Computational Tools for Neutron Crystallography

IKON15, Lund, September 2018

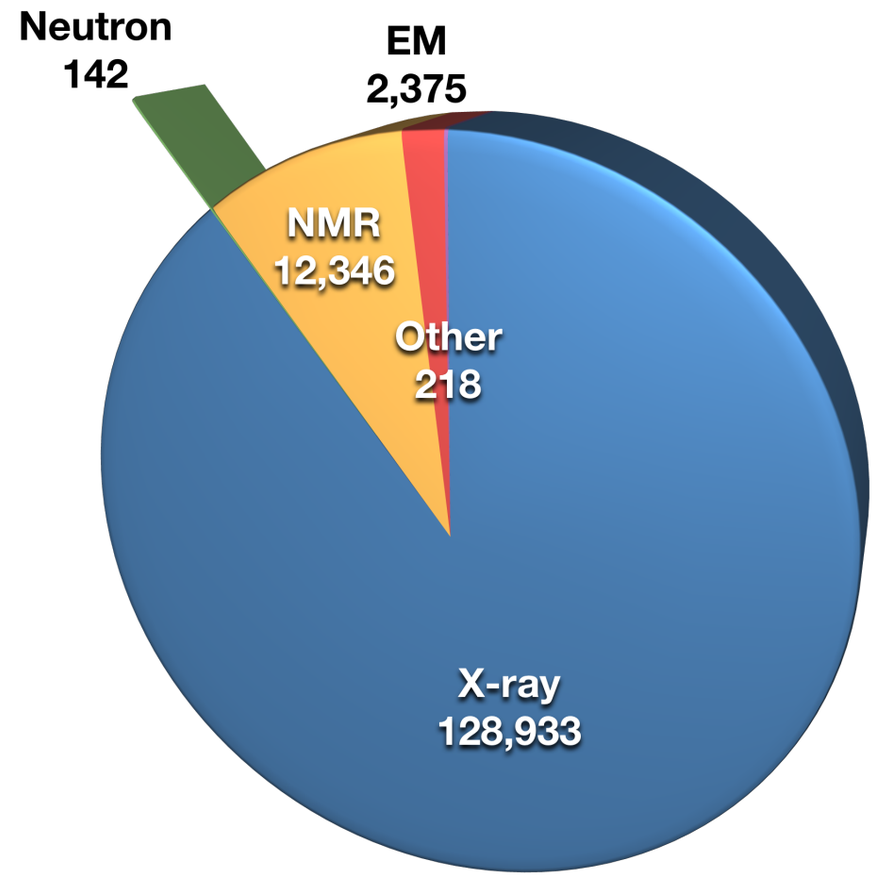
Paul Adams

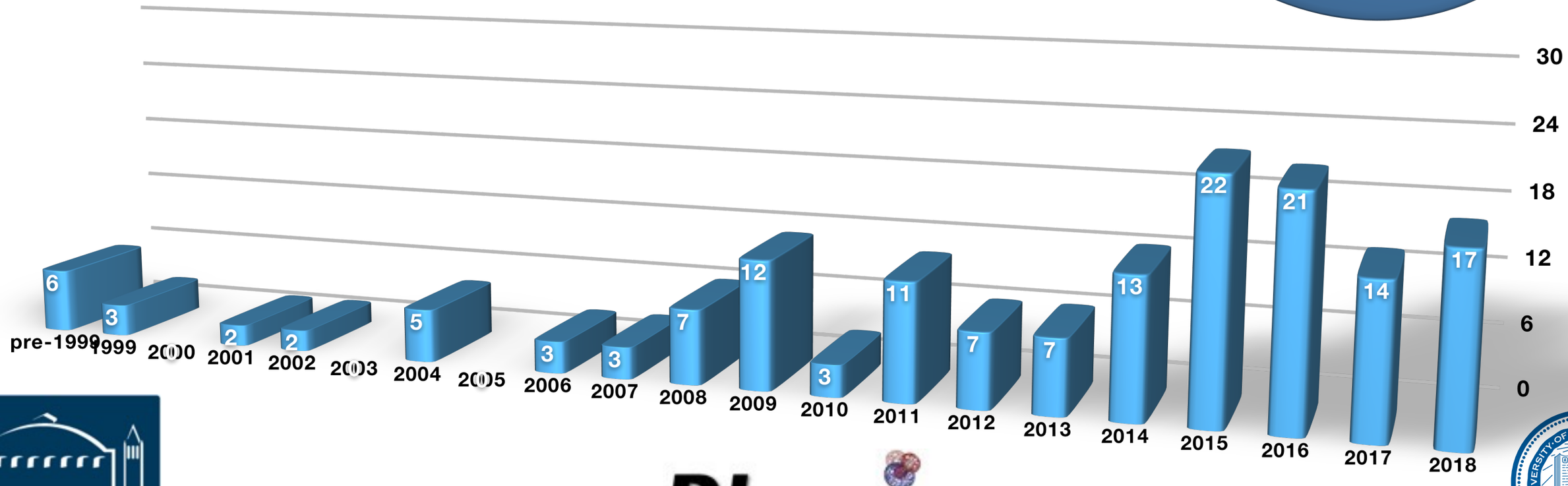
Lawrence Berkeley Laboratory and

Department of Bioengineering UC Berkeley

Growth of Neutron Crystallography

- Neutron crystallography has been challenging:
  - *Large crystals*
  - *Few data collection facilities*
  - *Long data collection times*
  - *Sample deuteration*
  - <u>Limited neutron software/algorithms</u>





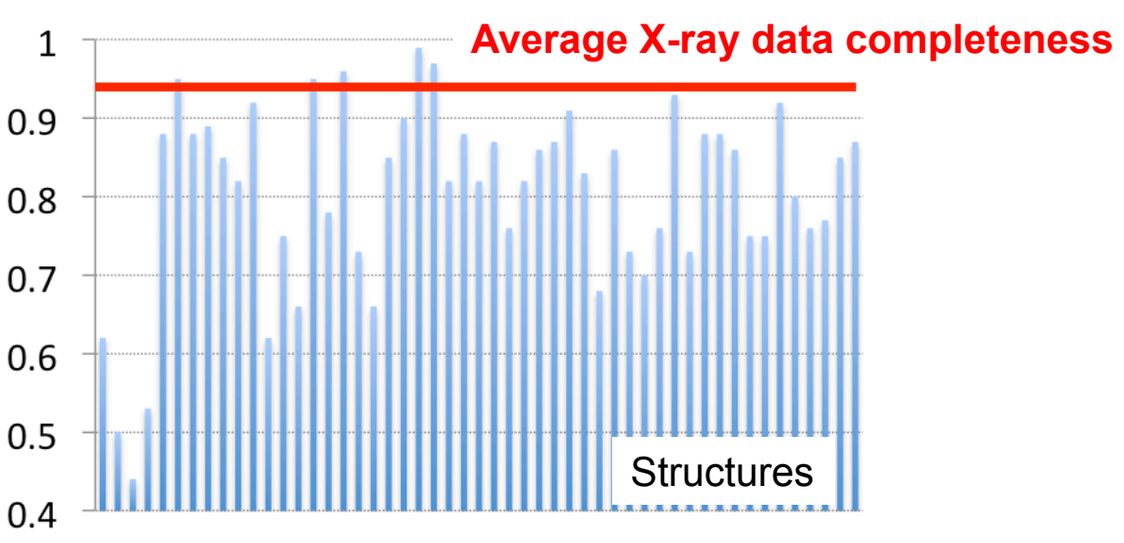
BERKELEY LAB
Lawrence Berkeley National Laboratory

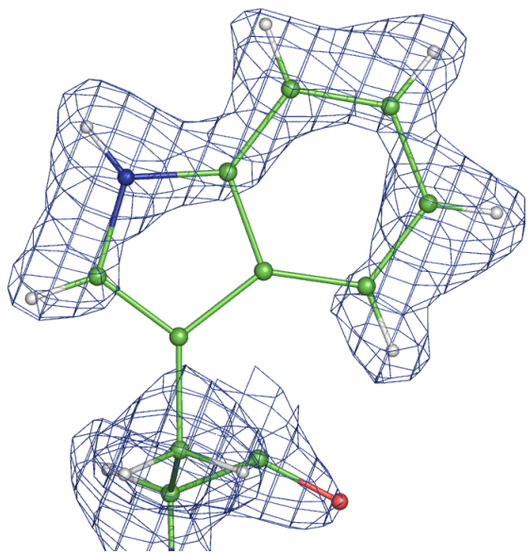
Phenix

THE UNIVERSITY OF CALIFORNIA BERKELEY

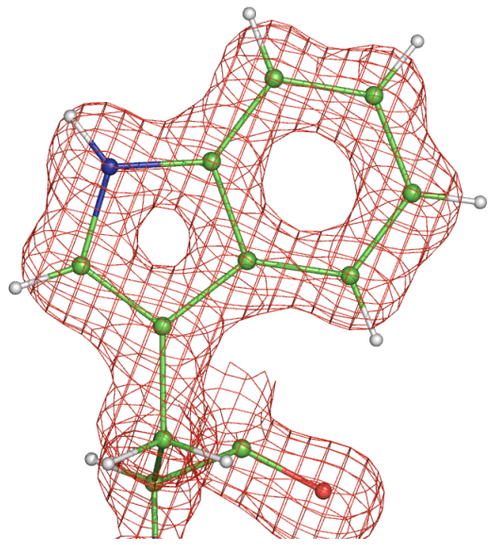
Challenges with Neutron Data

- Building H, D or H/D in the model, including water or ligands
- Optimizing the fit of water (DOD) into density
- Fit of rotatable X-H/D bonds into density
- Fewer data, more parameters to refine individually (H/D ~50% of the atoms)
- Cancellation effects make X-H species poorly defined in density
- Occupancy refinement of H/D sites
- Data quality: typically low overall and resolution bin completeness





Neutron 1.7 Å, 70% complete



X-ray 1.1 Å, 98% complete

BERKELEY LAB
Lawrence Berkeley National Laboratory

Phenix

THE UNIVERSITY OF CALIFORNIA BERKELEY

The Macromolecular Neutron Consortium

- Macromolecular Neutron Consortium (mnc.ornl.gov) is a collaboration between Lawrence Berkeley Lab and Oak Ridge National Lab:
  - Paul Langan, Vickie Lynch, Brendan Sullivan (ORNL), Marat Mustyakimov
  - Paul Adams, Pavel Afonine (LBNL)
  - Funded by NIH

- Goal is to create the computational tools to enable routine analysis of neutron diffraction data
- Initial developments made use of CNS, resulting in nCNS patches
- Current efforts are focused on Phenix

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, Gabor Bunkoczi, Rob Oeffner

Duke University

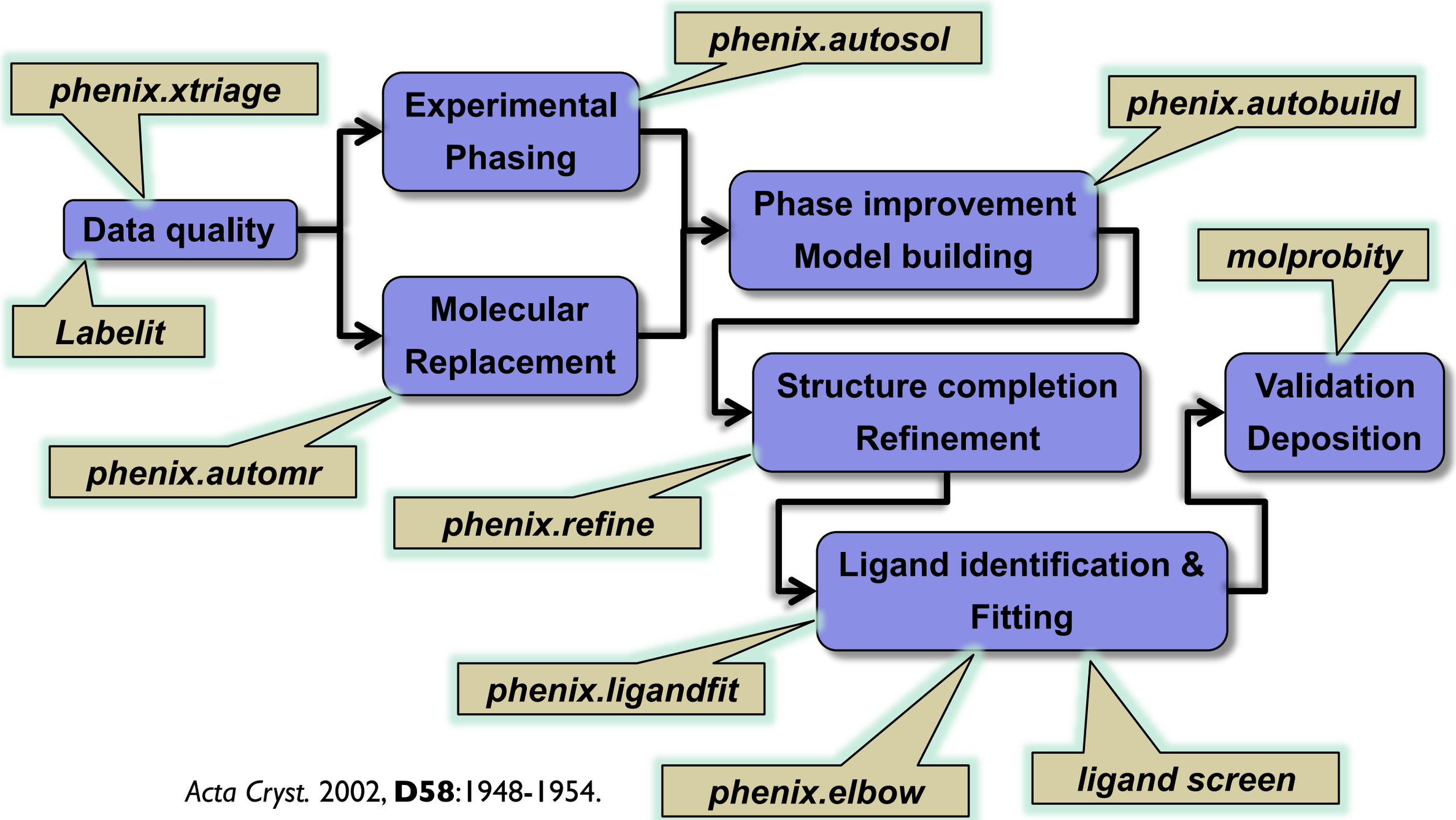
Jane & David Richardson, Vincent Chen, Bradley Hintze, Chris Williams

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

Automation of Structure Solution



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

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

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

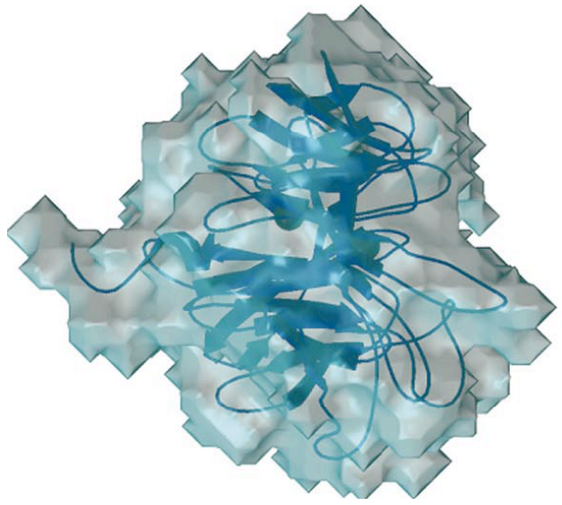
BERKELEY LAB
Lawrence Berkeley National Laboratory

