West Coast Structural Biology Workshop, March 2023

Using predicted models in Phenix

Dorothee Liebschner

Lawrence Berkeley Laboratory

The Phenix Project

Lawrence Berkeley Laboratory

Paul Adams, Pavel Afonine, Dorothee Liebschner, Nigel Moriarty, Billy Poon, Christopher Schlicksup, Oleg Sobolev

BERKELEY LAB

University of Cambridge

Randy Read, Airlie McCoy, Rob Oeffner

UNIVERSITY OF CAMBRIDGE

Los Alamos National Laboratory
New Mexico Consortium

Tom Terwilliger, Li-Wei Hung

New Mexico CONSORTIUM

Los Alamos NATIONAL LABORATORY

UTHealth

Matt Baker

UTHealth
The University of Texas
Health Science Center at Houston

Duke University

Jane & David Richardson, Christopher Williams, Vincent Chen
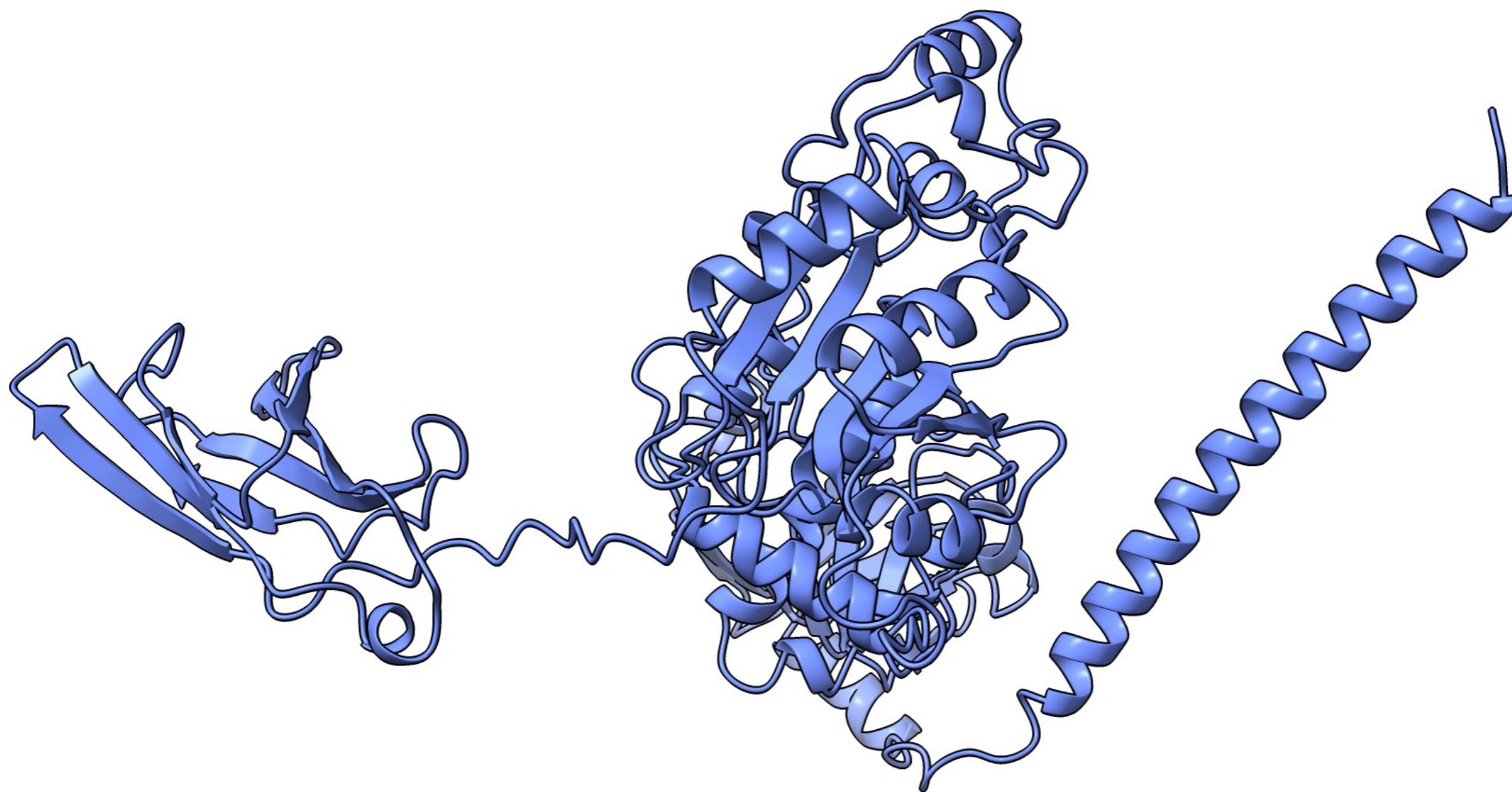
An NIH/NIGMS funded Program Project

Example for a prediction (POMGNT2)

https://alphafold.ebi.ac.uk/entry/Q8NAT1

(Example from Phenix AlphaFold MR tutorial)

Predictions come with confidence measures
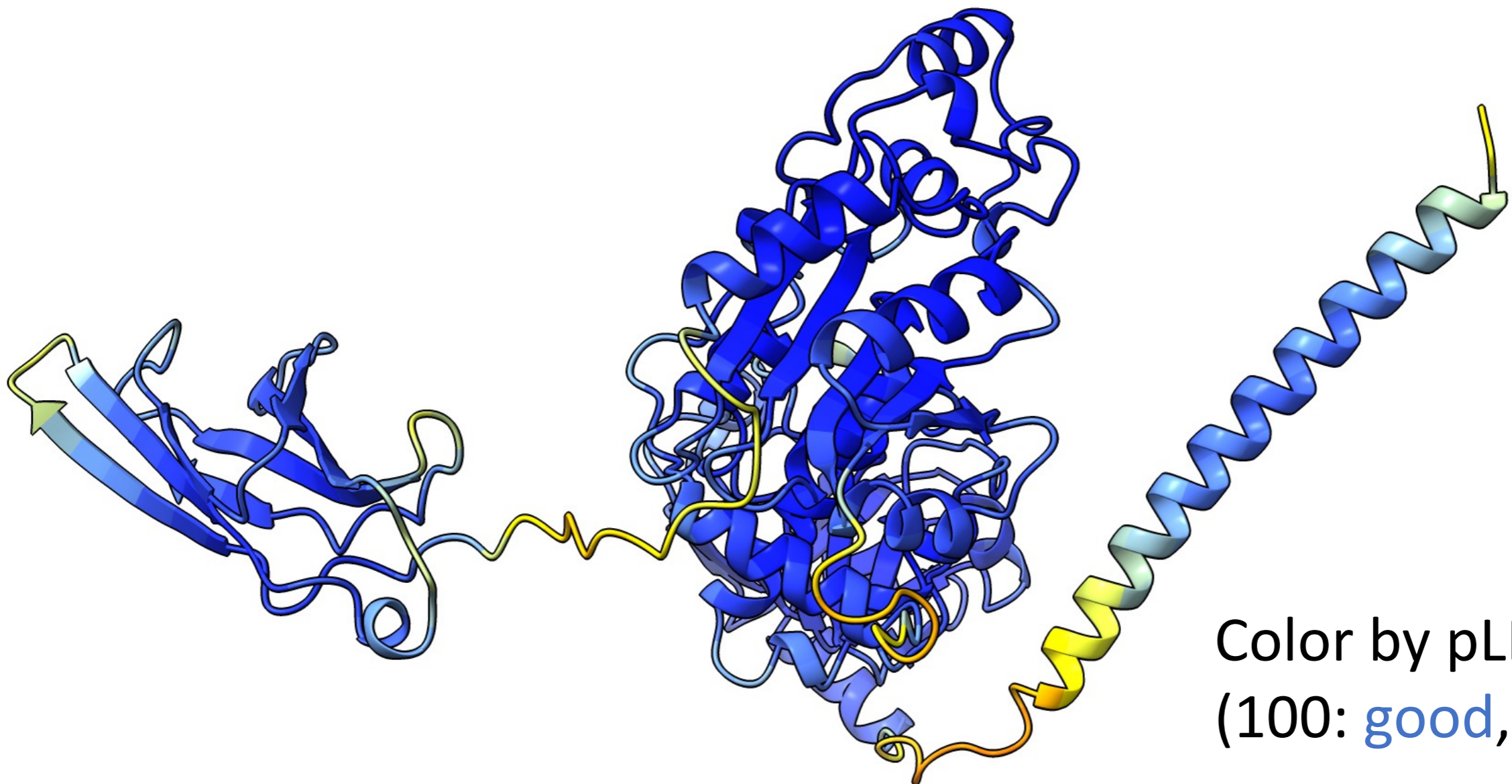
1. pLDDt (predicted Local Distance Difference Test)
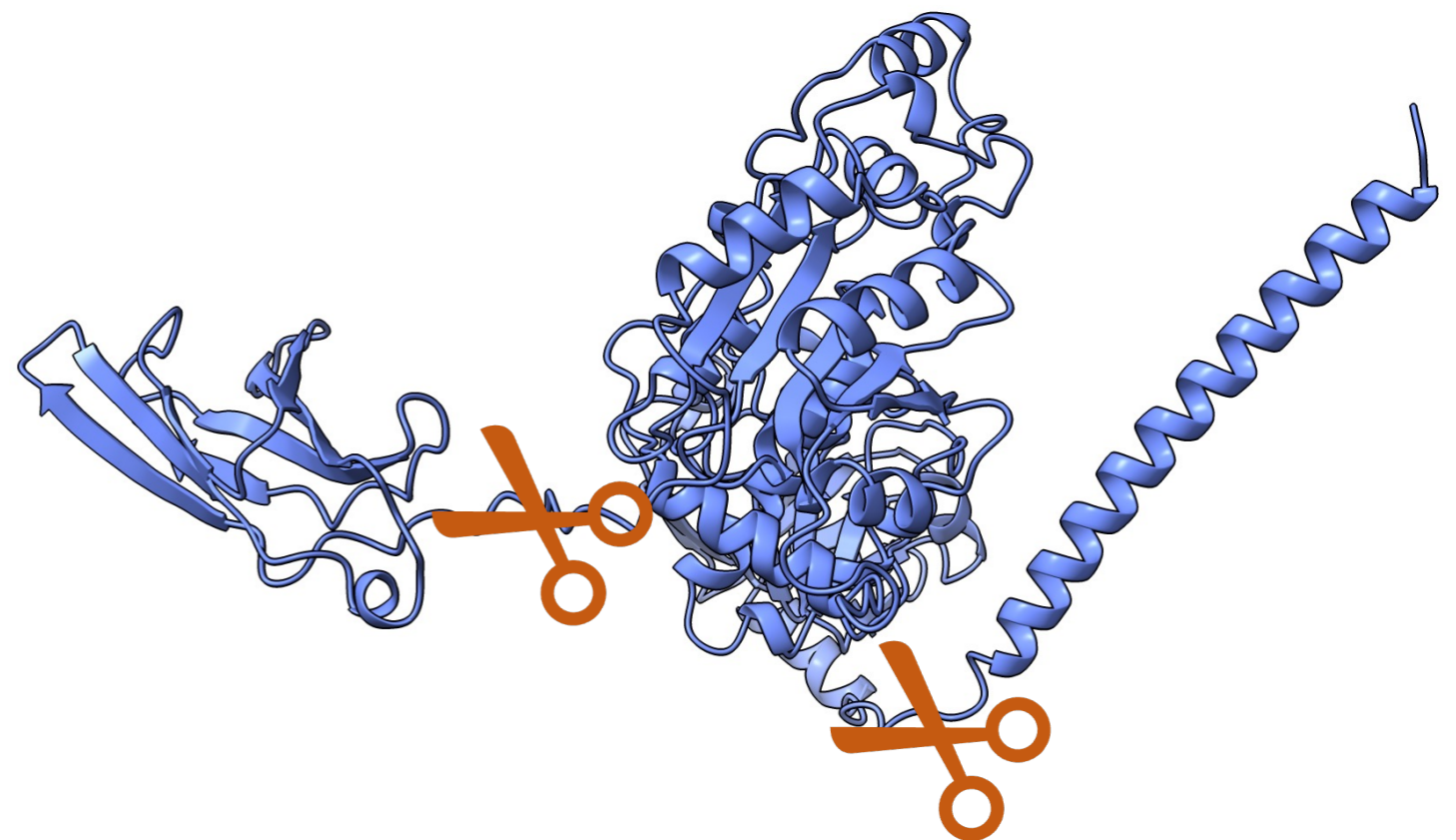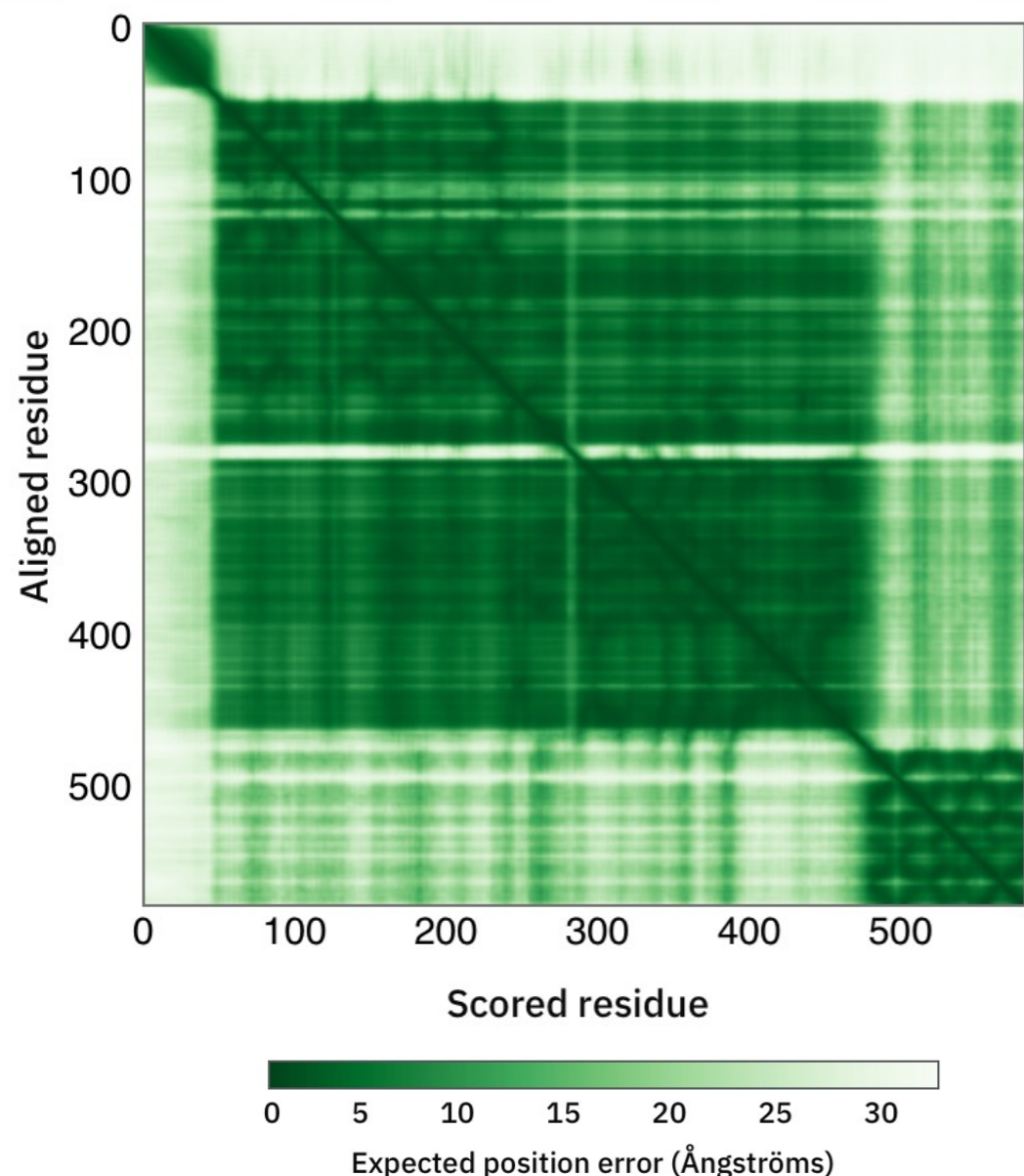
- per-residue confidence measure
- scales from 0 – 100
- pLDDt > 90: predicted with high accuracy

