



Phenix Tools for Cryo-EM

Christopher Schlicksup

Postdoc, Phenix Consortium
Lawrence Berkeley National Lab, California, USA

March 9, 2022
OpenEye CUP Conference

The Phenix Project

Lawrence Berkeley Laboratory

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



University of Cambridge

Randy Read, Airlie McCoy,
Tristan Croll, Claudia Millán Nebot,
Rob Oeffner, Massimo Sammito,
Duncan Stockwell



*An NIH/NIGMS funded
Program Project*

New Mexico Consortium

Los Alamos National Laboratory

Tom Terwilliger, Li-Wei Hung



Baylor College of Medicine

Matt Baker, Corey Hryc



Duke University

Jane & David Richardson,
Chris Williams, Vincent Chen



Liebschner et al., Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Cryst.* 2019 **D75**:861-877

Phenix



Search Phenix website

A comprehensive software package for macromolecular structure determination for crystallographic (X-ray, neutron and electron) and electron cryo-microscopy data.

Important Note:
Starting July 2019, the Protein Data Bank requires models to be in mmCIF for crystallographic structures. You should use the latest official release to generate these files for deposition.

Download

Getting started

Workshops & Tutorials

Documentation

Help

Developers

National Resource

Industrial Consortium

Metrics

Newsletter

Publications

CCTBX

Citing PHENIX:
Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix
D. Liebschner, P. V. Afonine, M. L. Baker, O. Bunkóczi, V. B. Chen, T. I. Croll, B. Hintze, L.-W. Hung, S. Jain, A. J. McCoy, N. W. Moriarty, R. D. Oeffner, B. K. Poon, M. G. Prisant, R. J. Read, J. S. Richardson, D. C. Richardson, M. D. Sammito, O. V. Sobolev, D. H. Stockwell, T. C. Terwilliger, A. G. Urzhumtsev, L. L. Videau, C. J. Williams, and P. D. Adams
Acta Cryst. (2019). D75, 861-877

Phenix Development, Maintenance and Distribution is Supported by:
NIH/NIGMS Program Project Grant ()
NIH/NIGMS R24 National Resource Grant ()
The Phenix Industrial Consortium

Contact
[Follow on Twitter](#)
[Email us](#)

Disclaimer
[Privacy & Security Notice](#)
[About this website](#)

[Back to top](#)

Website

PHENIX home

Quit Preferences Help Citations Coot PyMOL KING Other tools Ask for help

Actions Job history

Projects

Show group: All groups Manage...

Select Delete New project Settings


ID	Last modified	# of jobs	R-free
----	---------------	-----------	--------

Crystals: Data analysis and manipulation
Validation and map-based comparisons
Experimental phasing
Molecular replacement
Maps (create, manipulate, compare)
Enhanced maps (Polder, FEM, density-modified...)
Model building
Refinement
Ligands
Cryo-EM: Map analysis, symmetry, manipulation
Validation and map-based comparisons
Map improvement
Docking, model building and rebuilding
Refinement
Models: Superpose, search, compare, analyze symmetry
Modification, minimization and dynamics
PDB Deposition
Program search

Current directory: /Users/user/Dropbox/Mac/Documents Browse...

PHENIX version 1.20-4444-000 Project: test

GUI



STRUCTURAL BIOLOGY

ISSN 2059-7983

Received 26 July 2019
Accepted 15 August 2019

Edited by K. Diederichs, University of Konstanz, Germany

feature articles

Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in *Phenix*

Dorothee Liebschner,^a Pavel V. Afonine,^a Matthew L. Baker,^b Gábor Bunkóczi,^{c,*} Vincent B. Chen,^d Tristan I. Croll,^c Bradley Hintze,^{d,g} Li-Wei Hung,^e Swati Jain,^{d,g} Airlie J. McCoy,^c Nigel W. Moriarty,^a Robert D. Oeffner,^c Billy K. Poon,^a Michael G. Prisant,^d Randy J. Read,^c Jane S. Richardson,^d David C. Richardson,^d Massimo D. Sammito,^c Oleg V. Sobolev,^a Duncan H. Stockwell,^c Thomas C. Terwilliger,^{c,f} Alexandre G. Urzhumtsev,^{g,h} Lizbeth L. Videau,^d Christopher J. Williams^d and Paul D. Adams^{a,i,*}

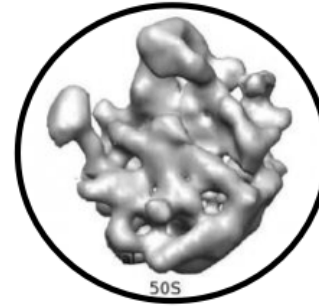
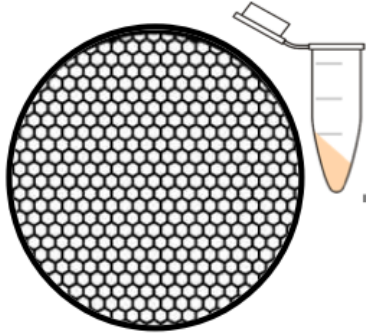
Most recent publication

Typical cryo-EM pipeline

Sample

Data collection

Data processing

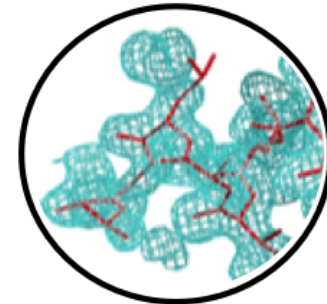
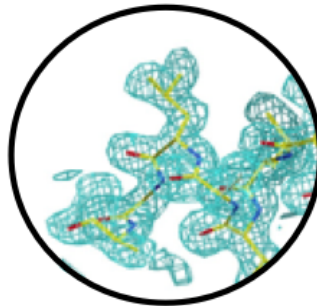
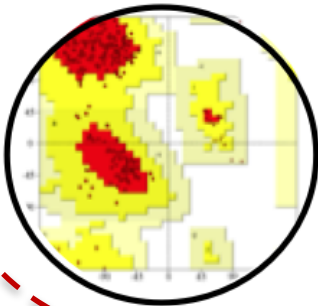


Map manipulations

Validation

Model refinement

Model building



Phenix

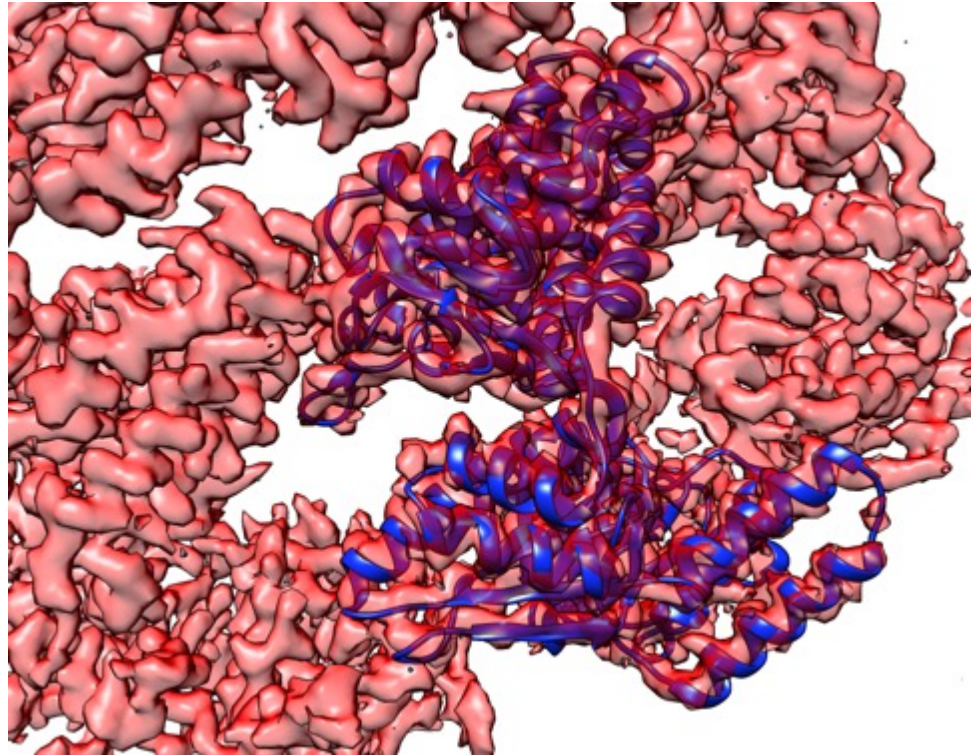
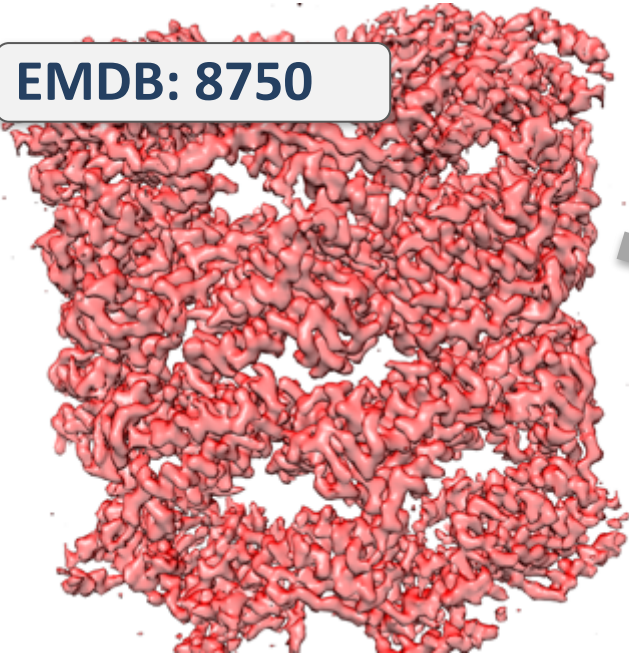
Agenda: A long list of cryo-EM tools within Phenix



