

# *Model Refinement*

**Pavel Afonine**



[phenix-online.org](http://phenix-online.org)



[lbl.gov](http://lbl.gov)

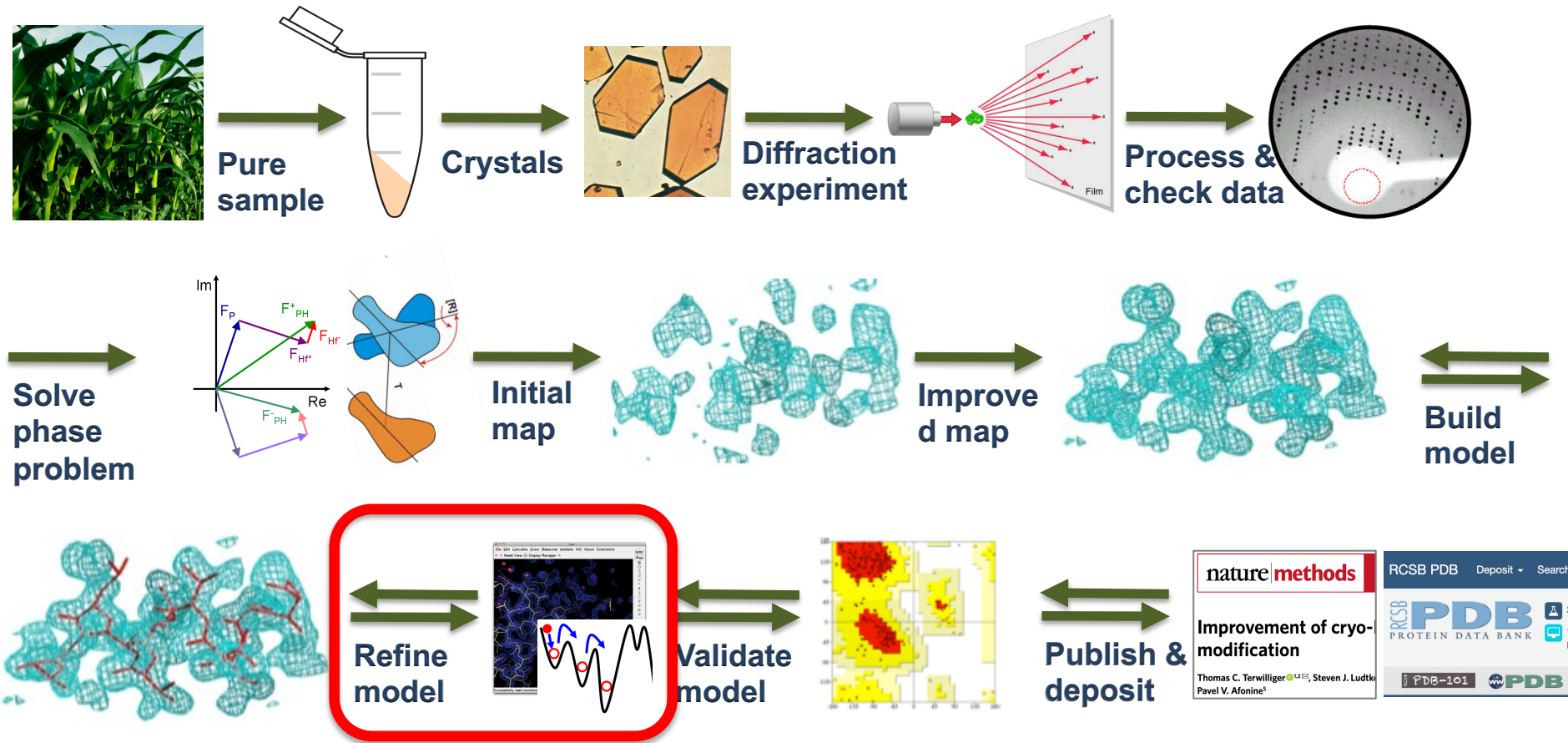


[qrefine.com](http://qrefine.com)

ACA, Denver, Colorado

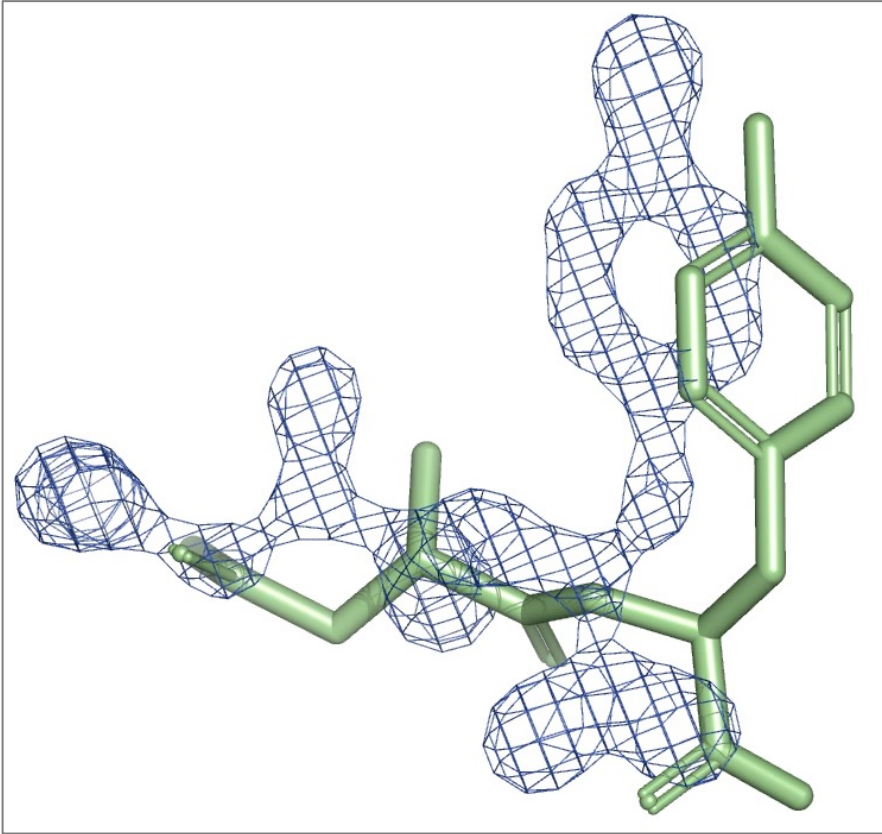
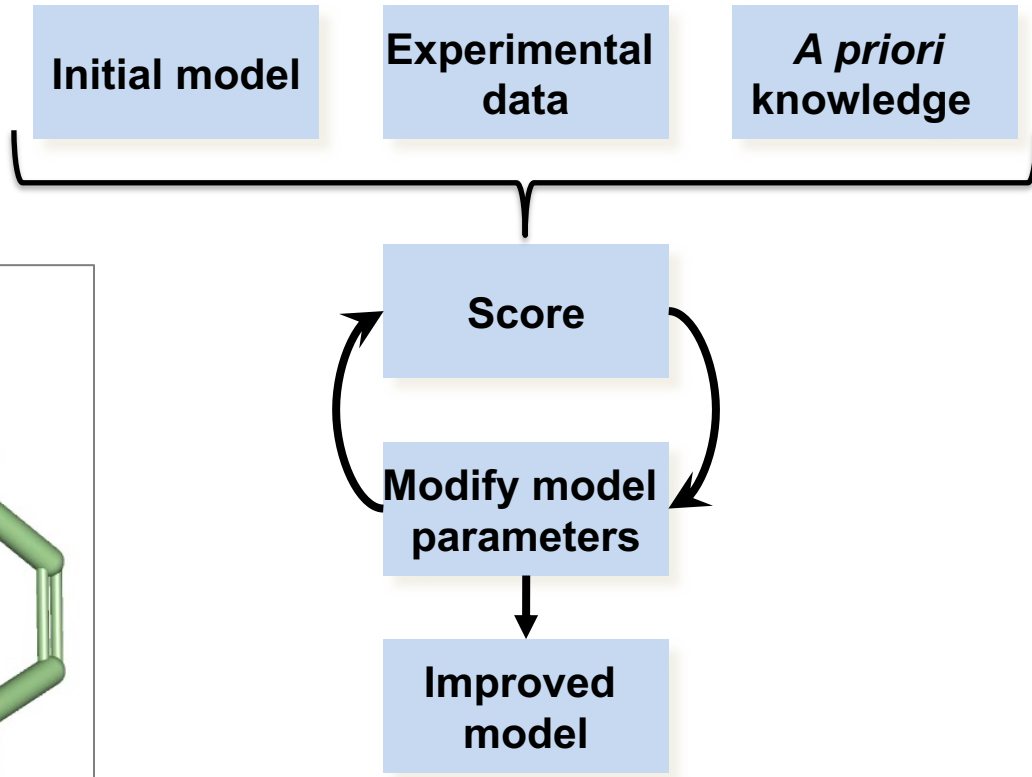
July 7<sup>th</sup> 2024

# Solving structure by crystallography



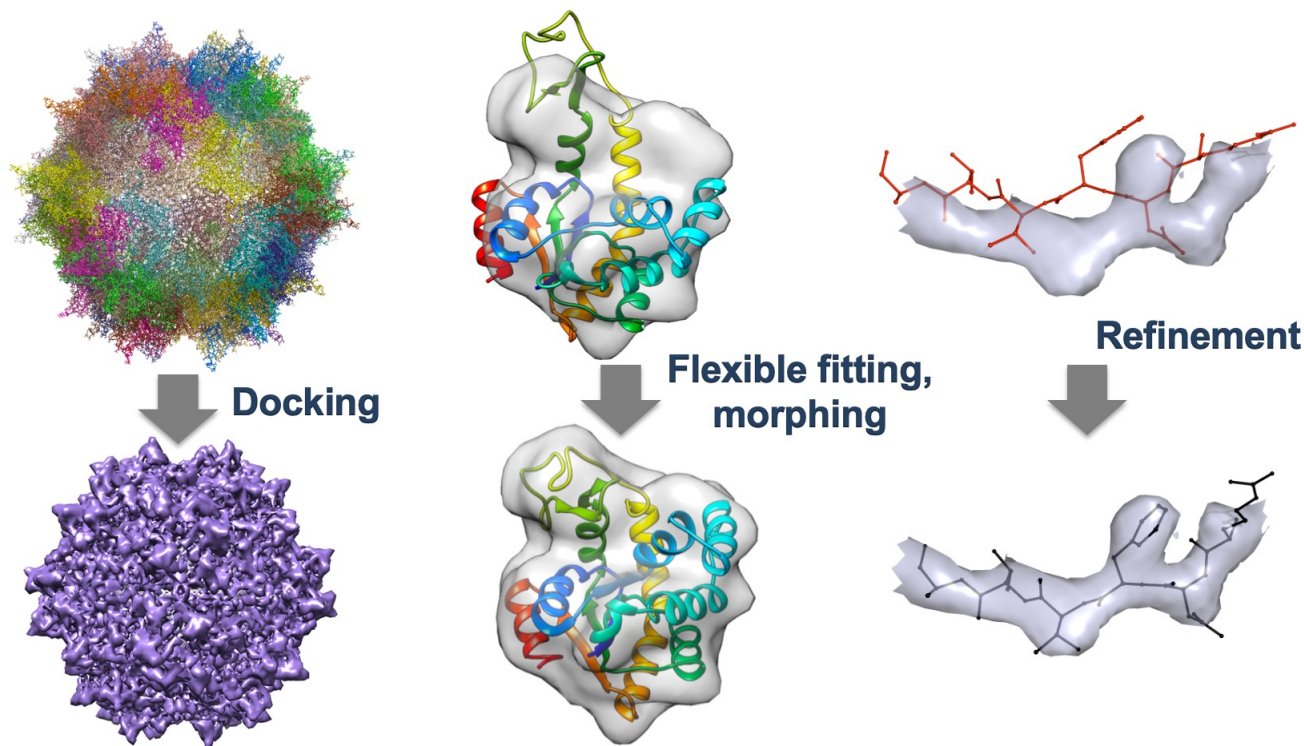
- Process is not as 'linear' as shown
- Each step has numerous sub-steps
- Crystals may not grow or exhibit pathologies
- Stuck solving phase problem

# Model refinement



Optimization process of fitting atomic model parameters to experimental data

# Not all model-to-data fitting is refinement



- Docking, flexible fitting, morphing are **not** refinement
- Refinement is to fine-tune an already fine atomic model
- Refinement does only small changes to the model (within *convergence radius of refinement*,  $\sim 1\text{\AA}$ )

# Solving structure in the past

- Familiar with many software packages (often with “orthogonal” philosophies)
- Mutually incompatible file formats for common data exchange
- Coding experience was a must (typically using arcane languages FORTRAN or C)
- No GUIs. Command line expertise (Unix)
- Reading thick books (no Google, YouTube or ChatGPT!)
- Limited online forums
- Don't expect your questions answered quickly by email
- Slow computers (with sometimes limited access)

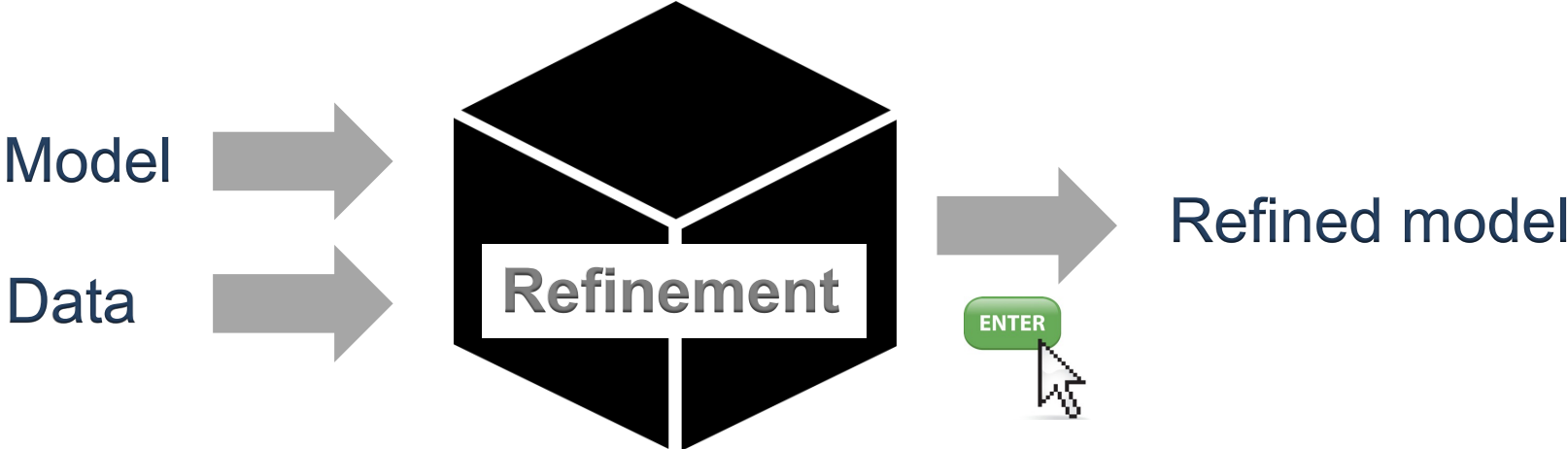
# Solving structure in the past

- From many months to years
  - Spend days on graphics (manual atomic model building)
  - Run computations overnight

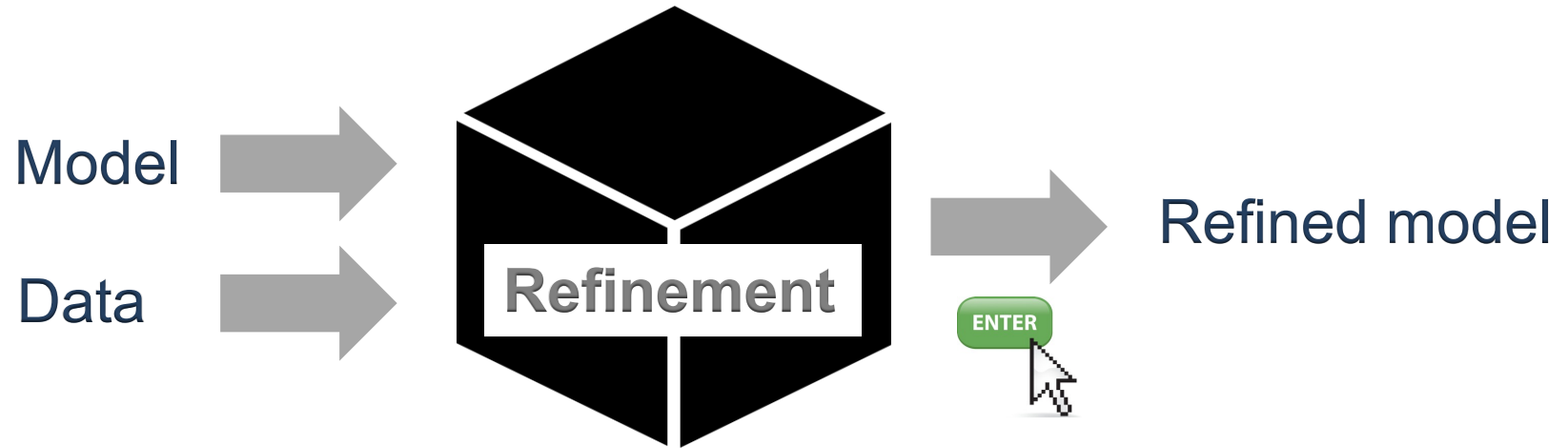


Solving my first structure back in 1997

# Model refinement: black box



# Model refinement: black box



- Does it always work?
- Is it always as easy as *poor model in, better model out*?



# Model refinement: black box

- **No.** Because:
  - Refinement parameterization isn't easy (next slide)
  - Default settings suit most common scenario
    - Typical resolution data, model reasonably fits data
  - Less typical situations need customizations
    - Low or high resolution data
    - Incomplete models
    - Final models
    - AlphaFold predicted models
    - Novel ligands

# Model refinement: lot of stuff to know...

Reference model?

TLS?

Rotamer fixing?

AltLocs?

ADP?

Group B vs individual?

Local minima?

tNCS?

Clashes?

NCS?

IAS?

Weights?

CDL?

SA?

Grid search?



Minimization?

Rama plot restraints?

$f'$  &  $f''$ ?

Hydrogens?

Restraints?

Bulk-Solvent?

Rigid body?

Rama-Z?

Anisotropy?

NQH flips?

SS restraints?

Twinning?

# Model refinement: black box

- What to do when the 'black box' does not work?
- **Your decision-making** is needed (and it is not always easy!)

