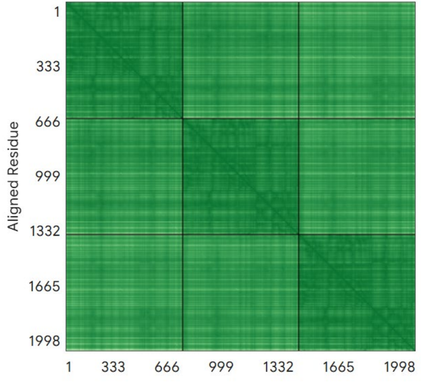


4

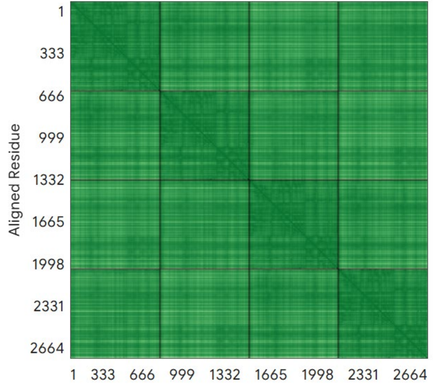
Let’s look at results



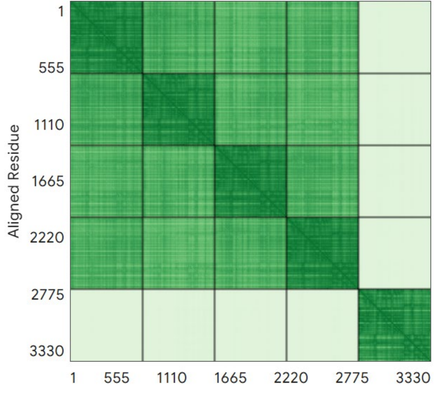
2



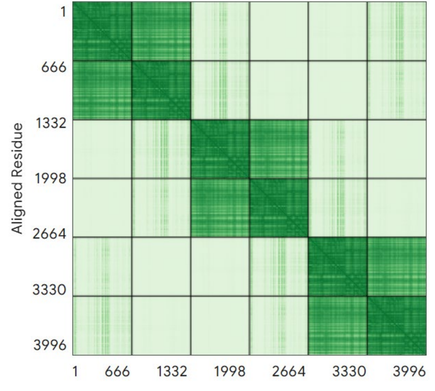
3



4



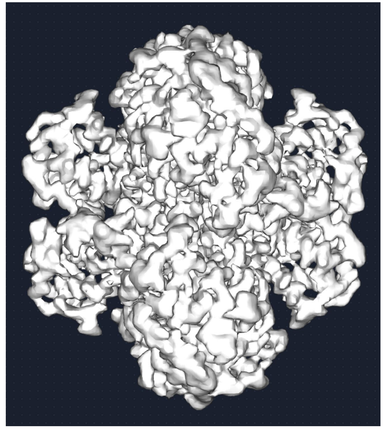
5

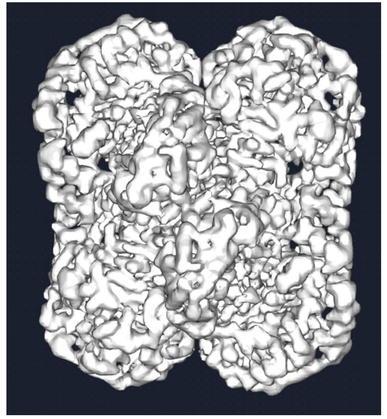


6

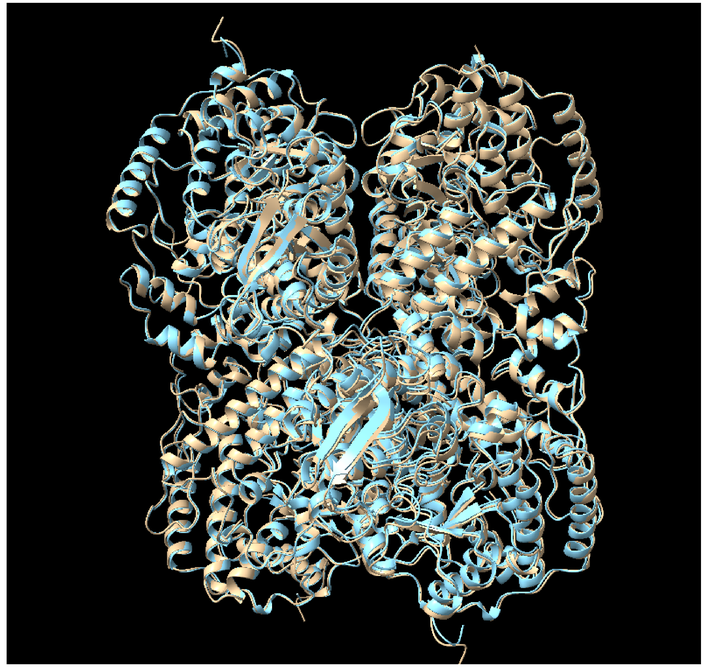
It’s a tetramer!

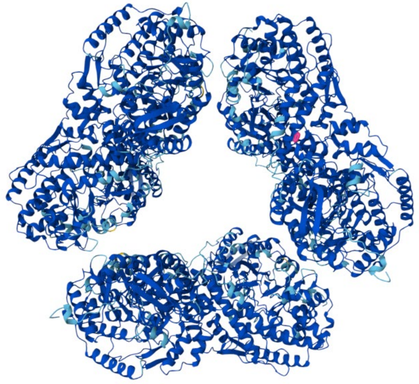
Both X-ray and Cryo-EM show a tetramer.



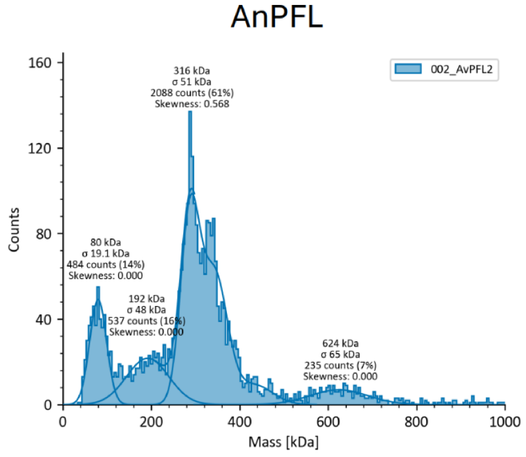


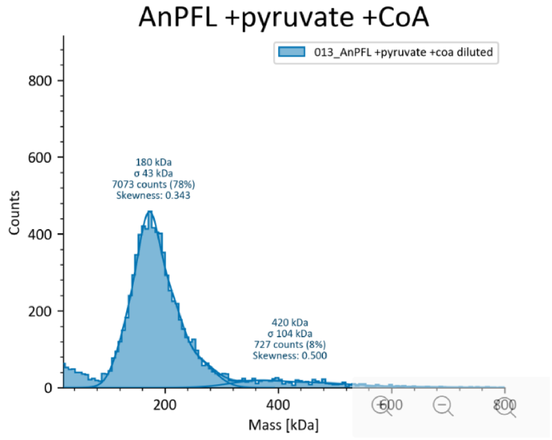
But…. Is it the tetramer predicted?

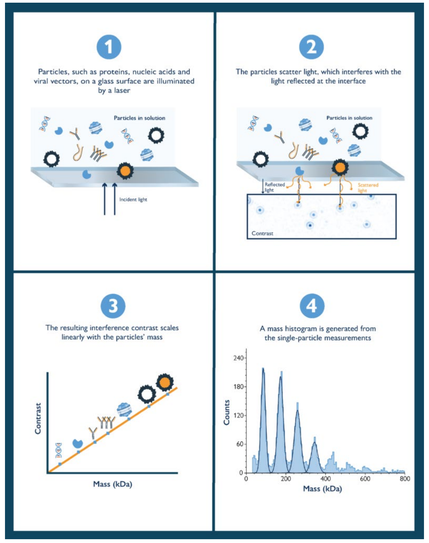




AmnuA.20507.a.A1







https://refeyn.com/post/how-does-mass-photometry-work

What have we learned?

Current computational methods of Protein-protein structures do not stand alone without biological data.

How do we make sure we can trust a prediction?

Build our skills in telling the difference.

Learn how to analyze prot-prot interactions.

Analyzing protein-protein interface

How do we analyze a protein-protein interface?

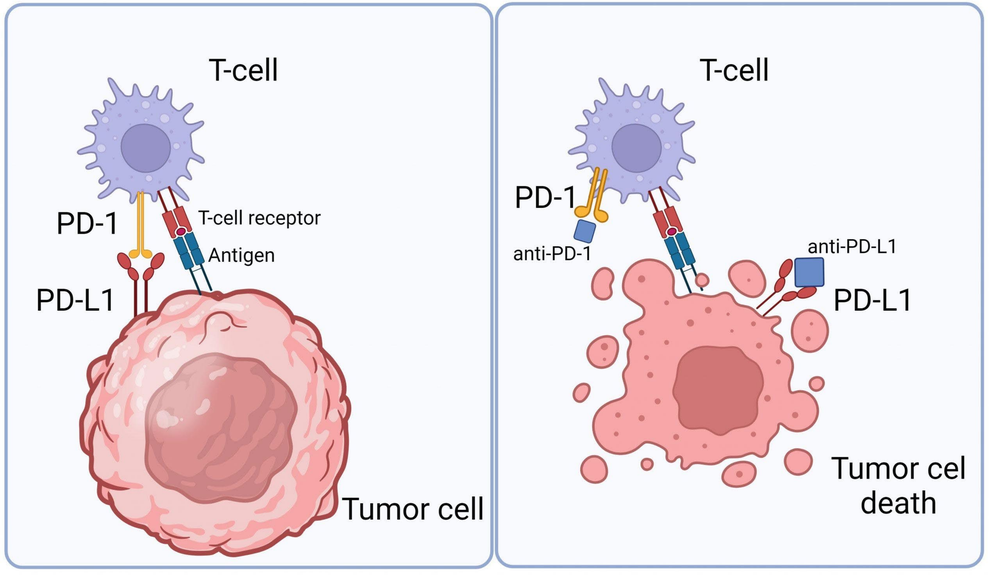
How do we know if a prediction makes sense?