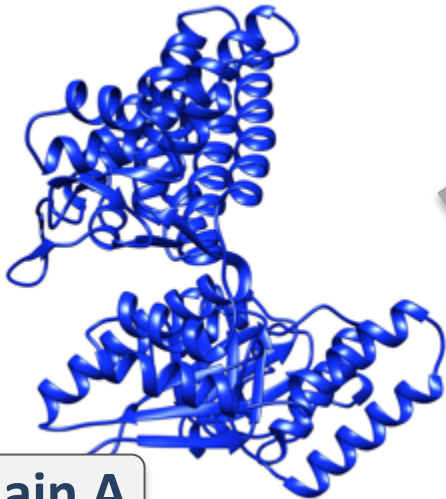
If a little is good, more must be better - Mae West (adapted)

Docking models with *phenix.dock_in_map*

EMDB: 8750



Chain A docked

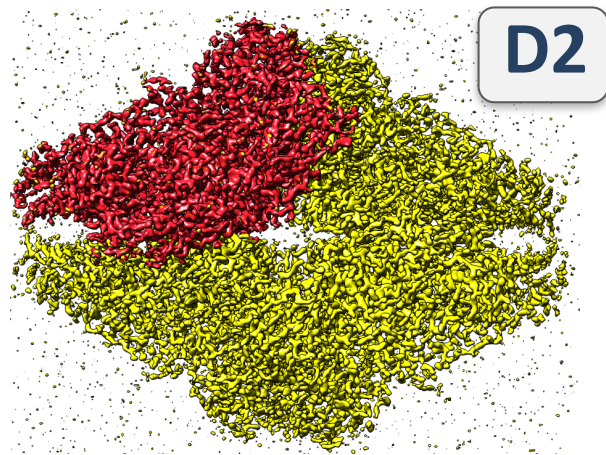
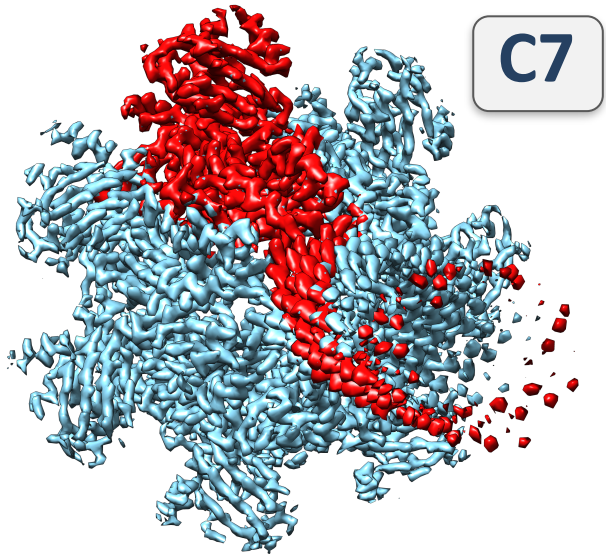


1ss8 chain A

Features:

- Multiple chains
- Density search
- Symmetry
- Multiprocessing

Finding map symmetry: *phenix.symmetry_from_map*



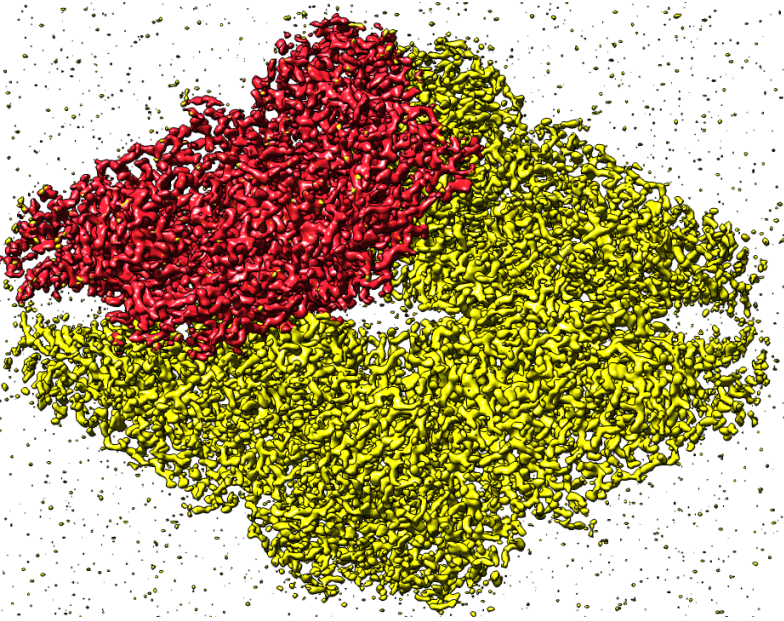
Procedure for finding symmetry:

- Test point group symmetries (e.g., C7, D2, I, O, T)
- Helical symmetry
- Score based on map correlation for symmetry-related points and number of operators

<http://phenix-online.org/newsletter/>

Tools for interpreting cryo-EM maps using models from the PDB

Extracting unique part of map: *phenix.map_box*



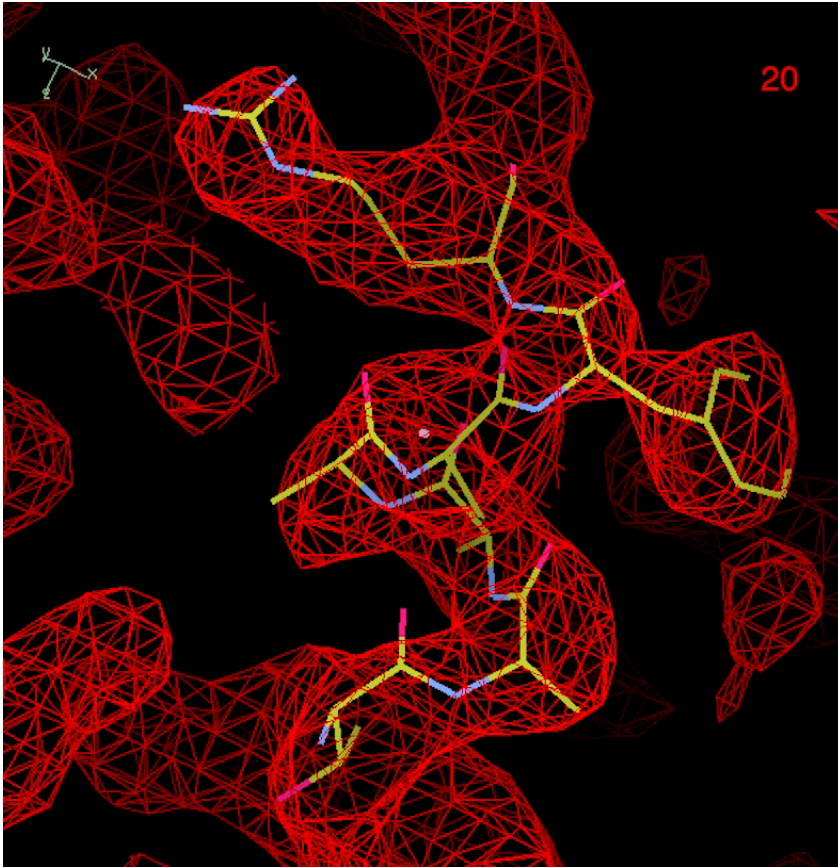
Procedure:

- Use symmetry of map
- Contour map at level that yields regions about 50 residues in size
- Group symmetry-related regions
- Choose one member of each group
- Optimize compactness and connectivity of unique part of map

<http://phenix-online.org/newsletter/>

Tools for interpreting cryo-EM maps using models from the PDB

Automated map sharpening: *phenix.auto_sharpen*



Maximize detail in the map

... and connectivity of map

Adjusted surface area



Optimally sharpened map

Fully automatic:

- No manual trial-and-error
- No parameters to adjust
- Only inputs: map and resolution



