AQuaRef: AI-enabled Quantum Refinement

Pavel Afonine

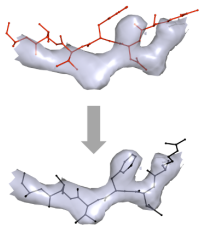
April 30, 2026
Lincoln, NE

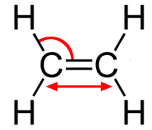
https://hbr.org/2023/08/ai-wont-replace-humans-but-humans-with-ai-will-replace-humans-without-ai

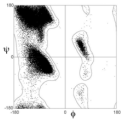
Restraints and data resolution

T = T_{DATA} + **w** * $T_{RESTRAINTS}$

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

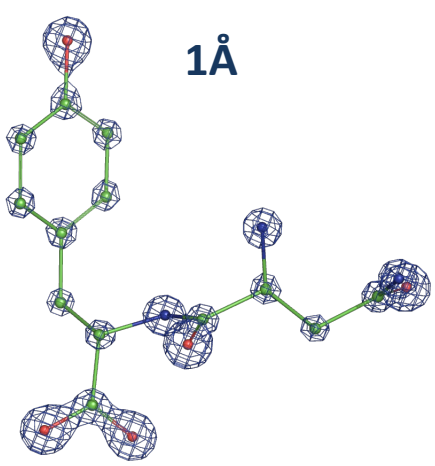


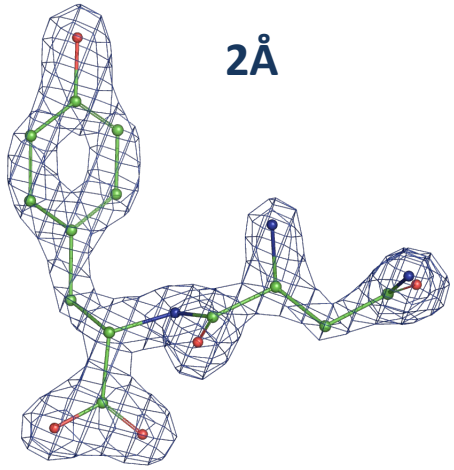


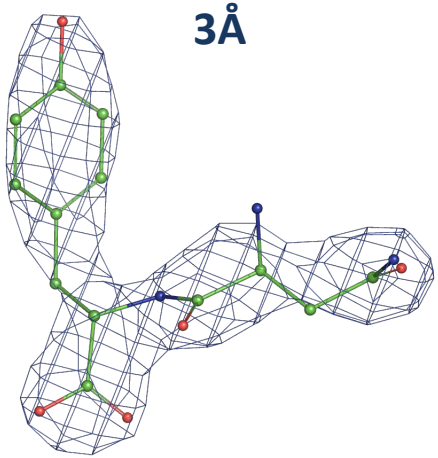


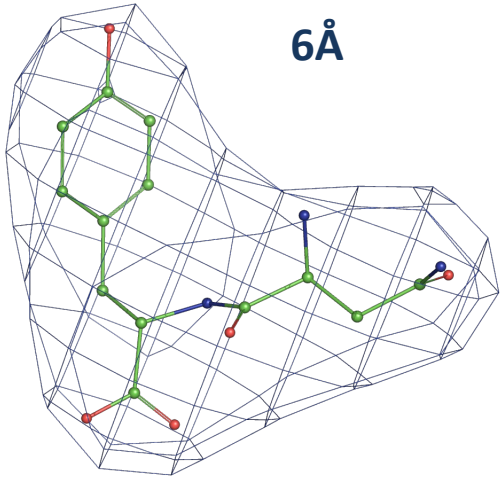
Bonds, angles, planes, torsions, chirality, non-bonded repulsion