Phenix

THE UNIVERSITY OF CALIFORNIA BERKELEY 1868

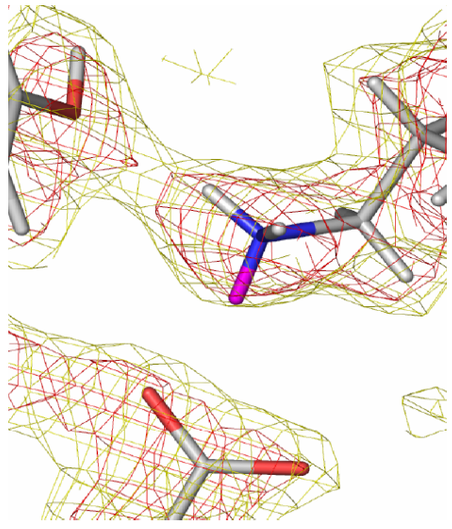
phenix.refine

Low



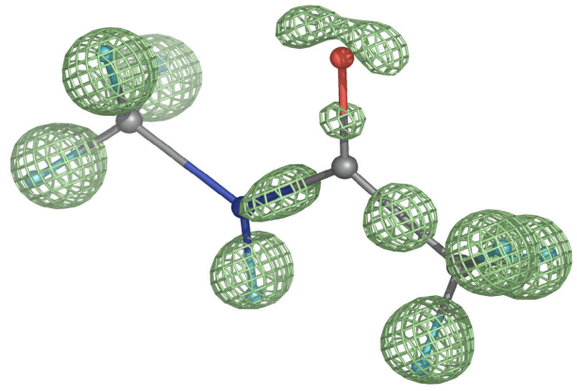
- Rigid body
- Group ADP
- Torsion angle constraints

Medium/High



- Restrained coordinates
- Restrained ADPs (iso/aniso)
- Automated water picking

Ultra-high



- Interatomic scatterers
- Unrestrained refinement
- Explicit hydrogens

- Simulated annealing
- NCS restraints (including automatic NCS determination and restraints generation)
- TLS refinement
- Occupancies (individual or group, automatically constrained for alternate side chains)
- Anomalous scattering factor refinement (individual or group)
- Twinned refinement target
- Refinement against X-ray and Neutron data

BERKELEY LAB
Lawrence Berkeley National Laboratory

Pavel Afonine, Ralf Grosse-Kunstleve, Nat Echols, Jeff Headd, Nigel Moriarty, Marat Mustyakimov, Tom Terwilliger, Sasha Urzhumtsev, Peter Zwart

Phenix

Acta Cryst. 2012, **D68**:352-367

THE UNIVERSITY OF CALIFORNIA BERKELEY 1868

Refinement with Neutron Data

- Similar to standard X-ray approach, with some key differences:
  - Addition of H or D to the model, including water or ligands
  - Addition of H and D as alternate conformations at exchangeable sites
  - Real space optimization of H/D positions w.r.t. density
  - Refinement of H/D atomic positions
  - Riding hydrogens are preferred at lower resolutions

Information from PDB file header				Information computed in *phenix.refine*				
PDB or NDB code	**Resolution (Å)**	**σ cutoff**	***R* (work/free) (%)**	***R* (work/free) (%)**	**H/D ratio**	**Data completeness (%)**	**No. of reflections/atoms**
1c57	2.4–15.79		27.0/30.1	20.5/25.7	90/10	87	8604/3677
2dxm	2.1–8.0	1	19.7/26.0	18.1/23.6	88/12	65	21242/9317
2efa	2.7–80.0	3	21.6/29.1	20.2/27.0	85/15	93	2198/788
2gve	2.2–10.0	3	27.1/31.9	24.5/30.4	0/100	80	19619/6185

BERKELEY LAB
Lawrence Berkeley National Laboratory

Phenix

Afonine PV, Mustyakimov M, Grosse-Kunstleve RW, Moriarty NW, Langan P, Adams PD: Joint X-ray and neutron refinement with phenix.refine. Acta Cryst. 2010, D66:1153-1163

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

Joint X-ray/Neutron Refinement

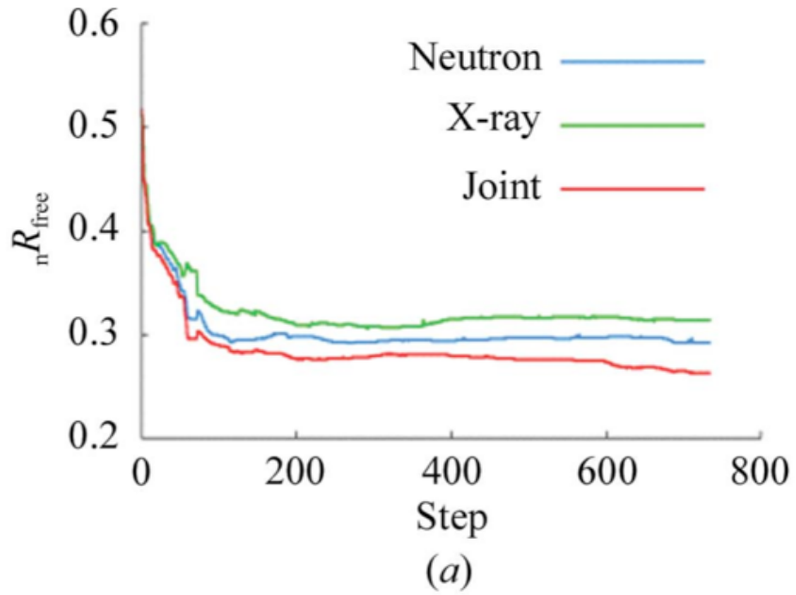
- Refine a single model against multiple data sets simultaneously
  - Introduced by Sheldrick et. (1978), Coppens et al. (1981)
  - Applied to macromolecules by Wlodawer & Sjölin (1982) and Wlodawer et al. (1982, 1989)
  - Implemented in nCNS *(Mustyakimov et al., 2009)*
  - Implemented in Phenix *(Afonine et al., 2010)*
  - Refinement Target:
    - $T_{Joint} = w_{xc} E_{Xray} + w_{nc} E_{Neutron} + w_c E_{Geom}$

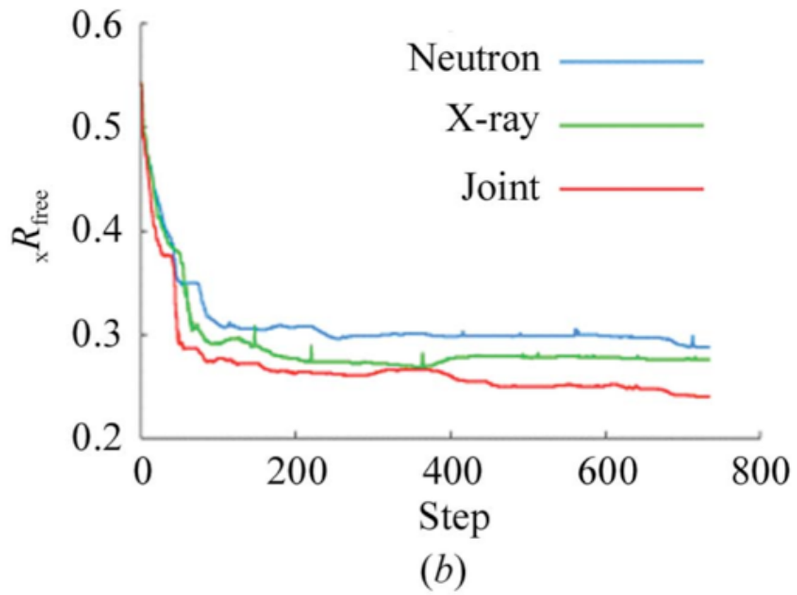
BERKELEY LAB
Lawrence Berkeley National Laboratory

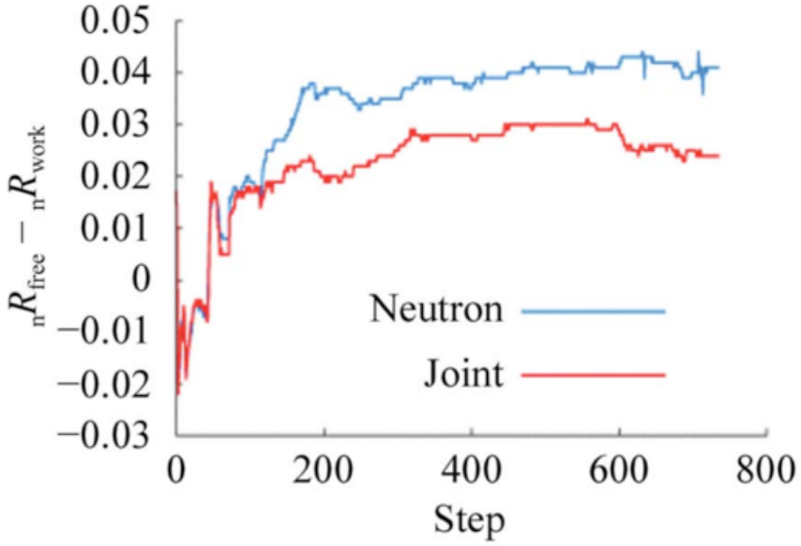
Phenix

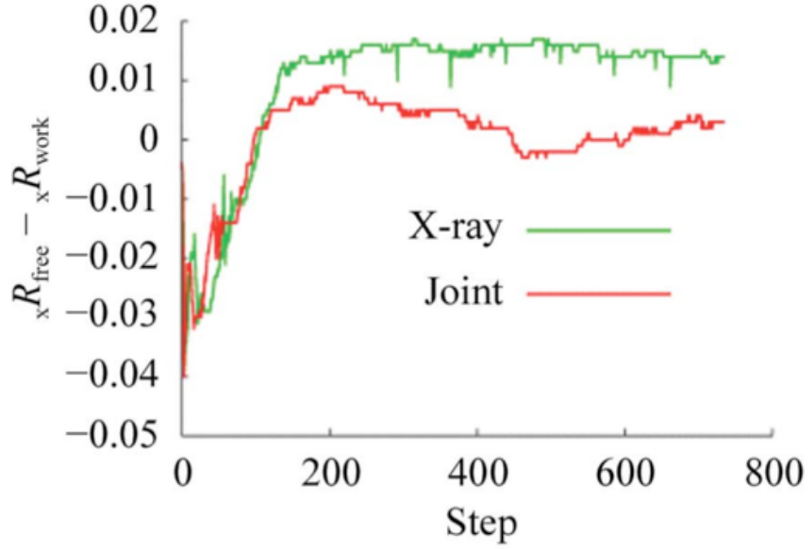
THE UNIVERSITY OF CALIFORNIA BERKELEY 1868

Joint X-ray/Neutron Refinement









- Tests with artificial and real data showed that both X-ray and neutron structures could be improved, with typically a larger effect on neutron derived models

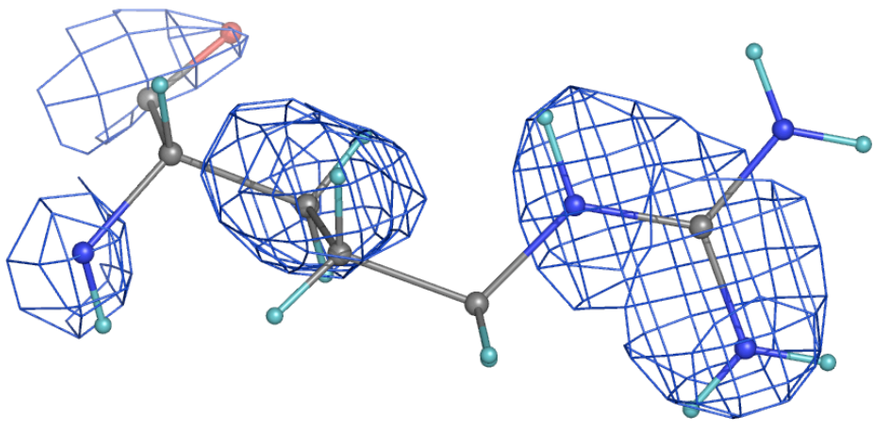
BERKELEY LAB
Lawrence Berkeley National Laboratory

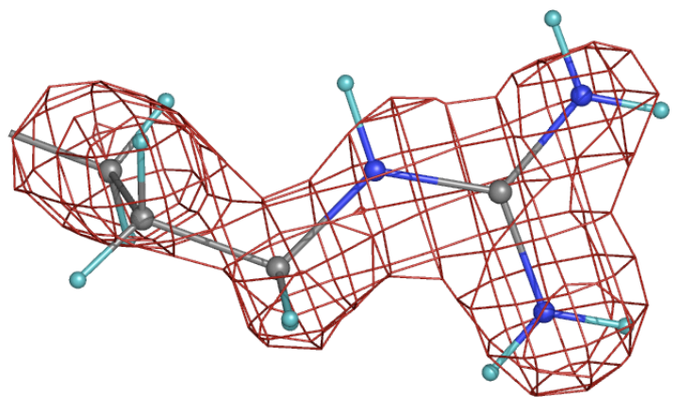
Phenix

THE UNIVERSITY OF CALIFORNIA BERKELEY 1868

Joint X-ray/Neutron Refinement

- Joint refinement of a model against both X-ray and Neutron data improves density maps





Neutron refinement alone

Neutron (Joint refinement)

PDB ID		Published R-work / R-free		phenix.refine R-work / R-free			
				Joint X + N		X-ray only	Neutron only
X	N	X	N	X	N		
1iu5	1iu6	18.7 / 20.3	19.4 / 25.4	13.1 / 16.9	15.2 / 21.0	12.7 / 17.2	19.1 / 21.6
1woe	1v9g	17.6 / 20.6	22.2 / 29.4	15.0 / 18.5	23.4 / 26.7	15.9 / 18.7	26.7 / 31.3
2er7	1gkt	14.2 /-	23.5 / 27.4	12.1 / 14.8	21.5 / 24.7	12.3 / 15.2	19.9 / 25.9

Adams PD, Mustyakimov M, Afonine PV, Langan P: Generalized X-ray and neutron crystallographic analysis: more accurate and complete structures for biological macromolecules. Acta Cryst. 2009, D65:567-573.