Please open up ChimeraX

PD-1 / PD-L1: An Immune Checkpoint Protein-Protein Interaction

PD-1 (Programmed Death-1) is an inhibitory receptor expressed on activated T cells.
PD-L1 (Programmed Death Ligand-1) is expressed on antigen-presenting cells and many tumors.
Binding of PD-L1 to PD-1 sends an inhibitory signal that reduces T-cell activity.

Therapeutic strategy
Antibodies block either partner:
•**Anti-PD-1:** Pembrolizumab, Nivolumab
•**Anti-PD-L1:** Atezolizumab, Durvalumab, Avelumab
Blocking the interaction restores T-cell function and anti-tumor immunity.



Current market size

The global PD-1/PD-L1 inhibitor market is currently estimated at roughly:

- **2024:** ~$55 billion
- **2025: $60-65 billion**
- **2026: $67-75 billion** (most analyst estimates)
- Expected to exceed **$140-200+ billion by the early 2030s**, depending on assumptions about biosimilars, new indications, and combination therapies.

The Lancet recently estimated that **PD-1/PD-L1 inhibitors generated about $52 billion in worldwide sales in 2023**, making them the largest class of immune checkpoint inhibitors.

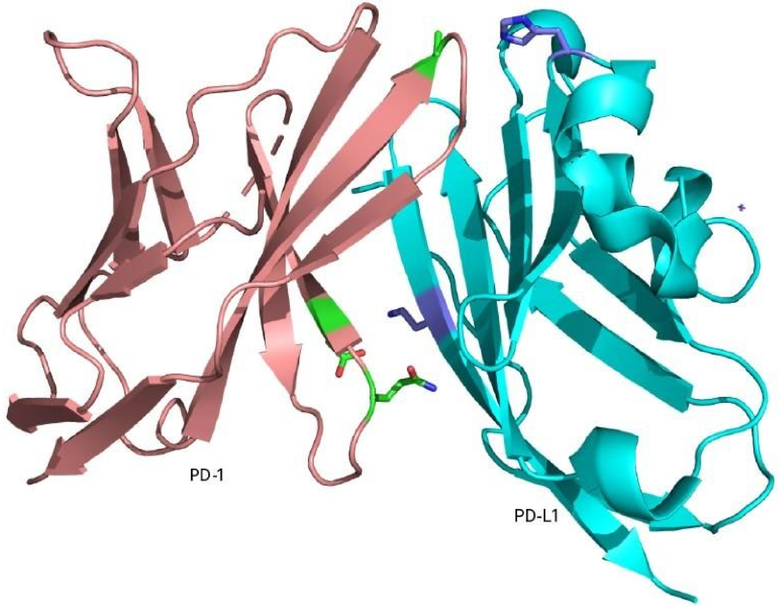
Drug	Target	Approximate annual sales
Keytruda (pembrolizumab)	PD-1	~$30B/year
Opdivo (nivolumab)	PD-1	~$10-12B/year
Tecentriq (atezolizumab)	PD-L1	~$4B/year
Imfinzi (durvalumab)	PD-L1	~$5B/year
Bavencio (avelumab)	PD-L1	<$1B/year
Libtayo (cemiplimab)	PD-1	~$1B/year

▸ eClinicalMedicine. 2025 May 24;84:103261. doi: 10.1016/j.eclinm.2025.103261

Learning Goal

By the end of the session, participants should be able to answer:

- identify the components of a protein complex;
- locate the protein-protein interface;
- identify interface residues;
- inspect hydrogen bonds and close contacts;
- measure buried surface area;
- propose mutations that could disrupt binding.



Exercise 1: Native Biology PD-1 : PD-L1 Question? How does the immune checkpoint naturally work?

Exercise 1: Analyze the PD-1:PD-L1 Interface in ChimeraX

Structure: PDB **4ZQK**, the 2.45 Å X-ray structure of the human PD-1:PD-L1 complex. In this entry, **chain A is PD-L1** and **chain B is PD-1**.

CONTENT B

Tasks:
Identify PD-1 and PD-L1 chains.
Color each protein.
Display the interface.
Measure buried surface area.
Count hydrogen bonds.
Identify hydrophobic patches.
Identify key interface residues.
Decide whether the interface appears "tight" or "loose.“

1. Open the structure

Launch **UCSF ChimeraX**.

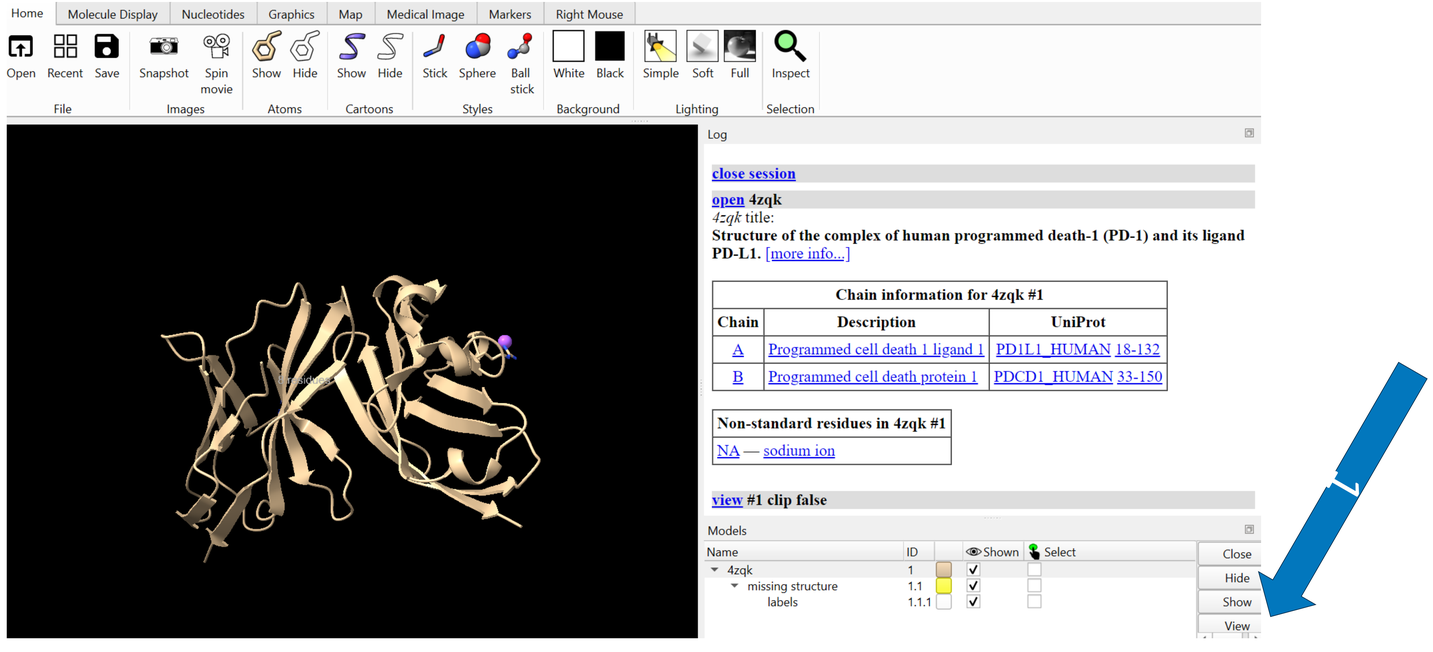
In the command line, enter:

```
open 4zqk
```

Checkpoint

Open the **Model Panel** or **Sequence Viewer** and confirm:

Chain	Protein
A	PD-L1
B	PD-1



2. Clean up the display

Hide atoms and show the proteins as cartoons:

```
hide atoms
cartoon
```

Remove solvent and the sodium ion:

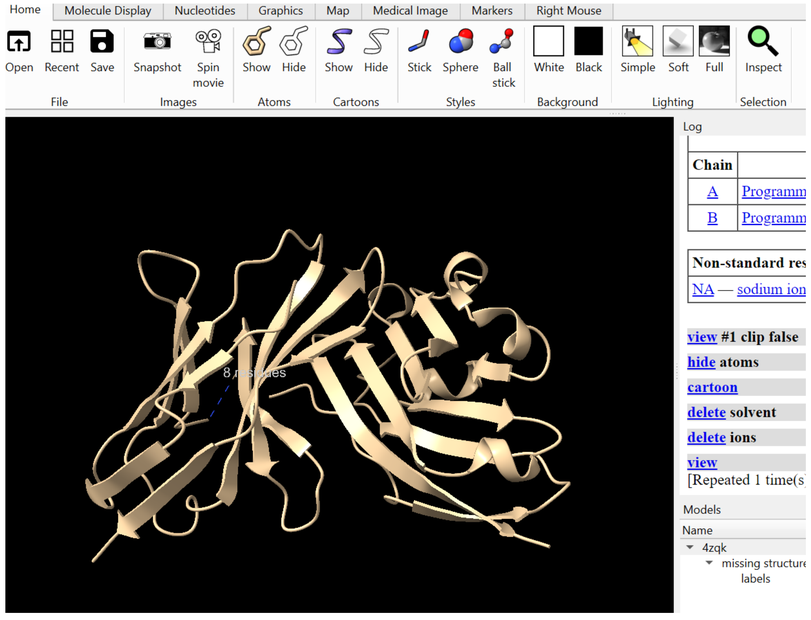
```
delete solvent
delete ions
```

Fit the complex to the window:

```
view
```

What should you see?

Two compact immunoglobulin-like domains contacting each other through broad opposing surfaces.



3. Color and label the proteins

Color PD-L1 blue and PD-1 pink:

```
color #1/A cornflowerblue
color #1/B hotpink
```

The model number is usually `#1`. Check the Model Panel if ChimeraX assigned another number.