STRUCTURAL
BIOLOGY

ISSN 2059-7983

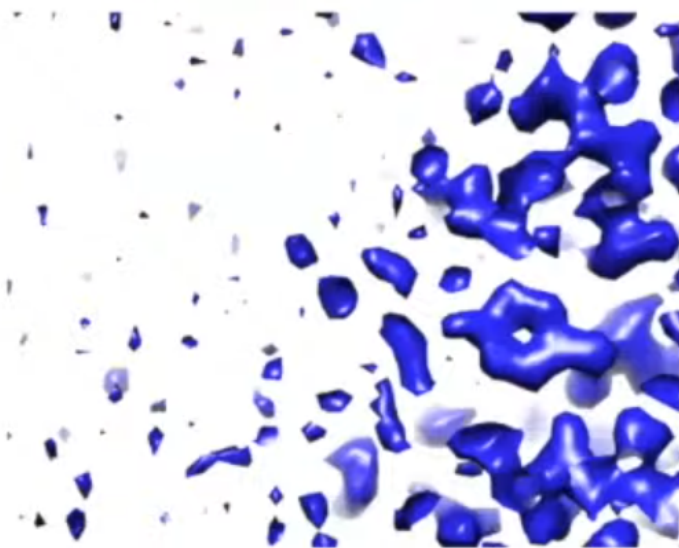
Automated map sharpening by maximization of
detail and connectivity

Thomas C. Terwilliger,^{a,b*} Oleg V. Sobolev,^c Pavel V. Afonine^{c,d} and
Paul D. Adams^{d,e}

Density modification: *phenix.density_modify_cryo_em*

Using expectations about **one** part of a map to improve **another** part of the map

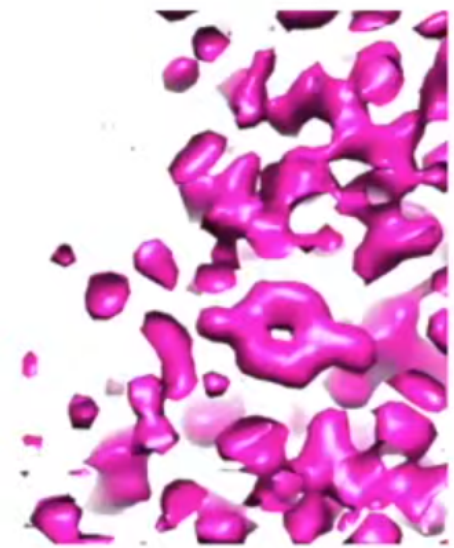
Original map



Solvent should be flat

Distribution of density (histograms) should match typical protein

Density modified map



nature **methods**

ARTICLES

<https://doi.org/10.1038/s41592-020-0914-9>

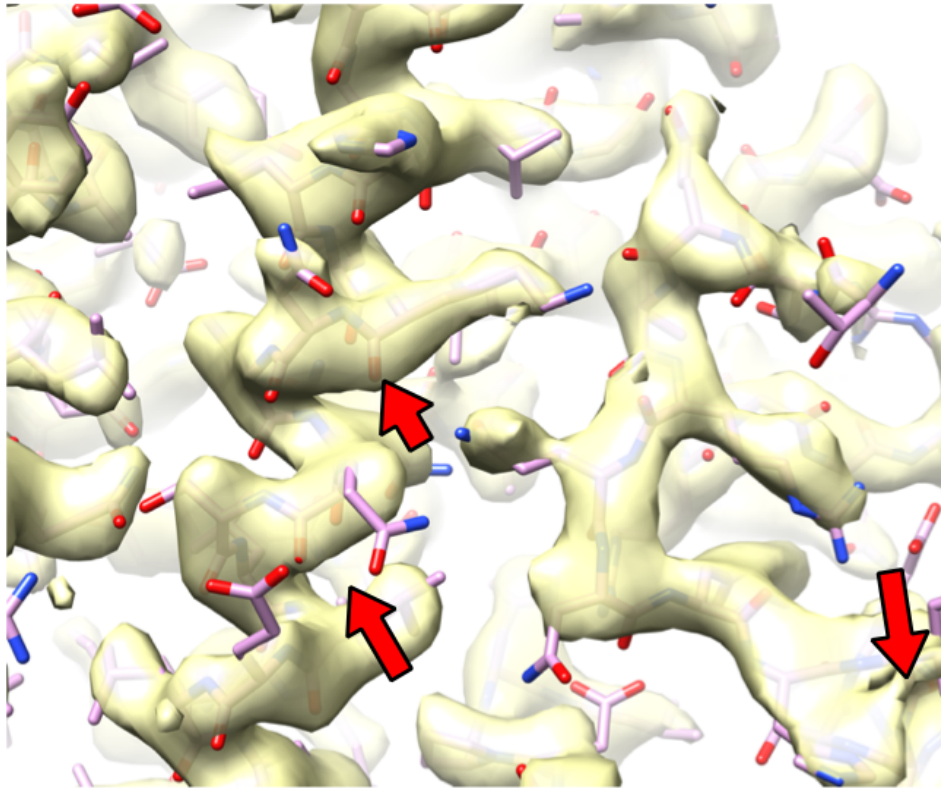
 Check for updates

Improvement of cryo-EM maps by density modification

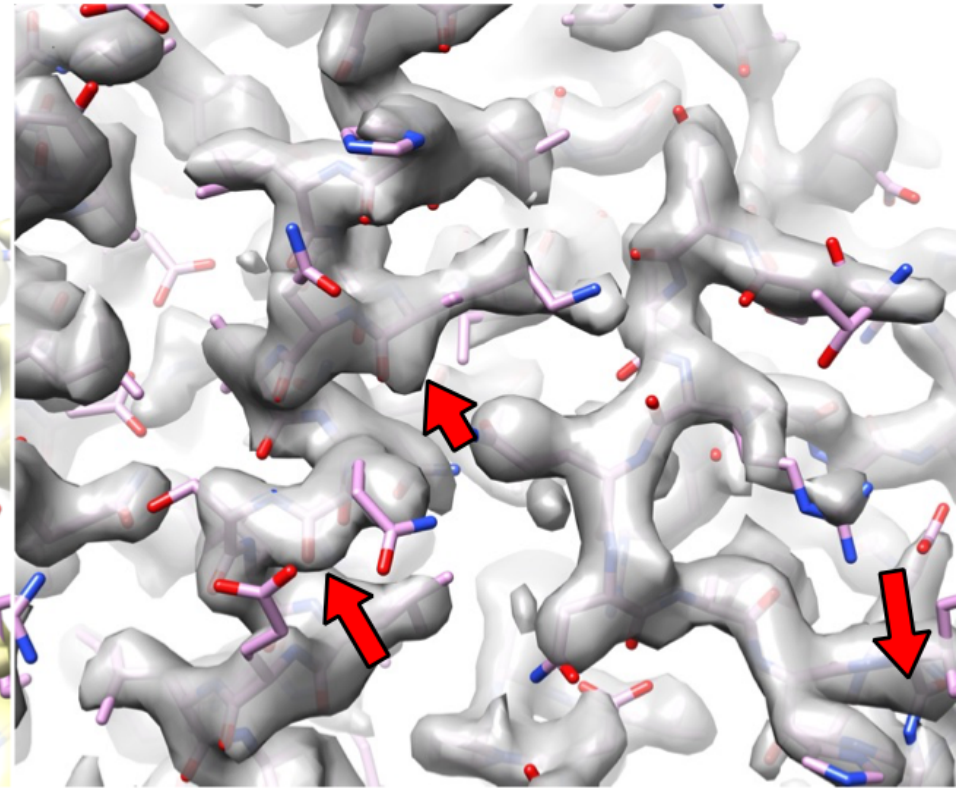
Thomas C. Terwilliger^{1,2}✉, Steven J. Ludtke³, Randy J. Read⁴, Paul D. Adams^{5,6} and Pavel V. Afonine⁵