Color by pLDDt
(100: good, 0: bad)
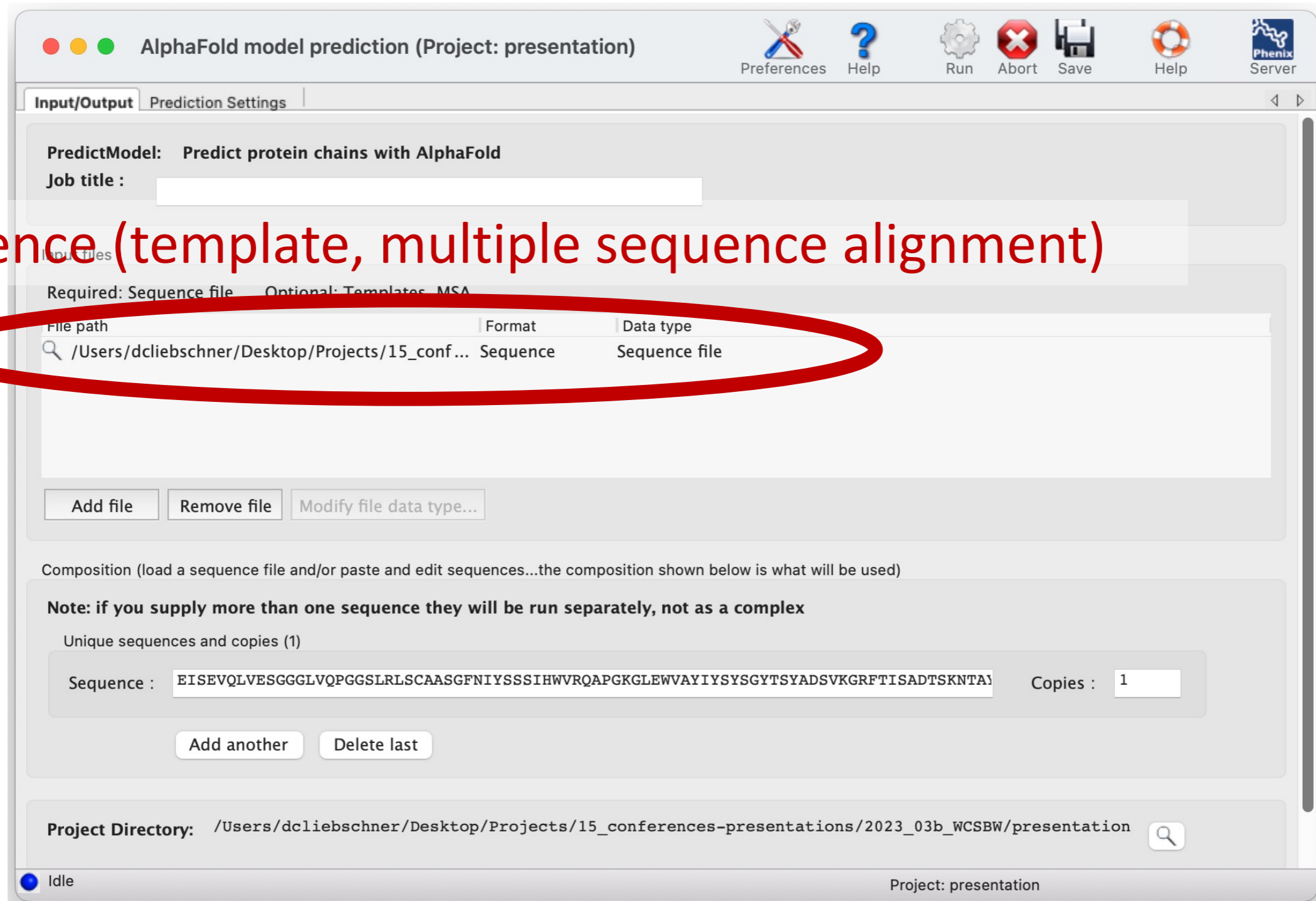
Predictions come with confidence measures

2. Predicted aligned error (PAE)

- Relative domain positions → Large-scale topology
- Dark green: the predicted relationship between this pair of residues is likely to be accurate

- Suggests 3 domains
- Mutual configuration is not clear

Predicting a model in Phenix

Sequence (template, multiple sequence alignment)

Do I need to install AlphaFold?