# Model refinement: decision-making variables

- **Crystal**

- Disorder
- Twinning, tNCS
- Solvent content
- Symmetry

- **Data**

- Resolution
- Errors
- Completeness
- Processing

- **Model**

- Stage
- Source
- Parameterization
- Fit to data

# How you know...

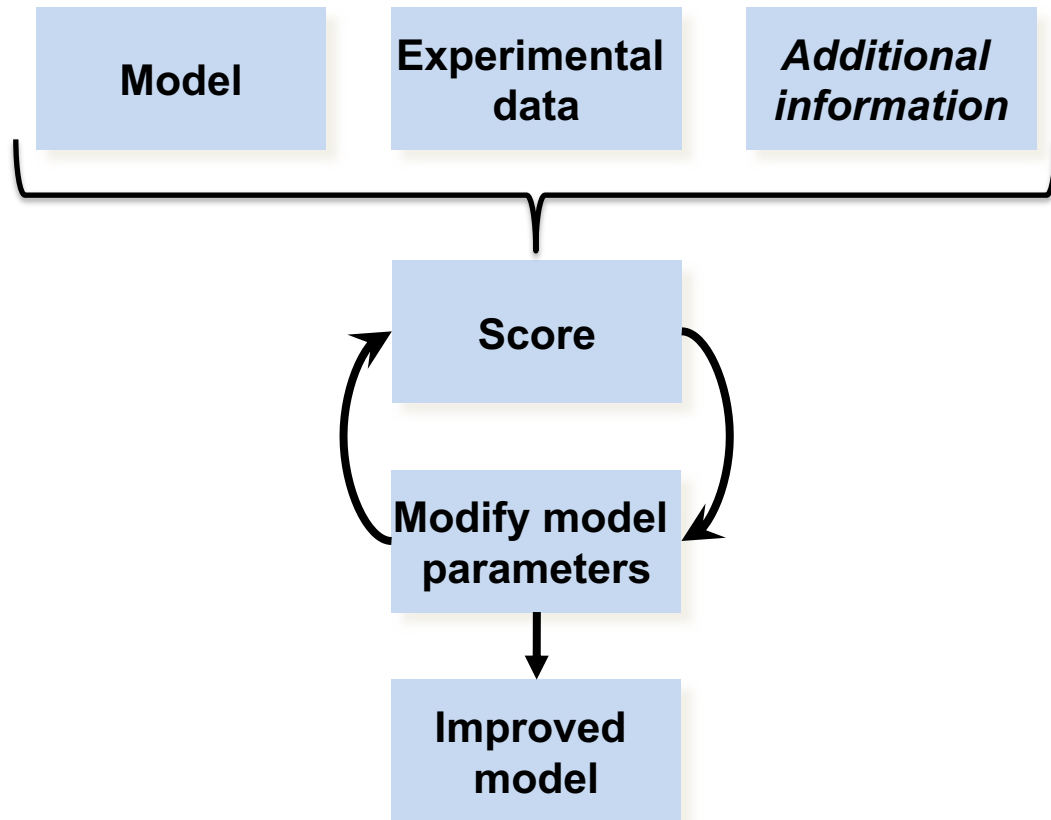
- ... refinement worked ?
- ... you did it correctly ?
- ... the model is good enough to publish ?

# How you know...

- ... refinement worked ?
  - ... you did it correctly ?
  - ... the model you got is good enough to publish ?
- 
- **Do validation!**
- Standard validation protocols are designed to answer these questions**

## **Refinement: a closer look**

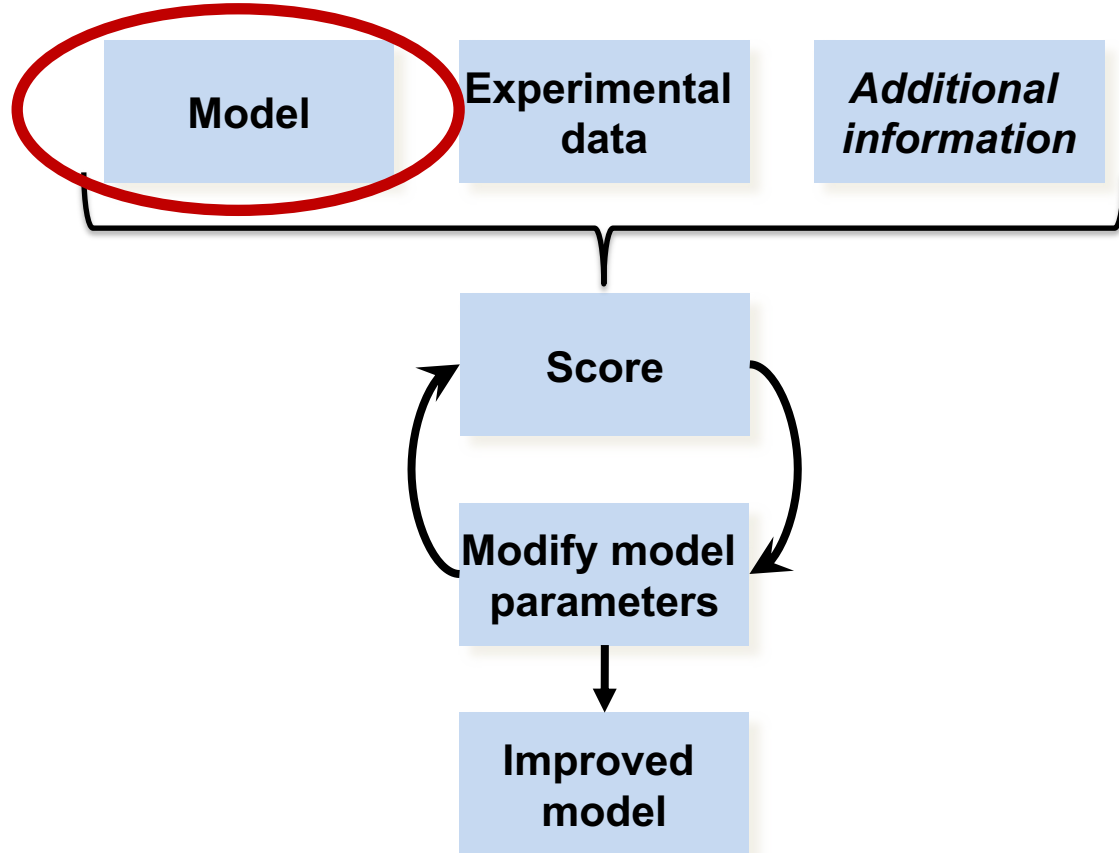
# Model refinement



Refinement – optimization process of fitting model parameters to experimental data

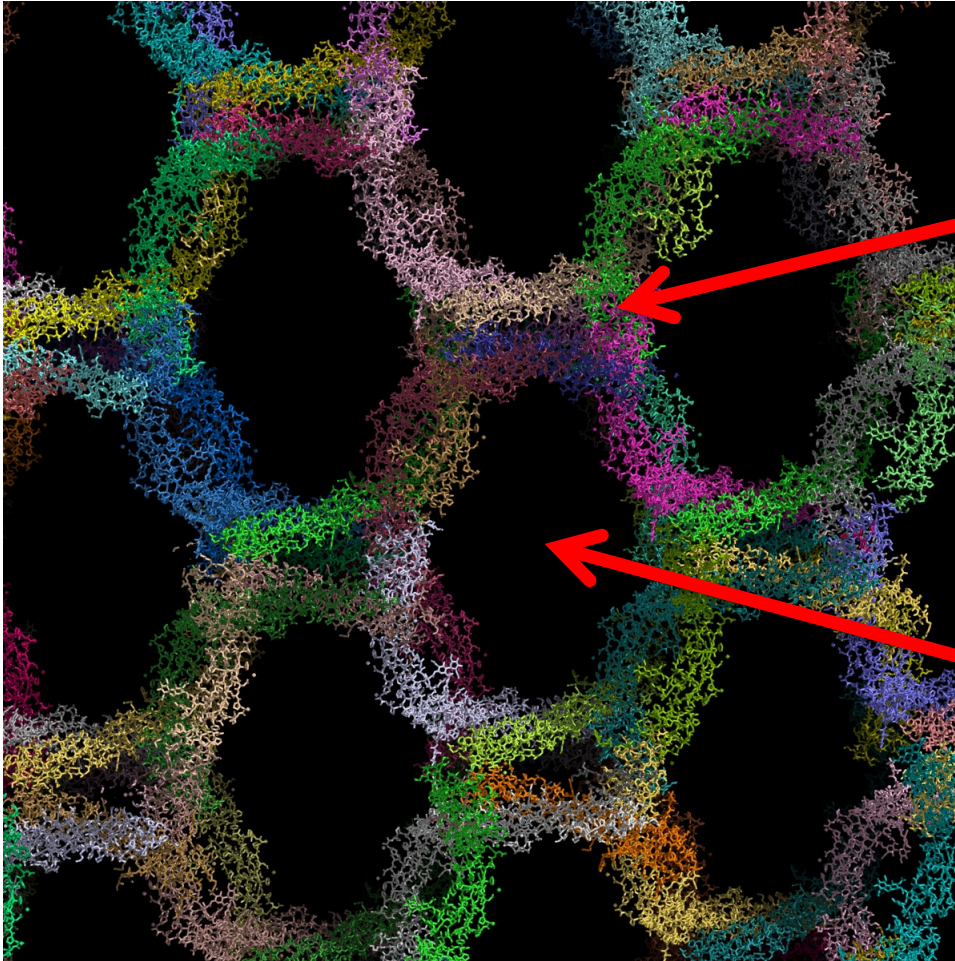


# Model refinement



# Crystal structure model

PDB code: 1QUB



**Macro-molecule**

**Bulk-solvent:  
~ 50% of unit cell  
volume**

**Crystal model:  $\rho_{\text{crystal}} = \rho_{\text{atoms}} + \rho_{\text{bulk solvent}}$**

# Atomic model

*Position*

*Larger-scale disorder*

ATOM	25	CA	PRO A	4	31.309	29.489	26.044	1.00	57.79	C	
ANISOU	25	CA	PRO A	4	8443	7405	6110	2093	-24	-80	C

*Local mobility (harmonic vibrations)*

$$\mathbf{F}_{\text{MODEL}} = k_{\text{OVERALL}} \left( \mathbf{F}_{\text{CALC (ATOMS)}} + \mathbf{F}_{\text{BULK}} \right)$$

Occupancy 1.00

57.79 ADP (B-factor)

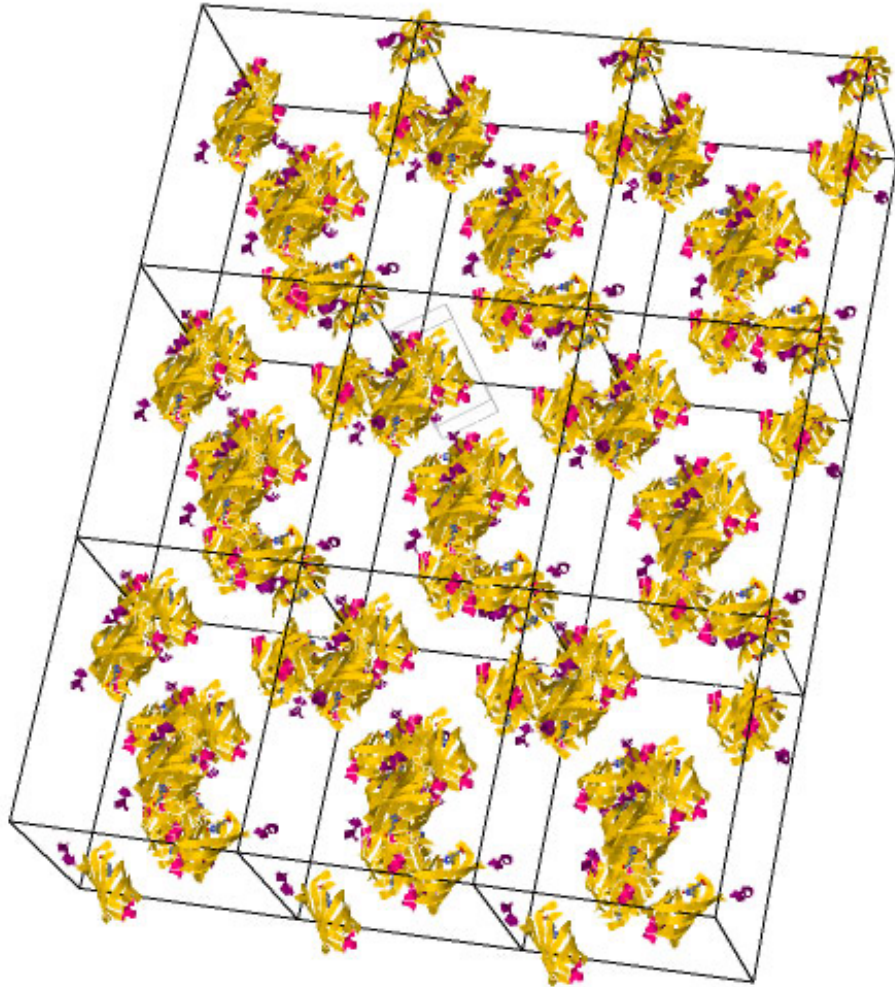
$$\mathbf{F}_{\text{CALC (ATOMS)}}(h, k, l) = \sum_{n=1}^{N_{\text{atoms}}} q_n f_n(s) \exp\left(-\frac{B_n s^2}{4}\right) \exp(2i\pi \mathbf{r}_n \cdot \mathbf{s})$$

Atom type C

31.309 29.489 26.044

Atomic coordinates

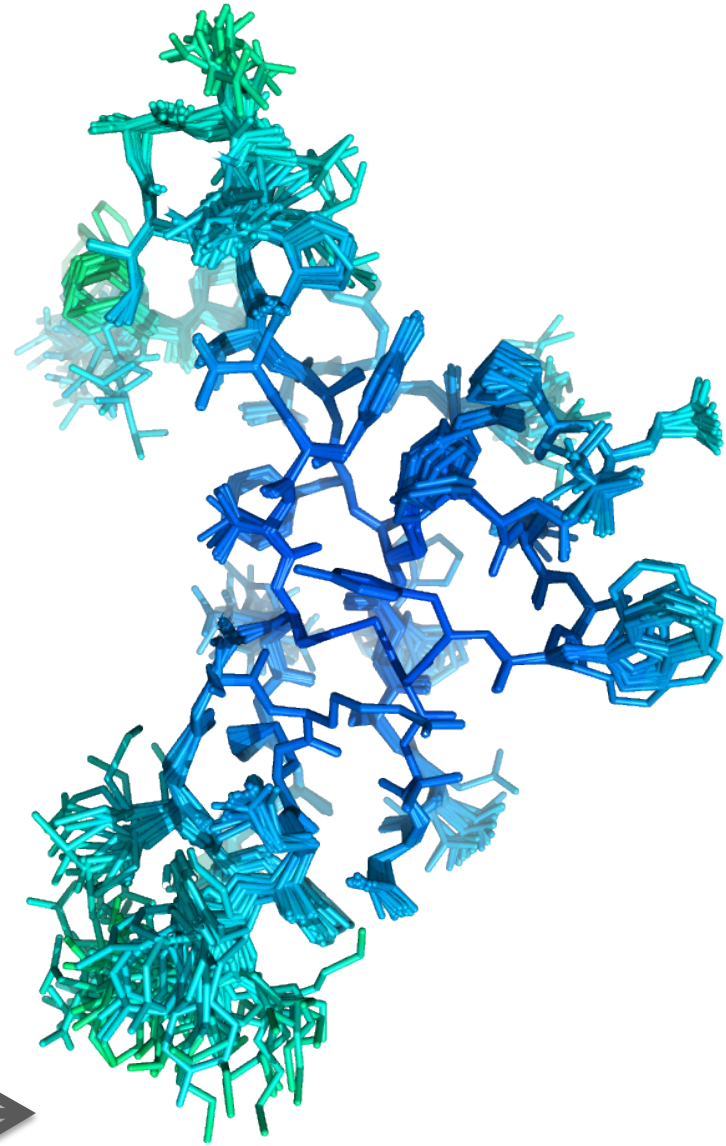
# Atomic model: disorder



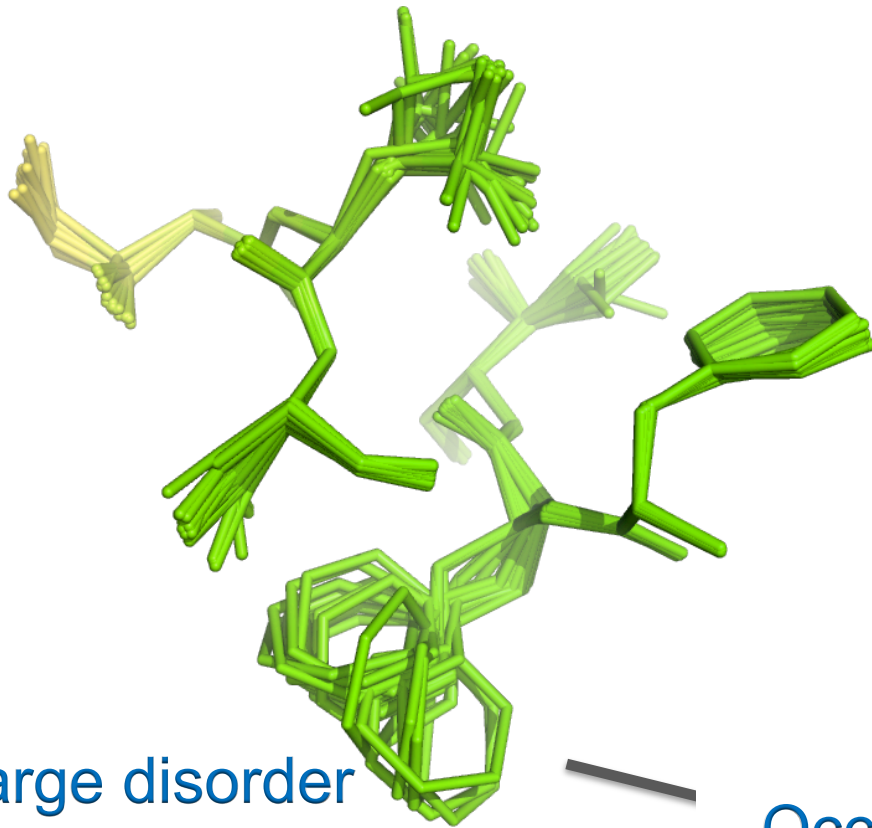
Crystal = many unit cells



Superpose all structures  
from each unit cell



# Atomic model: disorder



Small disorder

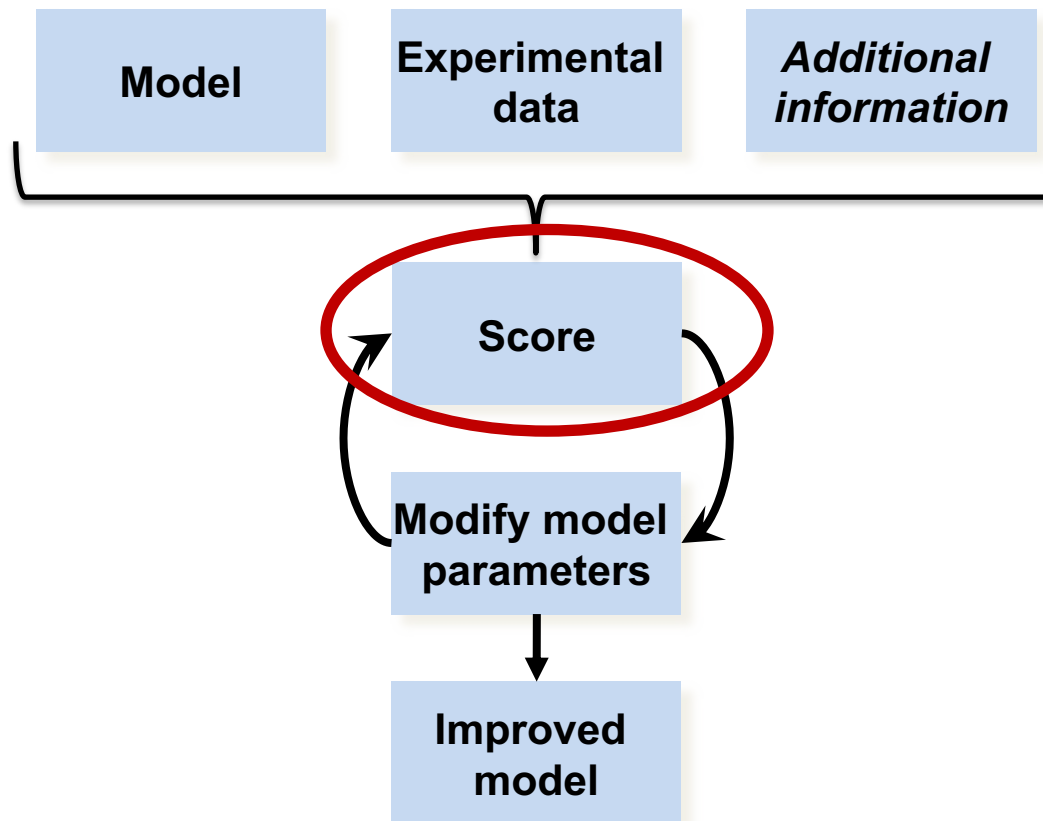
ADP (B-factor)

