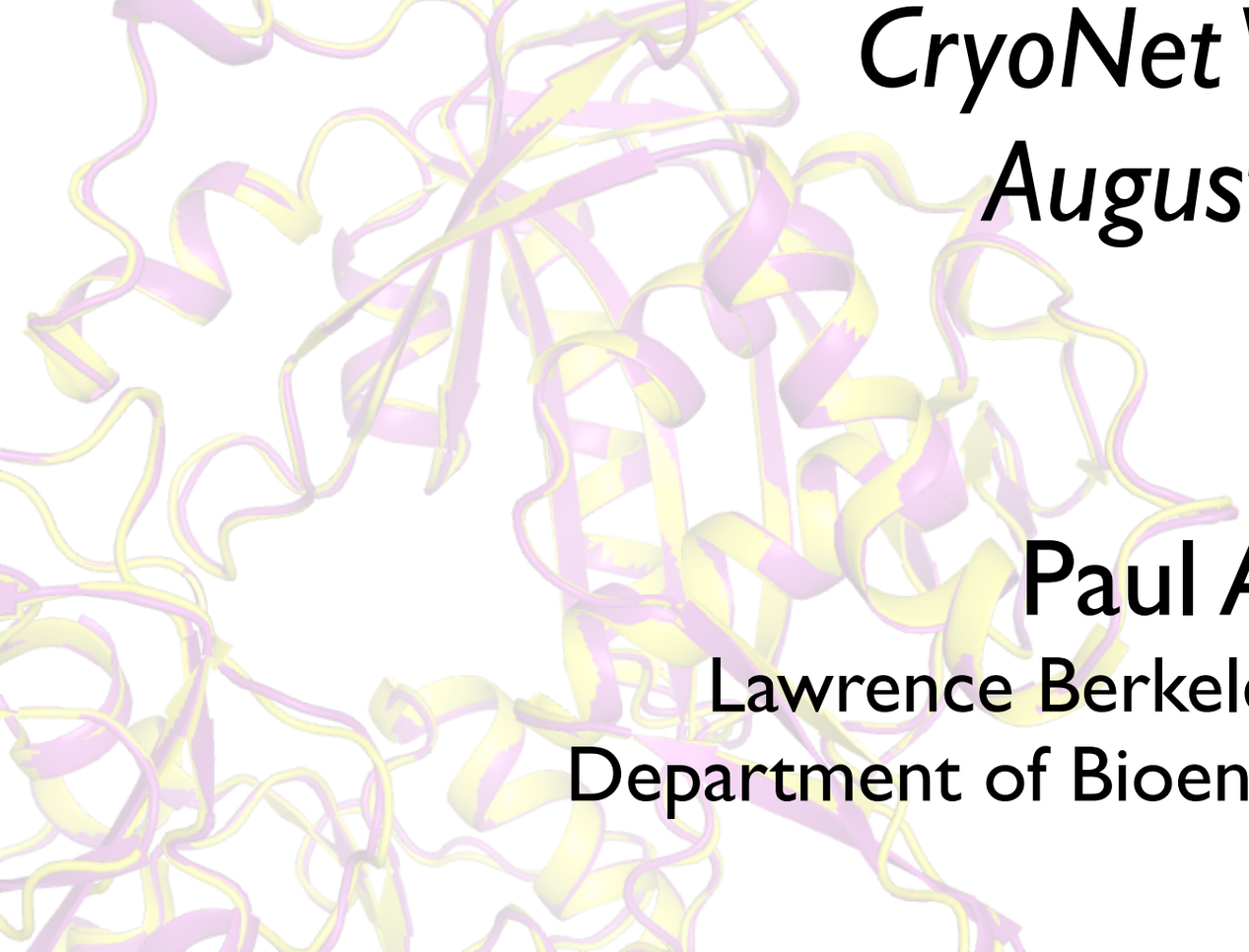
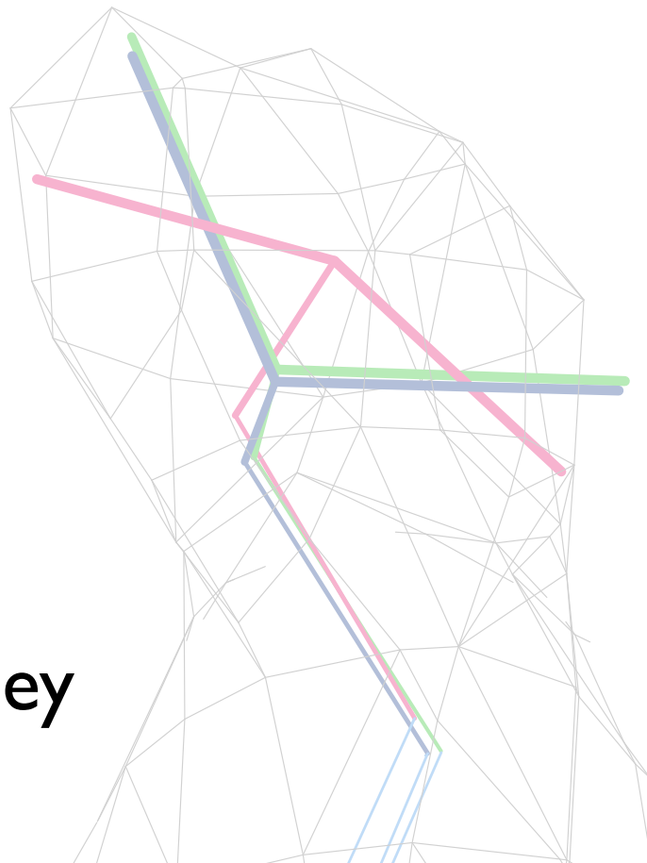
Models from Near Atomic Resolution Cryo-EM Data

CryoNet Workshop
August 2018

Paul Adams
Lawrence Berkeley Laboratory and
Department of Bioengineering UC Berkeley





The Phenix Project

Lawrence Berkeley Laboratory

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

Los Alamos National Laboratory

Tom Terwilliger, Li-Wei Hung

Cambridge University

Randy Read, Airlie McCoy, Tristan Croll, Rob Oeffner

Duke University

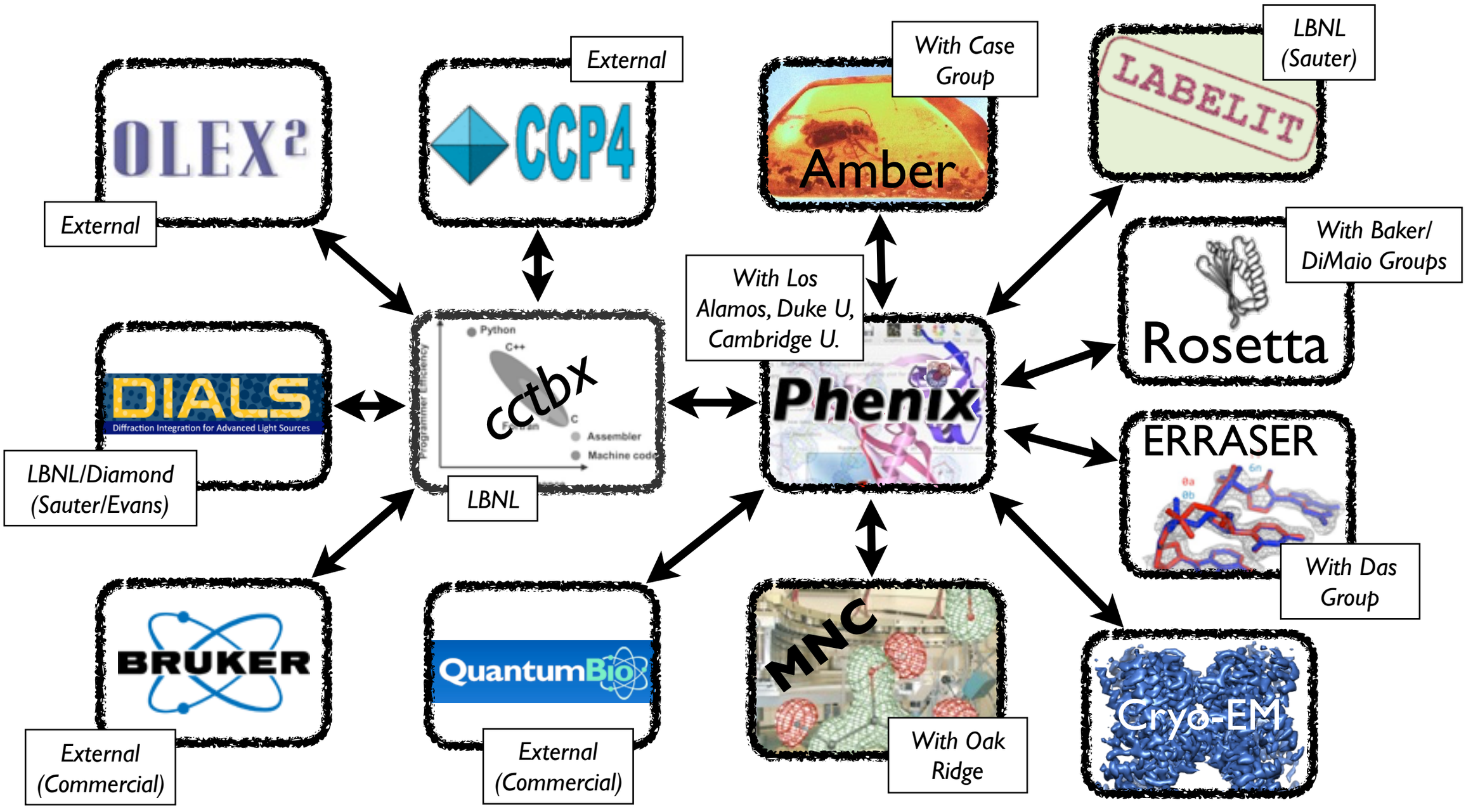
Jane & David Richardson, Chris Williams, Vincent Chen, Bradley Hintze

An NIH/NIGMS funded Program Project

Adams PD et al., PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Cryst.* 2010, **D66**:213-221.

Phenix

Phenix - a Structural Biology Hub

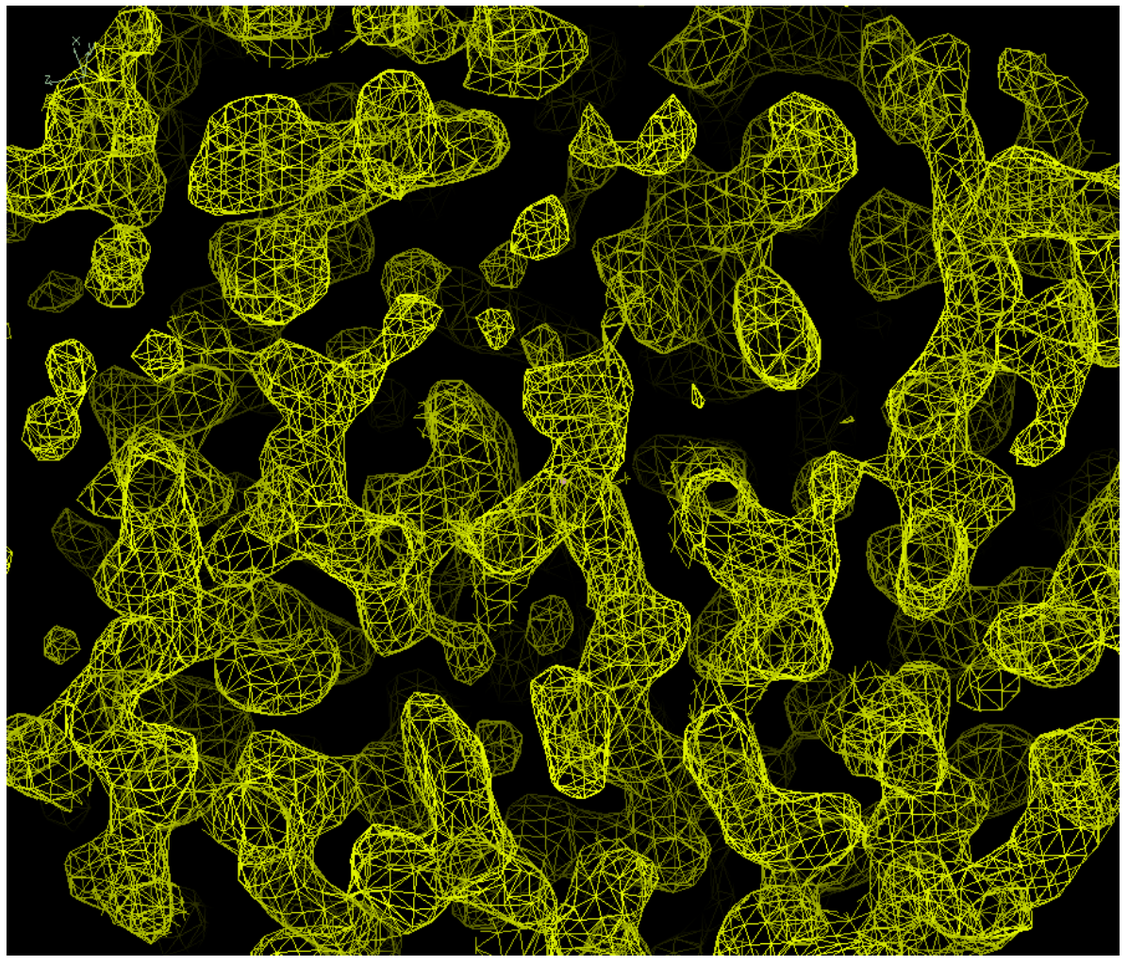


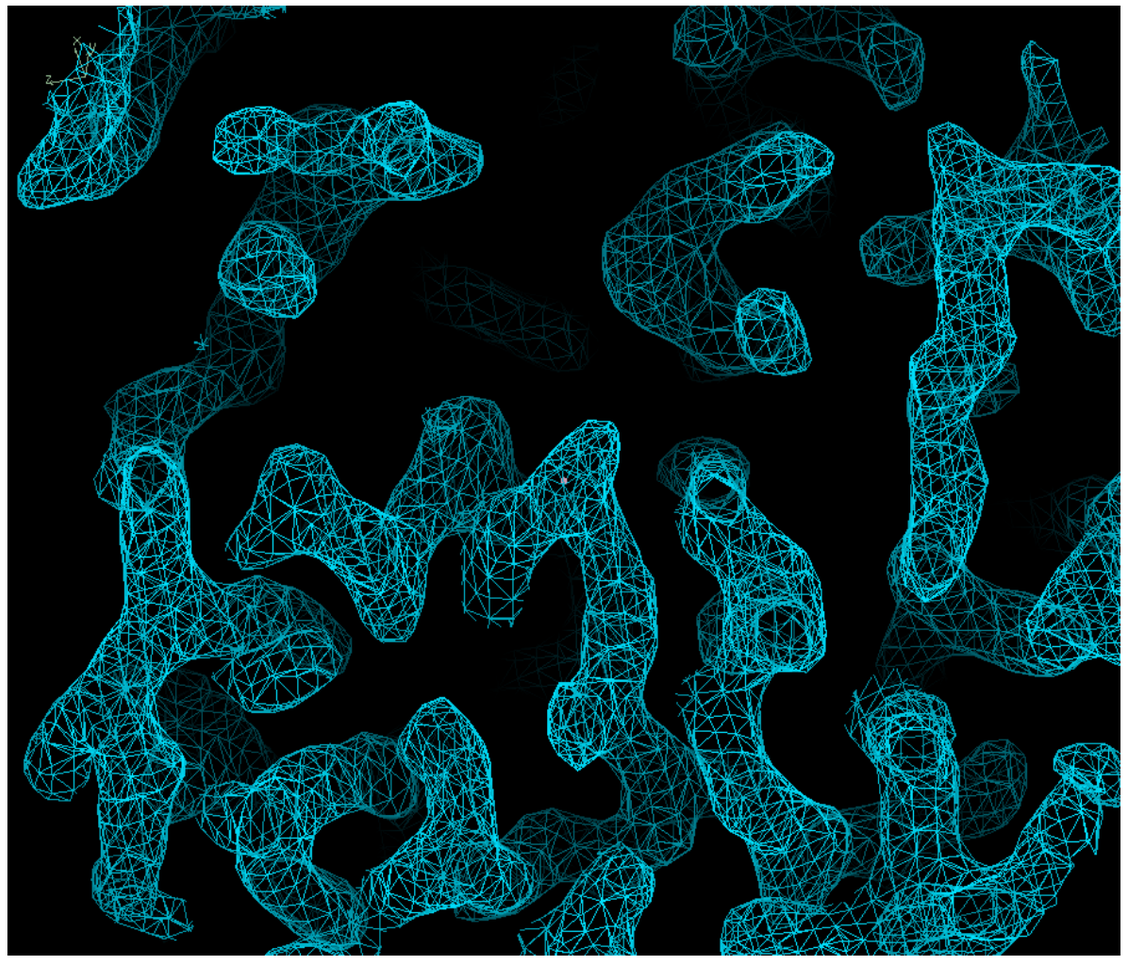
Phenix

BERKELEY LAB
Lawrence Berkeley National Laboratory

Crystallographic vs. Cryo-EM Maps

Beta galactosidase at 2.2 Å





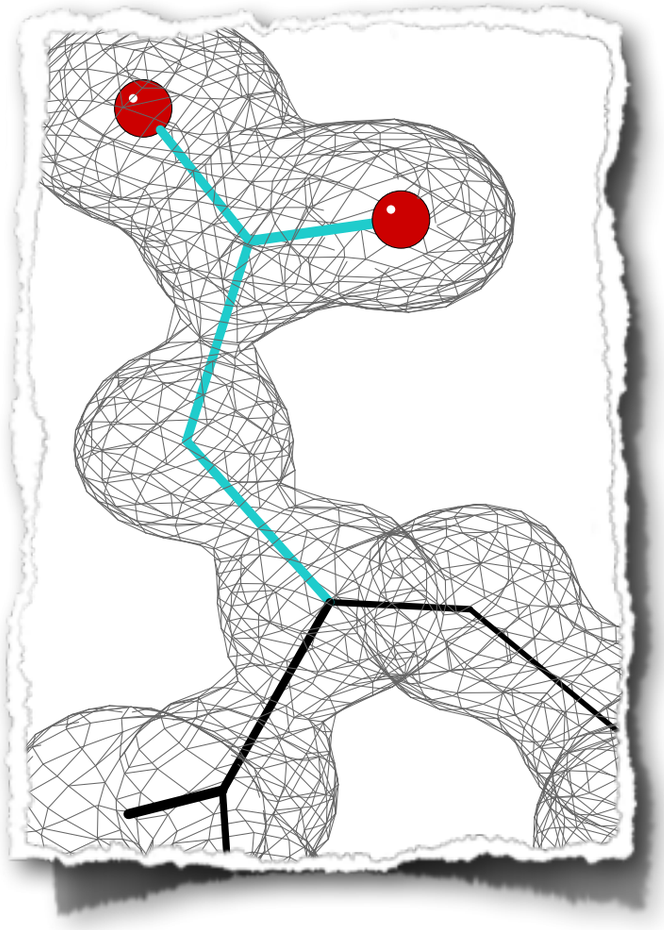
BERKELEY LAB
Lawrence Berkeley National Laboratory

Phenix

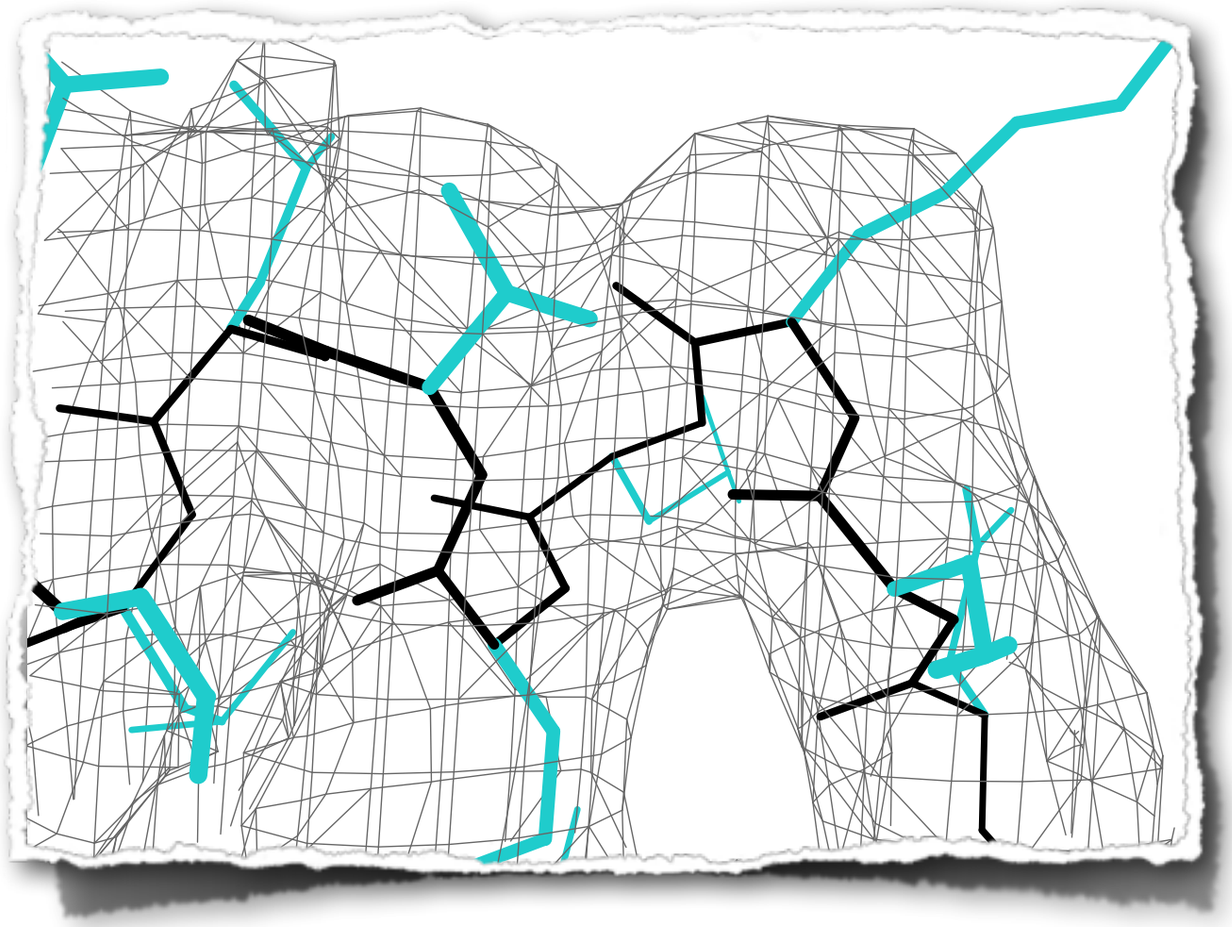
Tom Terwilliger, Los Alamos National Lab

Low Resolution

PDBID: 2gkg
Resolution: 1.00Å



PDB ID: 3k7a
Resolution: 3.80Å



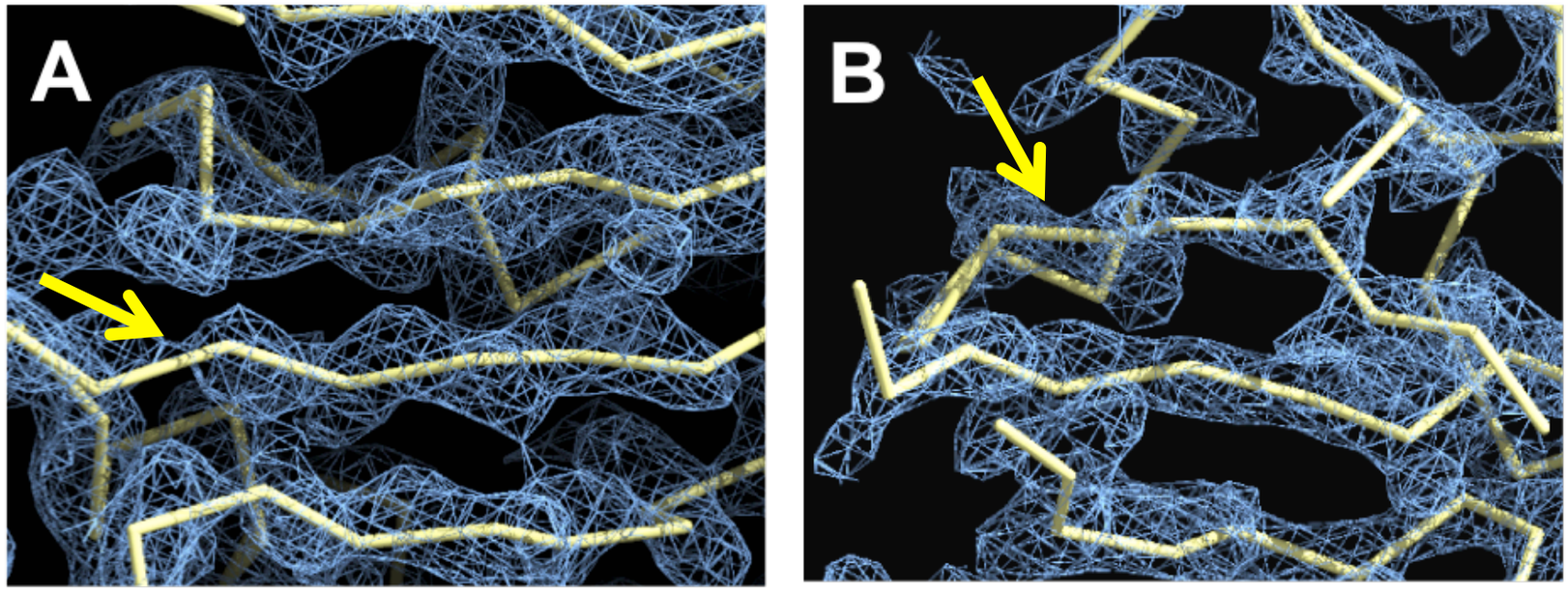
- Many challenges:
  - How to interpret "featureless" maps (pattern matching, chemical constraints)
  - How to optimize models with sparse data (prior information)

BERKELEY LAB
Lawrence Berkeley National Laboratory

Phenix

THE UNIVERSITY OF CALIFORNIA · BERKELEY · 1868

More Accurate Low Resolution Information in Cryo-EM Maps



Original

Phenix

Tom Terwilliger, Los Alamos National Lab

Challenges

- Automated model building
  - What is the magnification of the map? (can be 5% uncertainty)
  - What is the optimal sharpening of the map?
  - What is the region containing the molecule?
  - Low and variable resolution across maps
- Structure optimization
  - Variable resolution across maps
  - Large molecules
  - Poor initial models
- Validation
  - How to validate a model against moderate resolution maps

BERKELEY LAB
Lawrence Berkeley National Laboratory

Phenix

THE UNIVERSITY OF CALIFORNIA · BERKELEY · 1868

New Tools for Cryo-EM in Phenix

- Determine symmetry from a map
- **Automated map sharpening**
- **Map segmentation**
- **Rigid model docking**
- Flexible fitting with CryoFit (Karissa Sanbonmatsu, Los Alamos)
- **Automated model building**
- Identify sequence from map
- **Real space refinement**
- Combine focused maps
- **Comprehensive model and map validation**

BERKELEY LAB
Lawrence Berkeley National Laboratory

Phenix

THE UNIVERSITY OF CALIFORNIA BERKELEY 1868

Automated Model Docking

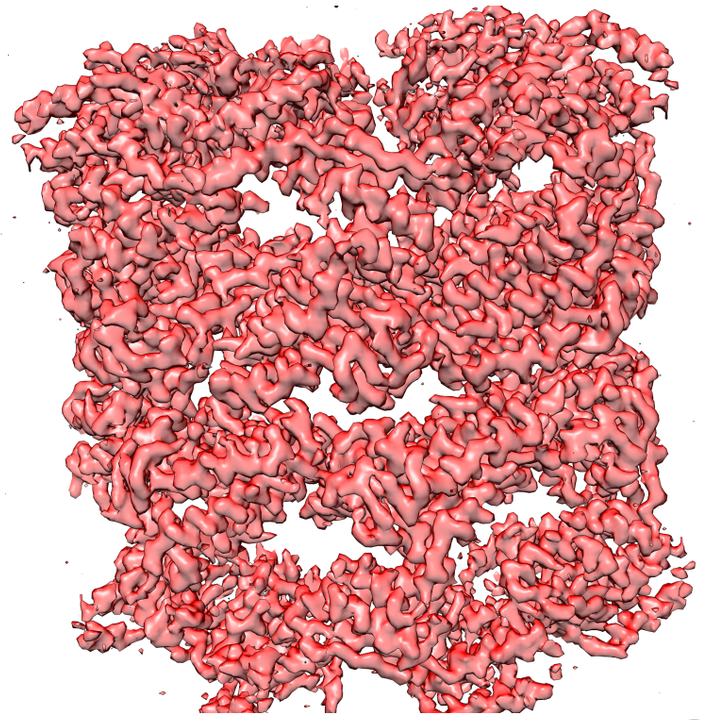
Tom Terwilliger
Los Alamos National Laboratory
Pavel Afonine, Oleg Sobolev
Lawrence Berkeley National Laboratory

BERKELEY LAB
Lawrence Berkeley National Laboratory

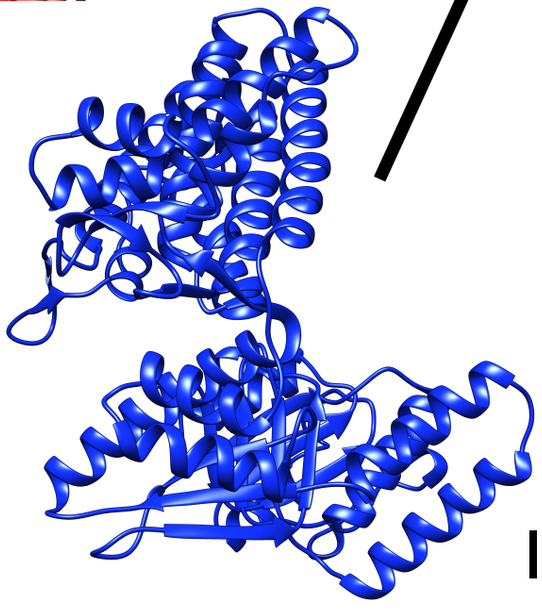
Phenix

Automated Model Docking

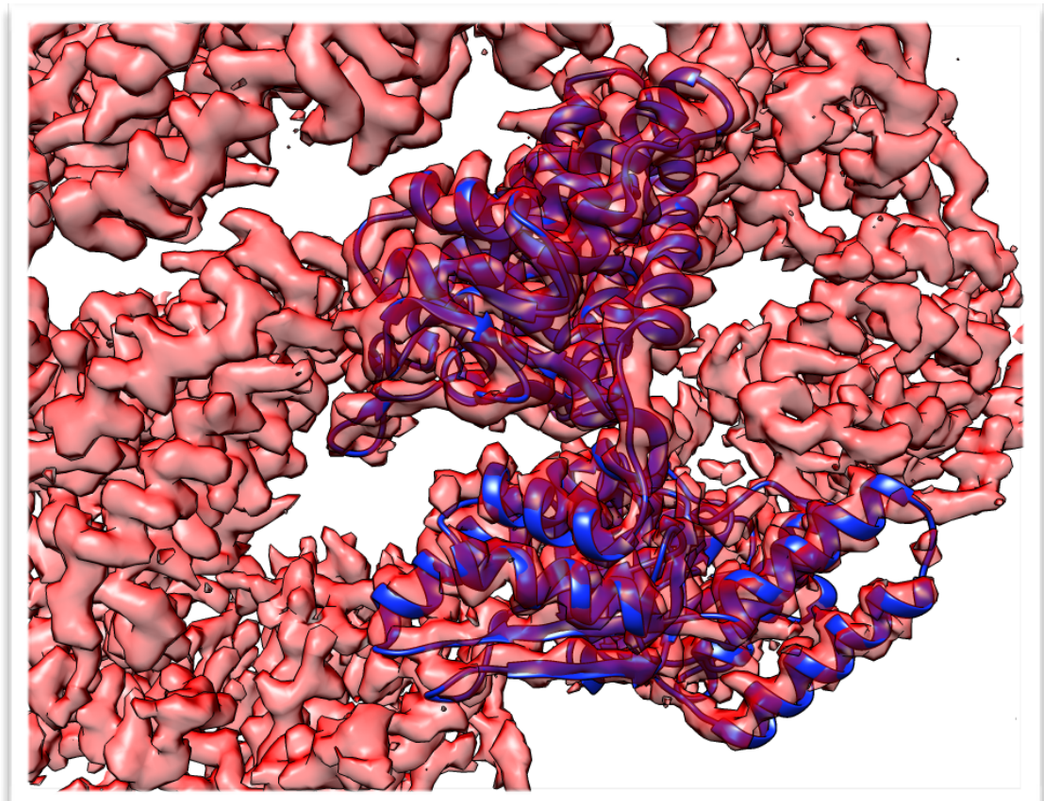
- Systematic cross correlation search of rotations and translations
- Performed in reciprocal space using FFT (very fast)
- Rigid body optimization of position