Density modification: *phenix.density_modify_cryo_em*

Original map

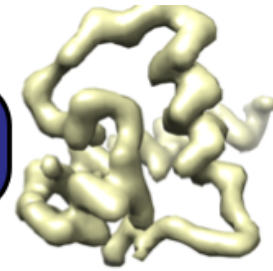


Density modified map

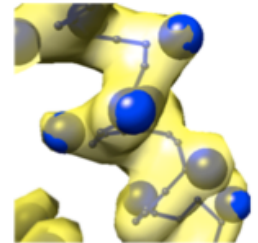


Automated model building: *phenix.map_to_model*

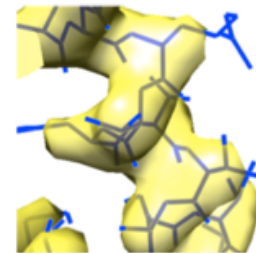
Isolate density for a chain



Identify C_α and C_β positions from side-chain density



Construct and refine all-atom model



nature **methods**

BRIEF COMMUNICATION

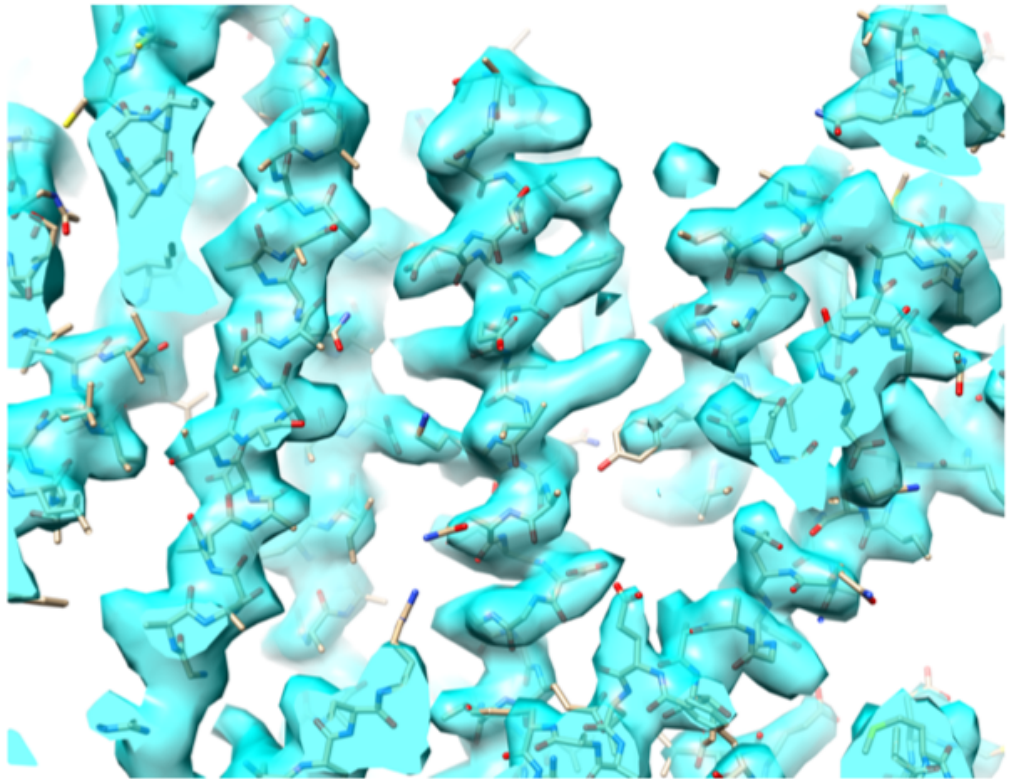
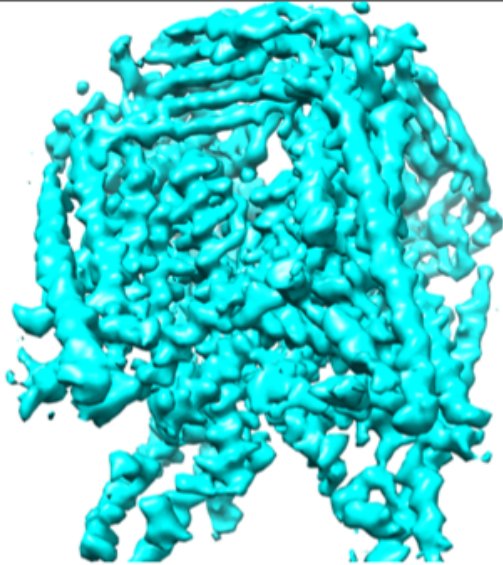
<https://doi.org/10.1038/s41592-018-0173-1>

A fully automatic method yielding initial models from high-resolution cryo-electron microscopy maps

Thomas C. Terwilliger^{1,2*}, Paul D. Adams^{3,4}, Pavel V. Afonine^{3,5} and Oleg V. Sobolev³

Automated model building: *phenix.map_to_model*

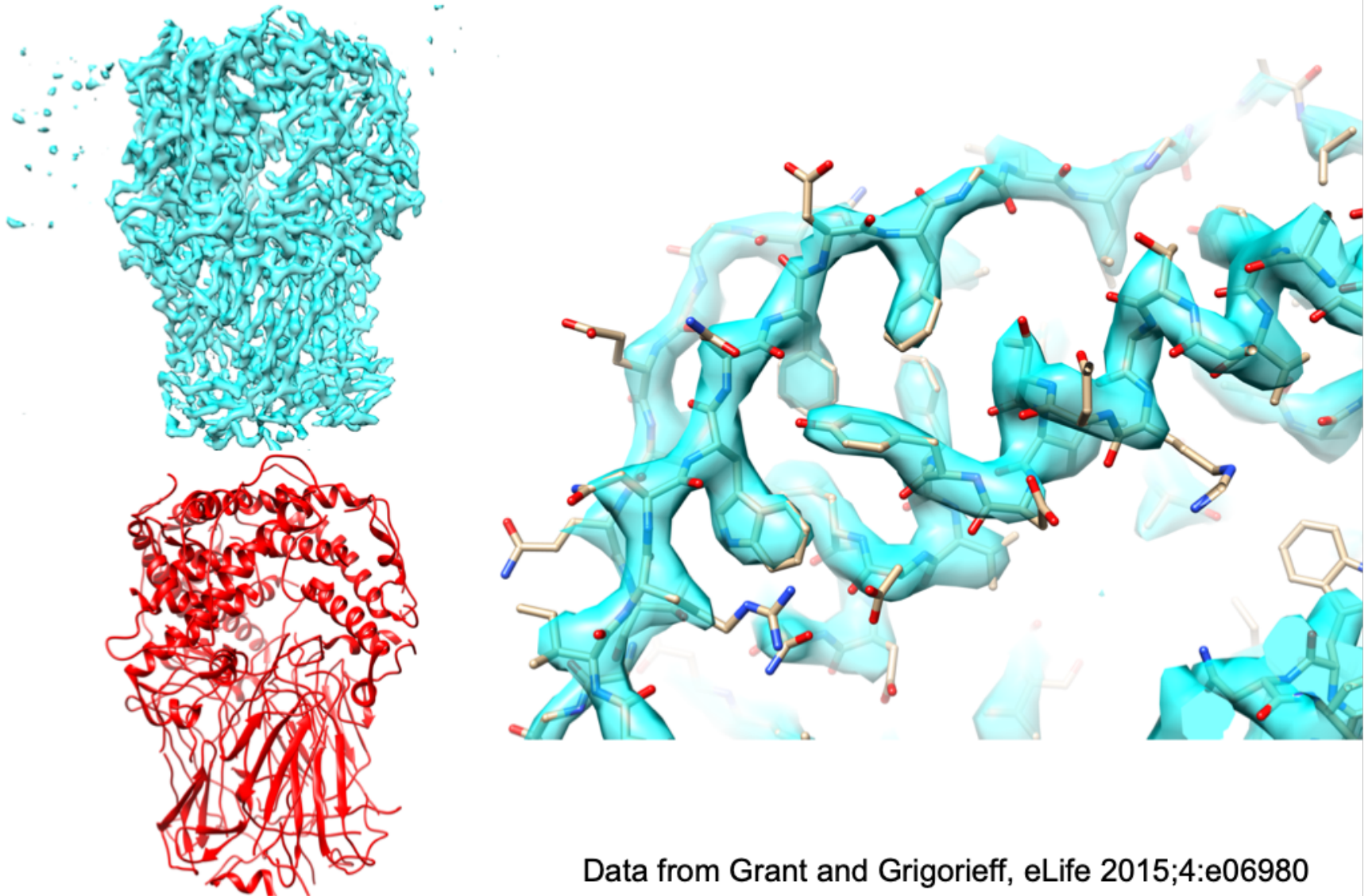
TRPML3 channel (4.1 Å, 78% built, 1.3 Å rmsd)



Data from Zhou, X. et al. (2017) Nat. Struct. Mol. Biol. 24: 1146

Automated model building: *phenix.map_to_model*

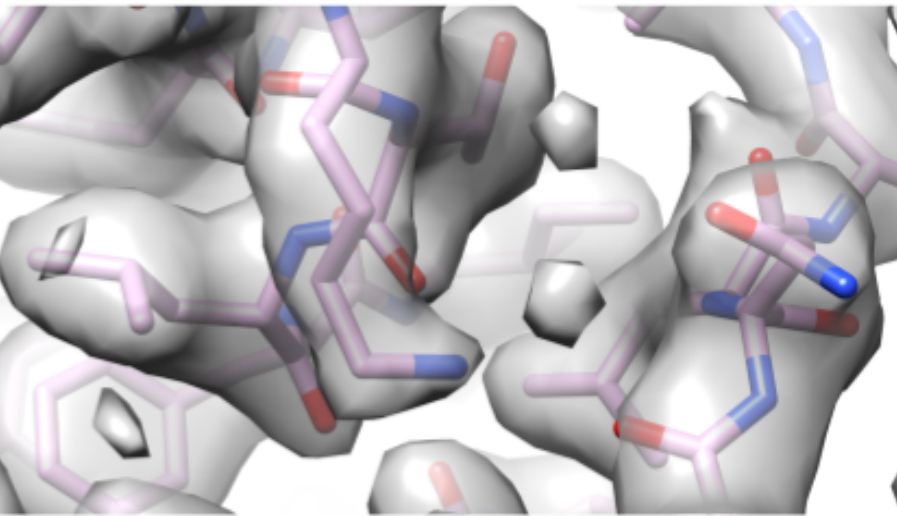
Rotavirus VP6 (2.6 Å, 100% built, 0.9 Å rmsd)



