

# Phenix Introduction

*ECM32 Vienna*

*August 2019*

**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 New Mexico Consortium Baylor College of Medicine

Tom Terwilliger, Li-Wei  
Hung, Matt Baker



Randy Read, Airlie McCoy,  
Tristan Croll, Rob Oeffner,  
Kaushik Hatti, Massimo  
Sammito, Duncan Stockwell

## Cambridge University



## Duke University

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

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



An NIH/NIGMS funded  
Program Project

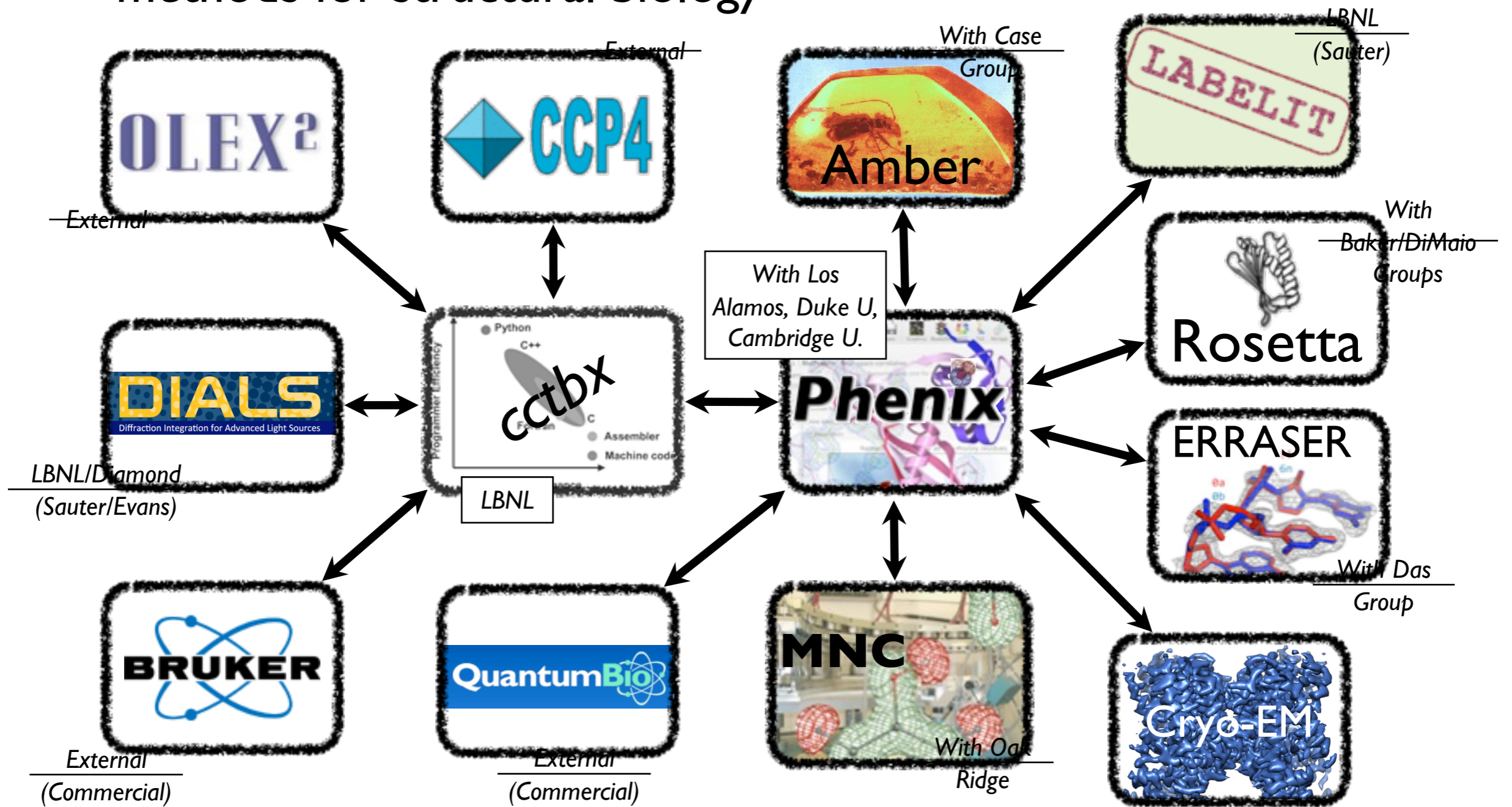


**Phenix**



# Phenix - a Structural Biology Hub

- We have nucleated the development of new computational methods for structural biology



**Phenix**

# What is PHENIX?

- Package for automated structure solution
- Modern programming concepts and new algorithm development
- Designed to be used by both novices and experienced users
- Long-term development and support
- Why is it called PHENIX?
  - Python Hierarchical Environment for Integrated Xtallography

