

TØ 15 – AI og modellering

Table of contents

1	Opgave 1. AlphaFold vs. Eksperimentel struktur - Confidence og fleksibilitet	2
1.1	Analysér B-faktorerne i strukturen	2
1.2	Beregn RMSD og analysér alignment	2
1.3	Fortolk pLDDT-værdier	2
1.4	Sammenlign B-factor og pLDDT	3
1.5	Sammenlign loopkonformationer	3
1.6	Forklar fjernelse af flexibelt loop	3
2	Opgave 2. Hvad fortæller PAE om domæne-orientering	4
2.1	Fortolk PAE-plottet	4
2.2	Identificér domæner i PAE	4
2.3	Analysér Parkins domænestruktur	4
2.4	Identificér det fleksible loop	4
3	Opgave 3. Modellering af et protein-kompleks med ligand og ioner	5
3.1	Fortolk pTM- og iPTM-værdier	5
3.2	Analysér pLDDT og PAE i kompleks	5
3.3	Find og sammenlign nukleotidbinding	5
3.4	Analysér calciumions koordination	5
4	Opgave 4. Protein-interaktion med Alphafold	6
4.1	Kortlæg interaktioner med AlphaFold	6
4.2	Beskriv kompleksets arkitektur	6
4.3	Fortolk pTM og iPTM-scorer	6
4.4	Design eksperiment med trunkeerede konstrukter	6

1 Opgave 1. AlphaFold vs. Eksperimentel struktur - Confidence og fleksibilitet

I denne opgave skal I lave en AlphaFold-forudsigtelse selv og sammenligne den med en eksperimentel krystalstruktur af det samme protein.

Protein: Humant Parkin er et 465 residue protein med 5 domæner:

- **Ubl (Ubiquitin-like) domain:** residues 1-76
- **RING0 domain:** residues 142-237
- **RING1 domain:** residues 238-293
- **IBR (In-Between-RING) domain:** residues 294-378
- **RING2 domain:** residues 389-449

Strukturforudsigtelse med Alphafold 3

- Gå til AlphaFold3 serveren: <https://alphafoldserver.com/>
- Opret en bruger eller log ind
- Indsæt sekvensen for humant Parkin (UniProt: O60260)
- Kør forudsigtelsen, det tager et par minutter (hold vinduet åbent og start på næste del)

Hent krystalstruktur af Parkin (1.8Å opløsning)

Åben krystalstrukturen af Parkin i PyMOL: **PDB-kode** 5C1Z (`fetch 5c1z, type=pdb1` for at hente den biologiske enhed). Farv strukturen efter B-factor.

Brug

```
spectrum b, blue_cyan_green_yellow_orange_red, 5c1z, minimum=20, maximum=100
```

1.1 Analysér B-faktorerne i strukturen

Hvad siger B-faktoren om strukturen? Hvilke regioner af strukturen have lav versus høj B-faktor? Giver det mening?

Tip

Hint: klik på enkelte atomer i PyMOL for at se B-faktoren, eller load .cif filen i en tekst editor og find søjlen med B-faktorer*

1.2 Beregn RMSD og analysér alignment

Nu skulle AF modellen være klar, download, pak ud og load den første ud af de 5 AF modeller i PyMOL (XXX_model_0.cif).

Align AF modellen og den experimentelle struktur.

Hvad er RMSD? Er der regioner af proteinet der aligner bedre end andre?

1.3 Fortolk pLDDT-værdier

Farv AF-modellen efter pLDDT (som ligger i B-faktor kolonnen i PDB/CIF filen)

```
spectrum b, orange_yellow_green_cyan_blue, model_0, minimum=50, maximum=90
```

udskift `model_0` med jeres objekt navn.

Hvad siger pLDDT om den forudsagte struktur? Hvilket område af strukturen har høj pLDDT og hvilke har lave værdier? Hvad betyder en pLDDT<50?

1.4 Sammenlign B-factor og pLDDT

Sammenlign nu B-factor for krystalstrukturen og pLDDT for AF-modellen. Er der en korrelation mellem B-factor og pLDDT?

1.5 Sammenlign loopkonformationer

Load de andre fire AF-modeller af Parkin (`*_model_1.cif` til `*_model_4.cif`) og align dem alle til `model_0`. Fokuser på det lange fleksible loop (residues 84-143).

Har loopet samme konformation i alle 5 modeller? Hvad fortæller det om AlphaFolds sikkerhed på denne regions struktur? Passer det med pLDDT værdierne I kiggede på i spørgsmål 3?

1.6 Forklar fjernelse af flexibelt loop

Bemærk at det flexible loop er fjernet i konstruktet brugt til krystalstrukturen (hvor residue 83 og 144 er forbundet). Hvorfor mon krystallograferne har gjort det?