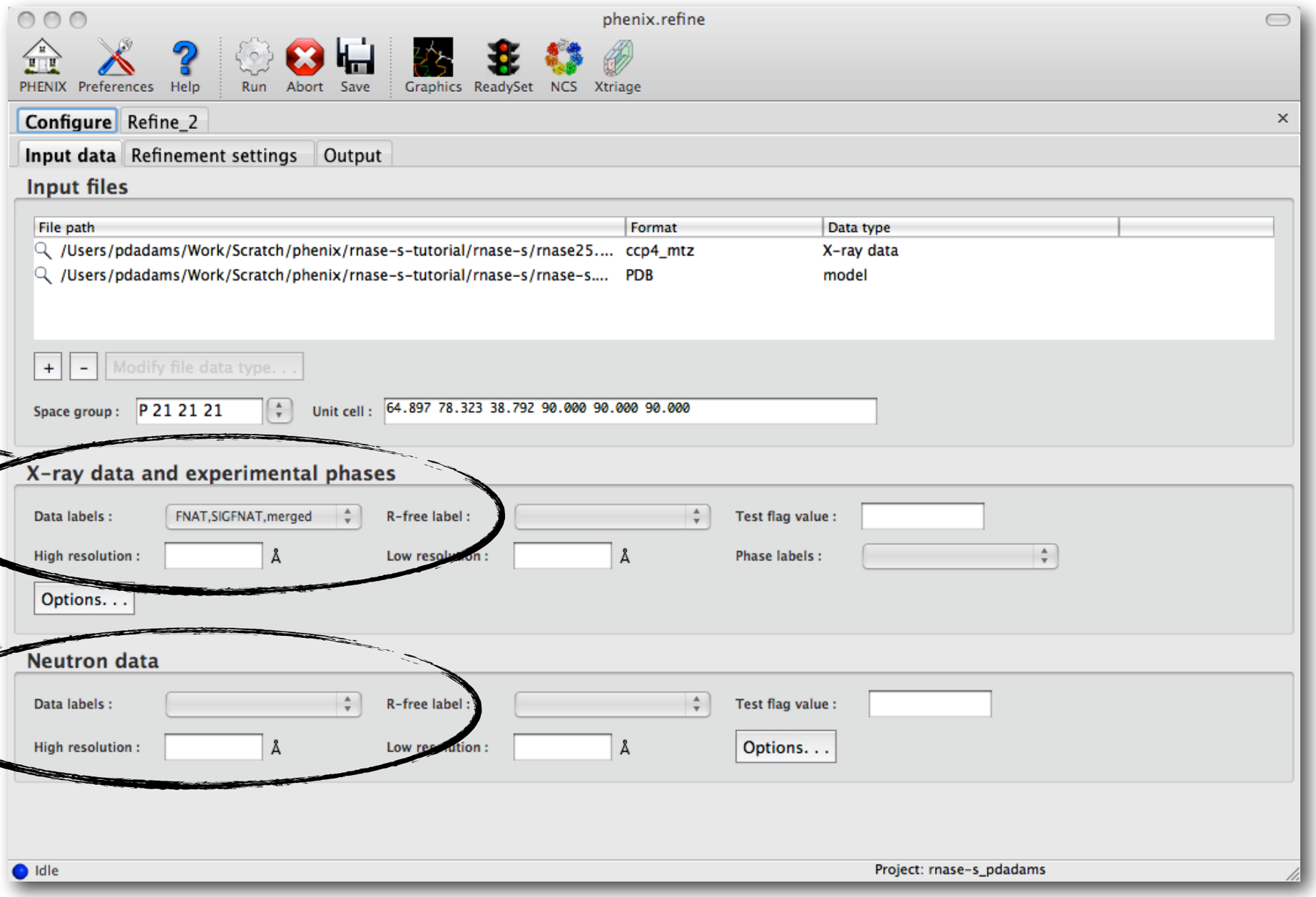
BERKELEY LAB
Lawrence Berkeley National Laboratory

Phenix

THE UNIVERSITY OF CALIFORNIA BERKELEY 1868

Enabling Joint Refinement

- The Phenix GUI provides easy access to joint refinement



X-ray data

Neutron data

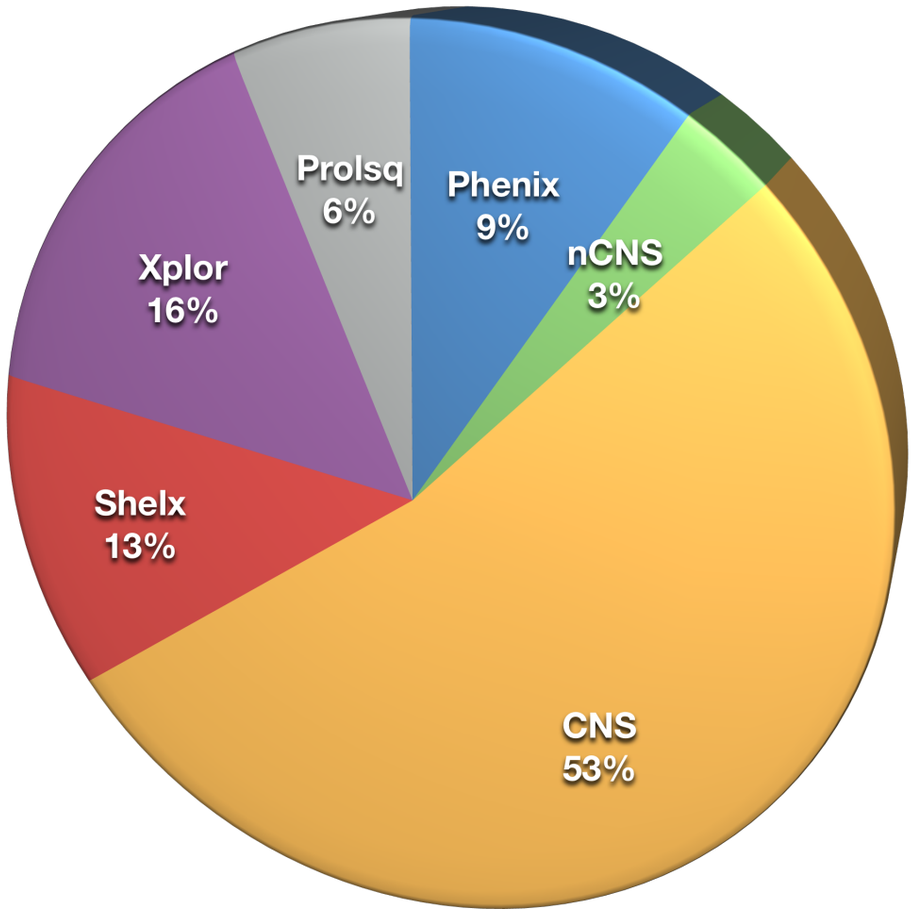
BERKELEY LAB
Lawrence Berkeley National Laboratory

Phenix

THE UNIVERSITY OF CALIFORNIA BERKELEY 1868

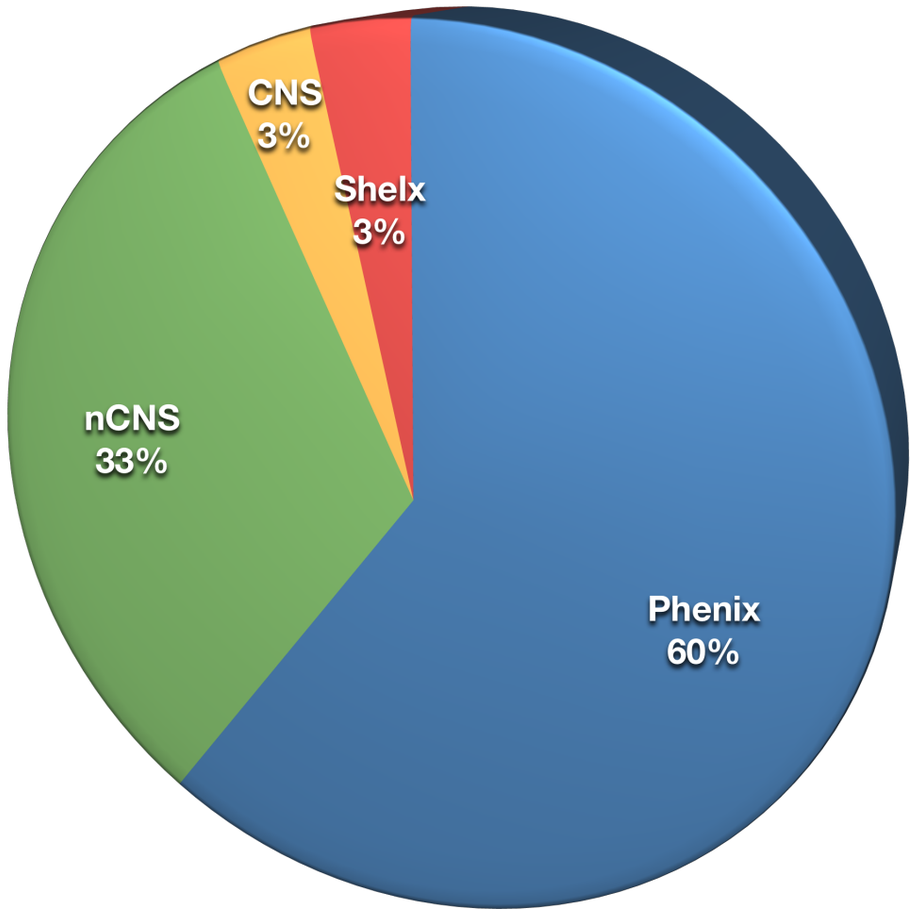
Impact on Structure Depositions

Before 2010



N=31

Since 2010



N=116

- Neutron structure refinement is increasingly making use of neutron-specific tools

BERKELEY LAB
Lawrence Berkeley National Laboratory

Phenix

THE UNIVERSITY OF CALIFORNIA BERKELEY 1868

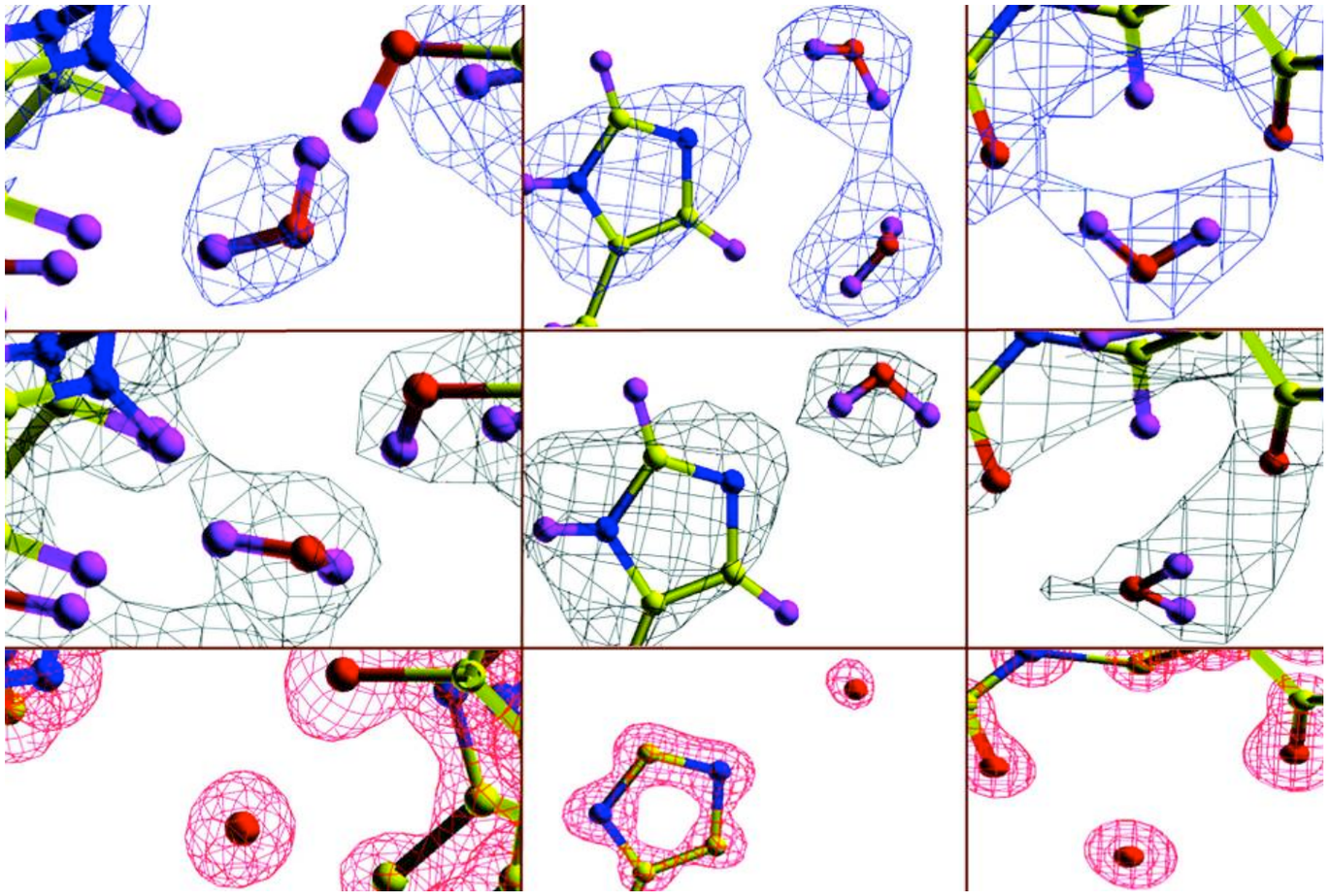
Challenges of Refining a Single Model

- Requires comparable datasets collected from isomorphous crystals and at the same temperature
- Structures still can have local differences in solvent structure, rotameric states
- Works well if data sets are comparable but relative weighting of E_{XRAY} and $E_{NEUTRON}$ is challenging if data sets are very different (e.g. 3Å neutron data, and 1Å X-ray data)

Neutron @15K

Neutron @293K

X-ray @110K



BERKELEY LAB
Lawrence Berkeley National Laboratory

Phenix

Blakeley et al., Proc Natl Acad Sci U S A. 2004 101: 16405–16410.