EMD8750



1SS8 chain A



BERKELEY LAB
Lawrence Berkeley National Laboratory

Automated Model Building

Tom Terwilliger
Los Alamos National Laboratory
Pavel Afonine, Oleg Sobolev
Lawrence Berkeley National Laboratory

BERKELEY LAB
Lawrence Berkeley National Laboratory

Phenix

Automated Model Building Procedure

Determine optimal sharpening of the map

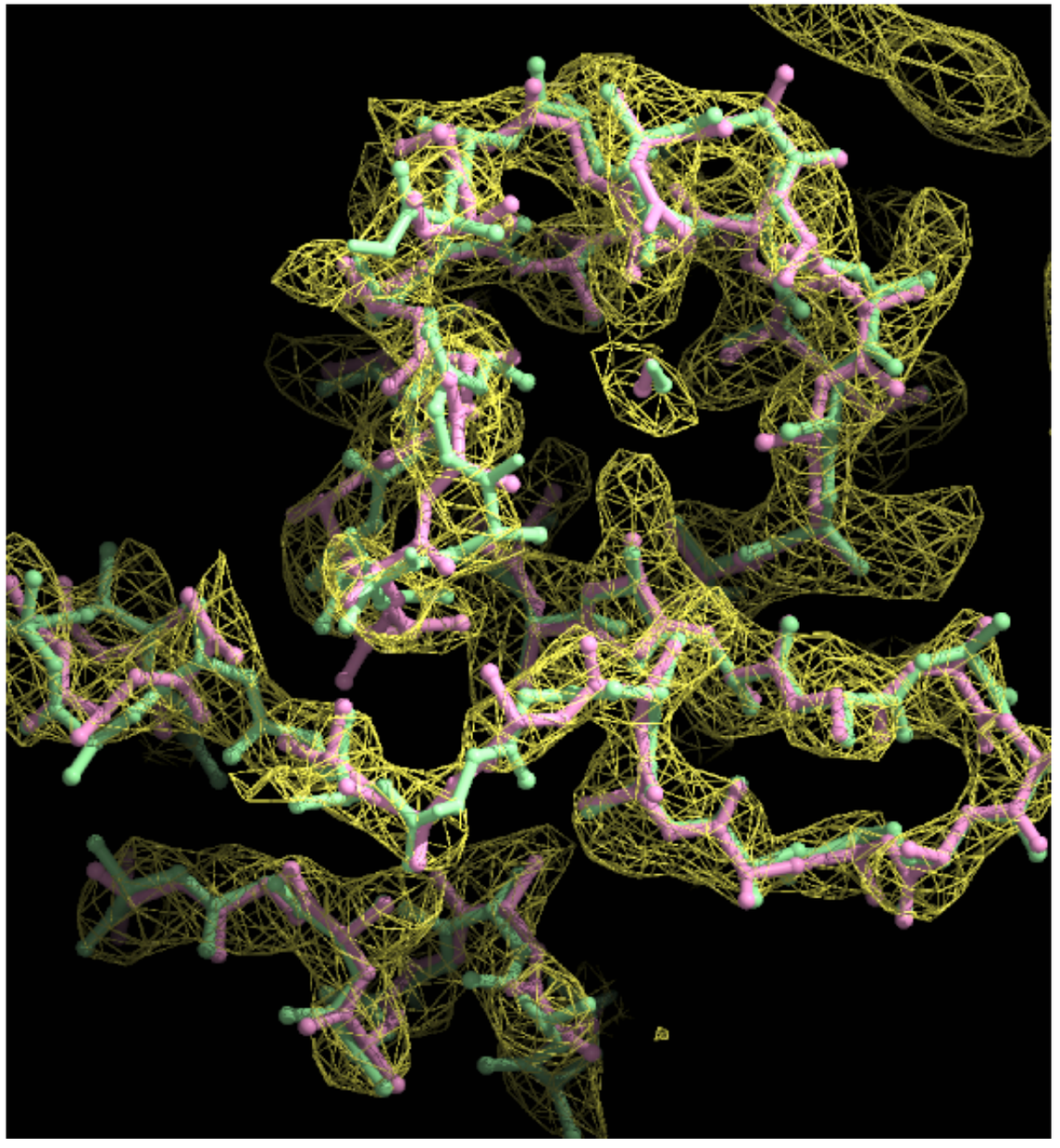
Cut out asymmetric unit of the map

Trace chain and build model

Idealize secondary structure and refine

Assemble and refine (protein/RNA/DNA)

Apply molecular symmetry and re-refine



Cryo-EM map from the yeast mitochondrial ribosome (chain I of large subunit, 3.2Å, Amunts *et al.*, 2014)

Autobuilt model (pink)
Deposited model (green)

Terwilliger et al. A fully automatic method yielding initial models from high-resolution electron cryo-microscopy maps. *Nature Methods*, in press

BERKELEY LAB
Lawrence Berkeley National Laboratory

Phenix

THE UNIVERSITY OF CALIFORNIA BERKELEY 1868

Automated Map Sharpening

Create series of maps with variable overall B-values

Analyze maps for detail and connectivity

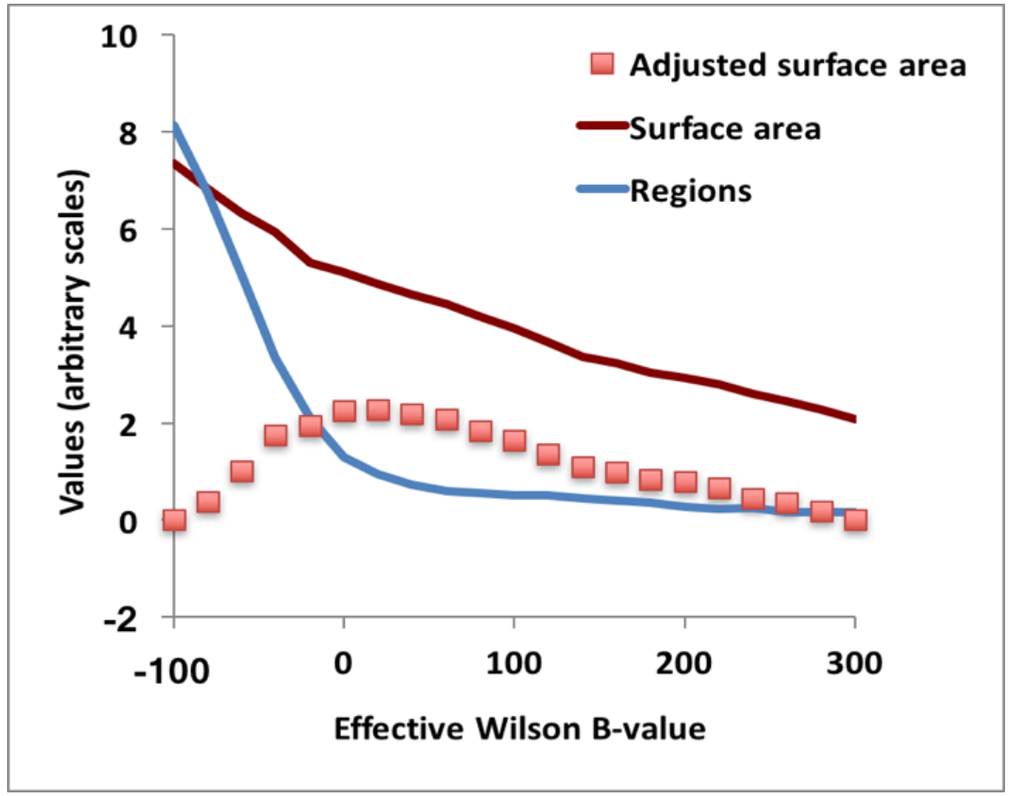
Set contour level enclosing 20% of molecular volume

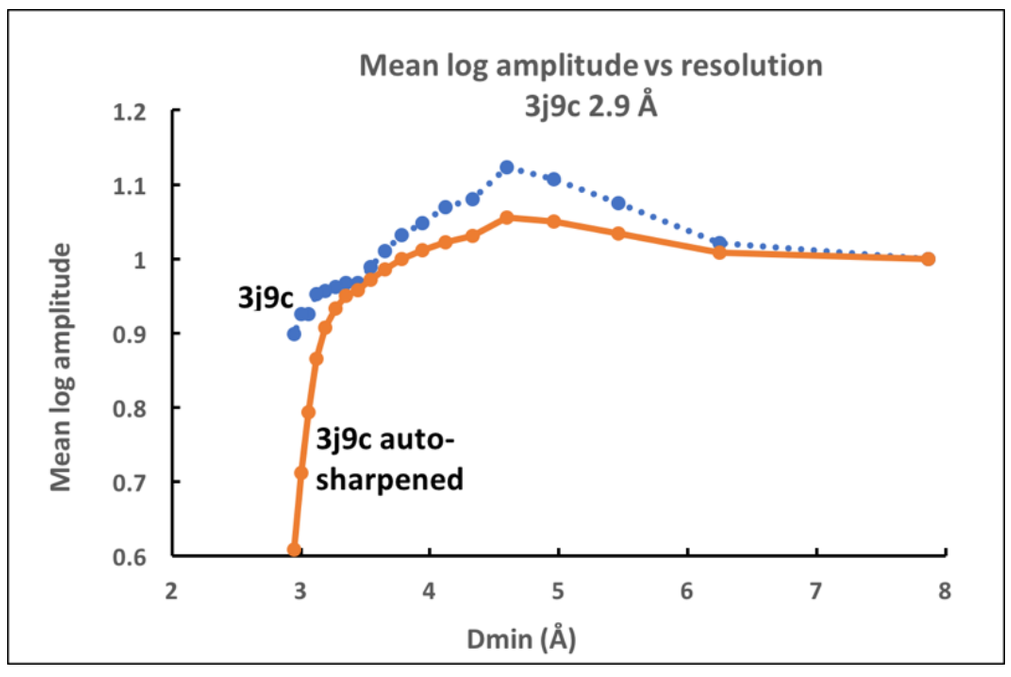
Calculate surface area of contours

Count number of distinct regions enclosed by contours

Choose map with maximum of adjusted surface area

adjusted area = surface area – weight * number of regions



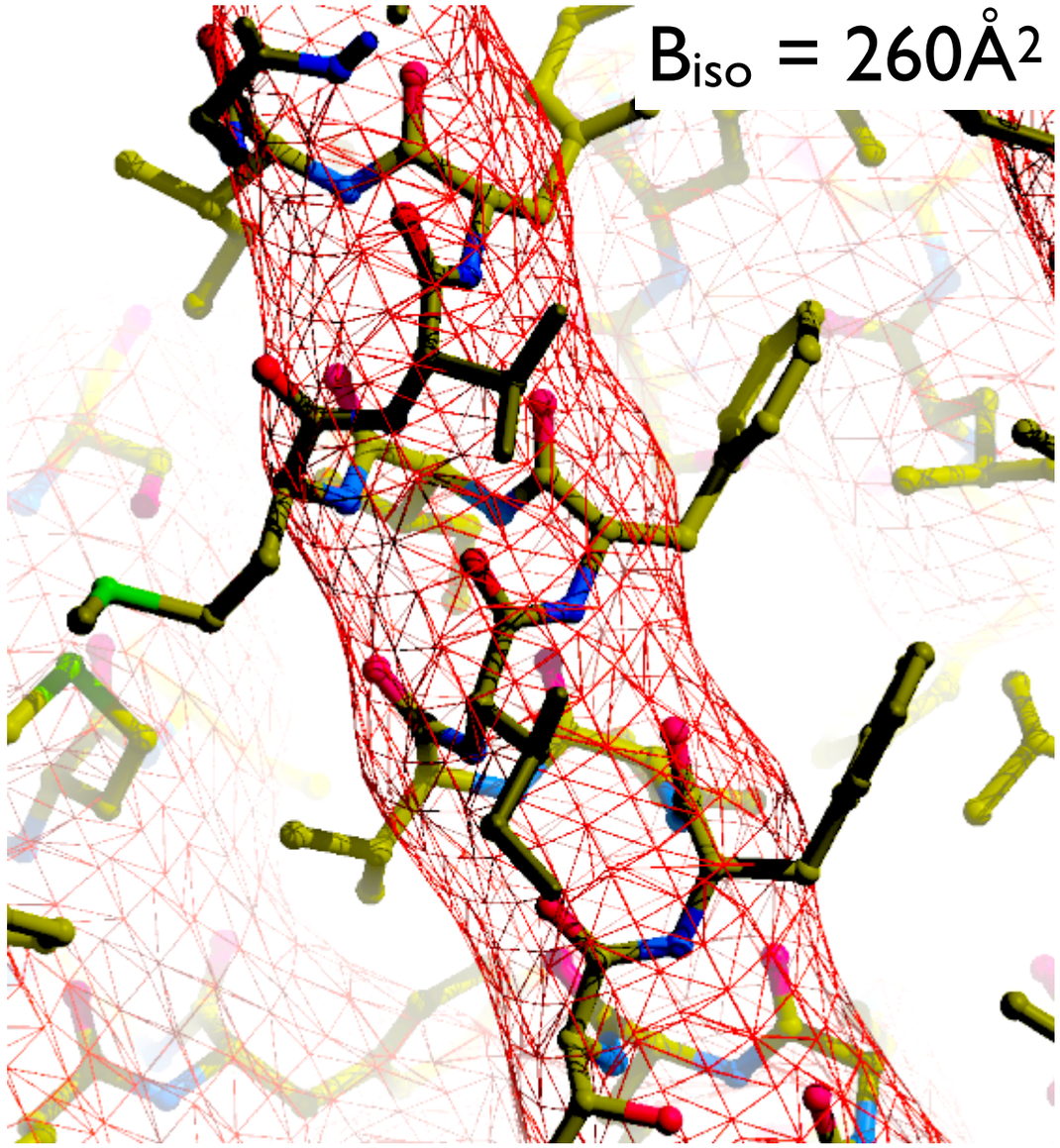


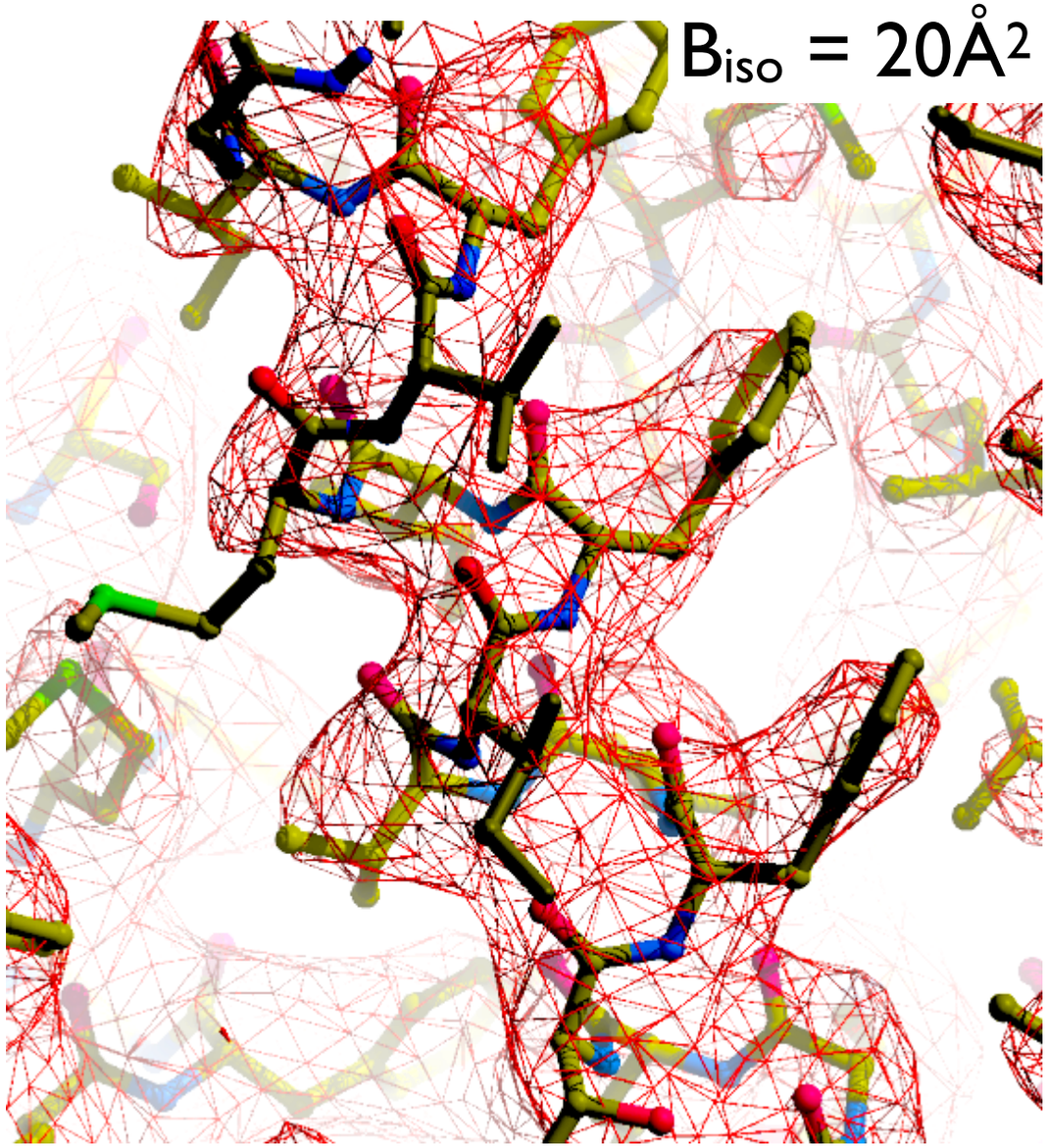
BERKELEY LAB
Lawrence Berkeley National Laboratory

Phenix

phenix.auto_sharpen

Automated Map Sharpening





Deposited Map *Autosharpened Map*

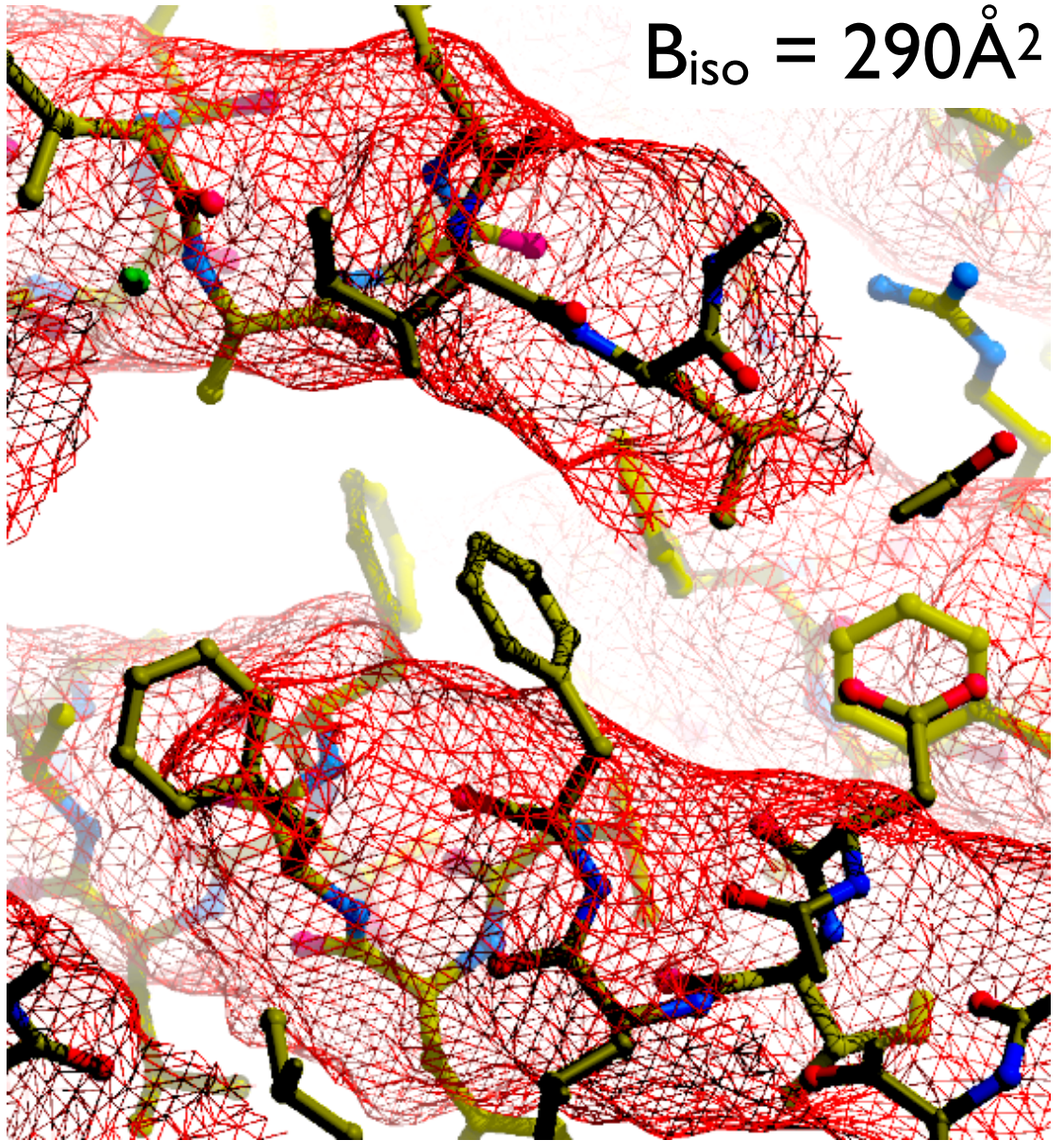
High-conductance Ca(2+)-activated K(+) channel (emd_8414 and PDB entry 5tji; Hite et al., 2017)

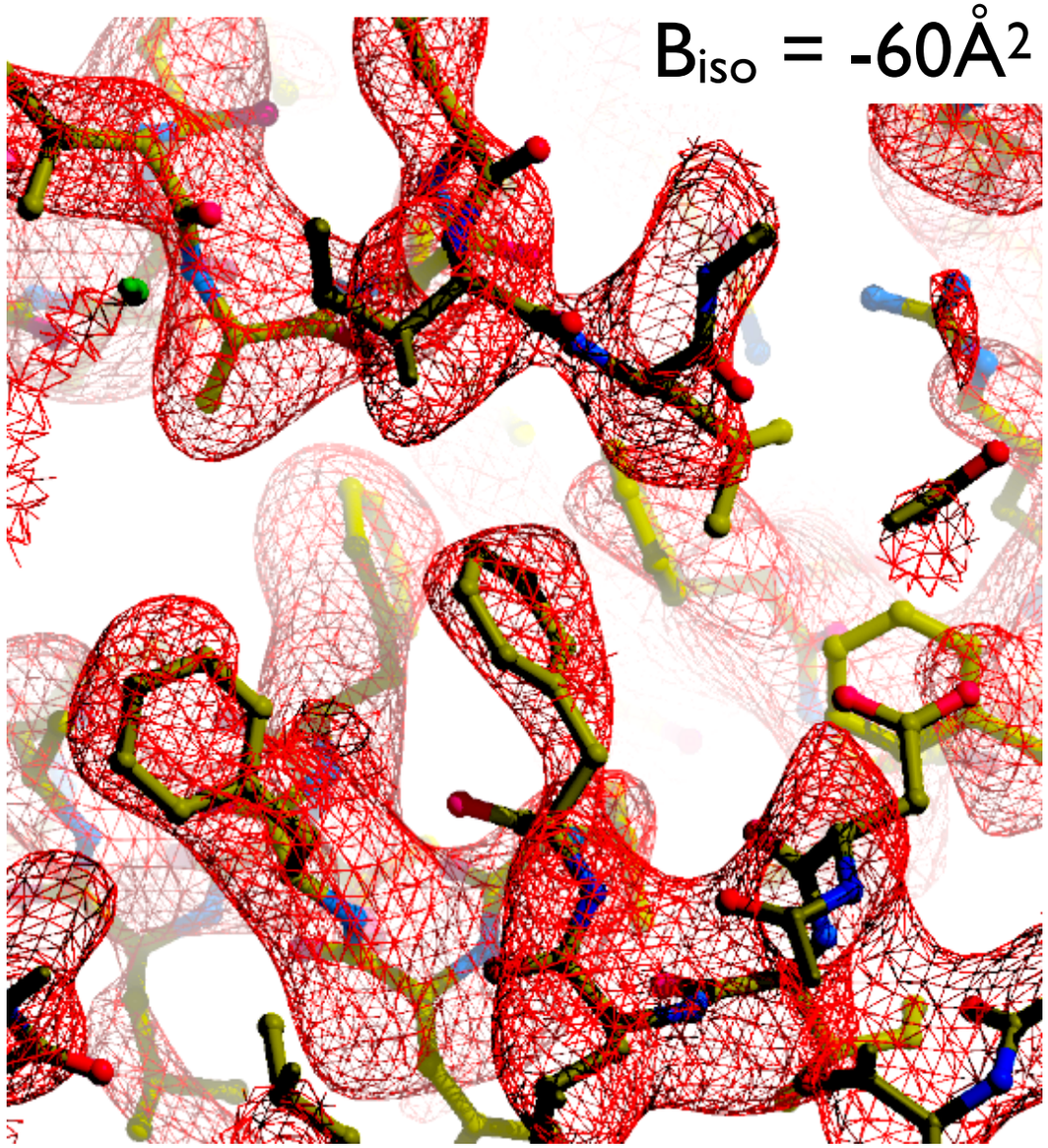
BERKELEY LAB
Lawrence Berkeley National Laboratory

Phenix

THE UNIVERSITY OF CALIFORNIA · BERKELEY · 1868

Automated Map Sharpening





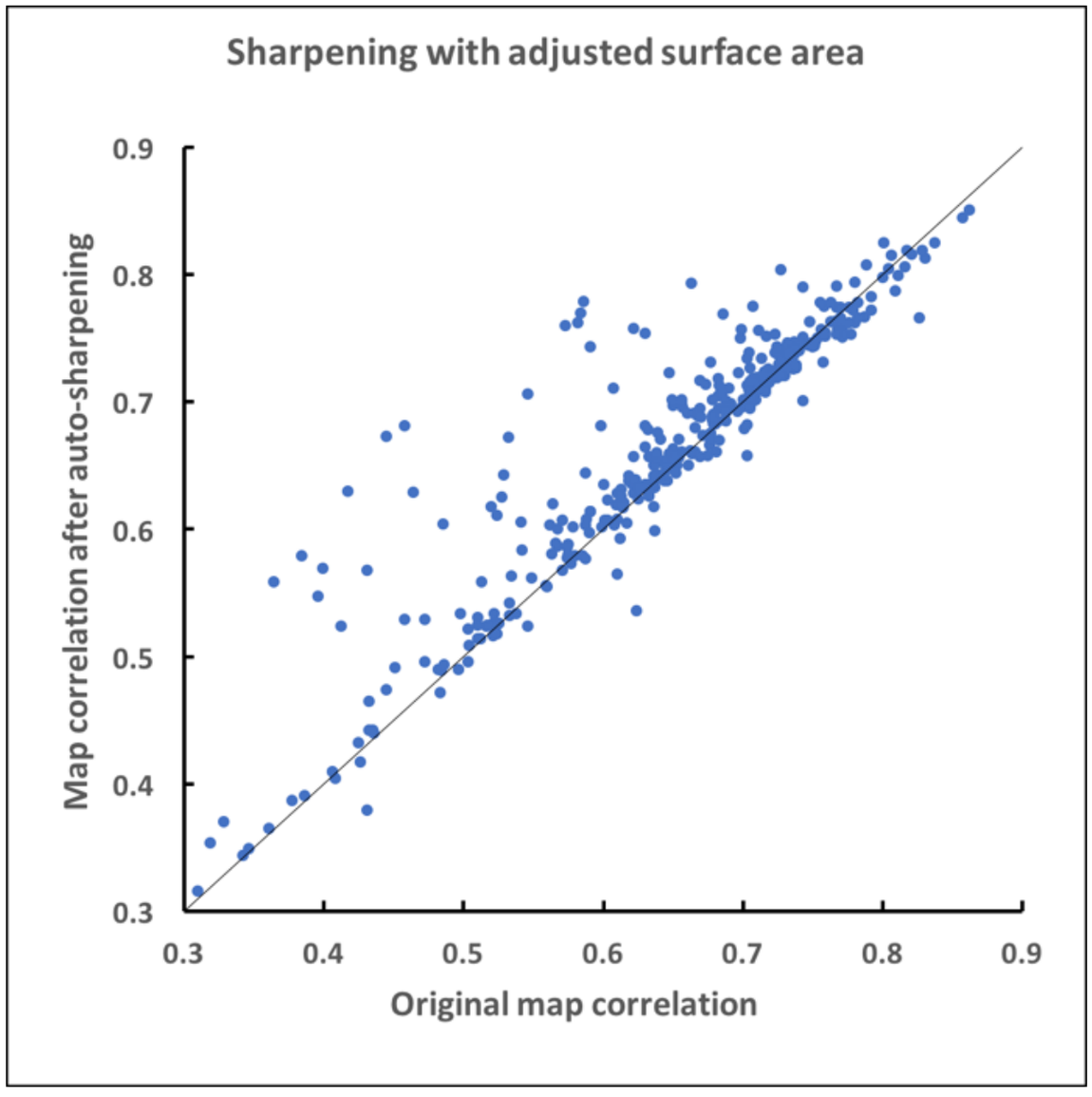
Deposited Map *Autosharpened Map*

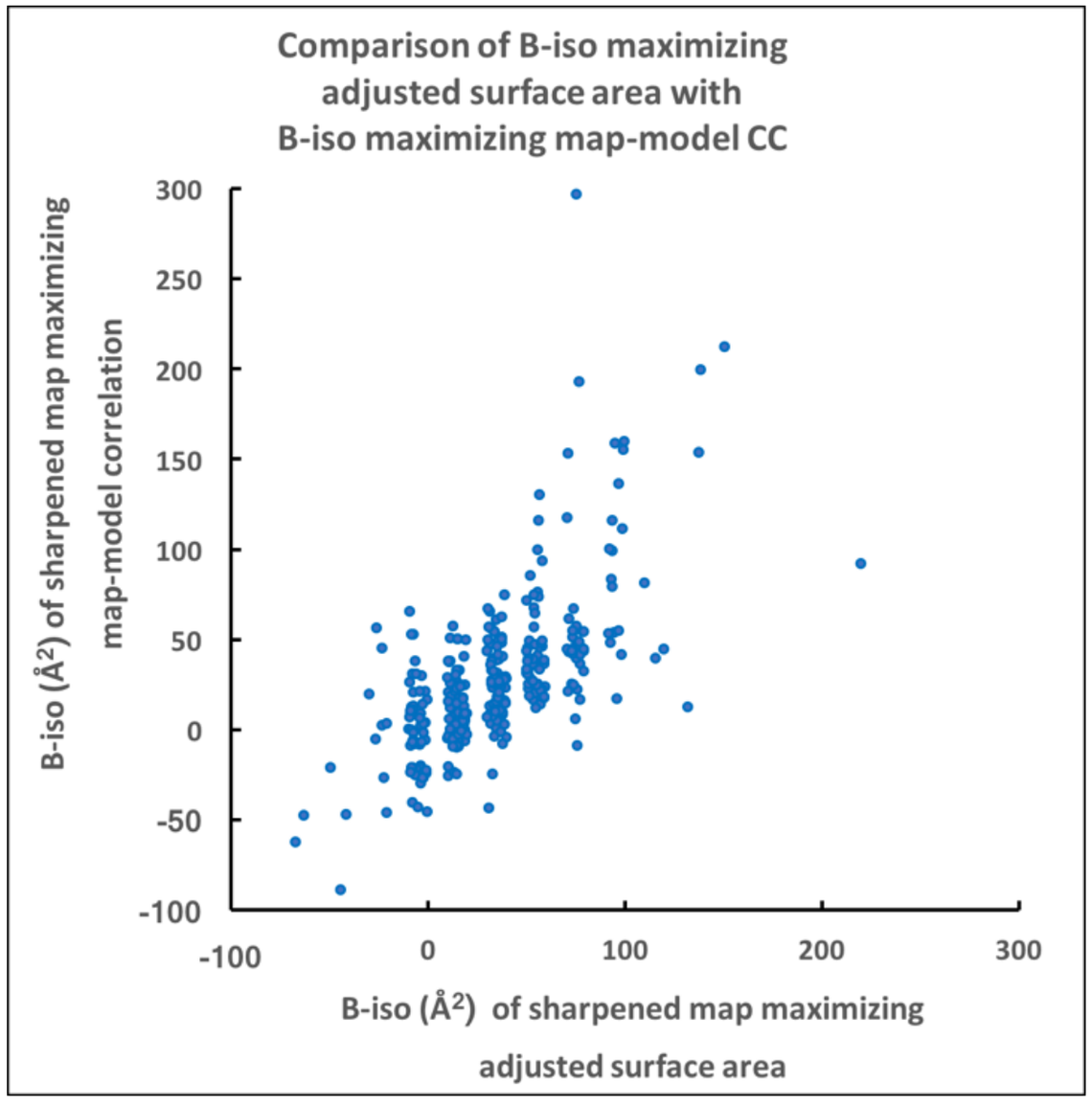
Cystic fibrosis transmembrane conductance regulator
(emd_8461 and PDB entry 5uar; Zhang and Chen, 2016)