Phenix Server or Colab for running AlphaFold
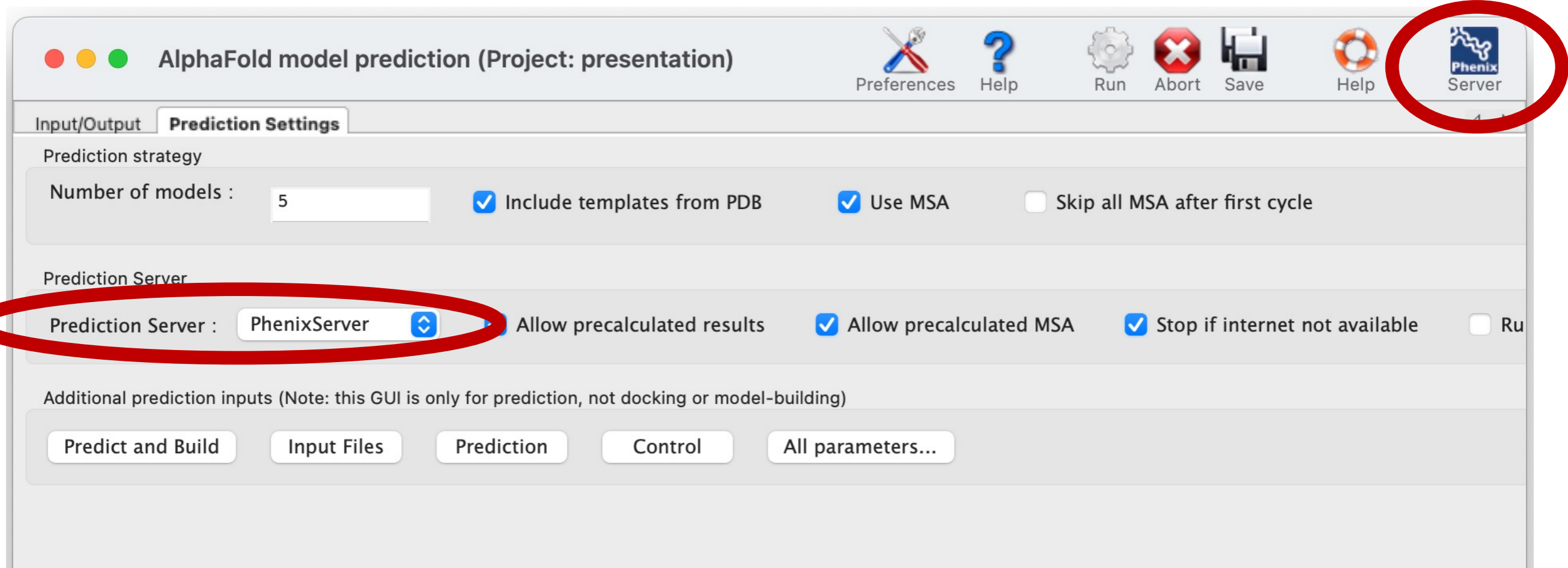
Run AlphaFold on
- Phenix Server
- Colab
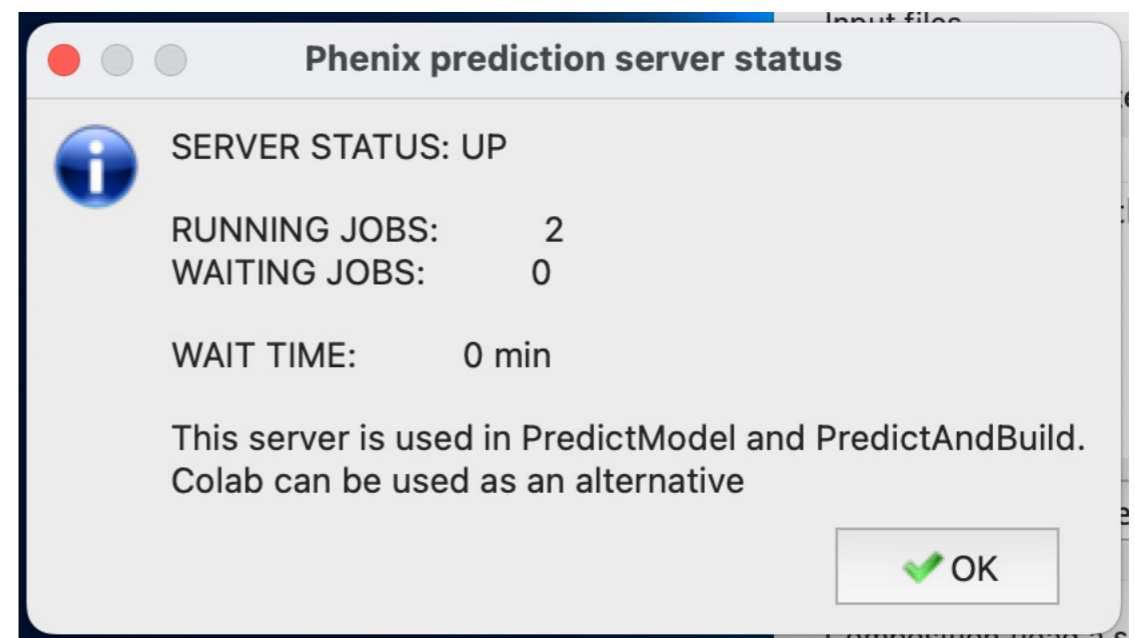
7s5L, 1.58 Å, X-ray diffraction
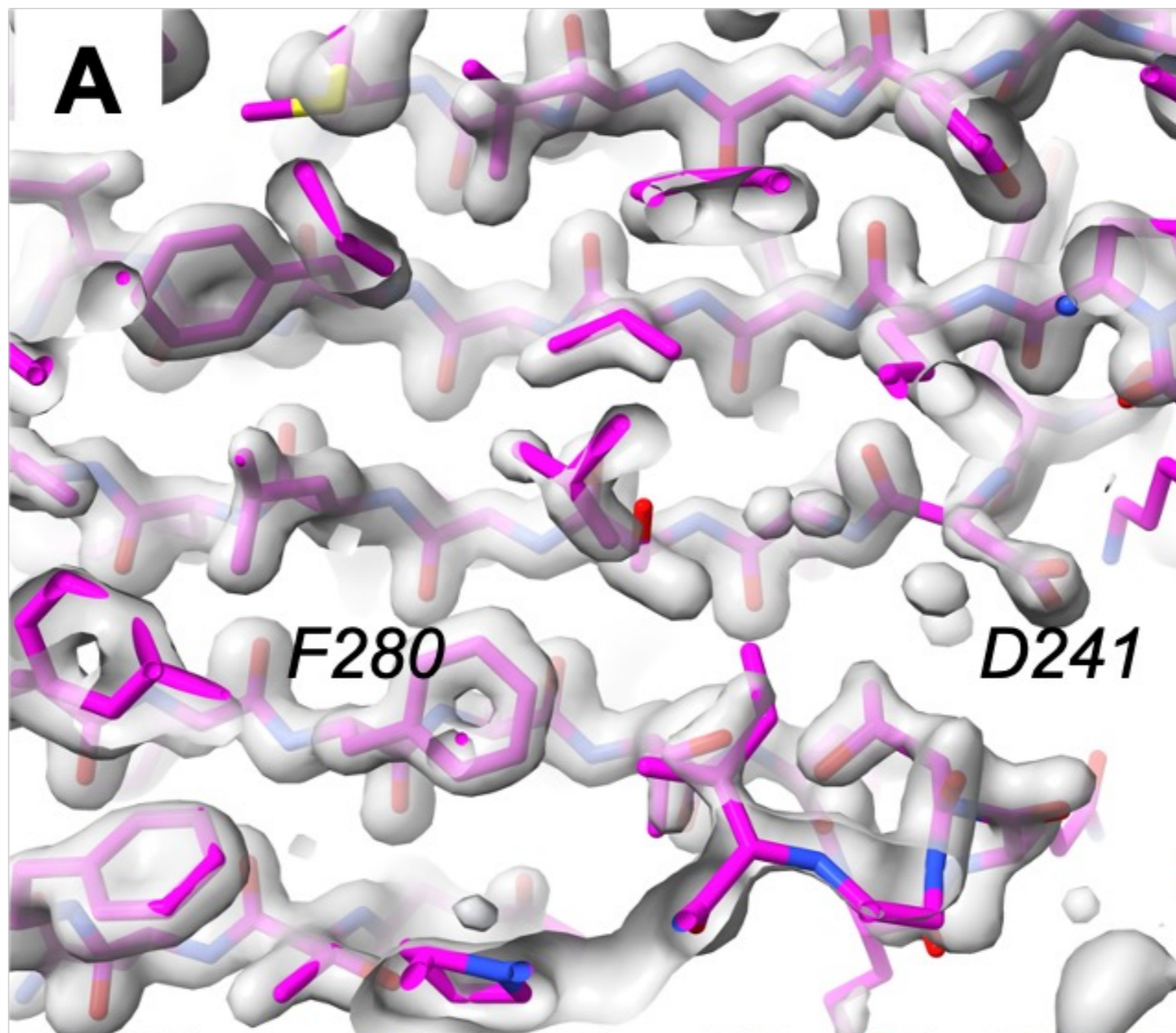
AlphaFold model

Predictions can be very accurate.

...but (still) no match for experiment

7s5L, 1.58 Å, X-ray diffraction
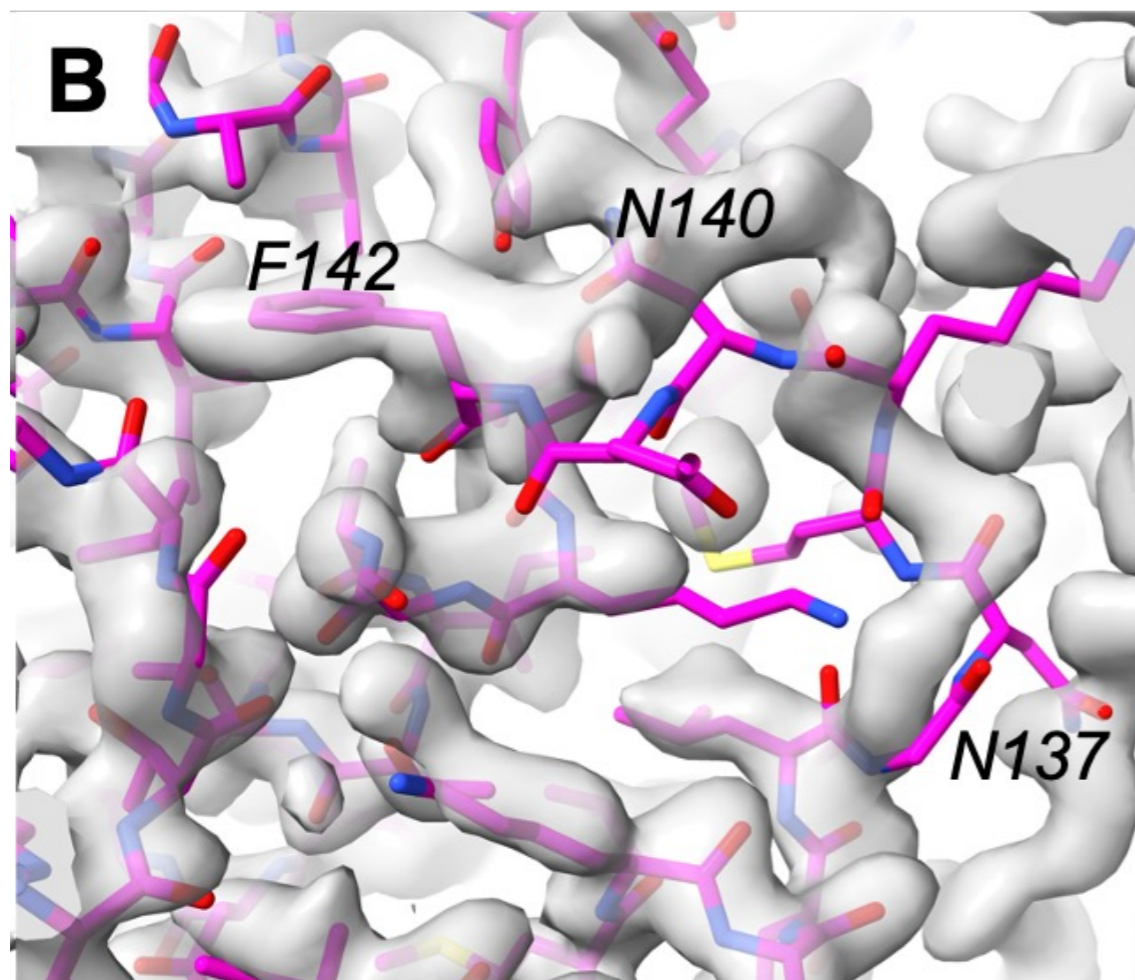
AlphaFold model

Some parts may be inaccurate
(even when predicted with high confidence)
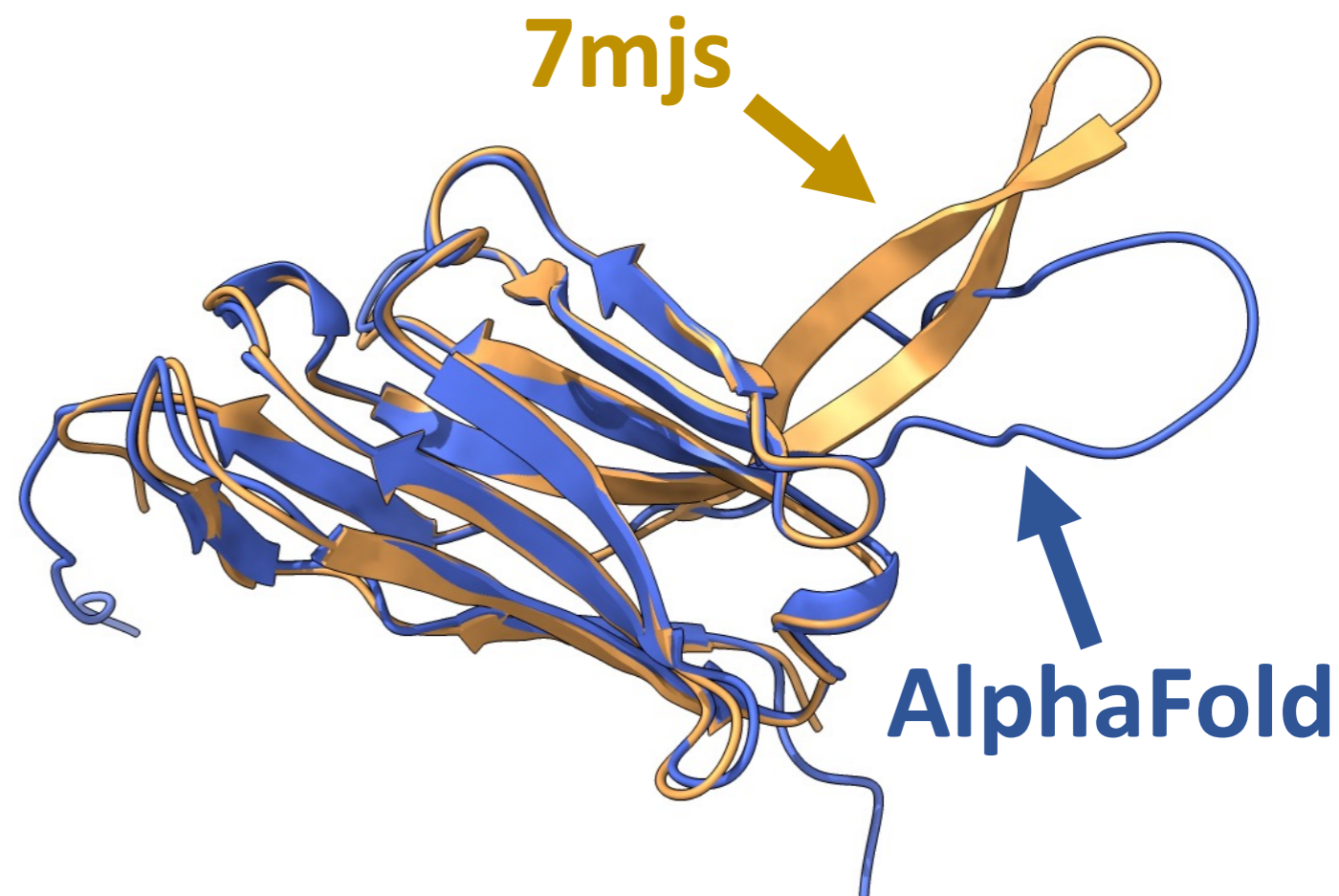
Parts may be distorted

Predicted models can disagree globally and locally with experimental structures.

Only ~40% of residues in the human proteome are modeled with high confidence (pLDDt).

How to use predictions?

Incorporate predictions into the "conventional" structure determination workflow.

