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

Cambridge University

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

Duke University

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

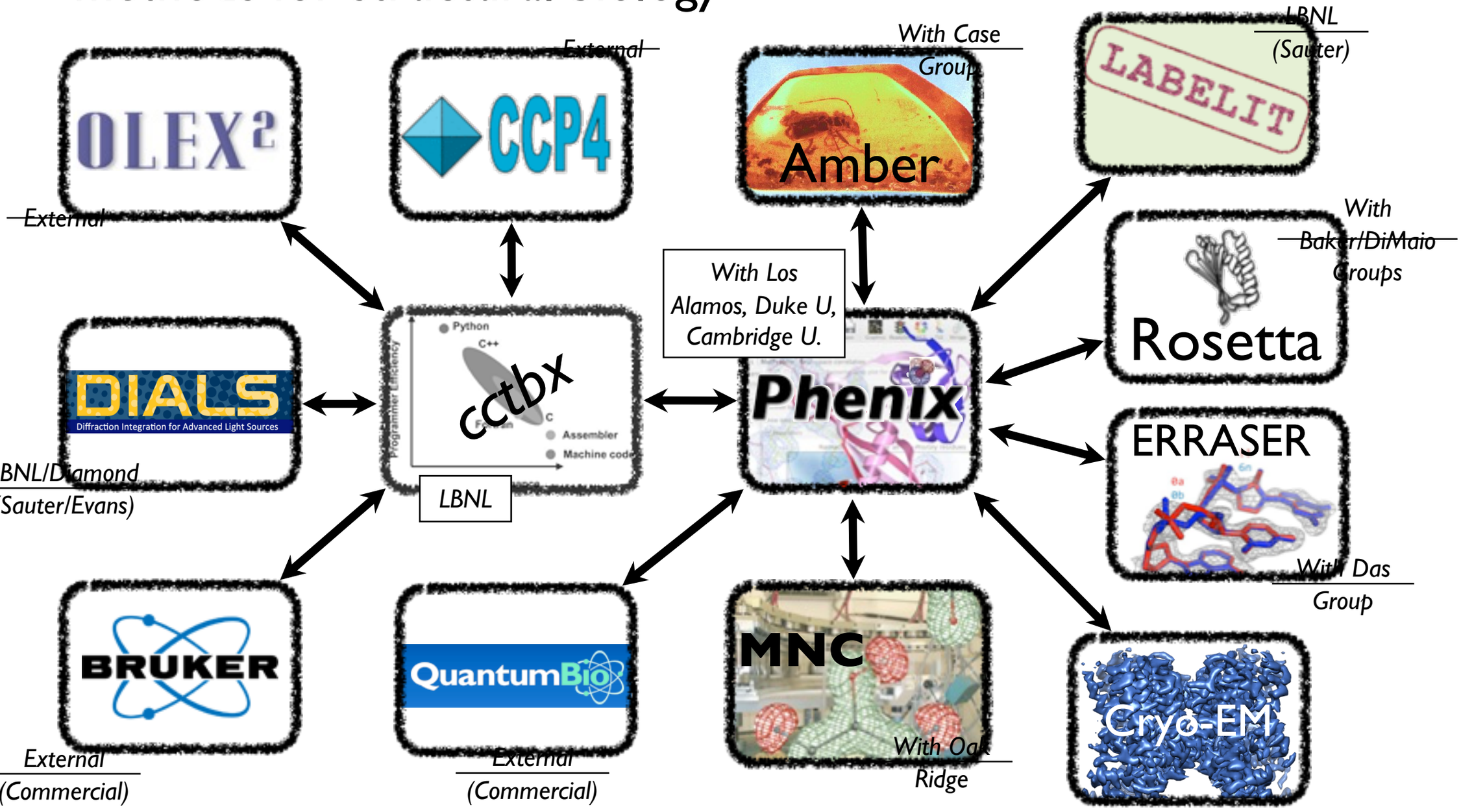
An NIH/NIGMS funded Program Project

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

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



BERKELEY LAB
Lawrence Berkeley National Laboratory

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

BERKELEY LAB
Lawrence Berkeley National Laboratory

Phenix

Key Features

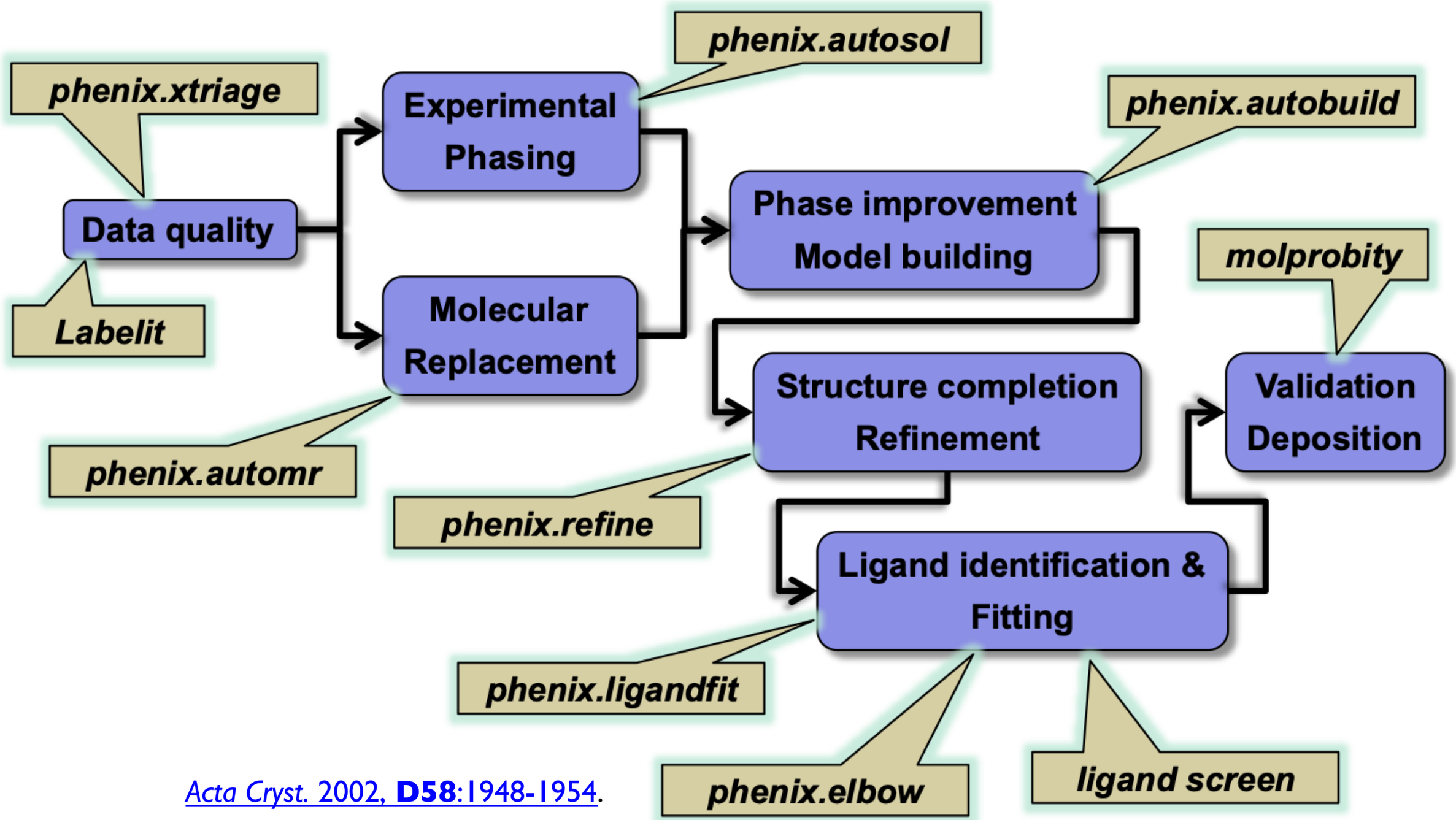
- Python
  - Easy scripting of repetitive tasks
  - Enables rapid prototyping and development
- Advanced algorithms
  - Experimental phasing