$$T_{RESTRAINTS} = T_{BOND} + T_{ANGLE} + T_{DIHEDRAL} + T_{PLANE} + T_{REPULSION} + T_{CHIRALITY}$$





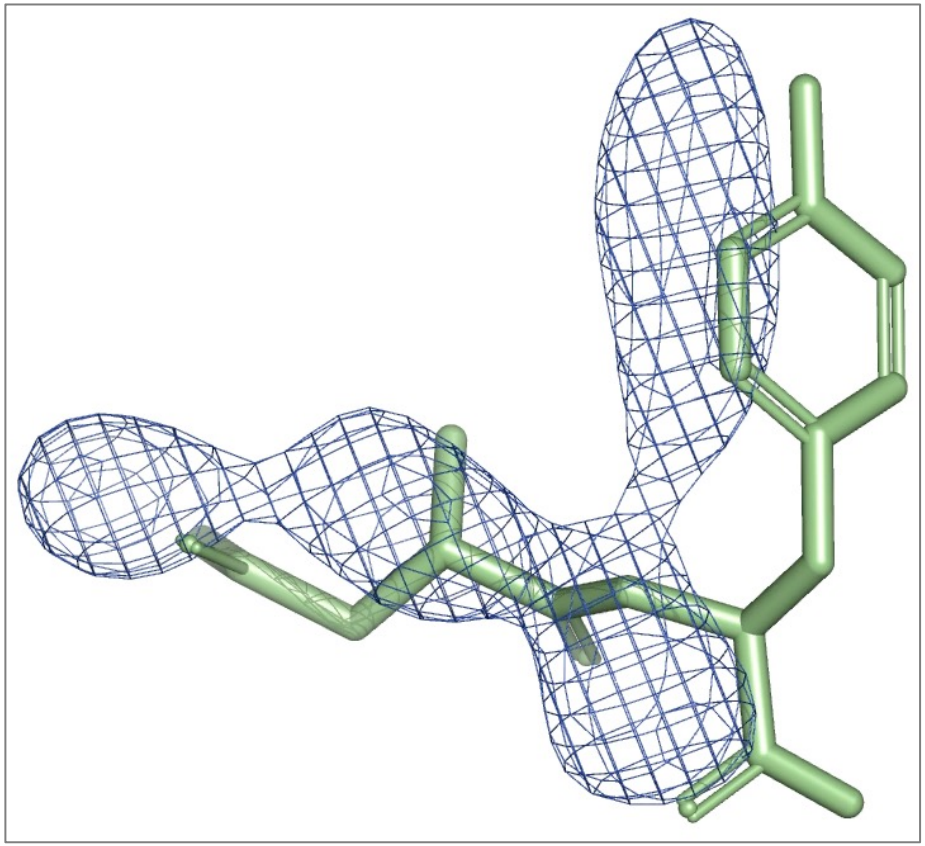




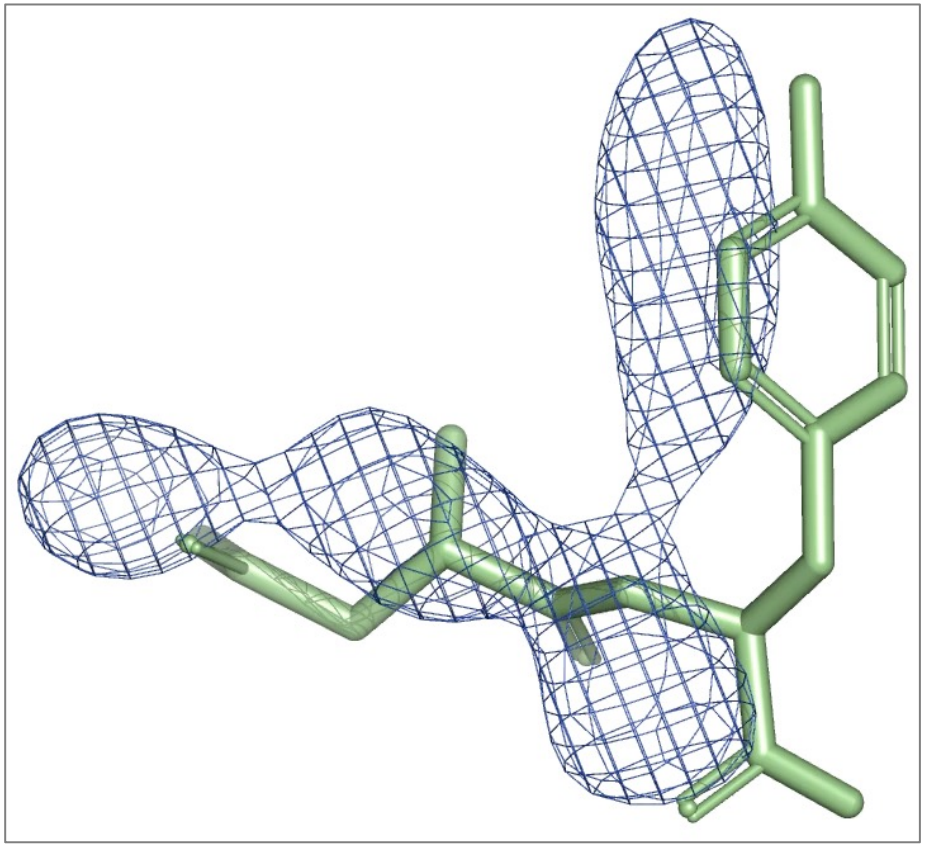
High Resolution Low

Model refinement with *vs* no restraints

$$T = T_{DATA} + w * T_{RESTRAINTS}$$



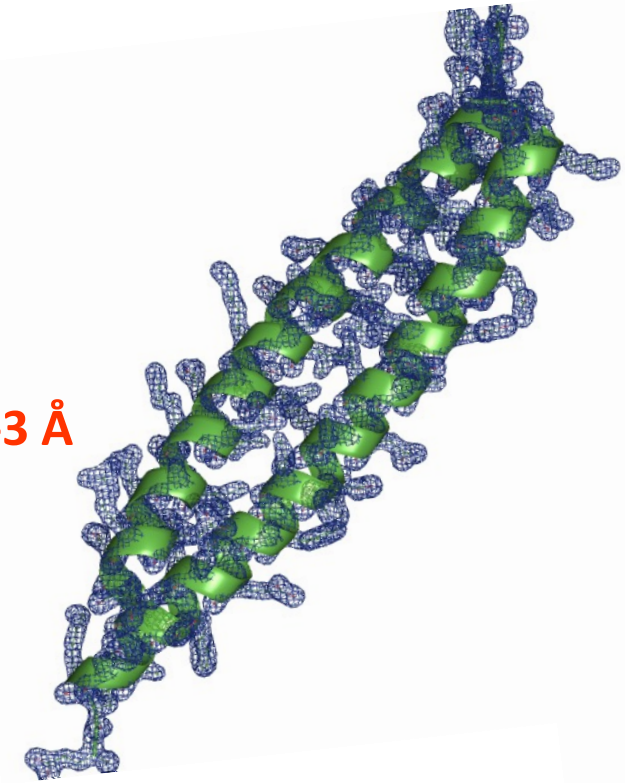
Using restraints



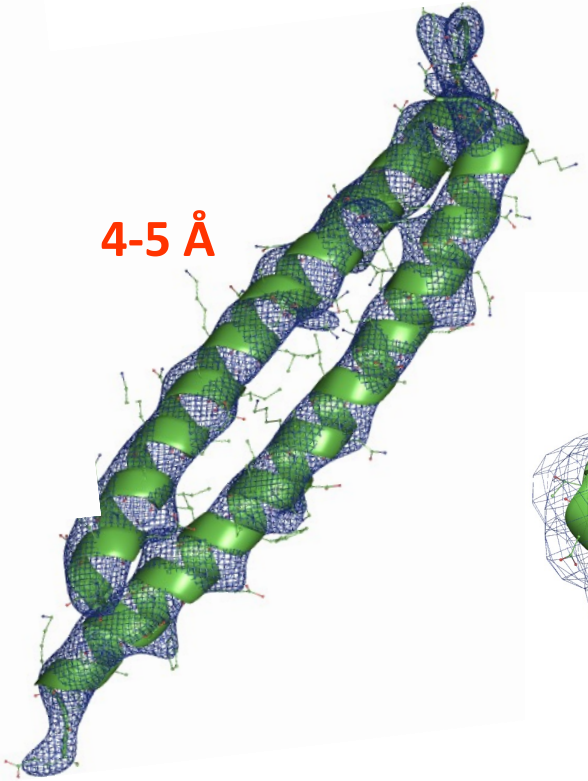
Not using restraints

Restraints: low resolution

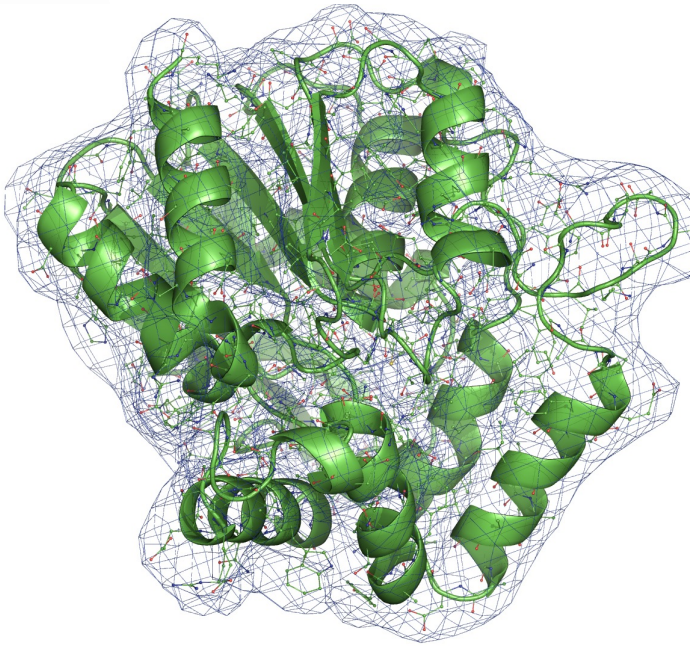
2-3 Å



4-5 Å

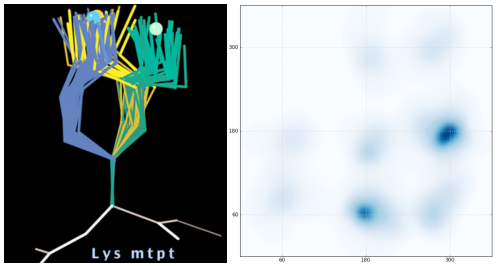


6Å-lower

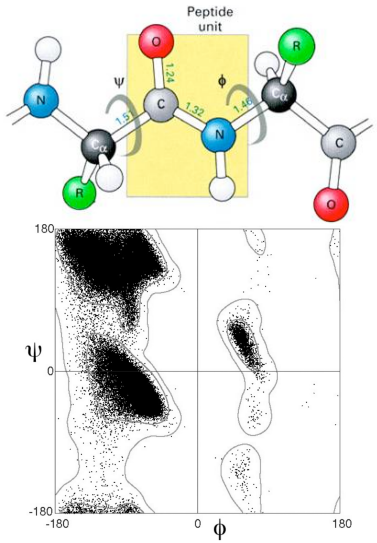


More restraints for low resolution

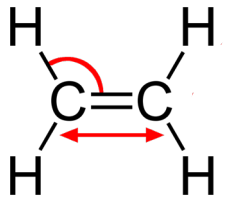
Side chain distributions

