Quick Facts About PHENIX, CCI APPS and Structure Refinement with phenix.refine

October – 2006

A collaboration between several groups:

CCI APPS	**Computational Crystallography Initiative (LBNL)** -*Paul Adams*, Ralf Grosse-Kunstleve, Pavel Afonine -Nigel Moriarty, Nicholas Sauter, Peter Zwart
SOLVE / RESOLVE	**Los Alamos National Lab (LANL)** -*Tom Terwilliger*, *Li-Wei Hung,* Thiru Radhakannan
PHASER	**Cambridge University** -*Randy Read*, Airlie McCoy, Laurent Storoni -Hamsapriye
TEXTAL	**Texas A&M University** -*Tom Ioerger*, *Jim Sacchettini*, Kreshna Gopal, Tod Romo -Reetal Pai, Erik McKee, Lalji Kanbi
MolProbity / REDUCE	**Duke University** - *Jane Richardson, David Richardson,* Ian Davis

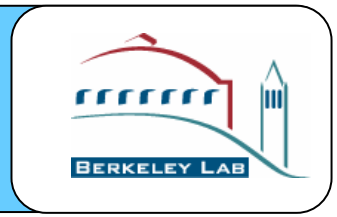
Project Director - Paul Adams

Tools for automated structure solution:

- **Automated analysis of data quality and integrity**
  - mmtbx.xtriage
- **Rapid substructure determination**
  - Hybrid Substructure Search (HySS)
- **Scoring of substructure enantiomorphs**
  - Analysis of maps calculated with each hand
- **Phasing**
  - Maximum likelihood – SOLVE, PHASER
- **Substructure completion**
  - Difference Fourier methods, soon PHASER log likelihood gradient maps
- **Density modification**
  - Statistical density modification (RESOLVE)
- **Automated model building**
  - Pattern matching methods (RESOLVE or TEXTAL)
  - Building and refinement of multiple models
- **Structure refinement**
  - phenix.refine (likelihood, minimization, annealing, ADP, TLS, water picking)
- **Structure validation**
  - MolProbity, Reduce (bad conformations and contacts, hydrogens)

CCI APPS: Subset of PHENIX components developed by the Computational Crystallography Initiative (CCI) at LBNL:

Computational Crystallography Initiative (LBNL)
-*Paul Adams*, Ralf Grosse-Kunstleve, Pavel Afonine
-Nigel Moriarty, Nicholas Sauter, Peter Zwart



CCI APPS: Command-line tools only. "One click" installation on most of platforms. Does not require installation of anything else. Can be obtained and installed separately from PHENIX package.

CCI APPS include:

phenix.refine	– Highly-automated state-of-the-art structure refinement
HySS	– Highly-automated location of anomalous scatterers
eLBOW	– Parameter and topology files builder for novel ligands
Xtriage	– Comprehensive analysis of experimental data

Below we focus on phenix.refine. For information about other components please refer to separate documentation.

► Refinement using X-ray, neutron or joint X-ray + neutron data. Any combination of the following refinement scenarios can be run in one step:

- Individual coordinates
- Individual isotropic ADP (Atomic Displacement Parameters)
- Rigid body
- Grouped isotropic ADP
- TLS
- Simulated Annealing (SA)
- NCS restraints on coordinates and ADP
- Occupancy
- Grouped occupancy