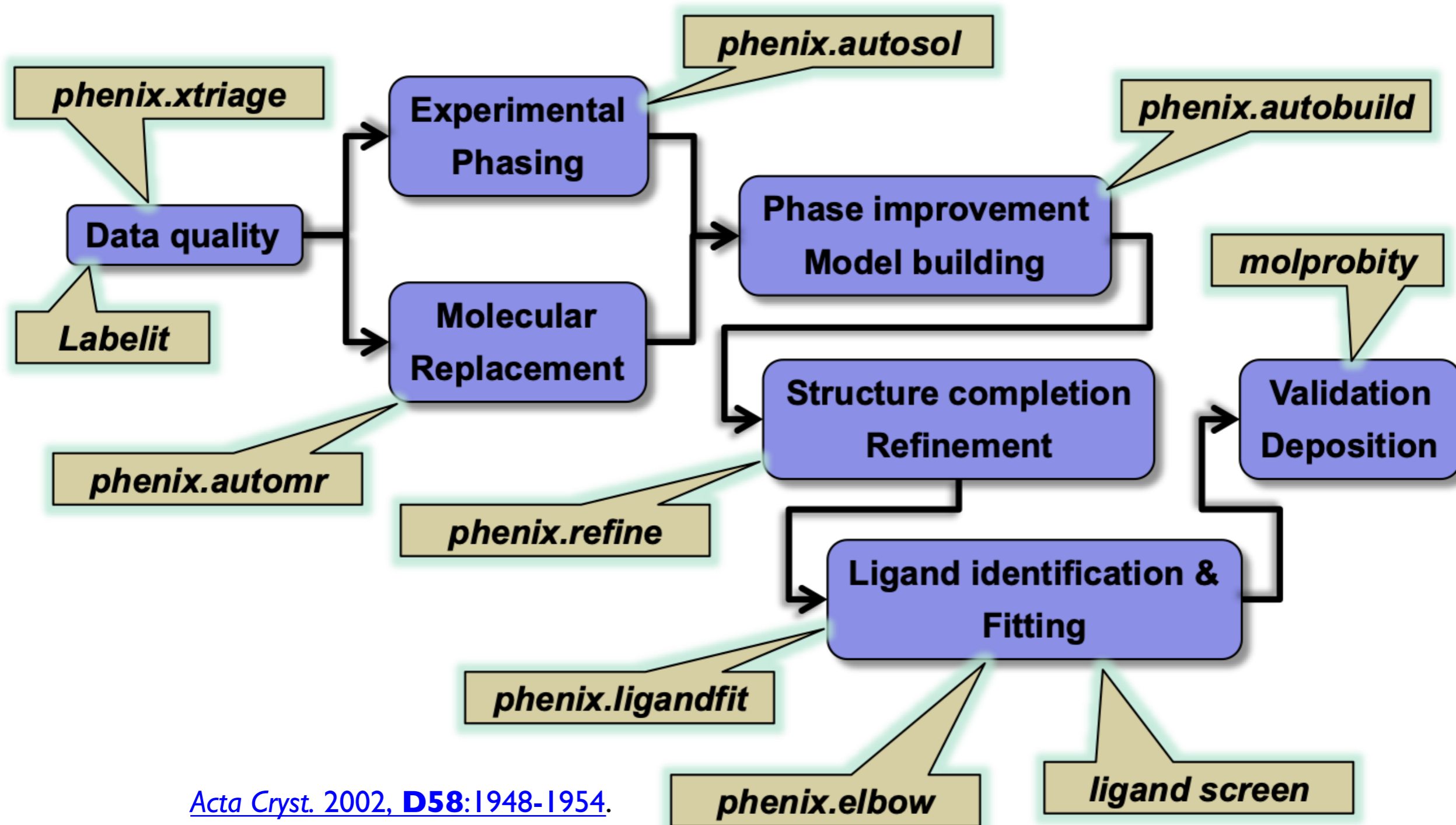


# Key Features

- Python
  - Easy scripting of repetitive tasks
  - Enables rapid prototyping and development
- Advanced algorithms
  - Experimental phasing
  - Molecular replacement
  - Automated model building and rebuilding
  - Structure refinement and validation
  - Ligand coordinate and restraint generation
- Rapid development and bug fixing



# Automation of Structure Solution



[Acta Cryst. 2002, D58:1948-1954.](#)

[J. Appl. Cryst. 2002, 35:126-136.](#)

[Acta Cryst. 2010, D66: 213-221](#)

**Phenix**

# Command Line Tools

- **Data validation**
  - `phenix.xtrriage porin_fp.mtz`
- **Automated structure solution**
  - `phenix.autosol data=peak.sca seq_file=nsf-d2.seq`
- **Automated model building**
  - `phenix.autobuild data=scale.mtz model=mr.pdb  
seq_file=correct.seq`
- **Automated ligand fitting**
  - `phenix.ligandfit data=nsf-d2.mtz model=noligand.pdb  
ligand=atp.pdb`
- **Structure refinement**
  - `phenix.refine nsf-d2.mtz nsf.pdb`
- **Building ligand coordinates and restraints**
  - `phenix.elbow --smiles="C12CC3CC(C2)CC(C1)C3"`

# Phenix GUI

- Central GUI to view job control and launch new jobs

The screenshot displays the Phenix GUI interface. At the top, there is a menu bar with icons for Quit, Preferences, Help, New project, Project settings, Job history, Citations, Reload last job, Coot, PyMOL, and KING. Below the menu bar is a status bar with the text: "Click or drag-and-drop files onto a program to launch it. To switch to a project, click the 'Choose this project' button."