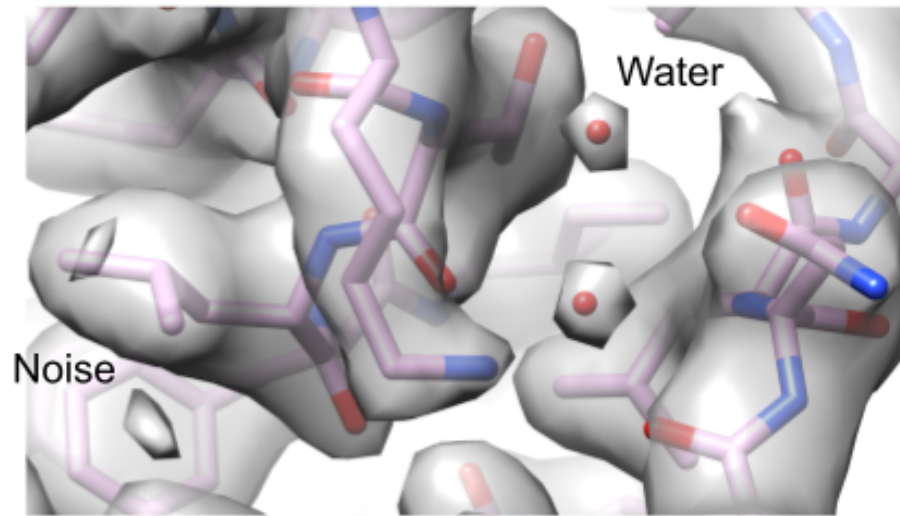
Data from Grant and Grigorieff, eLife 2015;4:e06980

Automated water building: *phenix.douse*

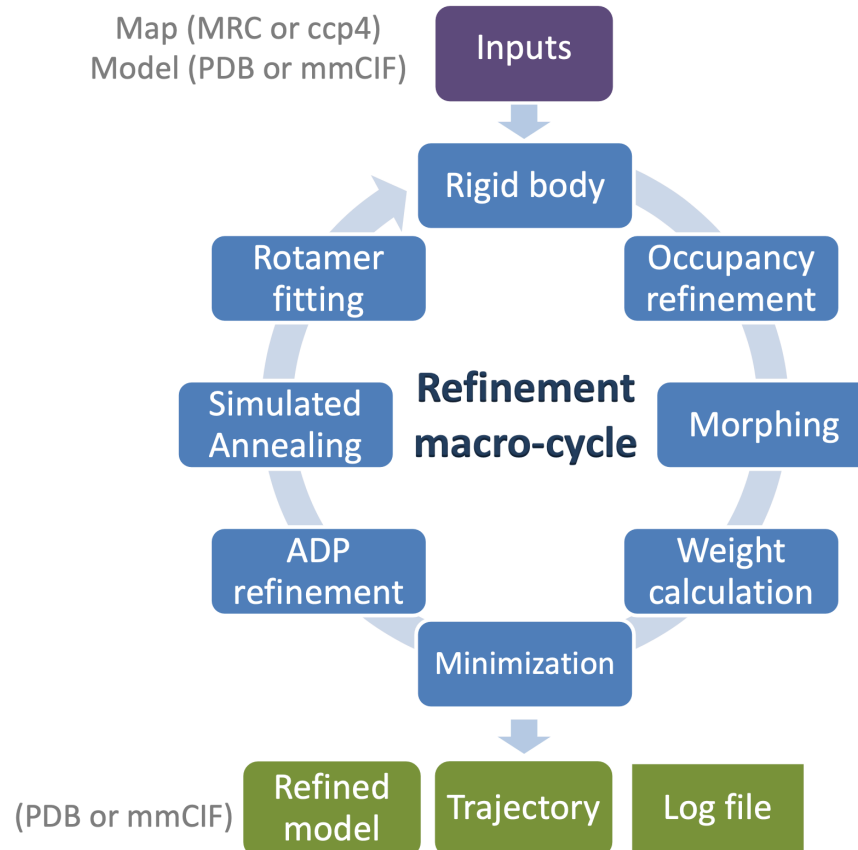
Beta-gal, isolated density peaks
near LYS 897



Automated water placement



Atomic model refinement: *phenix.real_space_refine*



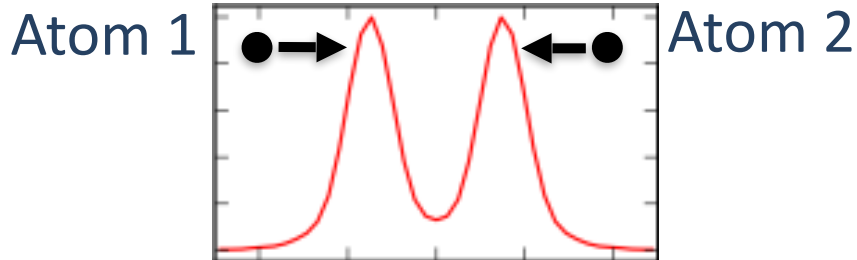
Refinement target

- Atom-centered:

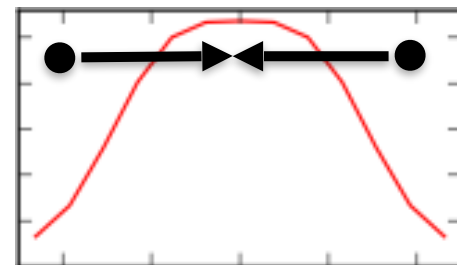
$$T = - \sum_{atoms} \rho_{obs}(x_{atom}, y_{atom}, z_{atom})$$

$x_{atom}, y_{atom}, z_{atom}$ = coordinates of atom center

High resolution

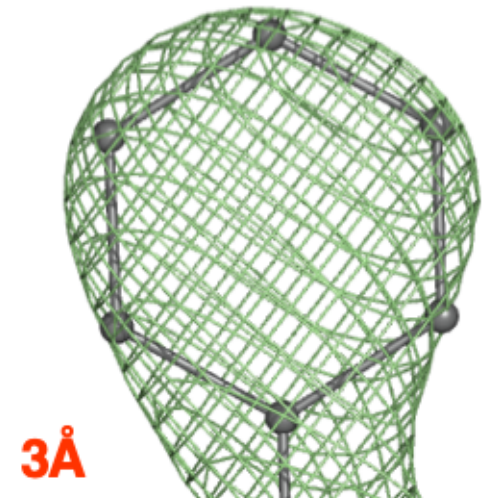
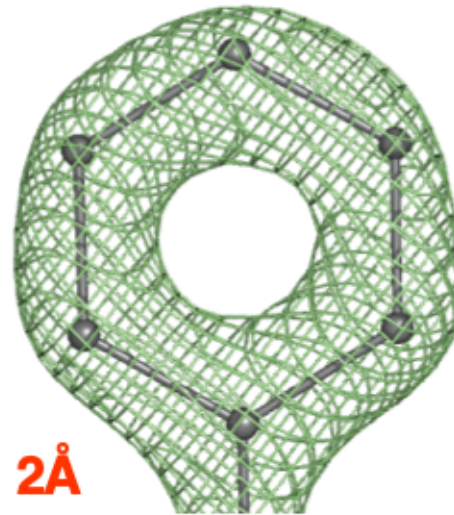
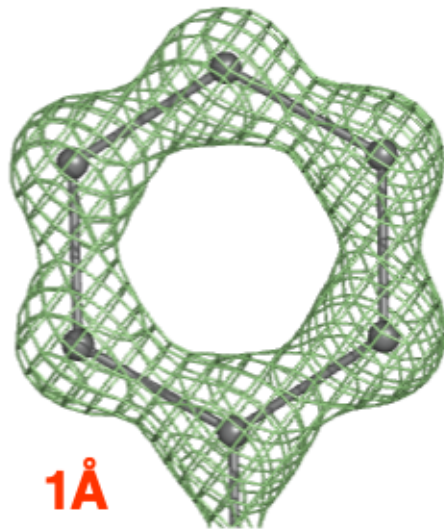