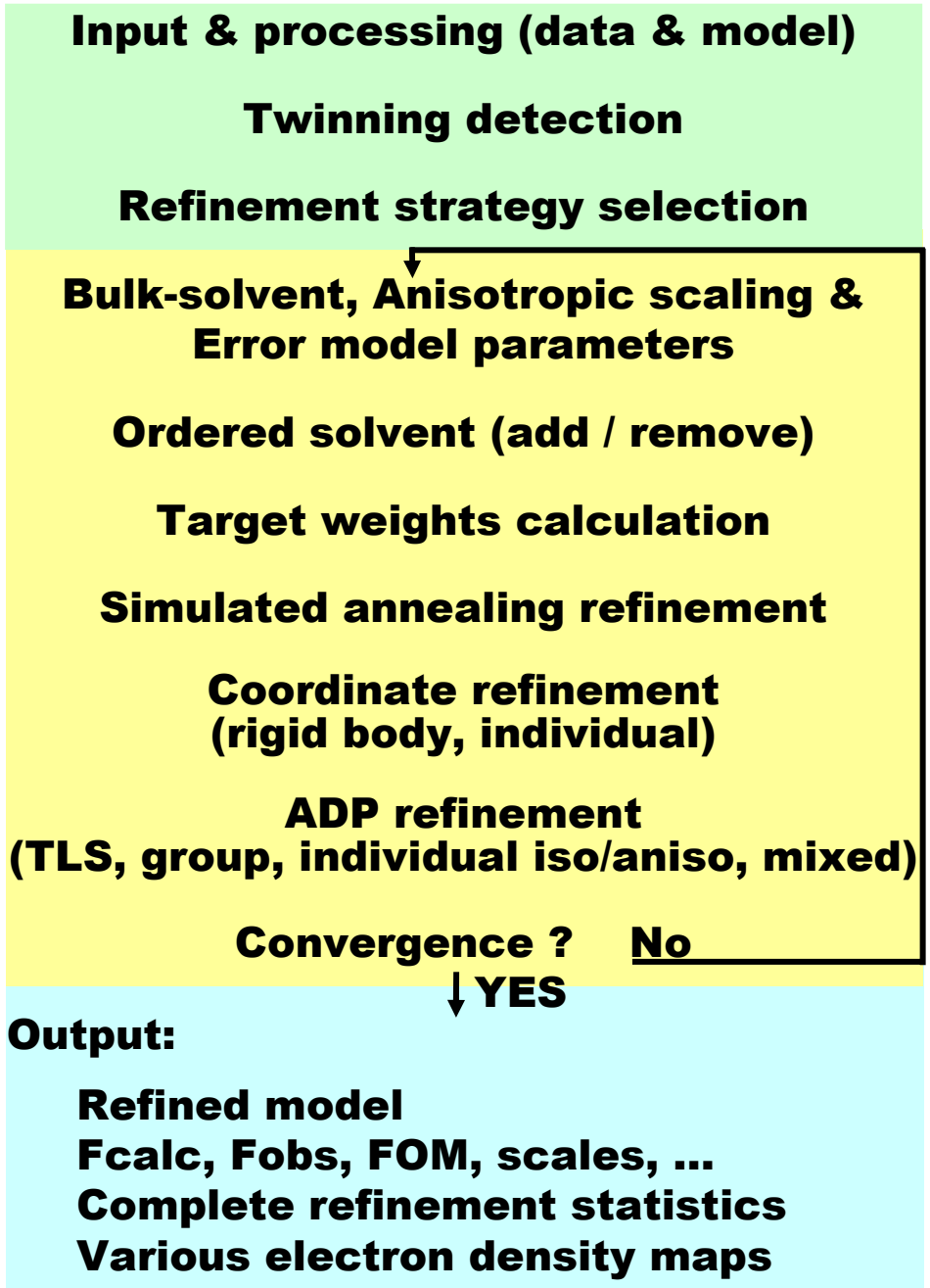
► Automatic water picking / removal combined with refinement

► Examples of more complex refinement scenarios:

- TLS + SA + individual coordinates and isotropic ADP
- SA + water picking + individual coordinates and isotropic ADP
- Rigid body + SA + individual coordinates and grouped ADP

► Automatic recognition of most of commonly used reflection file formats (including MTZ, SHELX, CNS/Xplor and more)

- Bulk solvent correction and anisotropic scaling
- Multiple target functions: LS, ML, MLHL, actively devel.: SAD and specific twin targets
- Various electron density map calculations, e.g. σ_A weighted maps
- Simple structure factor calculation
- Comprehensive model and data statistics
- Various utilities for research and development:
  - shaking initial coordinates and iso-/anisotropic ADP
  - setting ADP to given value (user defined or Wilson B)
  - shifting initial model as a rigid body
  - refinement using fake (calculated) Fobs (with simulated bulk solvent)
- Full control over refinement (> 200 parameters available for adjustment)
- Twinning detection (refinement with twinned data actively developed)
- Sophisticated atom selection tools
- Restraints using Monomer Library. For unknown ligands: easy and quick generation of topology and parameter files with eLBOW
- Automatic gradient-based target weights calculation
- FFT and direct summation based refinement


Input & processing (data & model)
Twinning detection
Refinement strategy selection
Bulk-solvent, Anisotropic scaling &
Error model parameters
Ordered solvent (add / remove)
Target weights calculation
Simulated annealing refinement
Coordinate refinement
(rigid body, individual)
ADP refinement
(TLS, group, individual iso/aniso, mixed)
Convergence ?
No
YES
Output:
Refined model
Fcalc, Fobs, FOM, scales, ...
Complete refinement statistics
Various electron density maps