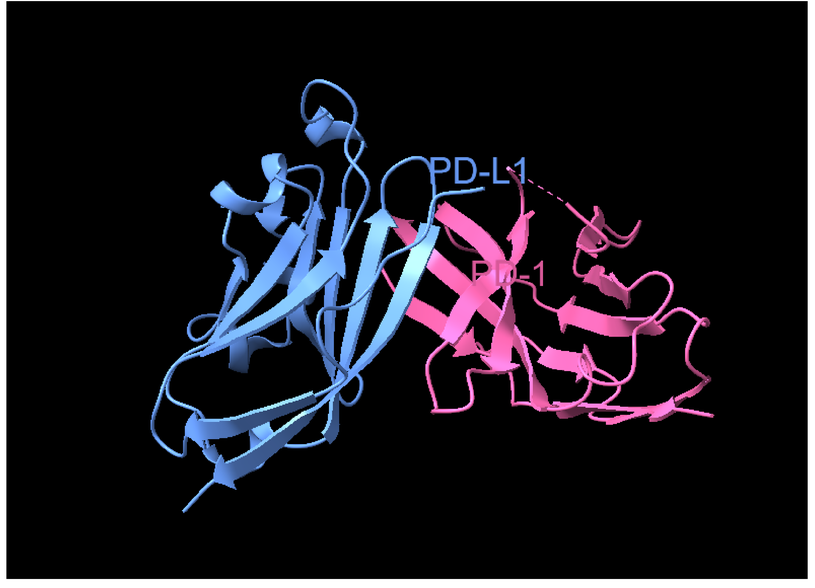
Add chain labels:

```
label #1/A models text "PD-L1"
label #1/B models text "PD-1"
```

Labels may land awkwardly, so omitting them is perfectly respectable. Molecular graphics sometimes behave like cats.

Optional: You can get rid of labels with

```
Label delete
```



4. Show the molecular surfaces

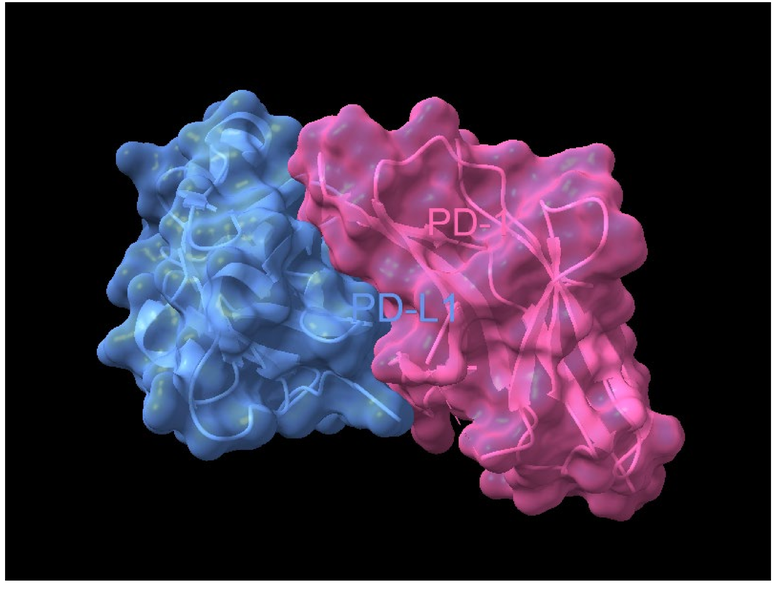
Display surfaces for both chains:

```
surface #1/A
surface #1/B
```

Make the surfaces partly transparent:

```
transparency #1/A 40 surfaces
transparency #1/B 40 surfaces
```

Rotate the complex and examine how the two surfaces fit together.



5. Identify interface residues from chain A

Select residues in PD-L1 within 4 Å of PD-1:

Select residues in chain A that have at least one atom within 4 Å of chain B

```
select #1/A & #1/B:<4
```

Show the selected atoms as sticks:

```
show sel atoms
style sel stick
```

Color them a contrasting color:

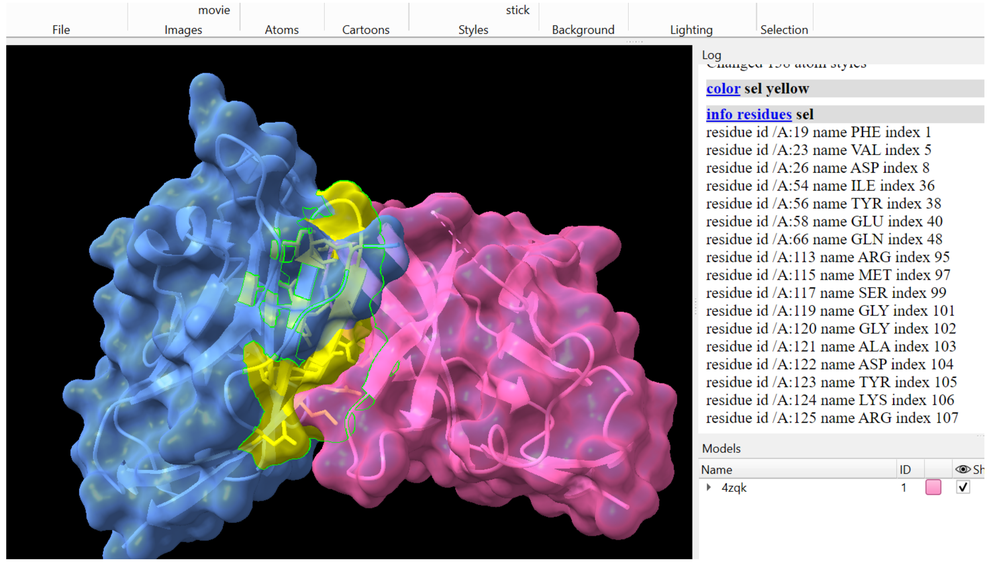
```
color sel yellow
```

Open:

Tools → General → Log

Then enter:

```
info residues sel
```



Alternative GUI method

Students can also use:

Tools → Structure Analysis → Contacts

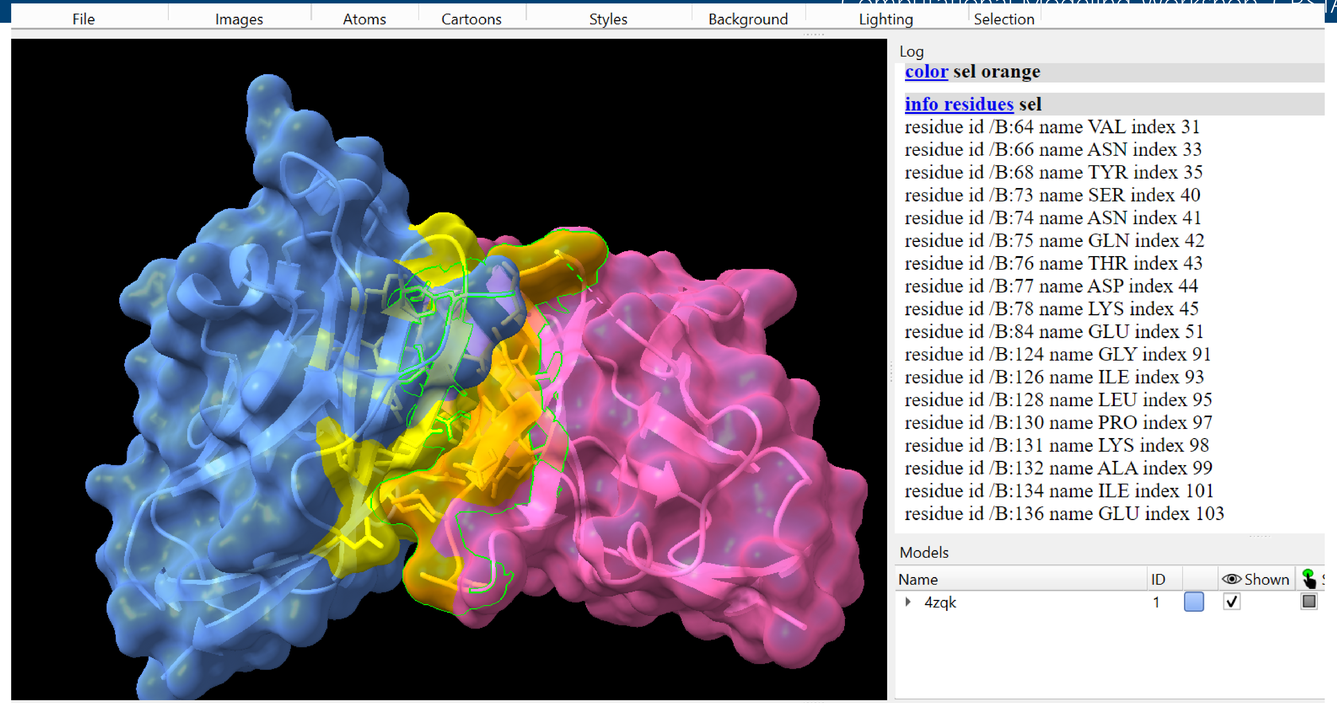
Specify chain A and chain B as the two interacting groups.

6. Identify interface residues from chain B

Now identify PD-1 residues within 4 Å of PD-L1:

```
select #1/B & #1/A:<4
show sel atoms
style sel stick
color sel orange
info residues sel
```

Students should record approximately **five important residues from each partner**.