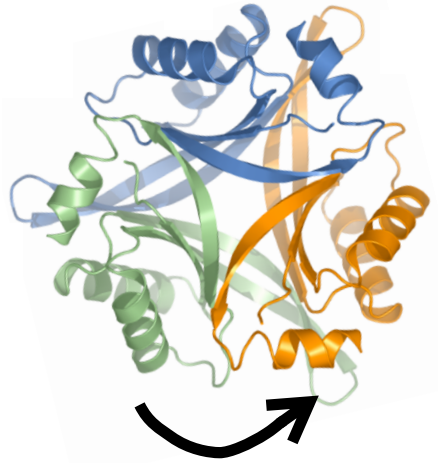


Main chain distributions

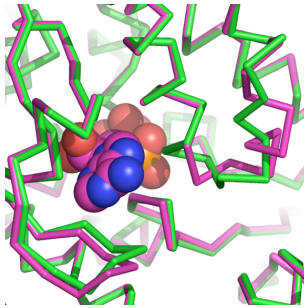


Covalent geometry

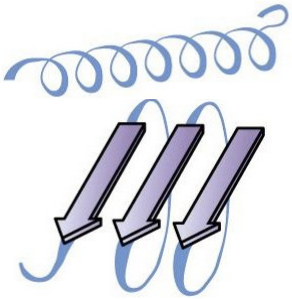




Internal symmetry (NCS)



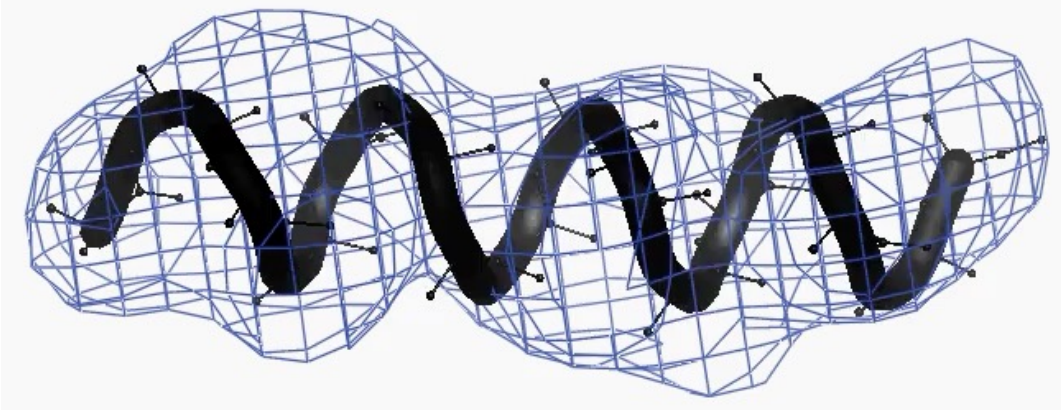
Similar (homologous) structures (reference model restraints)



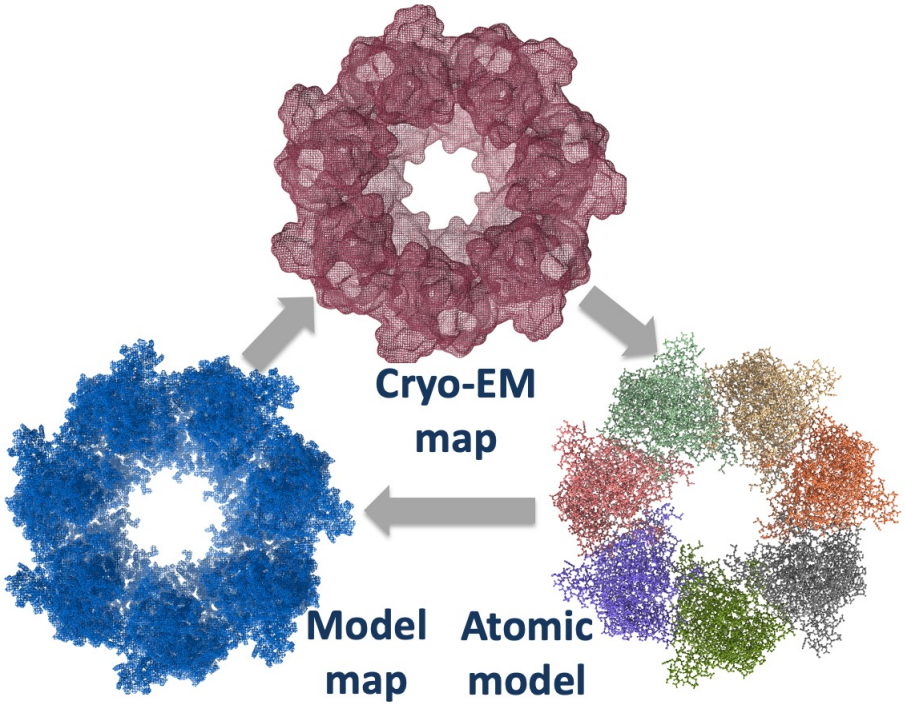
Secondary structure

Model refinement with insufficient restraints

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



Model-to-Map Fit: Refinement objective



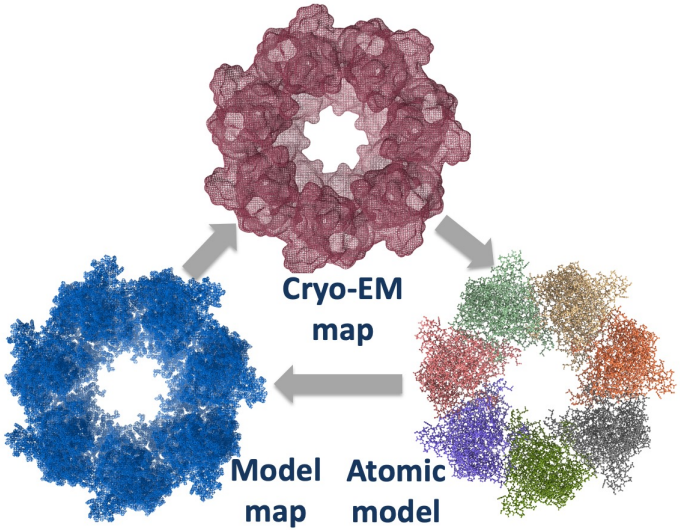
$$LS = \sum_{map} (\rho_{obs} - \rho_{calc})^2$$

$$\mathrm{CC} = \frac{\sum \rho_{obs}\rho_{calc}}{(\sum {\rho_{obs}}^2 \sum {\rho_{calc}}^2)^{1/2}}$$