BERKELEY LAB
Lawrence Berkeley National Laboratory

Phenix

Automated Map Sharpening





Terwilliger et al. Automated map sharpening by maximization of detail and connectivity. *Acta Cryst* 2018, **D74**:545-559

BERKELEY LAB
Lawrence Berkeley National Laboratory

Phenix

THE UNIVERSITY OF CALIFORNIA · BERKELEY · 1868

Automated Segmentation

Determine optimal sharpening of the map

Cut out asymmetric unit of the map

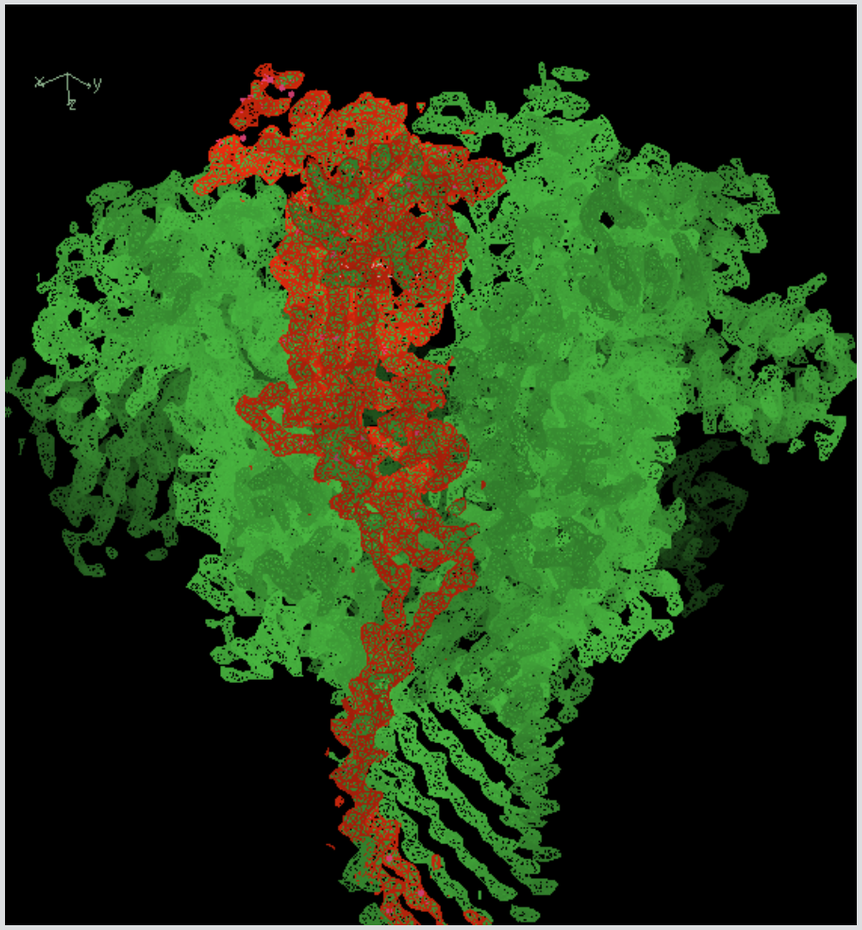
Trace chain and build model

Idealize secondary structure and refine

Assemble and refine (protein/RNA/DNA)

Apply molecular symmetry and re-refine

- Use the symmetry of the map
- Identify contiguous regions representing asymmetric unit of the map
- Choose symmetry-copies that make compact molecule



emd_6224 (anthrax toxin protective antigen pore at 2.9 Å; Jiang *et al.* 2015)

Terwilliger et al. Map segmentation, automated model-building and their application to the Cryo-EM Model Challenge. *J. Struct. Biol.* 2018, in press

BERKELEY LAB
Lawrence Berkeley National Laboratory

Phenix

THE UNIVERSITY OF CALIFORNIA BERKELEY 1868

Chain Tracing

Determine optimal sharpening of the map

Cut out asymmetric unit of the map

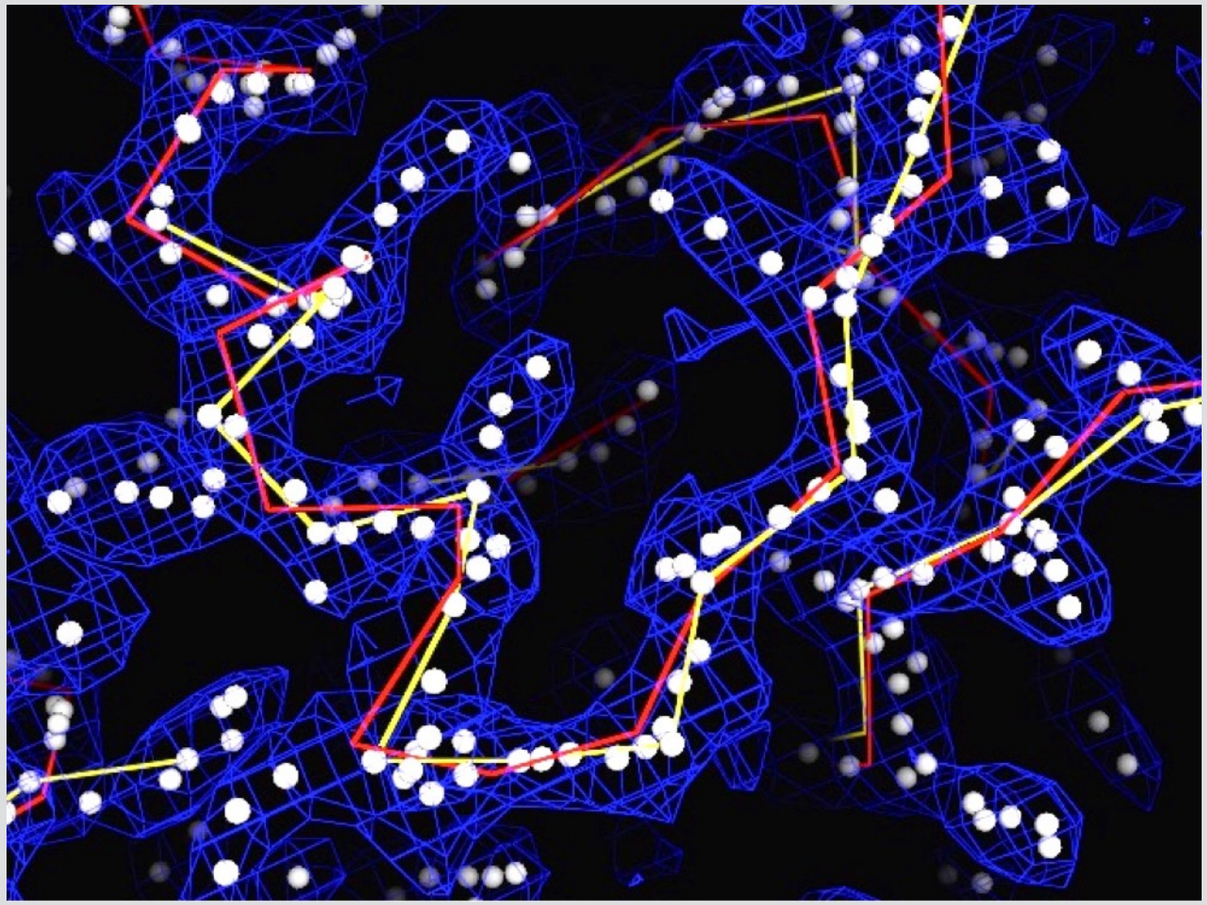
Trace chain and build model

Idealize secondary structure and refine

Assemble and refine (protein/RNA/DNA)

Apply molecular symmetry and re-refine

- Variable map thresholding
- Trace protein main chain
- Identify direction of main chain by fit to density



BERKELEY LAB
Lawrence Berkeley National Laboratory

Phenix

THE UNIVERSITY OF CALIFORNIA BERKELEY 1868

Idealization and Refinement

Determine optimal sharpening of the map

Cut out asymmetric unit of the map

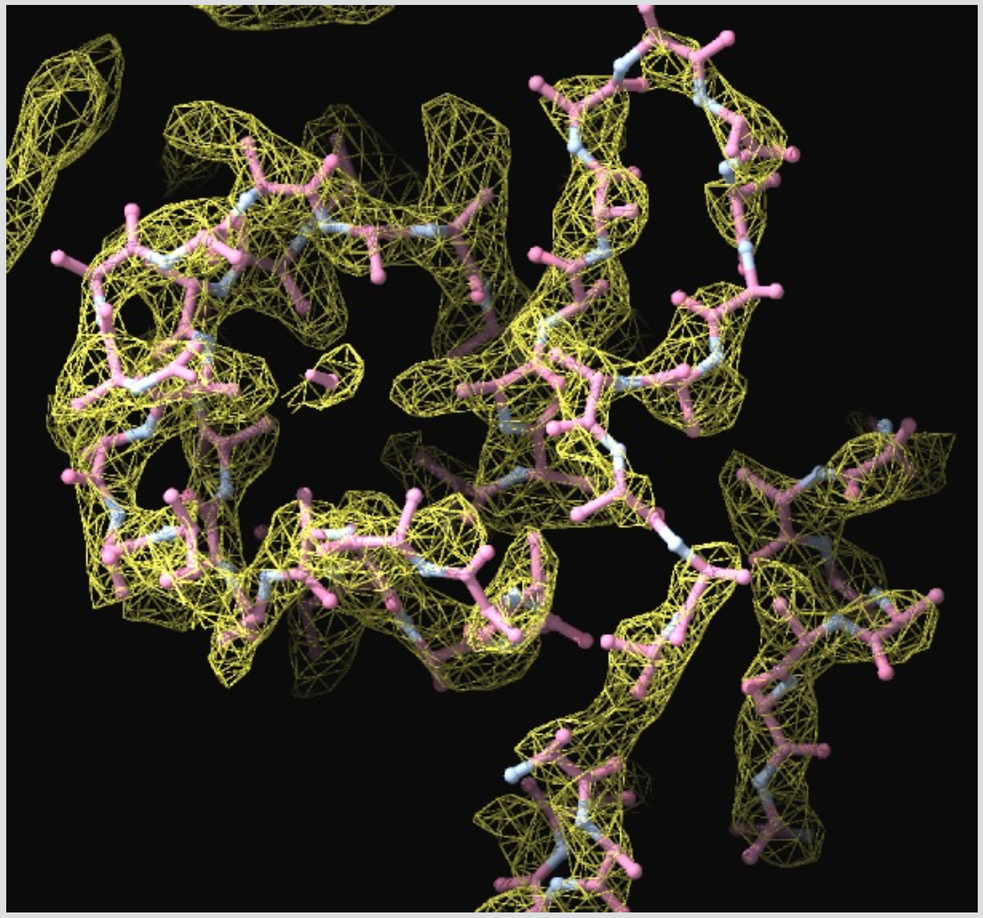
Trace chain and build model

Idealize secondary structure and refine

Assemble and refine (protein/RNA/DNA)

Apply molecular symmetry and re-refine

- Refine and rebuild model (simulated annealing, rebuilding and combination of best parts of each model)
- Replace segments with idealized structure
- Identify hydrogen-bonding (β-sheets, α-helices) and use them as restraints in real-space refinement



Chain I, yeast mitochondrial ribosome large subunit, 3.2 Å, 3j6b

BERKELEY LAB
Lawrence Berkeley National Laboratory

Phenix

THE UNIVERSITY OF CALIFORNIA · BERKELEY · 1868

Assembly and Polymer Recognition

Determine optimal sharpening of the map

↓

Cut out asymmetric unit of the map

↓

Trace chain and build model

↓

Idealize secondary structure and refine

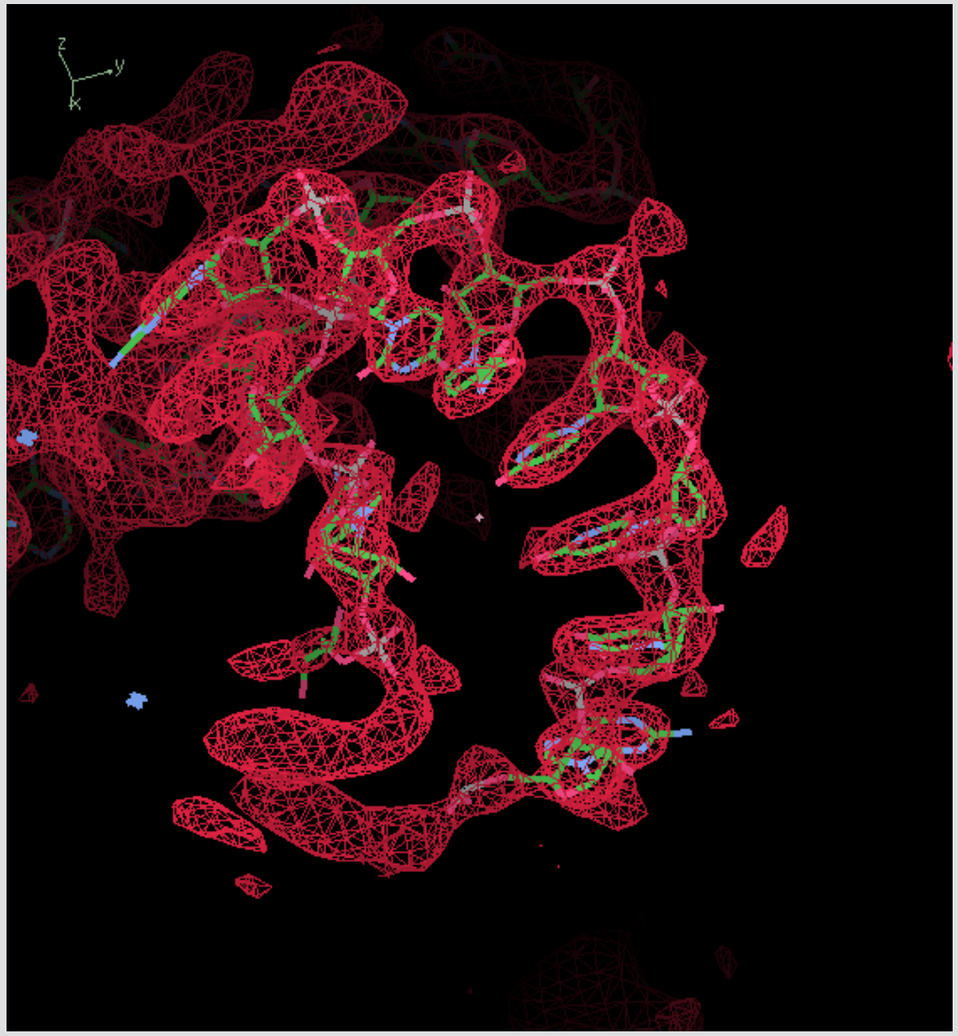
↓

Assemble and refine (protein/RNA/DNA)

↓

Apply molecular symmetry and re-refine

- Try building protein/RNA/DNA (whatever may be there)
- Choose segment type by map correlation



70S ribosome at 2.9 Å

BERKELEY LAB
Lawrence Berkeley National Laboratory

Phenix

THE UNIVERSITY OF CALIFORNIA BERKELEY 1868

The Final Model

Determine optimal sharpening of the map

Cut out asymmetric unit of the map

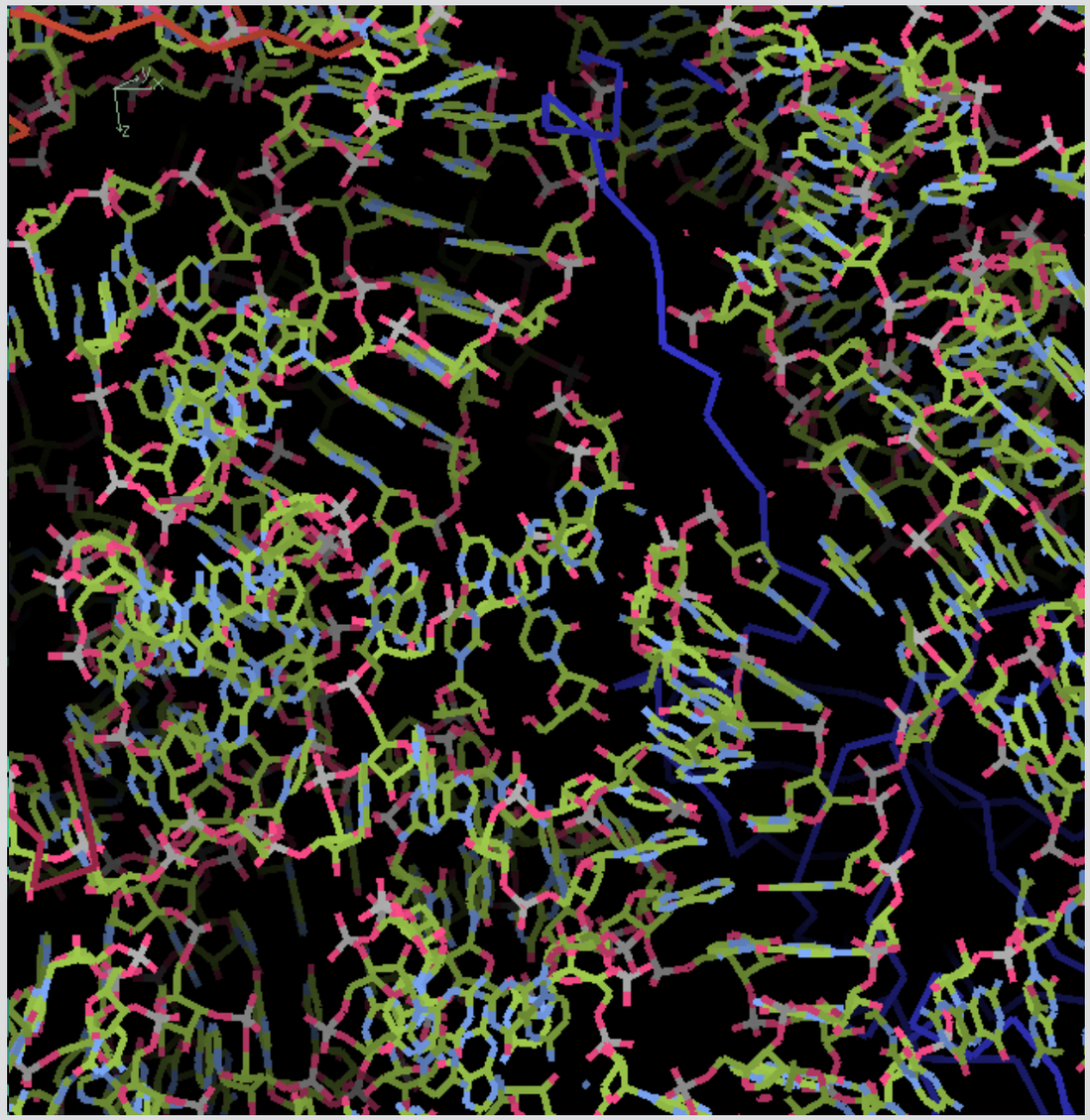
Trace chain and build model

Idealize secondary structure and refine

Assemble and refine (protein/RNA/DNA)

Apply molecular symmetry and re-refine

- `phenix.map_to_model`



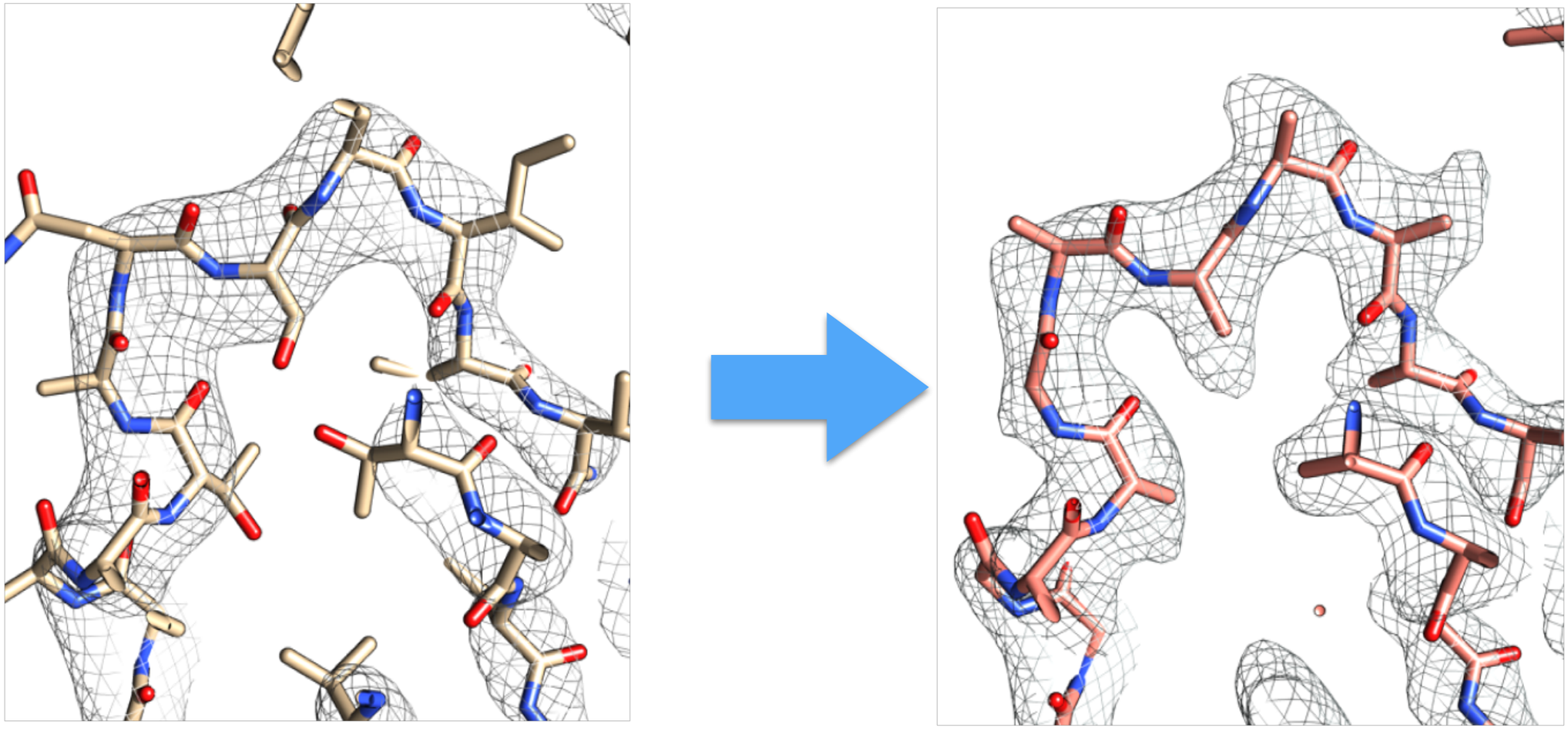
30S Ribosome (1j5e, 2.9 Å)

BERKELEY LAB
Lawrence Berkeley National Laboratory

Phenix

THE UNIVERSITY OF CALIFORNIA BERKELEY 1868

Automated Building - Sharpening



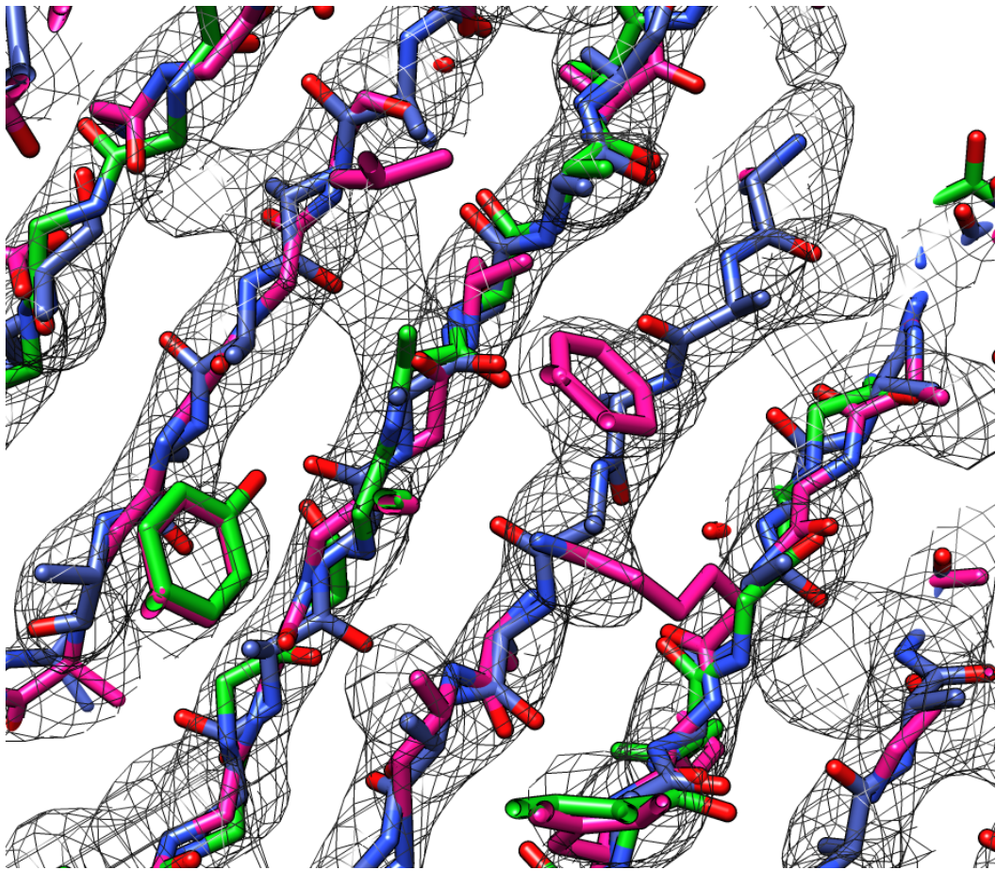
Original

Automatically Sharpened

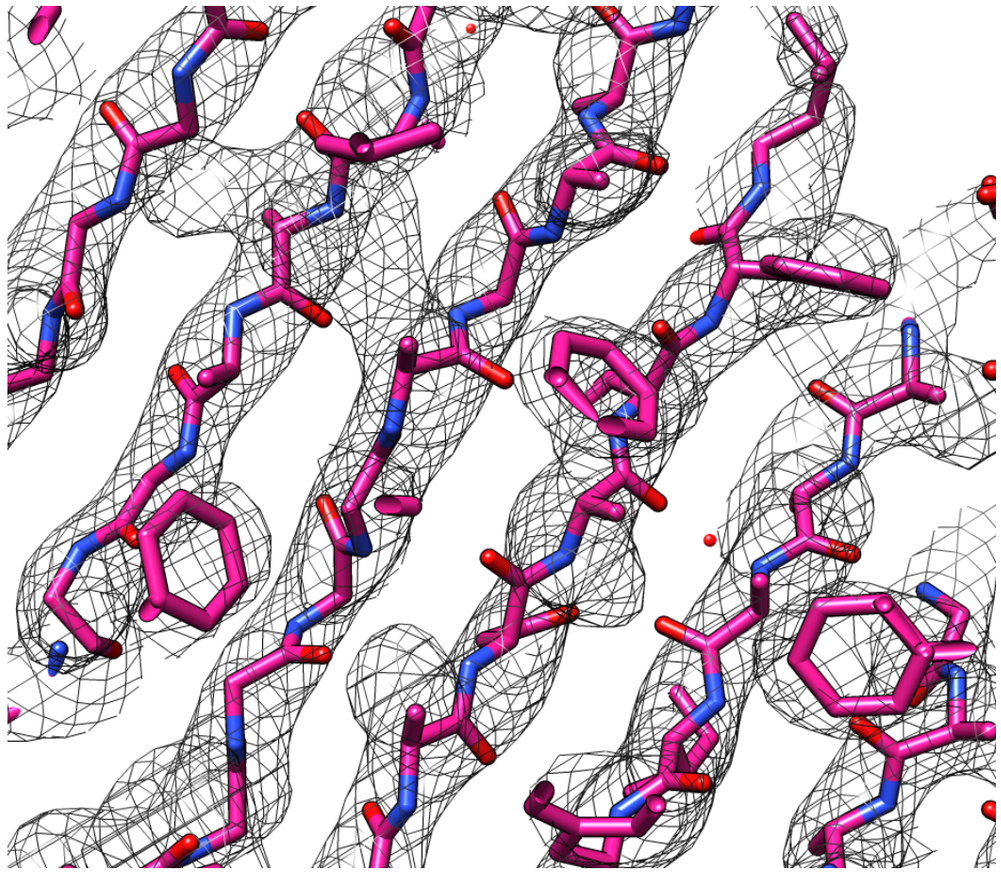
BERKELEY LAB
Lawrence Berkeley National Laboratory