7. Display inter-chain contacts

Clear the selection:

```
select clear
```

Show contacts between PD-L1 and PD-1:

```
contacts #1/A restrict #1/B
```

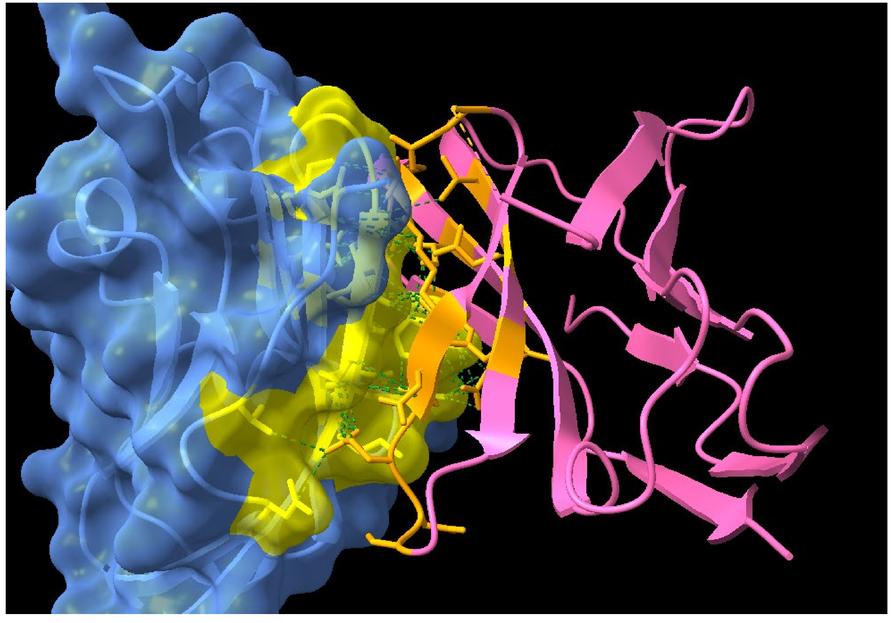
ChimeraX should display pseudobonds between atoms making close contacts.

Students should rotate the model and determine whether contacts are:

- evenly distributed;
- clustered in one hotspot;
- primarily backbone-mediated;
- primarily side-chain-mediated.

To remove the contact lines later:

```
~contacts
```



8. Identify hydrogen bonds

Display hydrogen bonds across the interface:

```
hbonds #1/A restrict #1/B
```

Hydrogen bonds should appear as pseudobonds.

For a clearer view, hide the surfaces temporarily:

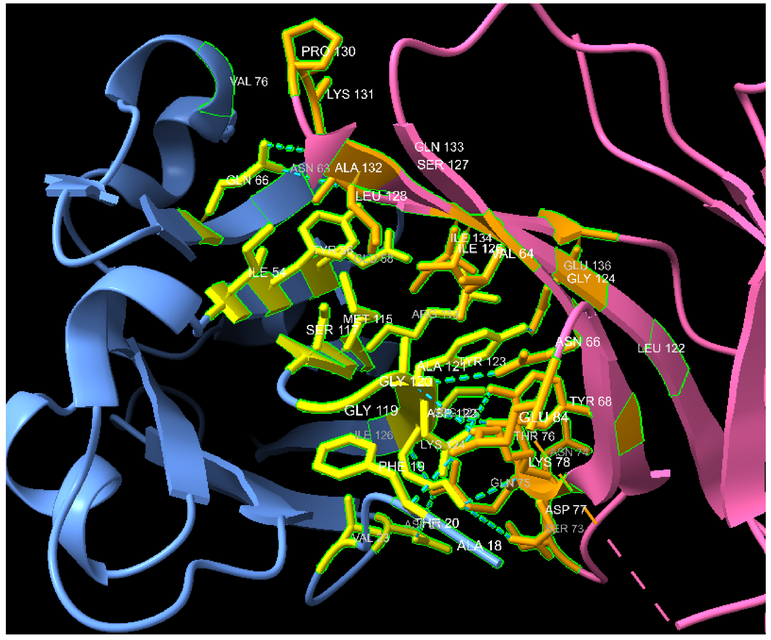
```
hide surfaces
```

Focus on the interface:

```
select (#1/A & #1/B:<5) | (#1/B & #1/A:<5)
view sel
```

Label residues participating in the interface:

```
Label sel residues
```



9. add hydrogens

For more detailed hydrogen-bond ana : :

```
addh
```

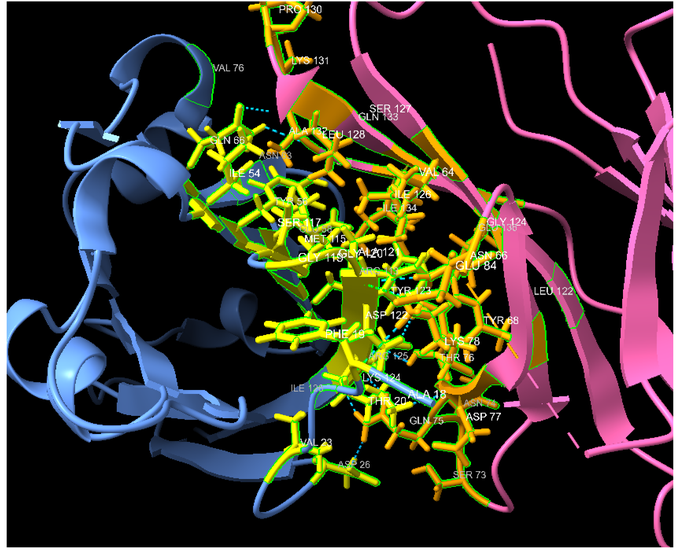
Then recalculate:

```
~hbonds
hbonds #1/A restrict #
```

Hydrogen positions in an X-ray ucture are lly modeled rather than directly observed at this resolution. The 4ZQK structure was determined at 2.45 Å resolution.

Better for calculations (usually) but makes it messy

Skip this slide



10. Measure buried surface area

Use ChimeraX's buried-area measurement:

```
measure buriedarea #1/A with #1/B
```

The result should appear in the Log.

Students should record:

- buried area contributed by PD-L1;
- buried area contributed by PD-1;
- total buried surface area.

hbonds #1/A restrict #1/B

The following atoms were skipped as donors/acceptors due to missing bond partners: /B ARG 139 NE

13 hydrogen bonds found

hide #1.1 models

hide #1.2 models

hide #1.3 models

measure buriedarea #1/A withAtoms2 #1/B

Buried area between #1/A and #1/B = 785.28

area #1/A = 6165.3, area #1/B = 5972.3, area both = 10567

Models

Name	ID	Shown	Select
4zqk	1	☑	☐
missing structure	1.1	☐	☐
4zqk_A SES surface	1.2	☐	☐
4zqk_B SES surface	1.3	☐	☐
hydrogen bonds	1.4	☑	☐

11. Examine interface chemistry

Show only interface residues:

```
hide atoms
select (#1/A & #1/B:<4) | (#1/B & #1/A:<4)
show sel atoms
style sel stick
```

Color atoms by element:

```
color sel byelement
```

Keep cartoons colored by chain:

```
color #1/A cornflowerblue cartoons
color #1/B hotpink cartoons
```

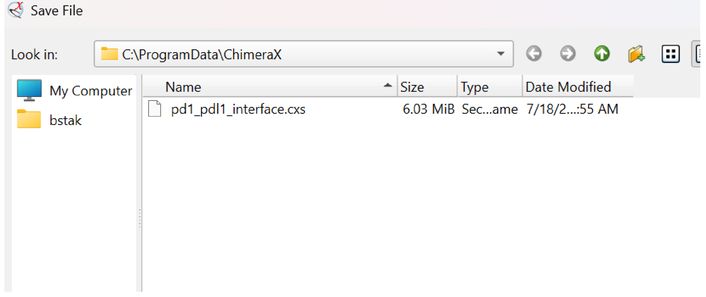
Save the session

Save the ChimeraX session:

```
save pd1_pdl1_interface.cxs
```

Save an image:

```
save pd1_pdl1_interface.png width 1800 height 1200
supersample 3
```



Let’s try virtual reality headsets!

`pd1_pdl1_interface.cxs`

12. Identify possible hotspot residues
choose:
•**one PD-L1 residue**, and
•**one PD-1 residue**
that you predict would substantially weaken binding if mutated to alanine.

Justify choices using at least two criteria:
•centrally located in the interface;
•makes multiple contacts;
•participates in hydrogen bonding;
•buries a large side chain;
•forms a salt bridge;
•supports local interface shape.

VR experience

Mutation proposal

Protein	Residue	Proposed mutation	Structural rationale
PD-L1		Ala	
PD-1		Ala	

13. Design an experimental test

How would you test whether the proposed mutation disrupts PD-1:PD-L1 binding?

13. Characterize the interface:

Feature	Observation
Hydrophobic residues	Low / moderate / extensive
Polar contacts	Low / moderate / extensive
Hydrogen bonds	Approximate number
Charged pairs	List one or two
Shape complementarity	Poor / moderate / strong
Interface distribution	Localized / broad

Exercise 2:
How does an antibody block the natural PD-1:PD-L1 interaction?

Structures:
- PDB **5XXY**: PD-L1 bound to the atezolizumab Fab
- PDB **4ZQK**: PD-1 bound to PD-L1

```
Close session
```

maybe quit and restart

Learning objectives

By the end of the exercise, students should be able to:
- identify the antigen, antibody heavy chain, and antibody light chain;
- locate and characterize an antibody-antigen interface;
- distinguish heavy-chain and light-chain contributions;
- compare an antibody epitope with the native PD-1 binding site;
- determine whether the antibody blocks PD-1 by direct competition or by an indirect mechanism;
- propose interface mutations and experimental validation strategies.