  - Molecular replacement
  - Automated model building and rebuilding
  - Structure refinement and validation
  - Ligand coordinate and restraint generation
- Rapid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

BERKELEY LAB
Lawrence Berkeley National Laboratory

Phenix

THE UNIVERSITY OF CALIFORNIA BERKELEY 1868

Automation of Structure Solution



Acta Cryst. 2002, **D58**:1948-1954.

J. Appl. Cryst. 2002, **35**:126-136.

Acta Cryst. 2010, **D66**: 213-221

BERKELEY LAB
Lawrence Berkeley National Laboratory

Phenix

Command Line Tools

- Data validation
  - `phenix.xtriage 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"`

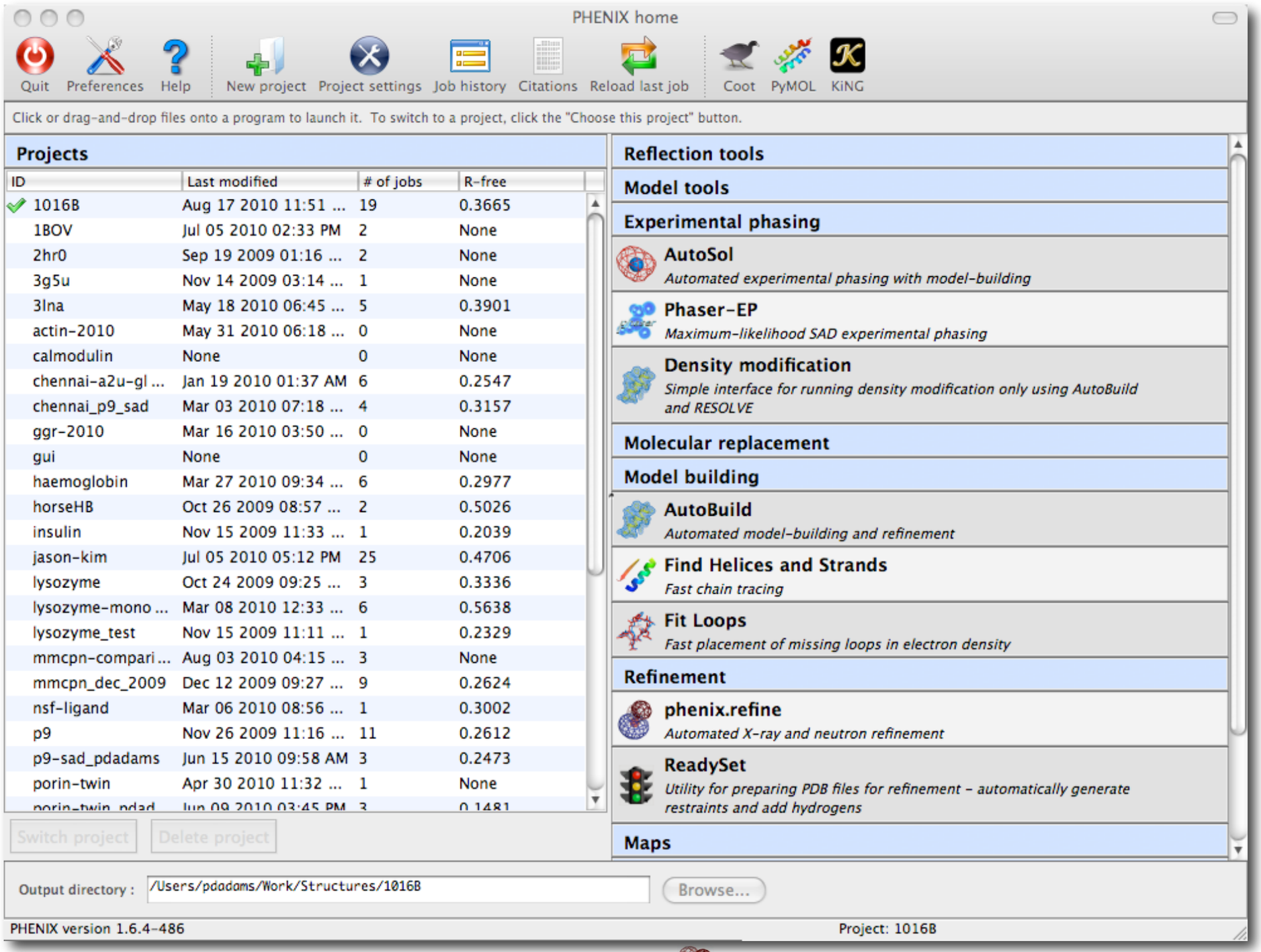
BERKELEY LAB
Lawrence Berkeley National Laboratory