Large disorder

Occupancy

ATOM	25	CA	PRO	A	4	31.309	29.489	26.044	1.00	57.79	C	
ANISOU	25	CA	PRO	A	4	8443	7405	6110	2093	-24	-80	C

# Refinement target function (score)



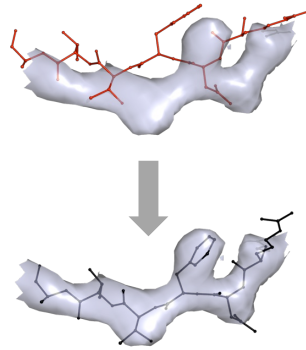
# Model refinement target (score)

**T**

Optimize  
consensus  
between model-  
to-data fit and...  
common sense

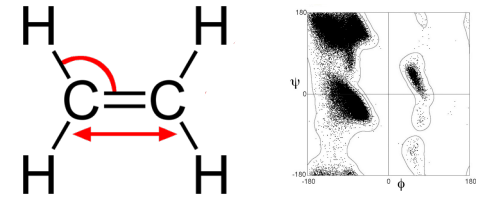
=

**T<sub>DATA</sub>**



+

**W \* T<sub>RESTRAINTS</sub>**



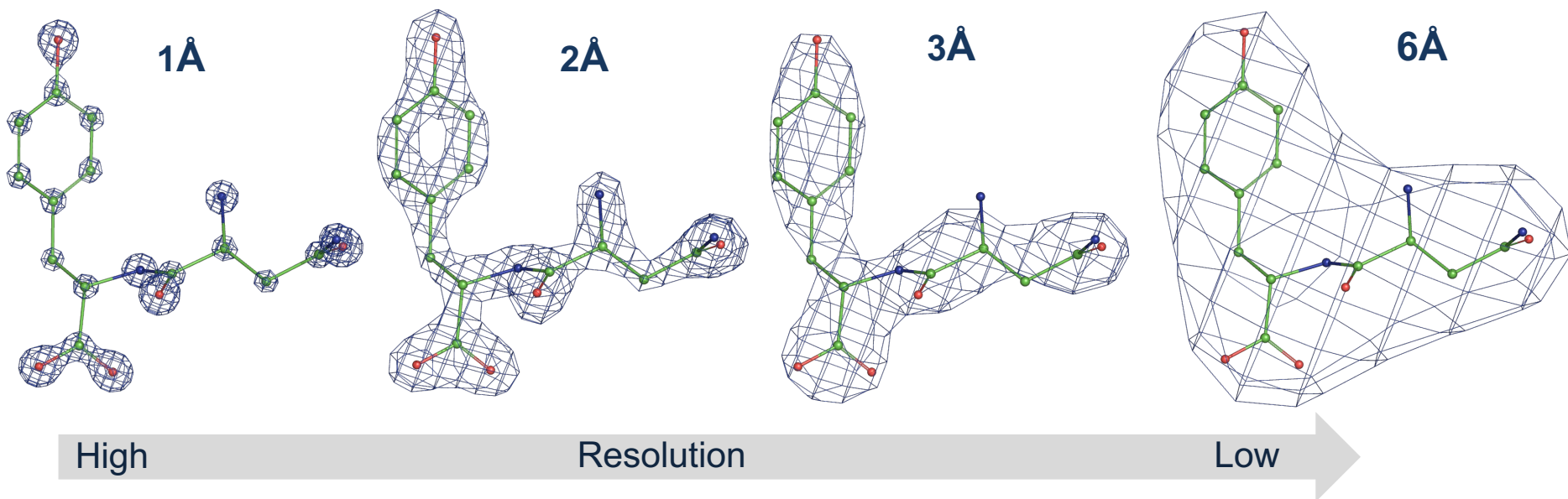
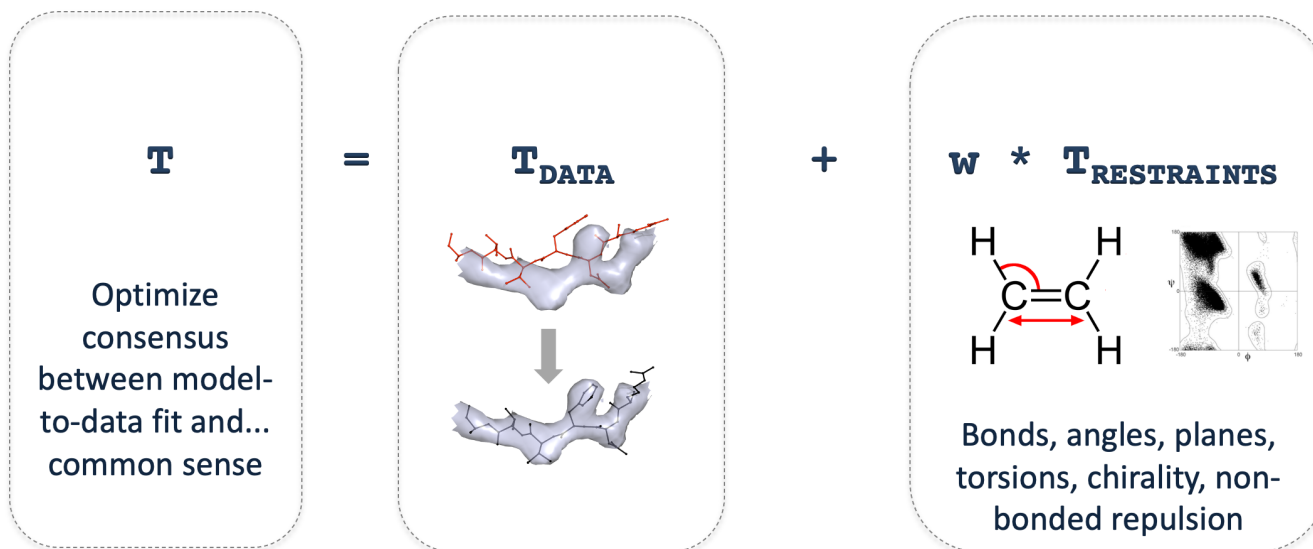
Bonds, angles, planes,  
torsions, chirality, non-  
bonded repulsion

$$\sum_{hkl} (F_{obs} - F_{model})^2$$

$$\sum_{hkl} \frac{||F_{obs}| - |F_{model}||}{|F_{obs}|}$$

Maximum-Likelihood

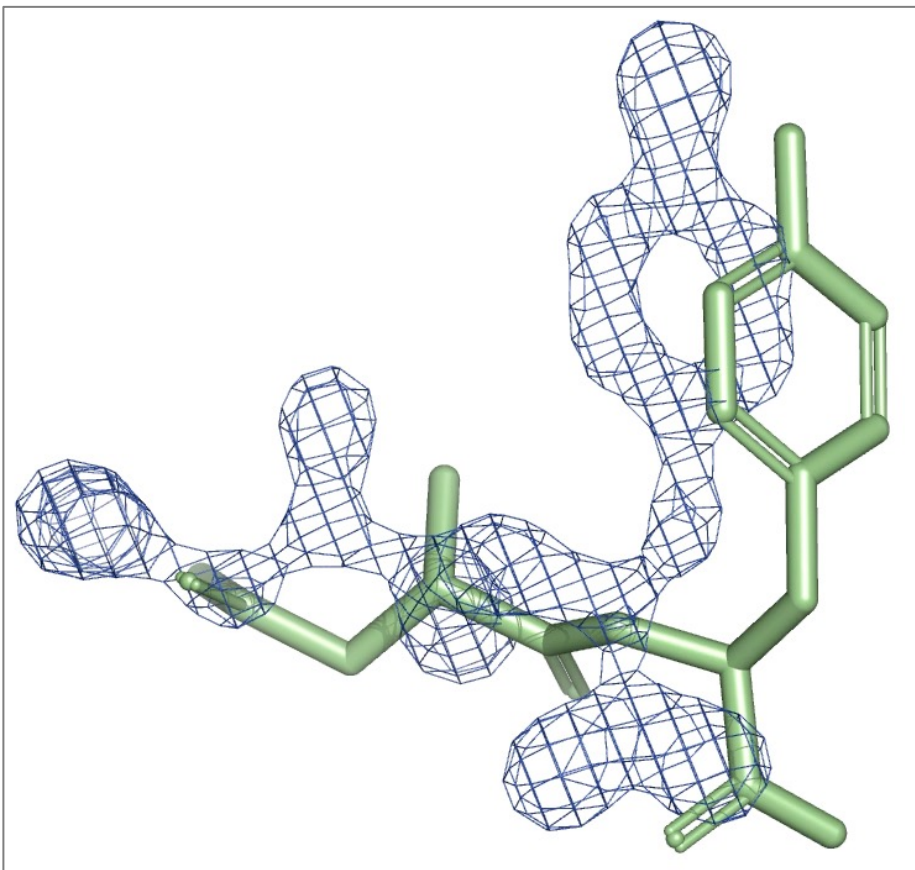
# Restraints and data resolution



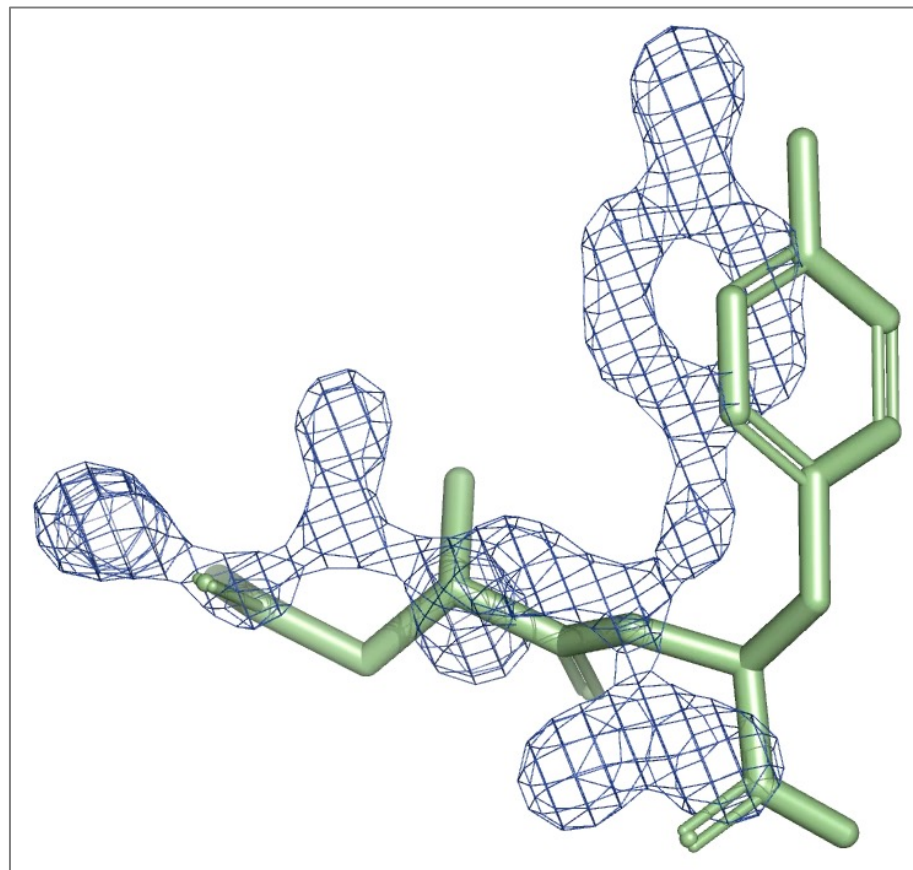


# Model refinement with vs no restraints: **high resolution**

$$\mathbf{T} = \mathbf{T}_{\text{DATA}} + \mathbf{W} * \mathbf{T}_{\text{RESTRAINTS}}$$



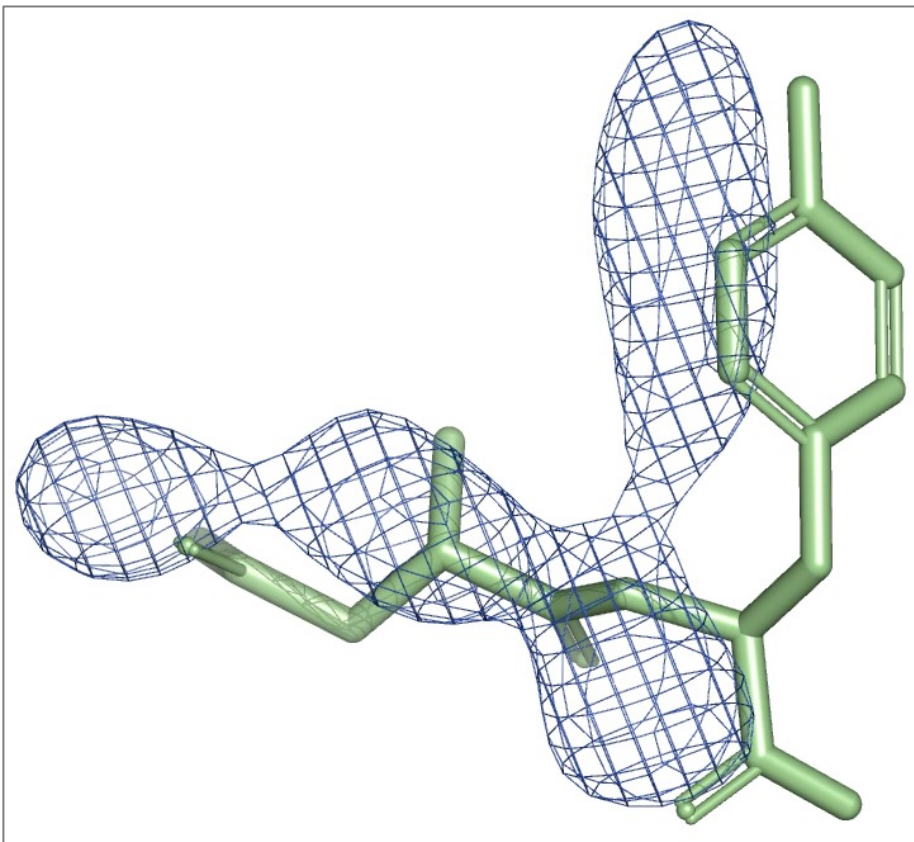
Using restraints



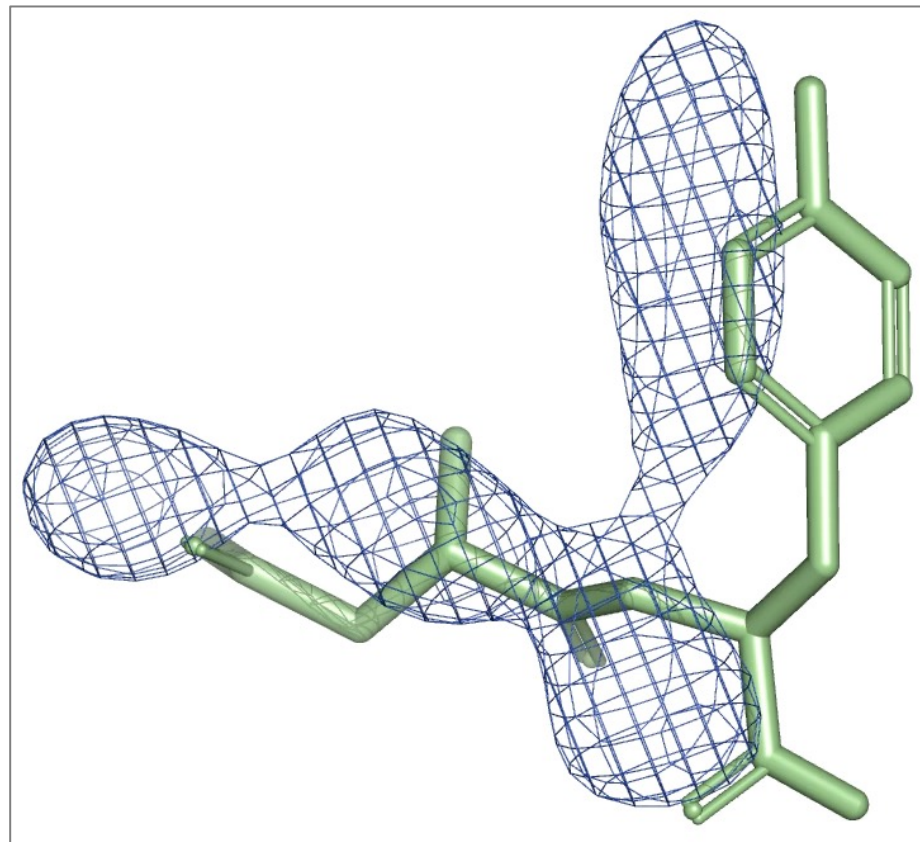
No restraints

# Model refinement with vs no restraints: **low resolution**

$$\mathbf{T} = \mathbf{T}_{\text{DATA}} + \mathbf{W} * \mathbf{T}_{\text{RESTRAINTS}}$$