Running phenix.refine from the command line:

Refinement of individual coordinates and B-factors:

```
% phenix.refine model.pdb data.hkl
```

Same as above plus Simulated Annealing (SA) and water picking:

```
% phenix.refine model.pdb data.hkl simulated_annealing=true ordered_solvent=true
```

Rigid-body + SA + grouped B-factors refinement:

```
% phenix.refine model.pdb data.hkl strategy=rigid_body+group_adp simulated_annealing=true
```

TLS + SA + individual coordinates and B-factors refinement:

```
% phenix.refine model.pdb data.hkl strategy=tls+individual_sites+individual_adp \
simulated_annealing=true
```

Refinement of individual coordinates and B-factors using neutron data:

```
% phenix.refine model.pdb data.hkl scattering_dictionary=neutron
```

To see all parameters (more than 200):

```
% phenix.refine model.pdb data.hkl  --show_defaults=ALL
```

See the documentation for more examples: http://www.phenix-online.org/

Running phenix.refine with parameter file:

% phenix.refine model.pdb data.hkl parameters_file

where the parameter_file contains following lines:

```
refinement.main {
   high_resolution = 2.0
   low_resolution = 15.0
   simulated_annealing = True
   ordered_solvent = True
   number_of_macro_cycles = 5
}
refinement.refine.adp {
    tls = chain A
    tls = chain B
}
```

This is equivalent to typing from the command line:

% phenix.refine model.pdb data.hkl main.high_resolution=2 main.low_resolution=15 simmulated_annealing=true ordered_solvent=True adp.tls="chain A" adp.tls="chain B" main.number_of_macro_cycles=5

Shortcuts are allowed, as long as they are unambiguous:

% phenix.refine model.pdb data.hkl main.high_res=2 main.low_res=15 simulated_annealing=true ordered_s=True adp.tls="chain A" adp.tls="chain B" main.number_of_m=5

When running:

% phenix.refine model.pdb data.hkl

each item in model.pdb is matched against the CCP4 Monomer Library to extract the topology and parameters and to automatically build corresponding restraints.

If model.pdb contains an item not available in CCP4 Monomer Library, e.g. a new ligand: use eLBOW program to generate topology and parameter file for this novel ligand:

% elbow.builder model.pdb --residue=XXX --overwrite --opt

This will produce the file elbow.XXX.model_pdb.cif which can be used for refinement:

% phenix.refine model.pdb data.hkl elbow.XXX.model_pdb.cif

No manual work!

MORE FEATURES, MORE AUTOMATION !

Twinning (detection+refinement: sophisticated target functions)

Composite SA Omit maps

Alternative conformations

Torsion angles parameterization for SA refinement

Occupancy refinement (alternative conformations)

Hydrogens in refinement

SAD likelihood function

Automatic & fast rigid body domain determination

Dual real+reciprocal space refinement (new approach is under development)

Ultra-high resolution refinement (DBE models)

Neutron specific water building

…

- Beta release in July 2006
- Supported on:
  - Linux (developed on Redhat, Fedora)
  - Tru64 Alpha
  - Irix SGI
  - In the future:
    - Windows, Mac OSX
- Freely available to academic (non-profit) groups
- Web download (http://www.phenix-online.org/)
- Consortium for commercial users

- Frequently updated once new features are available, based on bug fixes or feature requests (sometimes several times per month)

- Supported on:
  - Fedora 3, 4, 5
  - IRIX 6.5.21 MIPSpro 7.41
  - Mac OS 10.3.9
  - Red Hat 8.0, 9.0, WS 3
  - SuSE 9.2
  - Compaq Tru64 V5.1
  - Windows 2000/XP

- Freely available to academic (non-profit) groups

- Web download (http://www.phenix-online.org/)

Funding:

- Protein Structure Initiative of the National Institute of General Medical Sciences (NIH), grant P01GM063210

- US Department of Energy under Contract No. DE-AC02-05CH11231

- "Computational Tools for Neutron Protein Crystallography" grant R01GM071939

- PHENIX Industrial Consortium

More information about PHENIX and CCI APPS:

http://www.phenix-online.org/

http://cci.lbl.gov/

Questions, comments, bugs:

help@phenix-online.org

bugs@phenix-online.org

Questions / comments about this document: PAfonine@lbl.gov (Pavel Afonine)