Phenix

THE UNIVERSITY OF CALIFORNIA BERKELEY 1868

Phenix GUI

- Central GUI to view job control and launch new jobs



BERKELEY LAB
Lawrence Berkeley National Laboratory

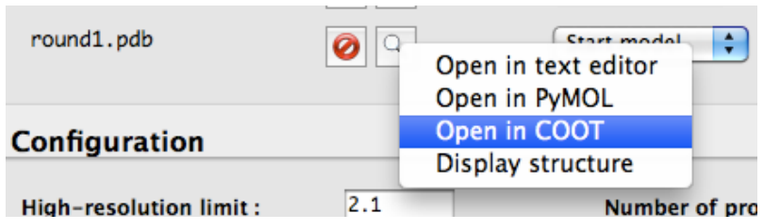
Phenix

GUI: Nat Echols (LBL)

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]*

BERKELEY LAB
Lawrence Berkeley National Laboratory

Phenix

THE UNIVERSITY OF CALIFORNIA BERKELEY 1868

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

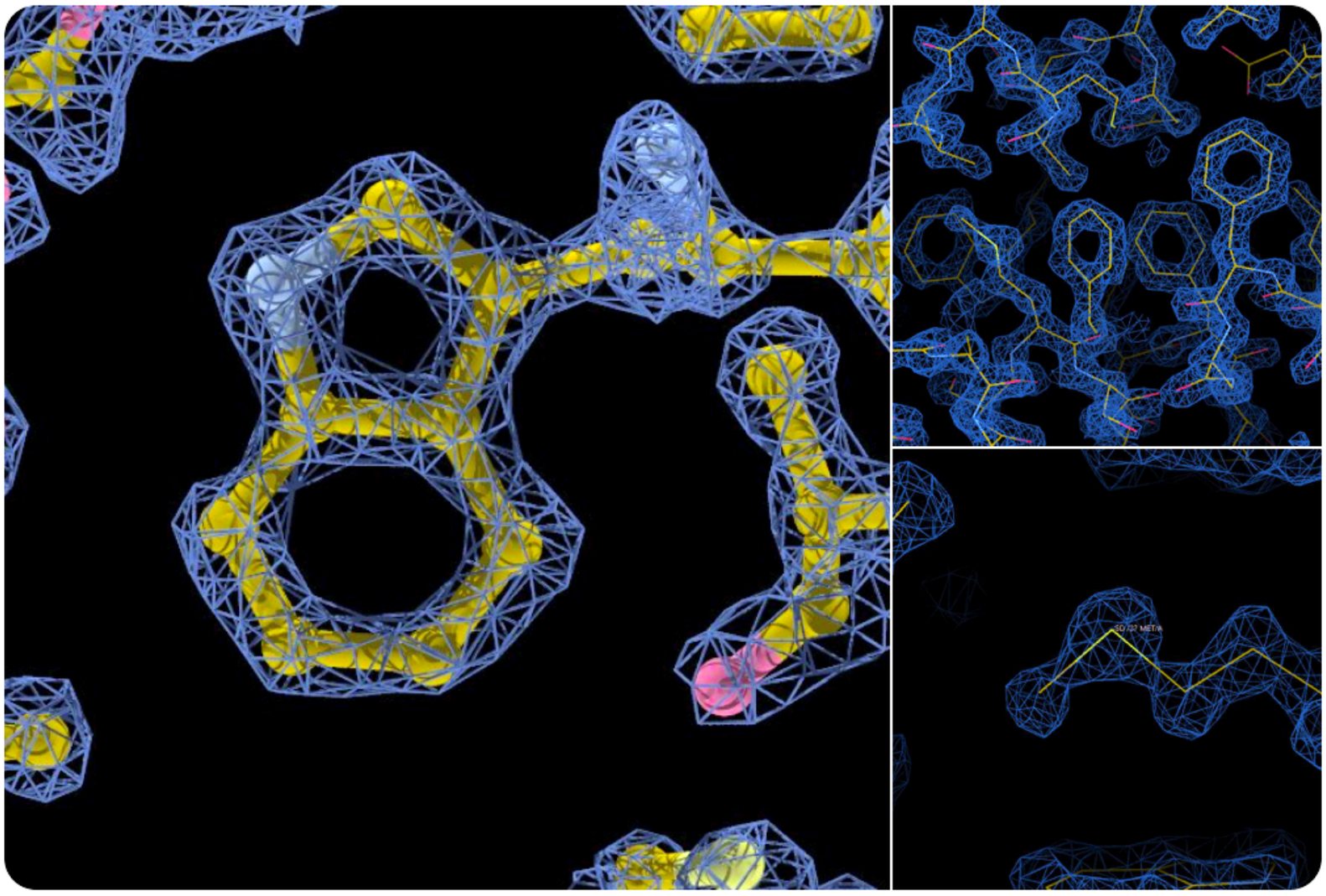
- Feature Enhanced Maps to improve interpretability
- Polder maps for better ligand density
- Ensemble refinement to understand dynamics and structural variability