Part 1. Open the PD-L1–atezolizumab structure

Launch ChimeraX and enter:

```
open 5xxy
```

Fit the structure to the window:

```
view
```

The structure contains three protein chains:

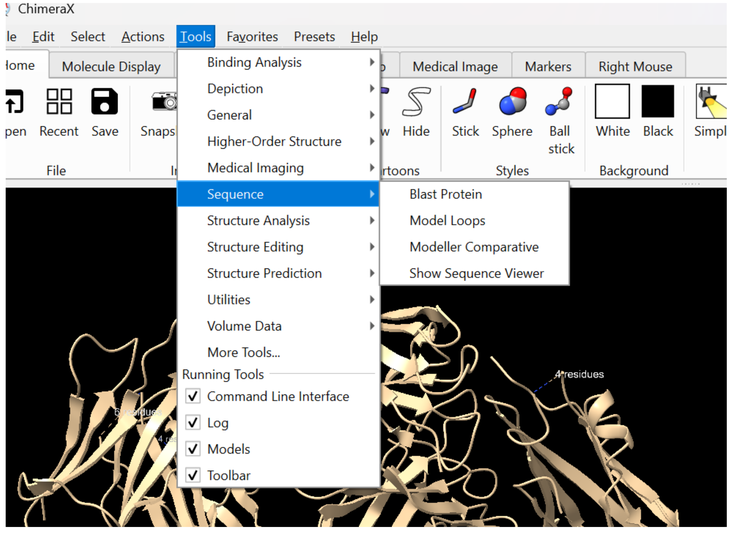
Checkpoint

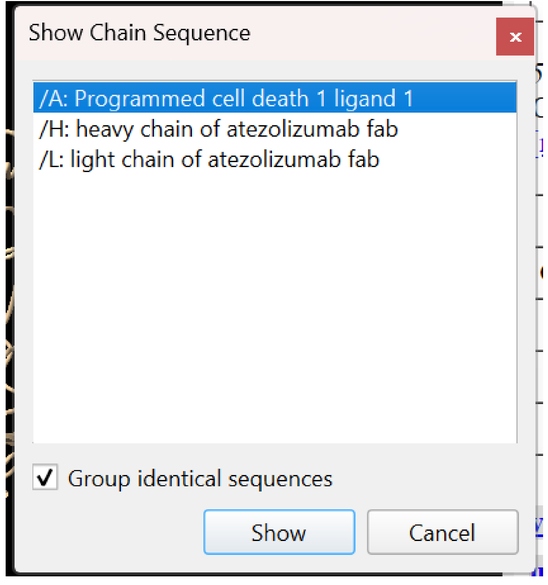
Open the Sequence Viewer:

Tools → Sequence → Sequence Viewer

Confirm that the model contains:

- one PD-L1 chain;
- one antibody heavy chain;
- one antibody light chain.





Part 2. Clean up the display

Hide atoms and show cartoons:

```
hide atoms
cartoon
```

Remove solvent and ions:

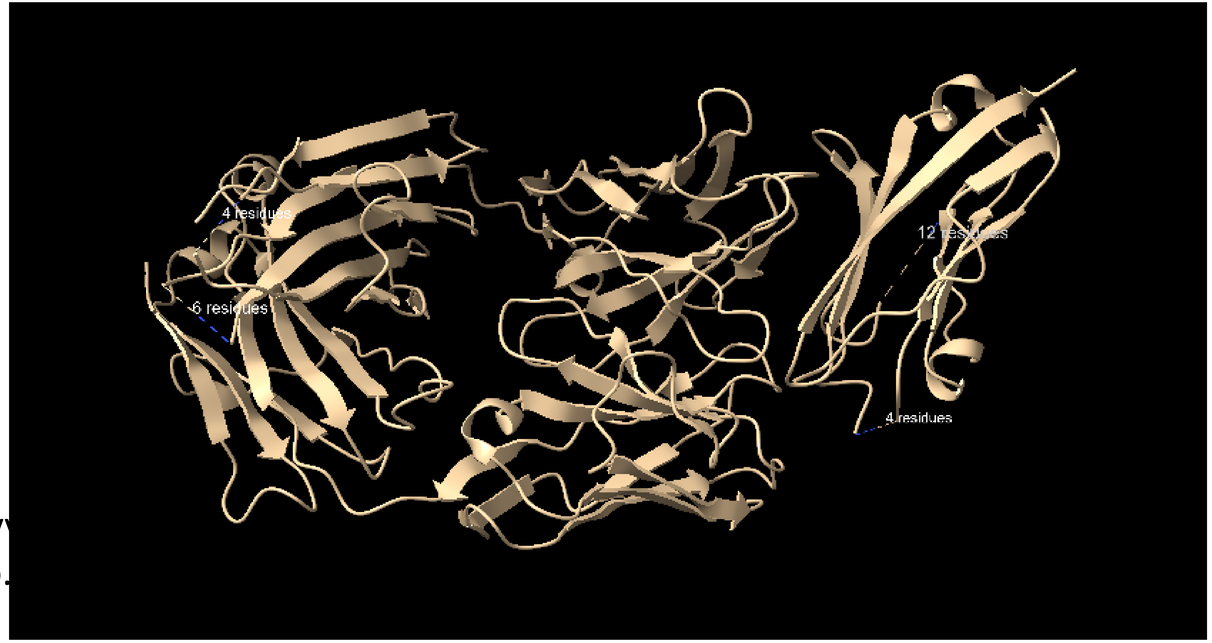
```
delete solvent
delete ions
```

Fit the model again:

```
view
```

What should you see?

The antibody Fab should resemble a two-armed clamp formed by the heav
and light chains, with PD-L1 contacting the variable-domain end of the Fab.



Part 3. Color the three components

Color the antibody heavy chain purple:

```
color #1/H purple
```

Color the antibody light chain cyan:

```
color #1/L cyan
```

Color PD-L1 gold:

```
color #1/A gold
```

If the model is not numbered `#1`, check the Model Panel and substitute the correct model number.

What should you see?

The antibody Fab should resemble a two-armed clamp formed by the heavy and light chains, with PD-L1 contacting the variable-domain end of the Fab.

Part 4. Show the antibody-antigen interface

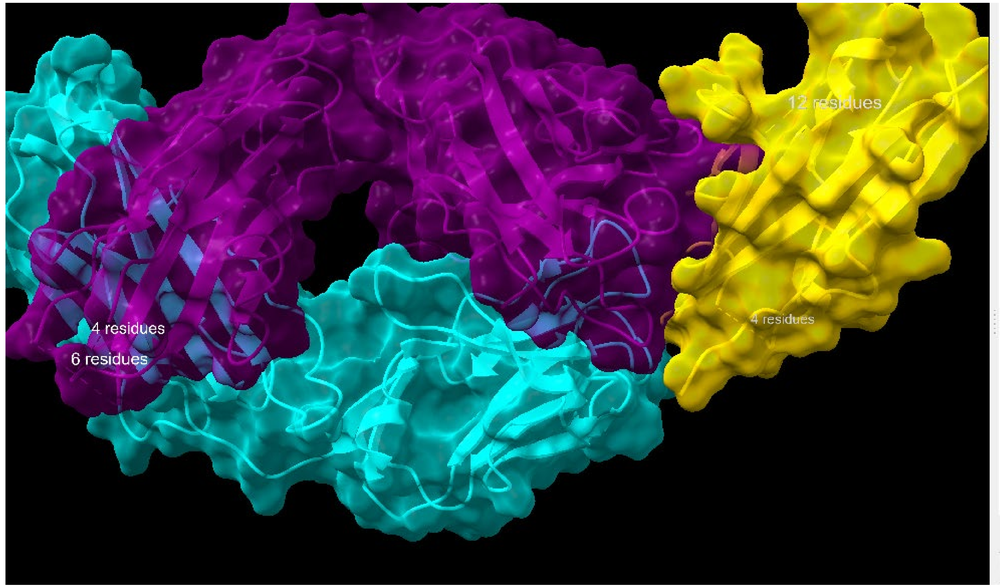
Display molecular surfaces:

```
surface #1/H
surface #1/L
surface #1/A
```

Make the surfaces partly transparent:

```
transparency #1/H 50 surfaces
transparency #1/L 50 surfaces
transparency #1/A 30 surfaces
```

Rotate the model and inspect the antibody-antigen contact.



See Content B

Part 5. Select PD-L1 residues contacting the antibody

Select PD-L1 residues within 4 Å of either antibody chain:

```
select #1/A & #1/H,L:<4
```

Show the selected atoms as sticks:

```
show sel atoms
style sel stick
```

Color the selected PD-L1 interface residues yellow:

```
color sel yellow
```

Now select antibody residues contacting PD-L1:

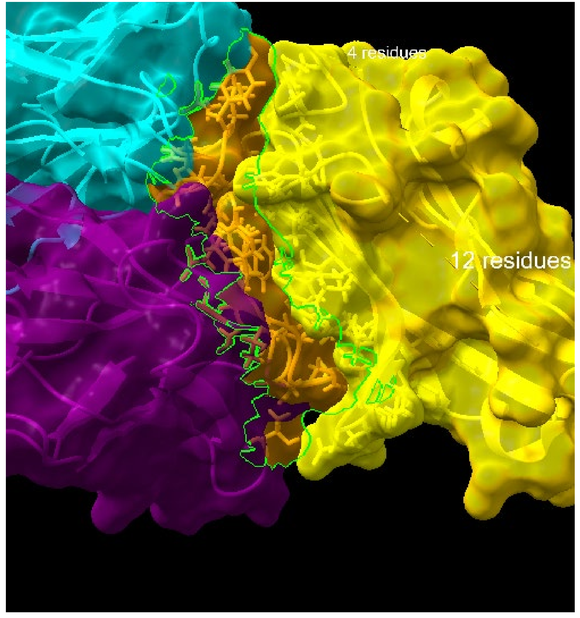
```
select #1/H,L & #1/A:<4
```

Show them as sticks:

```
show sel atoms
style sel stick
```

Color the antibody interface residues orange:

```
color sel orange
```



Part 6. Separate heavy-chain and light-chain contributions

First select heavy-chain residues within 4 Å of PD-L1:

```
select #1/H & #1/A:<4
```

List the residues in the Log:

```
info residues sel
```

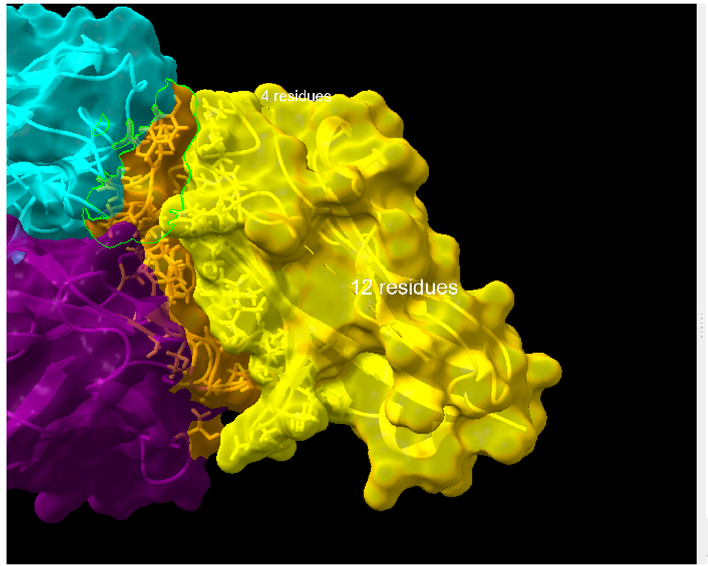
Record the number and identities of the contacting heavy-chain residues.