Low resolution



Moving atoms to nearest peaks \neq making correct model

Restraints



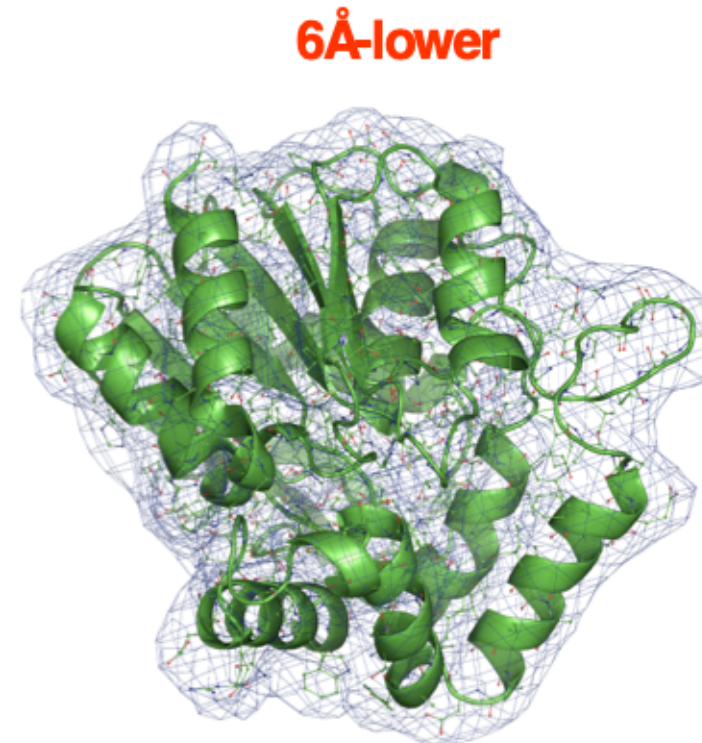
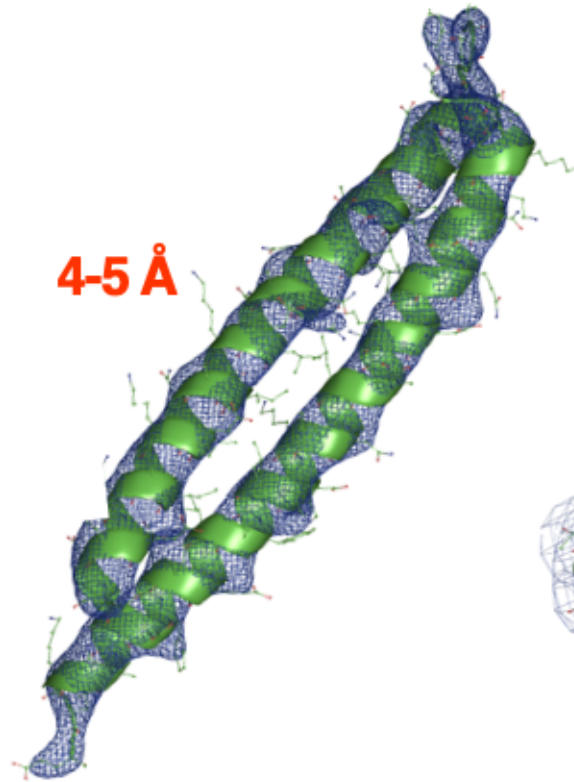
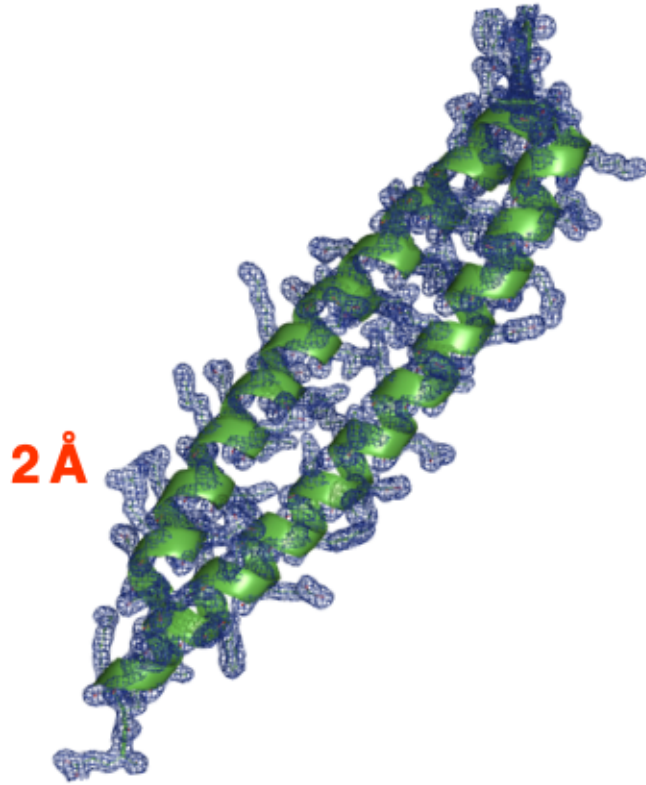
- Lower the resolution, less detailed the map
- Need extra information to keep correct geometry during refinement

$$T = T_{\text{DATA}} + wT_{\text{RESTRAINTS}}$$

$$T_{\text{RESTRAINTS}} = T_{\text{BOND}} + T_{\text{ANGLE}} + T_{\text{DIHEDRAL}} + T_{\text{PLANARITY}} + T_{\text{NONBONDED}} + T_{\text{CHIRALITY}}$$

Restraints

- Low resolution map is not sufficient to maintain secondary



- Secondary structure
- Ramachandran plot
- RNA/DNA parallelity
- Reference model
- Rotamer, starting position, etc

Ligand parameterization: *phenix.elbow*

What if you have a small molecule ligand?

1. We may ship restraints for that ligand:
 - Restraints are pre-calculated, archived in our GeoStd library
 - This is a subset of the CCD, molecules that have been previously appeared in macromolecular structures
2. Your ligand is new:
 - Need to use phenix.elbow to generate a restraints file
 - Restraints written to a cif file (eq. values for bonds, angles, torsions, etc)
 - Can use internal simplified force field, or external plugins (QM packages, Mogul, etc)

electronic Ligand Builder and Optimization Workbench (eLBOW): a tool for ligand coordinate and restraint generation

Nigel W. Moriarty,^{a*}  Ralf W. Grosse-Kunstleve^a and Paul D. Adams^{a,b}

^aLawrence Berkeley National Laboratory, One Cyclotron Road, Mailstop 64R0246, Berkeley, CA 94720, USA, and ^bDepartment of Bioengineering, UC Berkeley, CA 94720, USA

*Correspondence e-mail: nwmoriarty@lbl.gov



(Received 27 April 2009; accepted 23 July 2009)

The *electronic Ligand Builder and Optimization Workbench (eLBOW)* is a program module of the *PHENIX* suite of computational crystallographic software. It is designed to be a flexible procedure that uses simple and fast quantum-chemical techniques to provide chemically accurate information for novel and known ligands alike. A variety of input formats and options allow the attainment of a number of diverse goals including geometry optimization and generation of restraints.

<https://github.com/phenix-project/geostd>

Using AlphaFold via Phenix GUI

Sequence

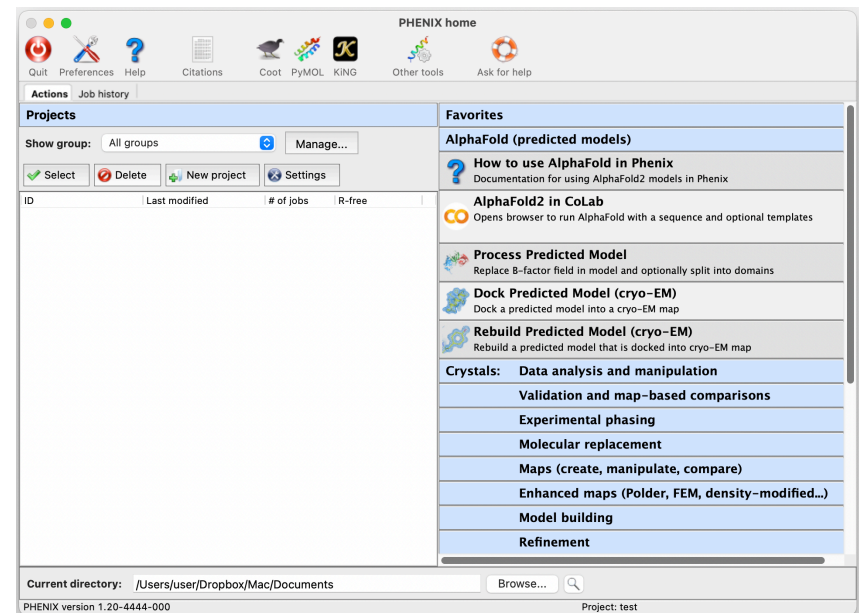
Templates from PDB

Templates you have created

AlphaFold Colab

AlphaFold

*Prediction
(hypothesis)*



Highly accurate protein structure prediction with AlphaFold

[John Jumper](#), [Richard Evans](#), ... [Demis Hassabis](#) [+ Show authors](#)

Nature **596**, 583–589 (2021) | [Cite this article](#)