BERKELEY LAB
Lawrence Berkeley National Laboratory

Phenix

THE UNIVERSITY OF CALIFORNIA BERKELEY 1868

Impressive Cryo-EM Achievements



Namba Lab, Osaka

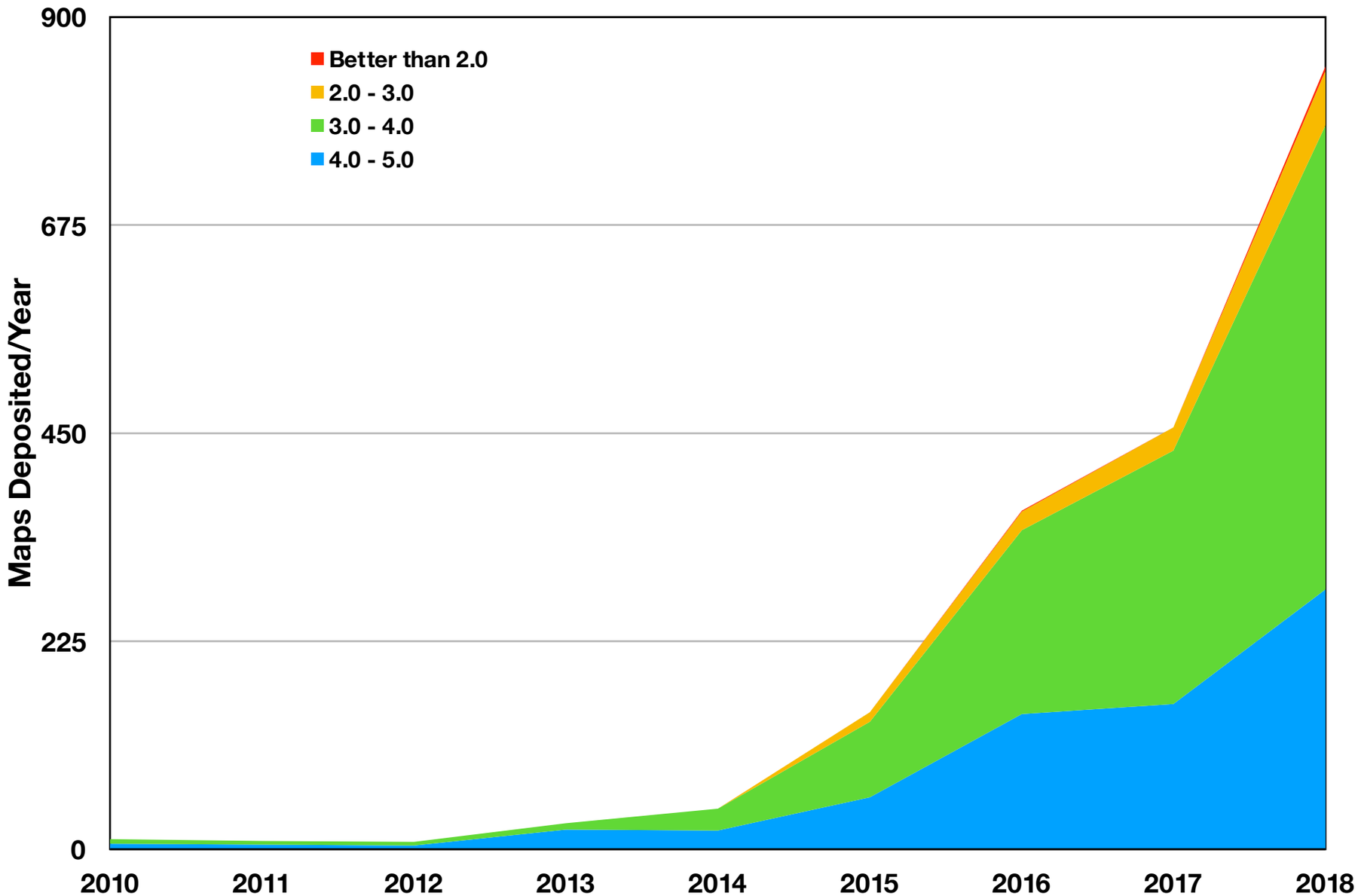
BERKELEY LAB
Lawrence Berkeley National Laboratory