Coupled Refinement

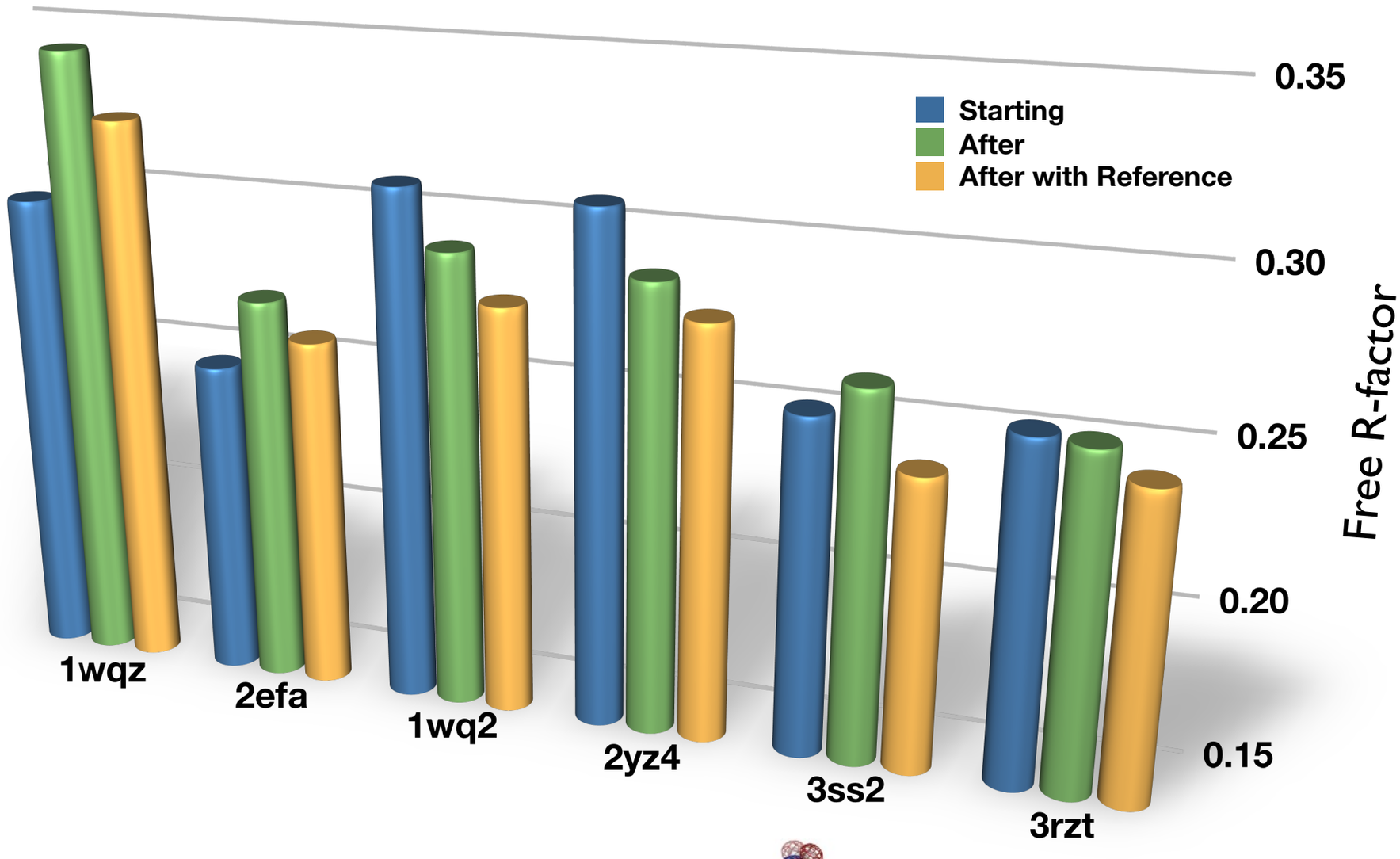
- Refine two separate models against the respective data sets
  - Eliminates the requirement for highly isomorphous crystals collected at the same temperature
- Use the higher quality/resolution structure as a source of information to inform the other refinement
  - Allows for structural variability and differences in water structure
  - Naturally accounts for differences in X-H/X-D bonds lengths

BERKELEY LAB
Lawrence Berkeley National Laboratory

Phenix

Enhanced Neutron Structure Refinement

- Initial tests suggest that use of the X-ray structure as a reference can produce improved neutron models
  - *It should be noted that refinement against neutron data alone is better than in the past as the algorithms have improved*



Phenix

Pavel Afonine (LBNL)

Coupled Refinement Developments

- Refinement workflows that are planned for the new neutron/X-ray refinement framework:
  - Given best possible refined X-ray model, use it as a source of reference restraints for refinement of neutron structure against neutron data only.
  - Refine X-ray and neutron models in one combined refinement run:
    - Use X-ray model as a source of restraints for neutron refinement
    - Use neutron model to adjust rotatable hydrogens in X-ray model
- Relies on new riding hydrogen model parameterization

BERKELEY LAB
Lawrence Berkeley National Laboratory

Phenix

THE UNIVERSITY OF CALIFORNIA · BERKELEY · 1868

Going Beyond Structure Refinement

- Refinement is one part of the problem
- As with X-ray crystallography, structure completion remains a significant bottleneck
  - We need to address solvent building (especially important for neutron data with rich information about waters)
- How can we make the best maps for interpretation?
- Structure deposition remains a serious problem, with many mistakes in the wwPDB

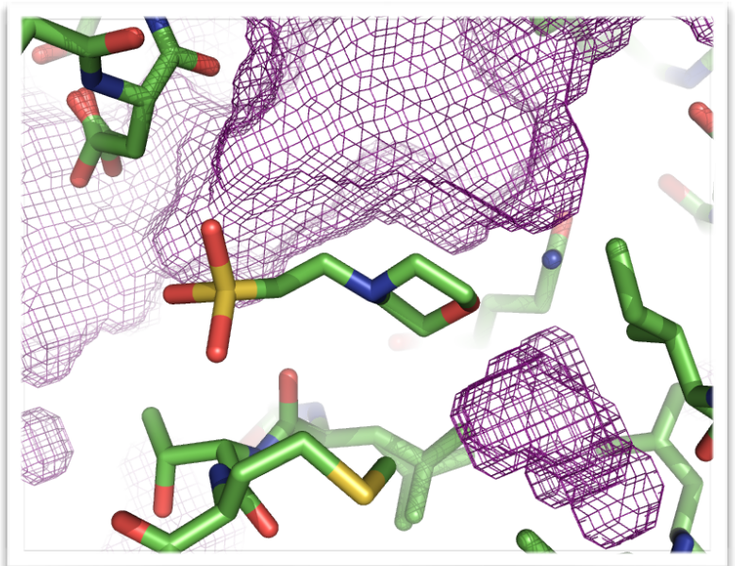
BERKELEY LAB
Lawrence Berkeley National Laboratory

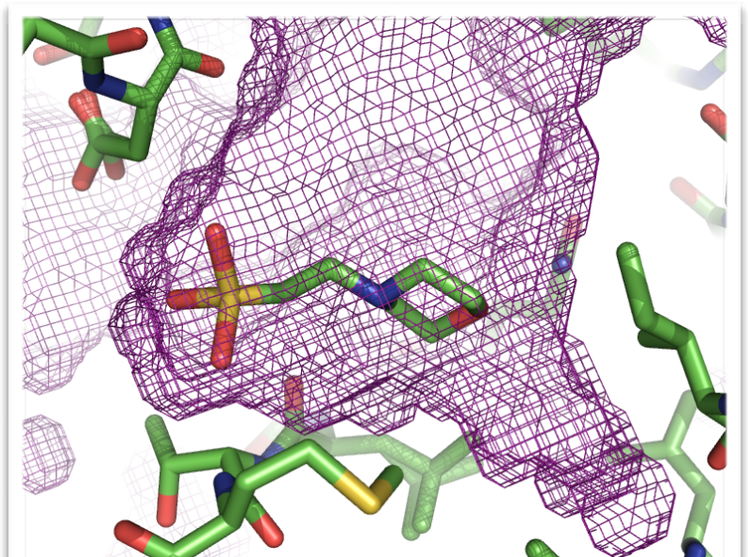
Phenix

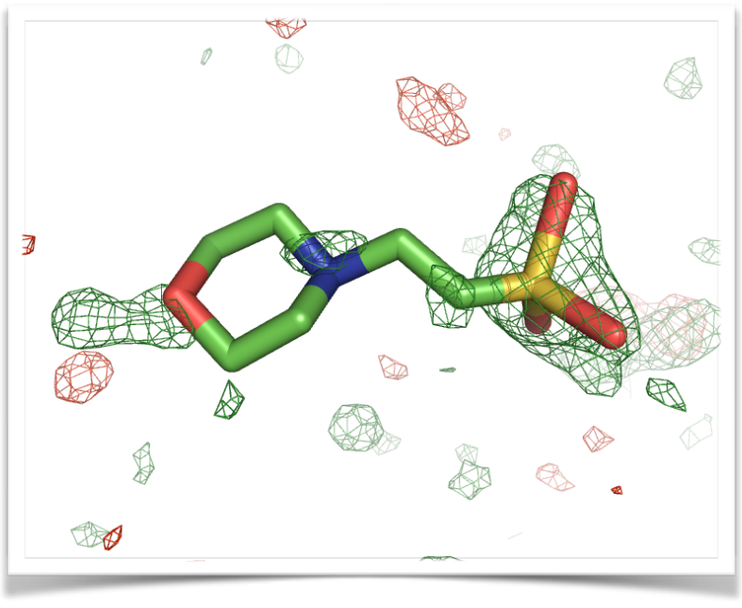
THE UNIVERSITY OF CALIFORNIA BERKELEY 1868

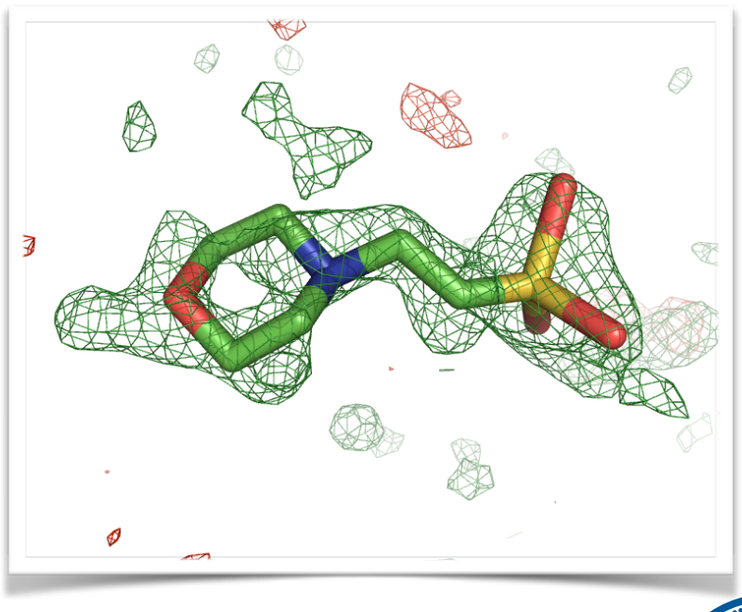
Better Omit Maps

- The bulk solvent model can obscure density for omitted regions in traditional omit maps.









Liebschner D, et al: Polder maps: improving OMIT maps by excluding bulk solvent. *Acta Cryst.* 2017, **D73**:148-157

BERKELEY LAB
Lawrence Berkeley National Laboratory

Phenix

THE UNIVERSITY OF CALIFORNIA · BERKELEY · 1868

Structure Deposition

- Tools for deposition of neutron structures lag X-ray
- Depositing joint X-ray/neutron structures has been challenging (1+ years to negotiate a format for phenix.refine with wwPDB)
- The original PDB format anticipated one model, one data set (X-ray data)
- The PDB format will be replaced with mmCIF
  - Allows for a much richer description of experiments, data and model
  - Deposition of joint refinements with multiple data sets
  - Deposition of coupled refinements with linked models
- Phenix now reads and writes mmCIF for many tasks (e.g. *phenix.refine*)

Liebschner et al: Evaluation of models determined by neutron diffraction and proposed improvements to their validation and deposition. *Acta Cryst.* 2018, **D74**:800-813

BERKELEY LAB
Lawrence Berkeley National Laboratory

Phenix

mmCIF Example

```
loop_
  _atom_site.group_PDB
  _atom_site.id
  _atom_site.label_atom_id
  _atom_site.label_alt_id
  _atom_site.label_comp_id
  _atom_site.auth_asym_id
  _atom_site.auth_seq_id
  _atom_site.pdbx_PDB_ins_code
  _atom_site.Cartn_x
  _atom_site.Cartn_y
  _atom_site.Cartn_z
  _atom_site.occupancy
  _atom_site.B_iso_or_equiv
  _atom_site.type_symbol
  _atom_site.pdbx_formal_charge
  _atom_site.label_asym_id
  _atom_site.label_entity_id
  _atom_site.label_seq_id
  _atom_site.pdbx_PDB_model_num
ATOM       1  N      .  ASP  L    1  A  11.31147   19.15776  20.12829  1.000  30.02111  N  ?  A  ?    1  1
ATOM       2  CA     .  ASP  L    1  A  10.10373   18.68105  19.44836  1.000  29.12464  C  ?  A  ?    1  1
ATOM       3  C      .  ASP  L    1  A   9.87214   17.15878  19.55617  1.000  26.73154  C  ?  A  ?    1  1
ATOM       4  O      .  ASP  L    1  A   8.81111   16.65235  19.17214  1.000  21.71081  O  ?  A  ?    1  1
ATOM       5  CB     .  ASP  L    1  A  10.14403   19.08618  17.97261  1.000  38.70228  C  ?  A  ?    1  1
ATOM       6  CG     .  ASP  L    1  A   8.75963   19.33936  17.39469  1.000  44.90010  C  ?  A  ?    1  1
ATOM       7  OD1    .  ASP  L    1  A   7.77712   19.32440  18.16915  1.000  44.50718  O  ?  A  ?    1  1
ATOM       8  OD2    .  ASP  L    1  A   8.66147   19.57786  16.17106  1.000  50.05253  O  ?  A  ?    1  1
```

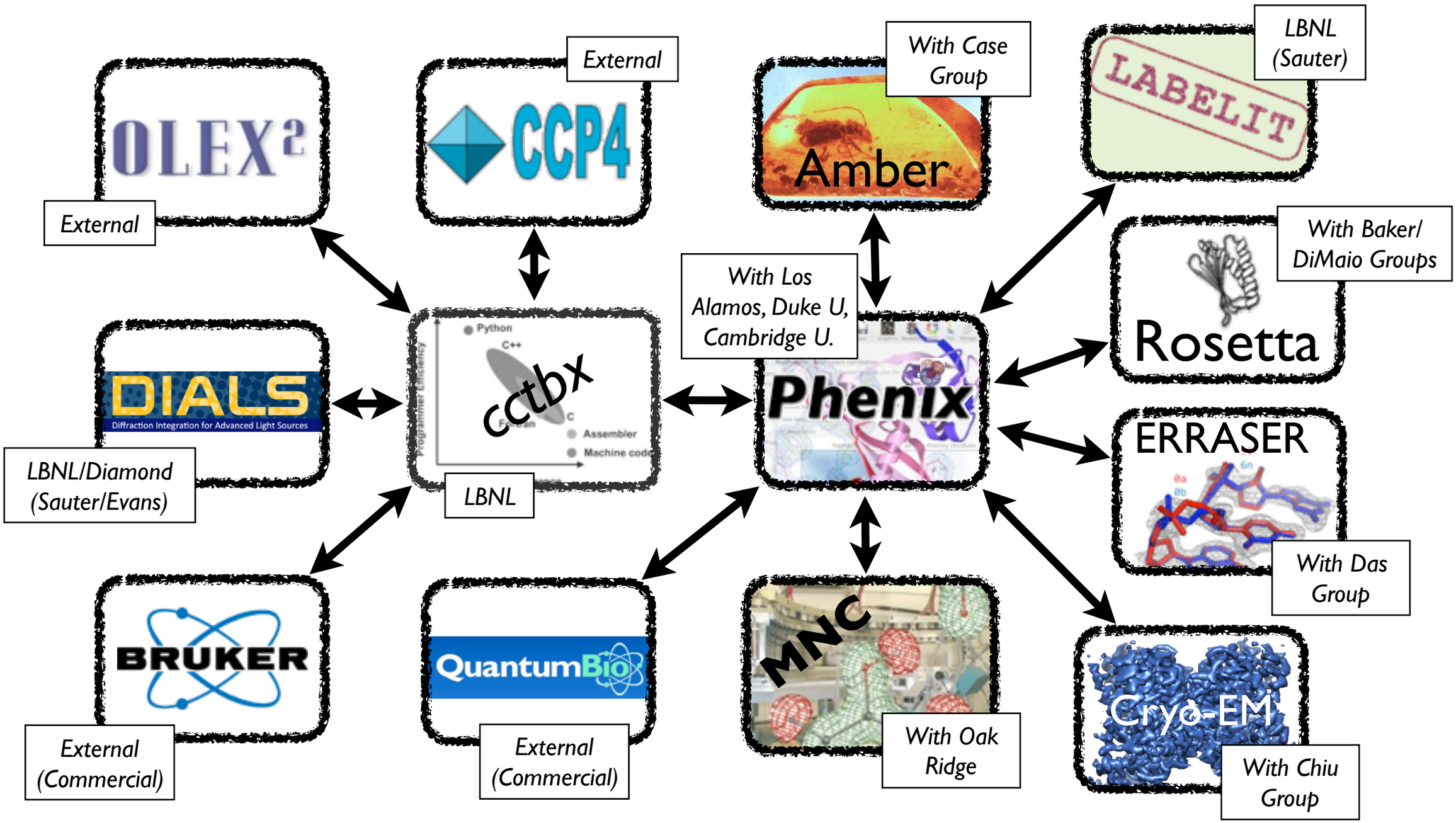
```
ATOM      1  N   ASP L   1A     11.311  19.158  20.128  1.00 30.02           N
ATOM      2  CA  ASP L   1A     10.104  18.681  19.448  1.00 29.12           C
ATOM      3  C   ASP L   1A      9.872  17.159  19.556  1.00 26.73           C
ATOM      4  O   ASP L   1A      8.811  16.652  19.172  1.00 21.71           O
ATOM      5  CB  ASP L   1A     10.144  19.086  17.973  1.00 38.70           C
ATOM      6  CG  ASP L   1A      8.760  19.339  17.395  1.00 44.90           C
ATOM      7  OD1 ASP L   1A      7.777  19.324  18.169  1.00 44.51           O
ATOM      8  OD2 ASP L   1A      8.661  19.578  16.171  1.00 50.05           O
```

- Efforts underway to represent diffraction data in mmCIF

Richard Gildea (LBNL)

Phenix

Phenix - a Structural Biology Hub



Adams PD, Afonine PV, Bunkóczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung L-W, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson JS, Terwilliger TC, Zwart PH: *PHENIX*: a comprehensive Python-based system for macromolecular structure solution. *Acta Cryst.* 2010, **D66**:213-221.