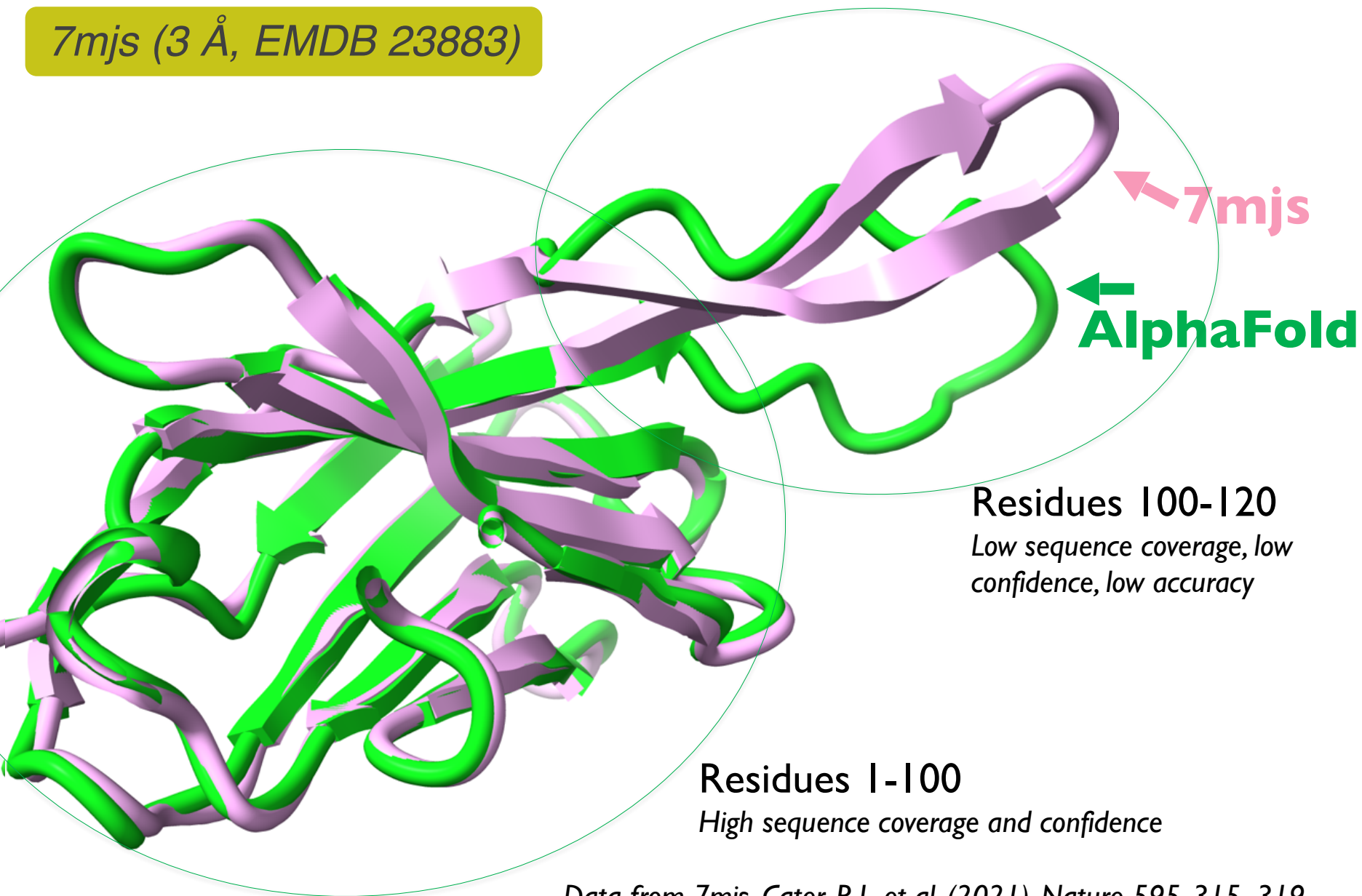
ColabFold - Making protein folding accessible to all

[Milot Mirdita](#), [Konstantin Schütze](#), [Yoshitaka Moriwaki](#), [Lim Heo](#), [Sergey Ovchinnikov](#), [Martin Steinegger](#)

doi: <https://doi.org/10.1101/2021.08.15.456425>

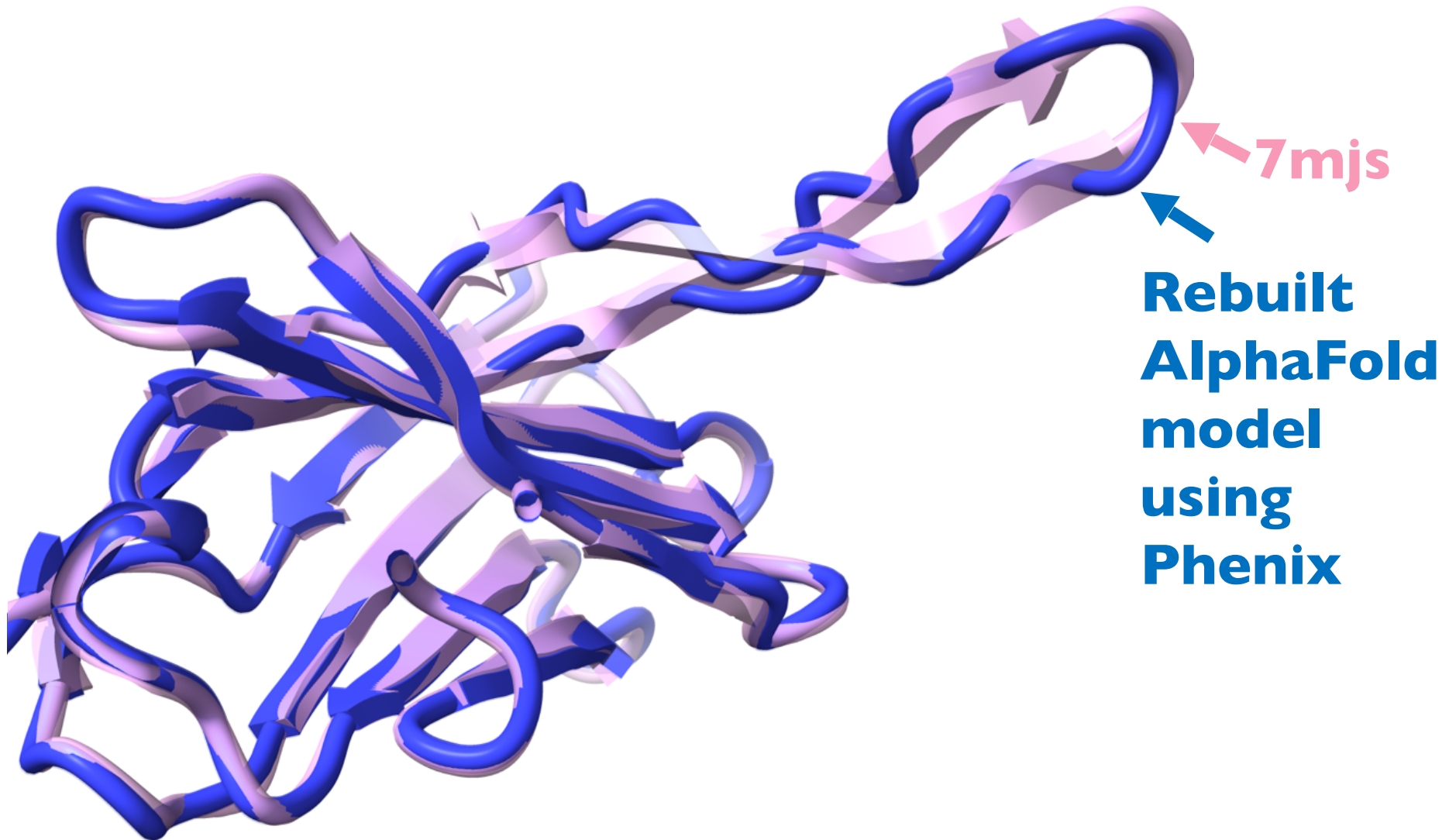
Models are accurate where sequence coverage is high

7mjs (3 Å, EMDB 23883)