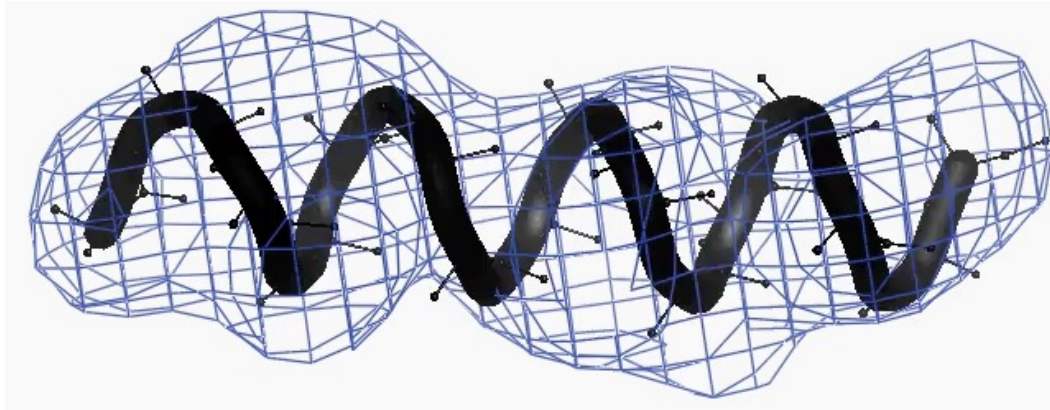
Using restraints



No restraints

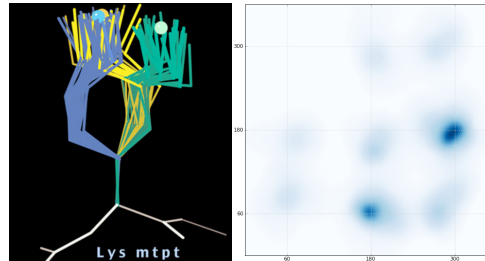
# Model refinement with insufficient restraints

- Refinement of a perfect  $\alpha$ -helix into low-res map
  - Using simplistic (standard) restraints on covalent geometry
    - Model geometry deteriorates as result of refinement

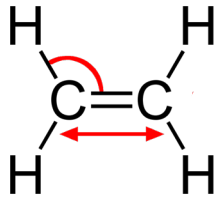


# More restraints for low resolution

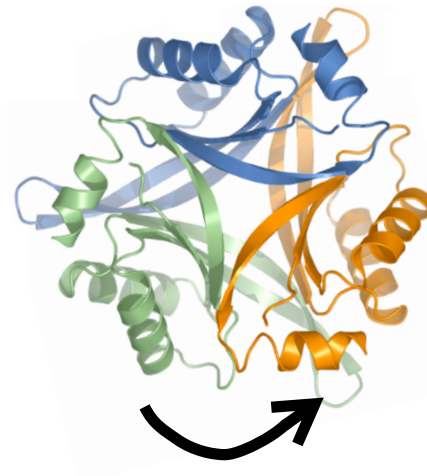
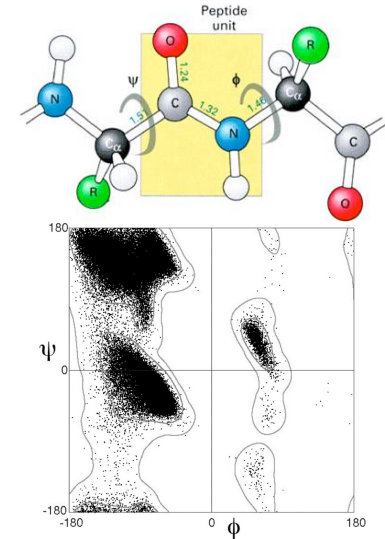
## Side chain distributions



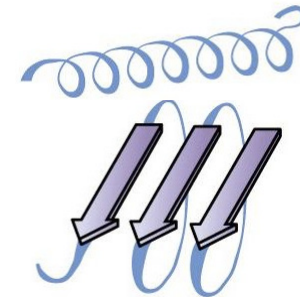
## Covalent geometry



## Main chain distributions

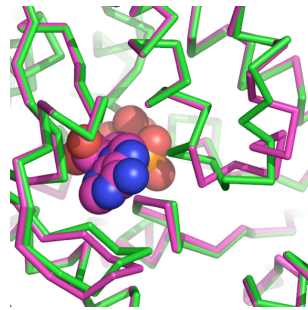


Internal symmetry (NCS)

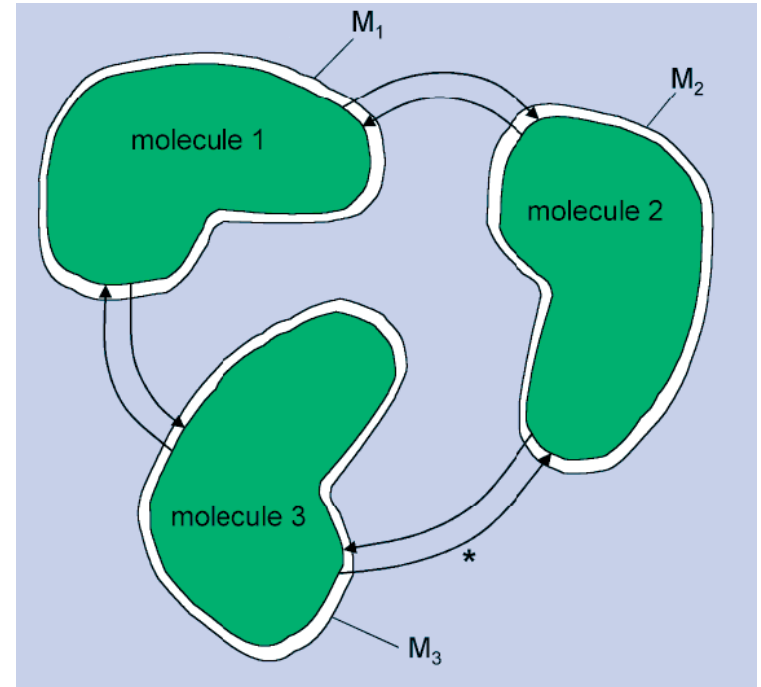
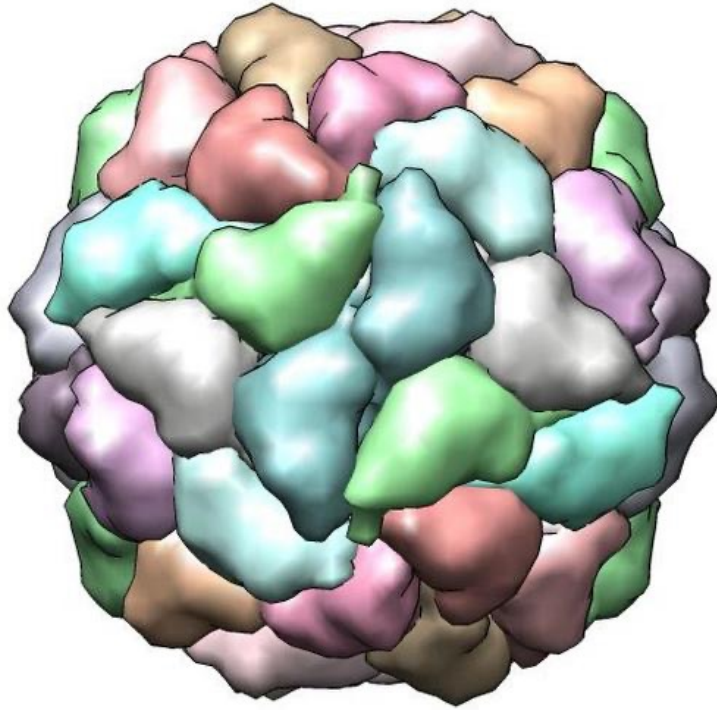


Secondary structure

Similar (homologous) structures (reference model restraints)



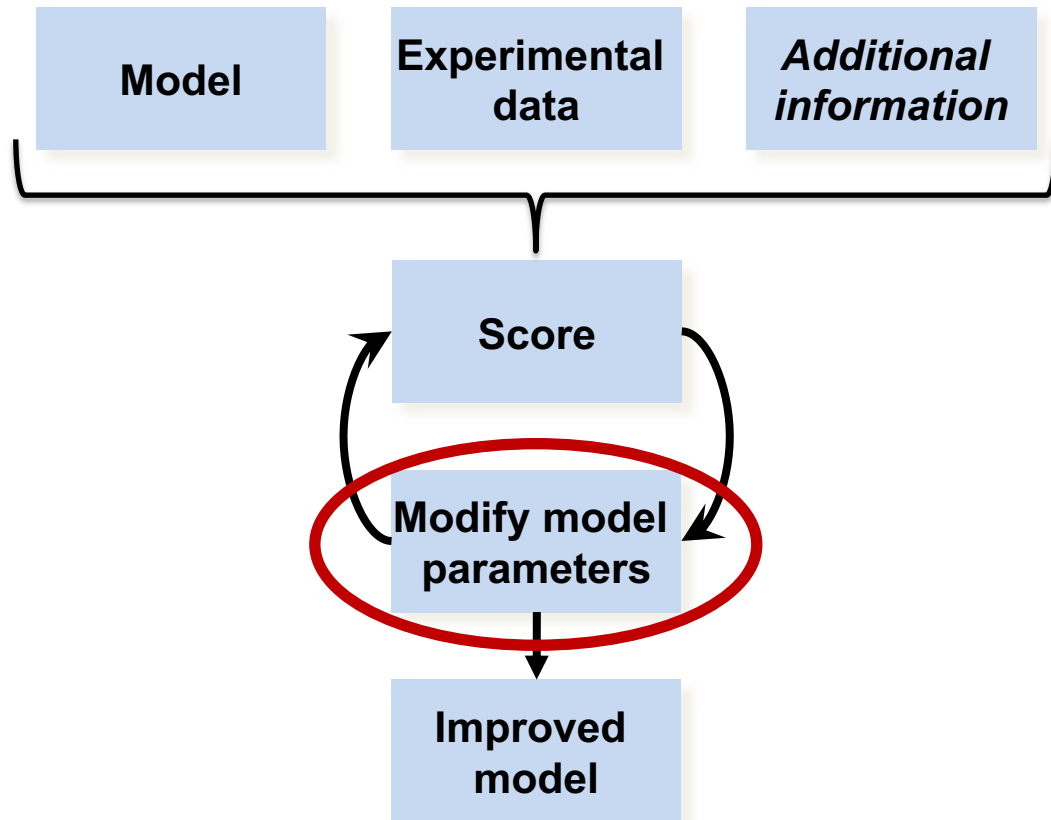
# NCS (internal symmetry): constraints vs restraints



Source: Internet

- **Constraints:** molecules 1, 2 and 3 are required to be **identical**
- **Restraints:** molecules 1, 2 and 3 are required to be **similar** but not necessarily identical

# Refinement

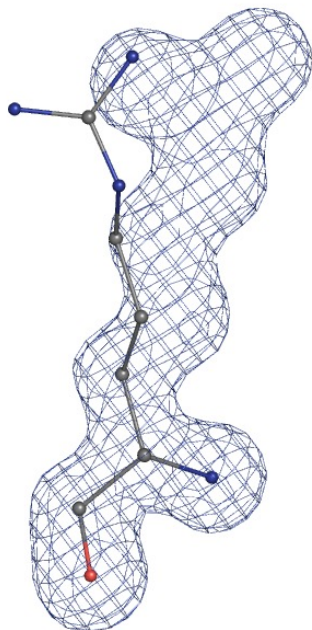


# Choices of optimization method

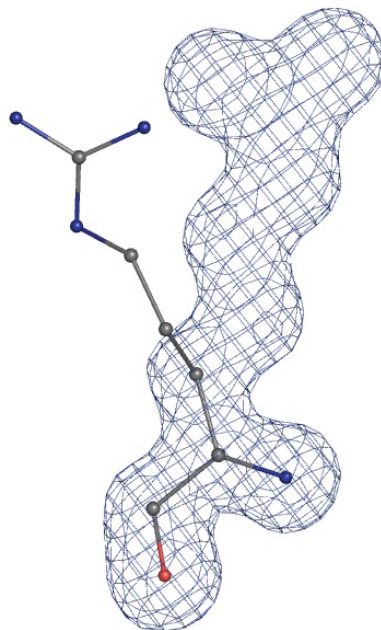
- Gradient-based minimization
- Simulated annealing
- Grid (systematic) searches
- Manual using molecular graphics programs (Coot, Chimera,...

# Choice of refinement method and refinement convergence

**Minimization**

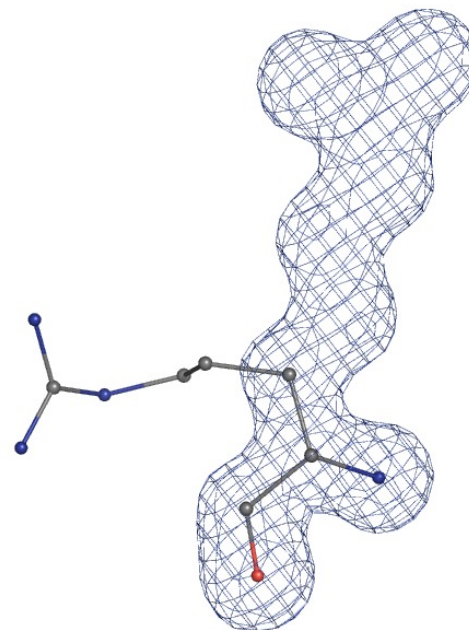


**Simulated Annealing**



**Beyond  
convergence radius  
of minimization**

**Real-space grid search**



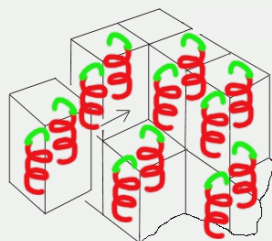
**Beyond convergence  
radius of  
minimization and SA**