Phenix

Automated Building - Combining Multiple Models



Three Independently Built Models



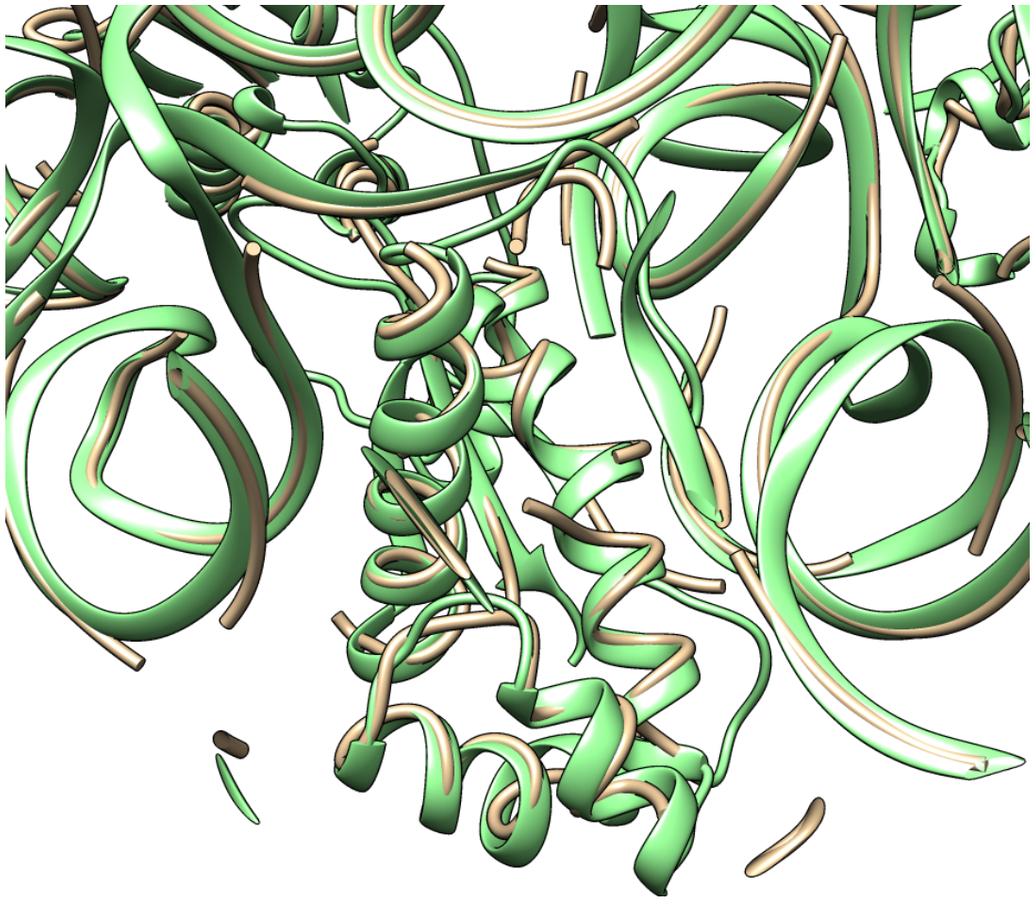
Composite Model

BERKELEY LAB
Lawrence Berkeley National Laboratory

Phenix

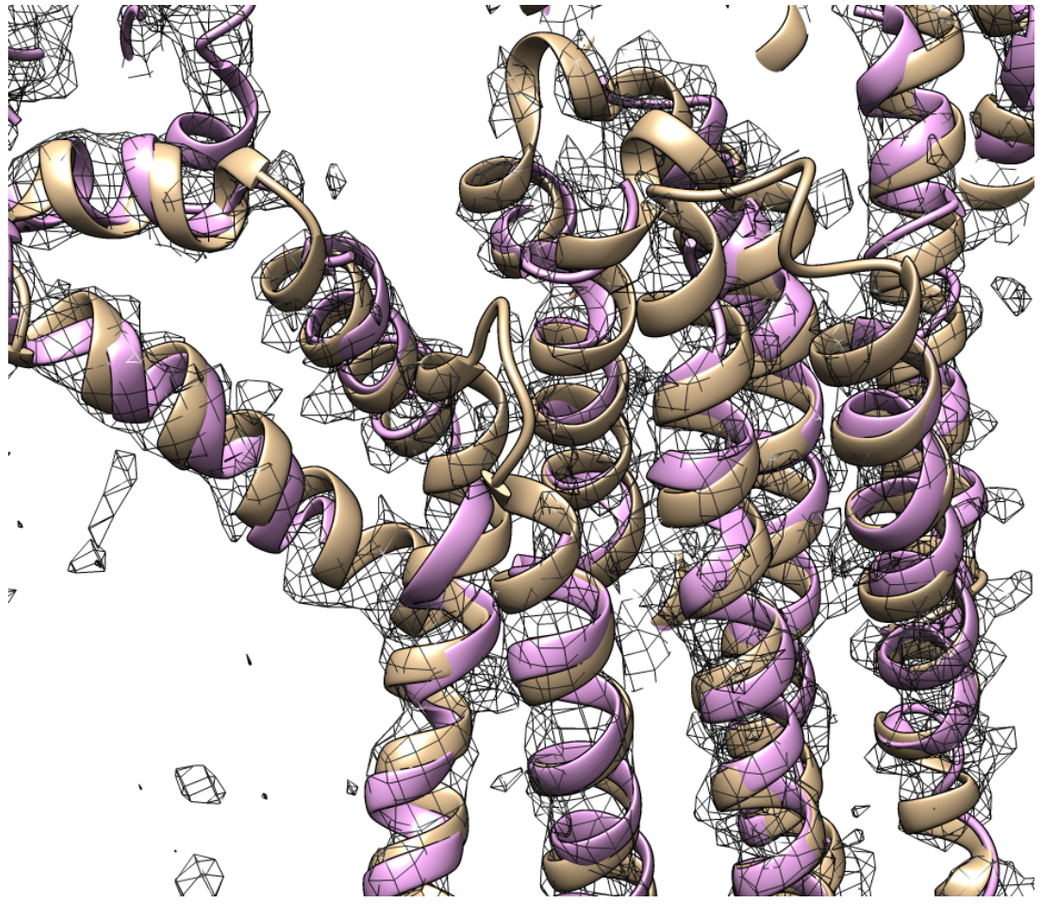
THE UNIVERSITY OF CALIFORNIA BERKELEY 1868

Automated Building - Assembling a Final Model



M. smegmatis ribosome

green: deposited
brown: auto built



ERAD Channel

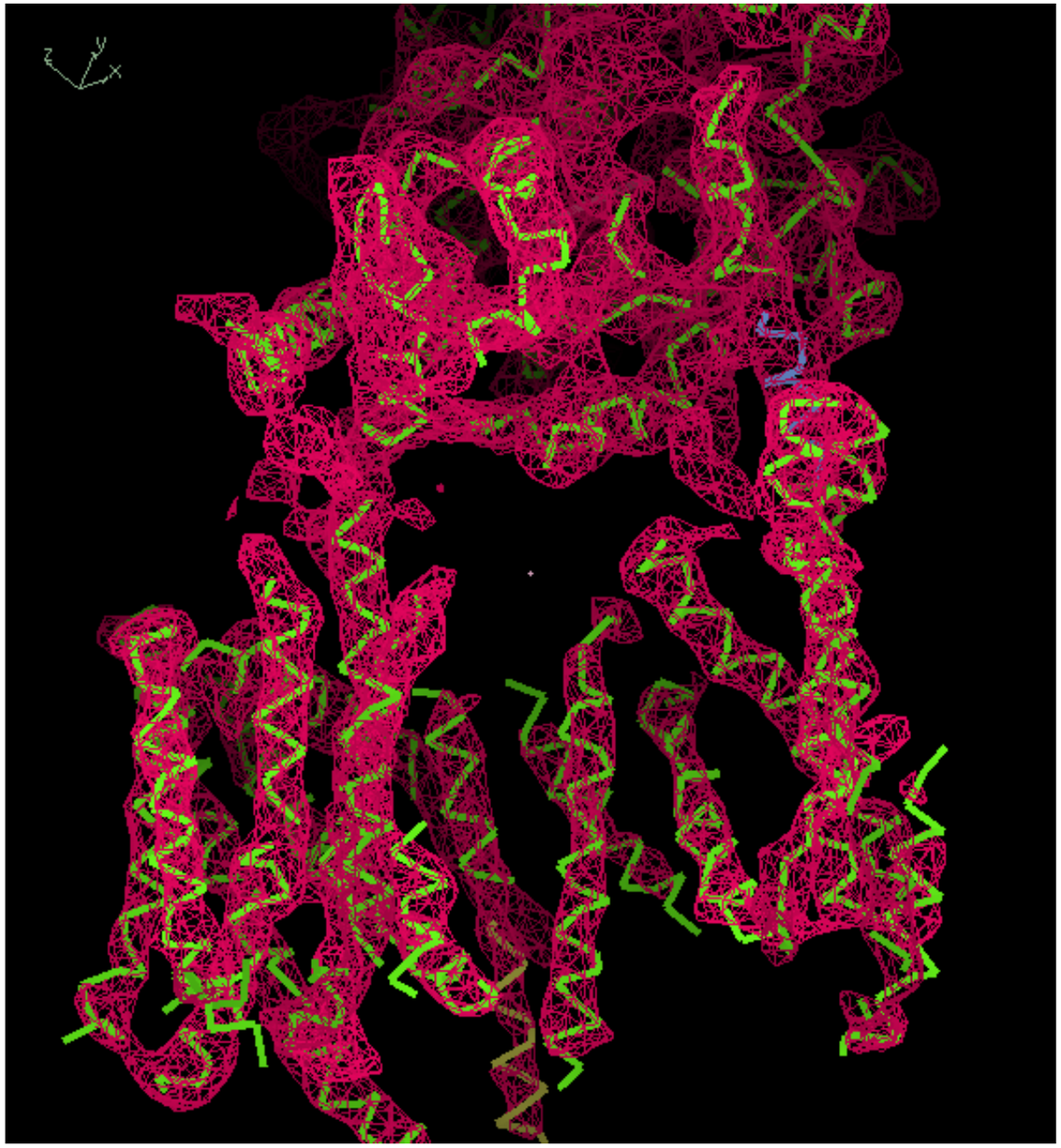
brown: deposited
purple: auto built

BERKELEY LAB
Lawrence Berkeley National Laboratory

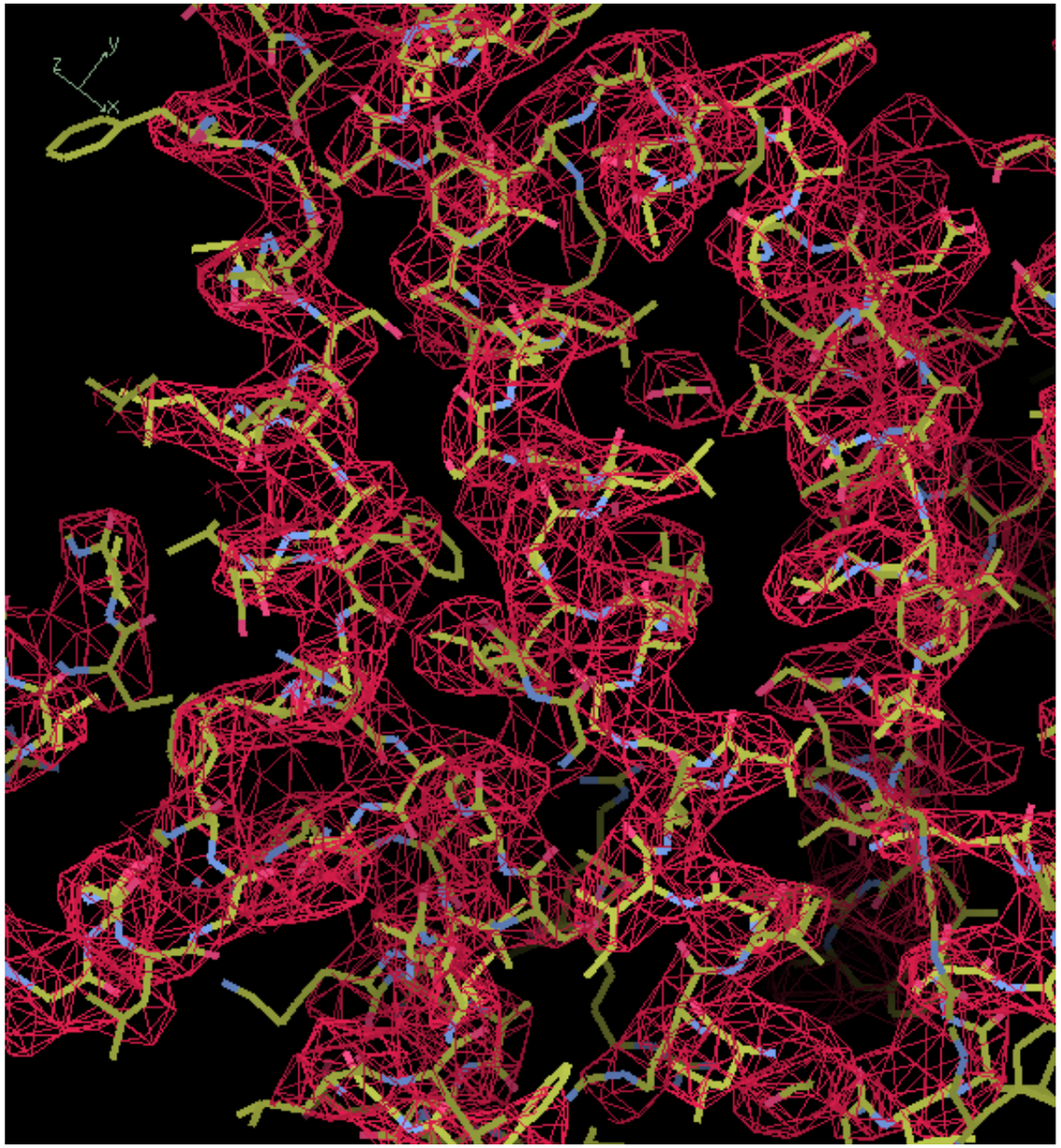
Phenix

THE UNIVERSITY OF CALIFORNIA BERKELEY 1868

Building at Low Resolution



Gamma-secretase at 4.5 Å
(autobuilt model; emd_2677)



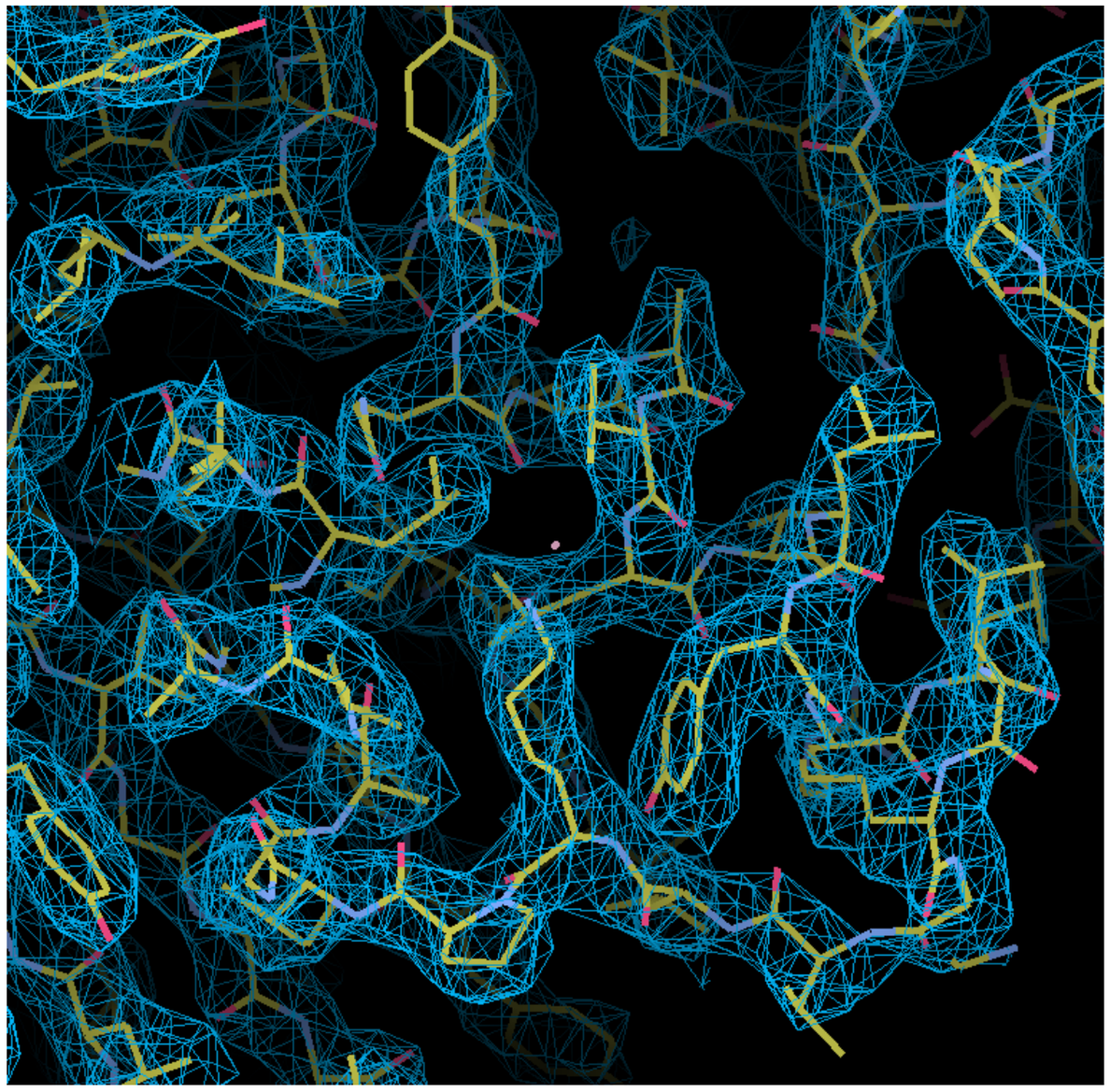
Gamma-secretase structure at 3.4 Å
(autobuilt model; emd_3061)

BERKELEY LAB
Lawrence Berkeley National Laboratory

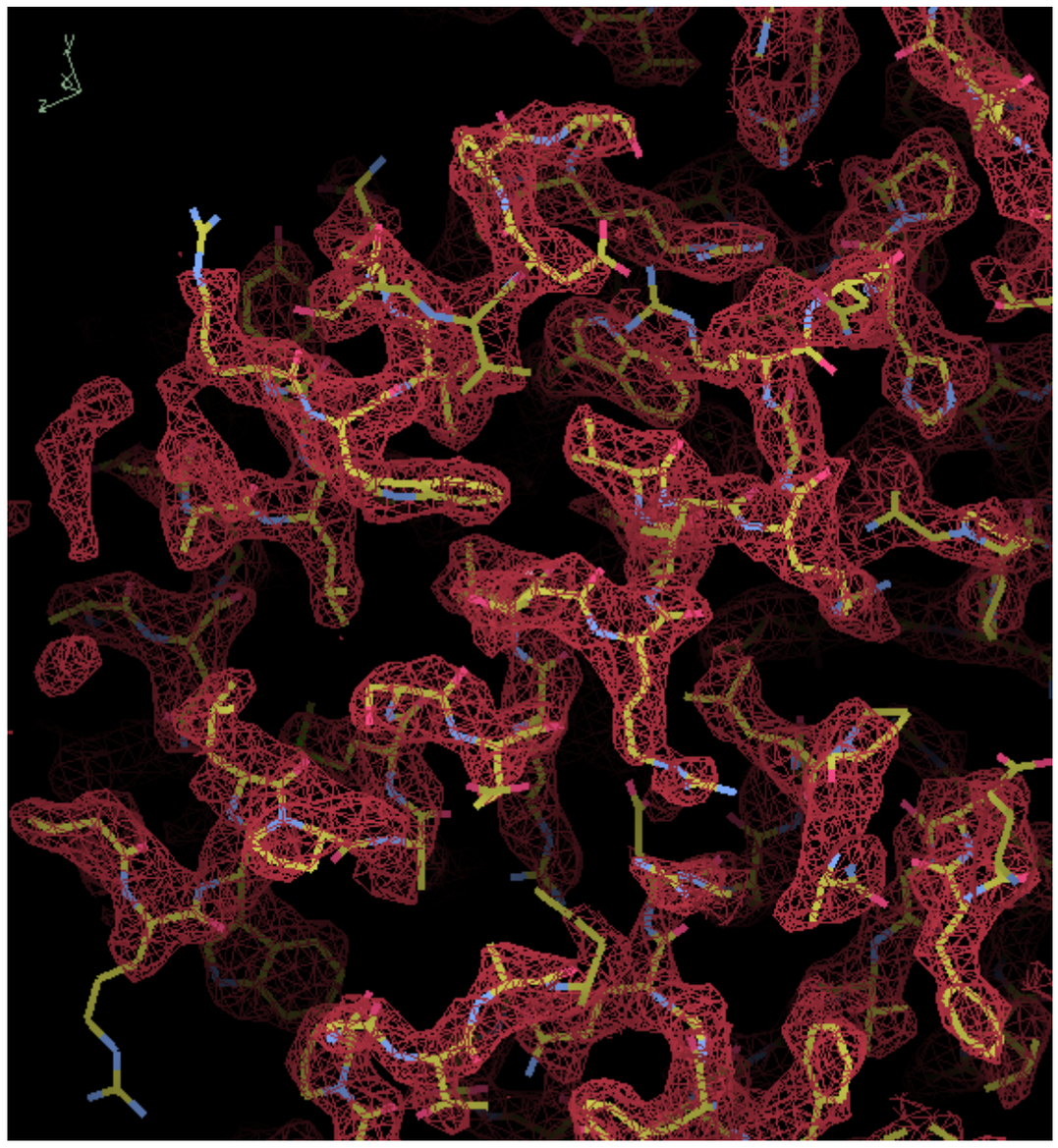
Phenix

THE UNIVERSITY OF CALIFORNIA · BERKELEY · 1868

Building at Medium/High Resolution



Proteasome at 2.8 Å
(autobuilt model; emd_6287)



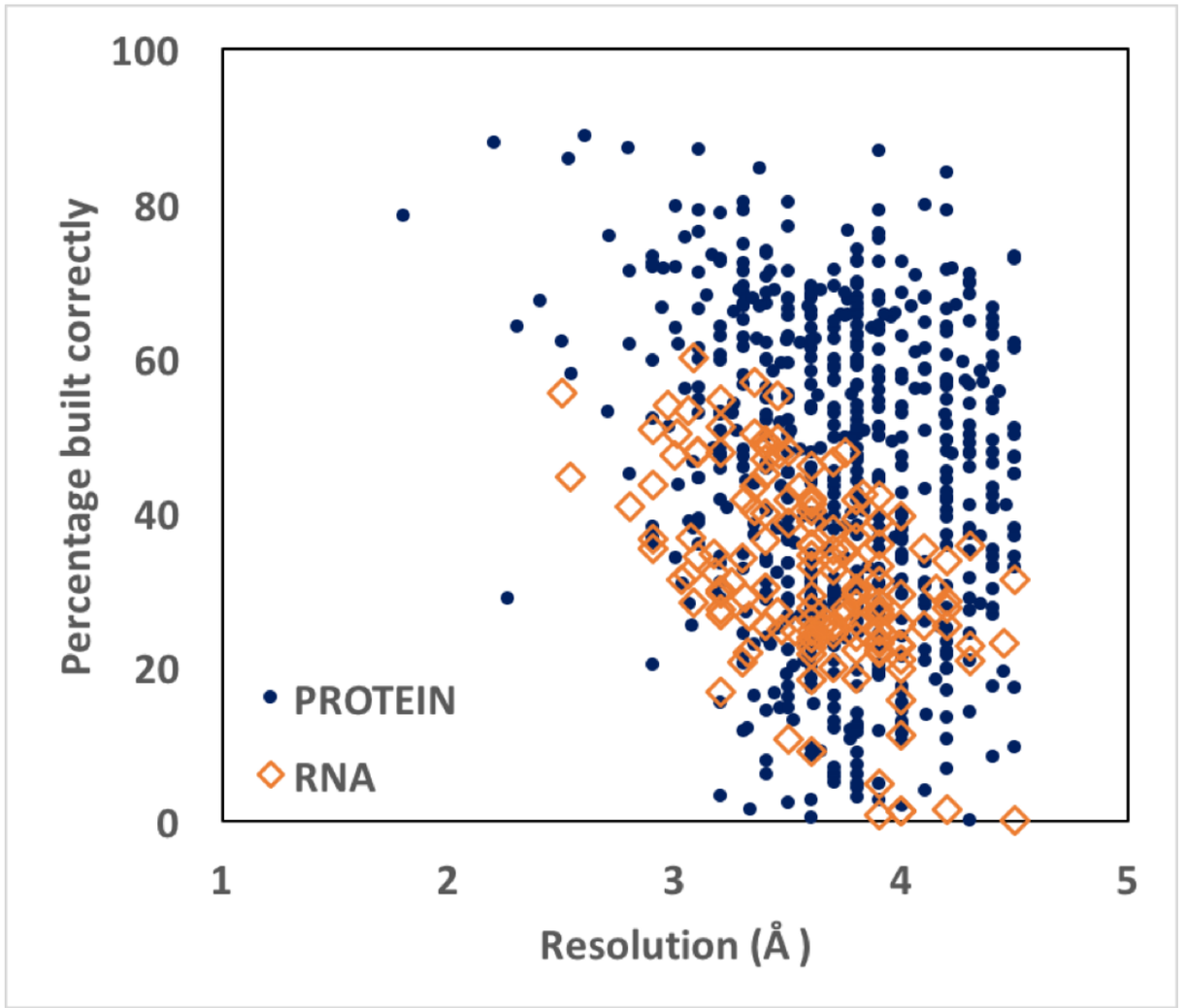
Beta-galactosidase at 2.2 Å
(autobuilt model; emd_2984)

BERKELEY LAB
Lawrence Berkeley National Laboratory

Phenix

THE UNIVERSITY OF CALIFORNIA BERKELEY 1868

Autobuilding Performance



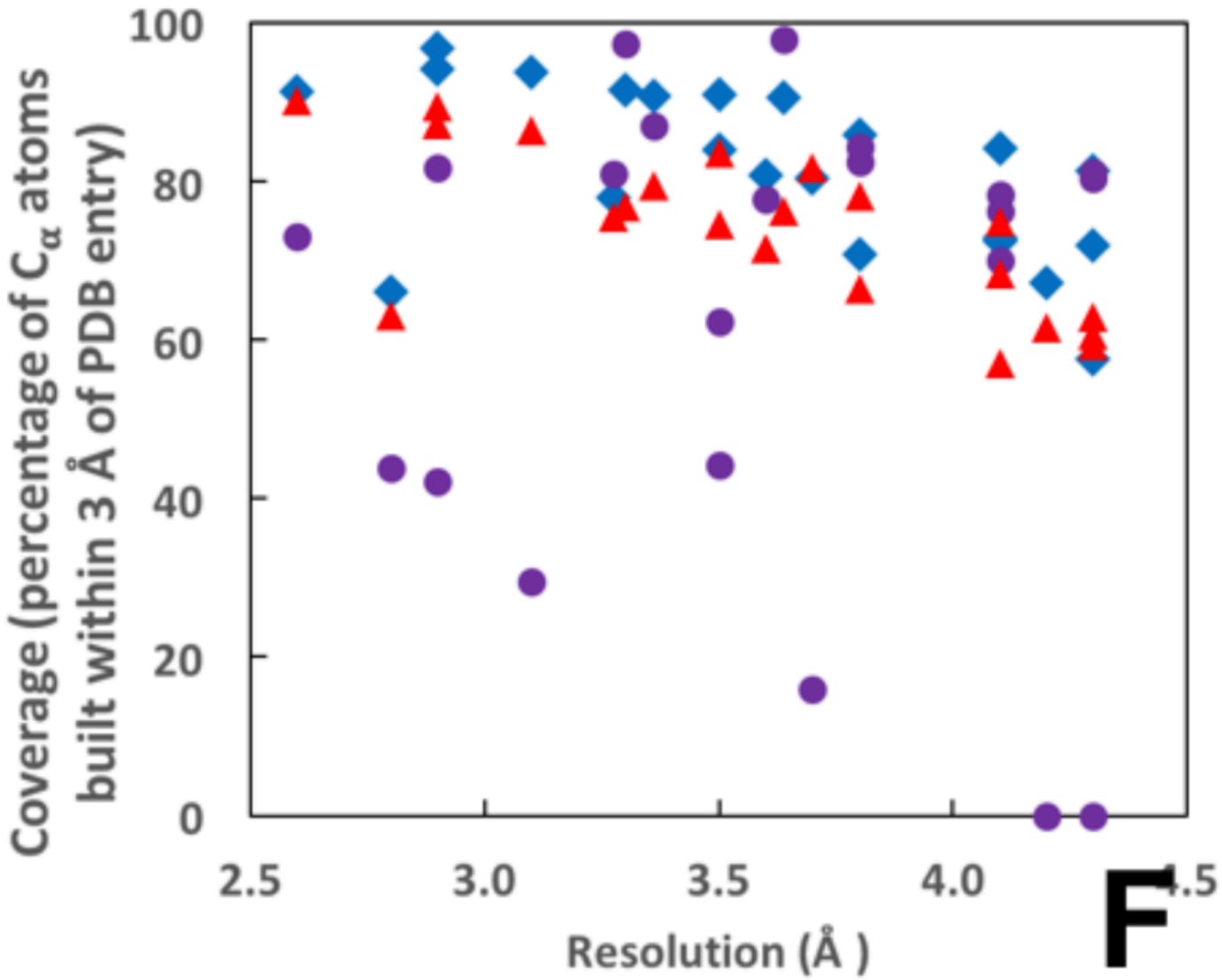
BERKELEY LAB
Lawrence Berkeley National Laboratory

Phenix

THE UNIVERSITY OF CALIFORNIA BERKELEY 1868

Building into Segmented Maps

- Model building is improved when building the correct sequence into a segmented volume



F

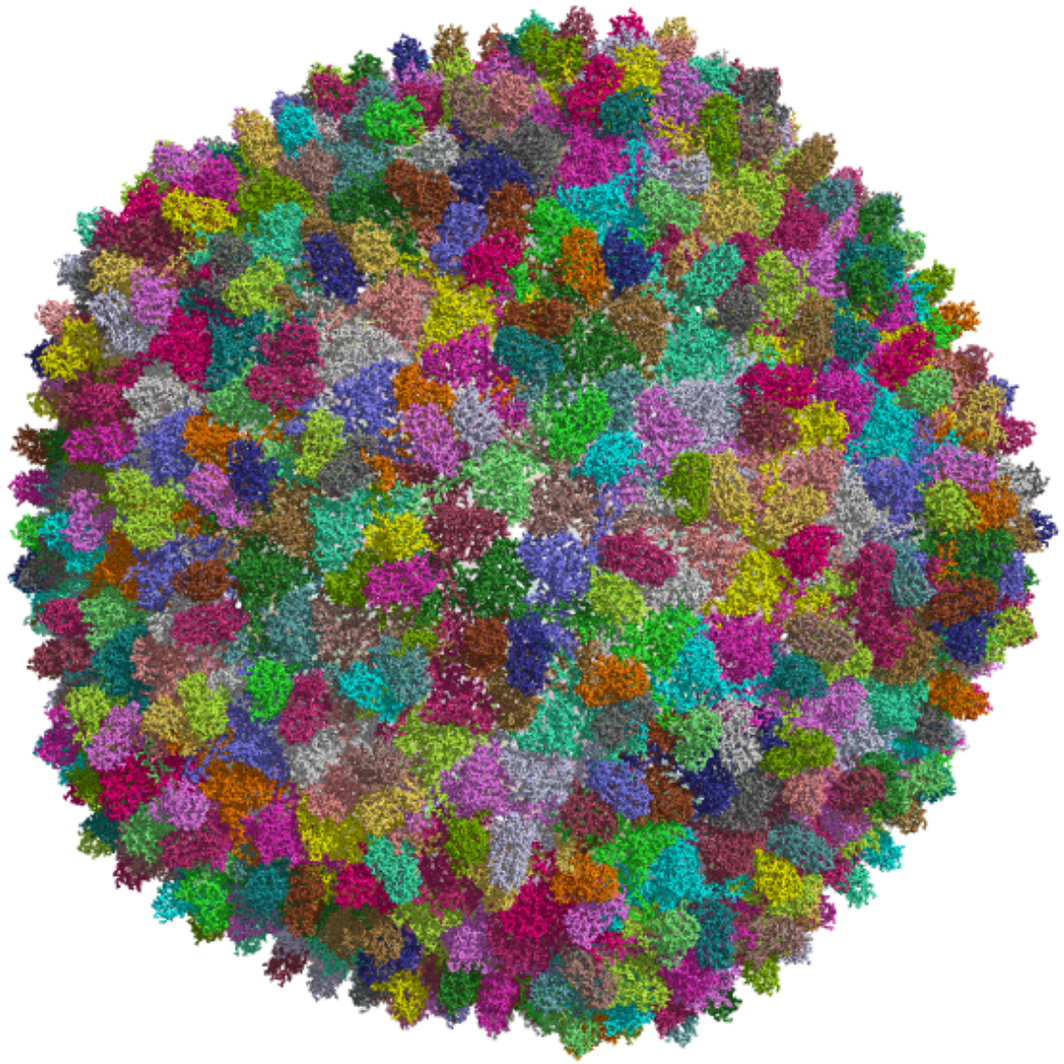
Atomic Model Optimization

Pavel Afonine, Oleg Sobolev, Youval Dar, Nat Echols, Jeff Headd, Nigel Moriarty
Lawrence Berkeley National Laboratory
Tom Terwilliger
Los Alamos National Laboratory