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

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

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

- Automated sharpening and model building for cryo-EM maps

- Real space refinement (for X-ray and cryo-EM)

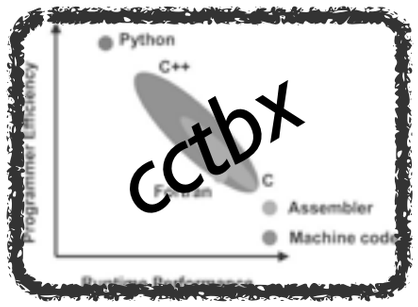
- New validation methods for cryo-EM models

BERKELEY LAB
Lawrence Berkeley National Laboratory

Phenix

THE UNIVERSITY OF CALIFORNIA BERKELEY 1868

Phenix and cctbx



- Open source
- General scientific tools (scitbx)
- Input/output, various formats (iotbx)
- Crystallographic tools (cctbx)
- Macromolecular-specific (mmtbx)
- Installation, testing (libtbx)

Python: 400,000
C++: 250,000

- Available source
- Graphic User Interface
- End-user programs
- Pipelines
- Some high-level algorithms

Python: 300,000

BERKELEY LAB
Lawrence Berkeley National Laboratory

Phenix

THE UNIVERSITY OF CALIFORNIA BERKELEY 1868

Computational Crystallography Toolbox Philosophy

- Fundamental crystallographic algorithms are required to build new software
- These should be available as a library
- Using these algorithms in the context of an interpreted language is the most efficient way to develop new software, and build automated "pipelines"
- Fundamental algorithms should be as general as possible
- Regression tests for every component

BERKELEY LAB
Lawrence Berkeley National Laboratory

Phenix

THE UNIVERSITY OF CALIFORNIA BERKELEY 1868

The Computational Crystallography Toolbox

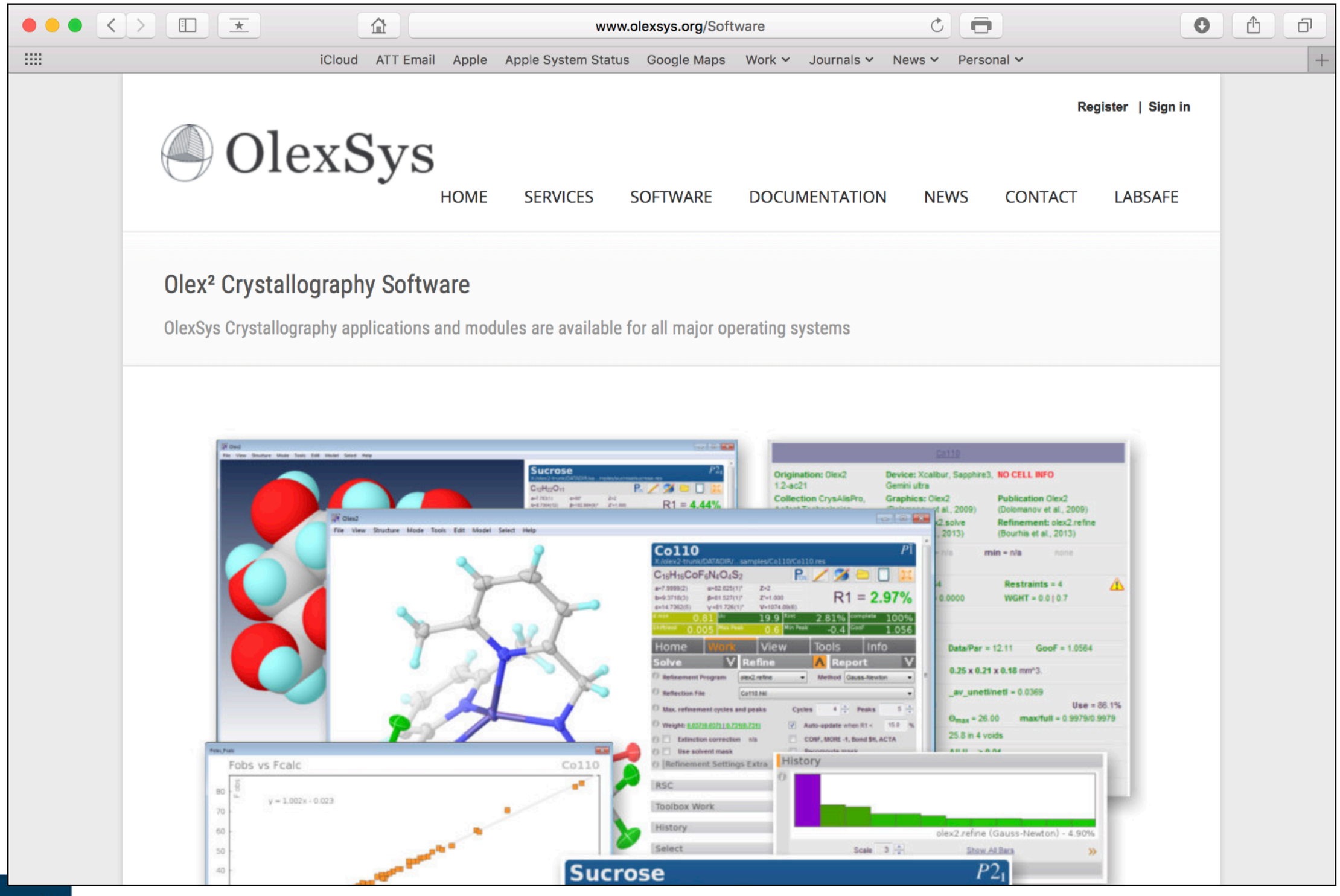
- Is an evolving C++/Python library of fundamental algorithms for computational crystallography (and more)
- Stable and tested
  - Phenix suite relies heavily on cctbx
  - Basis for Olex^2
  - Platform for DIALS data processing
- Contains a variety of tools for IO, model manipulation, refinement, restraints, etc
- Open source project on GitHub (https://github.com/cctbx)
  - ensures continued availability
  - rapid development
  - easy contribution by all developers

BERKELEY LAB
Lawrence Berkeley National Laboratory

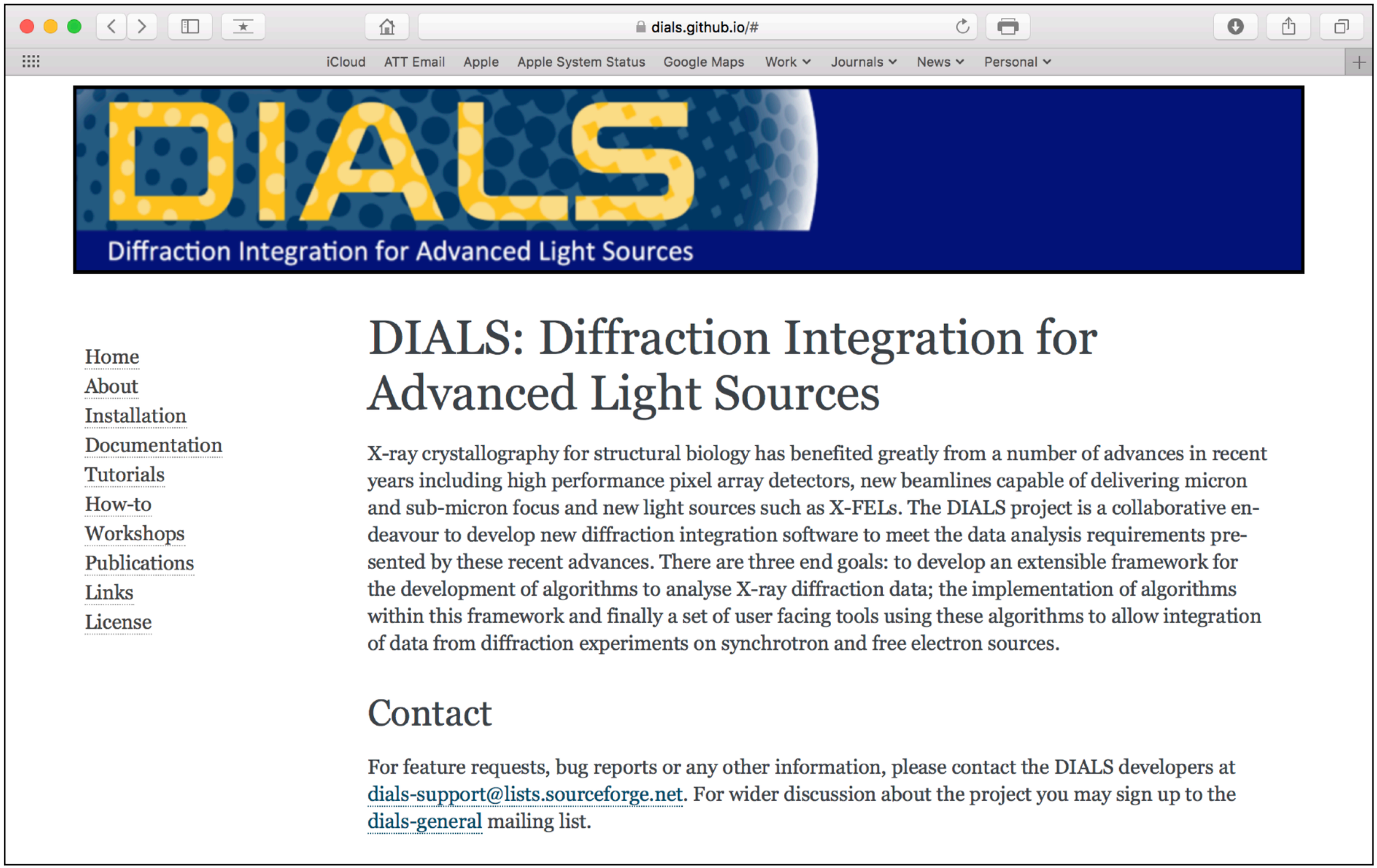
Phenix

THE UNIVERSITY OF CALIFORNIA · BERKELEY · 1868

Olex²


www.olexsys.org/Software
iCloud
ATT Email
Apple
Apple System Status
Google Maps
Work
Journals
News
Personal
Register | Sign in
OlexSys
HOME
SERVICES
SOFTWARE
DOCUMENTATION
NEWS
CONTACT
LABSAFE
Olex² Crystallography Software
OlexSys Crystallography applications and modules are available for all major operating systems
Sucrose
Co110
Fobs vs Fcalc
History
Phenix

DIALS



BERKELEY LAB
Lawrence Berkeley National Laboratory

Phenix

THE UNIVERSITY OF CALIFORNIA BERKELEY 1868

The Computational Crystallography Toolbox

A lot of functionality is written in C++ and exposed through Python

- Custom-made types for arrays, matrices, vectors, coordinates and much more with implemented operations with them
- Classes to work with symmetry, unit cell, structure factors
- Minimizers, FFT
- Restraints
- PDB, mmCIF, MTZ and other parsers
- Containers for structural and experimental information

BERKELEY LAB
Lawrence Berkeley National Laboratory

Phenix

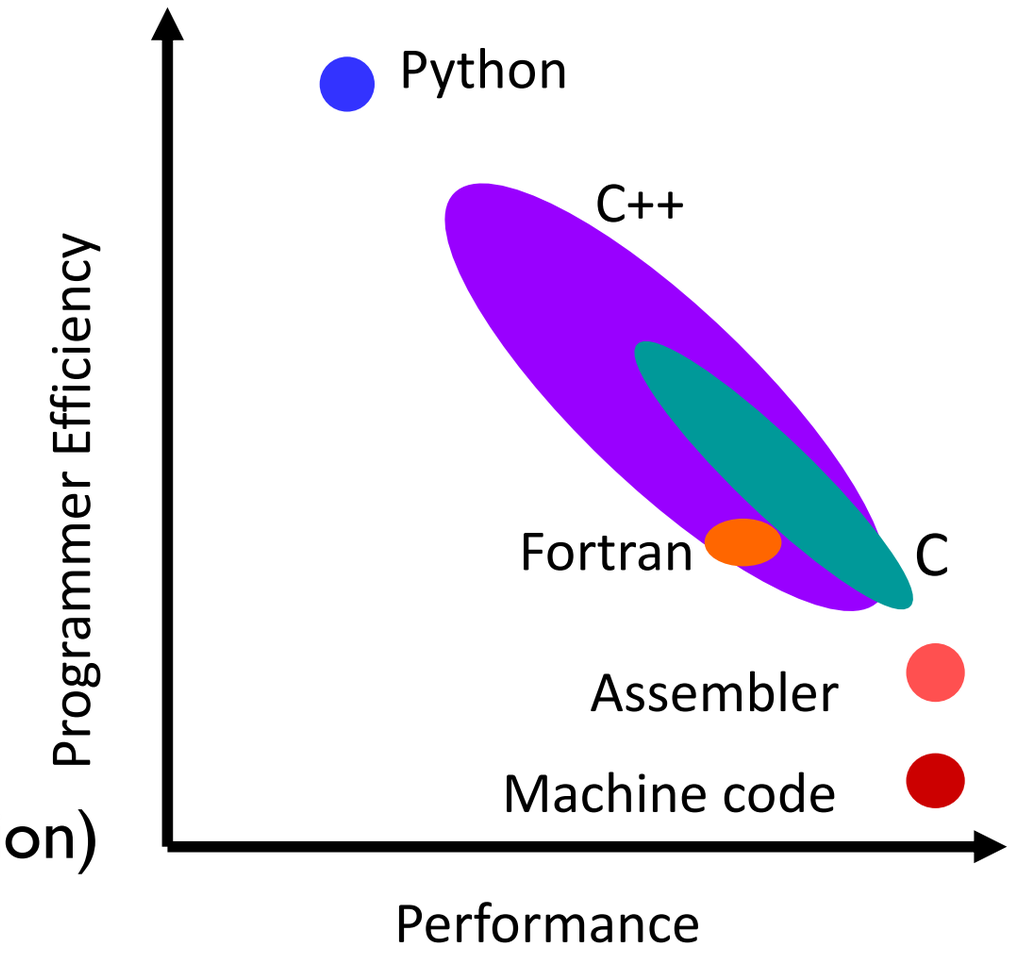
THE UNIVERSITY OF CALIFORNIA · BERKELEY · 1868