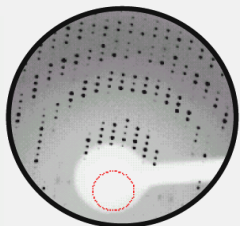
# **Phenix tools for model refinement**

# Refinement

## Crystallography



Initial model



Experimental data

*A priori* knowledge

Score

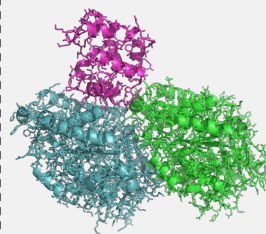
Modify model parameters

Improved model

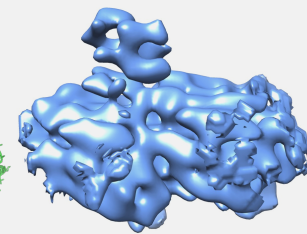
phenix.refine

Available since 2005

## Cryo-EM



Initial model



Experimental data

*A priori* knowledge

Score

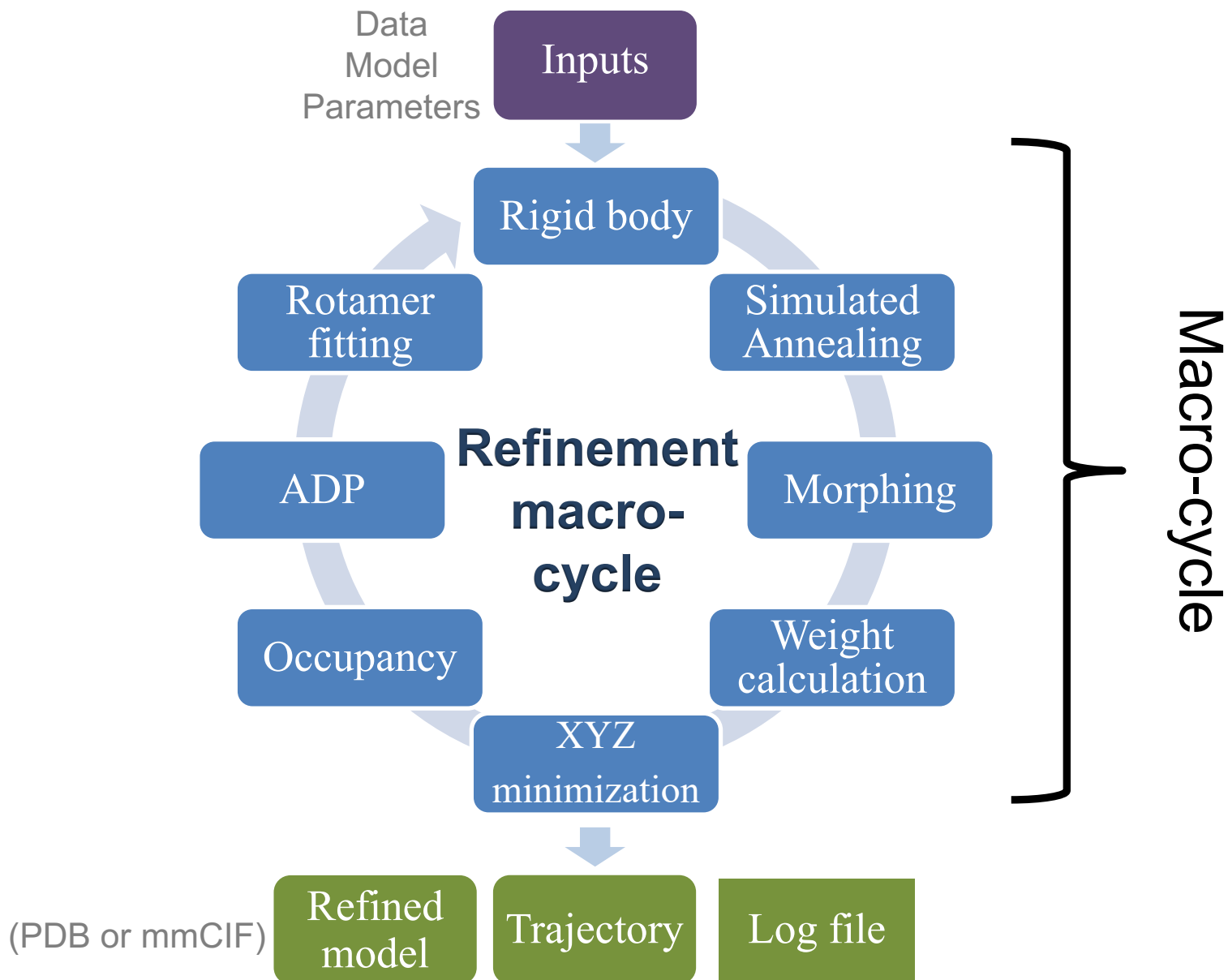
Modify model parameters

Improved model

phenix.real\_space\_refine

Available since 2013

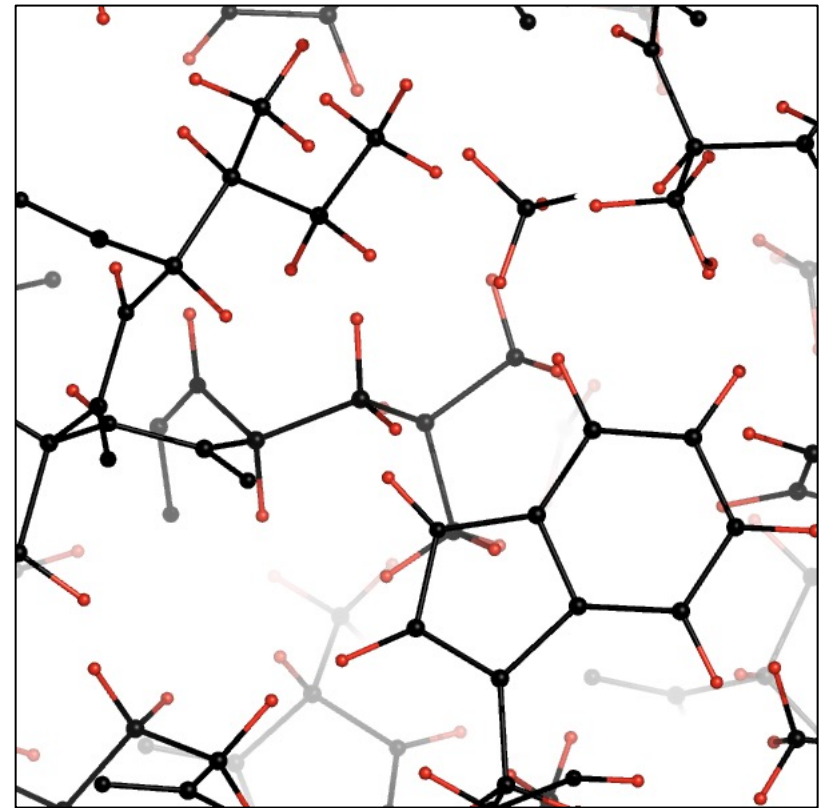
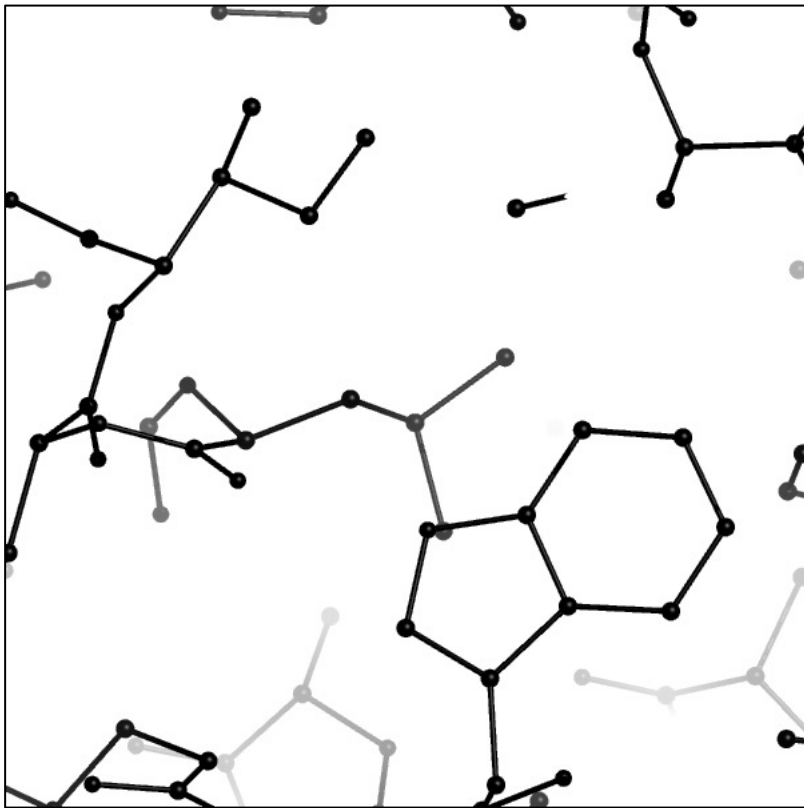
# Refinement protocol



# **Refinement: practical considerations**

# Use Hydrogen atoms

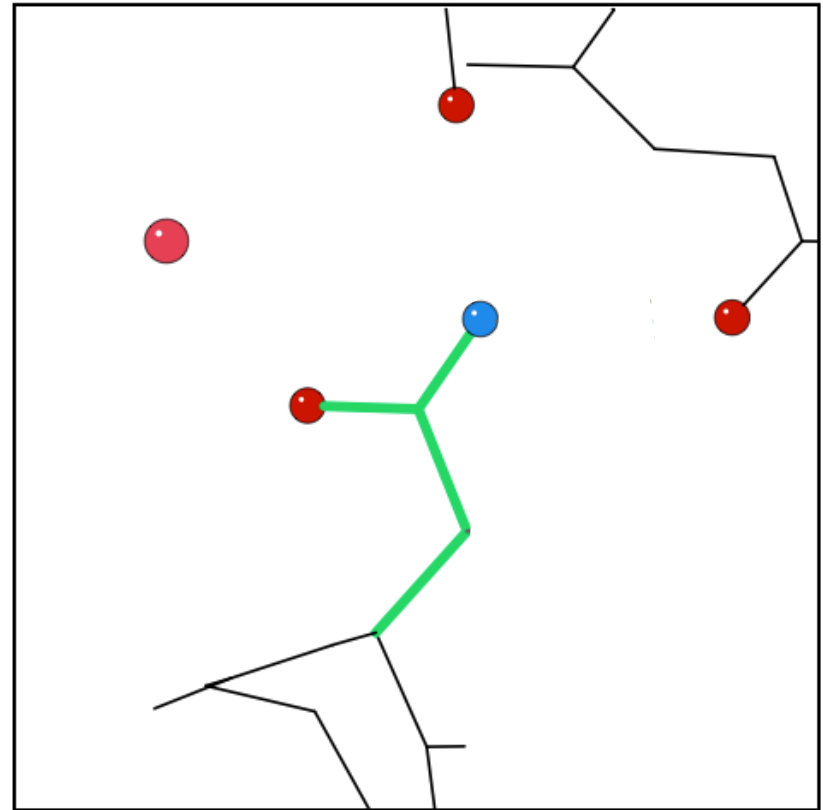
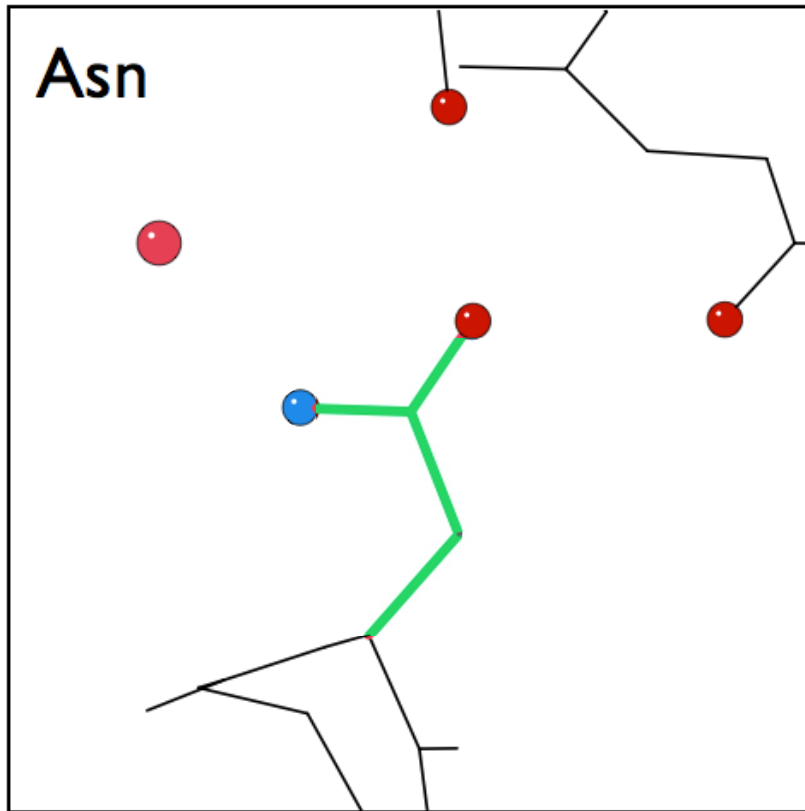
- Half of the atoms in a protein molecule
- Make most interatomic contacts
- Add to model towards the end, data resolution does not matter
- Once added, do not remove before the PDB deposition
- H do contribute to R-factors (expect 0.1-2% drop in R)



**A structure without (left) and with (right) hydrogen atoms**

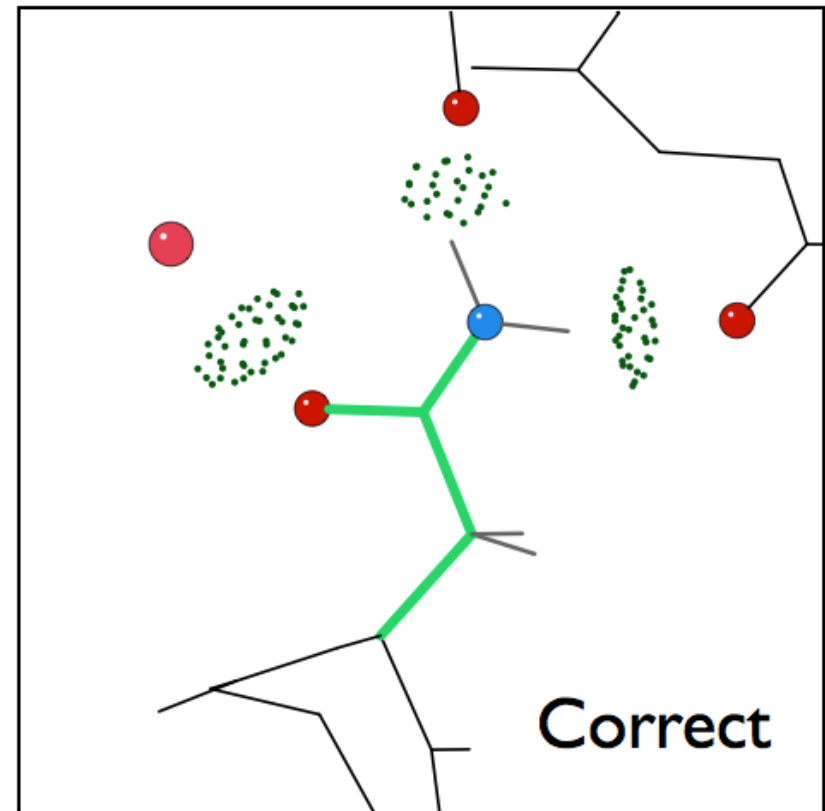
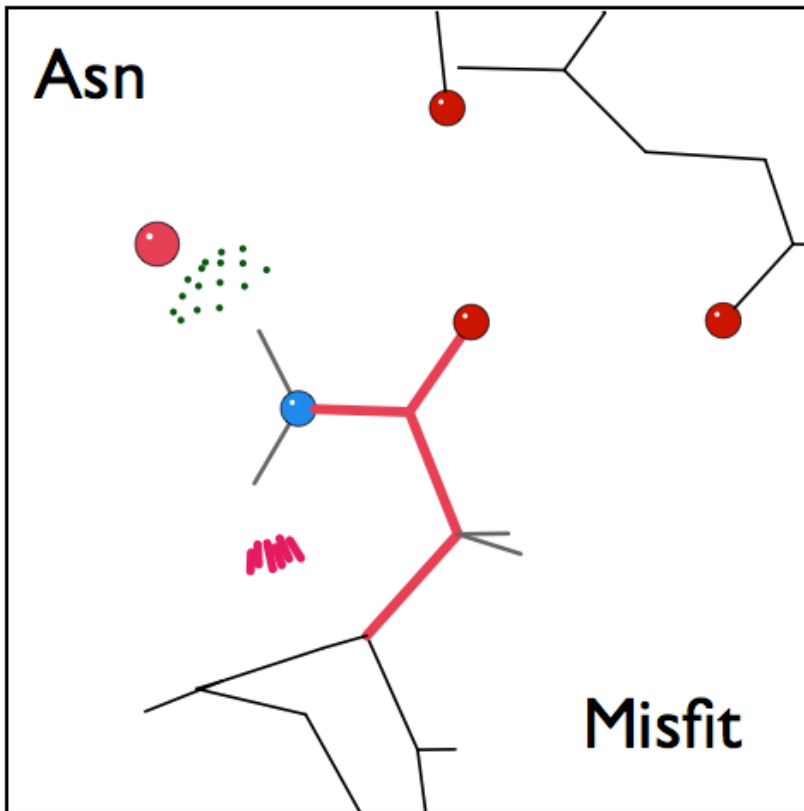
# Use Hydrogen atoms

- N/Q/H flips (asparagine/glutamine/histidine)
  - Based on clash analysis
  - Requires H present

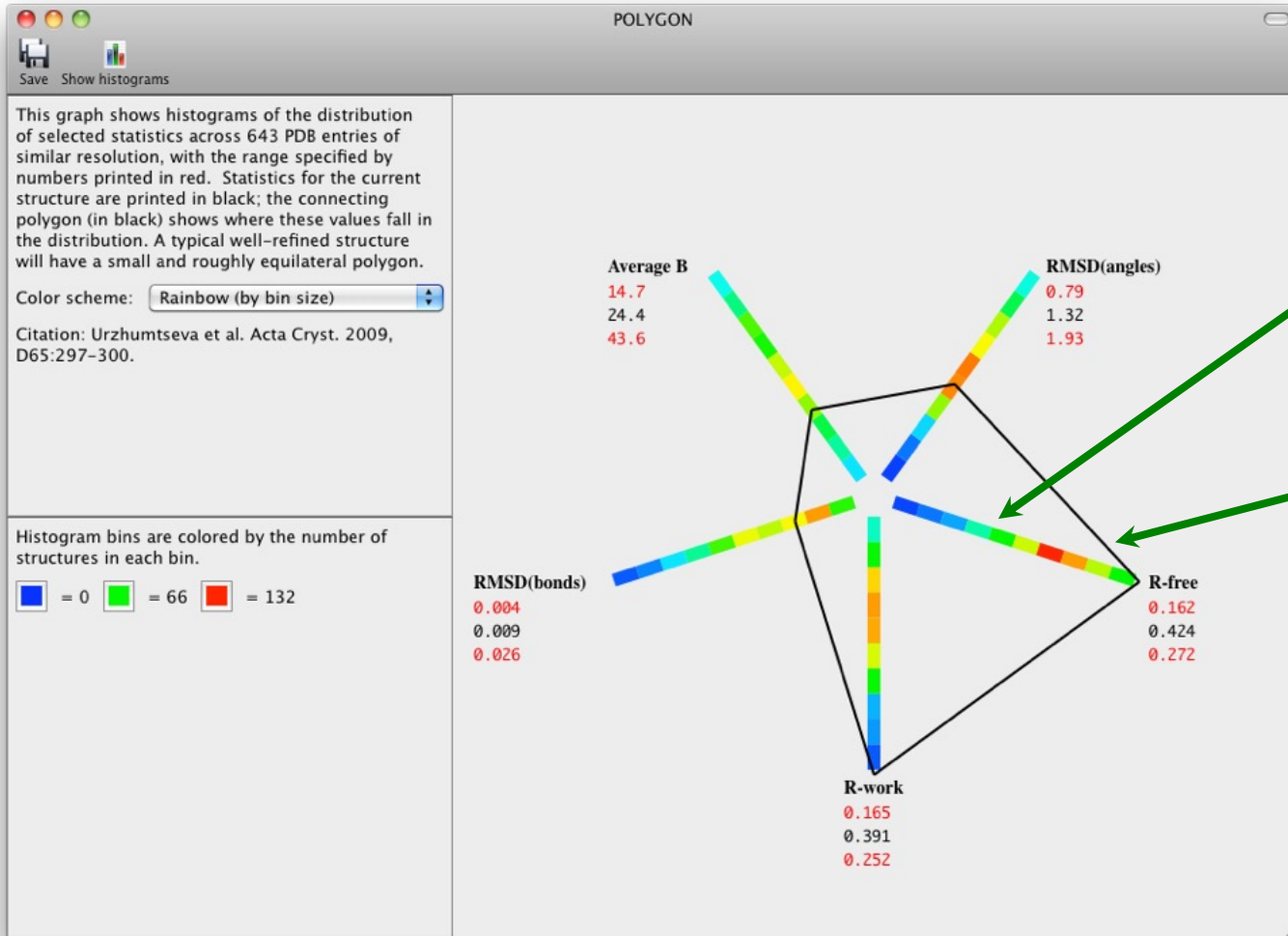


# Use Hydrogen atoms

- N/Q/H flips
  - Based on clash analysis
  - Requires H present



# Know when to stop



Colored bars are histograms showing distribution of values for structures at similar resolution

The black polygon shows where the statistics for the user's structure fall in each histogram

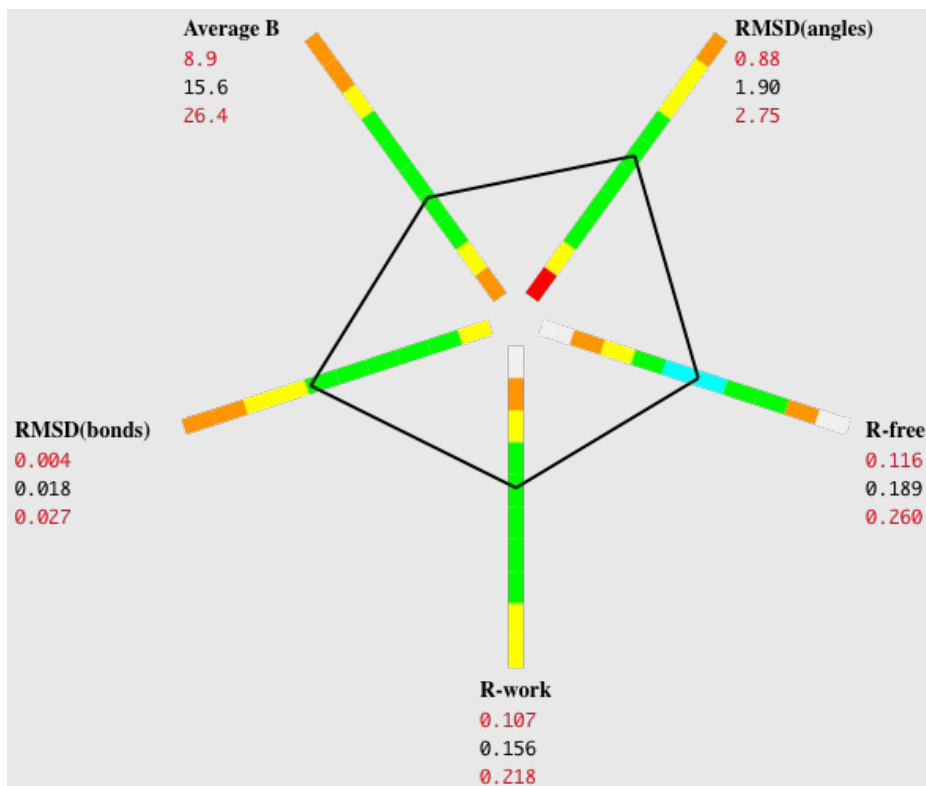
## Crystallographic model quality at a glance.