BERKELEY LAB
Lawrence Berkeley National Laboratory

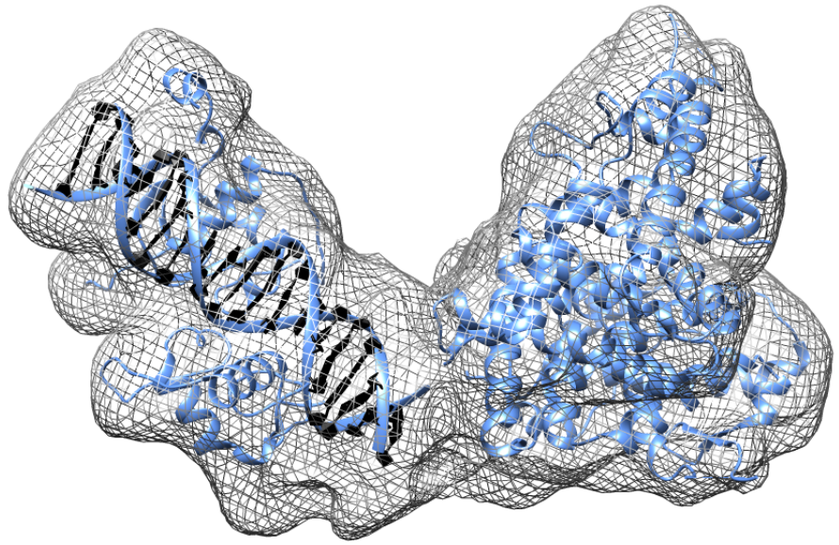
Phenix

Challenges



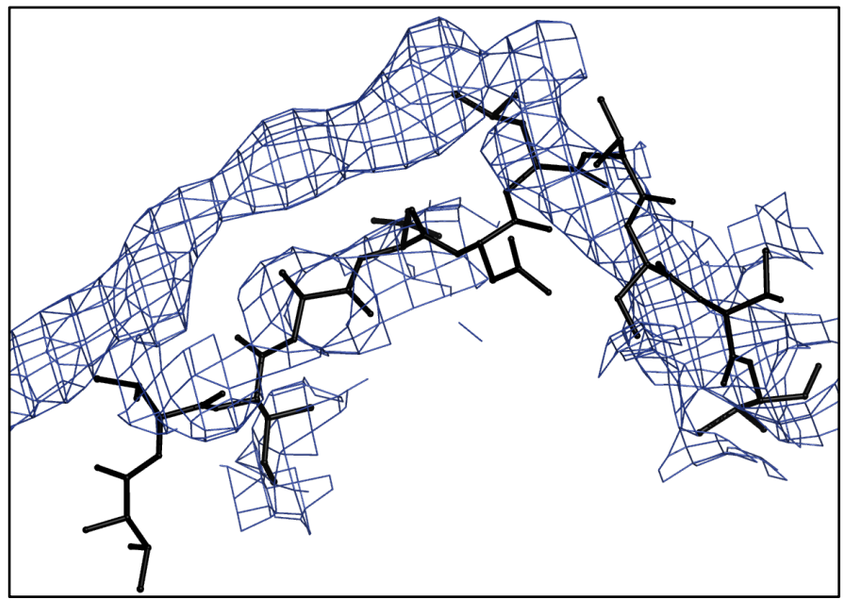
Resolution 4.5 Å
840 chains, 187,320 residues 1,443,960 atoms

Size



Resolution: 11.6 Å

Wide Resolution Range



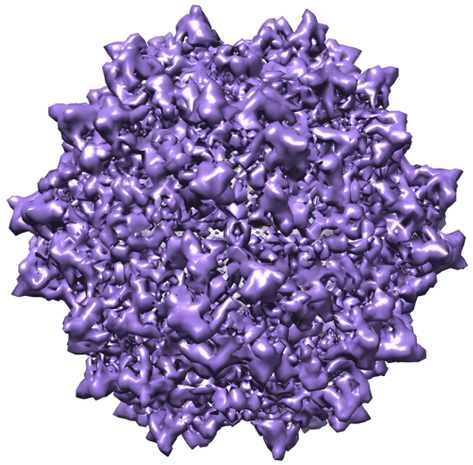
User data, resolution: 3.8 Å

Poor Initial Fit

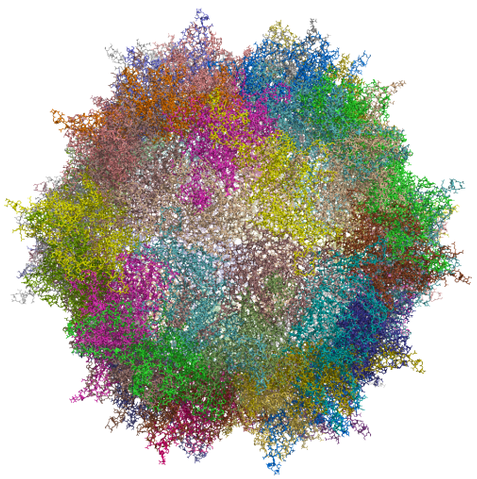
BERKELEY LAB
Lawrence Berkeley National Laboratory

Phenix

Direct Refinement Against the Map



Real space refinement



VOLUME FOUR

JULY MMXIII

COMPUTATIONAL CRYSTALLOGRAPHY NEWSLETTER

ENSEMBLE REFINEMENT, CABLAM

Table of Contents

Editor

Nigel W. Moriarty, NWMoriarty@LBL.Gov

PHENIX News

New programs

FEM: Feature Enhanced Maps (Pavel V. Afonine)

Interpretation of a crystallographic map is a means of obtaining an atomic representation of a crystal structure or the map itself may serve as the crystal model. There are number of factors that affect quality of crystallographic maps that in turn affect difficulty (or even feasibility) of their interpretation and quality of resulting model of crystal structure, and include:

- finite resolution of measured reflections;
- incompleteness of data (missing reflections within the resolution range of the measured data);
- experimental errors in measured reflections;
- errors in atomic model parameters.

These factors a) result in artificial peaks in the map that may be confused with the signal and therefore erroneously interpreted in terms of atomic model, b) introduce noise that may obscure the signal and c) may distort the signal in various ways.

Another fundamentally different contributor to the difficulty of map interpretation is that not all the signal has the same strength. For example, a strong signal arising from a heavy atom derivative may easily obscure a very weak signal (that may be at or below the noise level) arising from a partially occupied very mobile ligand or residue side chain alternative conformation or even hydrogen atoms.

BERKELEY LAB
Lawrence Berkeley National Laboratory

Phenix

THE UNIVERSITY OF CALIFORNIA BERKELEY 1868

Real Space Refinement

- Has a long history in both X-ray crystallography and cryo-EM
  - Early crystallographic refinement programs (Diamond)
  - Alternative to reciprocal space refinement, then applied to EM maps (Chapman)
  - Regularly used in model building (O, Coot)
- New structure fitting approaches make use of real space refinement
  - Molecular dynamics flexible fitting (MDFF)
  - Deformable elastic network fitting (DireX)
  - Rosetta model building and model refinement

Refinement

- An *optimization* algorithm is used to minimize a *target function* by changing the *parameters* of the model
- Parameters:
  - coordinates, atomic displacements, occupancies
- Optimization algorithm:
  - minimization, simulated annealing
- Target function (Objective function):
  - Function based on electron density (real-space refinement)
  - Function based on structure factors (reciprocal-space refinement)

$$E = E_{chem} + w_a \sum_{hkl} \frac{1}{\sigma^2} \left(|F_o| - |F_c| \right)^2$$

BERKELEY LAB
Lawrence Berkeley National Laboratory

Phenix

Goal for Cryo-EM Model Refinement

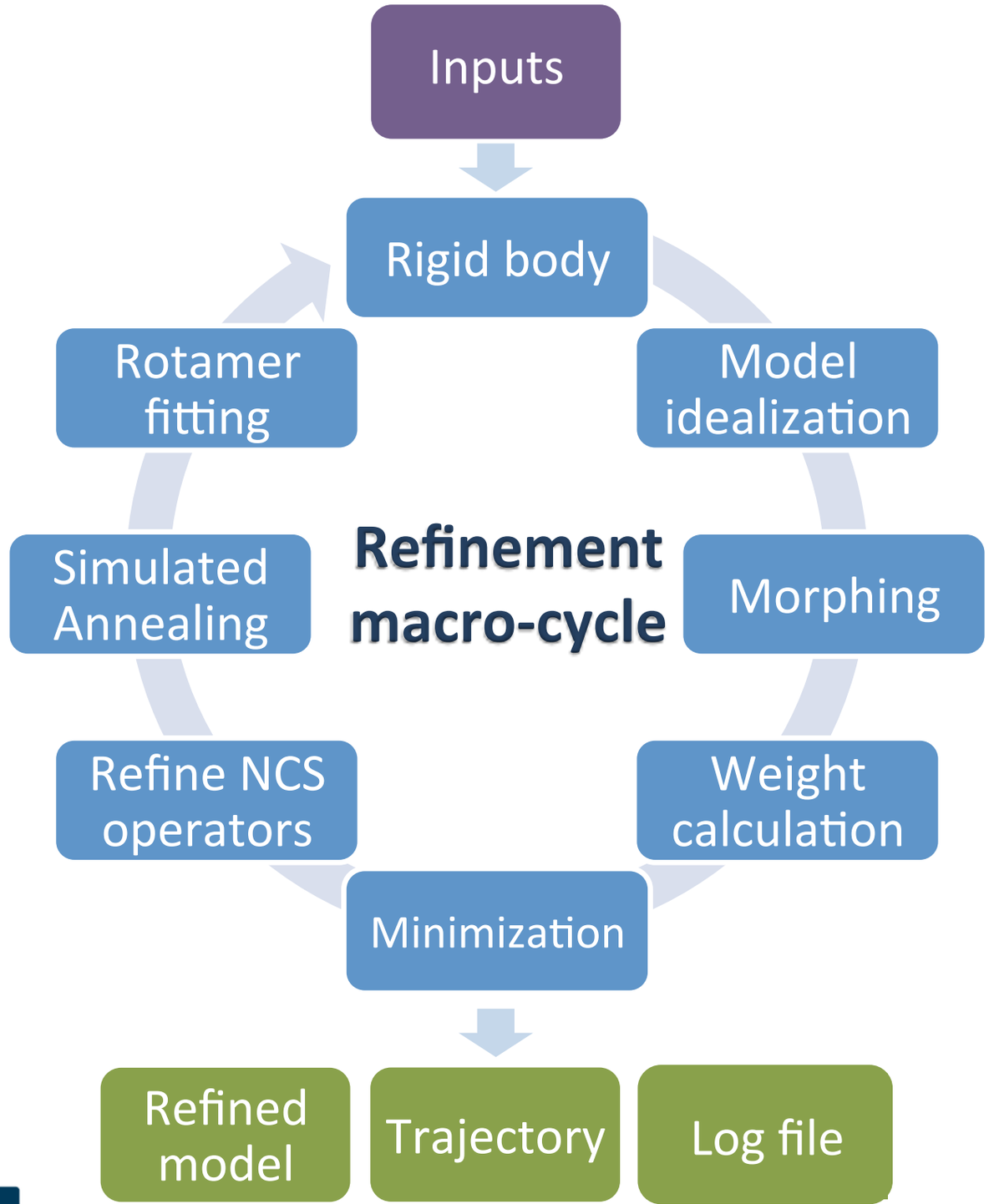
- Stable refinement against any density map (Cryo-EM or X-ray)
- End result should be an improvement in the model
- Large radius of convergence
- Final models with good fit to density and physically reasonable geometry (Ramachandran distribution, rotamers, packing)
- Fast: no more than one second per residue

$$E = E_{chemistry} + \omega\sum(\rho_o-\rho_c)^2$$

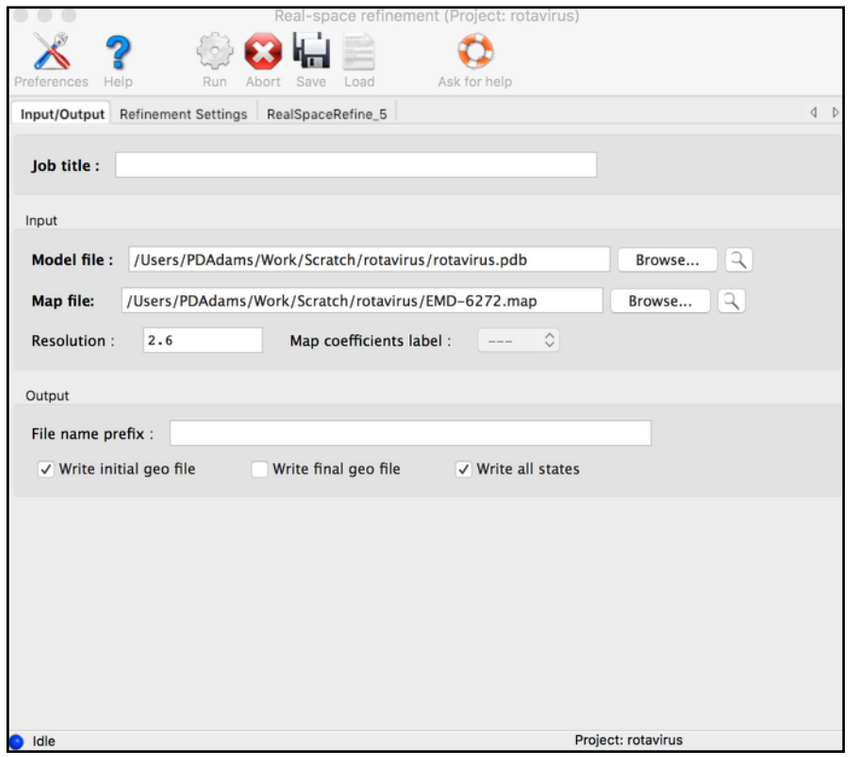
BERKELEY LAB
Lawrence Berkeley National Laboratory

Phenix

Real Space Refinement Procedure



- `phenix.real_space_refine`



Pavel Afonine, Oleg Sobolev, Billy Poon (LBNL), Tom Terwilliger (LANL)

Hryc et al. Accurate model annotation of a near-atomic resolution cryo-EM map. *Proc Natl Acad Sci U S A* 2017, **114**:3103-3108.

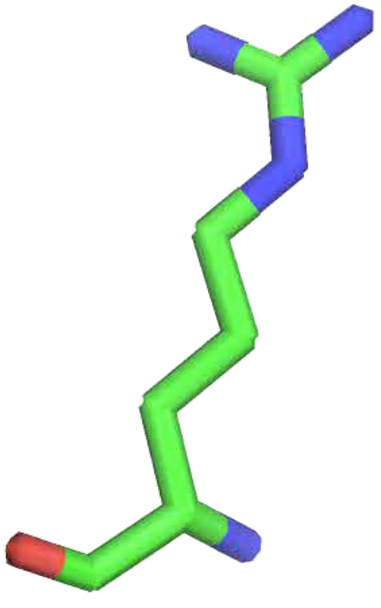
Afonine et al. Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta Cryst* 2018, **D74**:531-544.

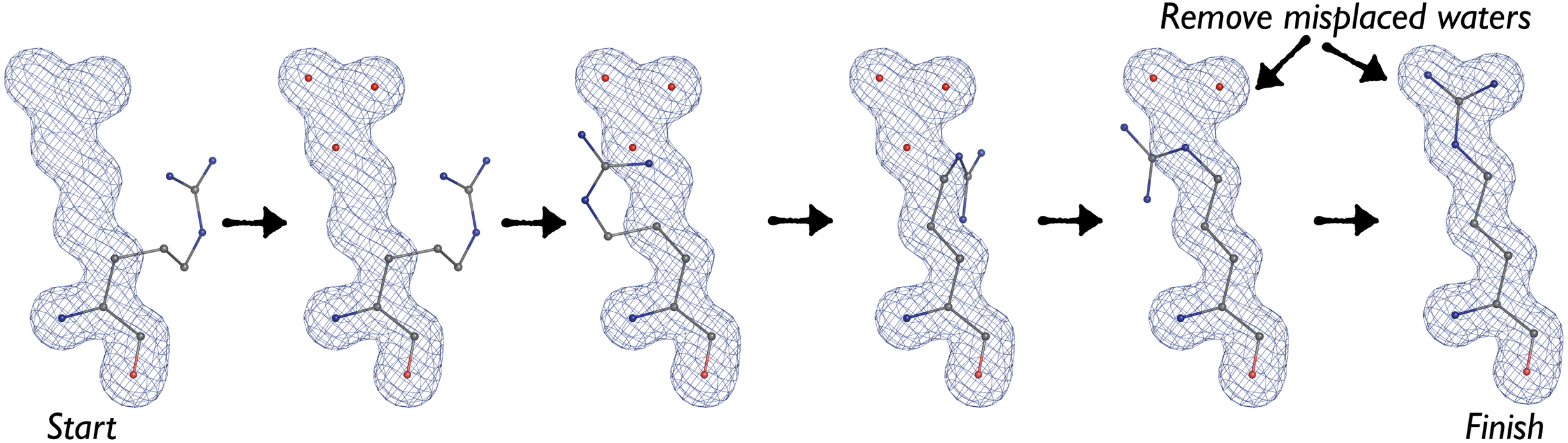
BERKELEY LAB
Lawrence Berkeley National Laboratory

Phenix

Systematic Searching of Rotamers

- In a protein structure 99% of the side chains obey known rotameric conformations
- Often errors are fixed manually but can now be fixed automatically following structure validation
- A systematic search through rotamer space is combined with a fit-to-density score





Fast: 0.01 – 1 second per residue

BERKELEY LAB
Lawrence Berkeley National Laboratory

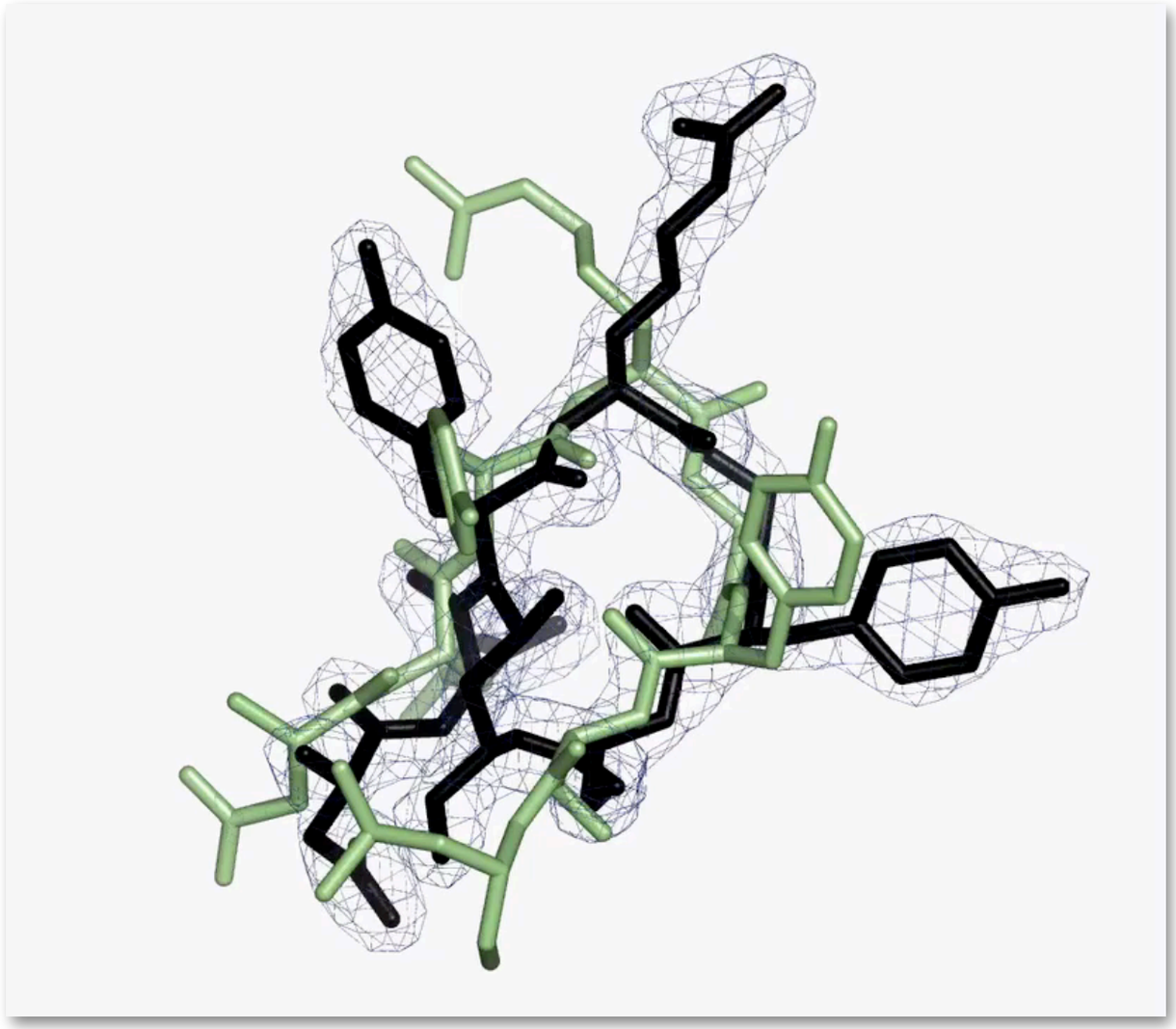
Phenix

Pavel Afonine, Jeff Headd, Nat Echols

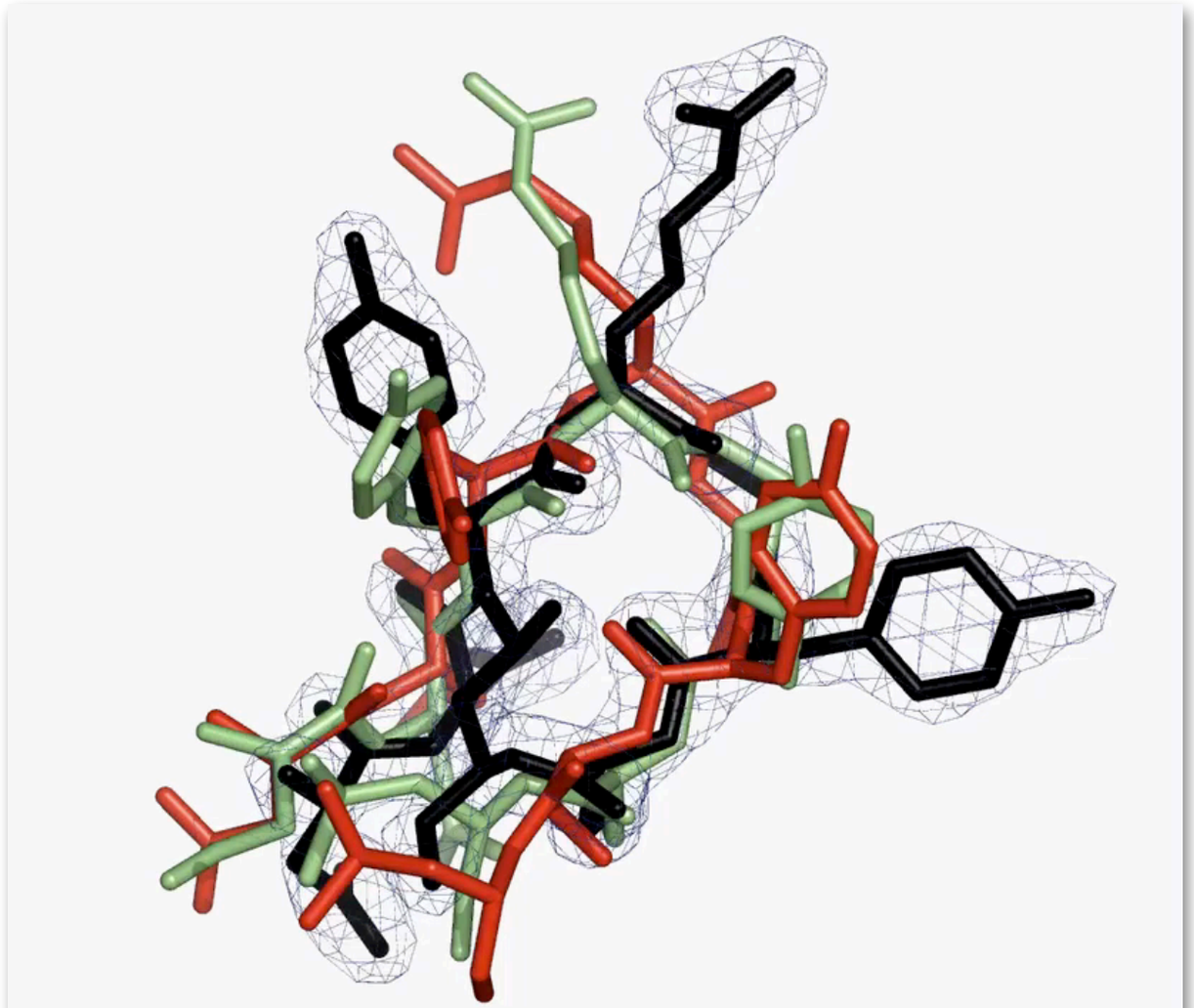
Afonine et al., Acta Cryst. 2012, **D68**:352-367

Optimization In Real Space

- Refinement against a map using minimization or other optimization method
- Minimization can get caught in local minima
- Simulated annealing is a method used to escape minima



Minimization



Simulated Annealing

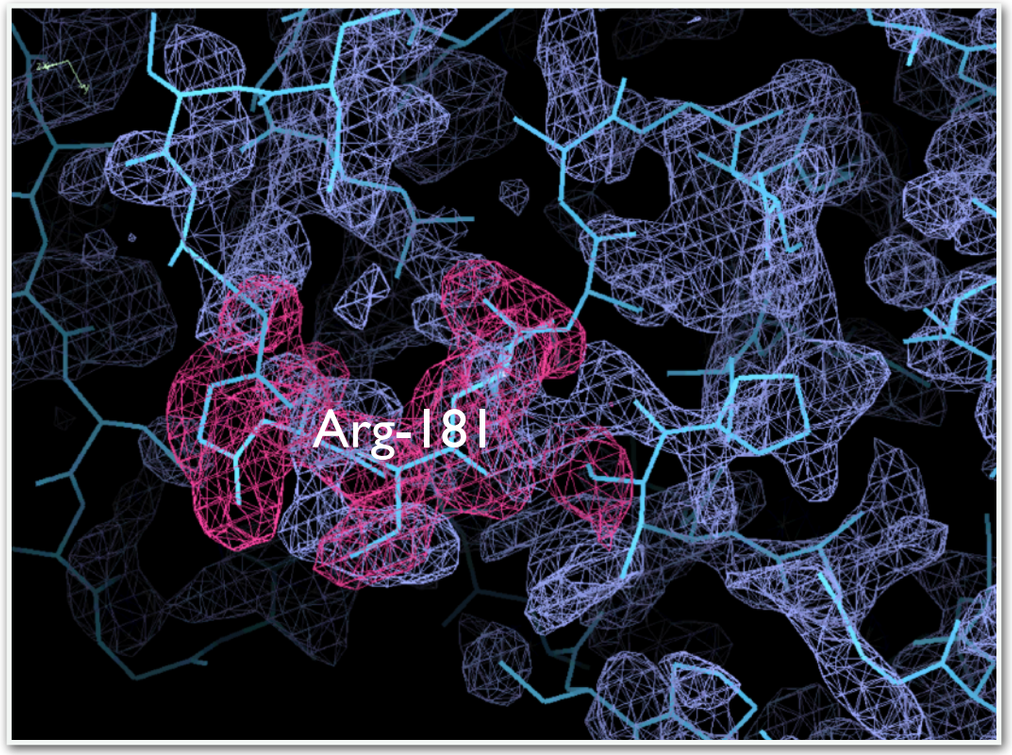
BERKELEY LAB
Lawrence Berkeley National Laboratory

Phenix

Pavel Afonine (LBNL)

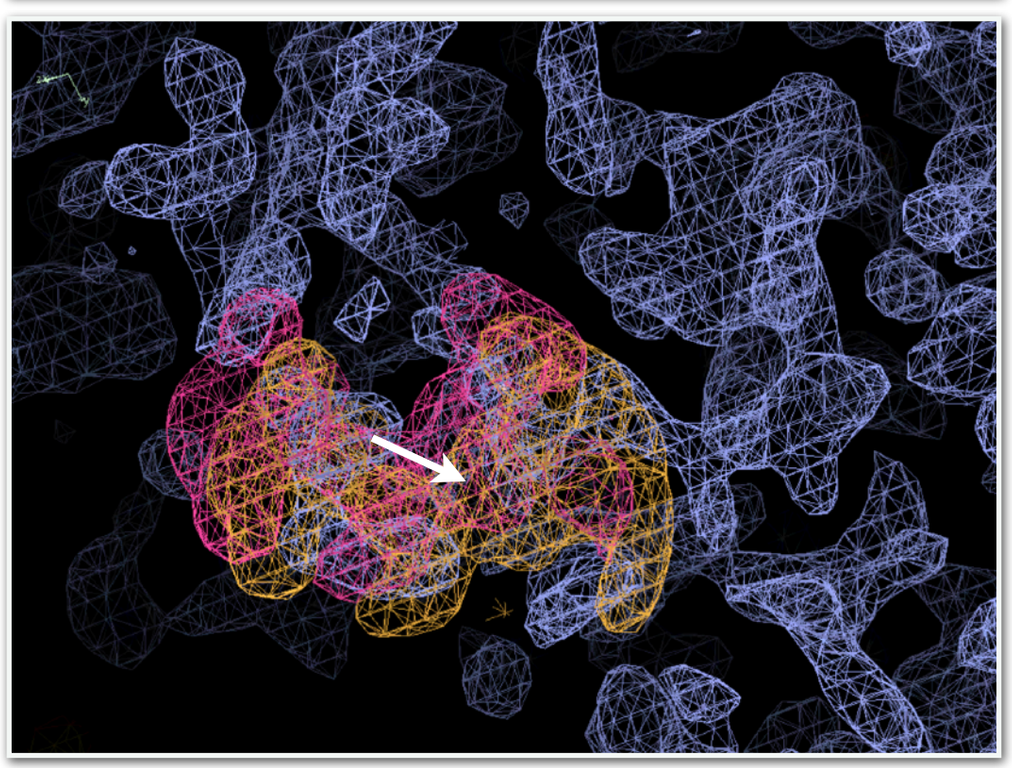
THE UNIVERSITY OF CALIFORNIA BERKELEY 1868