Choice of implementation languages in 2000

Desired features:

- Maintainability
- Reusability
- Modularity

- Python:
  - +Very high-level programming, OOP
  - +Easy to use
  - +Fast development cycle (no compilation)
  - -Too slow for certain tasks
- C++:
  - +High-level or medium-level programming, OOP
  - -Hard to use (many arcane details)
  - -Slow development cycle (needs compilation)
  - +Much faster than Python

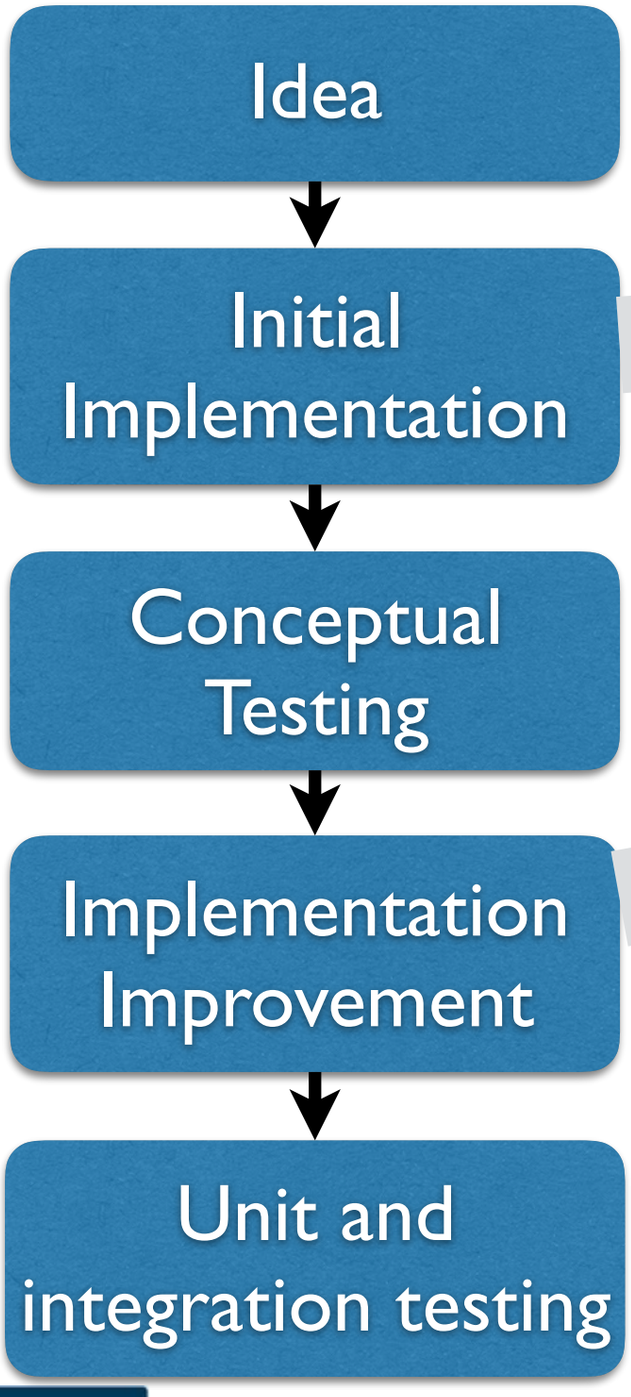


BERKELEY LAB
Lawrence Berkeley National Laboratory

Phenix

THE UNIVERSITY OF CALIFORNIA BERKELEY 1868

Development workflow in hybrid environment



It is more important to obtain a working implementation as fast as possible to be able to discard unsuccessful ideas. Better to use Python.

It is more important to be able speed up the initial implementation if needed. Ability to easily switch to C++ when needed is crucial.

Neither Python nor C++ alone can provide both fast initial implementation and good opportunities for speeding up later.

BERKELEY LAB
Lawrence Berkeley National Laboratory

Phenix

THE UNIVERSITY OF CALIFORNIA BERKELEY 1868

Boost.Python connects C++ and Python

- Mature and flexible library to bridge C++ and Python code
- Supports Python 3
- C++ objects are accessed directly, without making a copy
- C++ objects can be extended with Python-coded methods
- Supports many compilers (including very old ones)
- Possible to wrap CUDA code

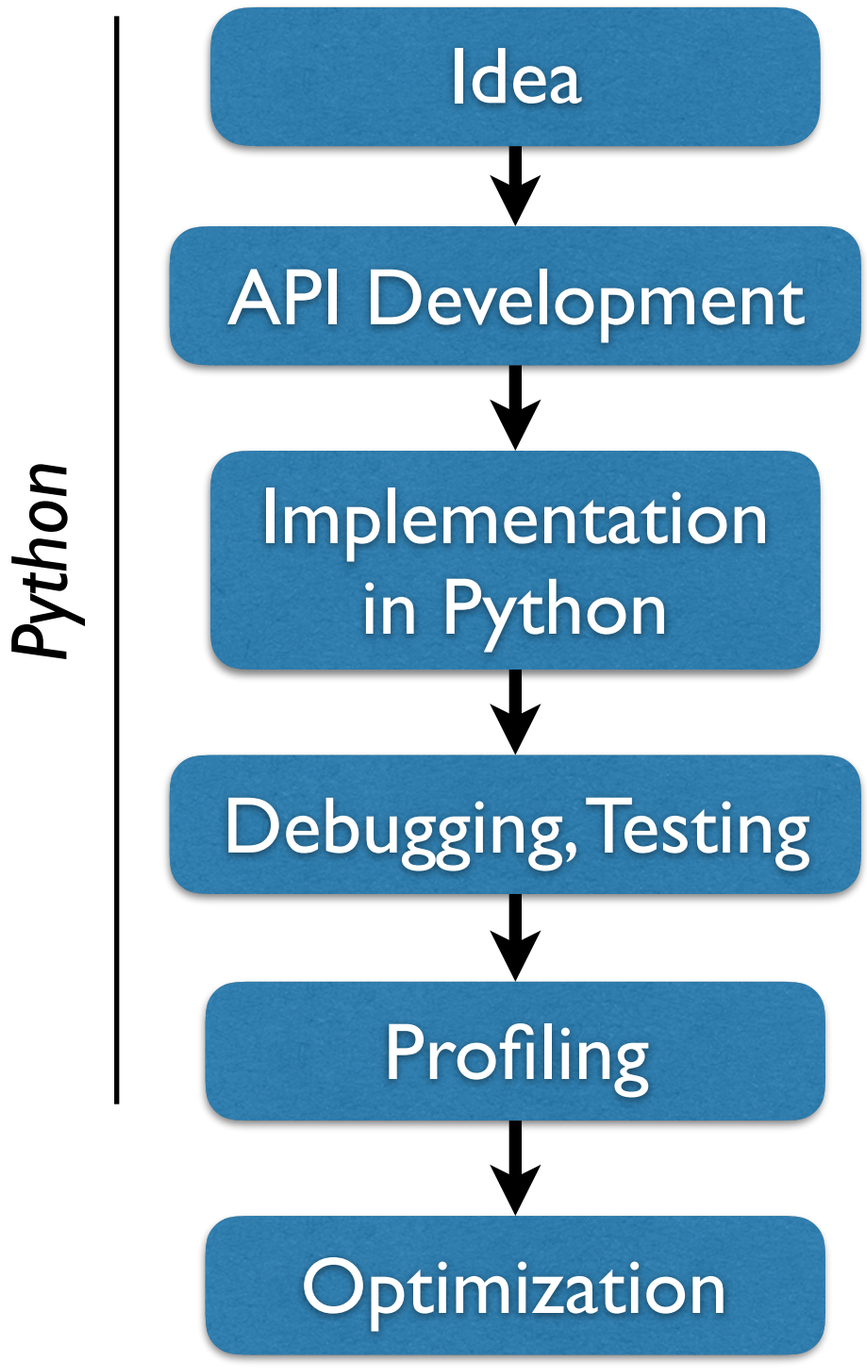
Abrahams D, Grosse-Kunstleve RW: Building Hybrid Systems with Boost.Python. *C/C++ Users Journal* 2003, **21**:29-36.

BERKELEY LAB
Lawrence Berkeley National Laboratory

Phenix

THE UNIVERSITY OF CALIFORNIA · BERKELEY · 1868

Idea Implementation in Practice



In many cases the performance of Python is sufficient

If the performance of C++ is necessary for optimizing part of program:

Time spent on API is not lost – whole class could be rewritten in C++ with exactly the same API

Time spent on writing testing is not lost, the same tests will work

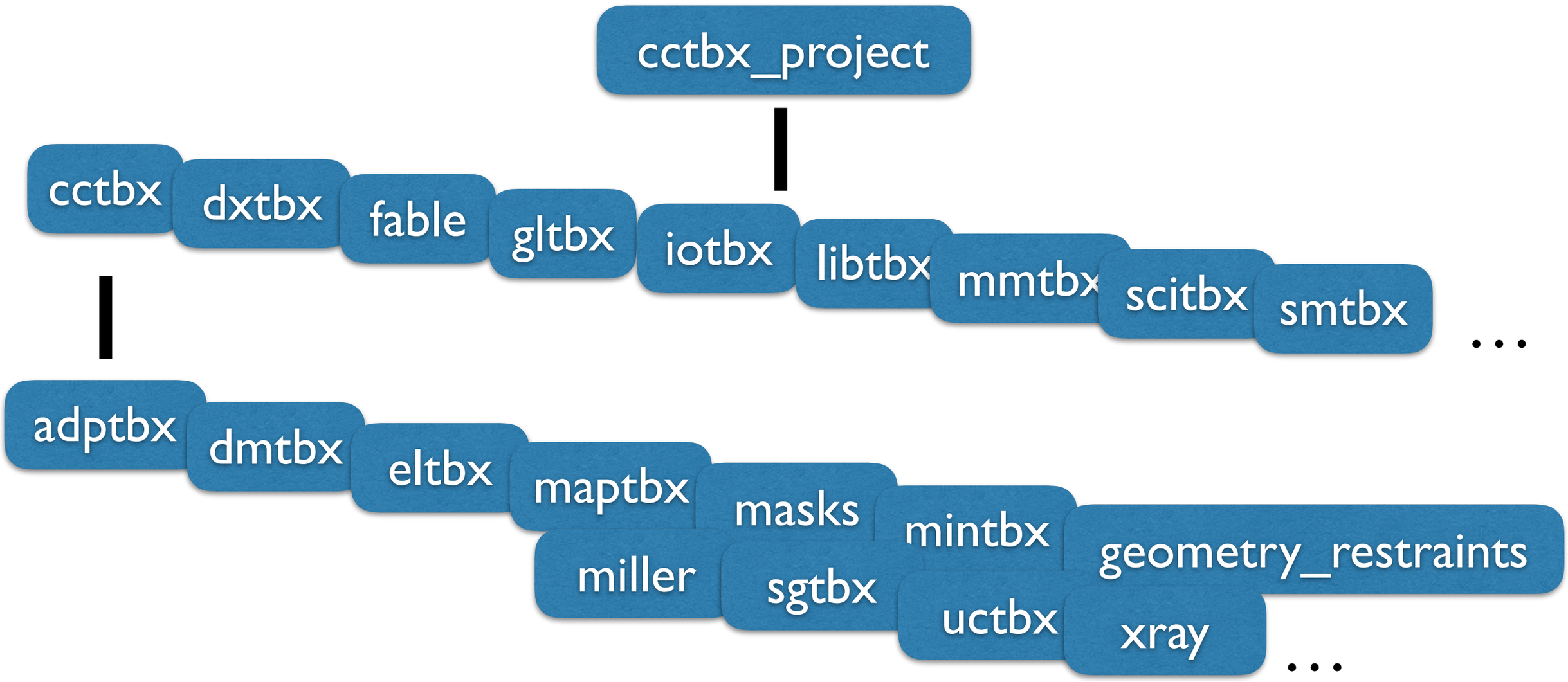
Often only one method of a class needs to be rewritten in C++, the rest remains in Python

BERKELEY LAB
Lawrence Berkeley National Laboratory

Phenix

THE UNIVERSITY OF CALIFORNIA · BERKELEY · 1868

Brief Tour of the cctbx



BERKELEY LAB
Lawrence Berkeley National Laboratory

Phenix

THE UNIVERSITY OF CALIFORNIA BERKELEY 1868

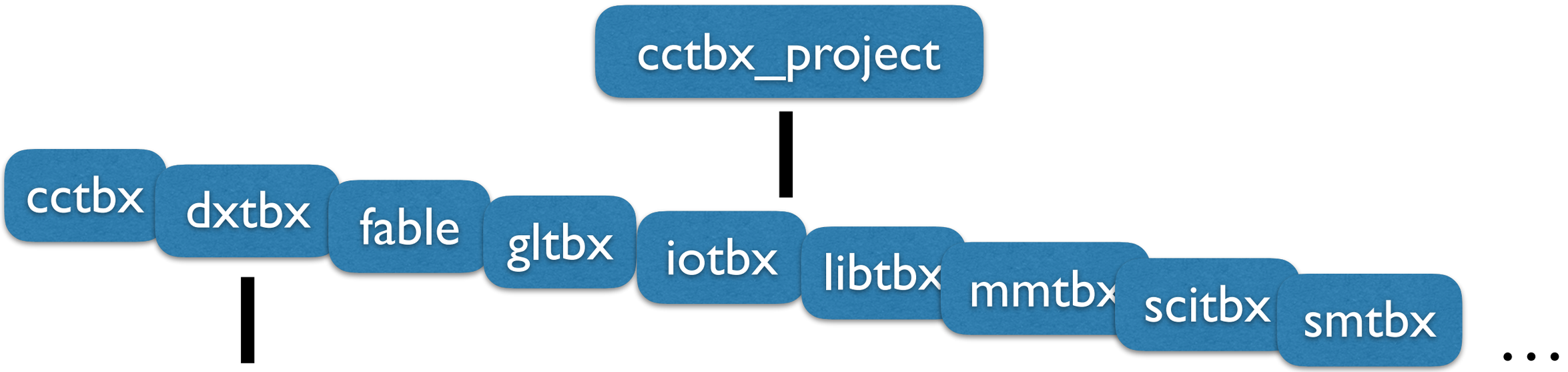
cctbx module

- uctbx – Unit cells
- sgtbx – Space groups
- miller – Structure factor algebra incl. selections and binning
- adptbx – Atomic anisotropic displacement parameters
- eltbx – Scattering factors, ionic radii, etc.
- xray – Structure factor toolbox incl. calc. of gradients
- mintbx – Minimization toolbox
- dmtbx – triplet generation for direct methods
- geometry_restraints – bonds, angles, etc.
- Euclidean model matching
- Generic map handling algorithms
- ….