150k atoms ~ 11 sec

Typical refinement: ~1000 evaluations

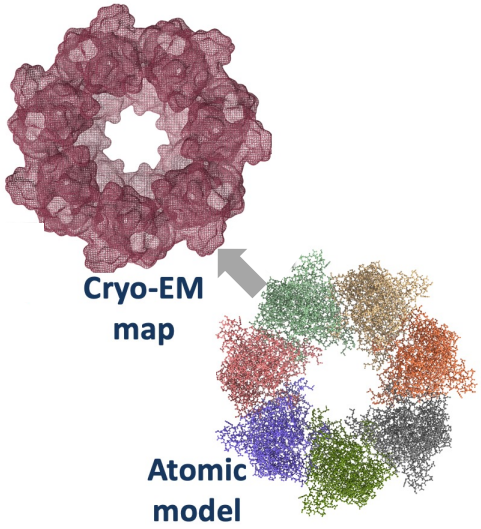
Model-to-Map Fit: Refinement objective



$$LS = \sum_{map} (\rho_{obs} - \rho_{calc})^2$$

$$\mathrm{CC} = \frac{\sum \rho_{obs}\rho_{calc}}{(\sum \rho_{obs}{}^2 \sum \rho_{calc}{}^2)^{1/2}}$$

150k atoms ~ 11 sec



$$T = -\sum_{atoms} \rho_{obs}(x_{atom}, y_{atom}, z_{atom})$$

Used in **`phenix.real_space_refine`**

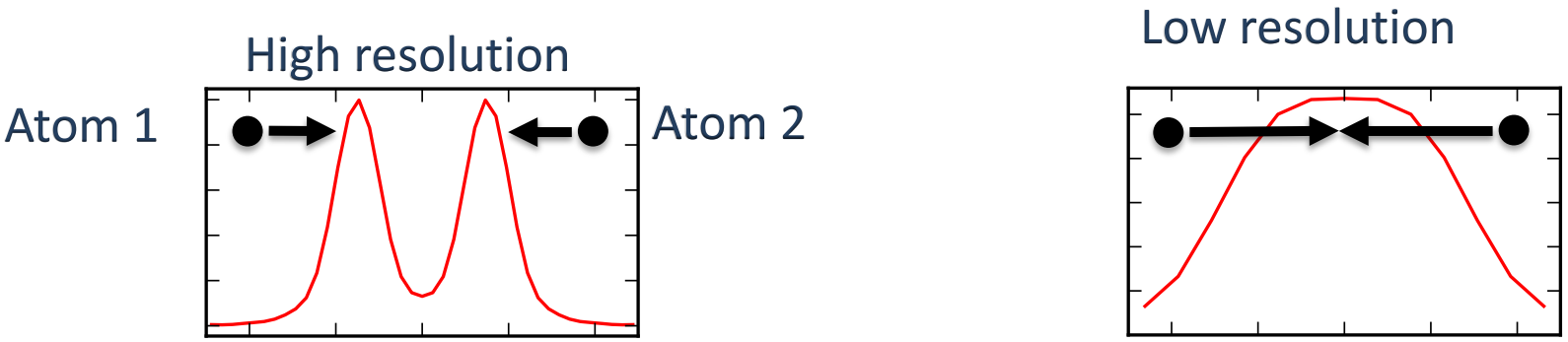
0.02 sec

Refinement target

$$T = -\sum_{atoms} \rho_{obs}(x_{atom}, y_{atom}, z_{atom})$$

is much less accurate than $$LS = \sum_{map} (\rho_{obs} - \rho_{calc})^2$$

because it aims at moving atoms towards nearest map peaks without assuming shape of the map:

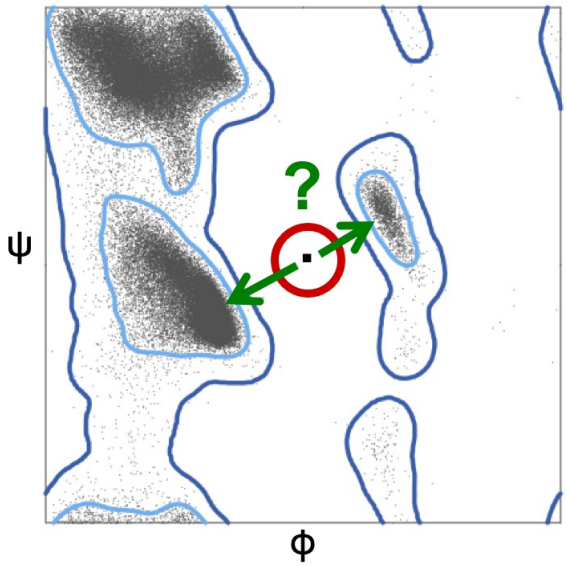


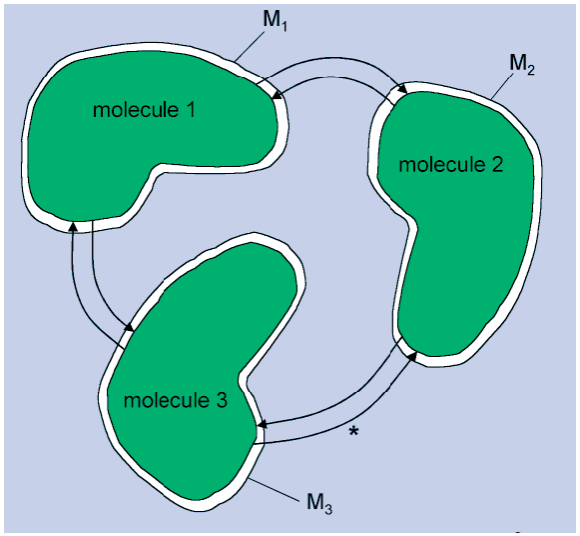
- Moving atoms to nearest peaks ≠ making correct model
- Lower resolution = less accurate
- Need to use a lot of geometric restraints in refinement

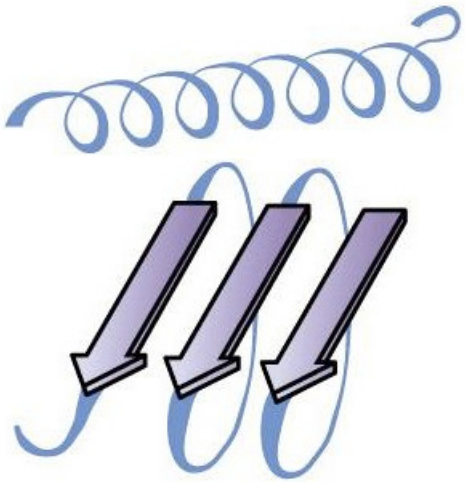
Validation and Refinement "conflict"