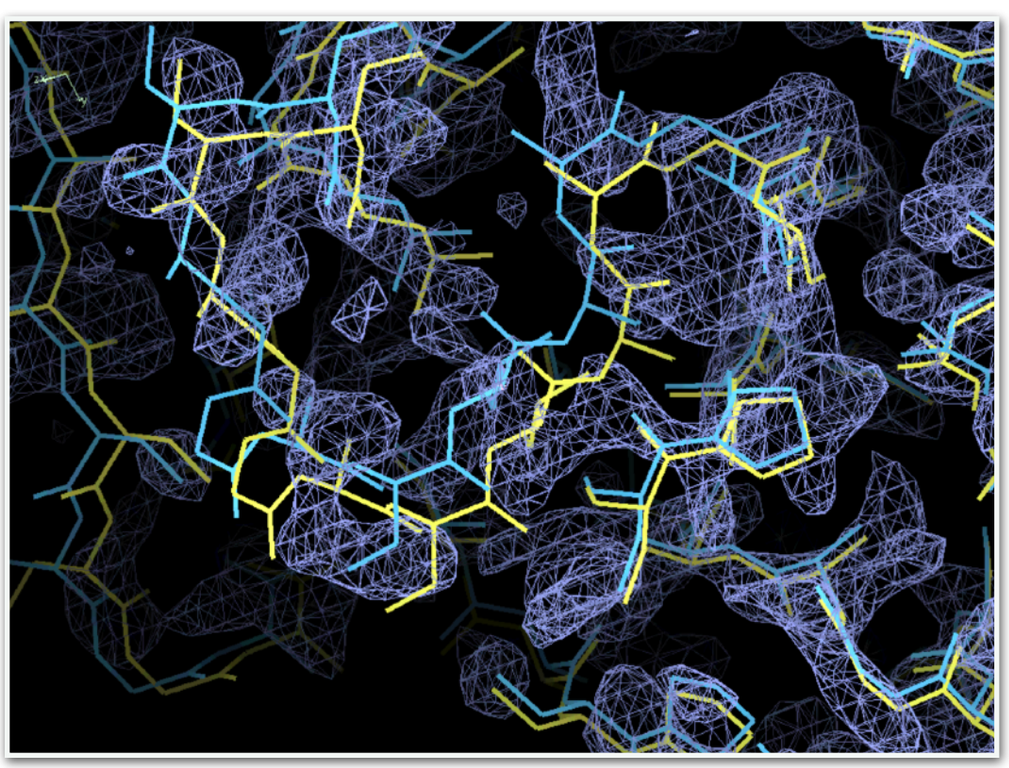
Morphing



- Identify local translation to apply to one C_α atom and nearby atoms
- Smooth the local translations in window of 10 residues
- Apply the smoothed translation to all atoms in the residue

Tom Terwilliger, Los Alamos National Laboratory





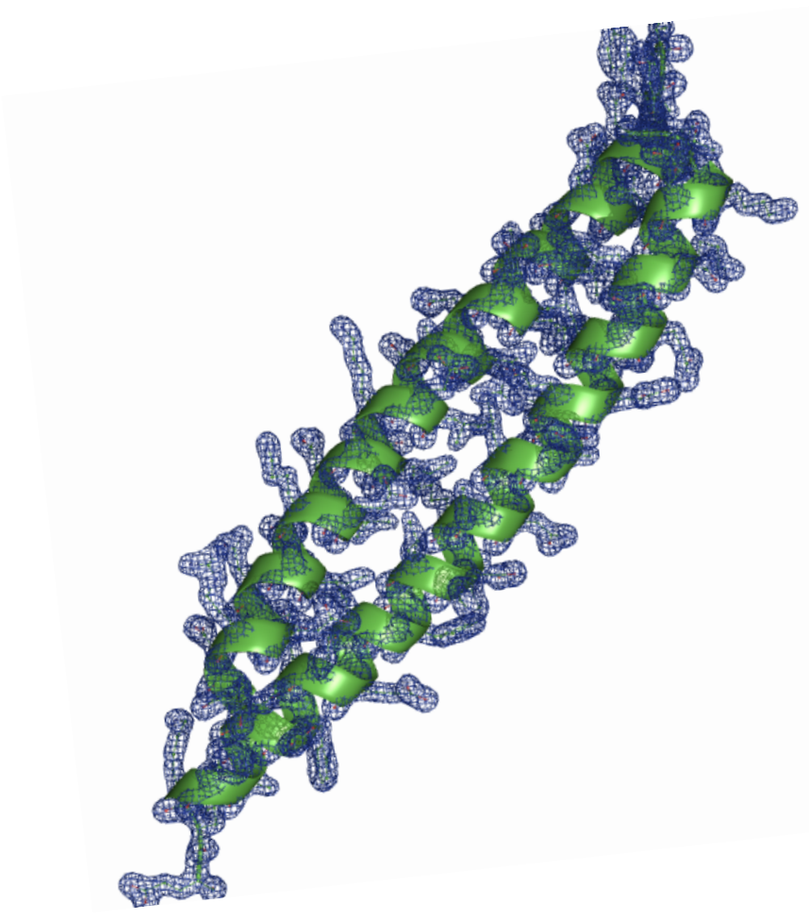
Terwilliger et al., Acta Cryst. 2012, **D68**:861-870

Phenix

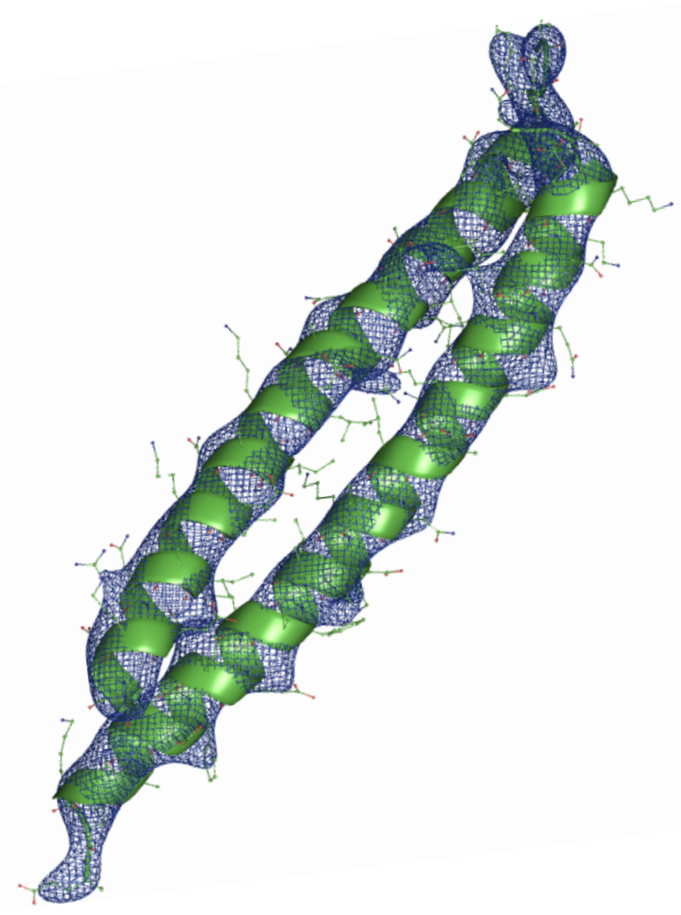
Terwilliger et al., Acta Cryst. 2013, **D69**:2244-2250

Lower Resolution Requires Additional Information

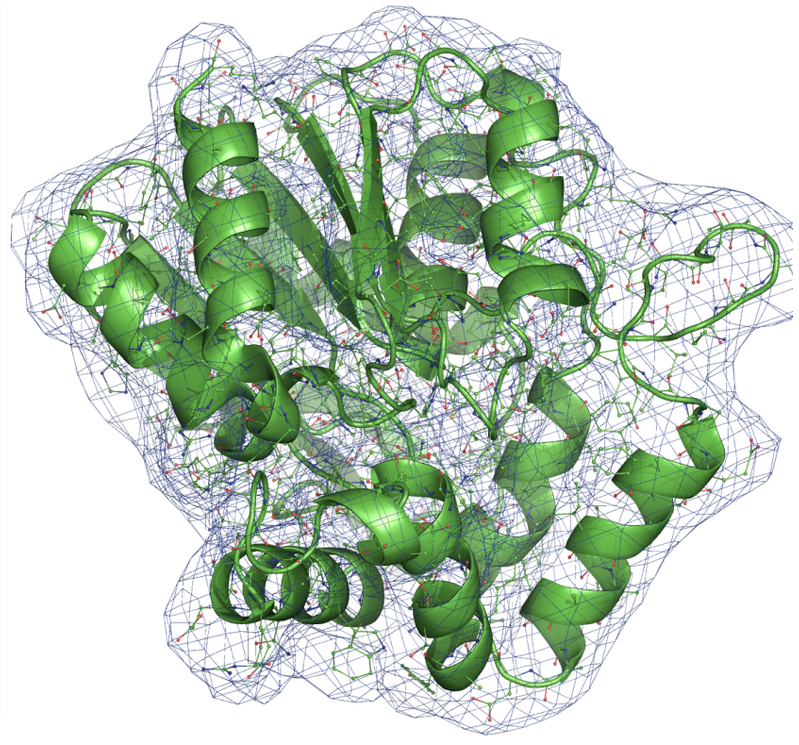
High Resolution **Low Resolution**



Side chains



Secondary Structure



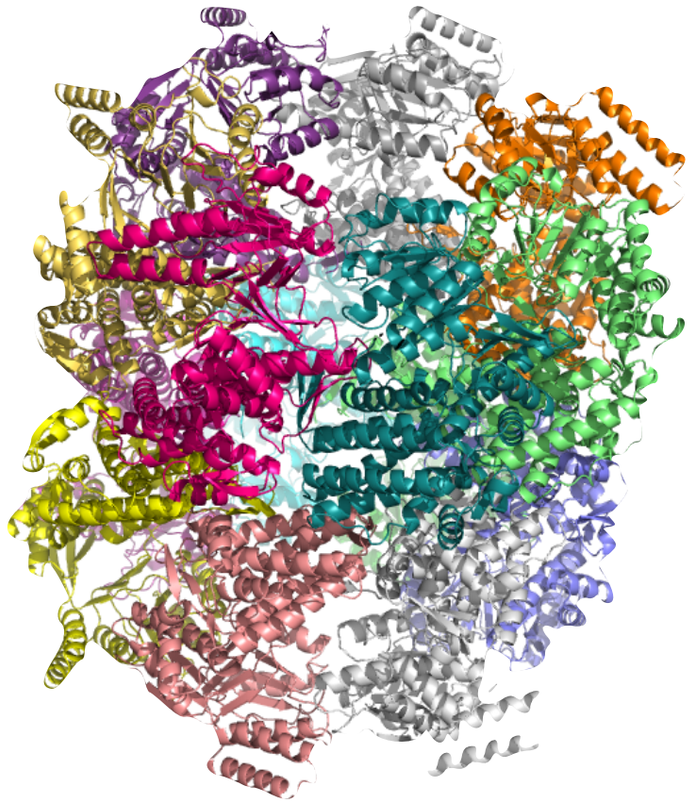
Molecule

BERKELEY LAB
Lawrence Berkeley National Laboratory

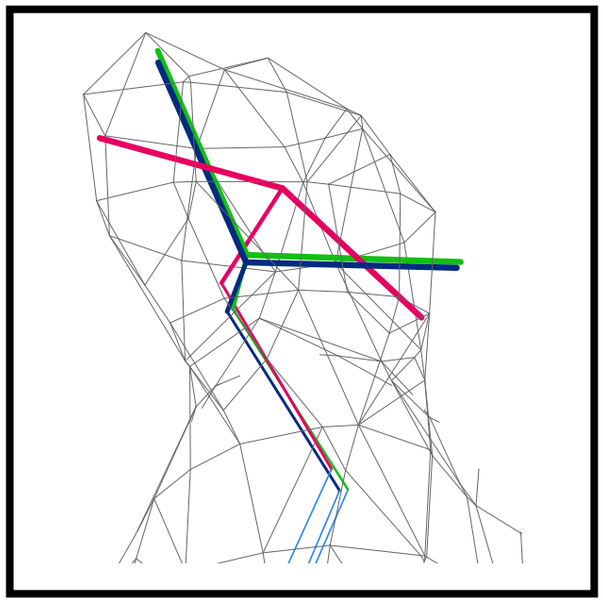
Phenix

THE UNIVERSITY OF CALIFORNIA · BERKELEY · 1868

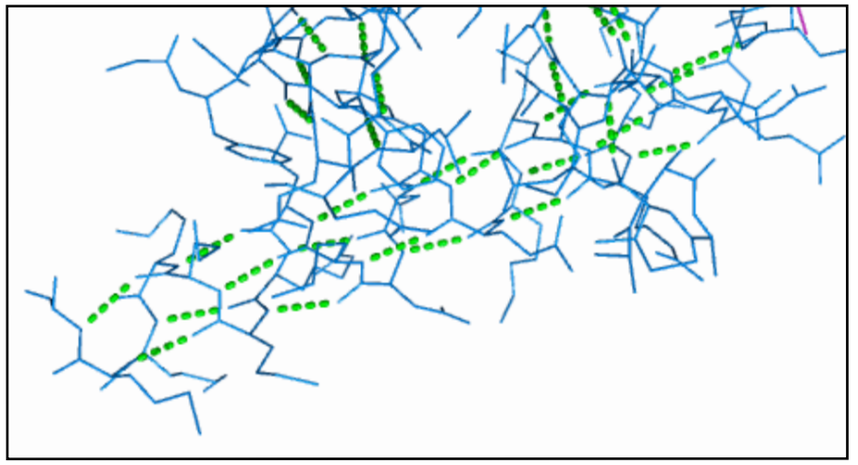
Other Model Restraints



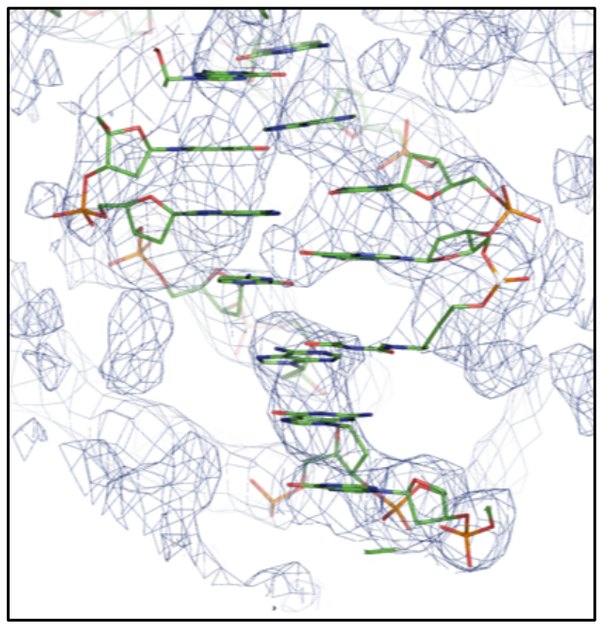
- Symmetry constraints
- Multiple symmetry groups
- Optimization of NCS operators (w.r.t density)
- Automatic expansion of monomer from MTRX records



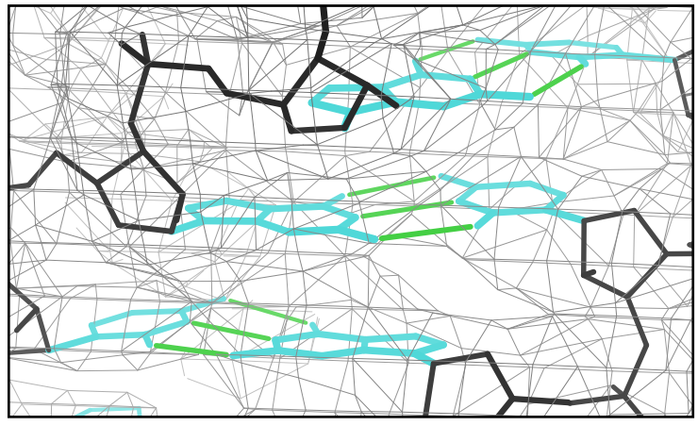
Reference model torsion angle restraints



Secondary structure restraints



Parallelity restraints



Base pairing restraints

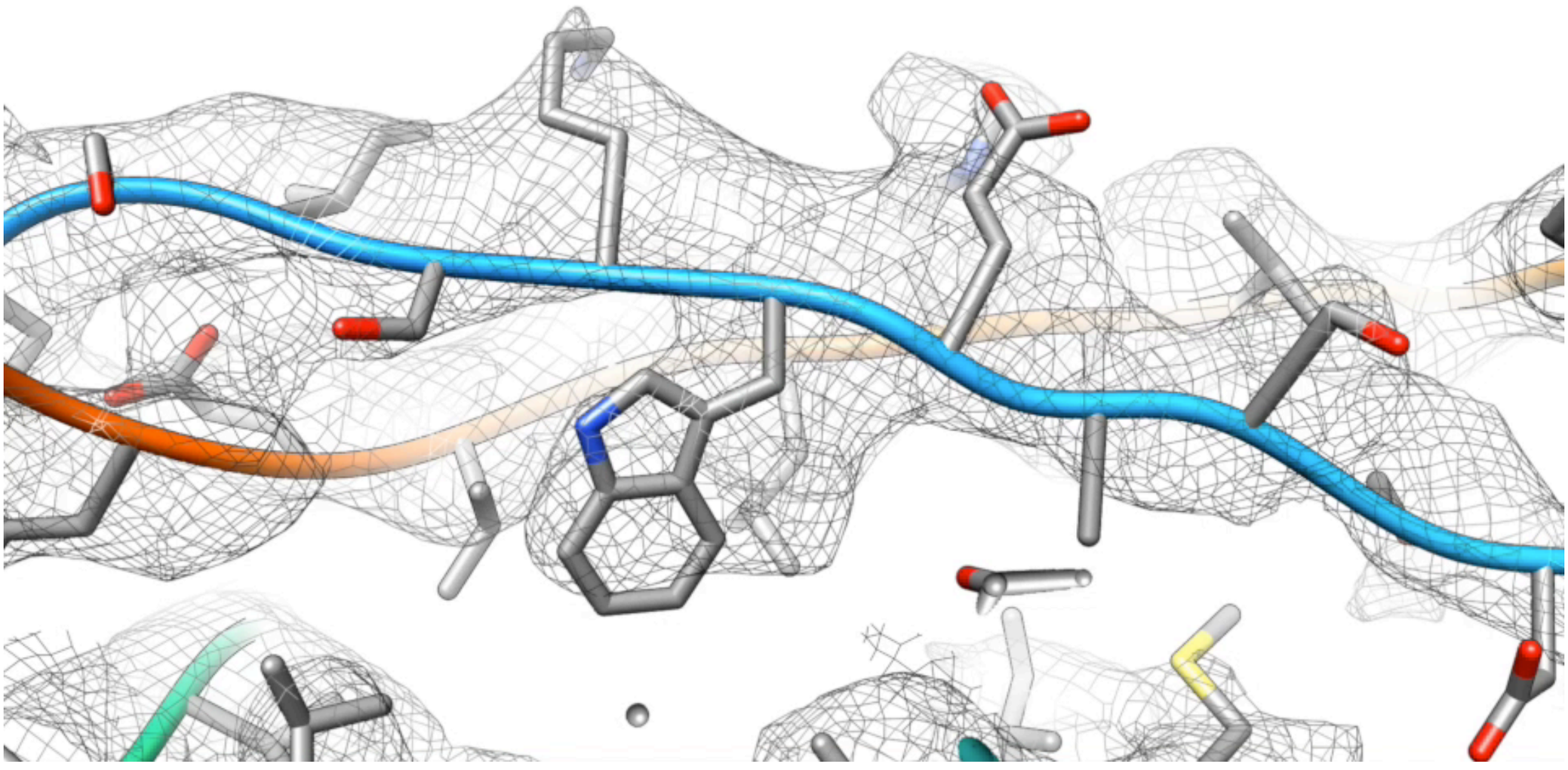
BERKELEY LAB
Lawrence Berkeley National Laboratory

THE UNIVERSITY OF CALIFORNIA BERKELEY 1868

Maintaining Stereochemistry

- After identification of a rotamer conformation:
  - Apply torsion angle restraints to maintain rotameric state
- Possible to apply restraints to the Ramachandran distribution
- Possible to apply restraints to maintain secondary structure elements
- Do not accept changes that generate poor geometry

Virus Structure Refinement



Collaboration with Wah Chiu, Zhao Wang, Corey Hryc, Matthew Baker (Baylor College of Medicine)

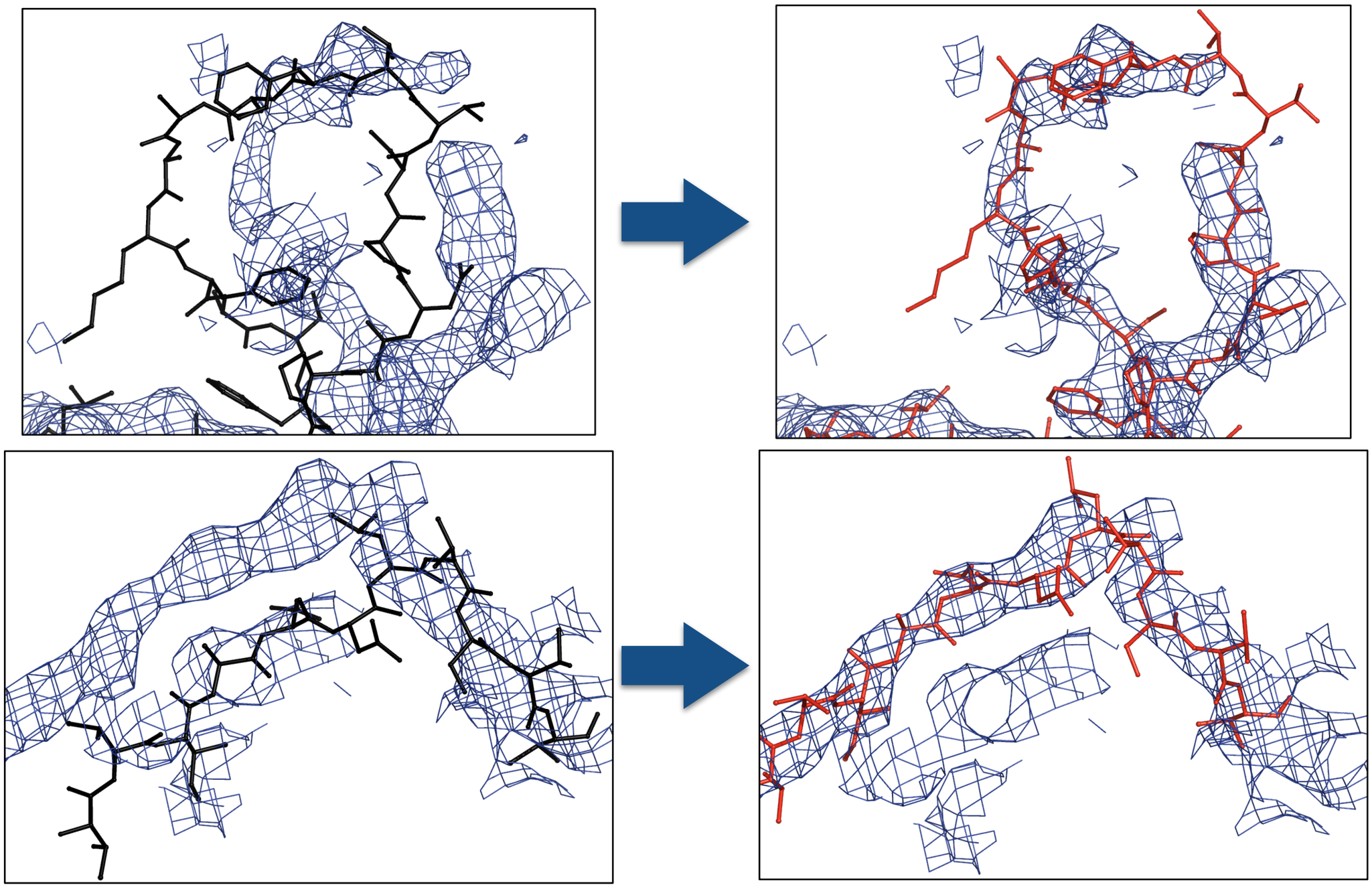
Pavel Afonine (LBNL), Corey Hryc (BCM)

Phenix

Wang et al., Nature Commun. 2014, **5:**4808

Real Space Refinement Improves Fit to Data

- Models are moved to better fit the Cryo-EM map



BERKELEY LAB
Lawrence Berkeley National Laboratory

Phenix

THE UNIVERSITY OF CALIFORNIA BERKELEY 1868

While Also Improving Stereochemistry

- Standard metrics (MolProbity) are all improved

	Starting Model	After RS Refinement
Clashscore	47	18
Ramachandran Outliers (%)	6	2
Ramachandran Allowed (%)	16	8
Ramachandran Favoured (%)	78	90
Rotamer Outliers (%)	20	0
C_β Deviations (%)	2	0

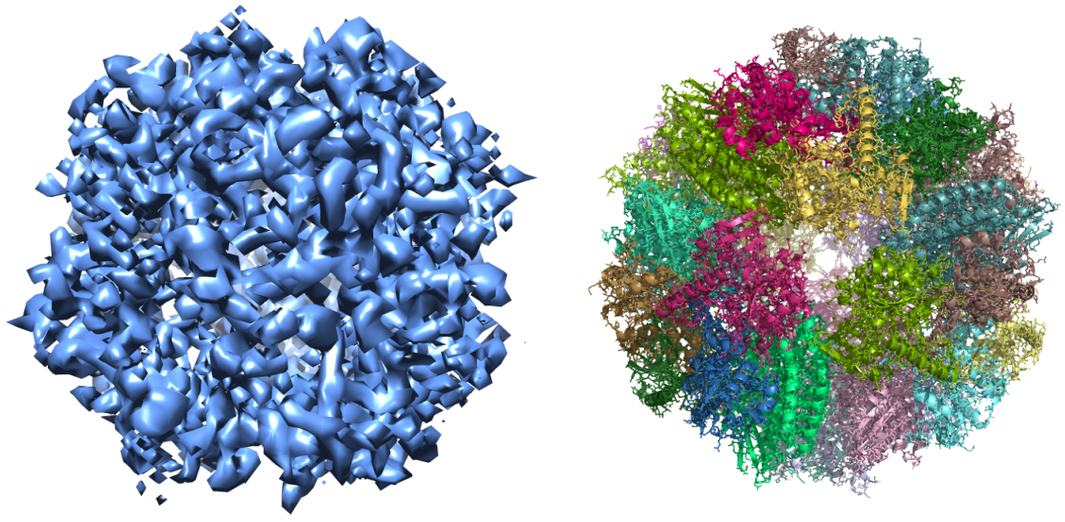
BERKELEY LAB
Lawrence Berkeley National Laboratory

Phenix

THE UNIVERSITY OF CALIFORNIA BERKELEY 1868

Typical Results at Higher Resolution

Resolution: 3.36 Å

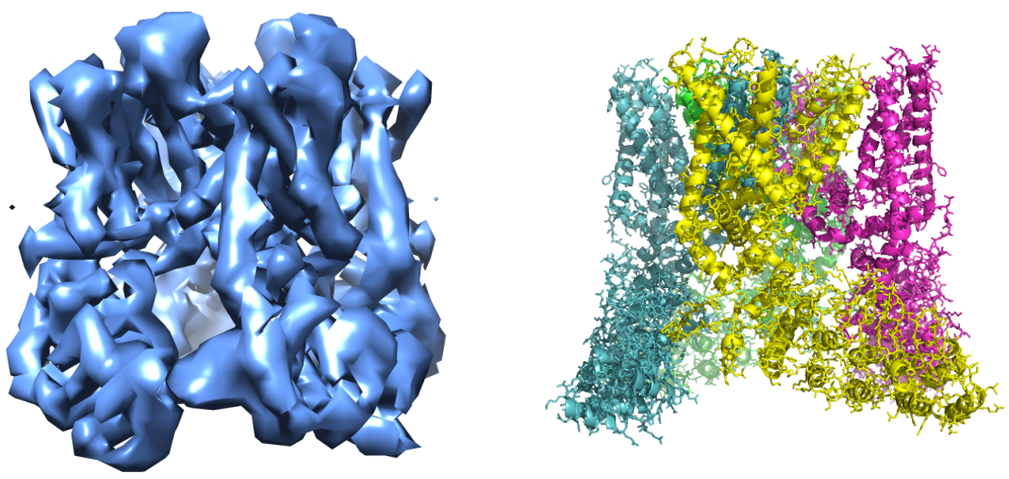


Residues/atoms: 10,716/82,404

Refinement: 173 min

METRIC	Original	*Phenix*
Map CC	0.645	0.783
RMSD (bonds/angles)	0.02/2.05	0.01/1.21
Clashscore	117.1	18.79
Rama. outl., %	0.11	0.11
Rotamer outl., %	35.51	0
C-beta deviations	24	0

Resolution: 3.8 Å



Residues/atoms: 2,324/17,424

Refinement: 20 min

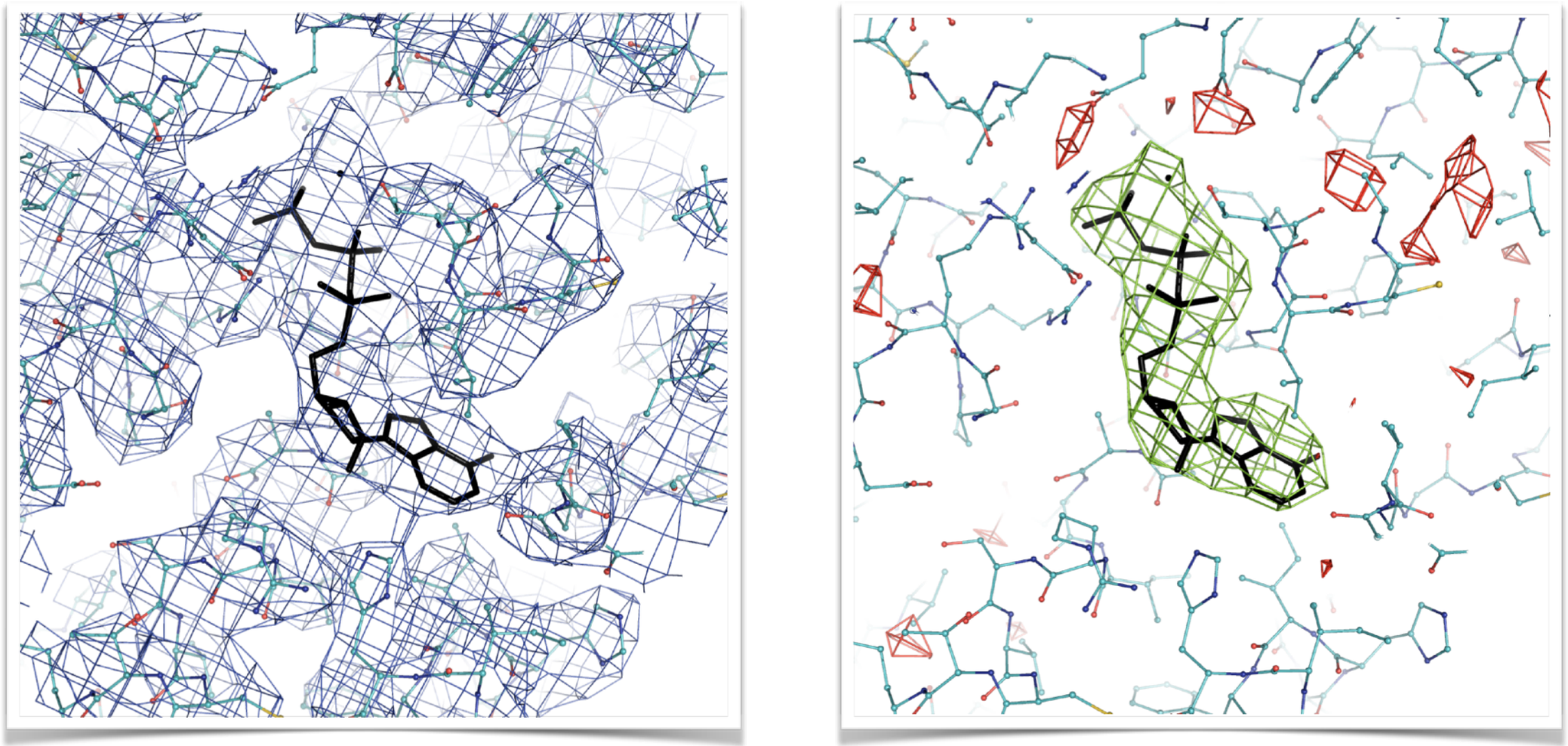
METRIC	Original	*Phenix*
Map CC	0.650	0.714
RMSD (bonds/angles)	0.01/1.34	0.01/1.31
Clashscore	100.9	32.84
Rama. outl., %	0.52	0
Rotamer outl., %	27.99	0
C-beta deviations	0	0

BERKELEY LAB
Lawrence Berkeley National Laboratory

Phenix

Difference Maps

- Local scaling of map and model density, real space subtraction
- Reveal features missing from the model



```
phenix.real_space_diff_map model.pdb map.ccp4 resolution=3.5
```

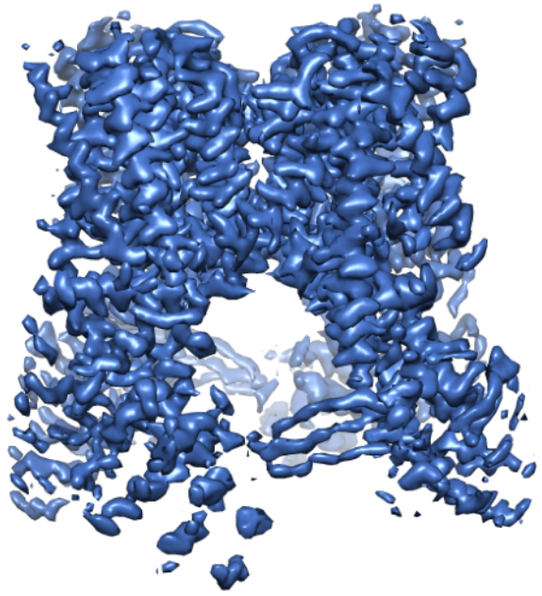
BERKELEY LAB
Lawrence Berkeley National Laboratory

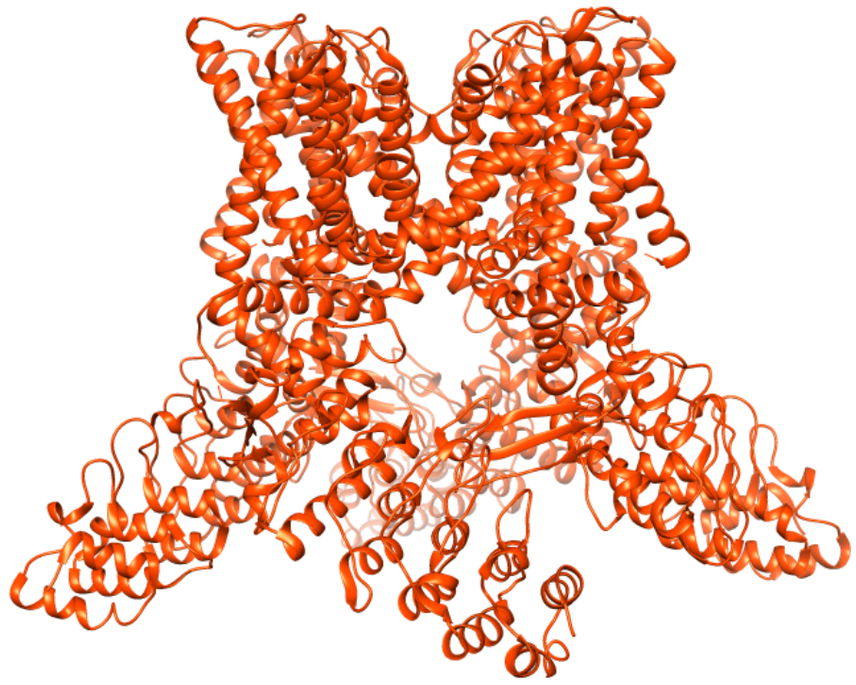
Phenix

THE UNIVERSITY OF CALIFORNIA BERKELEY 1868

Validation and Cryo-EM

- Does the map make sense?
  - Gold Standard FSC of half maps
- Does the model make sense?
  - MolProbity
- Does the model fit the map?
  - Overall and local density correlation
  - What about the detailed local fit?