BERKELEY LAB
Lawrence Berkeley National Laboratory

Phenix

Brief Tour of the cctbx



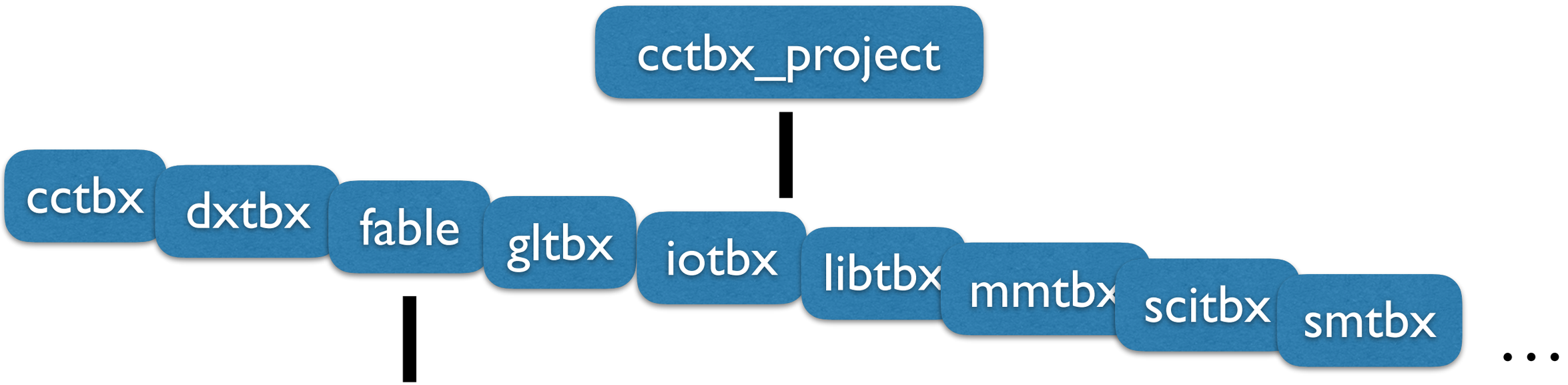
Diffraction Image Toolbox

BERKELEY LAB
Lawrence Berkeley National Laboratory

Phenix

THE UNIVERSITY OF CALIFORNIA BERKELEY 1868

Brief Tour of the cctbx



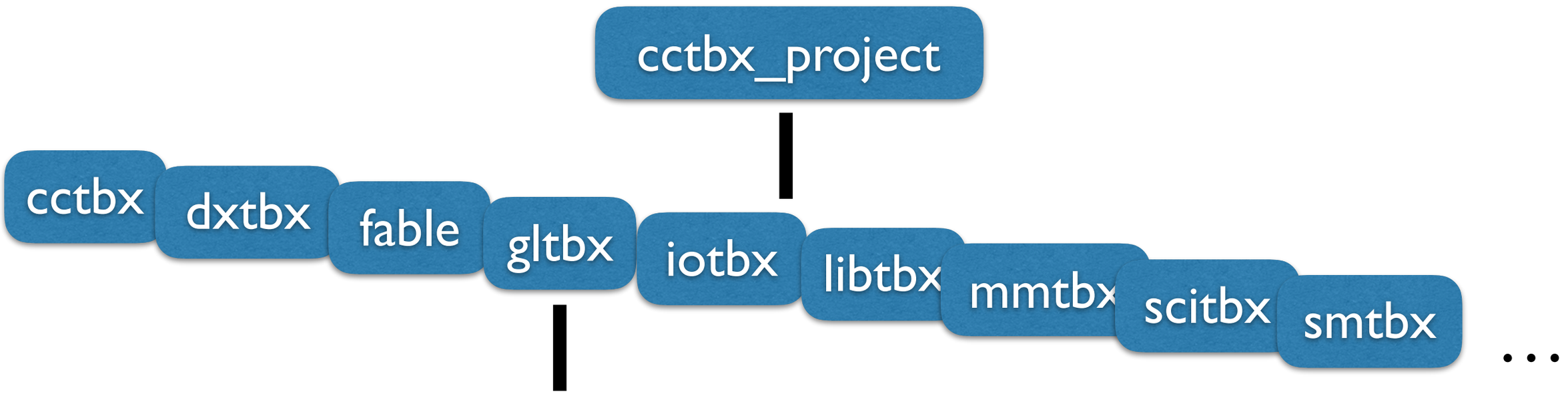
Automated Fortran77
to C++ conversion

BERKELEY LAB
Lawrence Berkeley National Laboratory

Phenix

THE UNIVERSITY OF CALIFORNIA BERKELEY 1868

Brief Tour of the cctbx



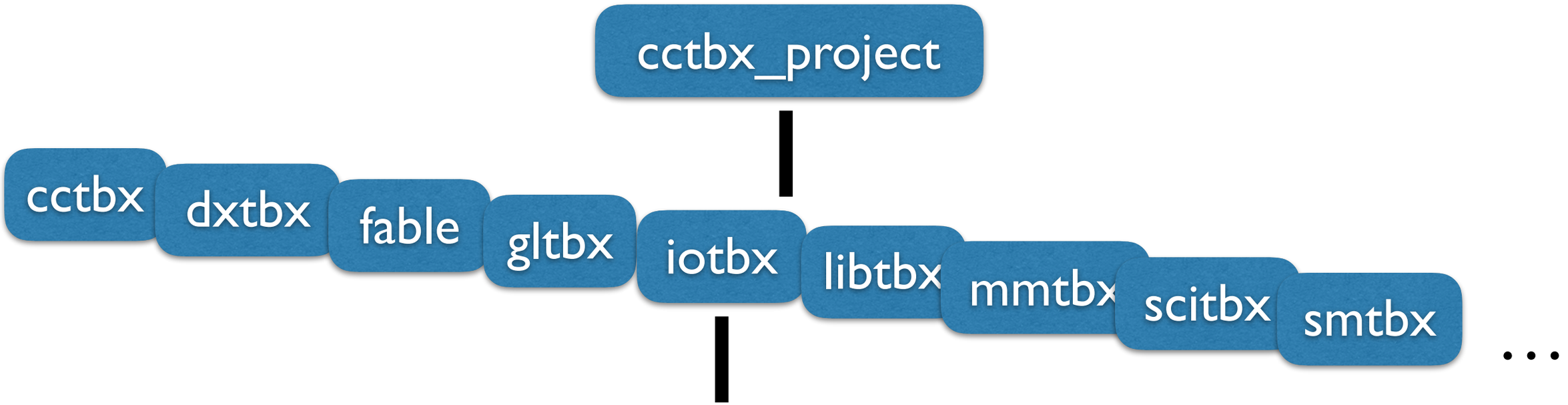
Graphics (GL)
toolbox

BERKELEY LAB
Lawrence Berkeley National Laboratory

Phenix

THE UNIVERSITY OF CALIFORNIA BERKELEY 1868

Brief Tour of the cctbx



Input/output toolbox

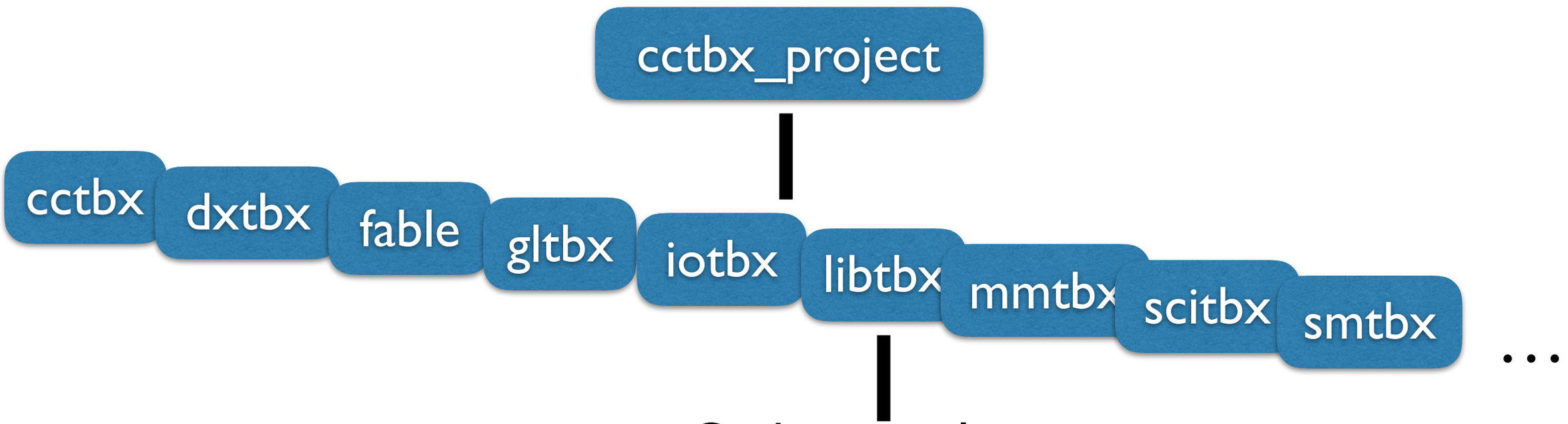
BERKELEY LAB
Lawrence Berkeley National Laboratory

Phenix

File Formats

- The cctbx supports a number of established input/ output formats
  - iotbx:
    - MTZ, CNS, XDS, SHELX, d*Trek, scalepack …
    - iotbx.cif
      - (mm)CIF
- The working format, isn't a format:
  - Object serialization using the Python pickle mechanism

Brief Tour of the cctbx



Code compilation
and installation

BERKELEY LAB
Lawrence Berkeley National Laboratory

Phenix

THE UNIVERSITY OF CALIFORNIA BERKELEY 1868

libtbx Module

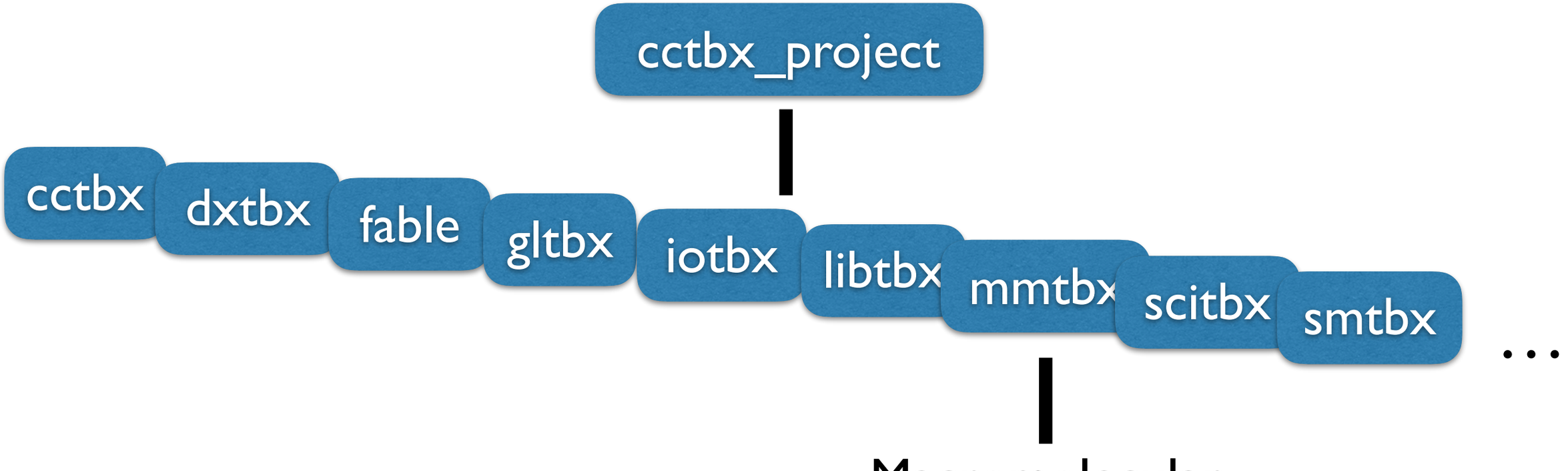
- Thin wrapper around SCons (<500 lines of Python code)
  - python $HOME/libtbx/configure.py scitbx cctbx
  - source setpaths.csh
  - libtbx_scons –j 12 .
- Scons – Software Construction Tool
  - Pure Python
  - No new syntax to learn
  - Completely replaces make
  - Uses MD5 check-sums instead of time-stamps
    - Check-sums include the compiler/linker options
  - Automatic global dependency analysis
  - Optimal utilization of multi-CPU machines (parallel compilation)

BERKELEY LAB
Lawrence Berkeley National Laboratory

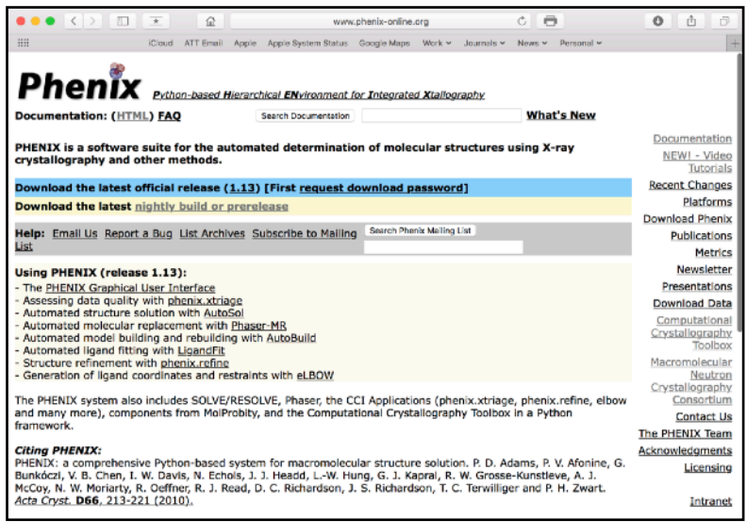
Phenix

THE UNIVERSITY OF CALIFORNIA BERKELEY 1868

Brief Tour of the cctbx



Macromolecular toolbox

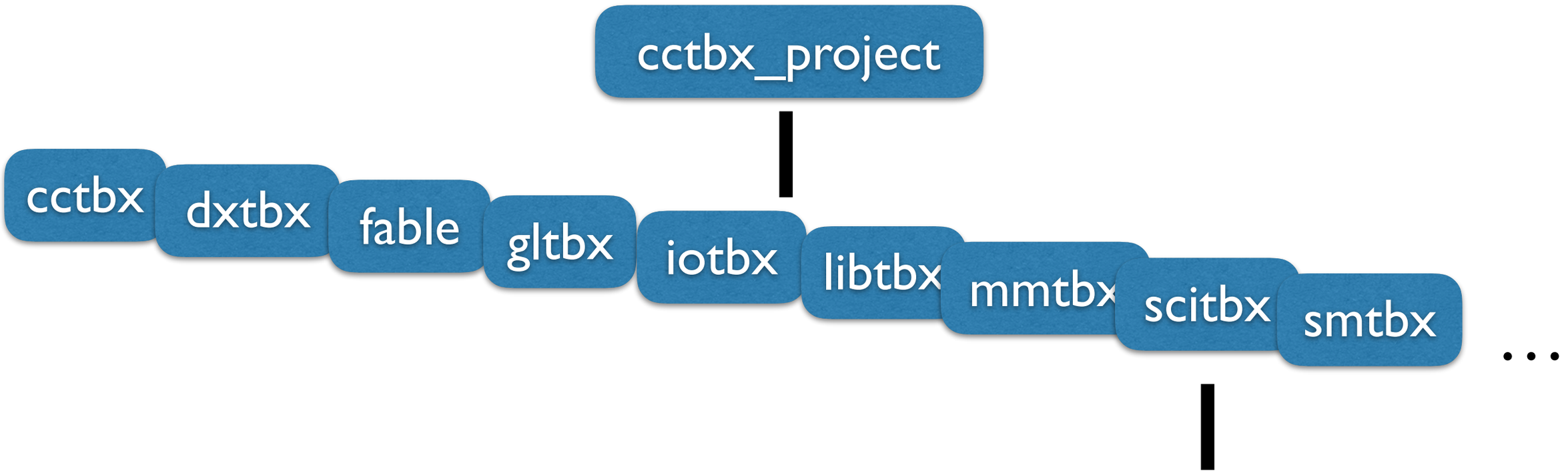


Phenix

BERKELEY LAB
Lawrence Berkeley National Laboratory

THE UNIVERSITY OF CALIFORNIA BERKELEY 1868

Brief Tour of the cctbx



Scientific computing toolbox

BERKELEY LAB
Lawrence Berkeley National Laboratory

Phenix

scitbx Module

- scitbx/array_family – Array family for scientific applications
- scitbx/fftpack – Fast Fourier transform toolbox
  - Port of FFTPACK (Fortran)
  - Pure, generic C++
- scitbx/lbfgs – LBFGS conjugate gradient minimizer
  - Port of Fortran LBFGS
  - Pure, generic C++ incl. Exception Handling