Millán C. et al.. Assessing the utility of CASP14 models for molecular replacement. Proteins. 2021 Dec;89(12):1752–69.
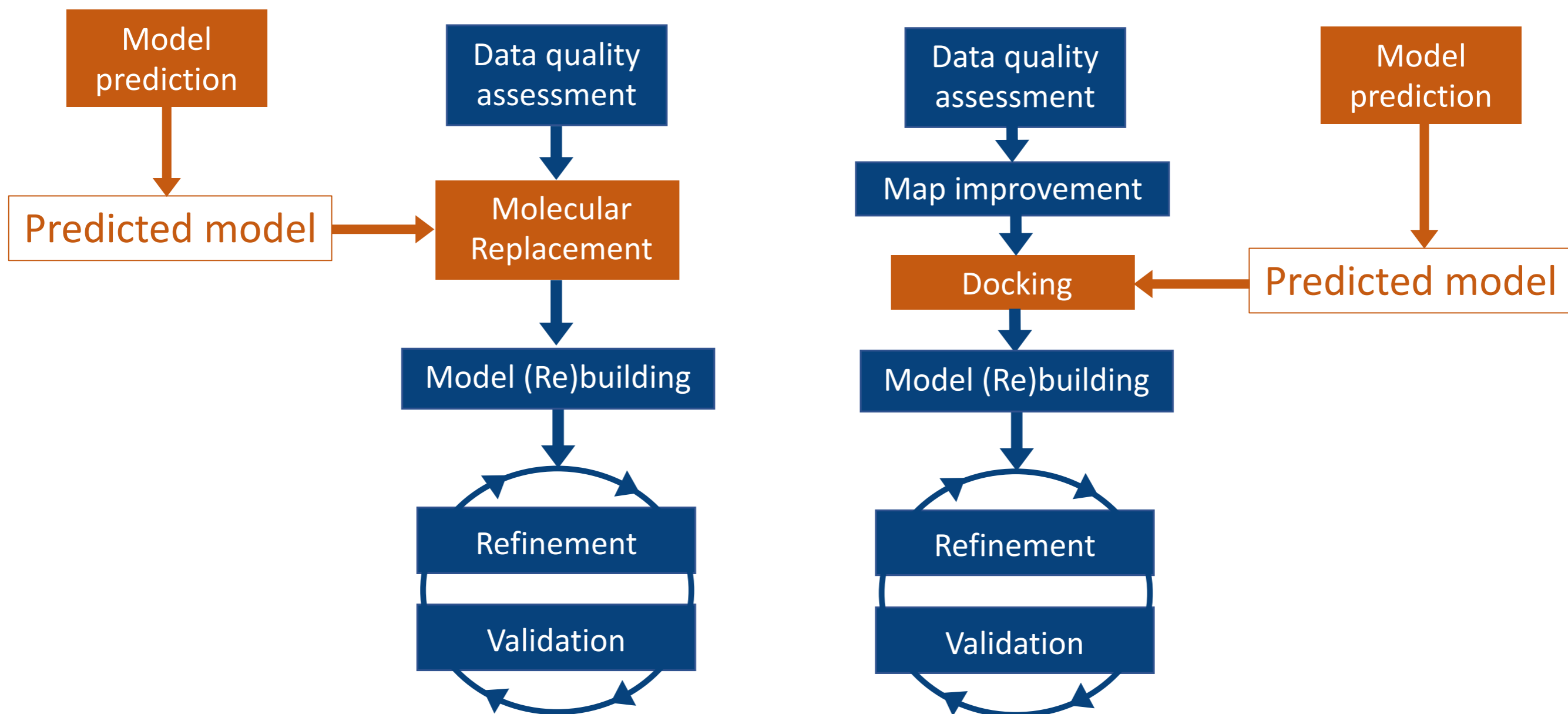
Incorporate predictions into the "conventional" structure determination workflow.

Example:
Cryo-EM map (30160 – 7brm)


```
>chain ' A'
XXXXXXXXXXXXXXXXCLTAPPKEAARPTLMPRAQSYKDLTHLPAP
TGKIFVSVYNIQDETGQFKPYPASNFSTAVPQSATAMLVTALKDS
RWFIPLERQGLQNLLNERKIIRAAQENGTVAINNRIPLQSLTAAN
IMVEGSIIGYESNVKSGGVGARYFGIGADTQYQLDQIAVNLRVVN
VSTGEILSSVNTSKTILSYEVQAGVFRFIDYQRLLEGEVGYTSNE
PVMLCLMSAIETGVIFLINDGIDRGLWDLQNKAERQNDILVKYRH
MS
```

sequence
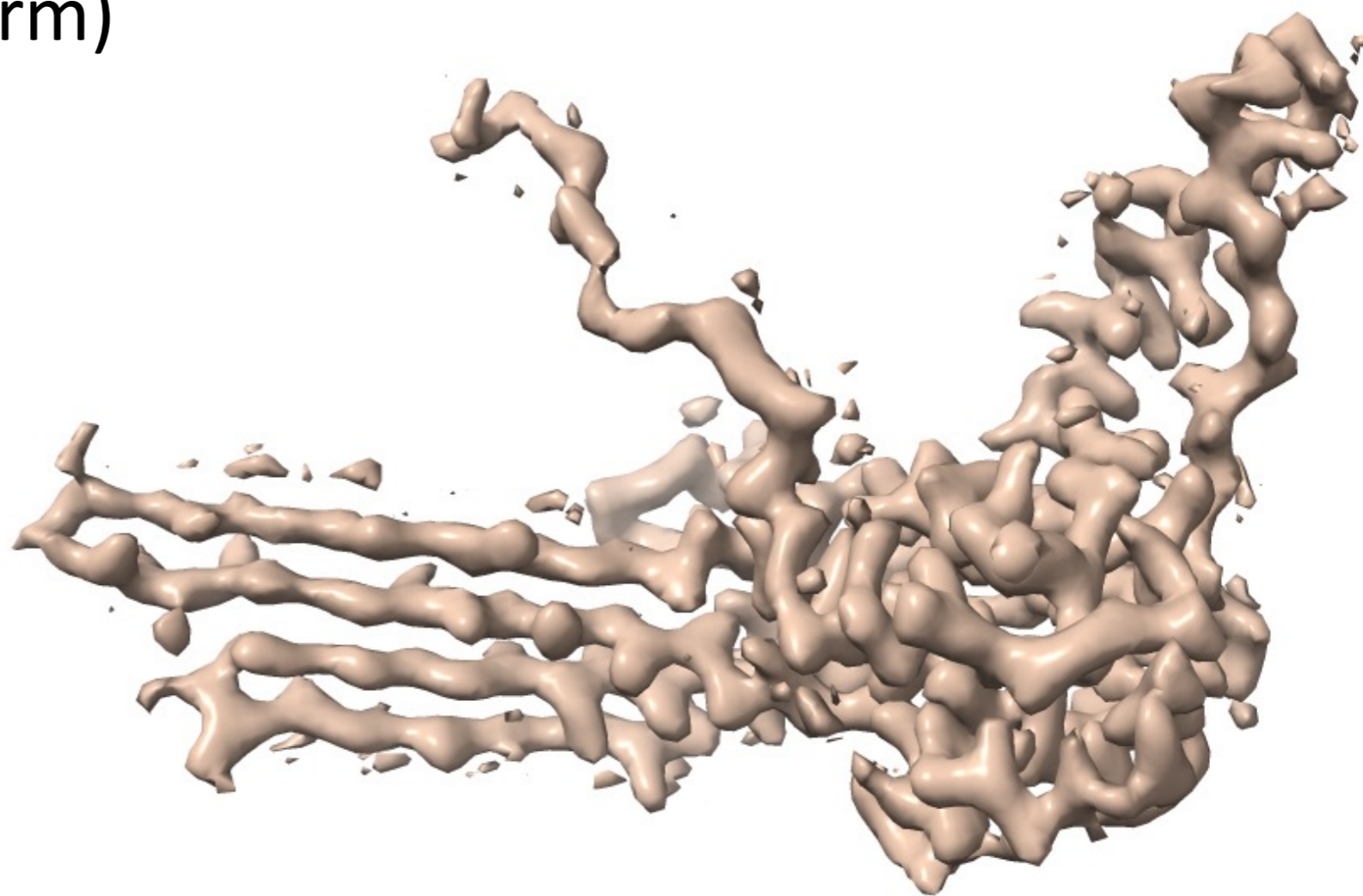
cryo-EM map (3.6 Å)

Get a prediction

sequence

```
>chain ' A'
XXXXXXXXXXXXXXXCLTAPPKEAARPTLMPRAQSYKDLTHLPAP
TGKIFVSVYNIQDETGQFKPYPASNFSTAVPQSATAMLVTALKDS
RWFIPLERQGLQNLLNERKIIRAAQENGTVAINNRIPLQSLTAAN
IMVEGSIIGYESNVKSGGVGARYFGIGADTQYQLDQIAVNLRVVN
VSTGEILSSVNTSKTILSYEVQAGVFRFIDYQRLLEGEVGYTSNE
PVMLCLMSAIETGVIFLINDGIDRGLWDLQNKAERQNDILVKYRH
MS
```
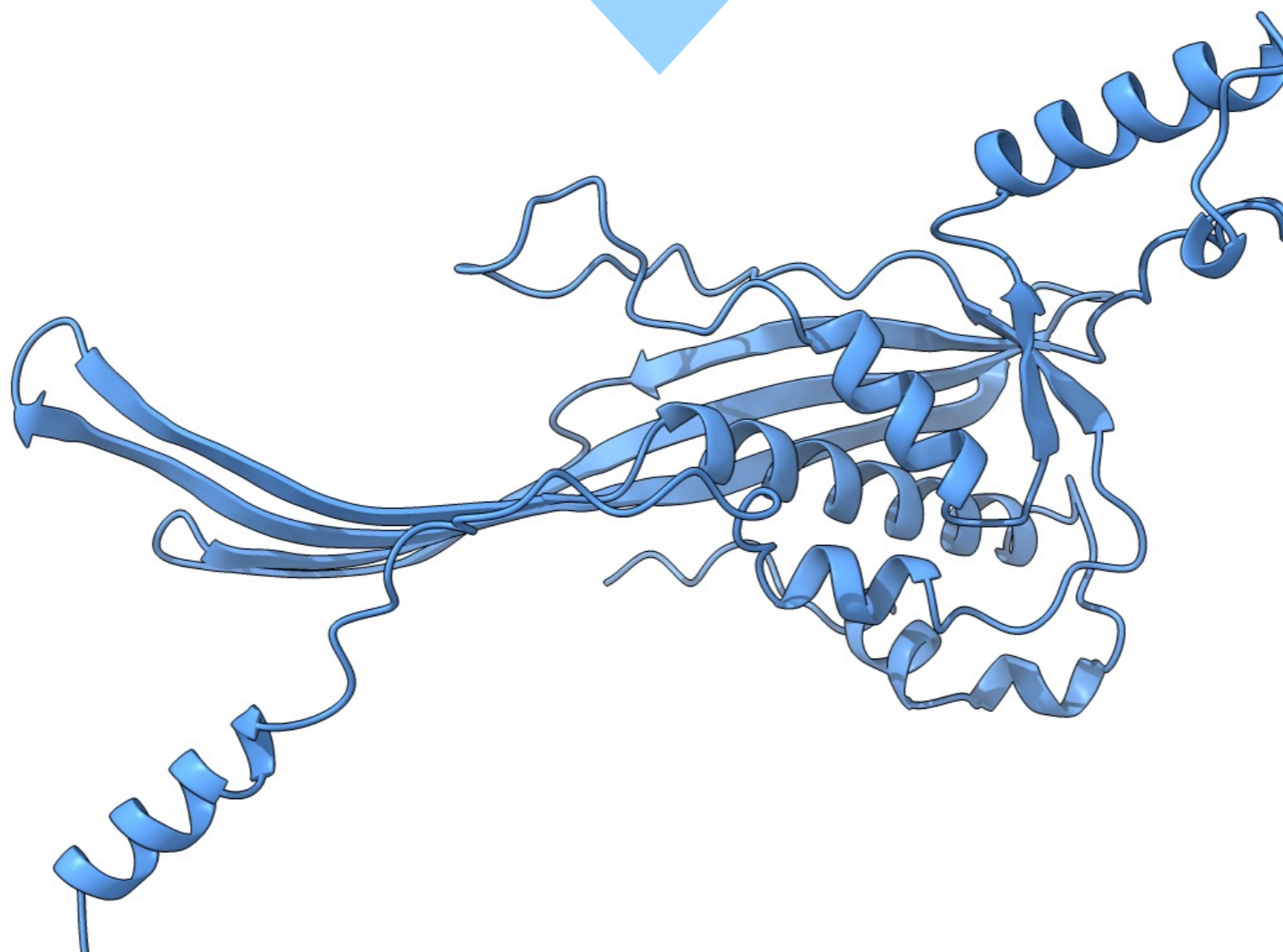
AlphaFold
model