Phenix

THE UNIVERSITY OF CALIFORNIA BERKELEY 1868

Map Resolution

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

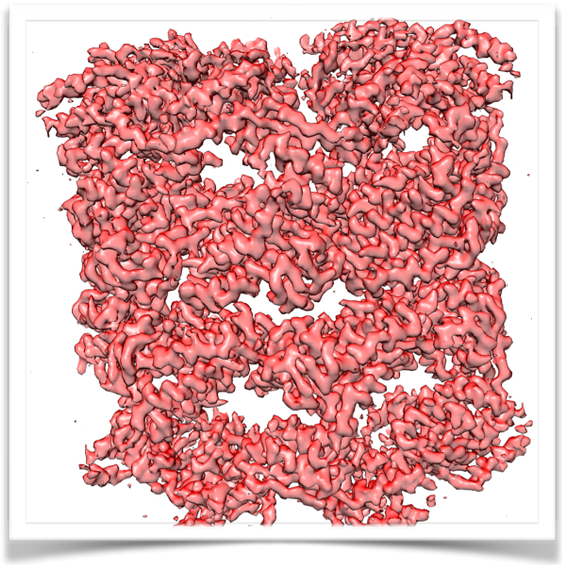


BERKELEY LAB
Lawrence Berkeley National Laboratory

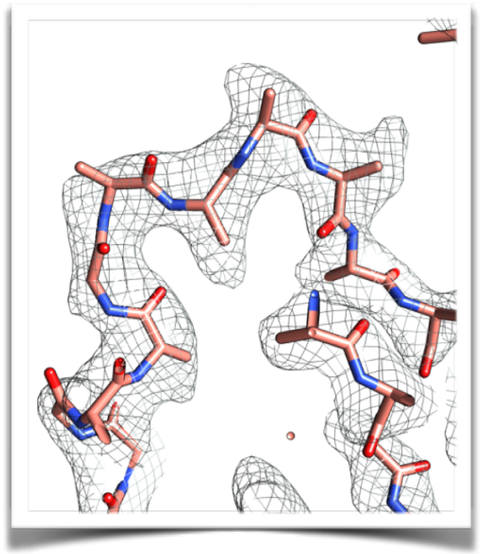
Phenix

** Not all maps have an associated model*

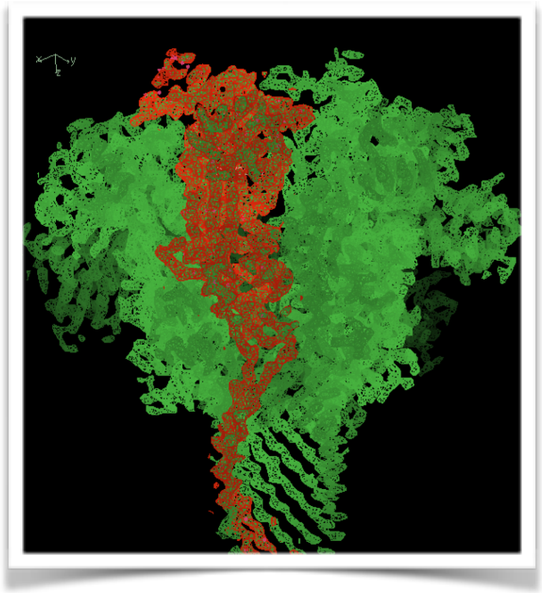
New Tools for Cryo-EM in Phenix



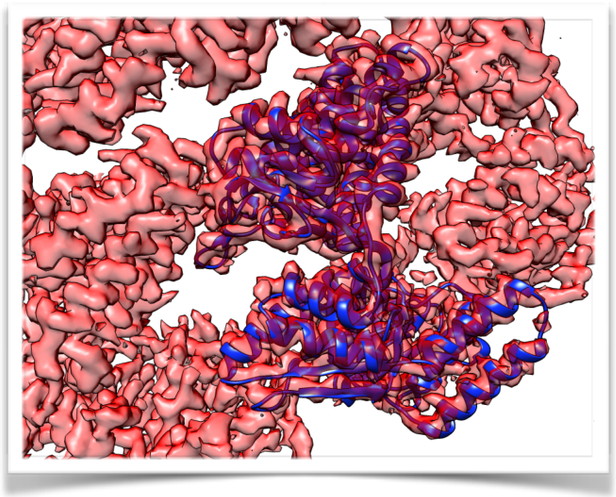
Symmetry from a map



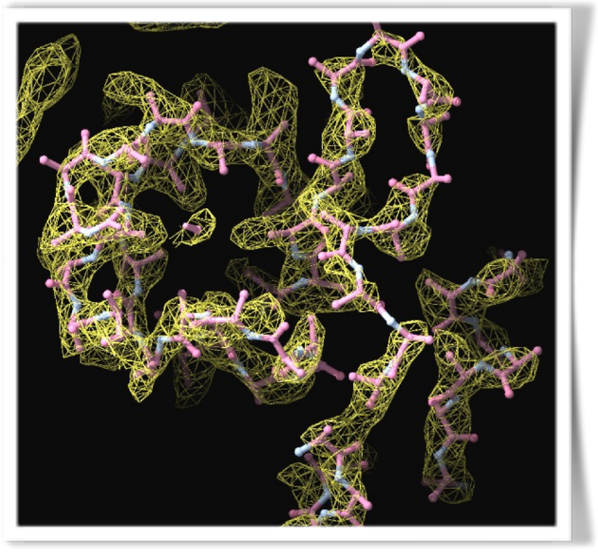
Automated map sharpening



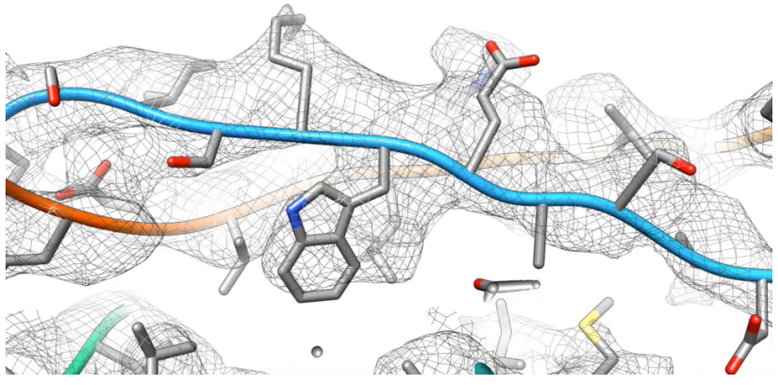
Map segmentation