- Validation metrics progressively become refinement goals
    - Ramachandran plot restraints
    - Cβ deviation restraints
    - Secondary structure restraints
    - Restraints on χ angles of amino-acid side-chain rotamers
- As result, validation becomes less capable of catching problems

Setting up extra restraints: manual work & very error-prone

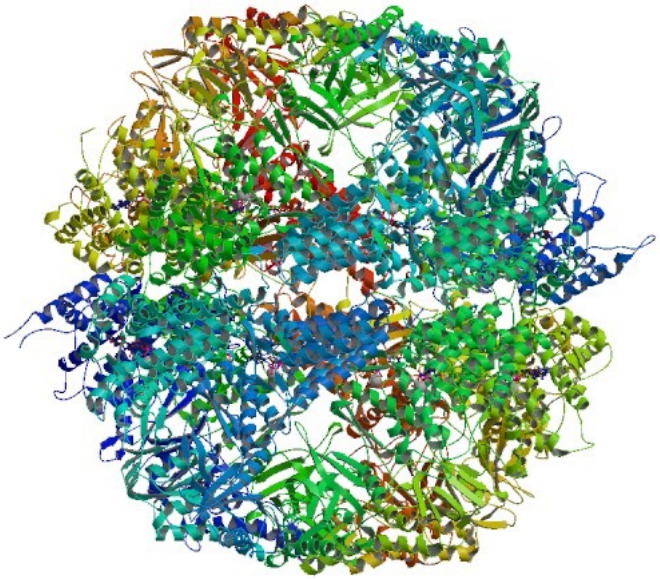






Example

PNAS, 2019 116 (39) 19513-19522

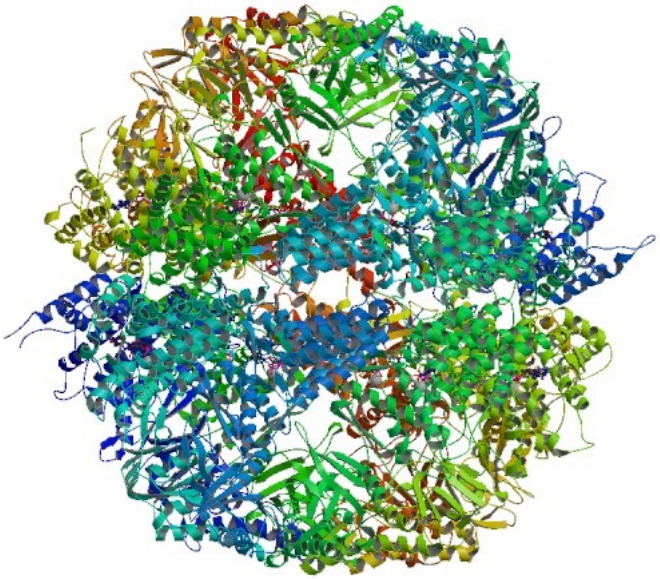


Metric / PDB code		6KS6
Clashscore		8.8
Rama. (%)	favored	96.4
	outliers	0.2
Rotamer outliers (%)		0
C_β deviations		0
RMSD	Bond (Å)	0.002
	Angle (°)	490
Resolution (Å)		3.0

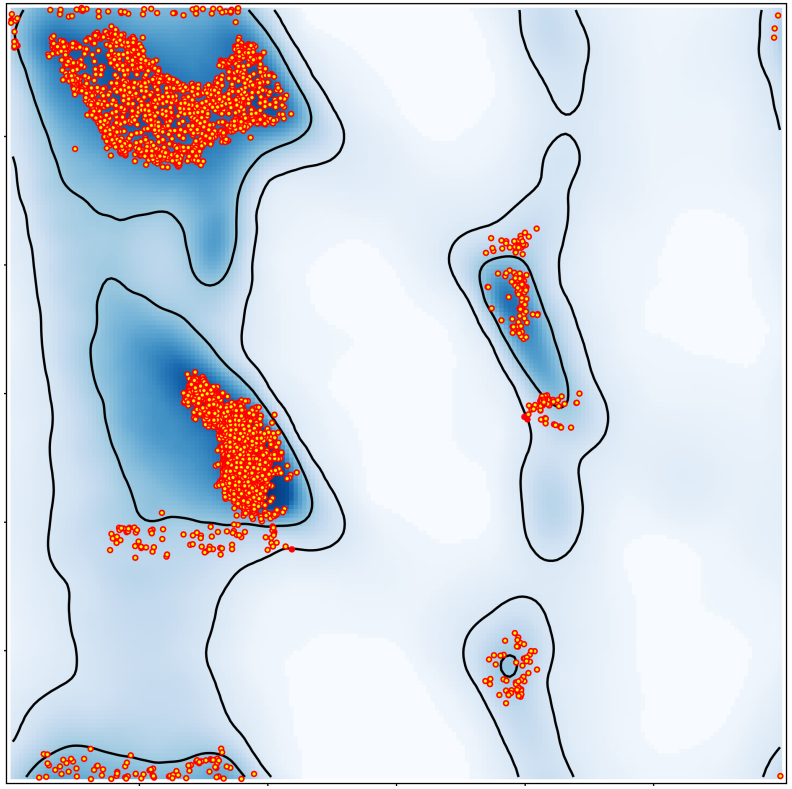
Perfect statistics! All looks just great!

Example

PNAS, 2019 116 (39) 19513-19522



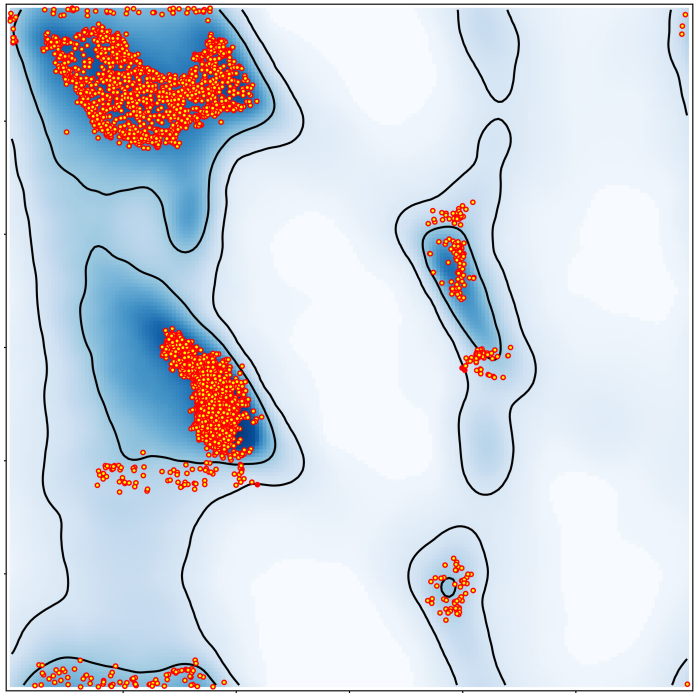
The plot looks very wrong!