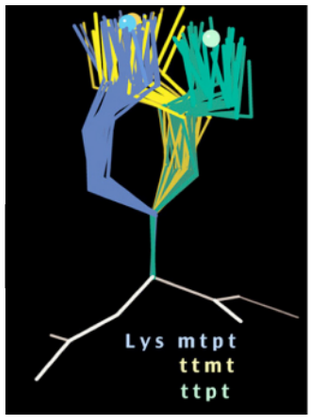


UCSF
University of California
San Francisco

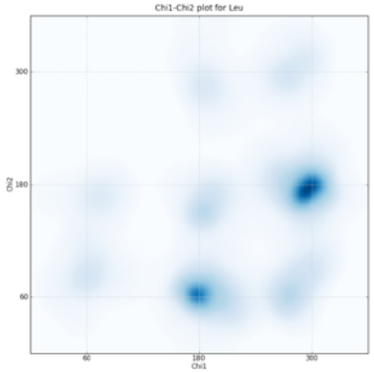
Phenix

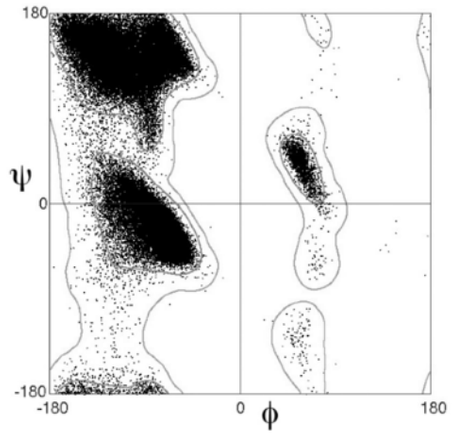
BERKELEY LAB
Lawrence Berkeley National Laboratory

Sources of Prior Information

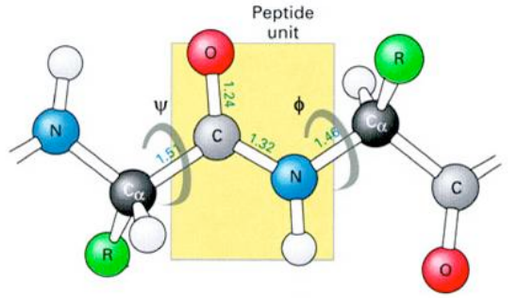


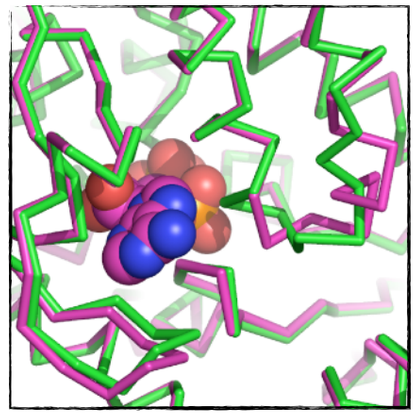
Sidechain distributions



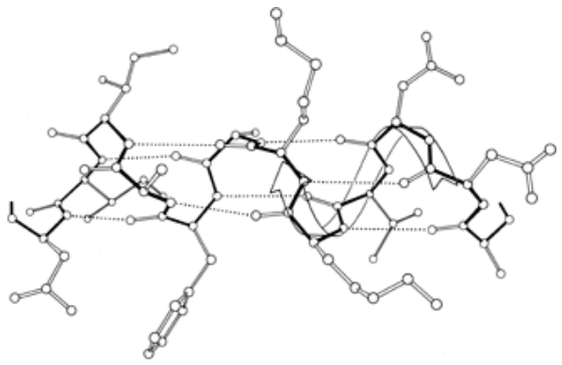


Mainchain distributions

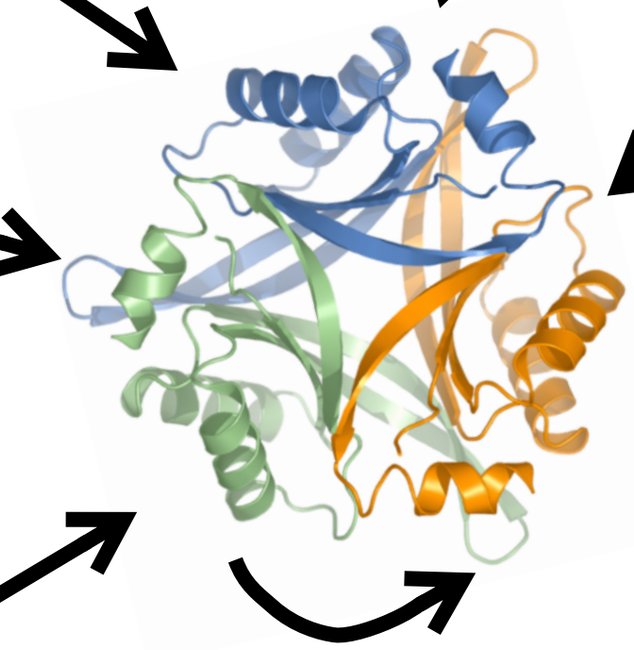


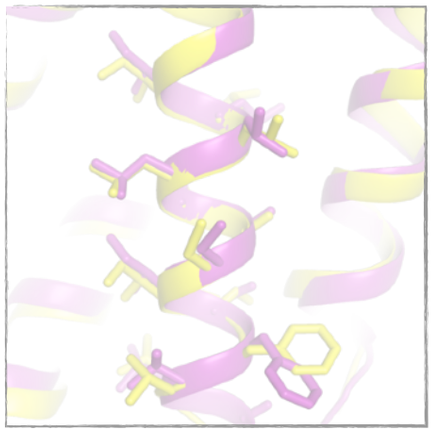


Related structures



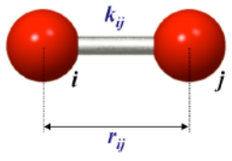
Secondary structure

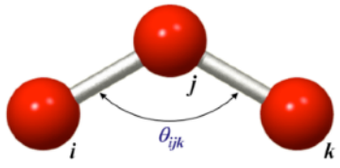




Physically Realistic Potentials

Covalent geometry





Images from PumMa web site (http://www.pumma.nl)

Internal symmetry

DiMaio et al., Nature Methods 2013, **10**:1102-1104

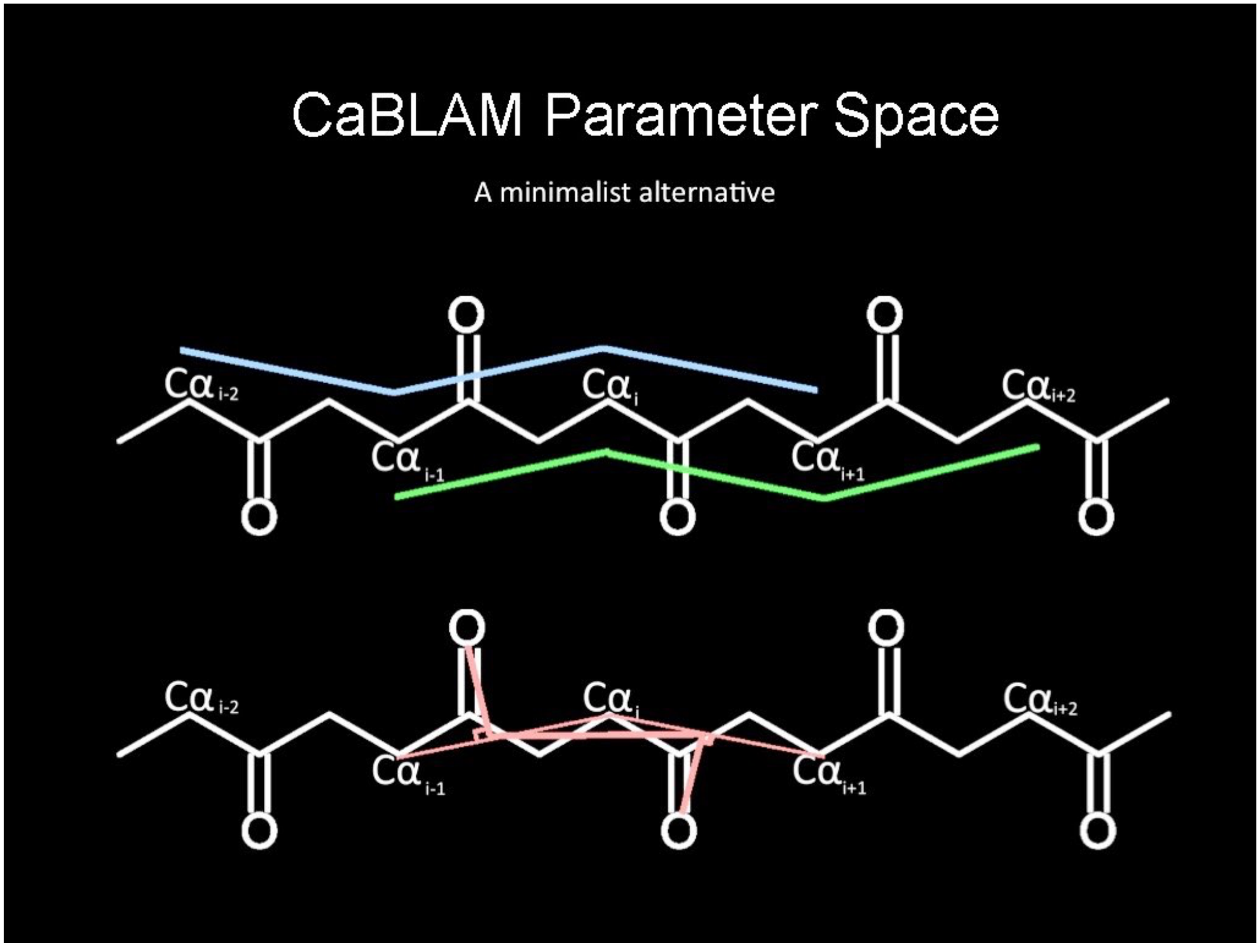
Headd et al., Acta Cryst. 2014, **D70:**1346-1356

Headd et al., Acta Cryst. 2012, **D68**:381-390

Phenix

BERKELEY LAB
Lawrence Berkeley National Laboratory

Validation Using Cα Atoms



BERKELEY LAB
Lawrence Berkeley National Laboratory

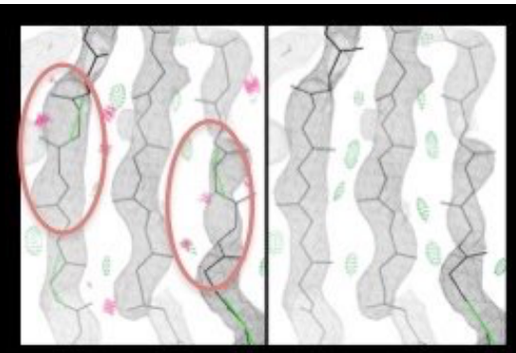
Phenix

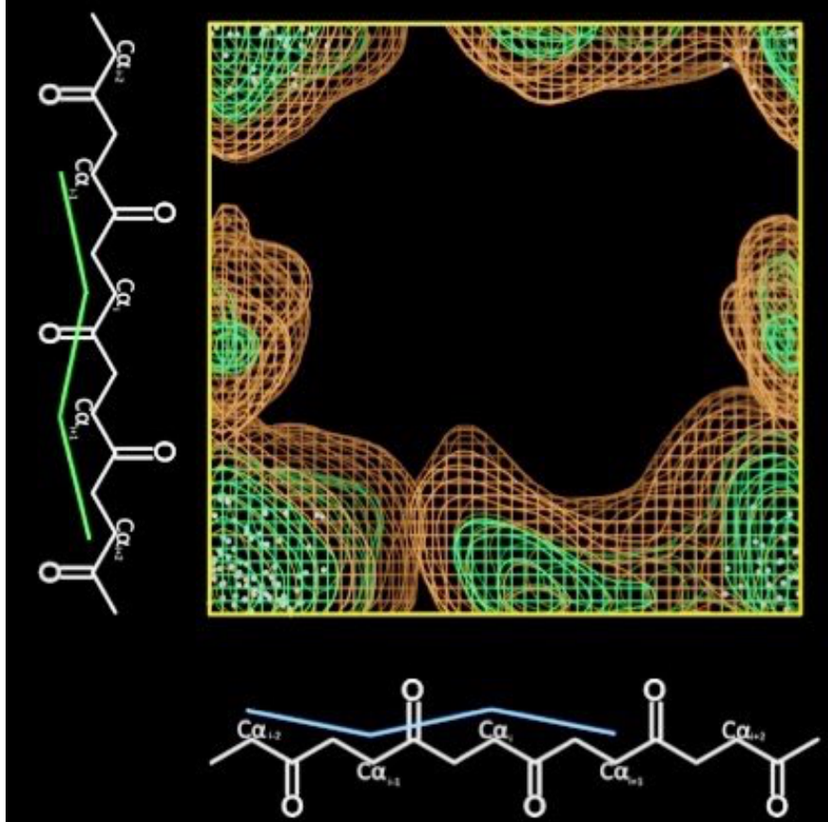
Christopher Williams,
Duke University

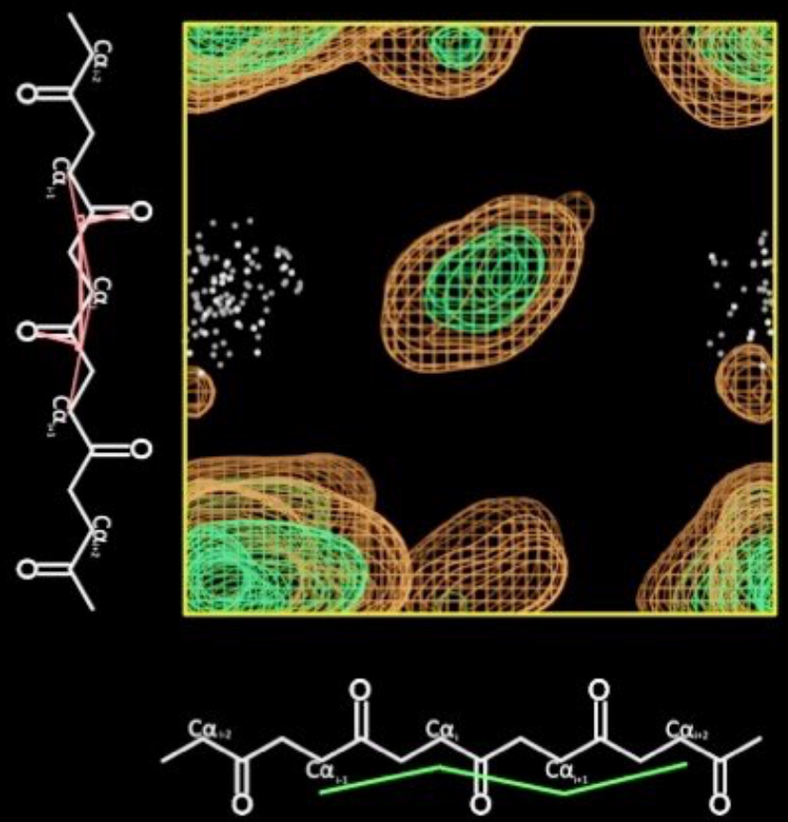
Identifying Distorted Secondary Structure

Diagnosing Strands

Pathological strands from 70S Ribosome



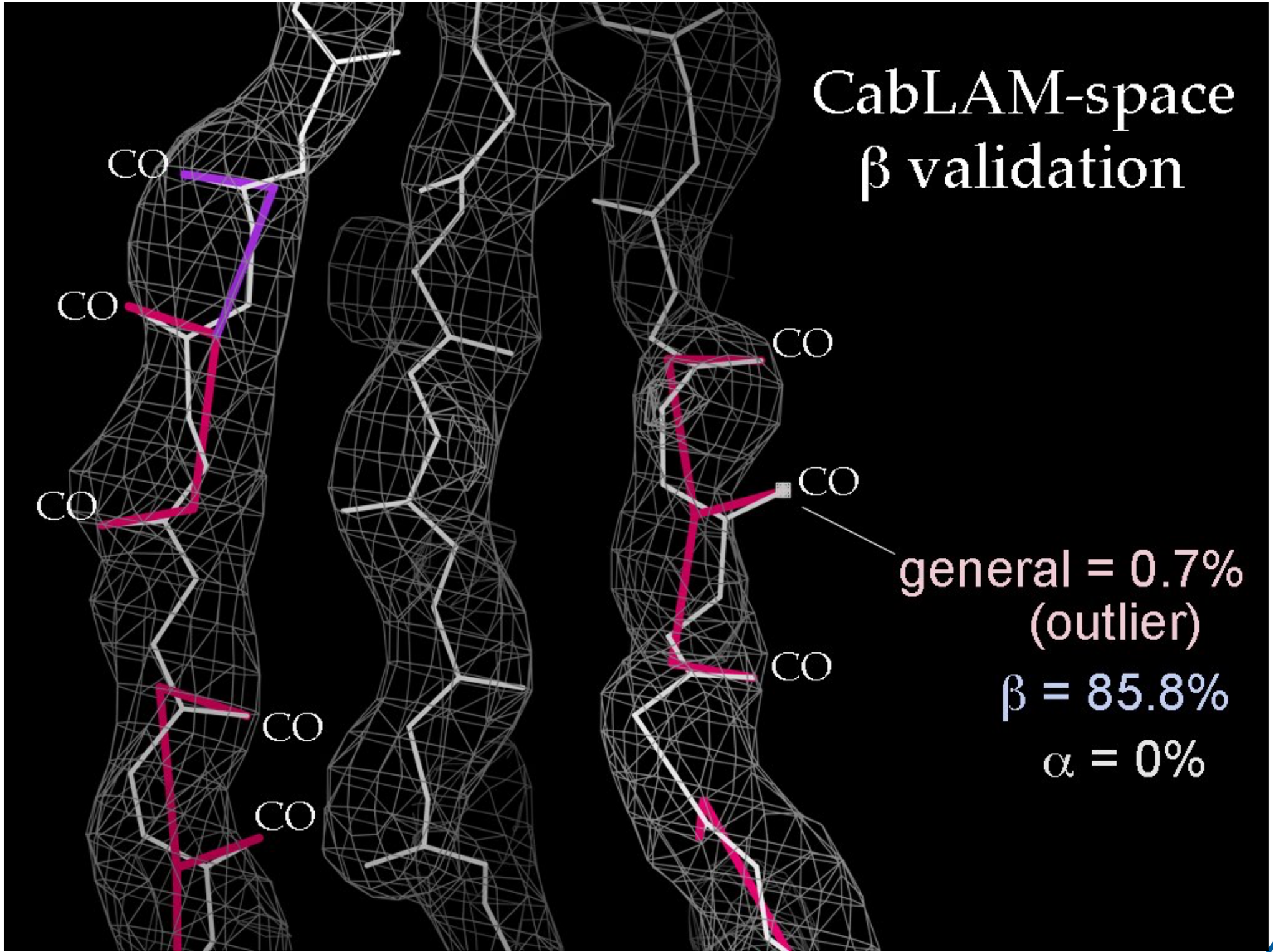




Phenix

Christopher Williams,
Duke University

Assessing Secondary Structure Probability

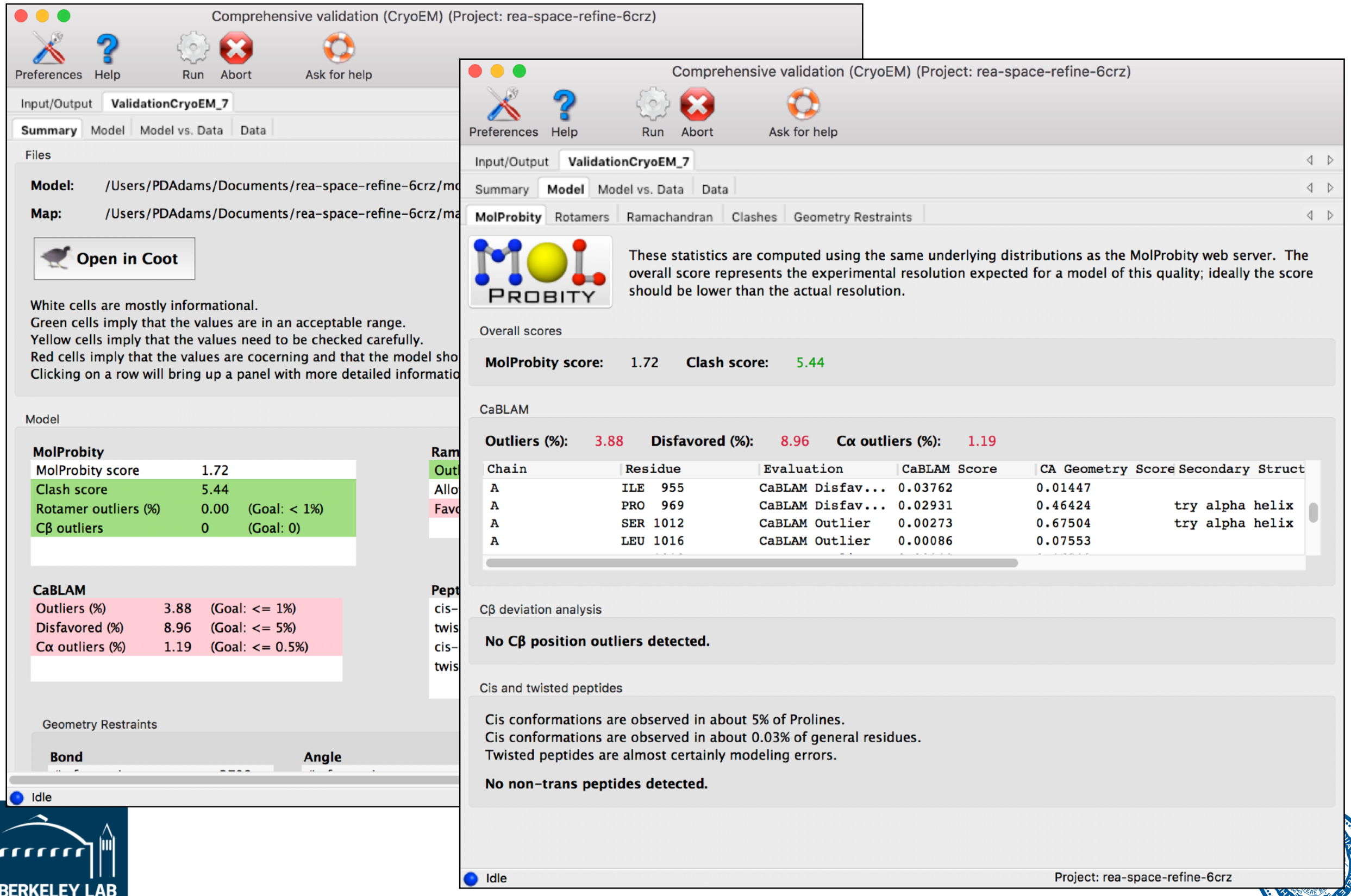


BERKELEY LAB
Lawrence Berkeley National Laboratory

Phenix

Christopher Williams,
Duke University

Comprehensive Validation



BERKELEY LAB
Lawrence Berkeley National Laboratory

Model Map Validation

Benjamin Barad, Yifan Cheng, Jaime Fraser
University of California San Francisco
Ray Yu-Ruei Wang, Frank DiMaio
University of Washington
Nat Echols
Lawrence Berkeley National Laboratory

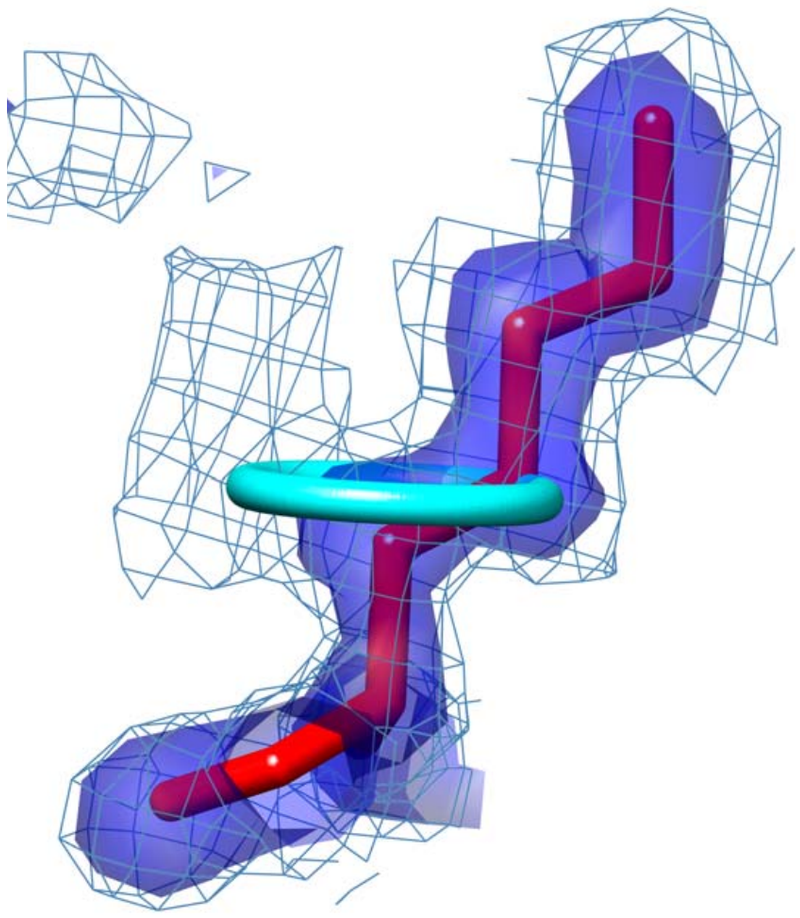
UCSF
University of California
San Francisco

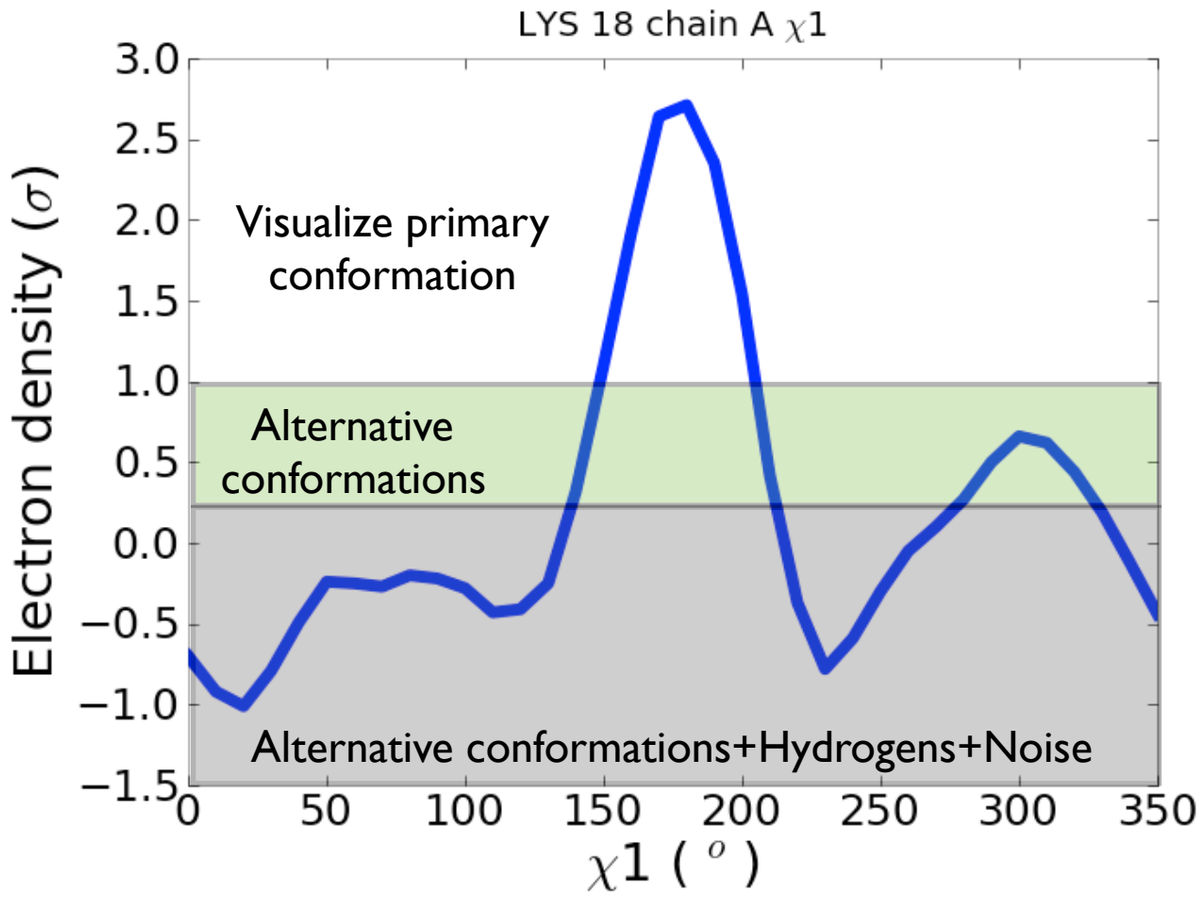
Phenix

BERKELEY LAB
Lawrence Berkeley National Laboratory

Look at the Density Around Sidechains

Ringer





Lang PT, et al. Automated electron-density sampling reveals widespread conformational polymorphism in proteins. *Protein Science*. 2010.