Now select light-chain residues within 4 Å of PD-L1:

```
select #1/L & #1/A:<4
```

List those residues:

```
info residues sel
```



residue id /H:56 name GLY index 55
residue id /H:57 name SER index 56
residue id /H:58 name THR index 57
residue id /H:59 name TYR index 58
residue id /H:74 name THR index 73
residue id /H:99 name ARG index 98
residue id /H:101 name TRP index 100
residue id /H:102 name PRO index 101
residue id /H:103 name GLY index 102

select #1/L & #1/A:<4
103 atoms, 105 bonds, 6 residues, 1 model selec

info residues sel
residue id /L:32 name ALA index 31
residue id /L:91 name TYR index 90
residue id /L:92 name LEU index 91
residue id /L:93 name TYR index 92
residue id /L:94 name HIS index 93
residue id /L:95 name PRO index 94

Models

See Content B

Part 7. Display close contacts

Clear the current selection:

```
select clear
```

Display contacts between PD-L1 and both antibody chains:

```
contacts #1/A restrict #1/H,L
```

Hide the surfaces temporarily for a clearer view:

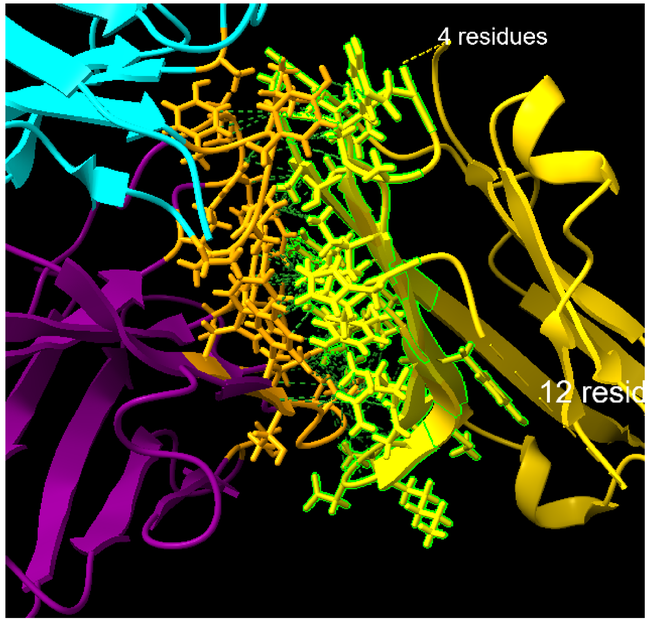
```
hide surfaces
```

Focus on the interface:

```
select #1/A & #1/H,L:<5
view sel
```

Remove contact pseudobonds when finished:

```
~contacts
```



See Content B

Part 8. Identify hydrogen bonds

Display hydrogen bonds between PD-L1 and the antibody:

```
hbonds #1/A restrict #1/H,L
```

```
hbonds #1/A restrict #1/H,L
```

Label the interface residues:

```
select #1/A & #1/H,L:<4
label sel residues
select #1/H,L & #1/A:<4
label sel residues
```

Rotate the structure and inspect the network.

See Content B

Part 9. Measure buried surface area

Measure the buried area between PD-L1 and the entire Fab:

```
measure buriedarea #1/A with #1/H,L
```

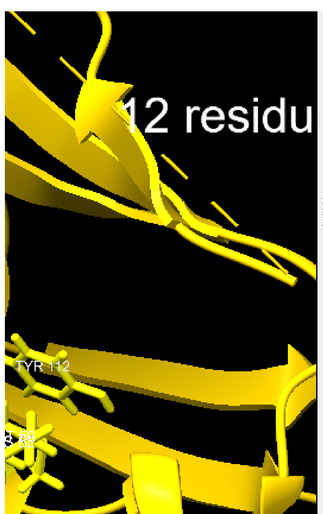
Record the result from the Log.

Now measure the heavy-chain contribution separately:

```
measure buriedarea #1/A with #1/H
```

Measure the light-chain contribution:

```
measure buriedarea #1/A with #1/L
```



360 atoms, 364 bonds, 23 residues, 1 model selected

label sel residues

measure buriedarea #1/A withAtoms2 #1/H,L
Buried area between #1/A and #1/H,L = 1016
area #1/A = 6296.4, area #1/H,L = 18894, area both = 23158

measure buriedarea #1/A withAtoms2 #1/H
Buried area between #1/A and #1/H = 823.01
area #1/A = 6296.4, area #1/H = 11049, area both = 15699

measure buriedarea #1/A withAtoms2 #1/L
Buried area between #1/A and #1/L = 260.83
area #1/A = 6296.4, area #1/L = 11141, area both = 16916

Models

Name	ID	Shown	Select
5xxy	1	✓	
missing structure	1.1	✓	

See Content B

Part 10. Examine interface chemistry

Show only the interface residues:

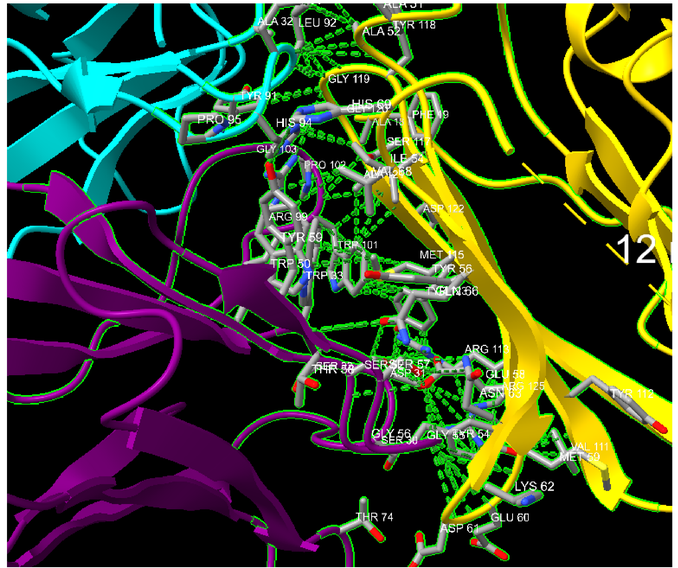
```
hide atoms
select (#1/A & #1/H,L:<4) | (#1/H,L & #1/A:<4)
show sel atoms
style sel stick
```

Color atoms by element:

```
color sel byelement atoms
```

Restore the chain colors on the cartoons:

```
color #1/H purple cartoons
color #1/L cyan cartoons
color #1/A gold cartoons
hide @H*
```



See Content B

Part 11. Open the native PD-1 complex

Keep 5XXY open and load the native complex:

```
open 4zqk
```

You should now have two models:

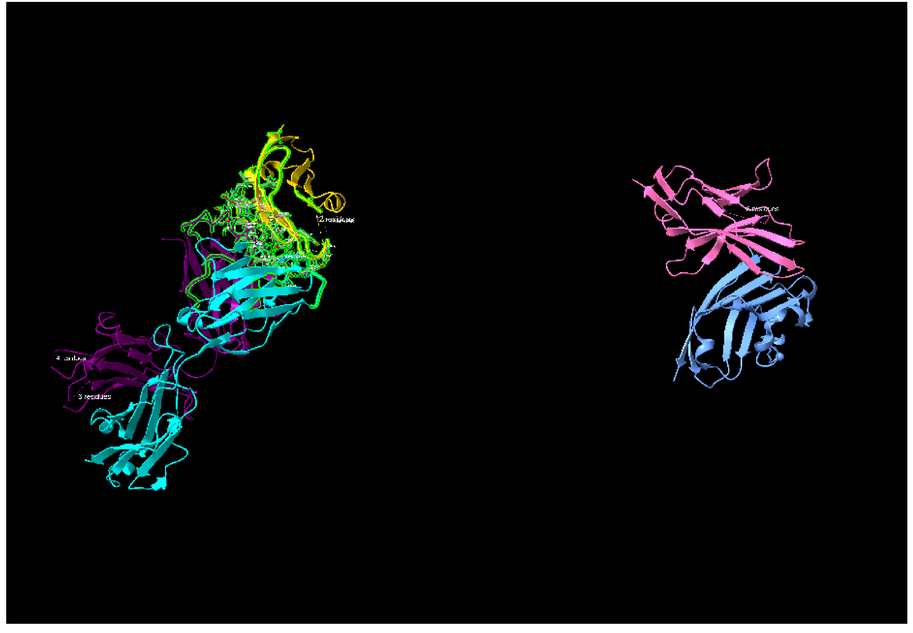
- `#1`: PD-L1–atezolizumab;
- `#2`: PD-1–PD-L1.

Check the Model Panel if your model numbering differs.