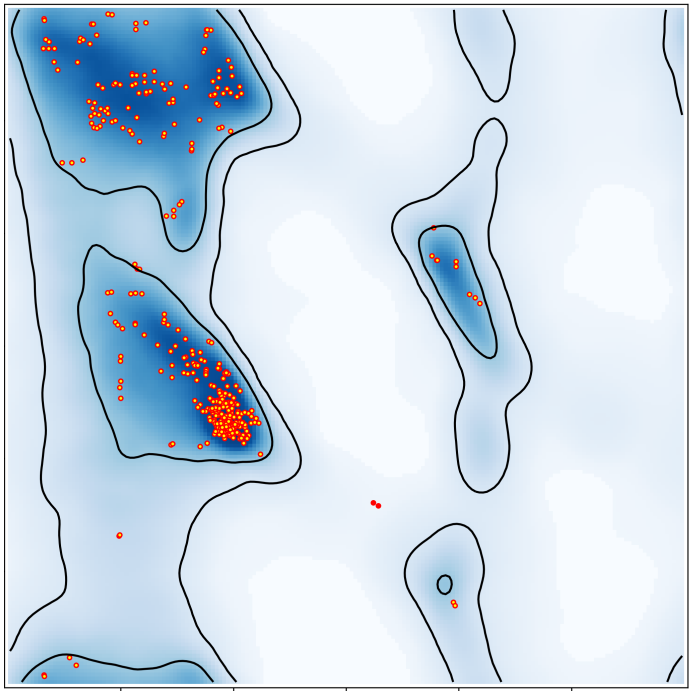
1. How we know the plot looks wrong?

2. How did that happen?

Q: How we know the plot looks wrong?



A: Because we know how good plot looks like!



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

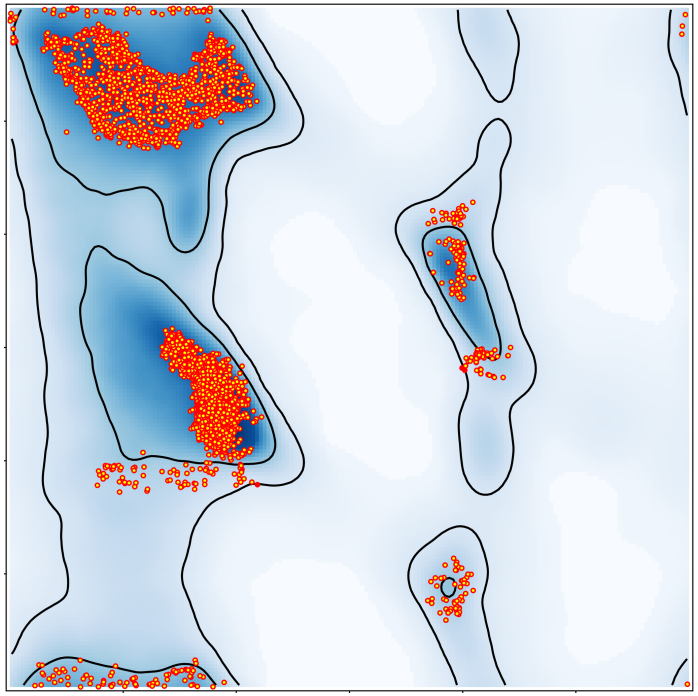
Resource

A Global Ramachandran Score Identifies Protein Structures with Unlikely Stereochemistry

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

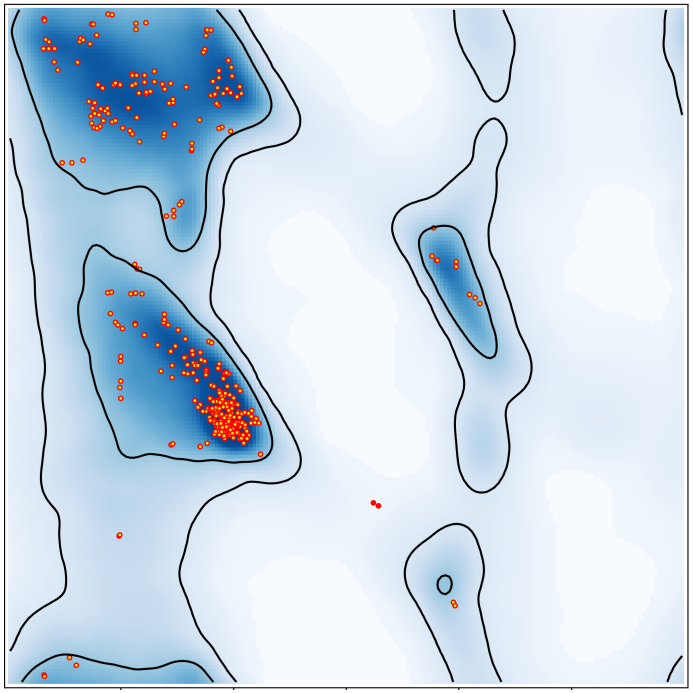
Ramachandran plot Z-score

Q: How we know the plot looks wrong?



RamaZ = -4.1 (Poor)

A: Because we know how good plot looks like!

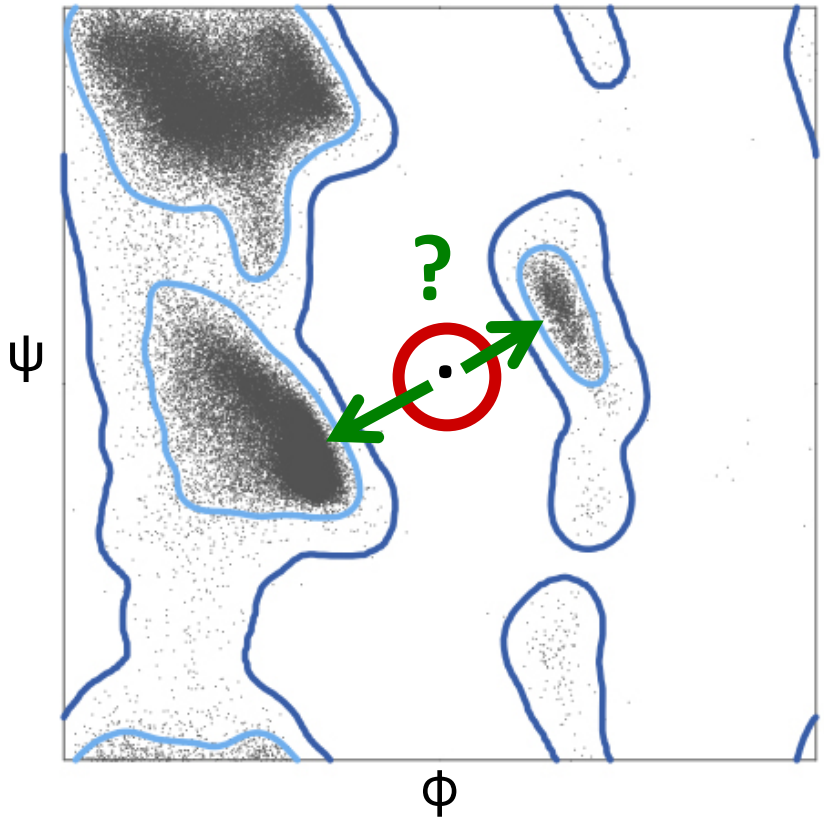


RamaZ = -1.9 (Good)