Process prediction

sequence

```
>chain ' A'
XXXXXXXXXXXXXXXXCLTAPPKEAARPTLMPRAQSYKDLTHLPAP
TGKIFVSVYNIQDETGQFKPYPASNFSTAVPQSATAMLVTALKDS
RWFIPLERQGLQNLLNERKIIRAAQENGTVAINNRIPLQSLTAAN
IMVEGSIIGYESNVKSGGVGARYFGIGADTQYQLDQIAVNLRVVN
VSTGEILSSVNTSKTILSYEVQAGVFRFIDYQRLLEGEVGYTSNE
PVMLCLMSAIETGVIFLINDGIDRGLWDLQNKAERQNDILVKYRH
MS
```
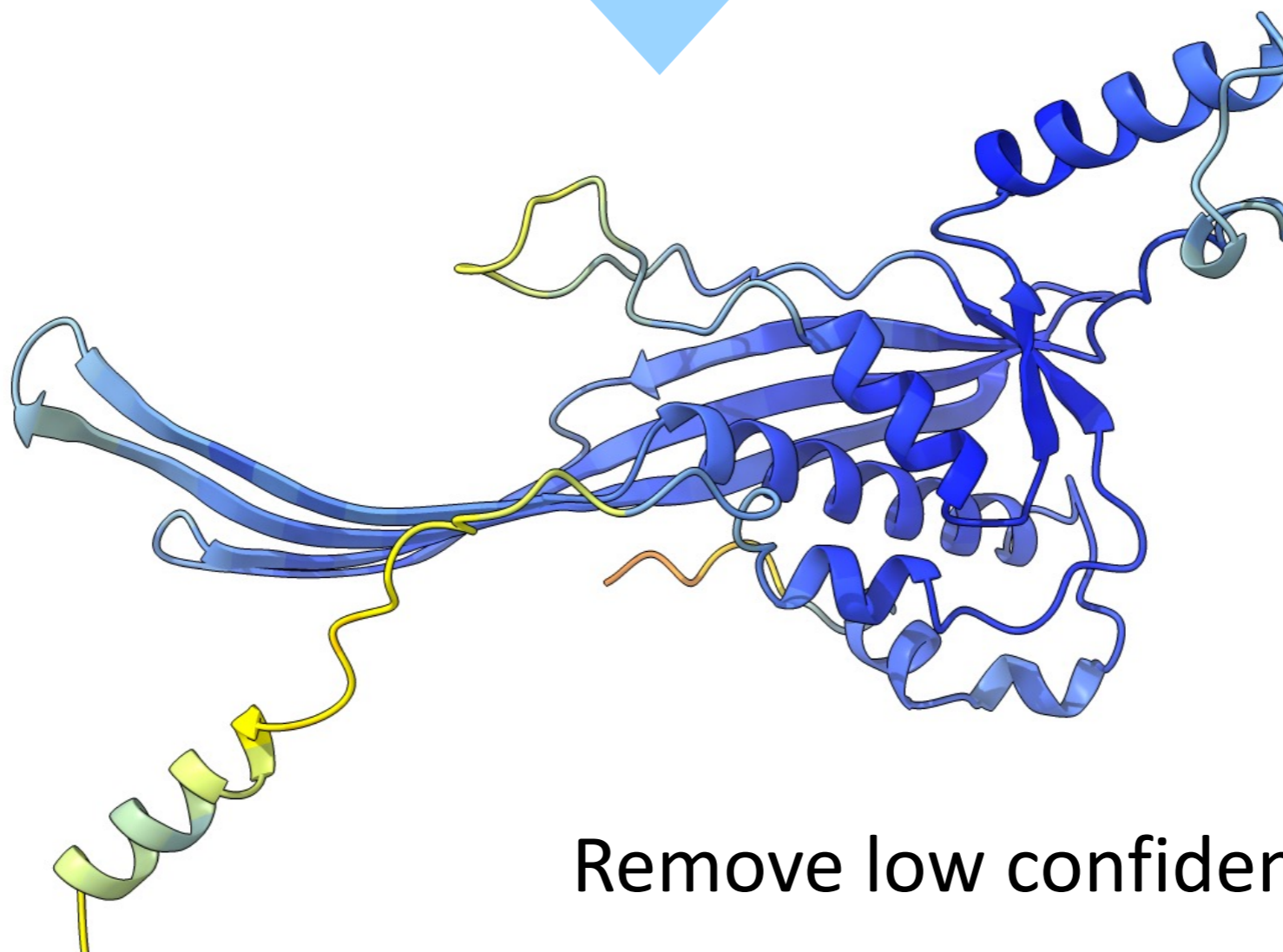
AlphaFold model

Remove low confidence parts

Processed
AlphaFold
model

Dock processed model

cryo-EM map

Processed
AlphaFold
model
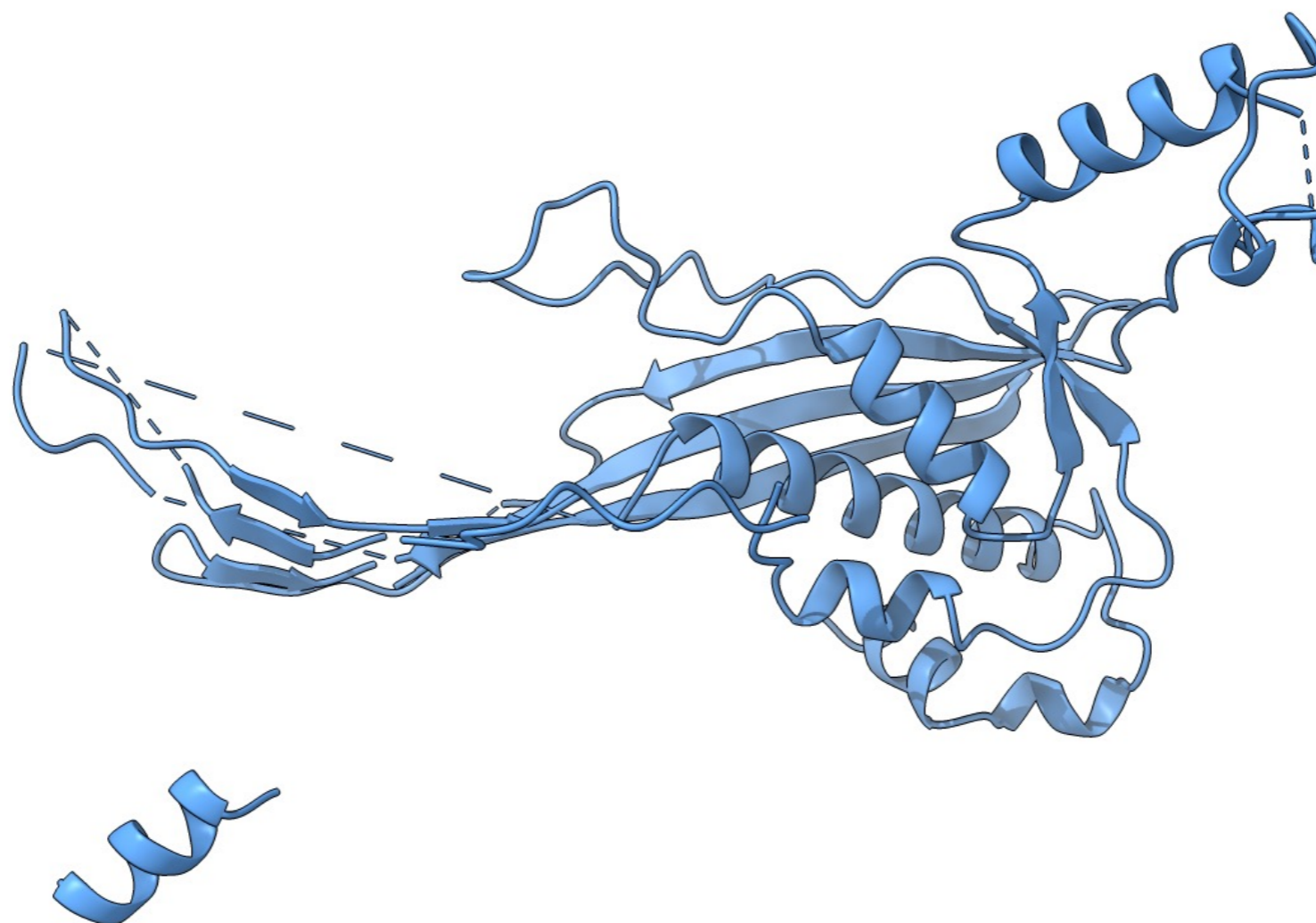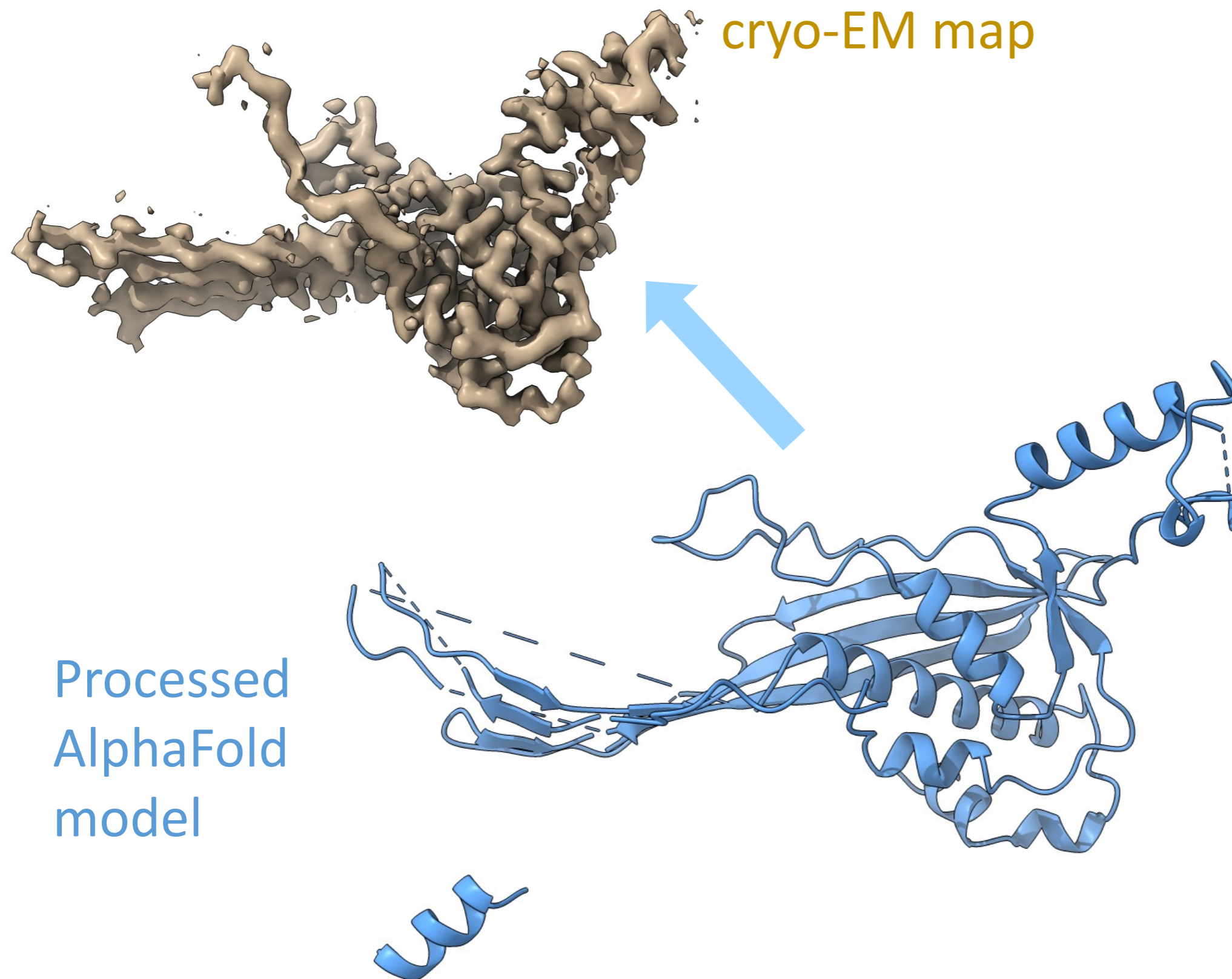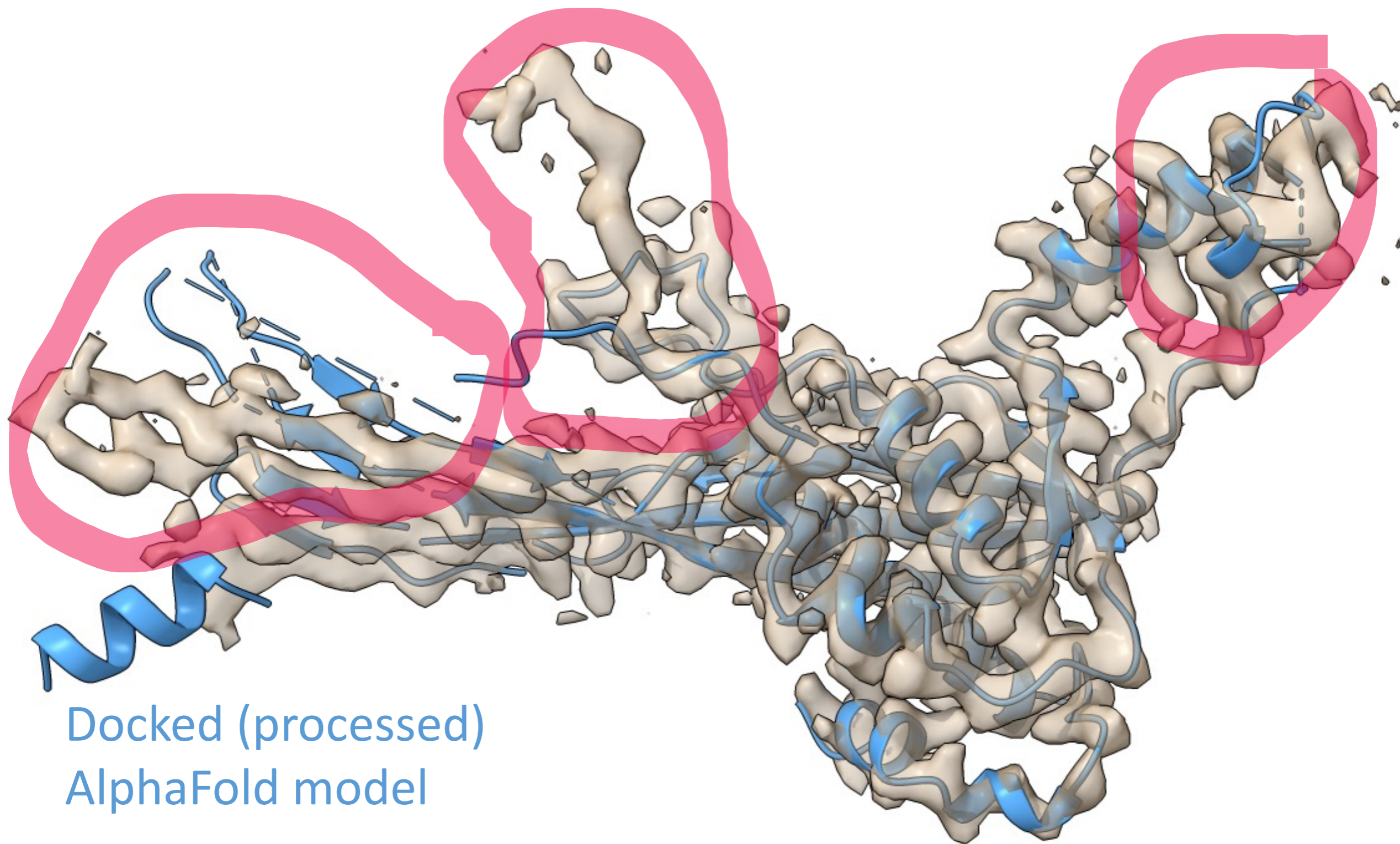
Dock processed model