For 4ZQK:

Color the native complex:

```
color #2/A cornflowerblue
color #2/B hotpink
```



Part 12. Superimpose the two PD-L1 structures

Align PD-L1 from the antibody complex with PD-L1 from the native complex:

```
matchmaker #1/A to #2/A
```

Fit both models to the window:

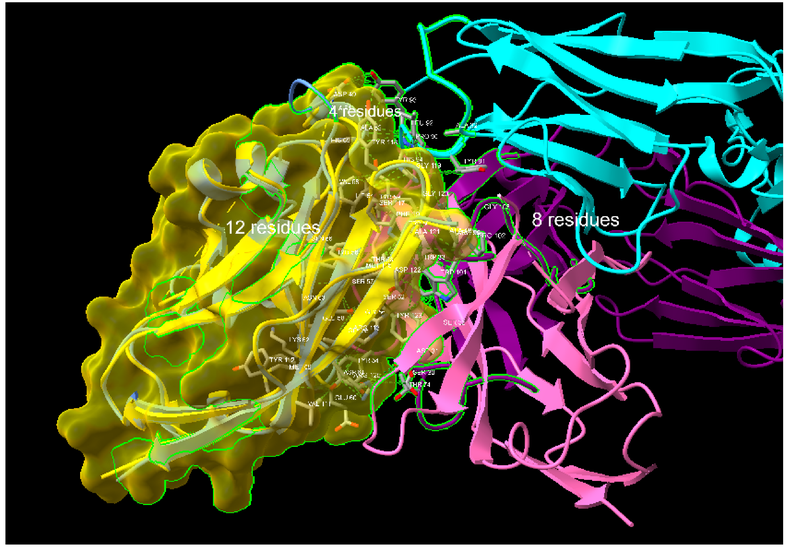
```
view
```

Hide the PD-L1 chain from one structure to reduce clutter:

```
hide #2/A cartoon
```

Or; Alternatively, keep both PD-L1 molecules visible and make one transparent.

```
surface #1/A
transparency #1/A 60 surfaces
```



Part 13. Determine whether the binding sites overlap

Show PD-1, PD-L1, and the antibody as cartoons:

```
show #1/H,L,A cartoons
show #2/B cartoons
```

Use the following colors:

```
color #1/H purple
color #1/L cyan
color #1/A gold
color #2/B hotpink
```

Rotate the superimposed complexes.

See Content B

Part 14. Compare the two PD-L1 epitopes

Select PD-L1 residues contacting atezolizumab:

```
select #1/A & #1/H,L:<4
```

Color those residues yellow:

```
color sel yellow
```

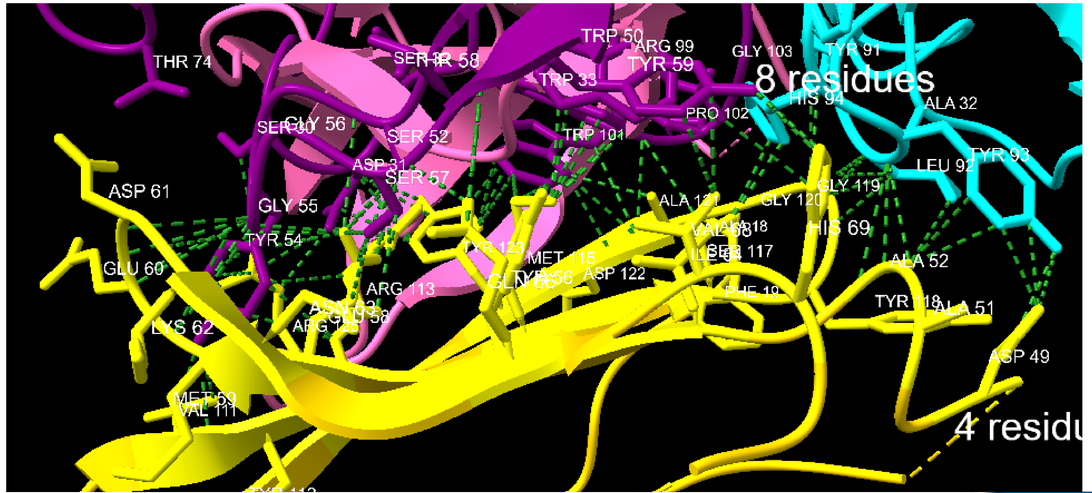
Select PD-L1 residues contacting PD-1 in the aligned native complex:

```
select #2/A & #2/B:<4
```

Color those residues lime:

```
color sel blue
```

Because the two PD-L1 structures are superimposed, the colored regions can be compared spatially.



See Content B

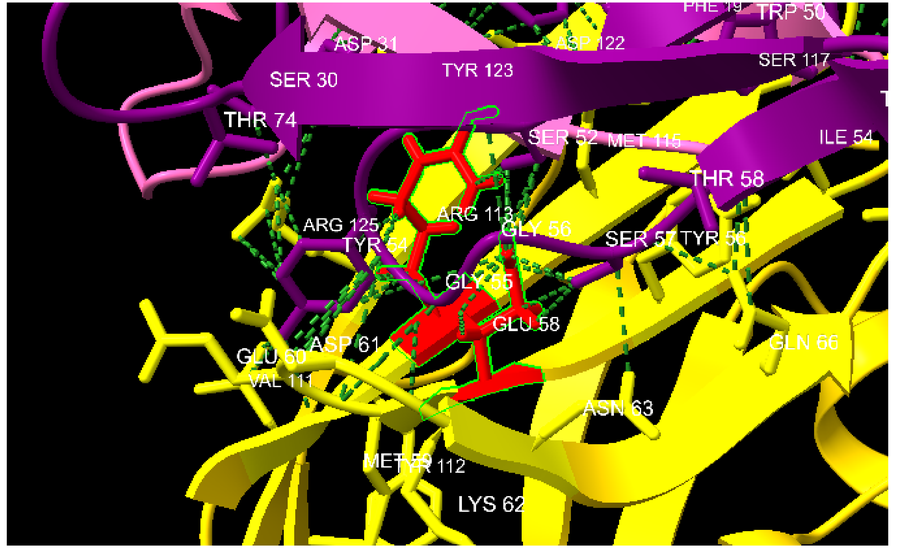
Part 15. Identify candidate hotspot residues

The structural study highlighted PD-L1 residues **Glu58 and Arg113** as important for atezolizumab binding.

Locate them:

```
select #1/A:58,113
show sel atoms
style sel stick
color sel red
label sel residues
```

Inspect their positions within the antibody interface.



Part 16. Propose mutations

Each group should propose:

- one PD-L1 mutation predicted to weaken atezolizumab binding;
- one antibody heavy-chain mutation predicted to weaken binding;
- optionally, one antibody light-chain mutation.

Students should avoid choosing a mutation solely because it is close to the antigen. The choice should be supported by structural evidence.

Mutation table

Protein	Residue	Proposed mutation	Structural rationale
PD-L1			
Antibody heavy chain			
Antibody light chain			

Part 18. Save images and the session

Save the ChimeraX session:

```
save exercise2_atezolizumab_pdl1.cxs
```

Save an image of the antibody-antigen complex: (you will have to adjust the view yourself)

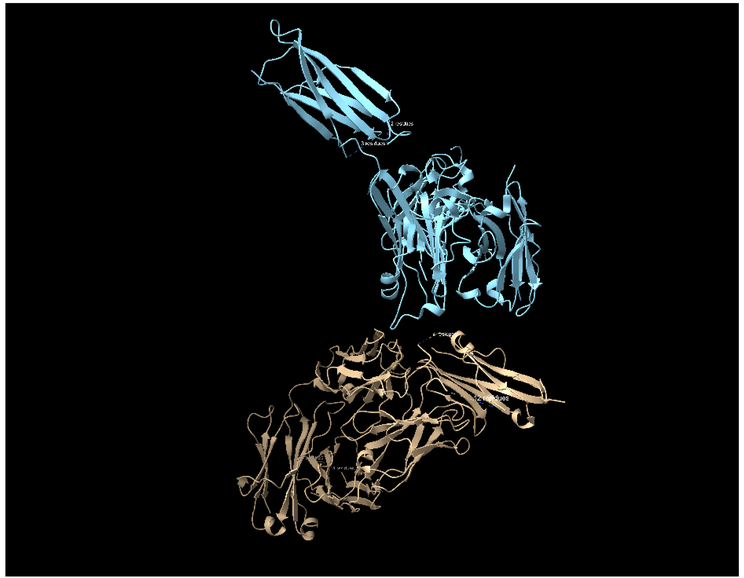
```
save atezolizumab_pdl1_interface.png width 1800
height 1200 supersample 3
```

Save an image of the PD-1 and antibody overlay:

```
save pdl1_epitope_comparison.png width 1800 height
1200 supersample 3

Close session
Restart chimera if there is time
```

Exercise 3: Comparing Two Therapeutic Antibodies
Estimated time: 35–45 minutes
Structures
•PDB **5XXY**: PD-L1 bound to atezolizumab
•PDB **5XJ4**: PD-L1 bound to durvalumab

Part 1. Open Both Structures
Open both complexes:
```
open 5xxy
open 5xj4
```
Fit both structures:
```
view
```
Use the Model Panel to determine the model numbers.



See Content B

Part 2. Color the Structures

Assign a different color scheme to each antibody.

Example:

Model 1 (Atezolizumab)

```
color #1/H purple
color #1/L cyan
color #1/A gold
```

Model 2 (Durvalumab)

```
color #2/H forestgreen
color #2/L skyblue
color #2/A tan
```

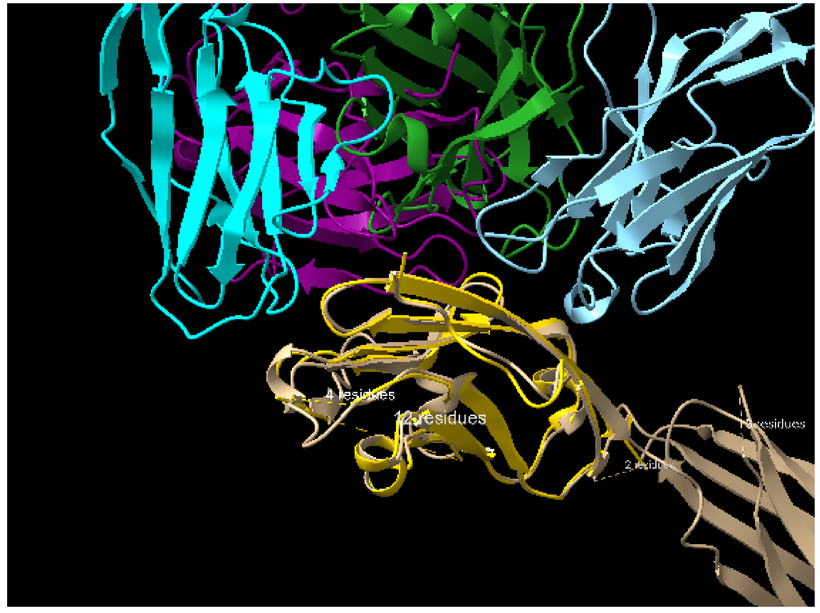
Part 3. Superimpose PD-L1

Align the PD-L1 molecules.

```
matchmaker #2/A to #1/A
```

Fit the models.

```
View
```



Part 4. Compare Overall Binding Orientation

Rotate the model.