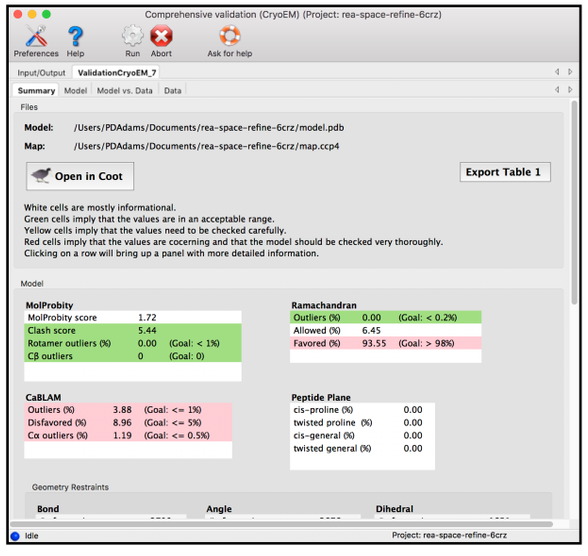
Rigid model docking



Automated model building



Real space refinement



Model and map validation

BERKELEY LAB
Lawrence Berkeley National Laboratory

Phenix

THE·UNIVERSITY·OF·CALIFORNIA·BERKELEY
1868

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

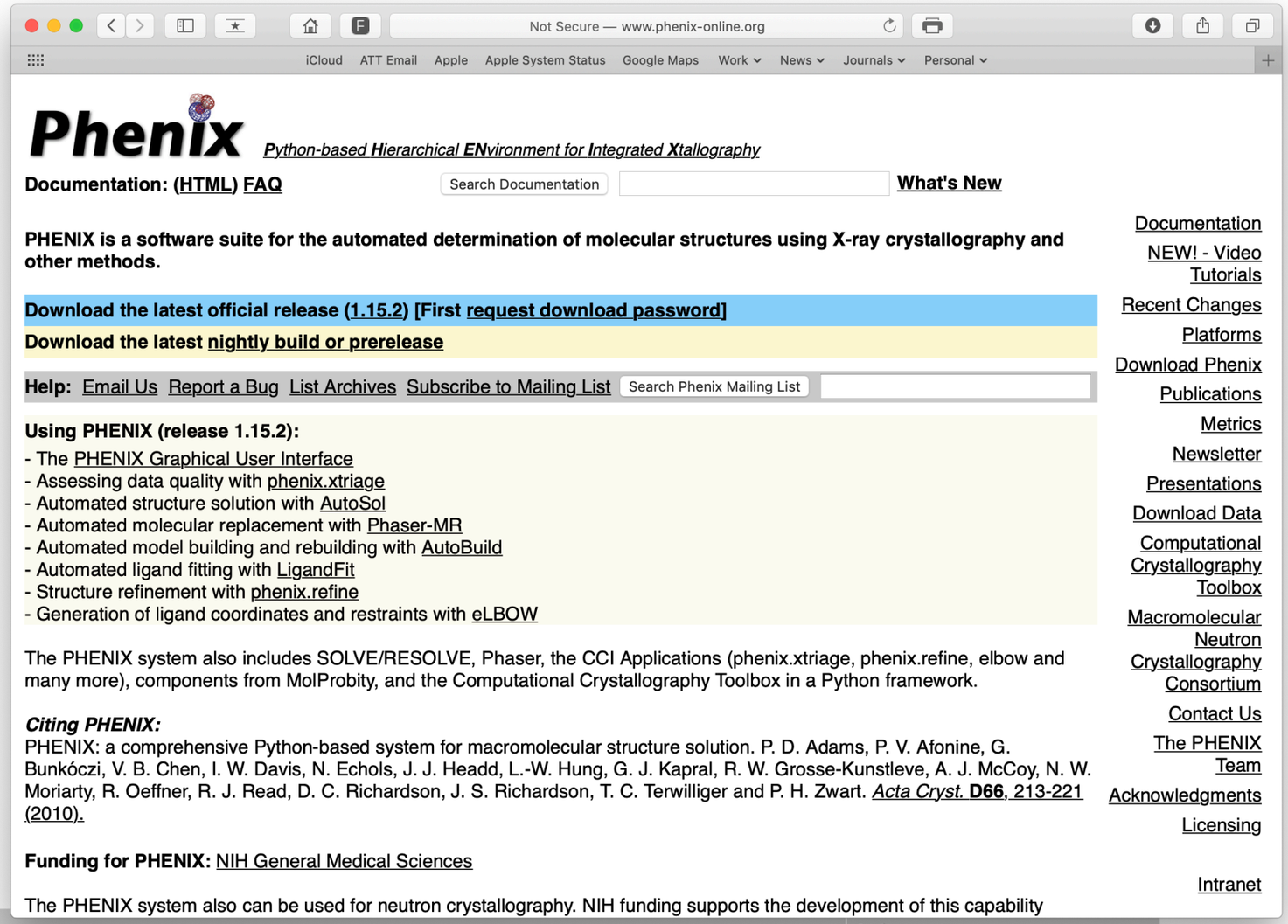
BERKELEY LAB
Lawrence Berkeley National Laboratory

Phenix

THE UNIVERSITY OF CALIFORNIA BERKELEY 1868

PHENIX Availability

- `phenix-online.org`



- Supported on:
  - Linux (RedHat, Fedora)
  - Mac OSX
  - Windows

- Extensive documentation

- Nightly builds:

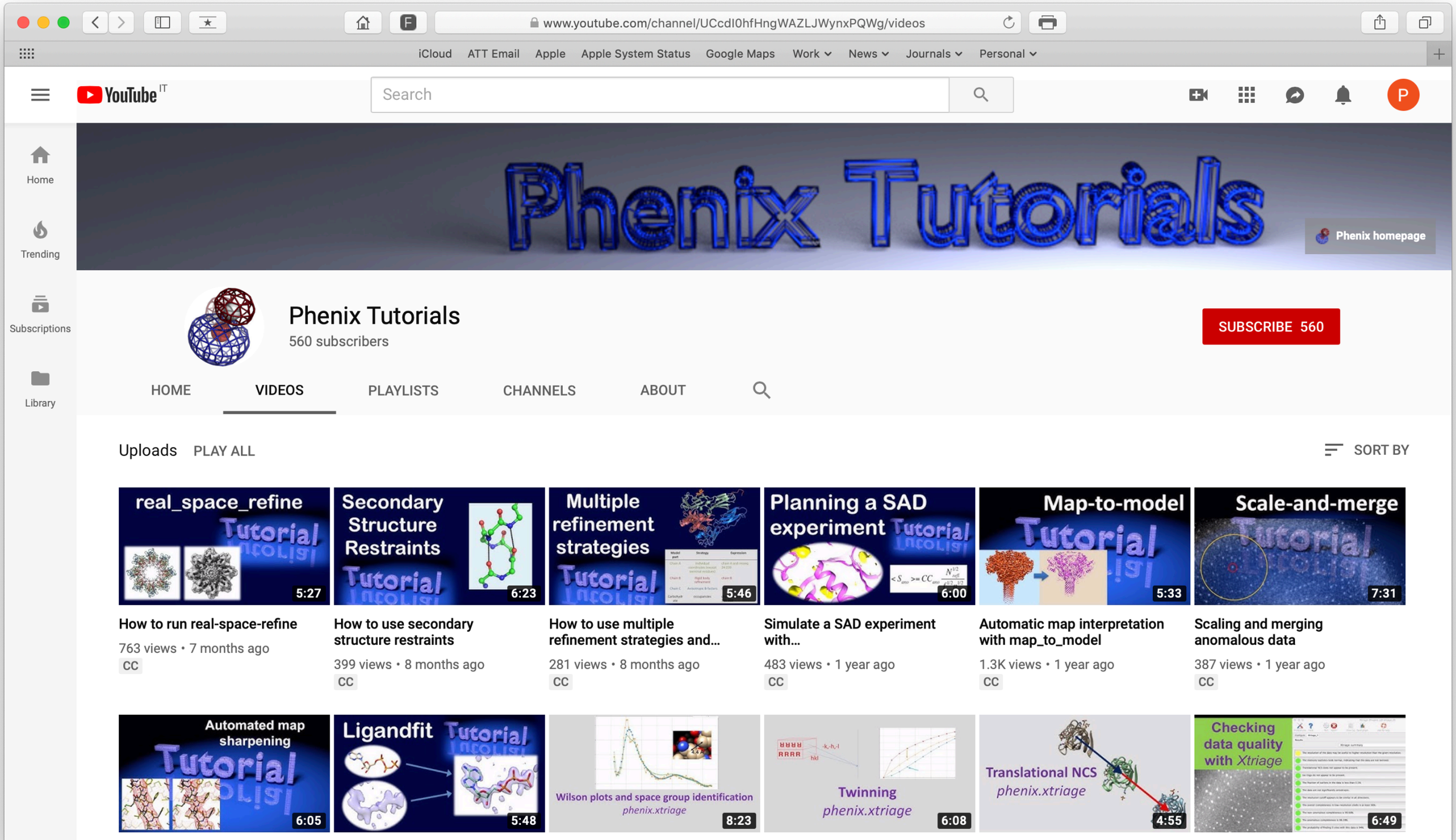
Version	Date	Status	Logs	Info
dev-3512	2019-05-21	successful	ci, intel-windows-x86_64, intel-linux-2.6-x86_64-centos6, mac-intel-osx-x86_64	docs; changelog
dev-3500	2019-05-09	successful	ci, intel-linux-2.6-x86_64-centos6, mac-intel-osx-x86_64, intel-windows-x86_64	docs; changelog
dev-3494	2019-05-03	successful	intel-linux-2.6-x86_64-centos6, ci, mac-intel-osx-x86_64, intel-windows-x86_64	docs; changelog
1.15.2-3472	**2019-04-11**	**successful**	**ci, intel-windows-x86_64, intel-linux-2.6-x86_64-centos6, mac-intel-osx-x86_64**	**Official 1.15.2 release; docs; changelog**

BERKELEY LAB
Lawrence Berkeley National Laboratory

Phenix

THE UNIVERSITY OF CALIFORNIA BERKELEY 1868

Video Tutorials



Dorothee Liebschner, Nigel Moriarty, Miffy Mifsud, Tom Terwilliger

Phenix

BERKELEY LAB
Lawrence Berkeley National Laboratory

THE UNIVERSITY OF CALIFORNIA BERKELEY 1868

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 Zwart

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

BERKELEY LAB
Lawrence Berkeley National Laboratory

Phenix

THE UNIVERSITY OF CALIFORNIA · BERKELEY · 1868 · LET THERE BE LIGHT

Tutorials

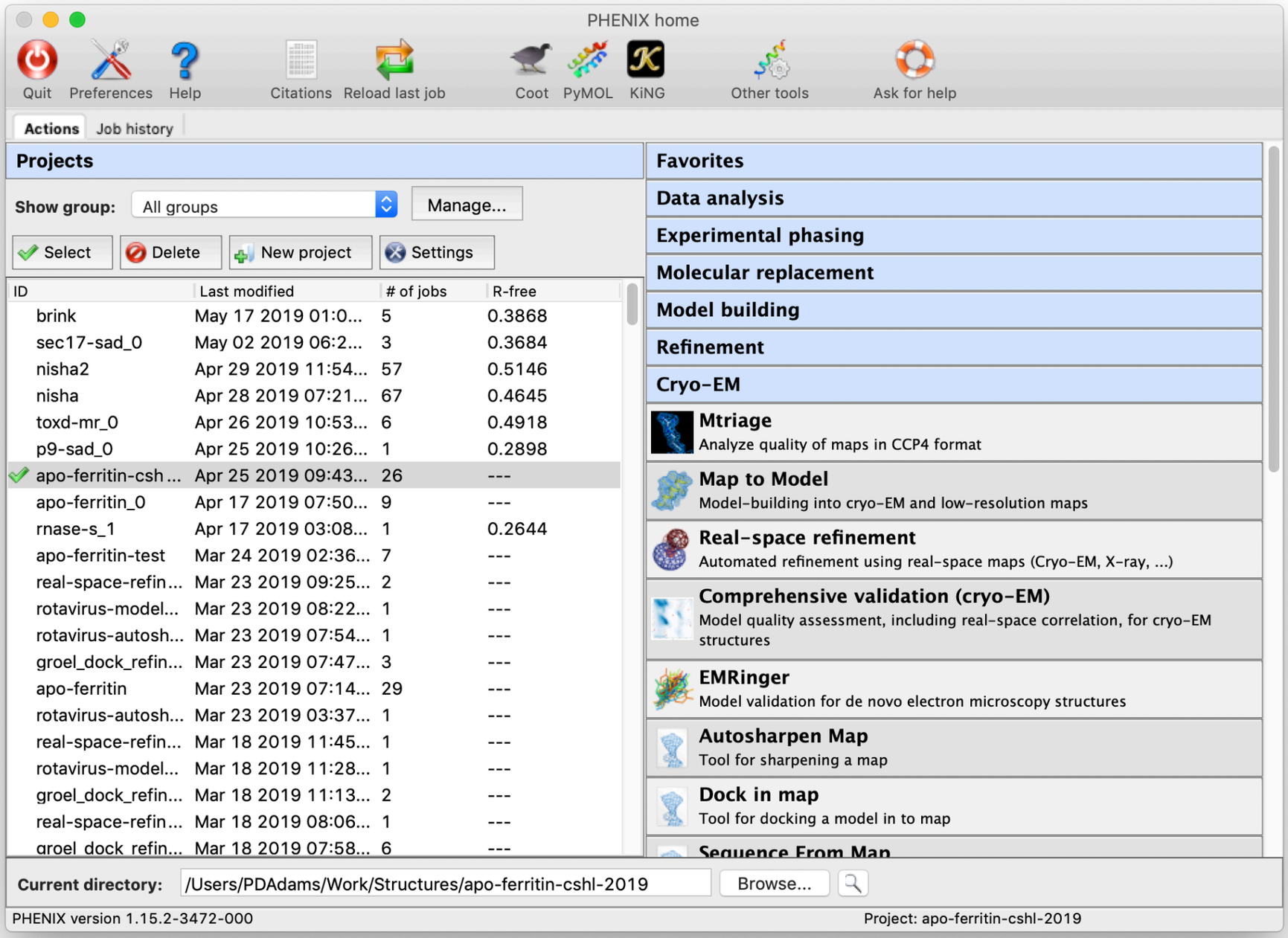
- Cryo-EM model placement and building
  - Symmetry determination
  - Rigid body model fitting
  - Map sharpening
  - Map segmentation
  - Automated model building
- Atomic model optimization and validation
  - Structure refinement (crystallography and cryo-EM)
  - Validation