Docked (processed)
AlphaFold model

Some parts don't fit into the map
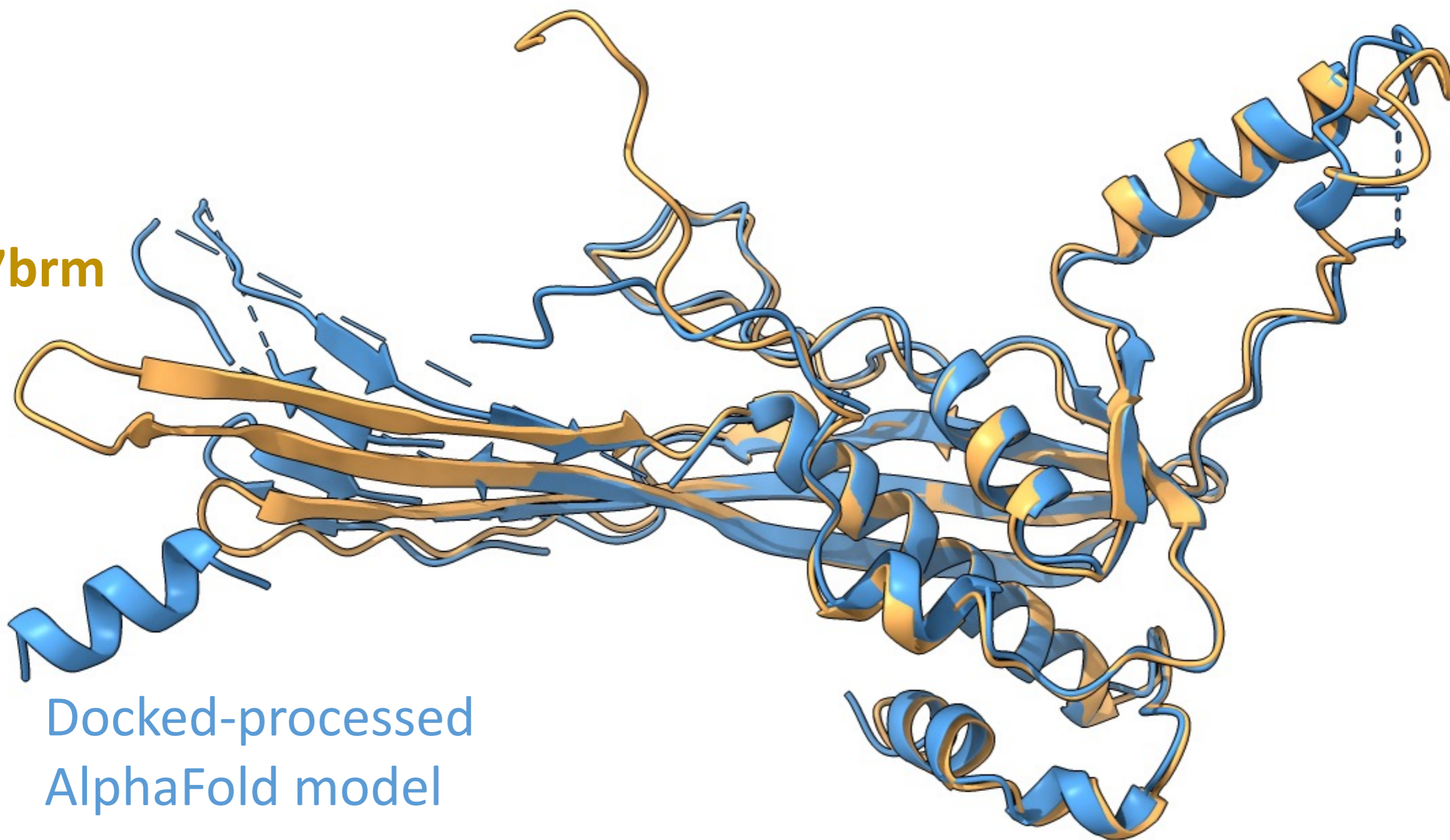
7brm

Docked-processed
AlphaFold model

Some parts don't fit into the map → fit loops and rebuild
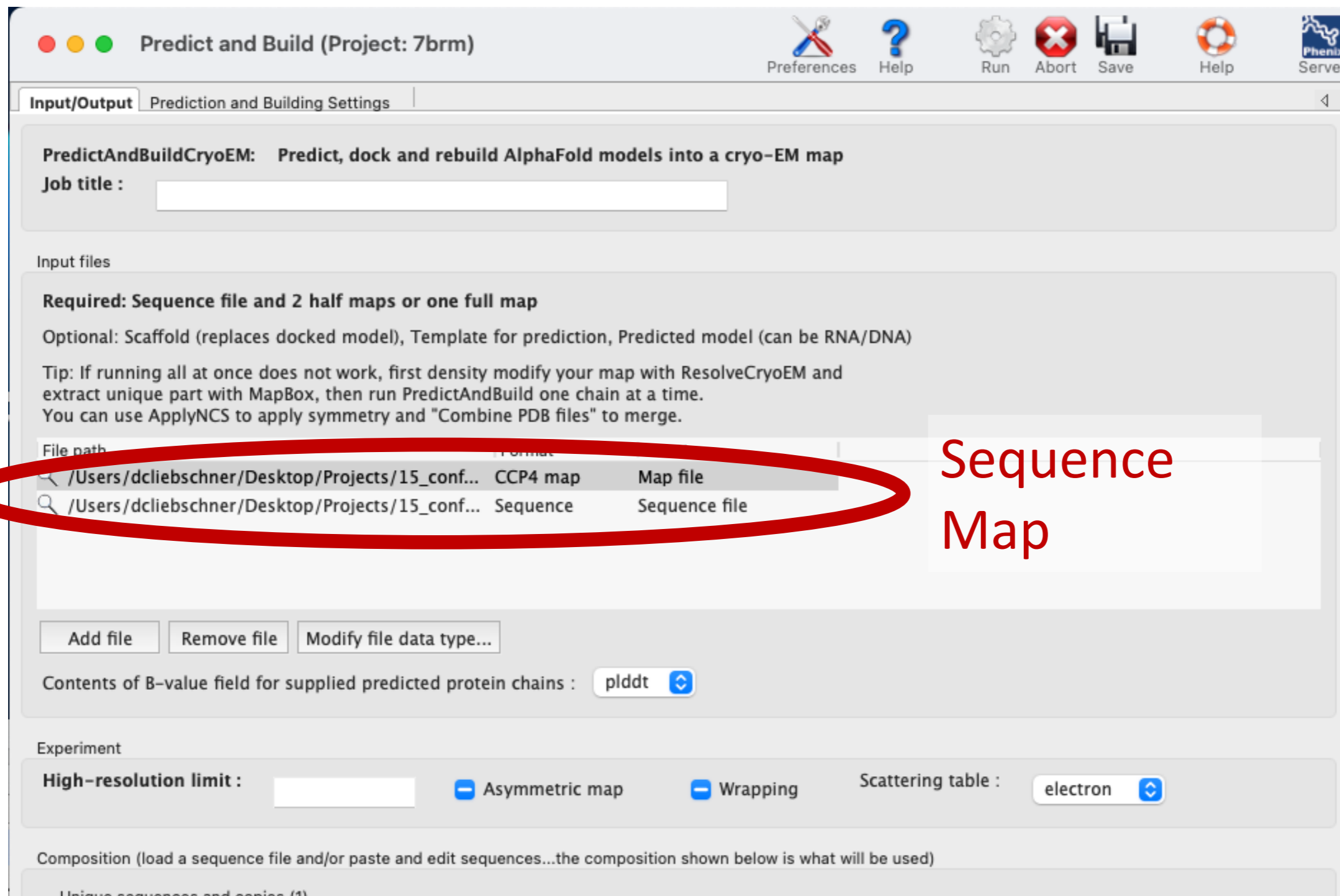
7brm
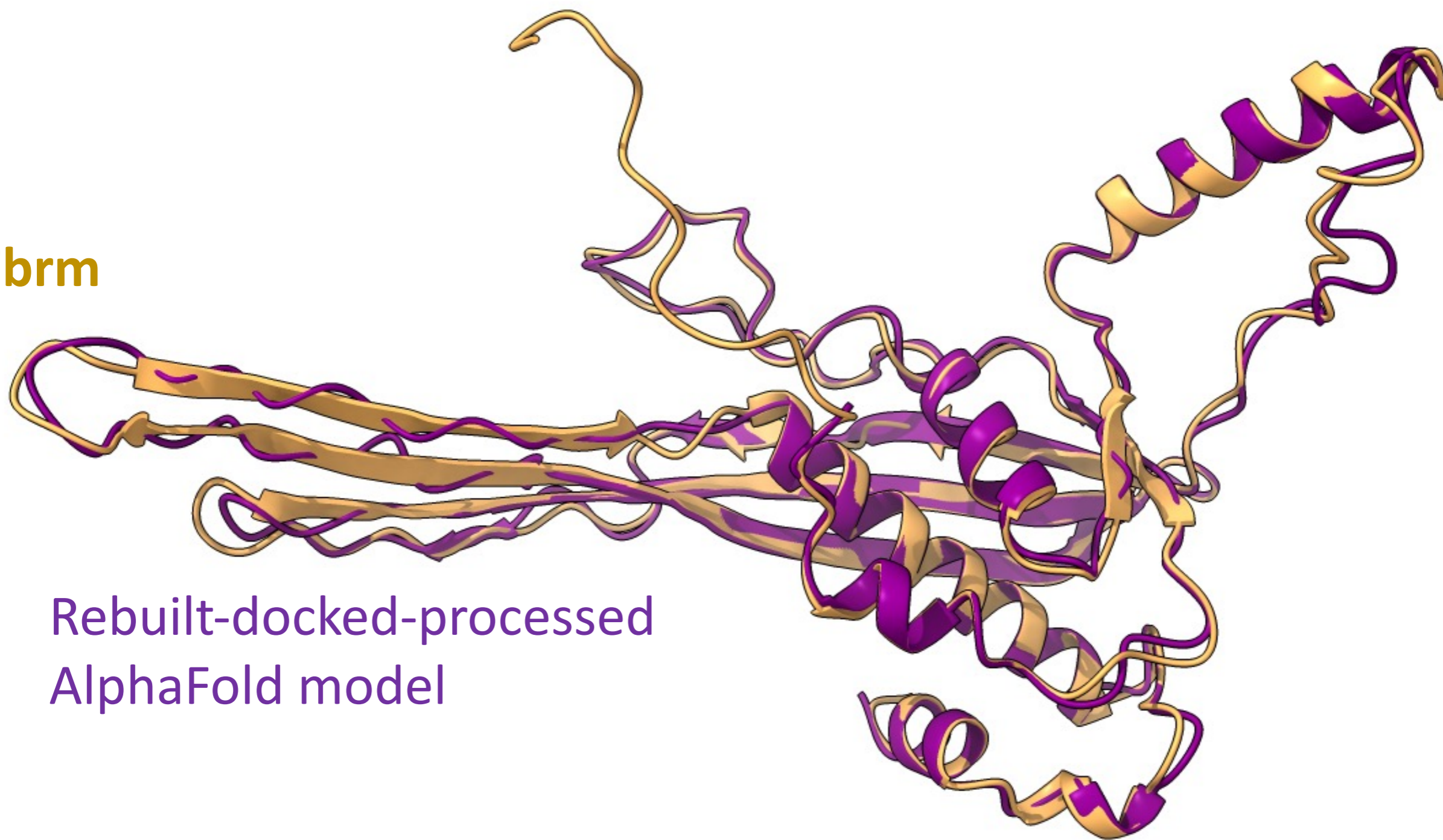
Rebuilt-docked-processed
AlphaFold model