L.Urzhumtseva, P.V.Afonine, P.D.Adams & A.Urzhumtsev. *Acta Cryst.* D65, 297-300 (2009)

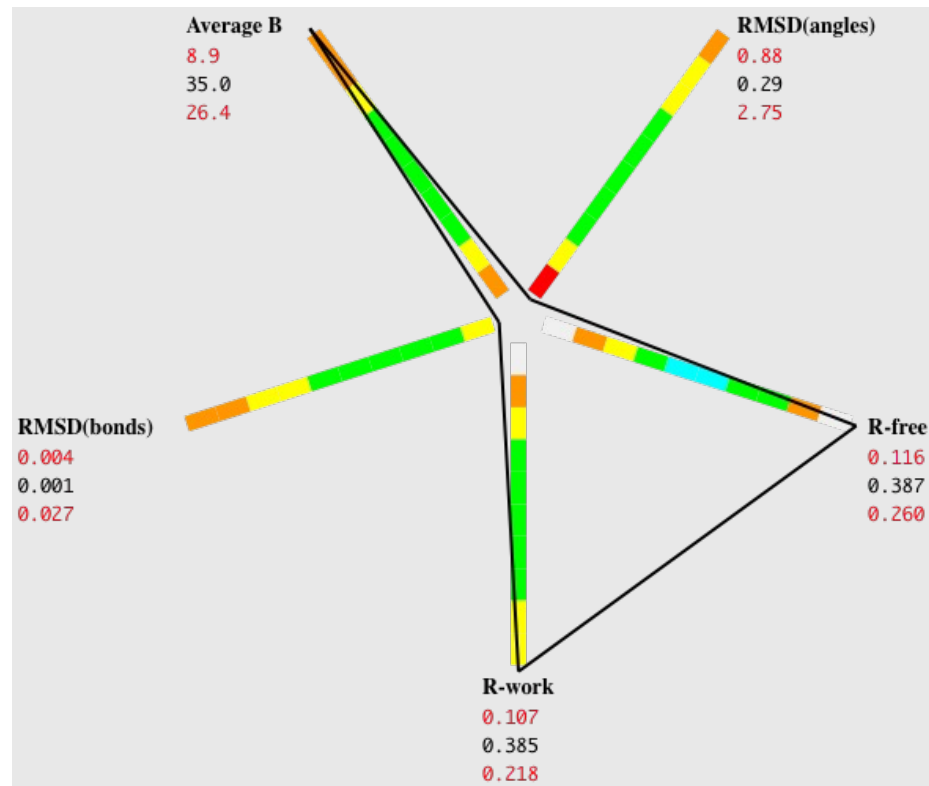


# Know when to stop

## Likely overall good model

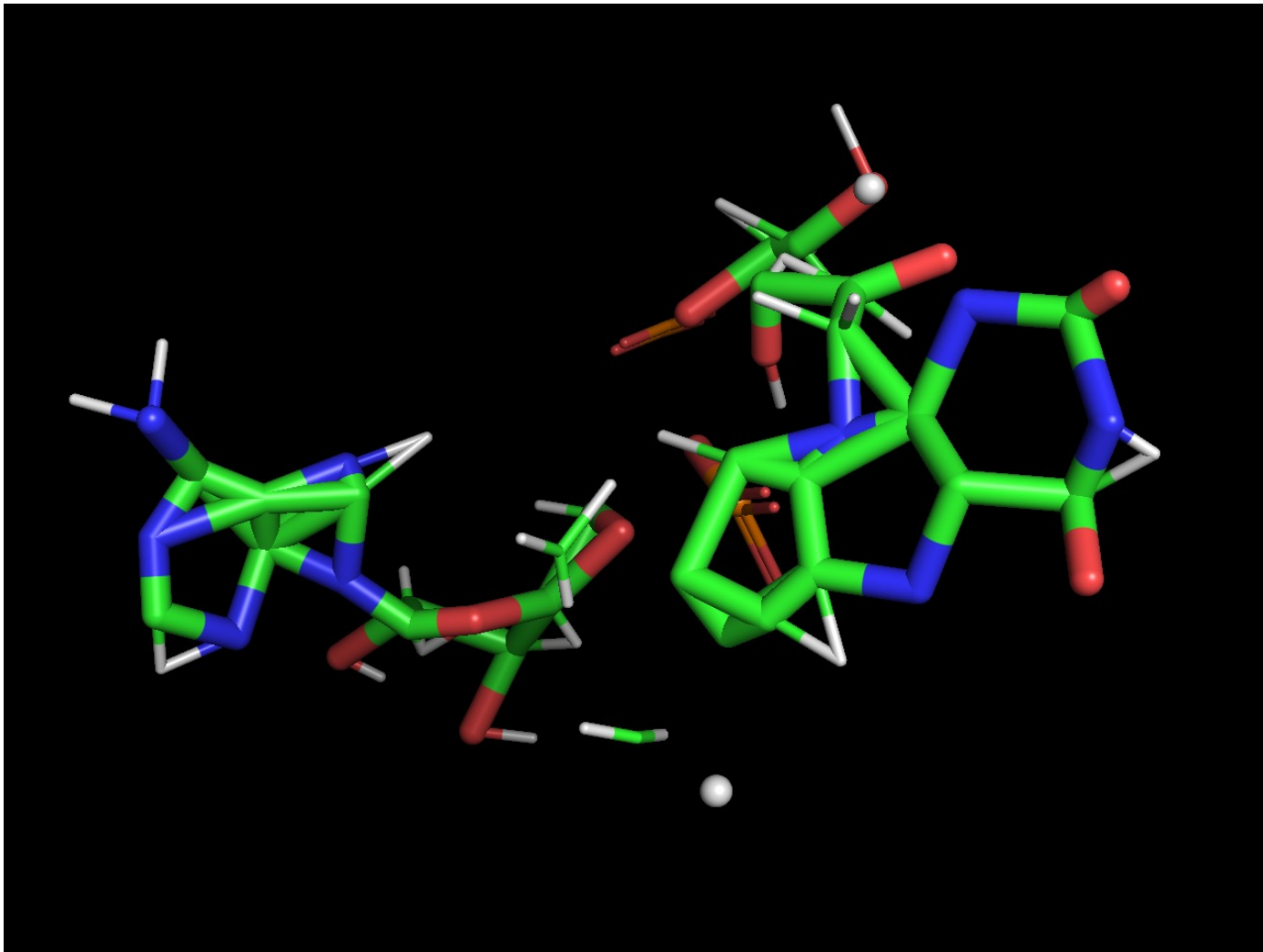


## Clearly there are problems

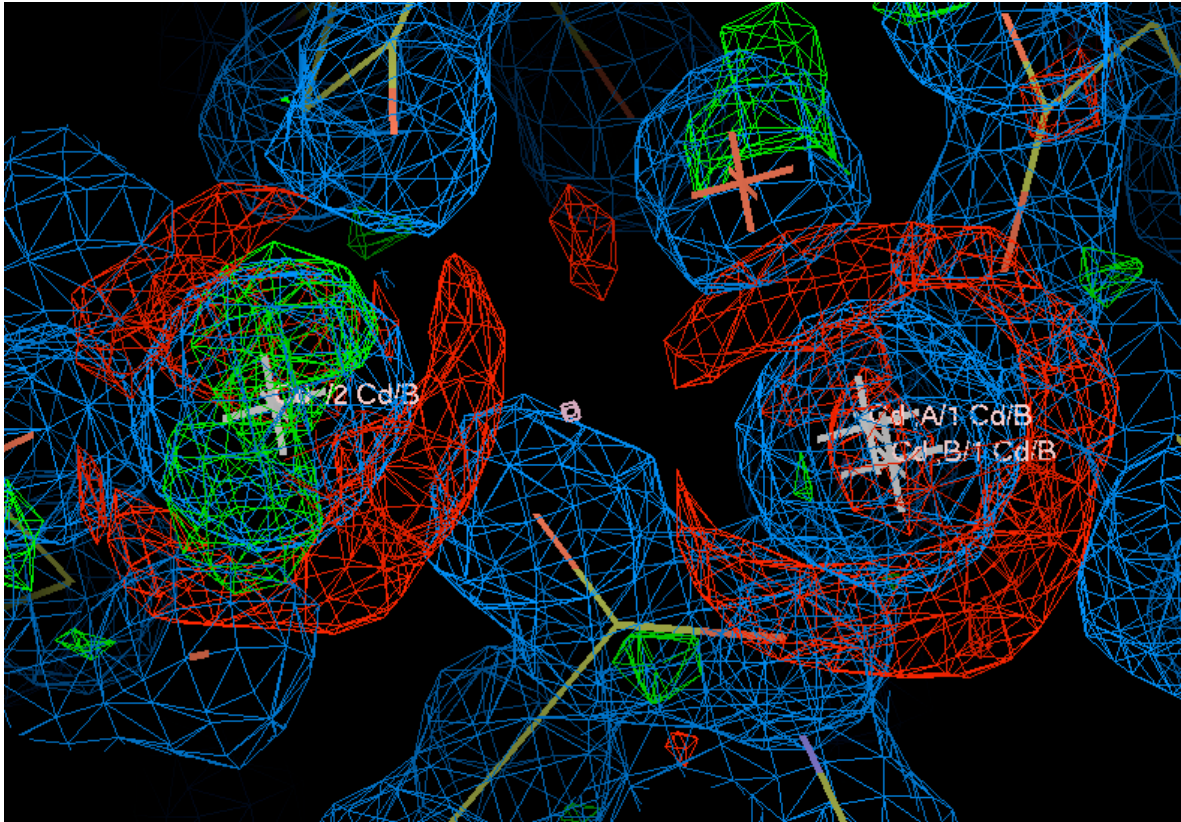


# Local vs Global

- $R_{\text{WORK}}/R_{\text{FREE}}$  , bond/angle RMSDs etc do not report on local errors



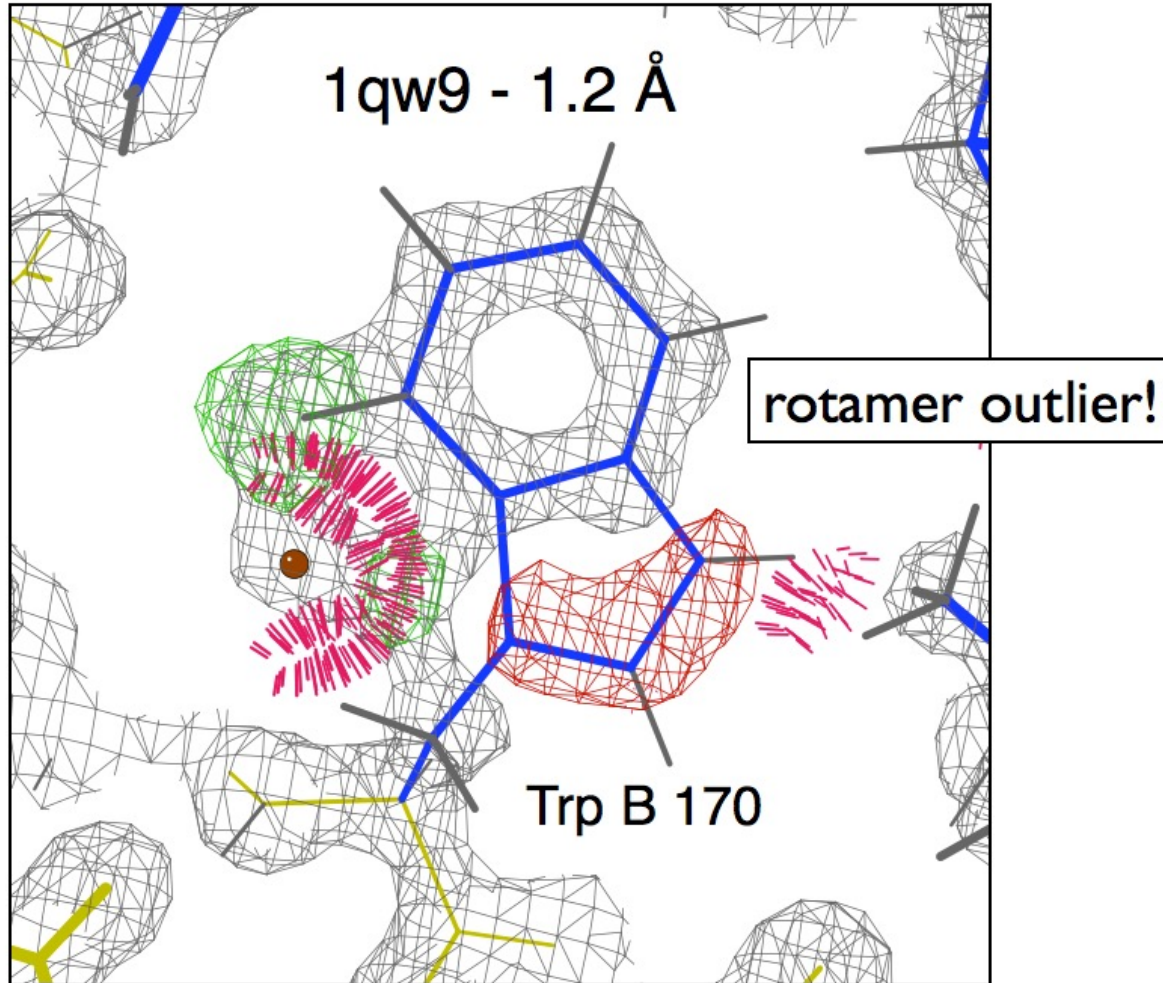
# Map and model errors



Reasons for +ve/-ve density:

- Suboptimal xyz, occupancy, ADP, anomalous  $f'$  &  $f''$ , charge
- Refinement has not reached convergence
- Wrong atom (ion)
- Suboptimal ADP (B-factor) type: isotropic vs anisotropic
- **NEW** phenix.oat is the new tool to help with this

# Not all modeling errors can be fixed by refinement



# Low resolution (3Å or worse)

- Use:
  - Ramachandran plot restraints
  - Secondary structure restraints
  - Reference model restraints (if quality homology model is available)
  - NCS (restraints or constraints)

# Aggressive optimization methods

- Simulated annealing (SA)
- Model morphing
  - Only use if model has gross errors (correction requires large movements)
  - Do not use if model is relatively good and only needs small corrections

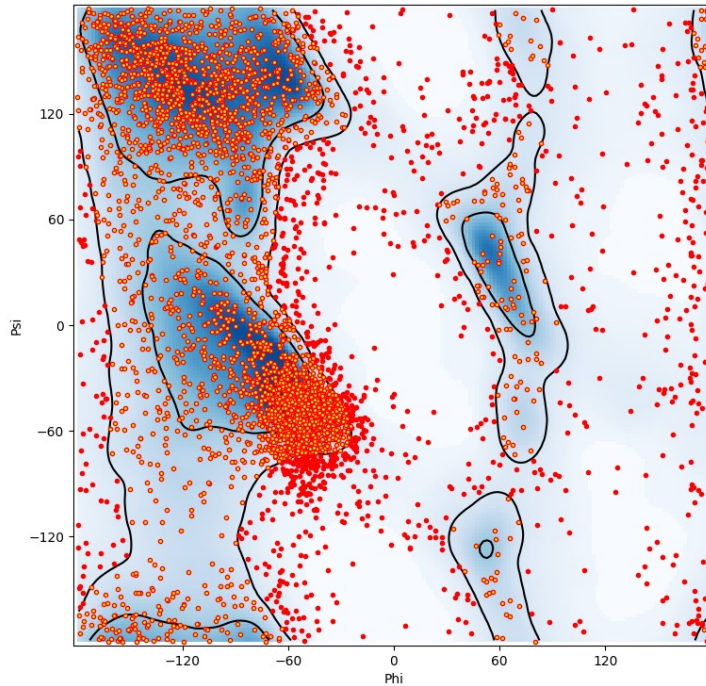
# Ramachandran plot restraints

- Likely need at about 3Å and worse
- Better than 3Å: use if needed (preserve good initial model from deterioration)
- Check Ramachandran plot regularly
- Don't use to fix outliers. Fix outliers first (manually), then use Ramachandran plot restraints to stop re-occurring outliers

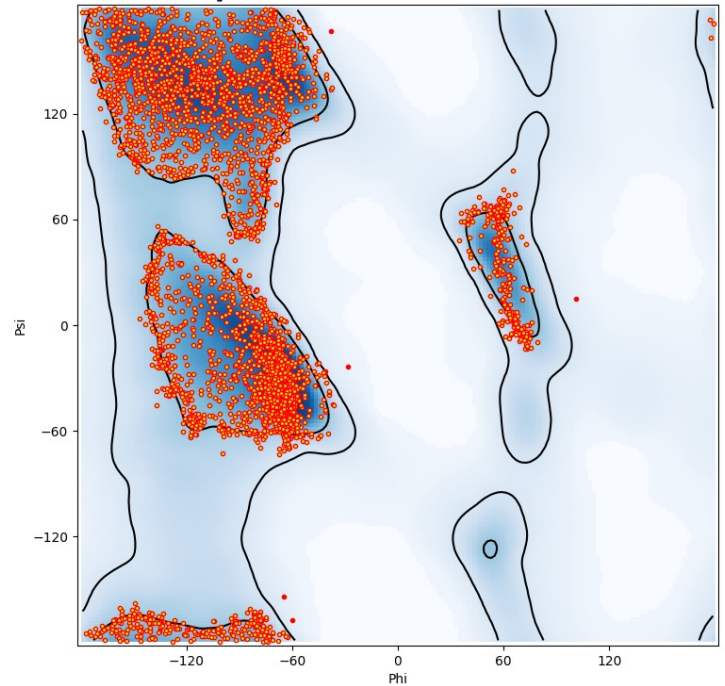
# Ramachandran plot restraints

PDB code: 5a9z

Original



Refined with Ramachandran plot restraints



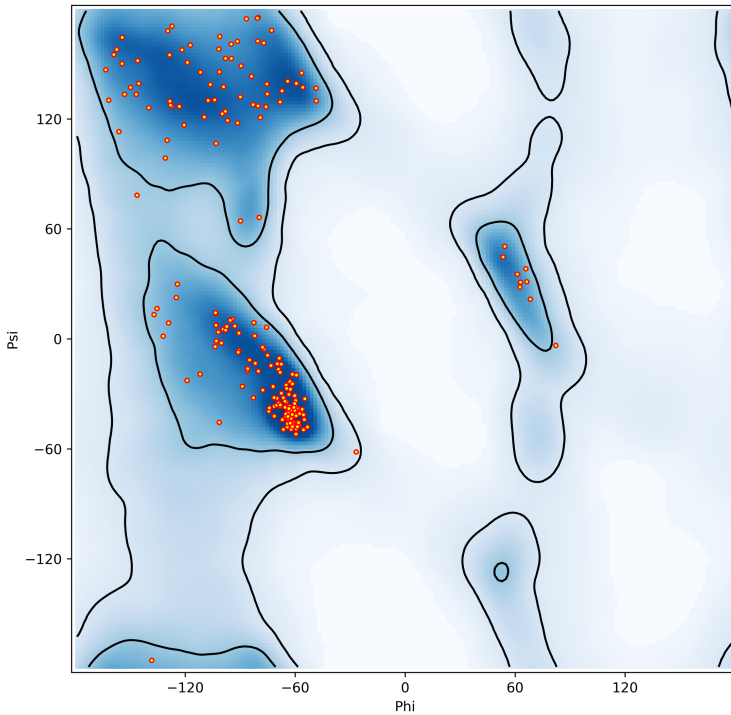
Bad idea to use Ramachandran plot restraints in this case. Fix outliers first!