Play with buttons

Observe:

- antibody approach angle;
- heavy-chain position;
- light-chain position;
- overall orientation of the Fab.

Questions

Does one antibody approach more vertically?

Does one approach from the side?

Which appears to cover more surface?

See Content B

Part 5. Compare the Epitopes

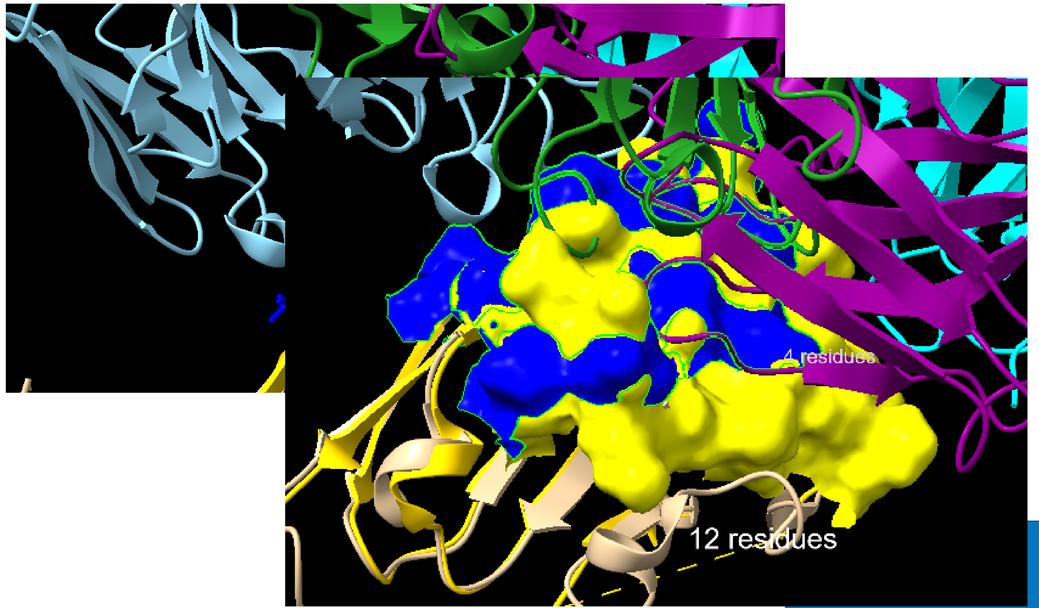
Identify interface residues for each antibody.

#For atezolizumab:

```
select #1/A & #1/H,L:<4
color sel yellow
show sel atoms
style sel stick
Surface sel
Sel clear
```

#For durvalumab:

```
select #2/A & #2/H,L:<4
color sel blue
show sel atoms
style sel stick
Surface sel
Sel clear
```



Content B

Part 6. Compare Buried Surface Area

Make PD-L1 semi-transparent.

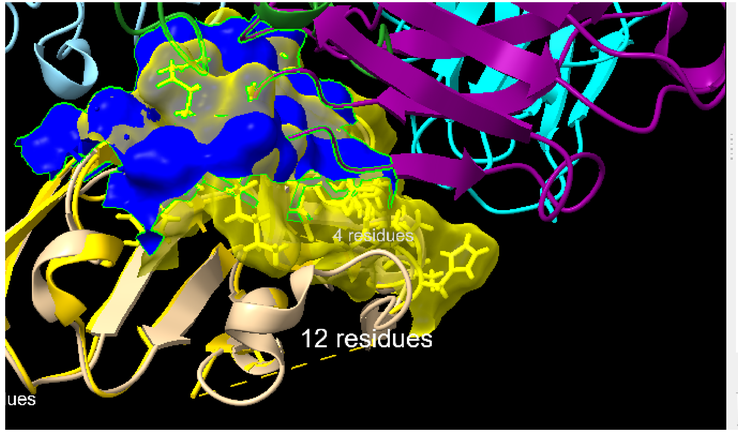
```
transparency #1/A 50 surfaces
```

#Measure the interface.

```
measure buriedarea #1/A with #1/H,L
```

#Repeat for Model 2.

```
measure buriedarea #2/A with #2/H,L
```



Changed 160 atom styles

Sel clear

Unknown command: Sel clear

transparency #1/A 50 surfaces

measure buriedarea #1/A withAtoms2 #1/H,L

Buried area between #1/A and #1/H,L = 1016

area #1/A = 6296.4, area #1/H,L = 18933, area both = 23198

measure buriedarea #2/A withAtoms2 #2/H,L

Buried area between #2/A and #2/H,L = 886.56

area #2/A = 12037, area #2/H,L = 10753, area both = 21016

Models

Name	ID	Shown	Select

Part 7. Compare Hydrogen Bonds

#Display hydrogen bonds.

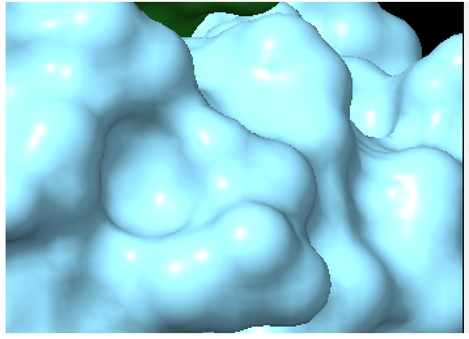
```
hbonds #1/A restrict #1/H,L
```

#Repeat for Model 2.

```
hbonds #2/A restrict #2/H,L
```

Estimate:

- number of hydrogen bonds;
- distribution;
- dominant chain.



Buried area between #2/A and #2/H, area #2/A = 12037, area #2/H,L = 1(

hbonds #1/A restrict #1/H,L

7 hydrogen bonds found

hbonds #2/A restrict #2/H,L

28 hydrogen bonds found

Models

Name ID

Part 8. Compare Interface Chemistry

#Show interface residues as sticks.

```
show sel atoms
style sel stick
```

#Color by element.

```
color sel byelement
```

#clear selection

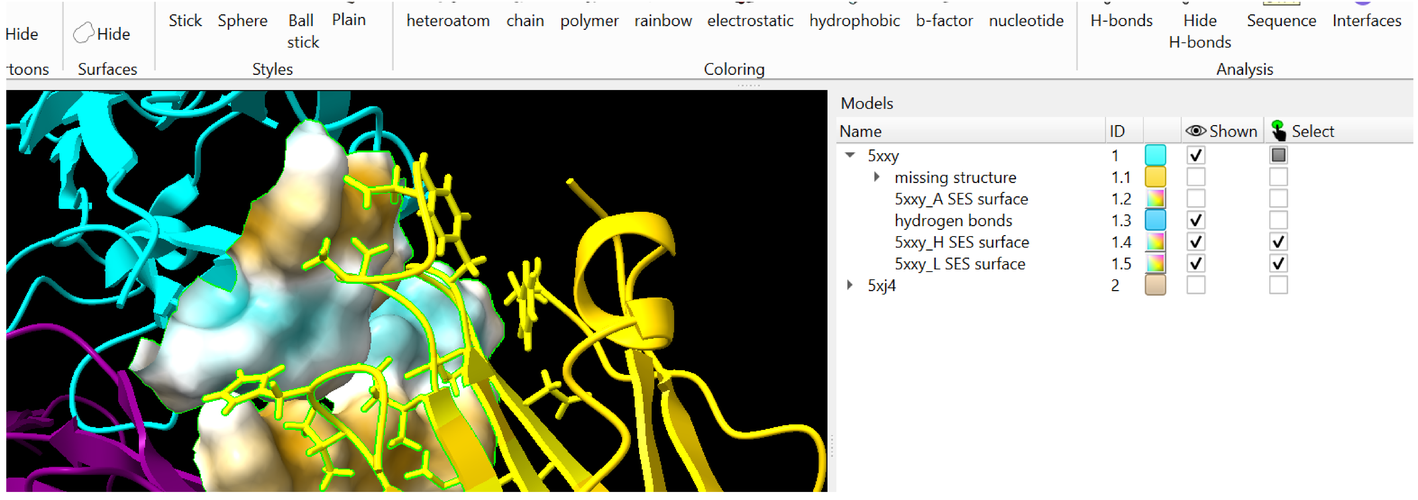
```
Sel clear
```

#selects A residues within 4 ang of H,L

```
select #1/A & #1/H,L:<4
```

#selects H,L residues within 4 ang of A

```
select #1/H,L & #1/A:<4
```



Part 9. Which Would You Engineer?

Imagine you are designing a 60-residue minibinder.
Which antibody provides the better structural template?
Support your answer using:

- interface size;
- hotspot residues;
- surface complementarity;
- complexity of the interface.

If time permits

Look at Ex 2 or 3 chimera files

Learning Objectives

By the end of this session participants will be able to:

- Predict candidate protein-protein interactions using modern computational tools.
- Critically evaluate confidence in predicted interactions.
- Analyze interfaces and identify likely binding determinants.
- Prioritize experimental validation strategies.
- Understand current capabilities and limitations of AI-based structural biology.

The end

Understanding Predictions

lorem

What Does AlphaFold Actually Predict?

Some content of the slide

Prediction Pipeline

Some content of the slide

Understanding pLDDT

Some content of the slide

Understanding PAE

Some content of the slide

Understanding ipTM

Some content of the slide

Looking Beyond the Scores

Some content of the slide

Inspecting Interfaces

Some content of the slide

Common Failure Modes

Some content of the slide