The main interface is divided into two main sections. On the left is a table titled "Projects" with columns for ID, Last modified, # of jobs, and R-free. On the right is a vertical list of tool categories and their descriptions.

ID	Last modified	# of jobs	R-free
✓ 1016B	Aug 17 2010 11:51 ...	19	0.3665
1BOV	Jul 05 2010 02:33 PM	2	None
2hr0	Sep 19 2009 01:16 ...	2	None
3g5u	Nov 14 2009 03:14 ...	1	None
3lna	May 18 2010 06:45 ...	5	0.3901
actin-2010	May 31 2010 06:18 ...	0	None
calmodulin	None	0	None
chennai-a2u-gl ...	Jan 19 2010 01:37 AM	6	0.2547
chennai_p9_sad	Mar 03 2010 07:18 ...	4	0.3157
ggr-2010	Mar 16 2010 03:50 ...	0	None
gui	None	0	None
haemoglobin	Mar 27 2010 09:34 ...	6	0.2977
horseHB	Oct 26 2009 08:57 ...	2	0.5026
insulin	Nov 15 2009 11:33 ...	1	0.2039
jason-kim	Jul 05 2010 05:12 PM	25	0.4706
lysozyme	Oct 24 2009 09:25 ...	3	0.3336
lysozyme-mono ...	Mar 08 2010 12:33 ...	6	0.5638
lysozyme_test	Nov 15 2009 11:11 ...	1	0.2329
mmcpn-compari ...	Aug 03 2010 04:15 ...	3	None
mmcpn_dec_2009	Dec 12 2009 09:27 ...	9	0.2624
nsf-ligand	Mar 06 2010 08:56 ...	1	0.3002
p9	Nov 26 2009 11:16 ...	11	0.2612
p9-sad_pdadams	Jun 15 2010 09:58 AM	3	0.2473
porin-twin	Apr 30 2010 11:32 ...	1	None
porin-twin_pdad	Jun 09 2010 03:45 PM	3	0.1481

Below the table are buttons for "Switch project" and "Delete project".

The right-hand side of the interface lists various tools under different categories:

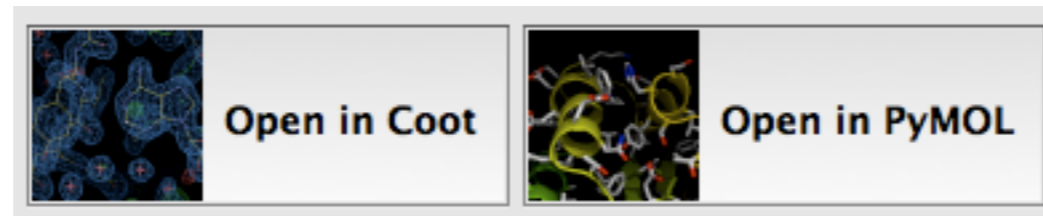
- Reflection tools**
- Model tools**
- Experimental phasing**
  - AutoSol**: Automated experimental phasing with model-building
  - Phaser-EP**: Maximum-likelihood SAD experimental phasing
  - Density modification**: Simple interface for running density modification only using AutoBuild and RESOLVE
- Molecular replacement**
- Model building**
  - AutoBuild**: Automated model-building and refinement
  - Find Helices and Strands**: Fast chain tracing
  - Fit Loops**: Fast placement of missing loops in electron density
- Refinement**
  - phenix.refine**: Automated X-ray and neutron refinement
  - ReadySet**: Utility for preparing PDB files for refinement – automatically generate restraints and add hydrogens
- Maps**

At the bottom of the window, there is an "Output directory:" field with the path "/Users/pdadams/Work/Structures/1016B" and a "Browse..." button. The status bar at the very bottom shows "PHENIX version 1.6.4-486" and "Project: 1016B".

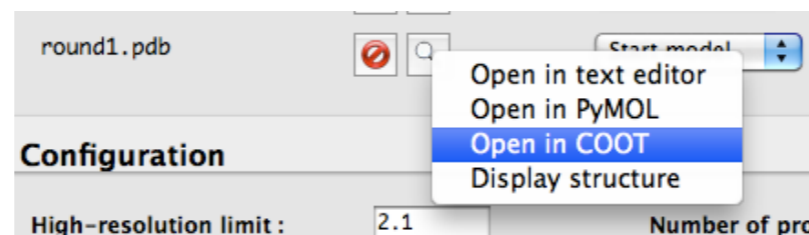


# Coot/PyMOL integration

- Most results can be opened directly in graphics apps



- Any PDB file listed in GUI can also be opened



- AutoSol, AutoBuild, and phenix.refine will update Coot continuously while running
- Coot must have Python support (default on Mac)
- Specific paths to executables usually required on Linux