BERKELEY LAB
Lawrence Berkeley National Laboratory

Phenix

Tutorial Format

- Use graphical user interface



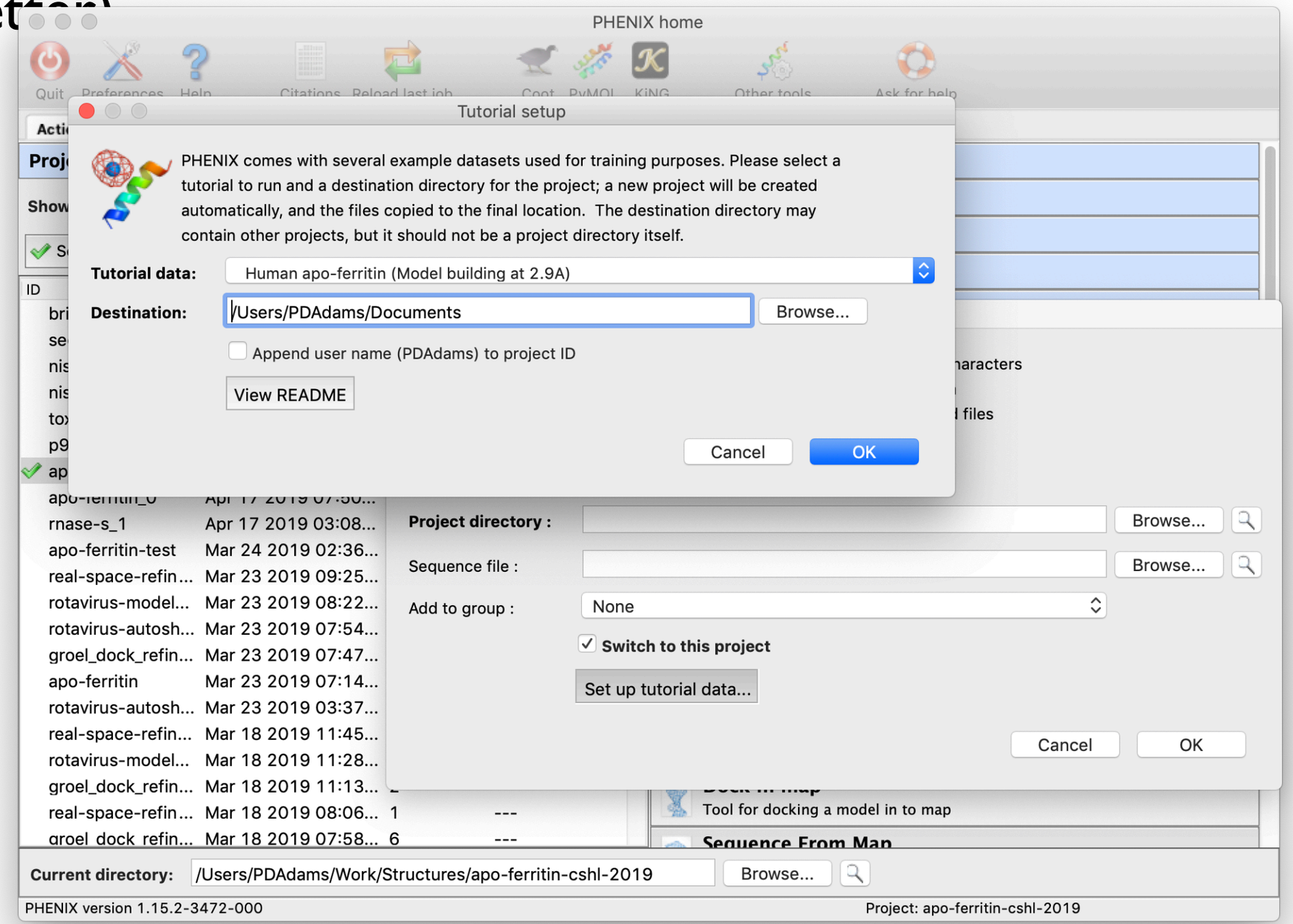
BERKELEY LAB
Lawrence Berkeley National Laboratory

Phenix

THE UNIVERSITY OF CALIFORNIA BERKELEY 1868

Tutorial Format

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



BERKELEY LAB
Lawrence Berkeley National Laboratory

Phenix

THE UNIVERSITY OF CALIFORNIA BERKELEY 1868