Poor: $|Z| > 3$ Suspicious: $2 < |Z| < 3$ Good: $|Z| < 2$

How did that happen?

$$E = \Sigma\, w * (\varphi_{model} - \varphi_{target})^2 + \Sigma\, w * (\psi_{model} - \psi_{target})^2$$

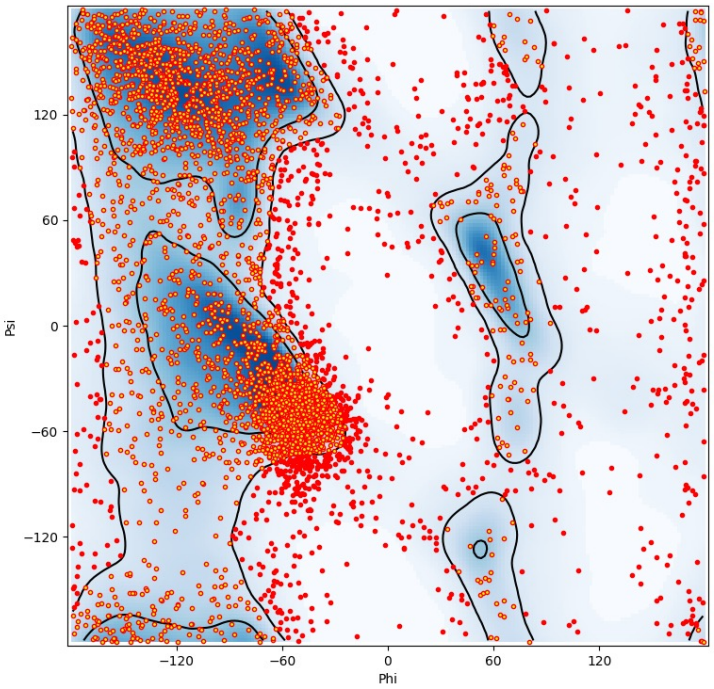


- Setting up Ramachandran plot, secondary structure, etc, restraints can be ambiguous and is error prone!

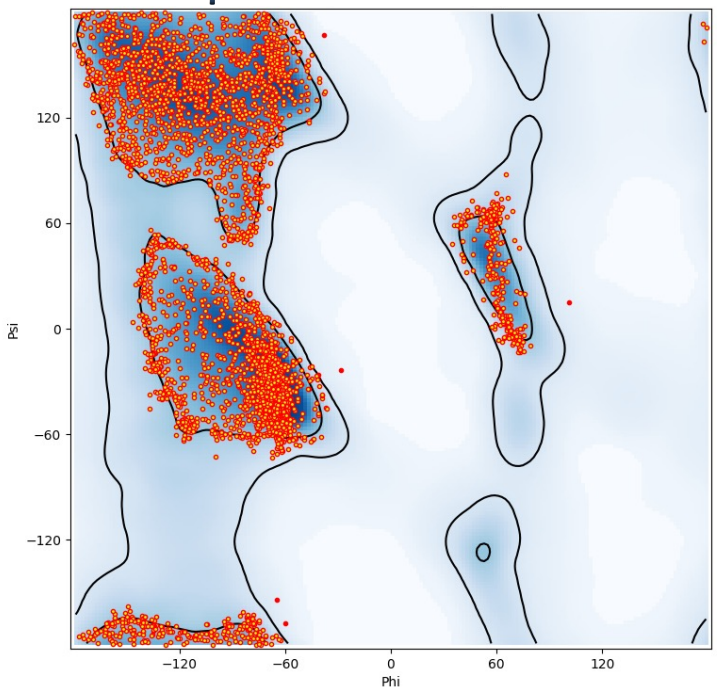
How did that happen?

PDB code: 5a9z

Original



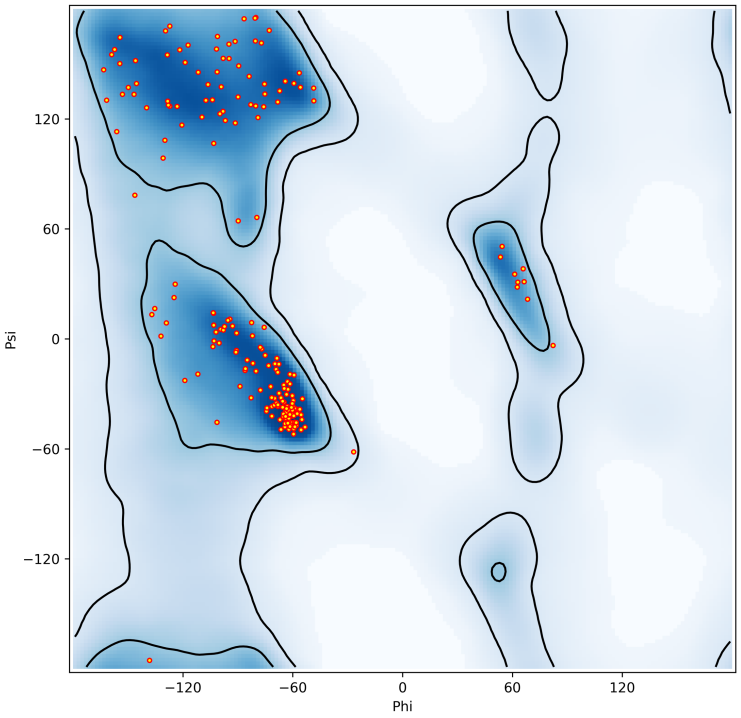
Refined with Ramachandran plot restraints



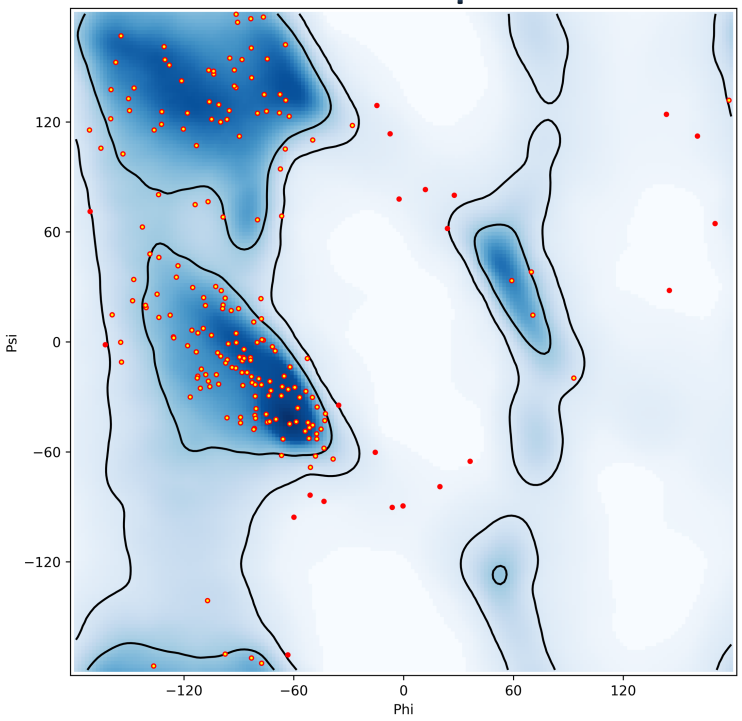
Don’t use Ramachandran plot restraints to remove outliers!