- *Preferences->Graphics->Full path to Coot [...PyMOL]*

**Phenix**

# Recent Developments

- Automated structure solution with weak anomalous data
- Translational NCS corrections in SAD phasing

- MR\_Rosetta and morphing for rescuing poor MR solutions
- Translational NCS corrections in MR

- New Rosetta methods for RNA rebuilding (ERRASER)

- Automated ion placement in refinement
- Automated ligand pipeline

- DIALS included in Phenix distribution

- Support for mmCIF format (for structure deposition)
- Video Tutorials

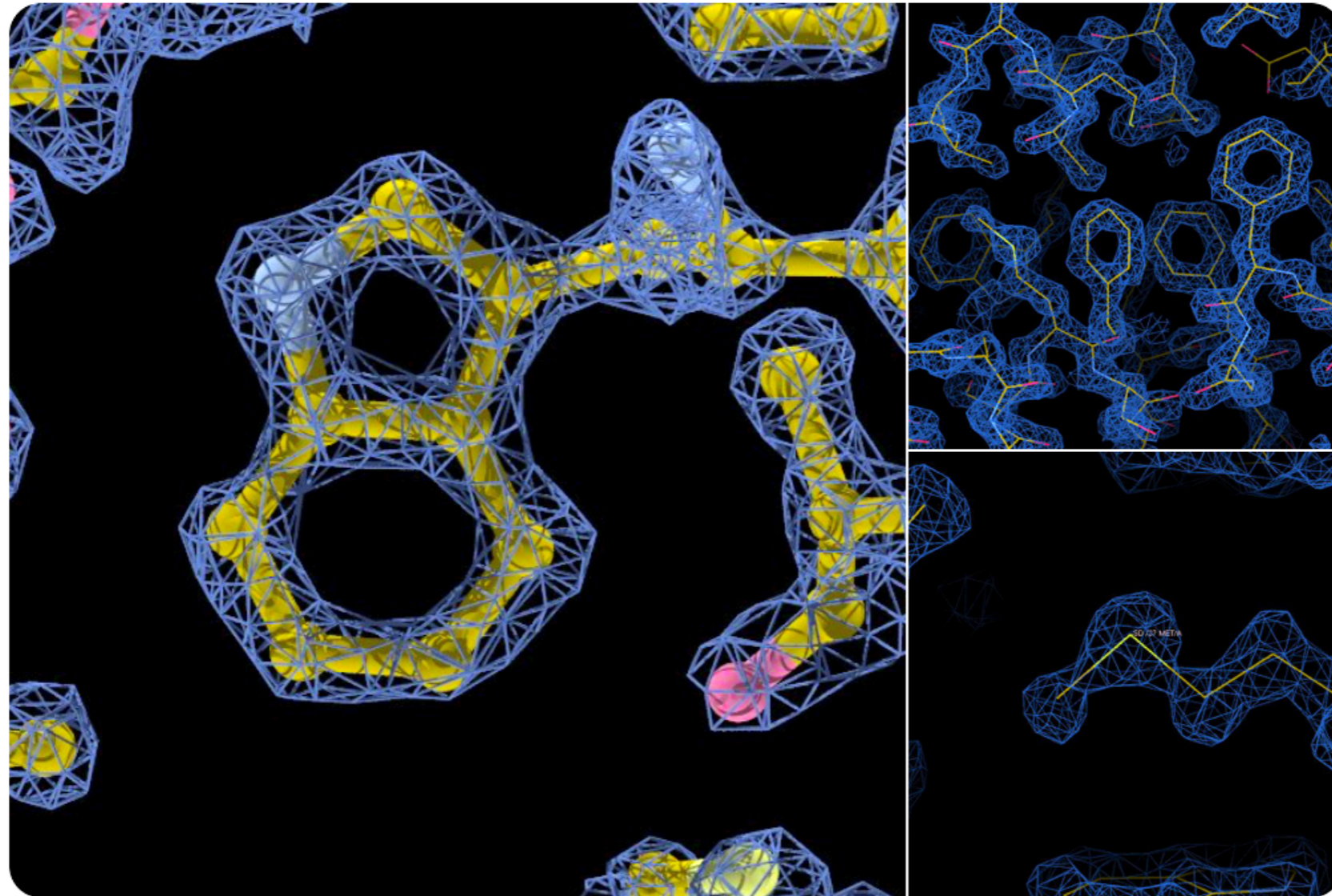
- Low resolution refinement algorithms:
  - Rosetta refinement
  - Reference model
  - Torsion NCS
  - Structure restraints

- New model validation metrics

- New hydrogen validation tools
- New riding hydrogen model

- Feature Enhanced Maps to improve interpretability
- Polder maps for better ligand density
- Ensemble refinement to understand dynamics and structural variability