UCSF University of California San Francisco

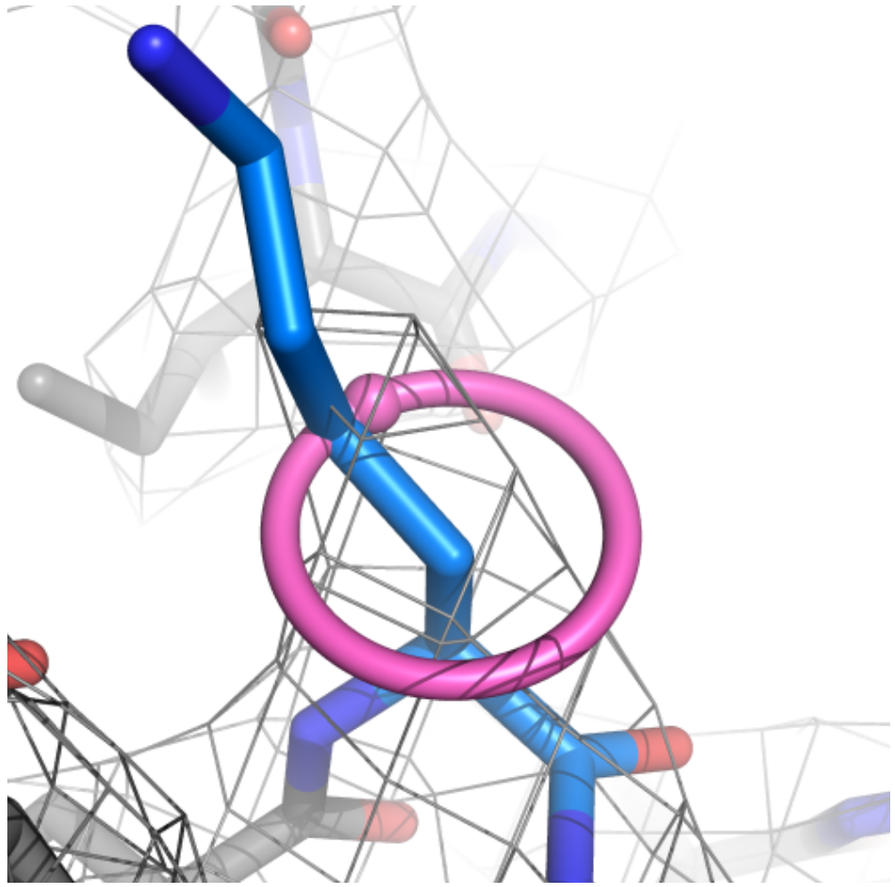
Phenix

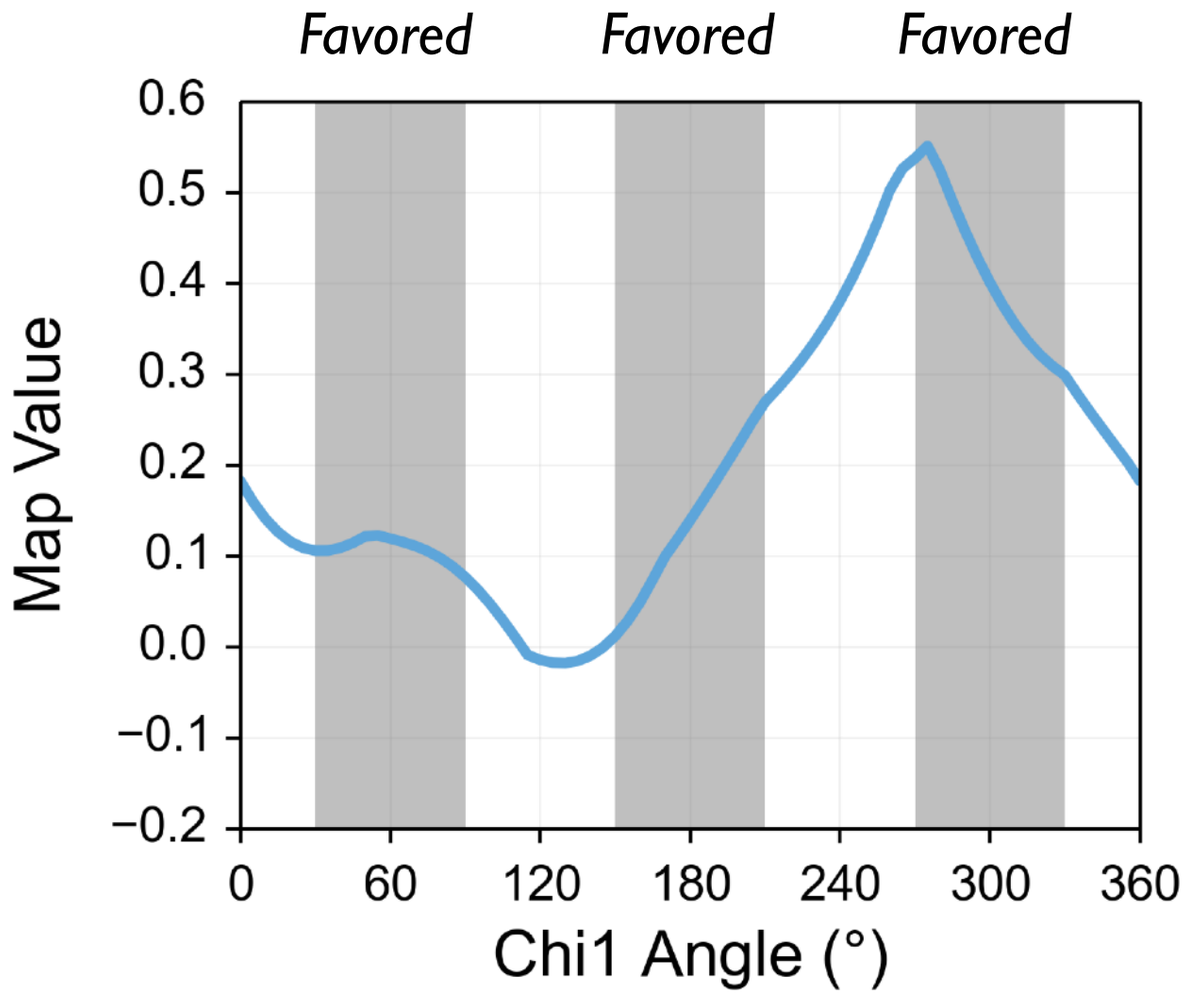
Ben Barad, Jaime Fraser, UCSF

BERKELEY LAB Lawrence Berkeley National Laboratory

Look at the Density Around Sidechains

EMRinger





Barad BA, et al. EMRinger: Side-chain-directed model and map validation for 3D Electron Cryomicroscopy. *Nature Methods*. 2015

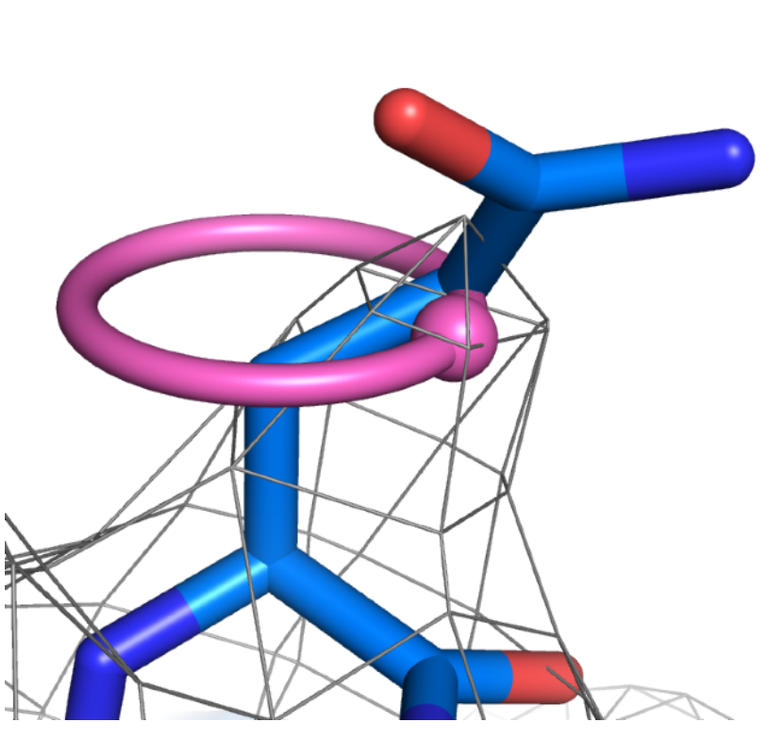
UCSF University of California San Francisco

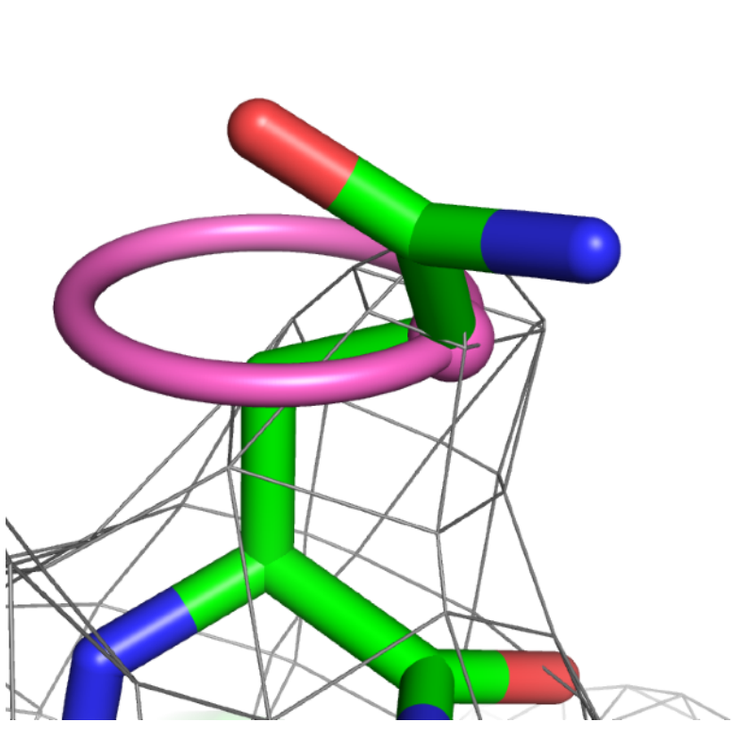
Phenix

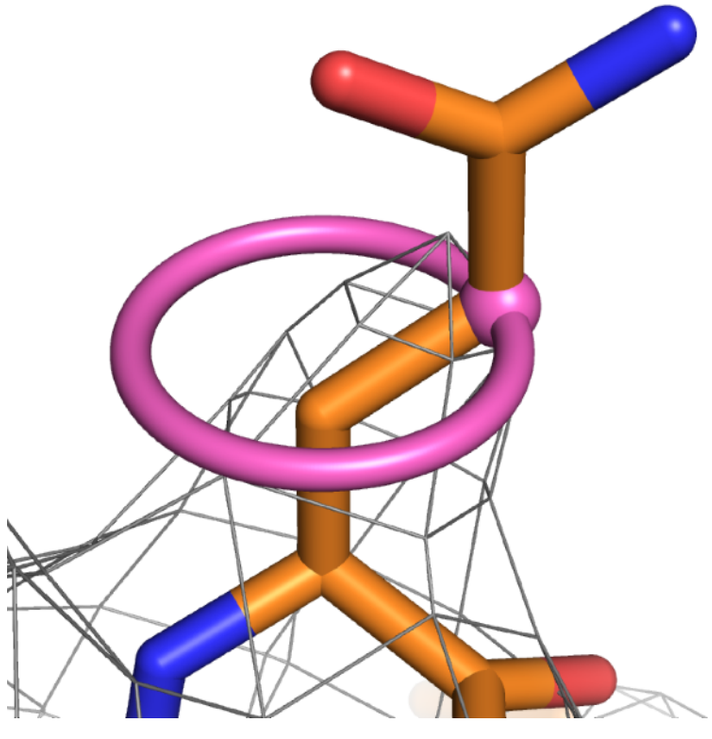
Ben Barad, Jaime Fraser, UCSF

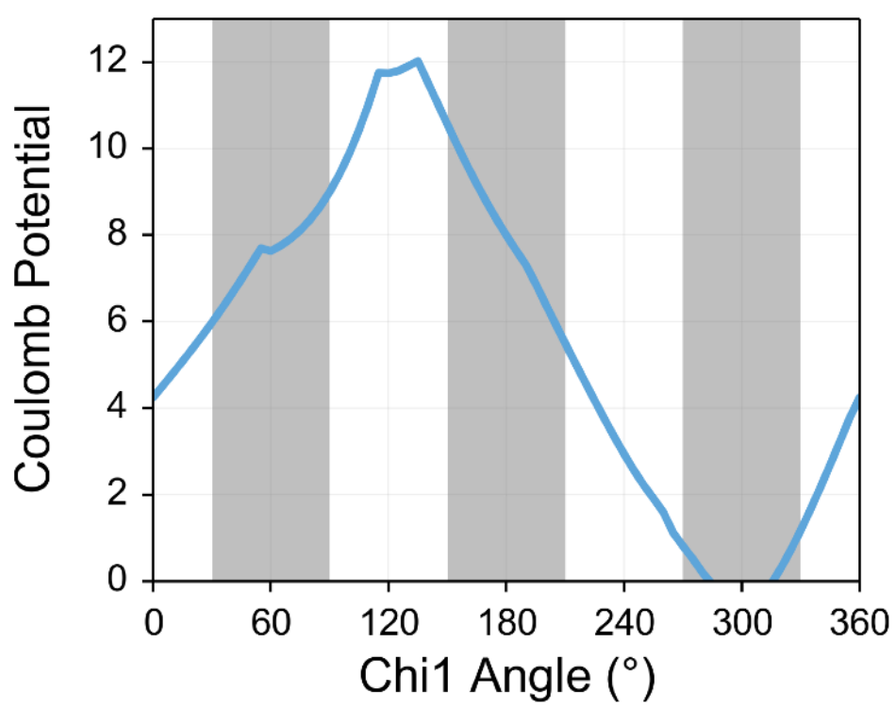
BERKELEY LAB Lawrence Berkeley National Laboratory

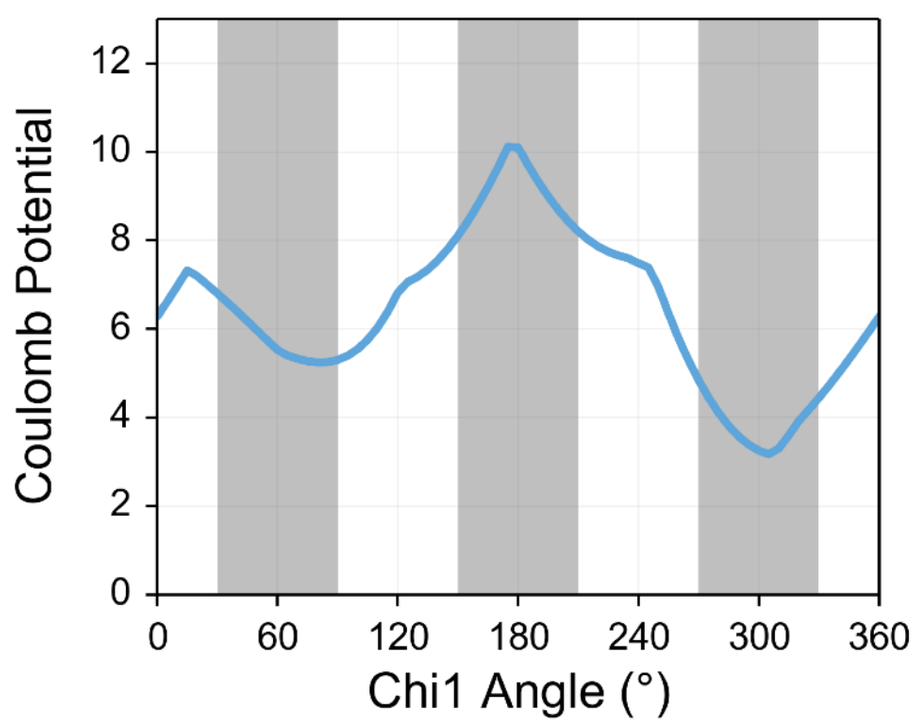
EMRinger reports on backbone placement











UCSF
University of California
San Francisco

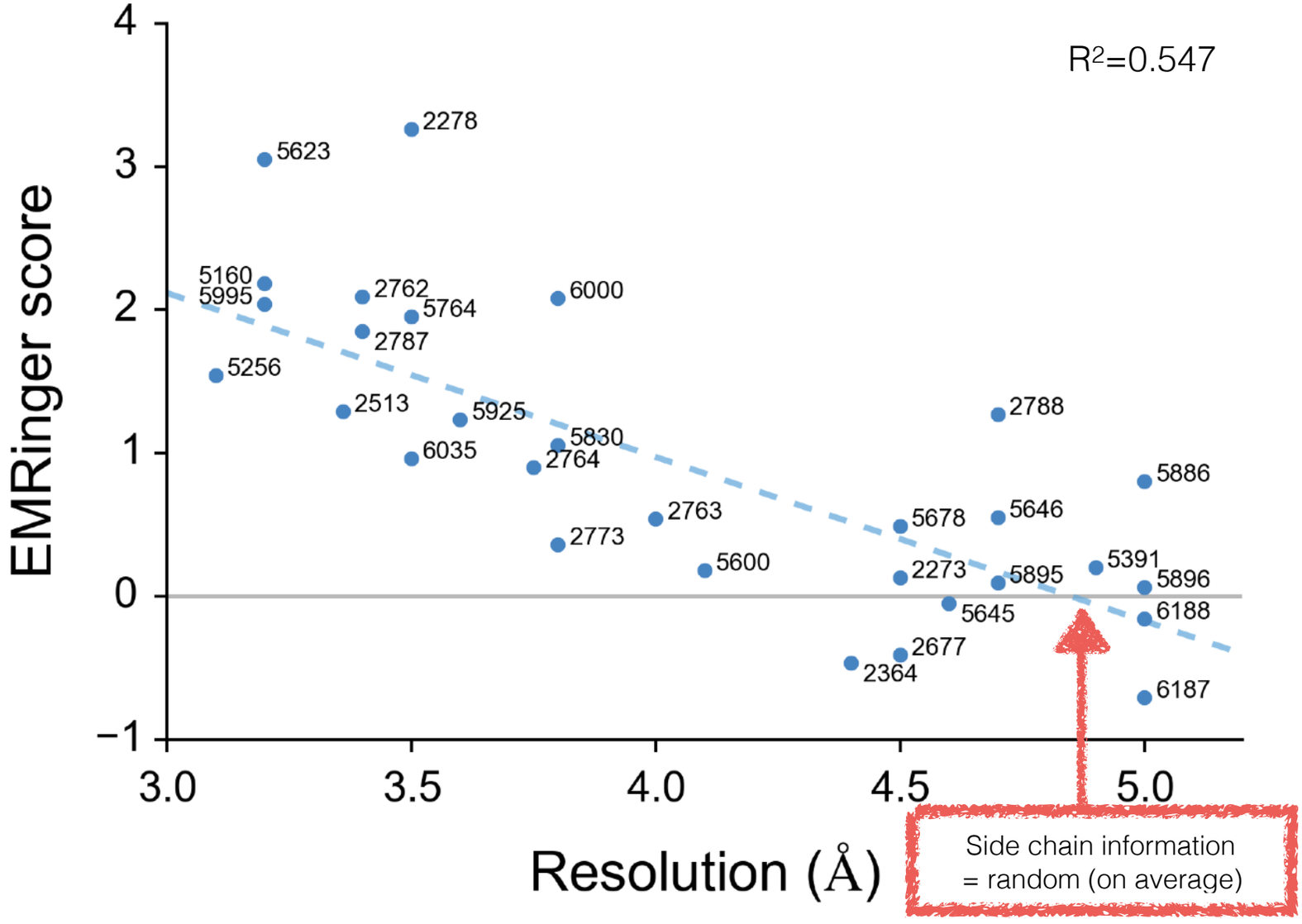
Phenix

Ben Barad, Jaime Fraser, UCSF

BERKELEY LAB
Lawrence Berkeley National Laboratory

EMRinger Score to Validate Model vs Data

- Quantify how well the model backbone puts side chains in places where there are density peaks consistent with rotameric conformations



http://emringer.com

Phenix

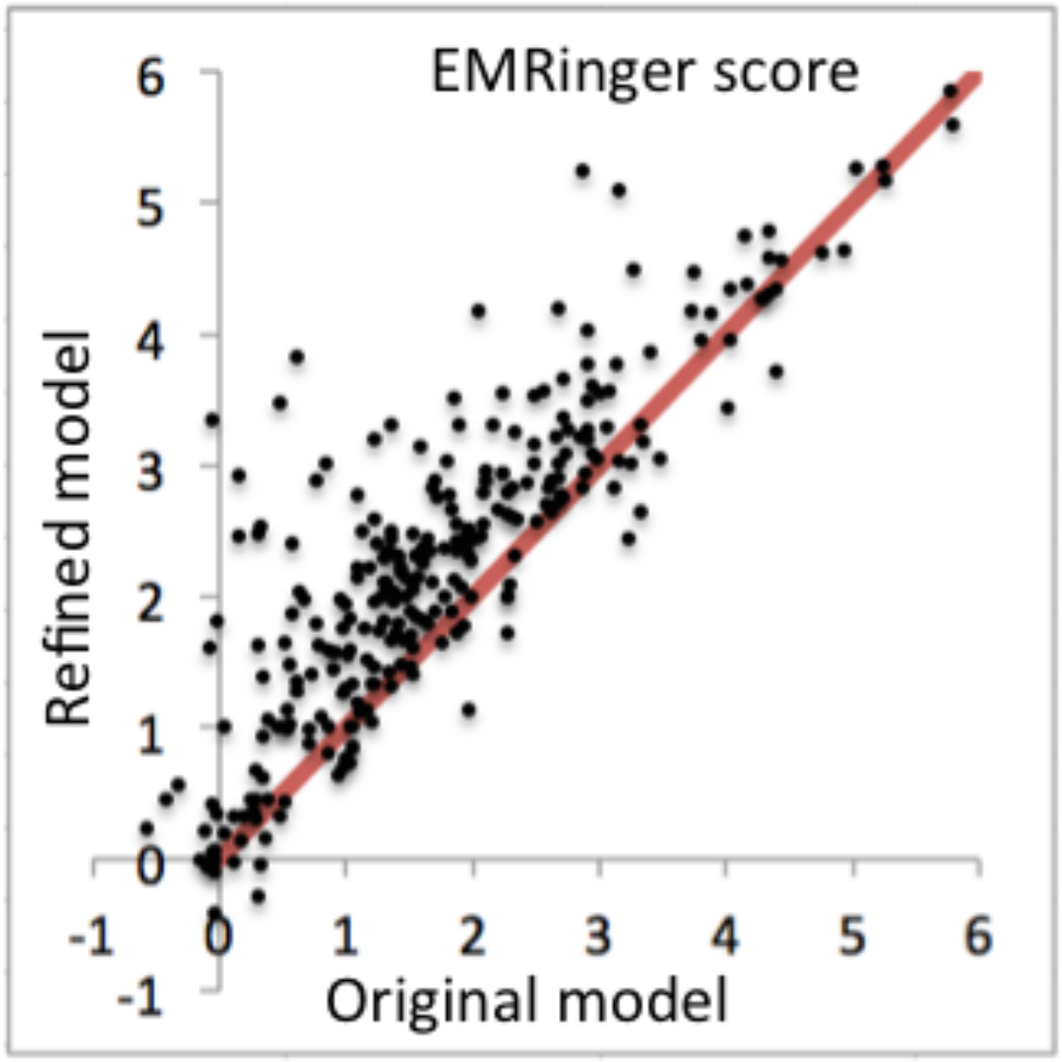
- Available in GUI and command line
- `phenix.emringer model.pdb map.ccp4`

UCSF University of California San Francisco

Phenix *Ben Barad, Jaime Fraser, UCSF*

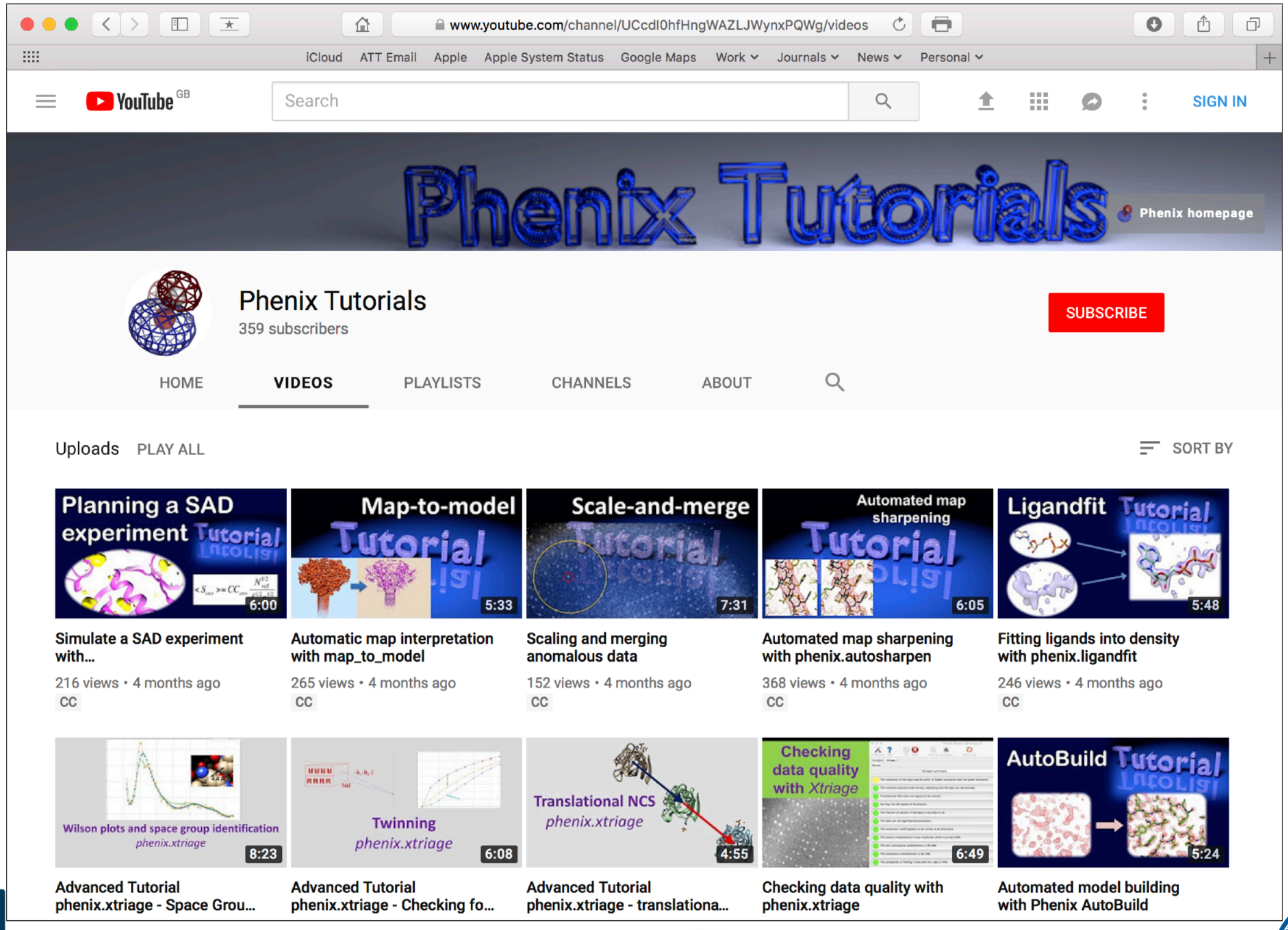
BERKELEY LAB Lawrence Berkeley National Laboratory

Improved Models from Real Space Refinement



BERKELEY LAB
Lawrence Berkeley National Laboratory

Video Tutorials and Lectures



BERKELEY LAB
Lawrence Berkeley National Laboratory

Phenix

Dorothee Liebschner, LBNL

Conclusions

- Automated model building is possible, but can be improved
  - Include information from secondary structure prediction, evolution etc.
  - Combine structure-modeling tools (Rosetta) with Phenix model-building
- The application of prior or complementary information improves refinement of Cryo-EM structures in real space
- Local analysis of side chain positions w.r.t. density provides a metric for model quality
- Many challenges remain:
  - Local variation in resolution leads to uncertainties in interpretation
  - Efficiently accounting for atomic displacements in models
  - Additional validation metrics for the model w.r.t. the data are needed
  - Reliably accounting for uncertainty in magnification

BERKELEY LAB
Lawrence Berkeley National Laboratory

Phenix

Acknowledgments

- **Lawrence Berkeley Laboratory**
  - Pavel Afonine, Youval Dar, Nat Echols, Jeff Headd, Richard Gildea, Ralf Grosse-Kunstleve, Dorothee Liebschner, Nigel Moriarty, Nader Morshed, Billy Poon, Ian Rees, Nicholas Sauter, Oleg Sobolev, Peter Zwart
- **Los Alamos National Laboratory**
  - Tom Terwilliger, Li-Wei Hung
- **Cambridge University**
  - Randy Read, Airlie McCoy, Laurent Storoni, Gabor Bunkoczi, Robert Oeffner
- **Duke University**
  - Jane Richardson & David Richardson, Ian Davis, Vincent Chen, Jeff Headd, Chris Williams, Bryan Arendall, Bradley Hintze, Laura Murray
- **EMRinger**
  - Ben Barad, Yifan Cheng, Jaime Fraser (UCSF)
  - Ray Wang, Frank DiMaio (UW)
- **Rosetta**
  - Frank DiMaio, David Baker (U Washington)
- **Oak Ridge National Laboratory**
  - Marat Mustyakimov, Paul Langan
- **Cryo-EM Structure Refinement**
  - Corey Hryc, Zhao Wang, Matthew Baker, Wah Chiu (Baylor College of Medicine)
  - Adam Frost (UC San Francisco), Gino Cingolani (Jefferson University, PA), Valerie Pye (Cancer Research UK), Darcie Miller (St. Jude Research Hospital), Anastasia Aksyuk (NIH), Mavis Agbandje-McKenna (University of Florida), Stuart Howes (Eva Nogales Lab), Ali Andres Malay (RIKEN, Japan)
- **Other Collaborators**
  - Pawel Janowski, David Case
  - Dale Tronrud, Donnie Berholz, Andy Karplus
  - Alexandre Urzhumtsev & Vladimir Lunin
  - Garib Murshudov & Alexi Vagin
  - Paul Emsley, Bernhard Lohkamp, Kevin Cowtan
  - David Abrahams
  - PHENIX Testers & Users
- **Funding:**
  - NIH/NIGMS:
    - *P01GM063210, P50GM062412, P01GM064692, R01GM071939*
  - PHENIX Industrial Consortium
  - Lawrence Berkeley Laboratory

Thank You