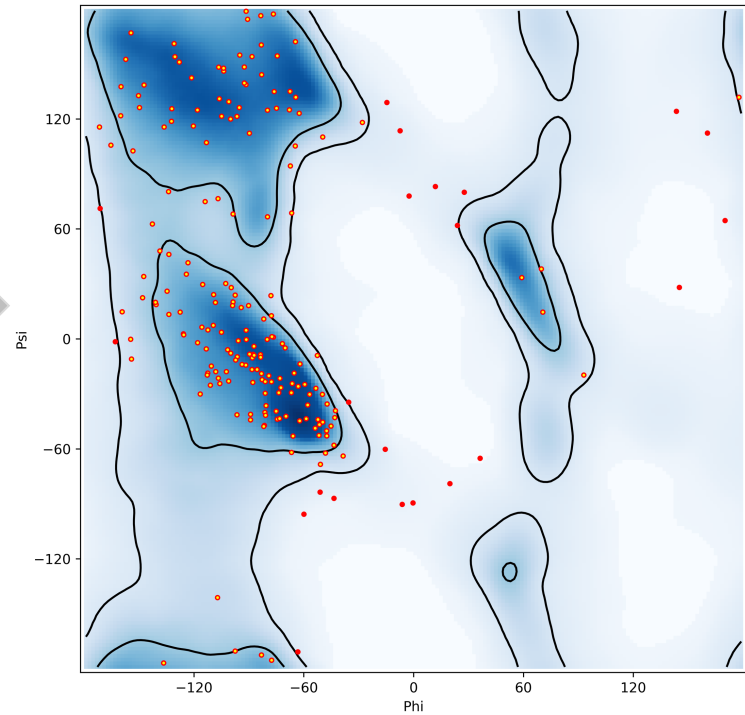
# Ramachandran plot restraints

- Ramachandran plot restraints
  - Use to stop outliers from occurring

Before refinement



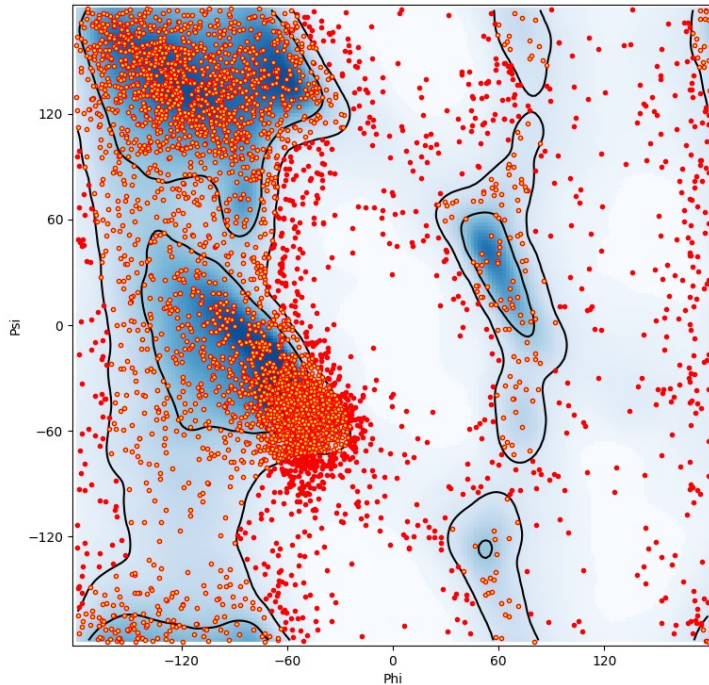
After refinement (No Ramachandran plot restraints)



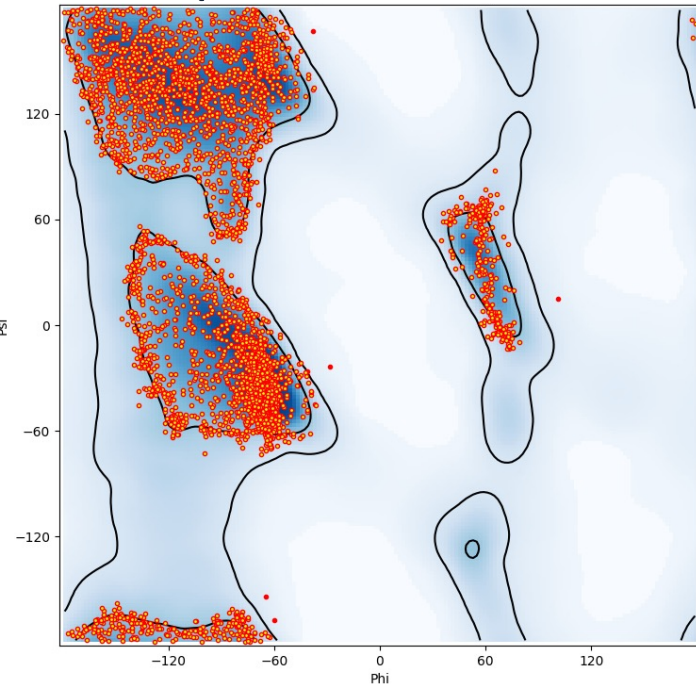
# Ramachandran plot restraints

- What is wrong with this plot?

Original

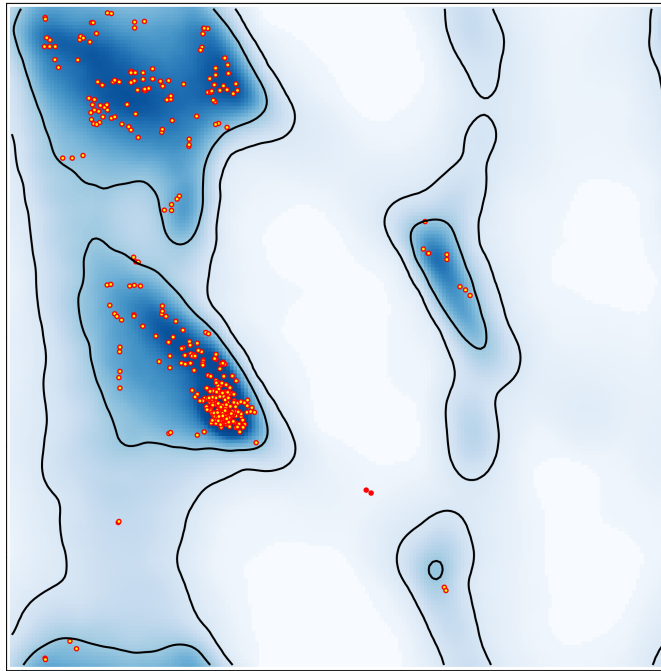


Refined with Ramachandran plot restraints



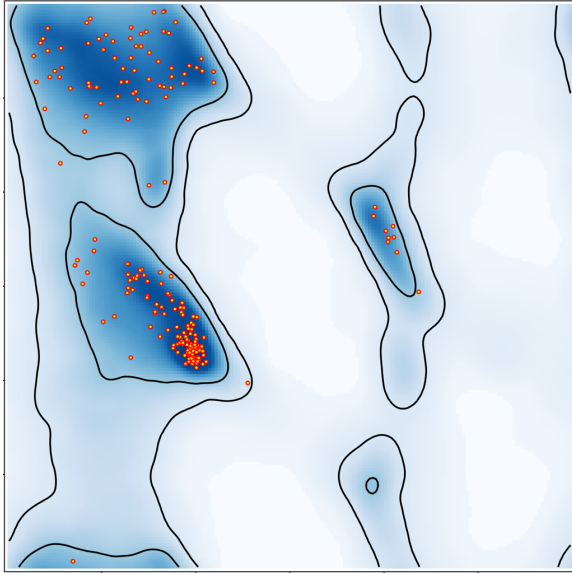
# Ramachandran plot restraints

- It is very different from what we expect!

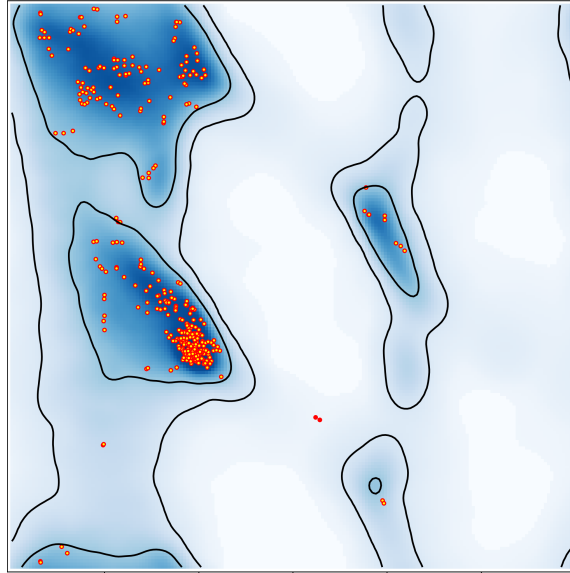


# How you can tell good vs bad plot?

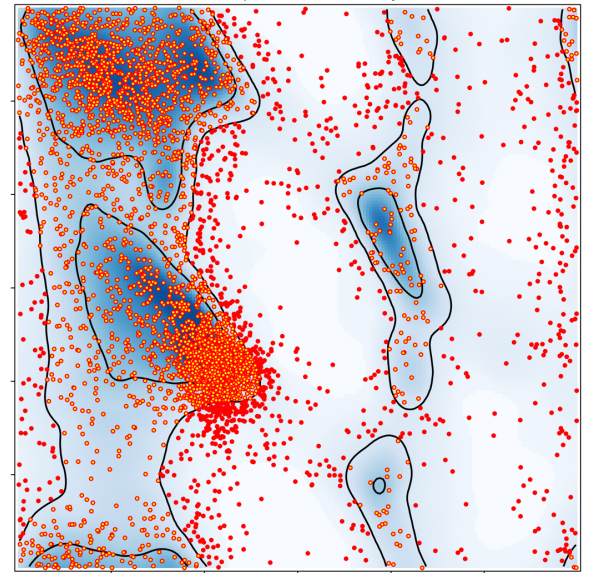
**Good**



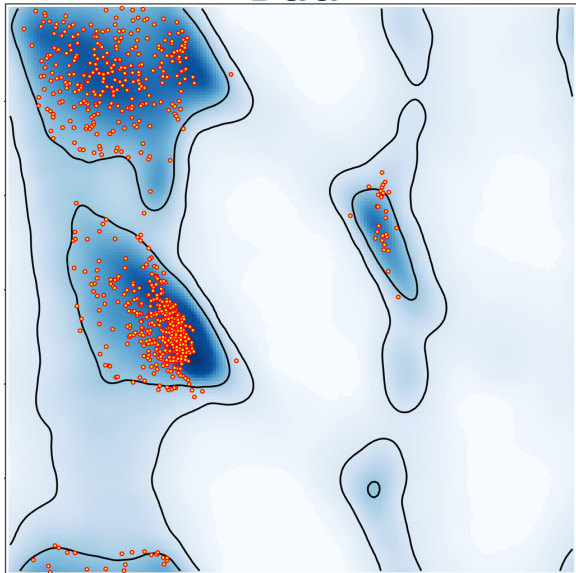
**Good**



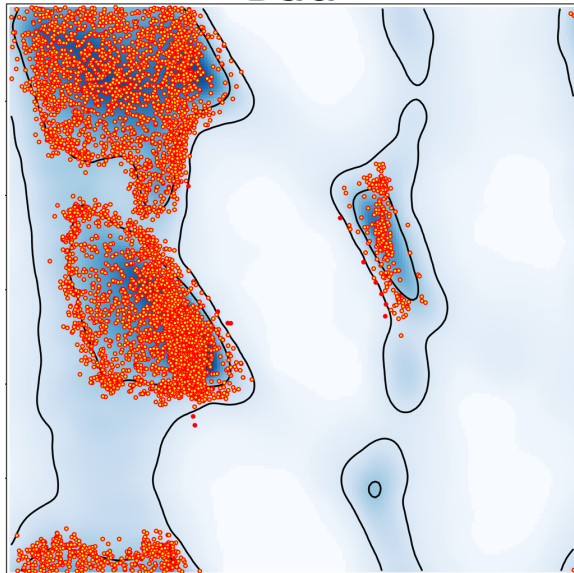
**Bad**



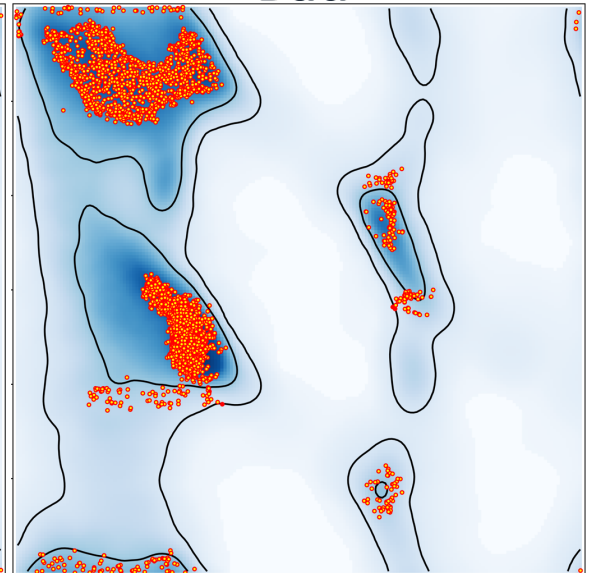
**Bad**



**Bad**



**Bad**



# Ramachandran plot Z-score

**CABIOS**

Vol. 13 no. 4 1997  
Pages 425–430

## ***Objectively judging the quality of a protein structure from a Ramachandran plot***

*Rob W.W.Hooft, Chris Sander and Gerrit Vriend*

- Good at spotting odd plots
- One number, simple criteria:
  - Poor:  $|Z| > 3$  Suspicious:  $2 < |Z| < 3$  Good:  $|Z| < 2$

**Structure**

 CellPress

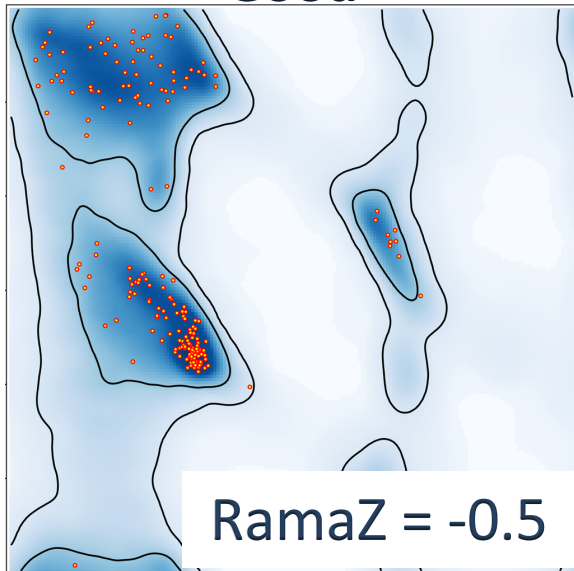
Resource

## **A Global Ramachandran Score Identifies Protein Structures with Unlikely Stereochemistry**

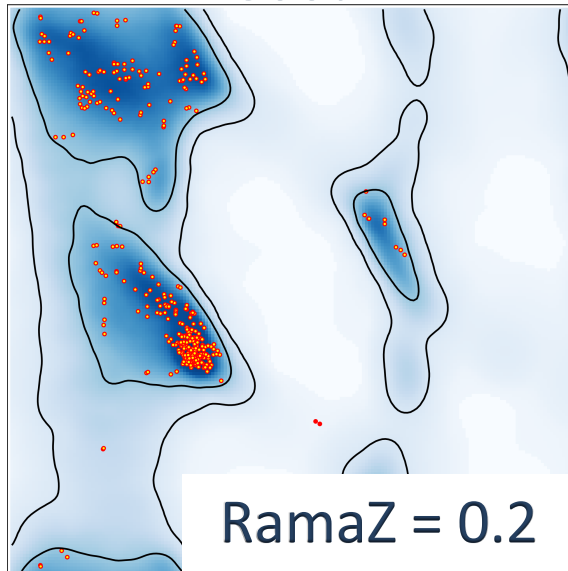
Oleg V. Sobolev,<sup>1,5,\*</sup> Pavel V. Afonine,<sup>1</sup> Nigel W. Moriarty,<sup>1</sup> Maarten L. Hekkelman,<sup>2,3</sup> Robbie P. Joosten,<sup>2,3,\*</sup> Anastassis Perrakis,<sup>2,3</sup> and Paul D. Adams<sup>1,4</sup>