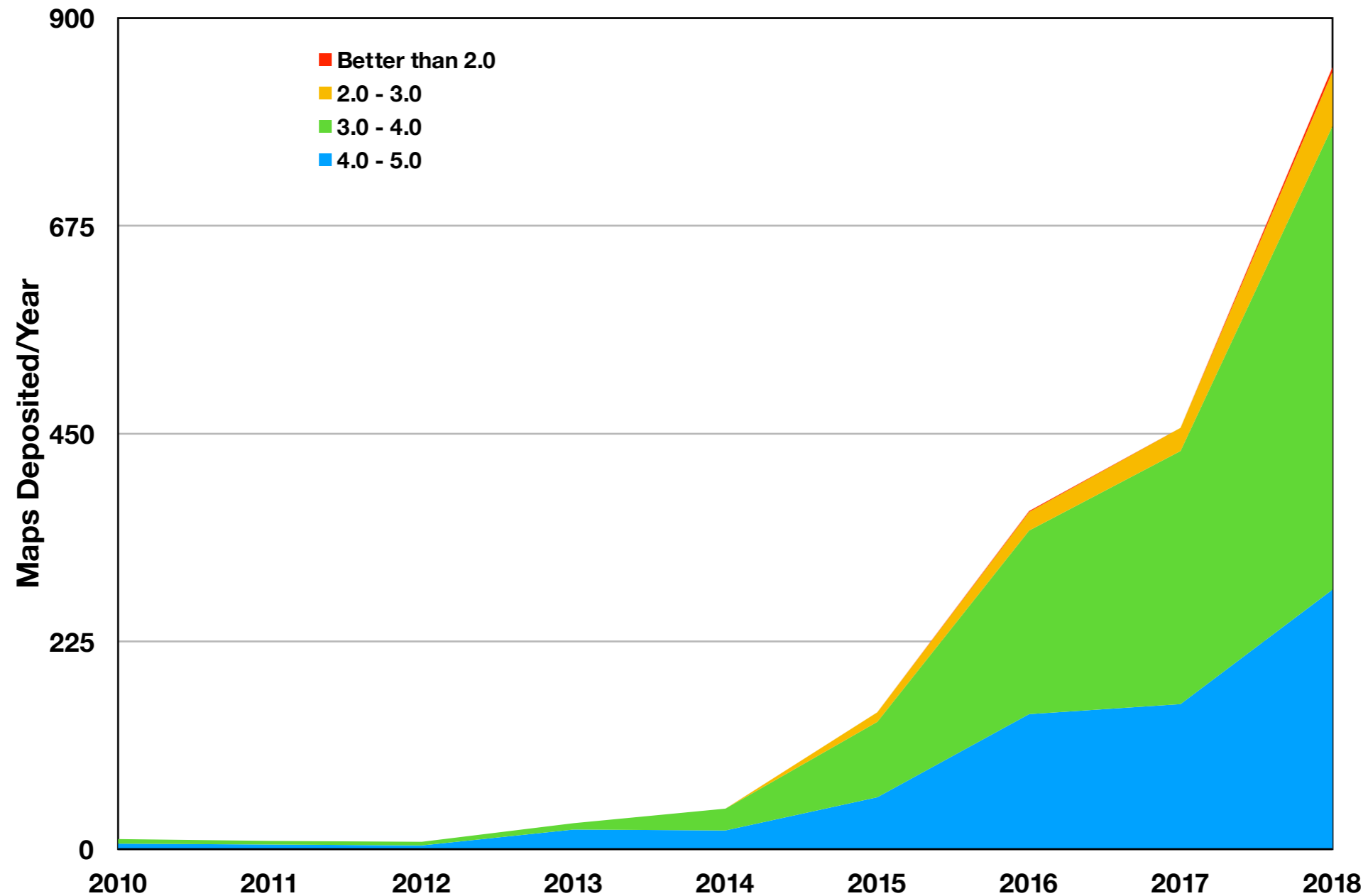
# Impressive Cryo-EM Achievements



*Namba Lab, Osaka*

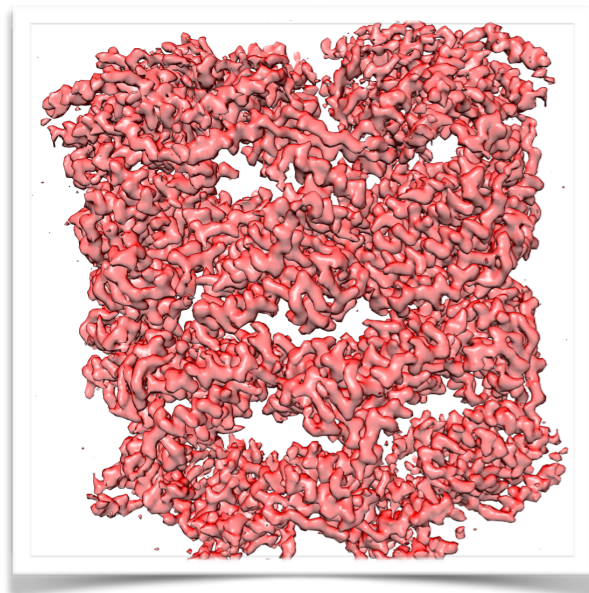
# Map Resolution

- Biggest growth is in the 3-4Å range
- Substantial number of maps in 4-5Å range

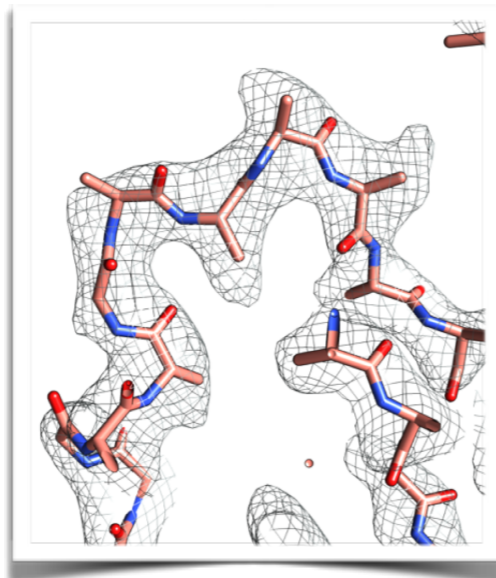


\* Not all maps have an associated model

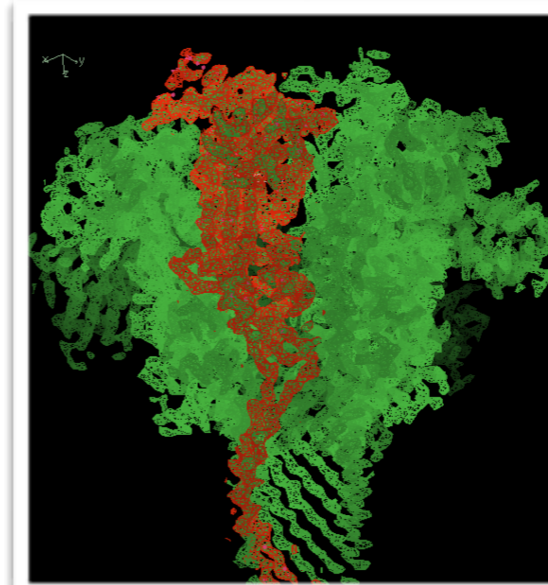
# New Tools for Cryo-EM in Phenix



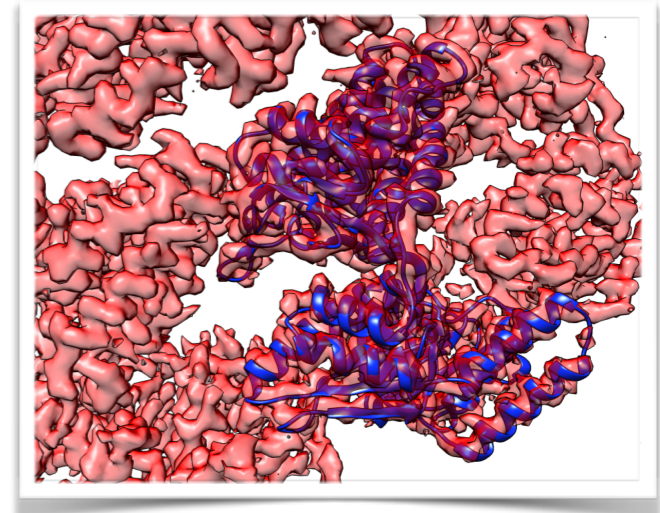
*Symmetry from a map*



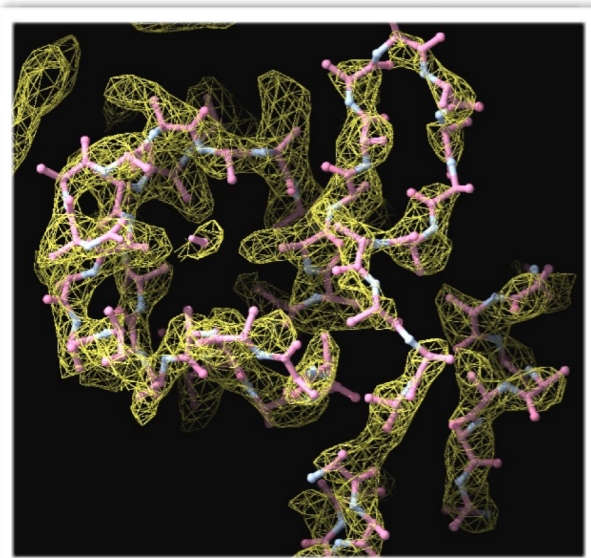
*Automated map sharpening*



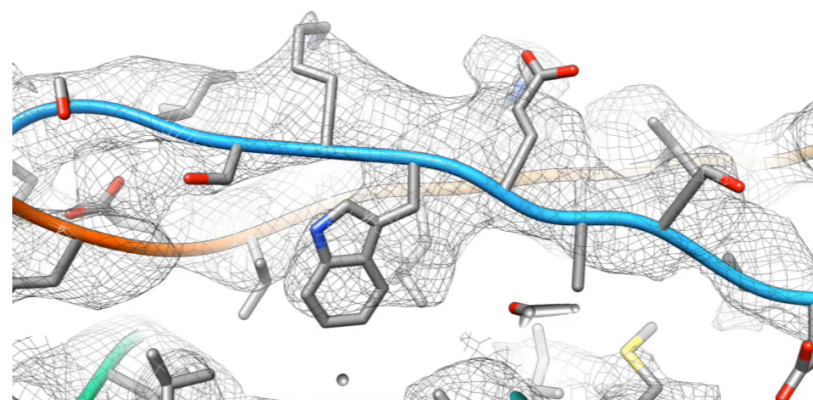
*Map segmentation*



*Rigid model docking*



*Automated model building*



*Real space refinement*

Model		Ramachandran	
MolProbity score	1.72	Outliers (%)	0.00 (Goal: < 0.2%)
Clash score	5.44	Allowed (%)	6.45
Rotamer outliers (%)	0.00 (Goal: < 1%)	Favored (%)	93.55 (Goal: > 98%)
CB outliers	0 (Goal: 0)		

CaBLAM		Peptide Plane	
Outliers (%)	3.88 (Goal: <= 1%)	cis-proline (%)	0.00
Disfavored (%)	8.96 (Goal: <= 5%)	twisted proline (%)	0.00
Ca outliers (%)	1.19 (Goal: <= 0.5%)	cis-general (%)	0.00
		twisted general (%)	0.00

*Model and map validation*

# Challenges

- Low resolution structure solution and refinement
- Structure completion
  - Automated identification, fitting and refinement of ligands, metals, ions, and water
  - Identification, fitting and refinement of discrete disorder (multiple conformations)
  - Representing other forms of disorder
- Automated parameterization of models in refinement
  - ADPs, TLS groups, NCS, hydrogens
- Handling different kinds of twinning and integrating it into the whole crystallographic process