This model is close to the deposited model.

Fully automatic: predict and build

Will perform all the steps automatically.

Using predicted models

New approach: Iterate prediction and model building
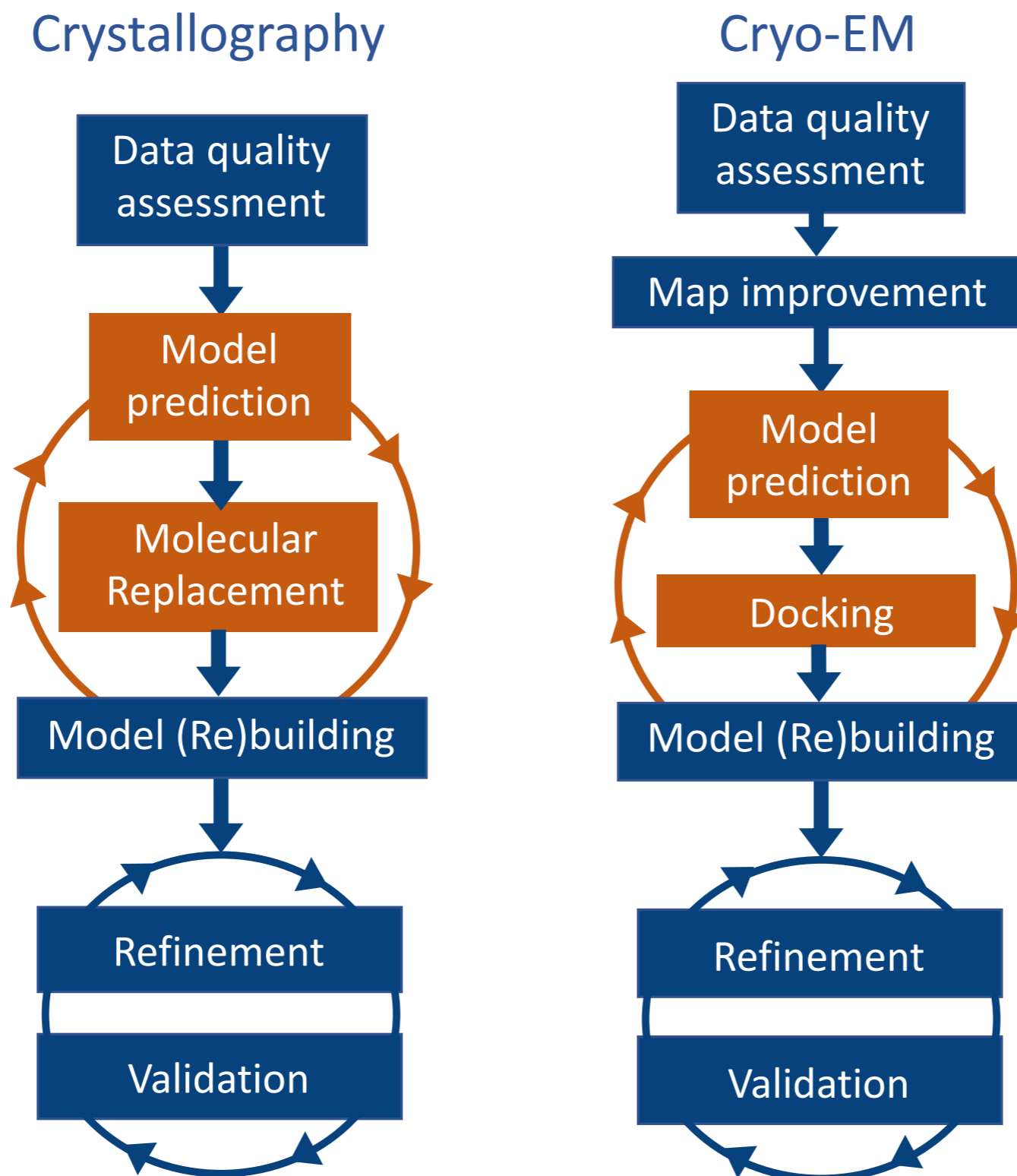
Improving prediction

Improving prediction

Rebuilt AlphaFold

7brm
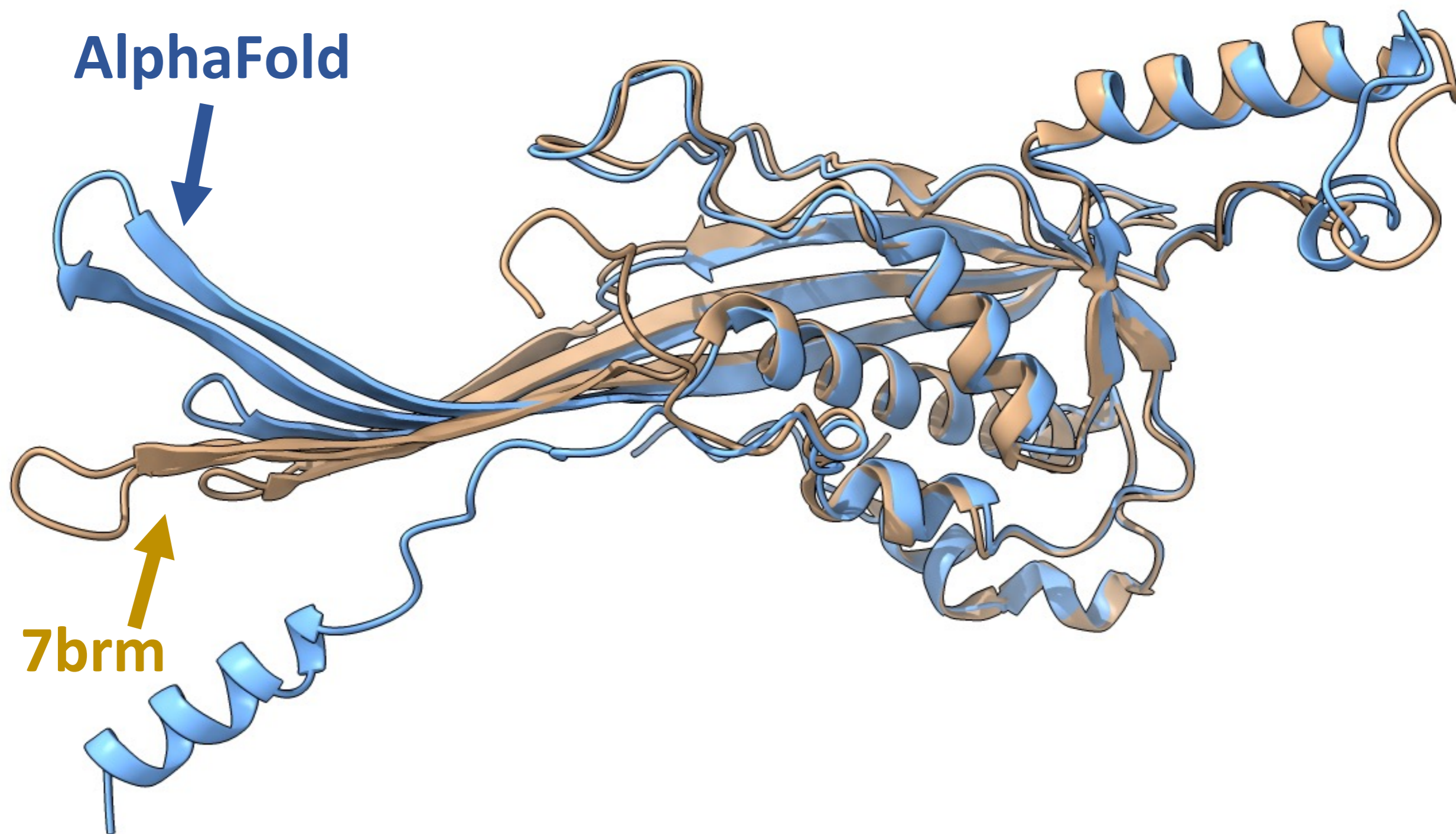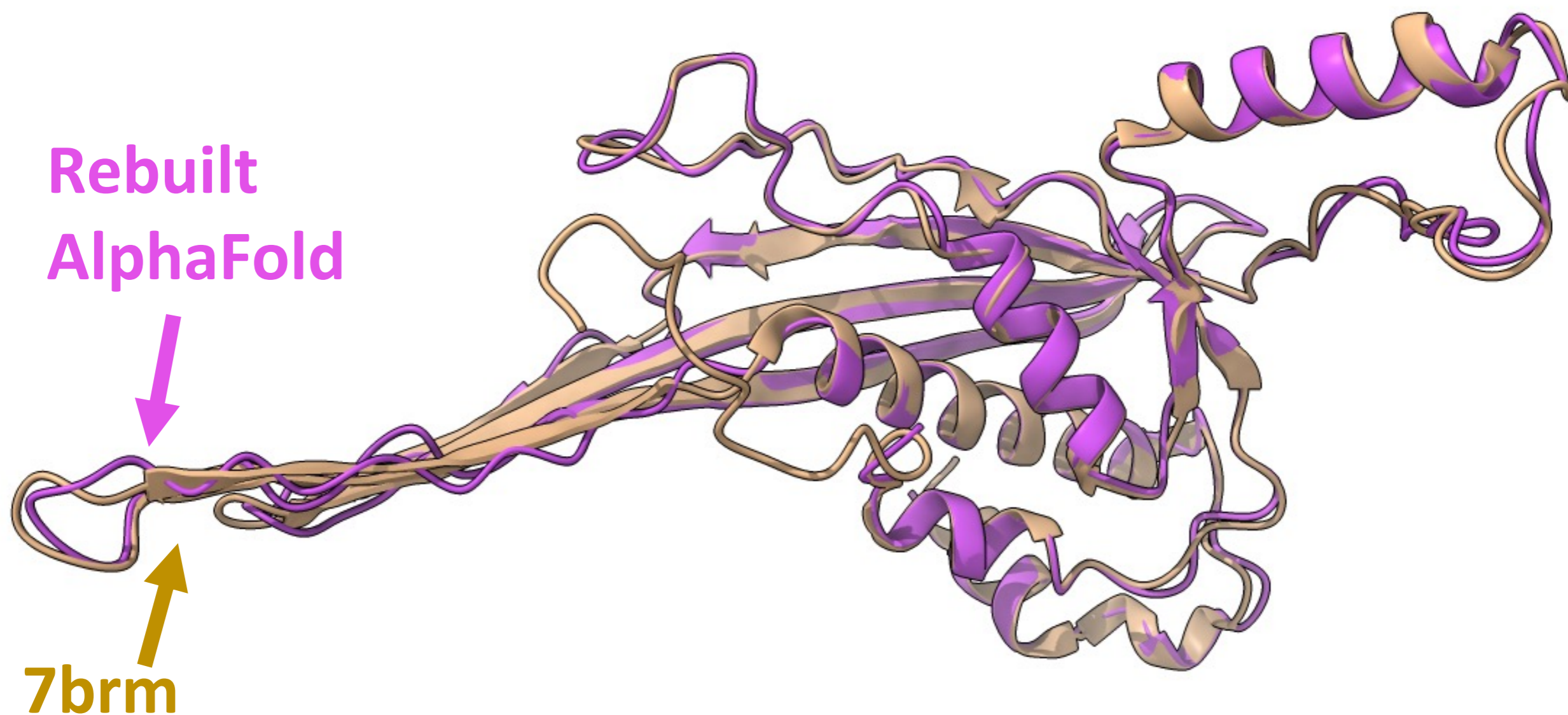
What happens if we use the rebuilt model as template for AlphaFold?