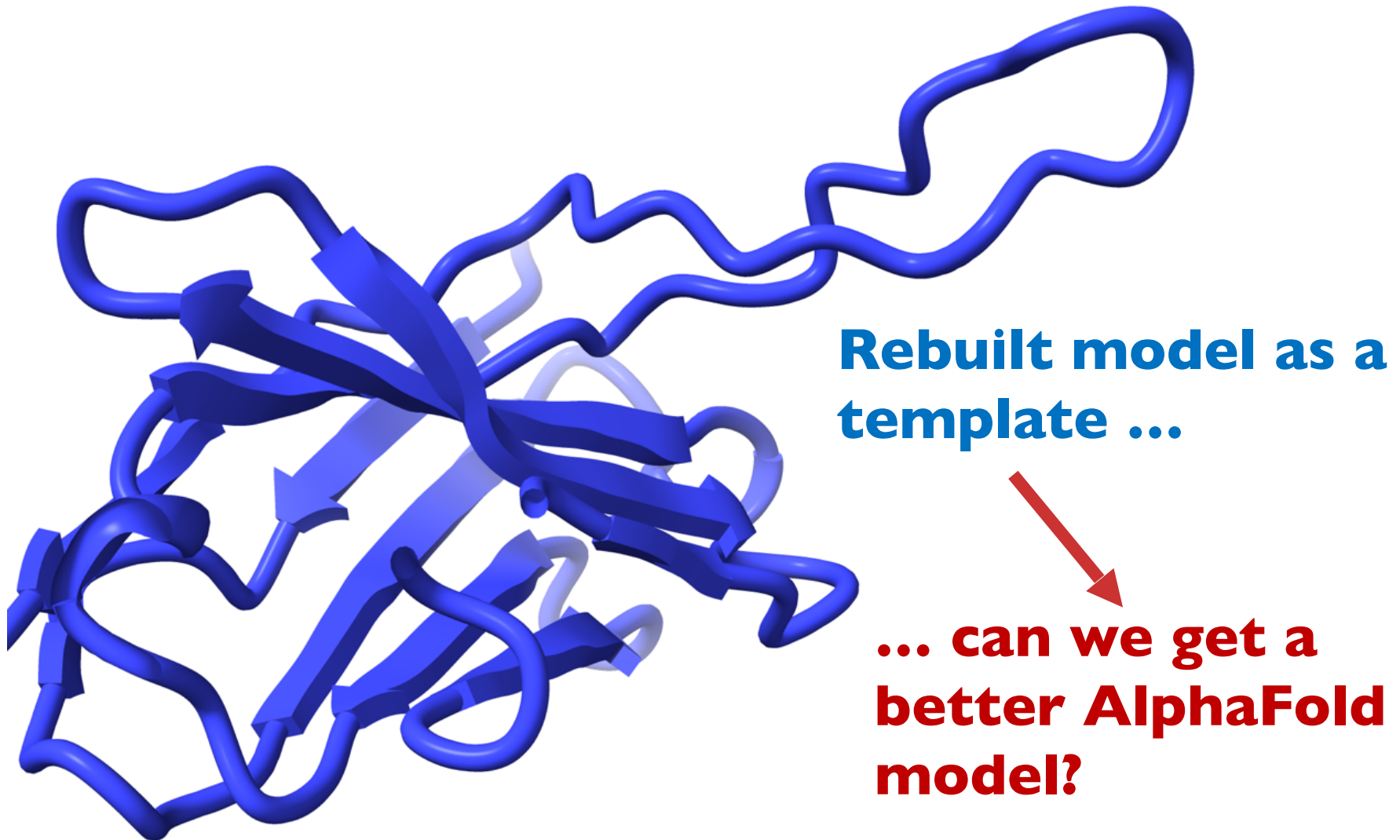


Data from 7mjs, Cater, R.J., et al. (2021). Nature 595, 315–319

AlphaFold as a starting model (7mjs 3 Å, EMDDB 23883)



Iterative AlphaFold prediction and rebuilding



Iterative AlphaFold prediction and rebuilding



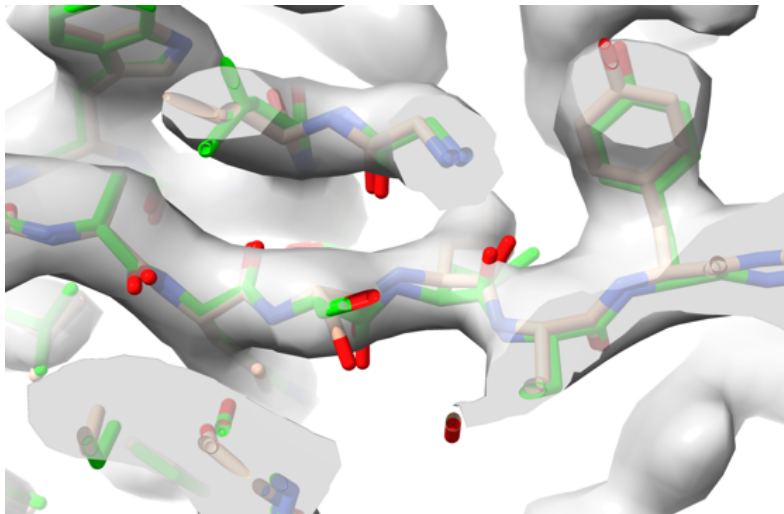
Rebuild AlphaFold model with density map



Rebuilt model improves next AlphaFold prediction

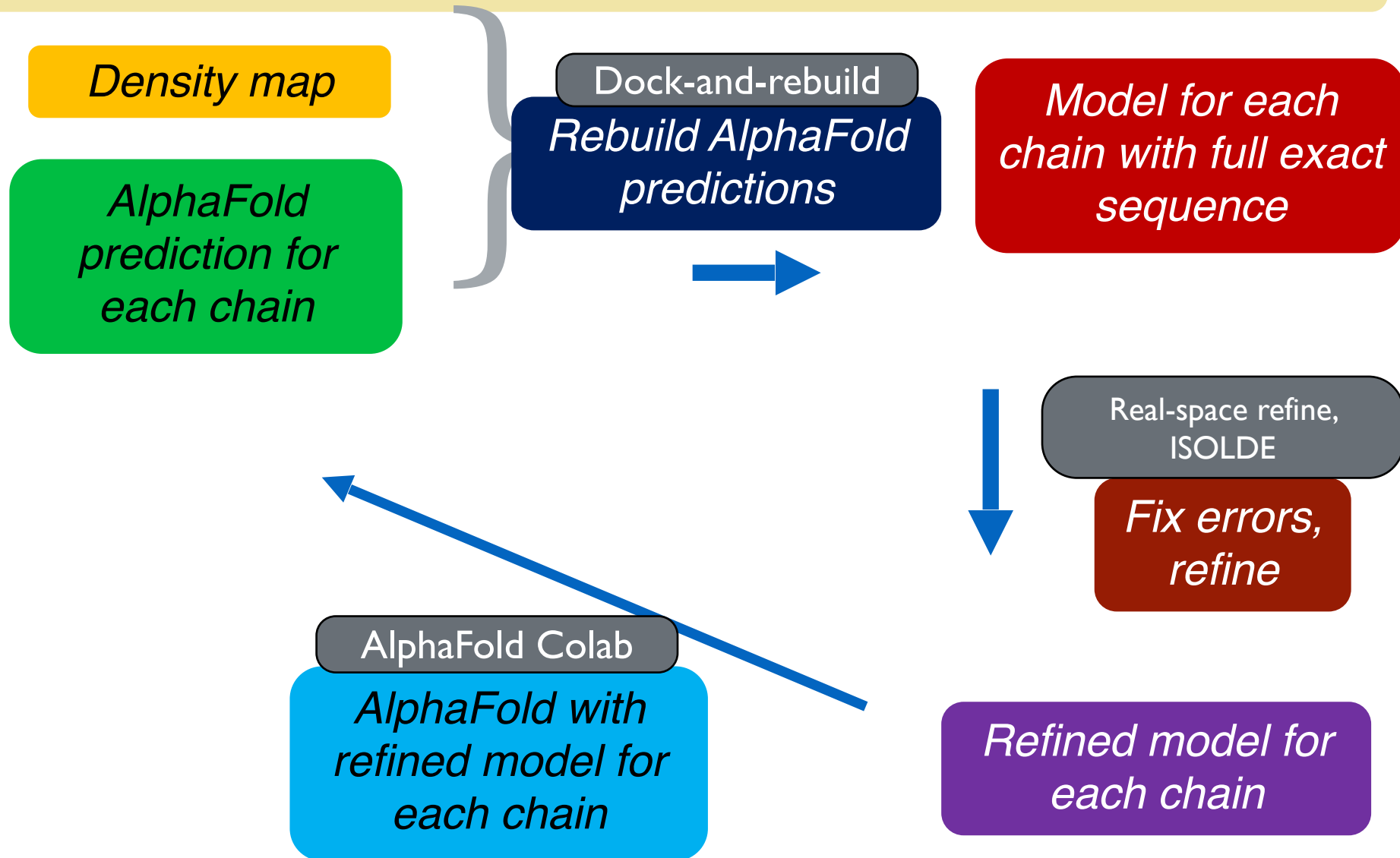


Iterate to improve model



Terwilliger et al. (2022). Improving AlphaFold modeling using implicit information from experimental density maps. *BioRxiv* 2022.01.07.475350

Workflow for Iterative Predict/Rebuild



User support

- **Feedback, questions, help**

Mailing list (all, developers and users): phenixbb@phenix-online.org

Bug reports (developers only): bugs@phenix-online.org

Ask for help (developers only): help@phenix-online.org

- **Reporting a bug or asking for help:**

- Please include a detailed description of the problem!
- Make sure the problem still exist using the latest *Phenix* version
- Send us all inputs (files, non-default parameters) and tell us steps that lead to the problem
- All data sent to us kept confidentially

Thanks!