- Automated understanding of chemistry

# PHENIX Availability

- phenix-online.org
- Supported on:
  - Linux (RedHat, Fedora)
  - Mac OSX
  - Windows
- Extensive documentation
- Nightly builds:

**Phenix** Python-based Hierarchical ENvironment for Integrated Xtallography

Documentation: [\(HTML\)](#) [FAQ](#)  [What's New](#)

PHENIX is a software suite for the automated determination of molecular structures using X-ray crystallography and other methods.

**Download the latest official release (1.15.2) [First request download password]**

**Download the latest nightly build or prerelease**

Help: [Email Us](#) [Report a Bug](#) [List Archives](#) [Subscribe to Mailing List](#)

**Using PHENIX (release 1.15.2):**

- The PHENIX Graphical User Interface
- Assessing data quality with phenix.xtriage
- Automated structure solution with AutoSol
- Automated molecular replacement with Phaser-MR
- Automated model building and rebuilding with AutoBuild
- Automated ligand fitting with LigandFit
- Structure refinement with phenix.refine
- Generation of ligand coordinates and restraints with eLBOW

The PHENIX system also includes SOLVE/RESOLVE, Phaser, the CCI Applications (phenix.xtriage, phenix.refine, elbow and many more), components from MolProbity, and the Computational Crystallography Toolbox in a Python framework.

**Citing PHENIX:**  
 PHENIX: a comprehensive Python-based system for macromolecular structure solution. P. D. Adams, P. V. Afonine, G. Bunkóczy, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L.-W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, A. J. McCoy, N. W. Moriarty, R. Oeffner, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger and P. H. Zwart. *Acta Cryst.* **D66**, 213-221 (2010).