2 Opgave 2. Hvad fortæller PAE om domæne-orientering

I denne opgave skal vi arbejde med Predicted Aligned Error (PAE).

2.1 Fortolk PAE-plottet

Fortolk PAE-plottet for jeres Parkin model:

Se på PAE-plottet i AlphaFold browser vinduet (findes også i jeres downloaded mappe som JSON-fil til data-behandling).

- Hvad betyder en mørk grøn firkant? (lav PAE = $\sim 0-5$ Å)
- Hvad betyder en lys/hvid firkant? (høj PAE = $\sim 20-30$ Å)
- Hvorfor er hele diagonalen mørk grøn?

2.2 Identificér domæner i PAE

Identificér domæne-grænser i PAE-plottet. Find de 5 domæner i plottet (se domæne-residue numre fra Opgave 1). Hvilke domæner har lav PAE (grøn) mellem hinanden, og hvilke har høj PAE (lys/hvid)?

Hint

Find Ubl-domænet (residues 1-76) i PAE-plottet. Kig på firkanten fra (1,1) til (76,76) - altså Ubl mod sig selv. Hvilken farve har denne firkant?

2.3 Analysér Parkins domænestructur

Hvad fortæller PAE om Parkins struktur? Baseret på PAE-plottet - hvilke dele af Parkin har en veldefineret relativ orientering, og hvilke dele er fleksibelt forbundet? Relaterer dette til pLDDT-værdierne I så i Opgave 1?

2.4 Identificér det fleksible loop

Identificér det lange fleksible loop i Parkin strukturen. Er positionen af dette loop veldefineret i forhold til resten af Parkin strukturen ifølge PAE-plottet?

3 Opgave 3. Modelling af et protein-kompleks med ligand og ioner

I denne opgave skal I bruge AlphaFold-serveren til at forudsige strukturen af en hetero-dimer mellem HsIFT25 (UniProt ID: Q9Y547) og HsIFT27 (UniProt ID: Q9BW83-1). IFT25 er et calcium-bindende protein, og IFT27 er en lille GTPase.

Tilføj begge proteiner samt relevante ligander/ioner til en strukturforudsigelse på AlphaFold-serveren. Download den resulterende model.

3.1 Fortolk pTM- og iPTM-værdier

Hvad fortæller henholdsvis pTM- og iPTM-værdierne om kvaliteten af den forudsagte model?

3.2 Analysér pLDDT og PAE i kompleks

Analyser pLDDT- og PAE-værdierne for jeres model. Baseret på disse metrics, kan man have tillid til kvaliteten af den forudsagte struktur af protein-komplekset? Er orienteringen af IFT25 og IFT27 veldefineret? Begrund jeres svar.

3.3 Find og sammenlign nukleotidbinding

Hvor er nukleotidet placeret i strukturen? Find en GTPase struktur i PDB med nukleotid bundet, og align denne med IFT27. Binder nukleotidet på samme måde i begge strukturer?

3.4 Analysér calciumions koordination

Hvor binder calcium-ionen, og hvordan er den koordineret? (Beskriv hvilke aminosyrerester der deltager i koordinationen)

4 Opgave 4. Protein-interaktion med Alphafold

Et affinity pull-down MS eksperiment med tagged UniProt: Q3Y8L7 (ODA16, bait protein) gav følgende to hits: UniProt: A2T2X4 (IFT46) og UniProt: A8JA42 (IFT56).

4.1 Kortlæg interaktioner med AlphaFold

Brug AlphaFold til at dissekere hvordan de 3 proteiner interagerer. Binder begge "hits" direkte til ODA16?

4.2 Beskriv kompleksets arkitektur

Hvad er arkitekturen af komplekset og hvilke residues af IFT46 interagerer med ODA16 og IFT56? Hvilke dele er velforudsagte? Kan I stole på de forudsagte strukturelle modeller for interaktionerne mellem proteinerne?

4.3 Fortolk pTM og iPTM-scorer

Hvad er pTM og iPTM værdierne? Hvorfor er de mon lavere end for IFT25/27 i opgave 3?

4.4 Design eksperiment med trunkeerede konstrukter

Baseret på pLDDT-værdierne for IFT46, design et eksperiment for at teste om de lave pTM/iPTM værdier skyldes IFT46's disordered regioner eller om de lokale interfaces faktisk er dårligt forudsagte. Beskriv hvilke AlphaFold kørsler I ville lave og hvad I ville forvente at se hvis problemet primært er disordered regioner. Hvis der er tid så kør gerne AlphaFold forudsigelserne.