Using predicted models in Phenix

Dorothee Liebschner
Lawrence Berkeley Laboratory
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 $\rightarrow$ 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?
Run AlphaFold on
- Phenix Server
- Colab
Predicted models are great hypotheses...

7s5L, 1.58 Å, X-ray diffraction

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

Data from 7mjs, Cater, R.J., et al. (2021). Nature 595, 315–319
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.

Incorporate predictions into the “conventional” structure determination workflow.

Example:
Cryo-EM map (30160 – 7brm)
Get a prediction

sequence

AlphaFold model
Process prediction

sequence

AlphaFold model

Remove low confidence parts
Processed
AlphaFold model
Dock processed model

Processed AlphaFold model

cryo-EM map
Dock processed model

Docked (processed)  
AlphaFold model

Some parts don’t fit into the map
Dock processed model

Docked-processed AlphaFold model

Some parts don’t fit into the map → fit loops and rebuild
Fit loops and rebuild

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

**Crystallography**
- Data quality assessment
- Model prediction
- Molecular Replacement
- Model (Re)building
- Refinement
- Validation

**Cryo-EM**
- Data quality assessment
- Map improvement
- Model prediction
- Docking
- Model (Re)building
- Refinement
- Validation
Improving prediction

AlphaFold

7brm
What happens if we use the rebuilt model as template for AlphaFold?
• The template improves prediction
• New prediction can be even better than the template

Iterate with Predict and Build

- Cycles: 10
- Number of processors: 4
- Prediction Server: PhenixServer

Number of prediction/building iterations
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**
- **Rescore a predicted model in a cryo-EM map**
- **Fitting and rebuilding a predicted model in a cryo-EM map**
- **Trimming parts of models**

---

**Video tutorials**

Tom Terwilliger: AlphaFold changes everything
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*

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

---

**Amino acid sequence**

```
EVQLVESGGGLVQPGGSLRLSCAASGFNIYSSSIHWVRQAPGKGLEWVAYI
```

Structural model