BERKELEY LAB
Lawrence Berkeley National Laboratory

Phenix

THE UNIVERSITY OF CALIFORNIA · BERKELEY · 1868

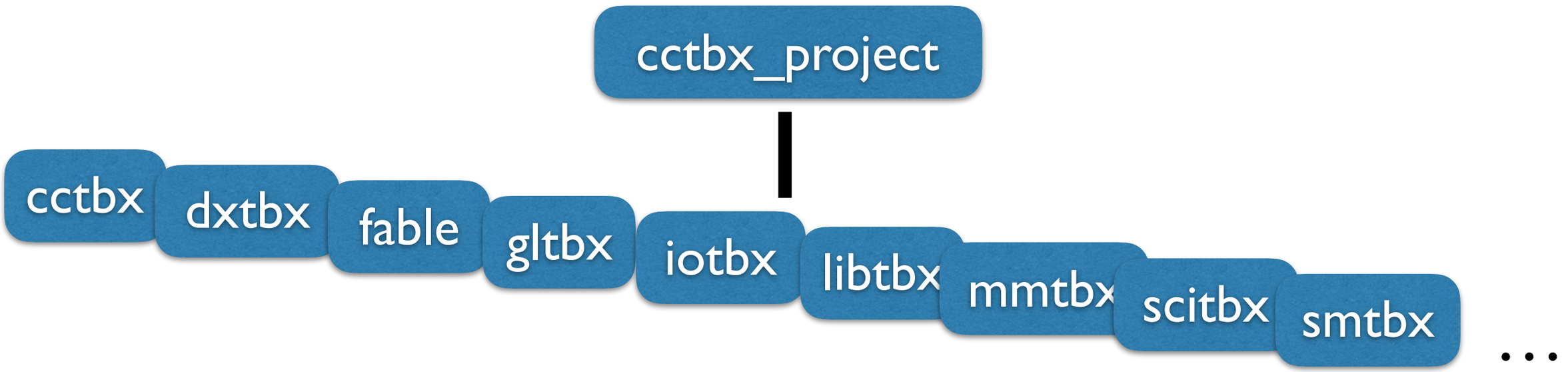
scitbx array family

- Comprehensive and uniform array family:
  - Selection of memory management models
    - Types: tiny, small, shared
  - Access scheme = parameter
    - Type: versa
- Algebras (+,-,*,/,sin,floor,etc.)
- Python: from scitbx.array_family import flex

BERKELEY LAB
Lawrence Berkeley National Laboratory

Phenix

Brief Tour of the cctbx



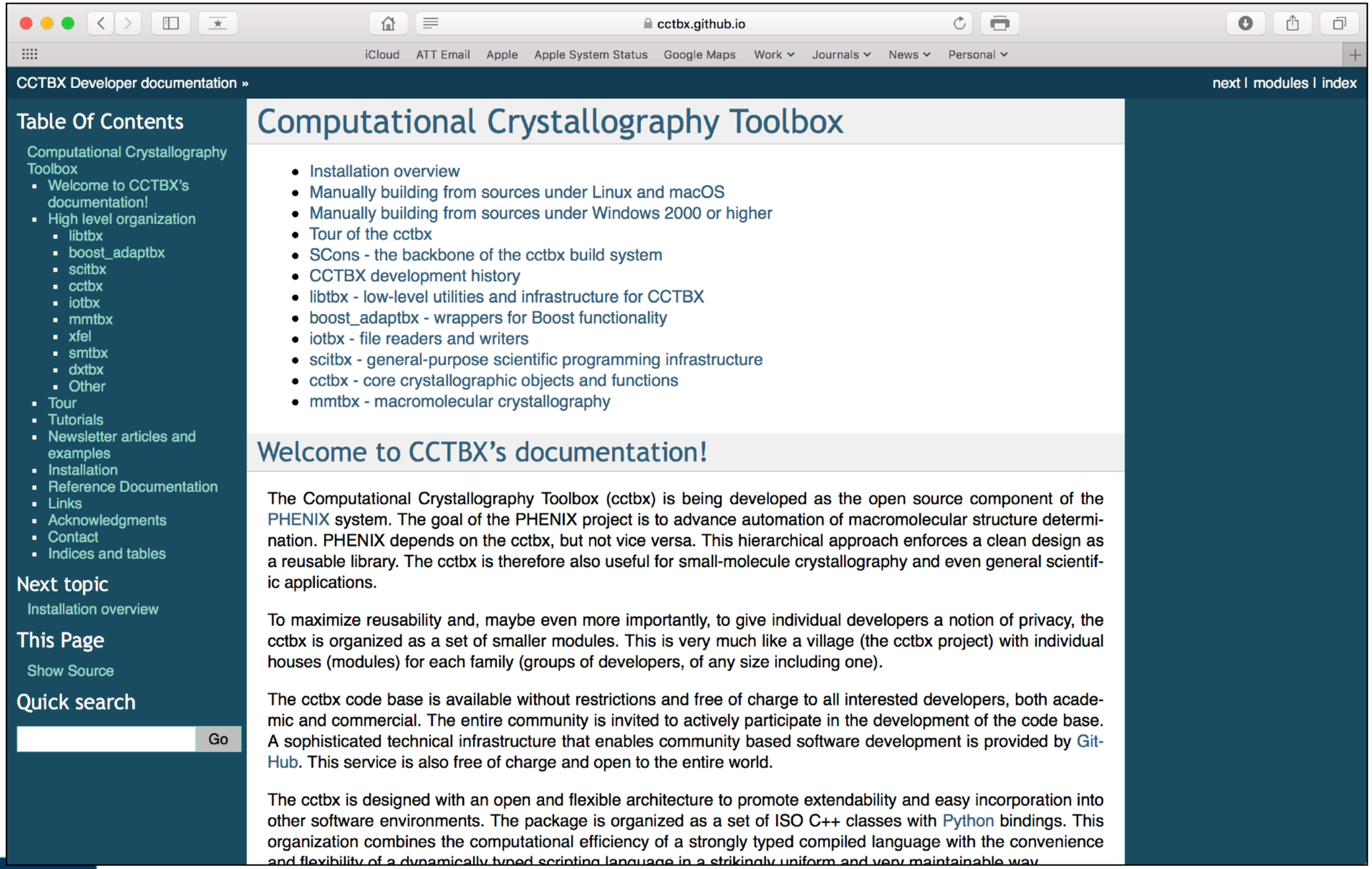
Small molecule toolbox

BERKELEY LAB
Lawrence Berkeley National Laboratory

Phenix

THE UNIVERSITY OF CALIFORNIA BERKELEY 1868

cctbx.github.io

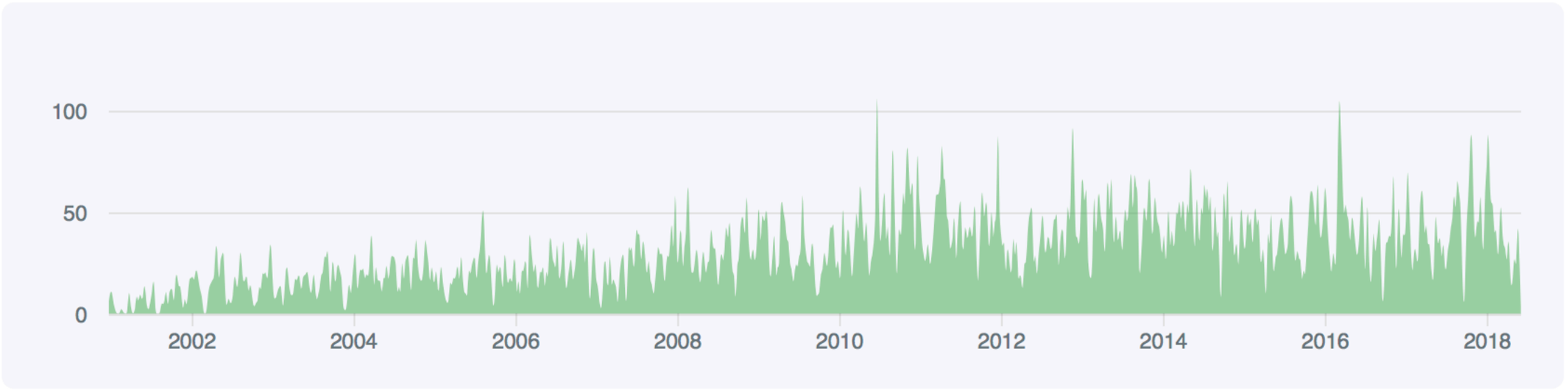


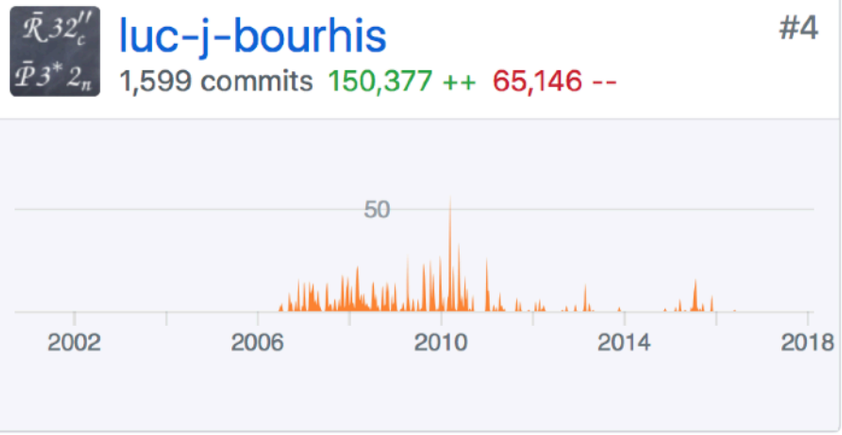
BERKELEY LAB
Lawrence Berkeley National Laboratory

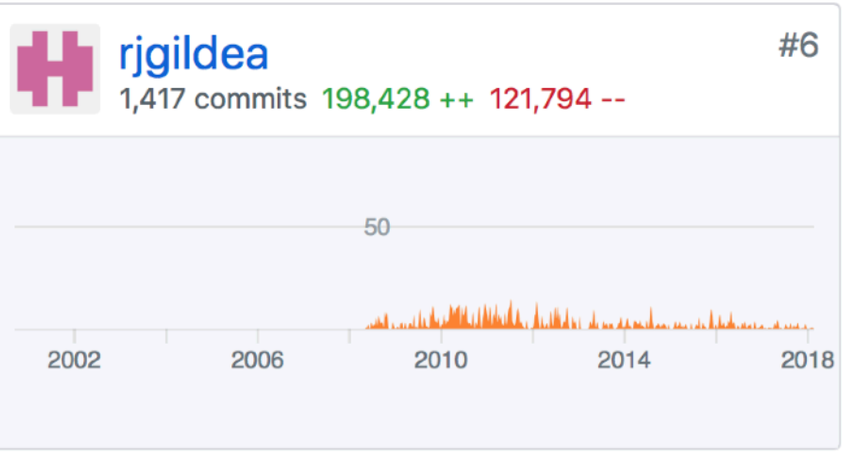
Phenix

THE UNIVERSITY OF CALIFORNIA · BERKELEY · 1868

Active Development







BERKELEY LAB
Lawrence Berkeley National Laboratory

Phenix

THE UNIVERSITY OF CALIFORNIA BERKELEY 1868

Contributing to the cctbx

- Contributors are actively encouraged
  - Discipline is required if changes are to core code (other packages need to be functional)
- Github provides mechanisms for code management and requests for changes
- There is guidance for developers:
  - https://github.com/cctbx/cctbx_project/wiki/cctbx-Developer-Guidance
- There are ground rules for contributing:
  - https://github.com/cctbx/cctbx_project/wiki/cctbx-contributors-guide

Conclusions

- There has been good progress in developing tools for neutron crystallography
  - But there is still much to be done
- New facilities can provide environments for structure solution and data curation
  - Neutron crystallography needs accurate capture of metadata for deposition
- The cctbx provides many tools for data analysis, structure solution and data management
- The cctbx has long term support and should provide a solid platform for future developments
- Don’t reinvent the wheel

BERKELEY LAB
Lawrence Berkeley National Laboratory

Phenix

THE UNIVERSITY OF CALIFORNIA BERKELEY 1868

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, Tristan Croll, Laurent Storoni, Gabor Bunkoczi, Robert Oeffner
- **Duke University**
  - Jane Richardson & David Richardson, Ian Davis, Vincent Chen, Jeff Headd, Christopher Williams, Bryan Arendall, Laura Murray, Gary Kapral, Dan Keedy, Swati Jain, Bradley Hintze, Lindsay Deis, Lizbeth Videau
- **University of Washington**
  - Frank DiMaio, David Baker
- **Oak Ridge National Laboratory**
  - Vickie Lynch, Brendan Sullivan, Marat Mustyakimov, Paul Langan
- **Others**
  - Alexandre Urzhumtsev & Vladimir Lunin
  - Garib Murshudov & Alexi Vagin
  - Kevin Cowtan, Paul Emsley, Bernhard Lohkamp
  - David Abrahams
  - PHENIX Testers & Users
- **Funding:**
  - NIH/NIGMS:
    - *P01GM063210, P50GM062412, P01GM064692, R01GM071939*
  - Lawrence Berkeley Laboratory
  - PHENIX Industrial Consortium

BERKELEY LAB
Lawrence Berkeley National Laboratory

Phenix