**Funding for PHENIX:** [NIH General Medical Sciences](#)

The PHENIX system also can be used for neutron crystallography. NIH funding supports the development of this capability

Version	Date	Status	Logs	Info
<a href="#">dev-3512</a>	2019-05-21	successful	<a href="#">ci</a> , <a href="#">intel-windows-x86_64</a> , <a href="#">intel-linux-2.6-x86_64-centos6</a> , <a href="#">mac-intel-osx-x86_64</a>	<a href="#">docs</a> ; <a href="#">changelog</a>
<a href="#">dev-3500</a>	2019-05-09	successful	<a href="#">ci</a> , <a href="#">intel-linux-2.6-x86_64-centos6</a> , <a href="#">mac-intel-osx-x86_64</a> , <a href="#">intel-windows-x86_64</a>	<a href="#">docs</a> ; <a href="#">changelog</a>
<a href="#">dev-3494</a>	2019-05-03	successful	<a href="#">intel-linux-2.6-x86_64-centos6</a> , <a href="#">ci</a> , <a href="#">mac-intel-osx-x86_64</a> , <a href="#">intel-windows-x86_64</a>	<a href="#">docs</a> ; <a href="#">changelog</a>
<a href="#">1.15.2-3472</a>	2019-04-11	successful	<a href="#">ci</a> , <a href="#">intel-windows-x86_64</a> , <a href="#">intel-linux-2.6-x86_64-centos6</a> , <a href="#">mac-intel-osx-x86_64</a>	Official 1.15.2 release; <a href="#">docs</a> ; <a href="#">changelog</a>



# Video Tutorials

The screenshot shows the YouTube channel page for 'Phenix Tutorials'. The channel has 560 subscribers and a 'SUBSCRIBE 560' button. The video grid includes:

- real\_space\_refine Tutorial** (5:27): How to run real-space-refine (763 views, 7 months ago)
- Secondary Structure Restraints Tutorial** (6:23): How to use secondary structure restraints (399 views, 8 months ago)
- Multiple refinement strategies Tutorial** (5:46): How to use multiple refinement strategies and... (281 views, 8 months ago)
- Planning a SAD experiment Tutorial** (6:00): Simulate a SAD experiment with... (483 views, 1 year ago)
- Map-to-model Tutorial** (5:33): Automatic map interpretation with map\_to\_model (1.3K views, 1 year ago)
- Scale-and-merge Tutorial** (7:31): Scaling and merging anomalous data (387 views, 1 year ago)
- Automated map sharpening Tutorial** (6:05)
- Ligandfit Tutorial** (5:48)
- Wilson plots and space group identification phenix.xtriage** (8:23)
- Twinning phenix.xtriage** (6:08)
- Translational NCS phenix.xtriage** (4:55)
- Checking data quality with Xtriage** (6:49)

**Dorothee Liebschner, Nigel**

**Moriarty, Miffy Mifsud, Tom Terwilliger**

**Phenix**





# Acknowledgements

## **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

## **Los Alamos Laboratory/New Mexico Consortium**

Tom Terwilliger, Li-Wei Hung

## **Baylor College of Medicine**

Matt Baker

## **Cambridge University**

Randy Read, Airlie McCoy, Gabor Bunckozi, Tristan Croll, Rob Oeffner, Kaushik Hatti, Massimo Sammito, Duncan Stockwell, Laurent Storoni

## **Duke University**

Jane Richardson & David Richardson, Ian Davis, Vincent Chen, Jeff Headd, Chris Williams, Bryan Arendall, Bradley Hintze, Laura Murray

## **UC San Francisco**

Ben Barad, Yifan Cheng, Jaime Fraser

## **University of Washington**

Frank DiMaio, Ray Wang, David Baker

## **Oak Ridge National Laboratory**

Marat Mustyakimov, Paul Langan

## **Other Collaborators**

Corey Hryc, Zhao Wang, Wah Chiu  
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



# Tutorials

- Cryo-EM model placement and building
  - Symmetry determination
  - Rigid body model fitting
  - Map sharpening
  - Map segmentation
  - Automated model building
- Atomic model optimization and validation
  - Structure refinement (crystallography and cryo-EM)
  - Validation

# Tutorial Format

- Use graphical user interface

The screenshot displays the PHENIX graphical user interface. At the top, there is a menu bar with icons for Quit, Preferences, Help, Citations, Reload last job, Coot, PyMOL, KiNG, Other tools, and Ask for help. Below the menu bar, there are tabs for 'Actions' and 'Job history'. The main area is divided into two panels. The left panel, titled 'Projects', shows a table of project entries with columns for ID, Last modified, # of jobs, and R-free. The right panel, titled 'Favorites', lists various tools and actions such as Data analysis, Experimental phasing, Molecular replacement, Model building, Refinement, Cryo-EM, Mtriage, Map to Model, Real-space refinement, Comprehensive validation (cryo-EM), EMRinger, Autosharpen Map, Dock in map, and Sequence From Man. At the bottom, there is a 'Current directory' field showing the path /Users/PDAdams/Work/Structures/apo-ferritin-cshl-2019 and a 'Browse...' button. The status bar at the very bottom indicates 'PHENIX version 1.15.2-3472-000' and 'Project: apo-ferritin-cshl-2019'.

ID	Last modified	# of jobs	R-free
brink	May 17 2019 01:0...	5	0.3868
sec17-sad_0	May 02 2019 06:2...	3	0.3684
nisha2	Apr 29 2019 11:54...	57	0.5146
nisha	Apr 28 2019 07:21...	67	0.4645
toxd-mr_0	Apr 26 2019 10:53...	6	0.4918
p9-sad_0	Apr 25 2019 10:26...	1	0.2898
✓ apo-ferritin-csh ...	Apr 25 2019 09:43...	26	---
apo-ferritin_0	Apr 17 2019 07:50...	9	---
rnase-s_1	Apr 17 2019 03:08...	1	0.2644
apo-ferritin-test	Mar 24 2019 02:36...	7	---
real-space-refin...	Mar 23 2019 09:25...	2	---
rotavirus-model...	Mar 23 2019 08:22...	1	---
rotavirus-autosh...	Mar 23 2019 07:54...	1	---
groel_dock_refin...	Mar 23 2019 07:47...	3	---
apo-ferritin	Mar 23 2019 07:14...	29	---
rotavirus-autosh...	Mar 23 2019 03:37...	1	---
real-space-refin...	Mar 18 2019 11:45...	1	---
rotavirus-model...	Mar 18 2019 11:28...	1	---
groel_dock_refin...	Mar 18 2019 11:13...	2	---
real-space-refin...	Mar 18 2019 08:06...	1	---
groel dock refin...	Mar 18 2019 07:58...	6	---

# Tutorial Format

- Use tutorial datasets distributed with Phenix
- Should run on most laptops (2GB RAM, multiple CPUs better)