# Model validation: *Ramachandran plot Z-score*

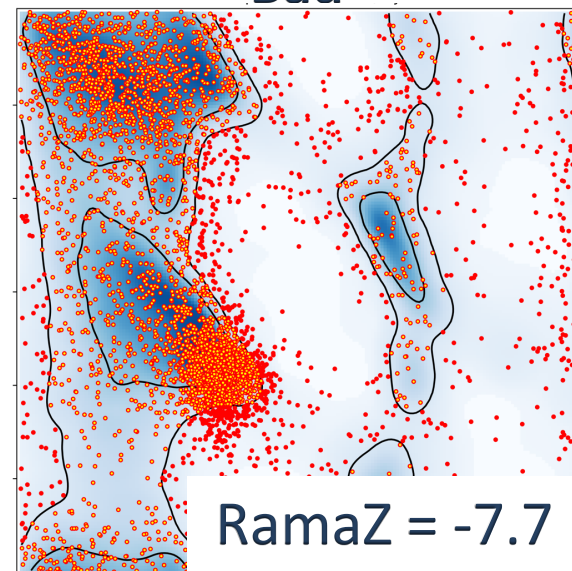
Good



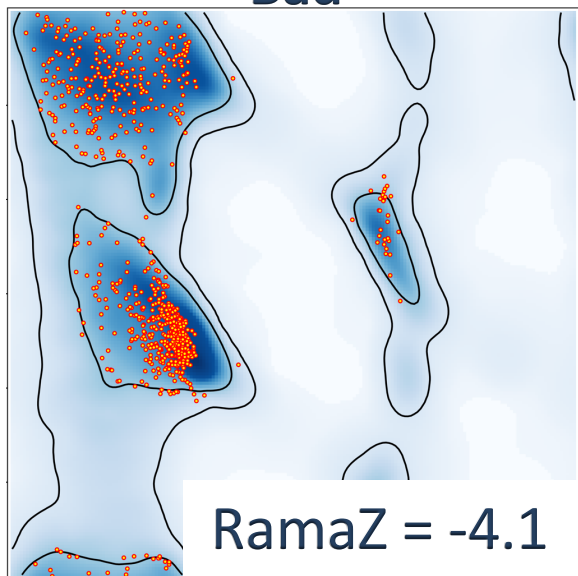
Good



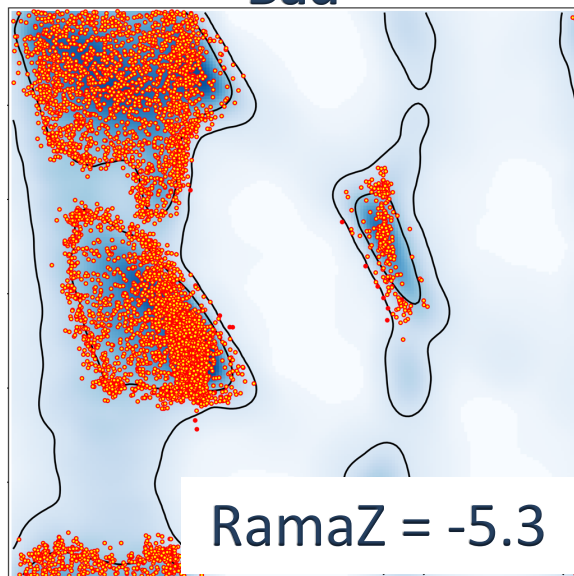
Bad



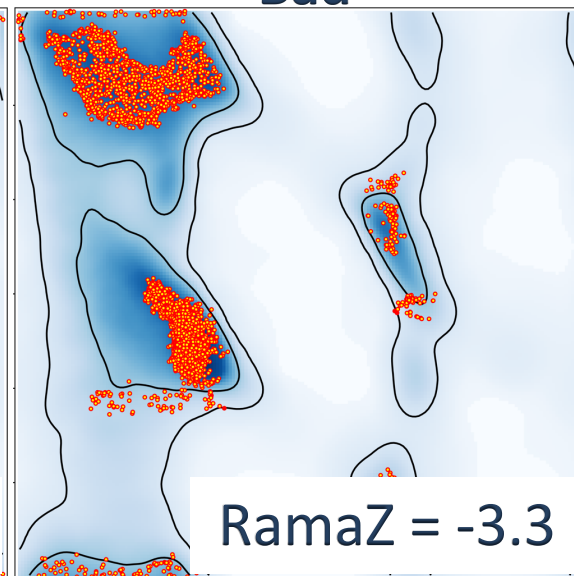
Bad



Bad

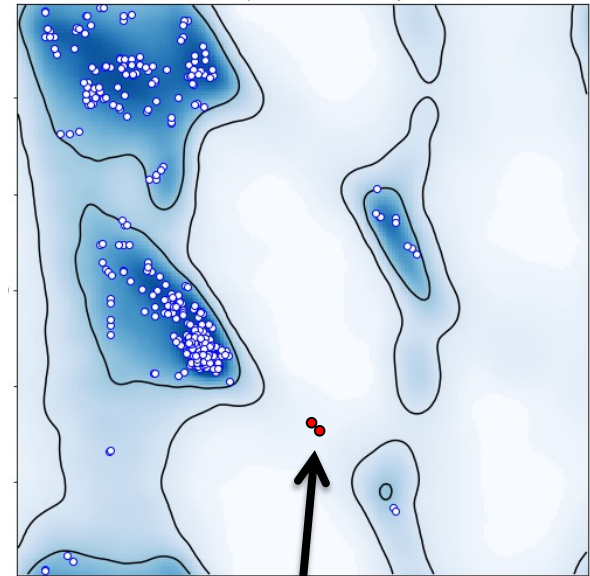
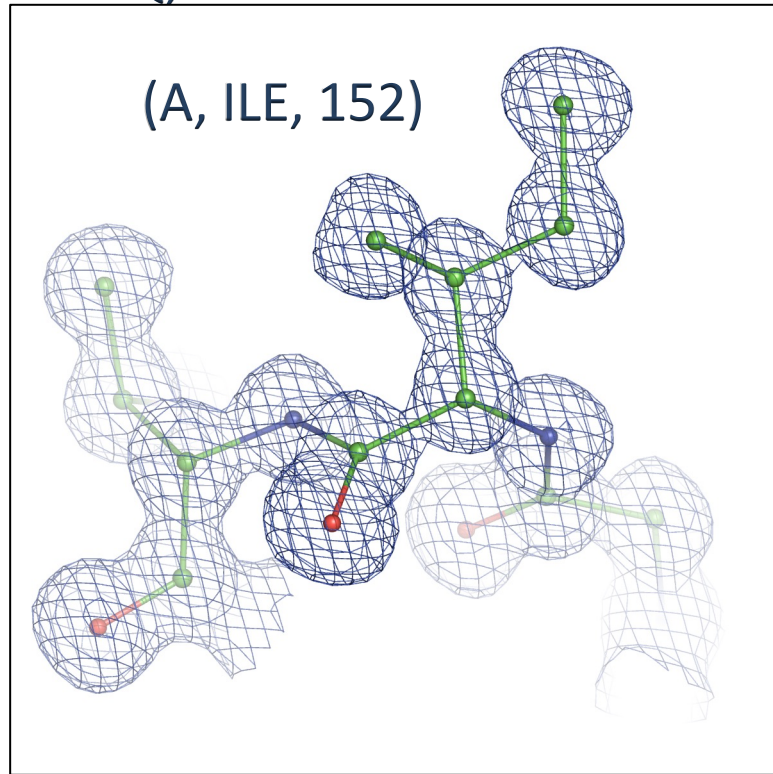


Bad



# An outlier $\neq$ wrong

3NOQ, 1 Å



Outliers:

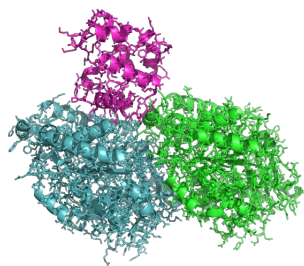
(A, ILE, 152), (B, ILE, 154)

- All outliers need to be explained (supported by the data)

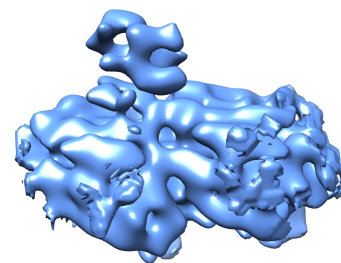
# Refinement success is function of data quality

- Do validation

**Model**

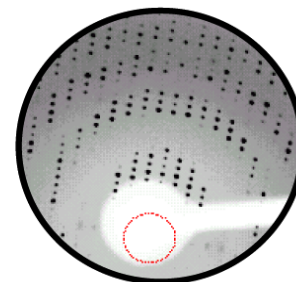


**Data**



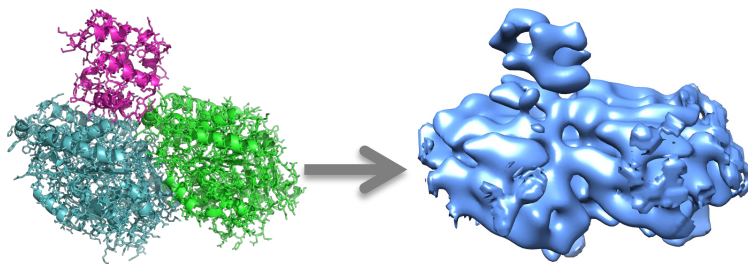
**Cryo-EM**

or



**Diffraction**

**Model to data fit**





# Validation tools in Phenix

PHENIX home

Quit Preferences Help Citations Coot PyMOL KING Other tools Ask for help

Actions Job history

### Projects

Show group: All groups Manage...

Select Delete New project Settings

ID	Last modified	# of jobs	R-free
✓ ChrisF	Apr 13 2020 09:42...	28	0.1944
real-space-refin...	Apr 03 2020 07:42...	2	---
zzz1	Mar 21 2020 09:10...	1	---
chris	Mar 12 2020 12:27...	11	0.1890
dan	Mar 11 2020 05:44...	1	---
3j63	Mar 11 2020 02:28...	1	---
jason	Mar 11 2020 11:36...	1	---
rt6	Mar 11 2020 10:31...	1	0.2459
mate	Mar 10 2020 01:36...	1	---
emily	Mar 09 2020 03:52...	3	---
—	Mar 05 2020 08:25...	3	0.1923
alex	Feb 27 2020 11:33...	6	---
rt20201	Feb 18 2020 12:50...	4	0.2213
1f8t	Feb 03 2020 09:00...	1	0.1977
real-space-refin...	Jan 30 2020 02:38...	2	---
real-space-refin...	Jan 29 2020 10:56...	1	---
ion_channel_den...	Jan 27 2020 07:36...	3	---
10101	Jan 27 2020 12:38...	2	---
demos	Jan 27 2020 10:57...	3	---
ion_channel_den...	Jan 27 2020 10:03...	2	---
malcolm	Jan 22 2020 10:22...	14	0.1748
real-space-refin...	Jan 16 2020 04:28...	3	---
3NIR	Dec 05 2019 10:2...	1	---
leighton	Sep 02 2019 05:1...	2	---
5pti	Aug 27 2019 03:4...	3	---

### Favorites

#### Data analysis

- Xtrriage**  
Analysis of data quality and crystal defects
- Merging statistics**  
Calculates a variety of statistics for unmerged intensities, including I/sigma, R-merge, R-meas, and CC1/2.
- Mtrriage**  
Analyze quality of maps in CCP4 format

#### Experimental phasing

#### Molecular replacement

#### Model building

#### Refinement

#### Cryo-EM

#### Validation

- Comprehensive validation (X-ray/Neutron)**  
Model quality assessment, including real-space correlation and geometry inspection using MolProbity tools
- Comprehensive validation (cryo-EM)**  
Model quality assessment, including real-space correlation, for cryo-EM structures

#### Structure comparison

Identify differences between multiple structures of the same protein, using multiple criteria

#### Calculate CC\*

Comparison of unmerged data quality with refined model, as described in Karplus & Diederichs (2012)

#### EMRinger

Model validation for de novo electron microscopy structures

### Ligands

Current directory: /Users/pafonine/Desktop/all/people/ChrisF Browse... 🔍

PHENIX version dev-svn-000 Project: ChrisF

# Xtrriage: all about your diffraction data

- Matthews coefficient probabilities
- Completeness by resolution
- Wilson plot sanity
- Detection of translational NCS (tNCS)
- Analysis of systematic absences and combination of tNCS with current space group
- Anomalous signal from measurability analysis
- Symmetry and twinning analyses
- Alternative point-group symmetry (can be detected on the basis of an R-value analyses)

# Xtrriage

Xtrriage (Project: porin-twin)

Preferences Help Run Abort View log Save graph Ask for help

Configure **Xtrriage\_1**

Run status **Results**

Xtrriage summary

- Intensity statistics suggest twinning (intensities are significantly different from expected for normal data) and one or more twin operators show a significant twin fraction.
- Translational NCS does not appear to be present.
- Ice rings do not appear to be present.
- The fraction of outliers in the data is less than 0.1%.
- The data are not significantly anisotropic.
- The resolution cutoff appears to be similar in all directions.
- The overall completeness in low-resolution shells is at least 90%.
- Overall completeness is above 90%.

# PDB deposition

The screenshot shows the Phenix software interface. The top menu bar includes 'Phenix home' and various tool icons. The main window is divided into a 'Projects' panel on the left and a main menu on the right. The 'PDB Deposition' section of the main menu is highlighted with an orange oval. The 'Current directory' at the bottom is set to '/Users/dcliebschner/Documents/AF\_POMGNT2\_1'.

**Phenix home**

Quit Preferences Help Citations Reload last job ChimeraX Coot PyMOL KiNG Tools Help Server

**Actions** Job history

**Projects**

Show group: All groups Manage...

Select Delete New project Import project Settings

ID	Last modified	# of jobs	R-free
✓ AF_POMGNT2_1	Jun 05 2024 11:46...	3	---
bugs	May 30 2024 02:38...	12	---
02_test_comma...	May 24 2024 01:20...	17	---
tests	May 22 2024 11:15...	67	0.2650
AF_bromodomai...	May 16 2024 10:37...	1	---
AF_7mjs_H_Pre...	Mar 19 2024 09:54...	1	---
groel_dock_refine	Mar 19 2024 09:28...	4	---
bugs_playground	Mar 07 2024 04:43...	13	---
fmodel	Feb 28 2024 02:44...	30	---
SEACOAST	Feb 13 2024 01:09...	7	---
AF_7mjs_H_Pre...	Jan 03 2024 10:19 ...	4	---
joint_XN	Nov 02 2023 03:49...	50	0.0989
AF_7mjs_H_Pre...	Apr 13 2023 02:18 ...	20	---
AF_7mjs_H_Pre...	Apr 13 2023 09:35 ...	0	---
AF_POMGNT2_0	Mar 31 2023 07:07...	3	---
AF_POMGNT2	Mar 30 2023 09:07...	6	---
7brm	Mar 17 2023 11:39...	25	---
7mjs_wcsbw	Mar 17 2023 09:31...	33	---
presentation	Mar 15 2023 02:00...	17	---
bughaton	Mar 06 2023 03:23...	8	---
-----	-----	-----	-----

**maps (create, manipulate, compare)**

**Enhanced maps (Polder, FEM, density-modified...)**

**Model building**

**Refinement**

**Ligands**

**Cryo-EM: Map analysis, symmetry, manipulation**

**Validation and map-based comparisons**

**Map improvement**

**Docking, model building and rebuilding**

**Refinement**

**Models: Superpose, search, compare, analyze symmetry**

**Modification, minimization and dynamics**

**PDB Deposition**

- Prepare model for PDB deposition**  
Finalize mmCIF files for deposition to the PDB
- Get PDB validation report**  
Retrieve a validation report from the PDB
- Generate "Table 1" for journal**  
Extraction of final model statistics for publication

**Program search**

Current directory: /Users/dcliebschner/Documents/AF\_POMGNT2\_1 Browse...

Phenix version 1.21.1-5286-000 Project: AF\_POMGNT2\_1

# PDB deposition

mmCIF format is mandatory for deposition as of 2019



STRUCTURAL  
BIOLOGY

ISSN 2059-7983

## Announcing mandatory submission of PDBx/mmCIF format files for crystallographic depositions to the Protein Data Bank (PDB)

**Paul D. Adams,<sup>a,b</sup> Pavel V. Afonine,<sup>a</sup> Kumaran Baskaran,<sup>c</sup> Helen M. Berman,<sup>d</sup> John Berrisford,<sup>e</sup> Gerard Bricogne,<sup>f</sup> David G. Brown,<sup>g</sup> Stephen K. Burley,<sup>d,h,i,\*</sup> Minyu Chen,<sup>j</sup> Zukang Feng,<sup>d</sup> Claus Flensburg,<sup>f</sup> Aleksandras Gutmanas,<sup>e</sup> Jeffrey C. Hoch,<sup>k,\*</sup> Yasuyo Ikegawa,<sup>j</sup> Yumiko Kengaku,<sup>j</sup> Eugene Krissinel,<sup>l</sup> Genji Kurisu,<sup>j,\*</sup> Yuhe Liang,<sup>d</sup> Dorothee Liebschner,<sup>a</sup> Lora Mak,<sup>e</sup> John L. Markley,<sup>c,\*</sup> Nigel W. Moriarty,<sup>a</sup> Garib N. Murshudov,<sup>m</sup> Martin Noble,<sup>n</sup> Ezra Peisach,<sup>d</sup> Irina Persikova,<sup>d</sup> Billy K. Poon,<sup>a</sup> Oleg V. Sobolev,<sup>a</sup> Eldon L. Ulrich,<sup>c</sup> Sameer Velankar,<sup>e,\*</sup> Clemens Vonrhein,<sup>f</sup> John Westbrook,<sup>d</sup> Marcin Wojdyr,<sup>f,l</sup> Masashi Yokochi<sup>j</sup> and Jasmine Y. Young<sup>d</sup>**

Received 21 February 2019  
Accepted 3 April 2019

Edited by R. J. Read, University of Cambridge,  
England

# PDB deposition: mmCIF facts

- Contains a lot more information than PDB
- Not intended to be human editable
  - You can read it but it is (much) harder than PDB
- Phenix tools generally produce output in mmCIF format
- Avoid editing by hand
  - Easy to make hard-to-recover mistakes

# PDB deposition: CIF file confusion

- CIF is a file format
- CIF file can contain:
  - Ligand information
  - Atomic model
  - Reflection data
  - Any mixture of three above

# PDB deposition: dos and don'ts

- Do not change the content of files from refinement for any reason:
  - Add/remove atoms (hydrogens, water)
  - Edit labels, header information
- Run Comprehensive validation (Phenix GUI) to address all outstanding issues before deposition
- Don't panic if validation statistics reported by Phenix does not match PDB validation report
  - If that happens and presents a problem – start conversation with PDB stuff and involve Phenix developers
- Once all is deposited and up on the web – check everything: mistakes at PDB end happen



# User support

- **Feedback, questions, help**

Mailing list (anyone signed up):

[phenixbb@phenix-online.org](mailto:phenixbb@phenix-online.org)

Bug reports (developers only):

[bugs@phenix-online.org](mailto:bugs@phenix-online.org)

Ask for help (developers only):

[help@phenix-online.org](mailto:help@phenix-online.org)

- **Reporting a bug or asking for help:**

- We can't help you if you don't help us to understand your problem
- Make sure the problem still exist using the latest *Phenix* version
- Send us all inputs (files, non-default parameters) and tell us steps that lead to the problem
- All data sent to us is kept confidentially

# The Project



## Lawrence Berkeley Laboratory

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



## University of Cambridge

Randy Read, Airlie McCoy,  
Tristan Croll, Claudia Millán Nebot,  
Rob Oeffner



## Los Alamos National Laboratory New Mexico Consortium

Tom Terwilliger, Li-Wei Hung



## UTHealth

Matt Baker, Corey Hyrc



## Duke University

Jane & David Richardson,  
Christopher Williams,  
Vincent Chen



An NIH/NIGMS funded  
Program Project

Liebschner D, *et al.*, Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in *Phenix*. Acta Cryst. 2019 **D75**:861–877