Improving prediction

- The template improves prediction
- New prediction can be even better than the template

Terwilliger TC et al. Improved AlphaFold modeling with implicit experimental information. Nat Methods. 2022 Nov;19(11):1376–82.

Iterate with Predict and Build

Further reading/material

Documentation:
https://www.phenix-online.org/documentation/

Working with AlphaFold2, RoseTTAFold and other predicted models

You can use the predicted models from AlphaFold and other prediction software in Phenix. Using these models can be very helpful in structure determination because the models can be very accurate over much of their length and the models come with accuracy estimates that allow removal of poorly-predicted regions.

How to use predicted models in Phenix

Use Predict_and_build to incorporate predicted models in the structure solution workflow.

Predict_and_build generates predicted models and uses them to solve an X-ray structure by MR or to interpret a cryo-EM map. The tool then carries out iterative model rebuilding and prediction to improve the models. The iterative procedure allows creating more accurate predicted models than can be obtained with a simple prediction.
Predict and build can automatically generate a fairly accurate model starting from just a sequence file and either cryo-EM half-maps or X-ray data. Additionally, it provides morphed versions of unrefined predicted models that can be useful as reference models for refinement.

Other tools for using predicted models:

- Overview: AlphaFold and Phenix
- Processing a predicted model
- ssed predicted model in a cryo-EM map
- ked predicted model in a cryo-EM map
- king and rebuilding a predicted model in a cryo-EM map
- pping parts of models

Phenix homepage

Phenix Tutorials
346 subscribers

SUBSCRIBED 346

HOME | **VIDEOS** | PLAYLISTS | CHANNELS | ABOUT

Uploads | PLAY ALL | SORT BY

Planning a SAD experiment Tutorial 6:00
Simulate a SAD experiment with...

Map-to-model Tutorial 5:33
Automatic map interpretation with map_to_model

Scale-and-merge Tutorial 7:31
Scaling and merging anomalous data

Automated map sharpening Tutorial 6:05
Automated map sharpening with phenix.autosharpen

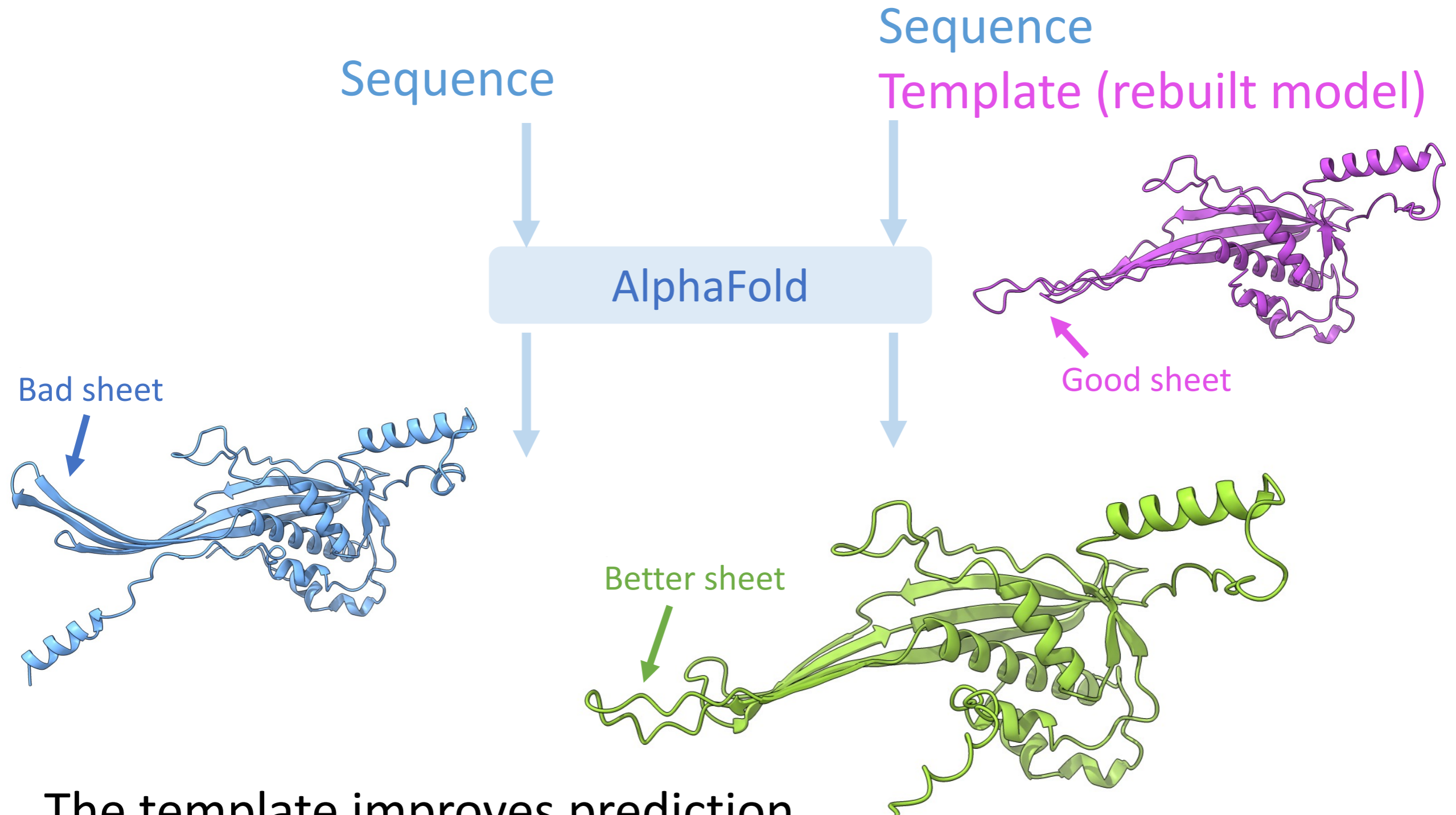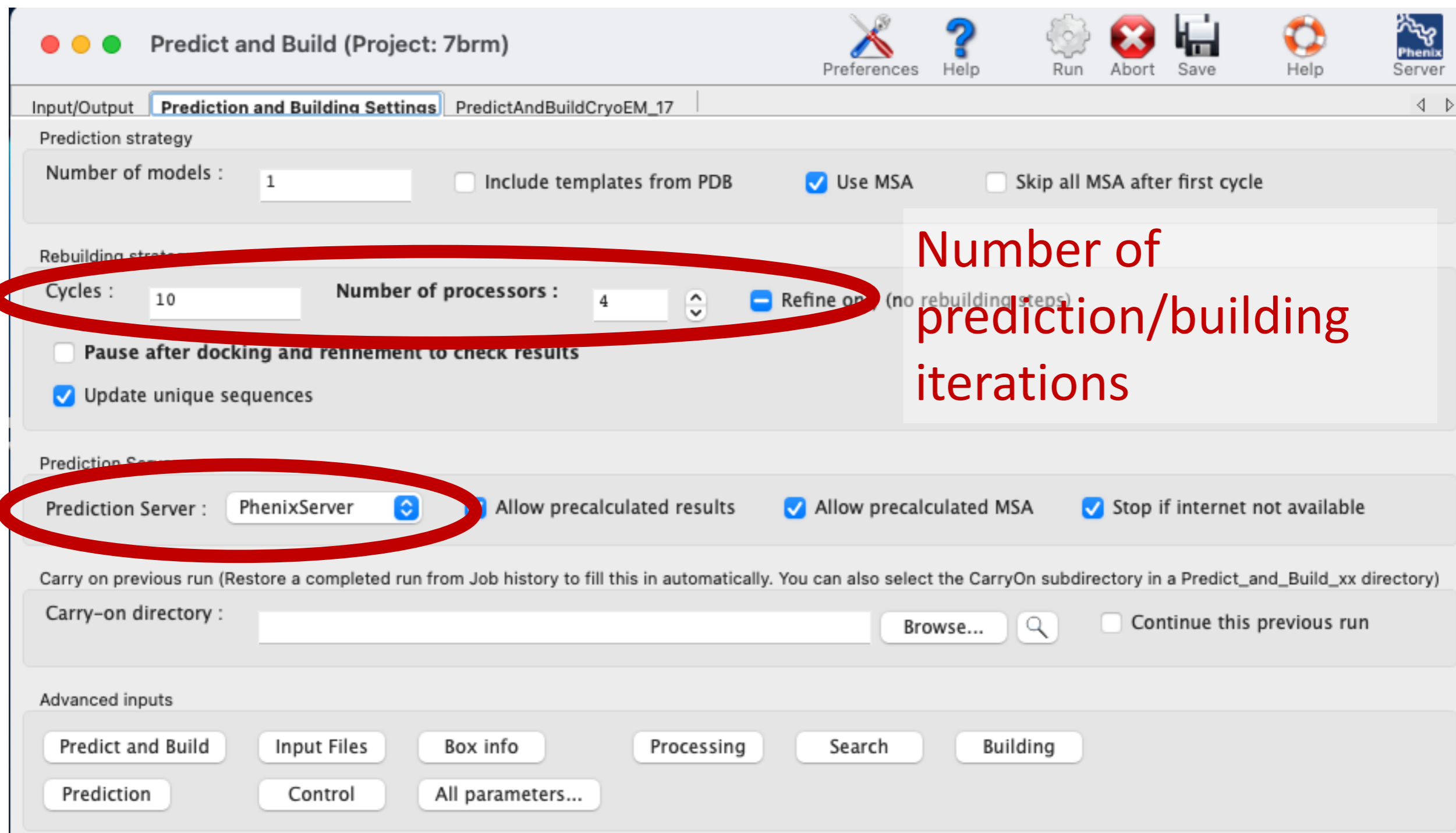**Video tutorials**

AlphaFold changes everything:
Incorporating predicted models in X-ray and Cryo-EM structure determination

Presentation available at:
phenix-online.org/presentations

January 18, 2022
RéNaFoBiS Webinar

Tom Terwilliger

The New Mexico Consortium
Los Alamos National Laboratory

How you can benefit from AlphaFold ...

... by determining structures more easily

Randy Read, Tristan Croll, Claudia Millán, (Cambridge, University), Paul Adams, Billy Poon, Pavel Afonine, Christopher J. Schlicksup (Lawrence Berkeley National Laboratory); Jane Richardson (Duke University)

Phenix · Los Alamos · Berkeley Lab · UNIVERSITY OF CAMBRIDGE

Tom Terwilliger: AlphaFold changes everything

Tom Terwilliger: AlphaFold changes everything

Predicting a model (AlphaFold)

Amino acid sequence

1) **Multiple sequence alignment (MSA)**
   → find AA more likely to mutate
   → find similar structures ("templates")
   → guess which AA are likely to be in contact with each other

*Residues that **co-vary** are probably close in 3D structure*

```
EVQLVESGGGLVQPGGSLRLSCAASGFNIYSSSIHWVRQAPGKGLEWVAYI
...............................F...M........Q........
........................K........Y...L.......A........
............A.......................................V......
...................A.................................
.............................L...V....E..............
.........A.............................Q.....
```

2) **Neural networks:**
Which parts of the MSA are more informative?
Build a model

Structural model