Ramachandran plot restraints

Before refinement



After refinement
(No Ramachandran plot restraints)



Use Ramachandran plot restraints to prevent outliers from occurring!

Restraints and limitations

$$T = T_{DATA} + w * T_{RESTRAINTS}$$

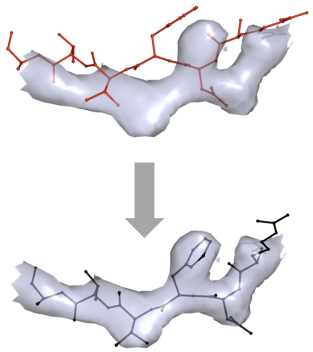
$$T_{RESTRAINTS} = T_{BOND} + T_{ANGLE} + T_{DIHEDRAL} + T_{PLANE} + T_{REPULSION} + T_{CHIRALITY}$$

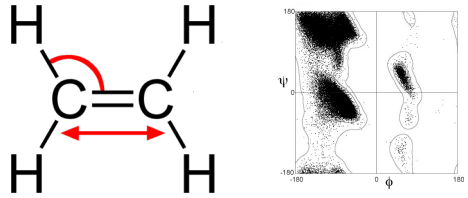
- Restraints are too limited:
  - No attraction terms (electrostatics, etc)
  - Not using information about protein structure (secondary structure, rotamers)
  - Limited to tabulated entities in the libraries (e.g., Monomer Library, GeoStd)

Restraints from QM

T = T_{DATA} + $w * T_{RESTRAINTS}$

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





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

Replace with energies/gradients from QM calculations

NEW: AQuaRef – QM based refinement in *Phenix*

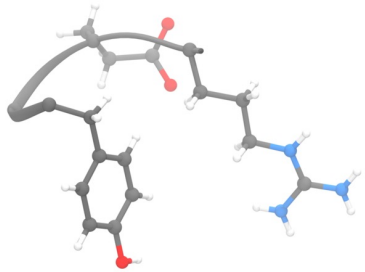
History of progress

Impossible

2010

QM Calculations

Impossible for proteins. Limited to small molecules only

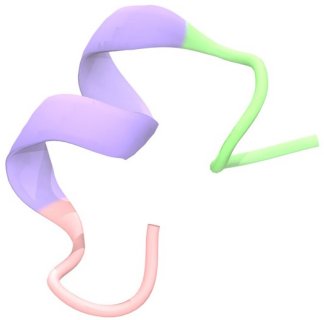


Limited

2012

GPU Accelerated QM

Limited to peptides and very small proteins (~hundreds of atoms)

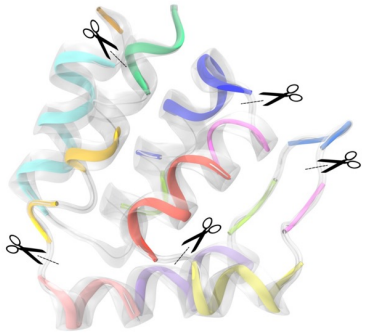


Possible

2017

Q|R with Fragmentation

QM-based protein refinement. Slow, resource-intensive, no inherent size limit

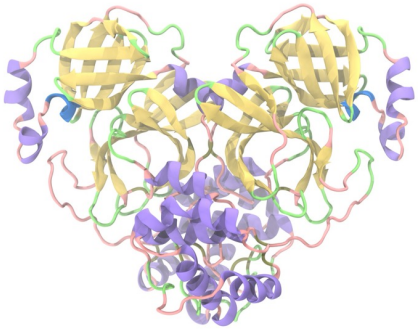


Practical

2024

ML Potentials

Fast, rivaling classical force fields, with QM-level accuracy and no fragmentation required



The Team

Carnegie Mellon University


Olexandr Isayev


Roman Zubatyuk


Hatice Gökcan

Uni. of Florida


Adrian Roitberg

Blend of expertise and background

Crystallography methods + Software development + QM expertise + ML / AI expertise

Pending AI


Marl Waller


Holger Kruse

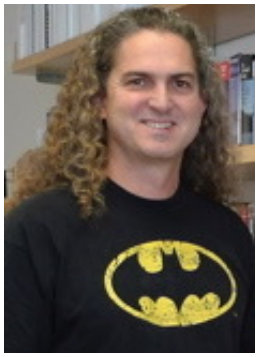
Uni of Wrocław


Malgo Biczysko

LBNL


Billy Poon


Pavel Afonine


Nigel Moriarty

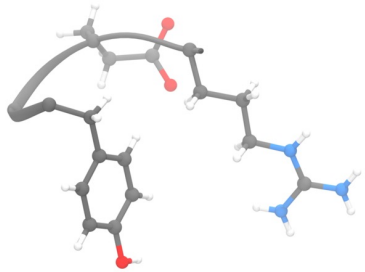
History of progress

Impossible

2010

QM Calculations

Impossible for proteins. Limited to small molecules only

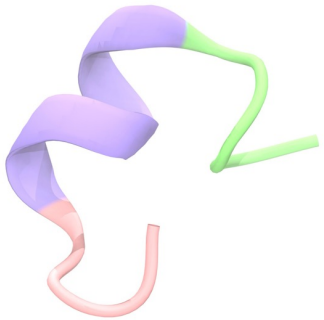


Limited

2012

GPU Accelerated QM

Limited to peptides and very small proteins (~hundreds of atoms)

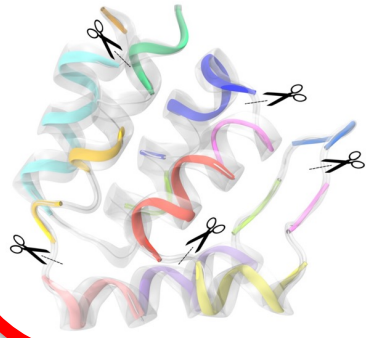


Possible

2017

Q|R with Fragmentation

QM-based protein refinement. Slow, resource-intensive, no inherent size limit

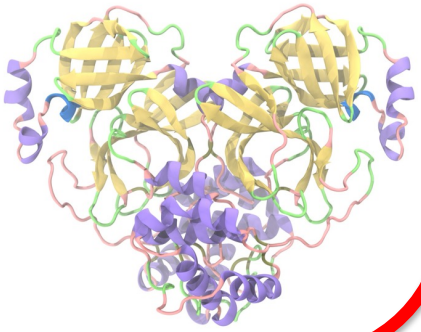


Practical

2024

ML Potentials

Fast, rivaling classical force fields, with QM-level accuracy and no fragmentation required



Machine Learning potential (AIMNet2)

Standard amino-acids

- Generate all possible 1-, 2-, 3-, and 4-peptides (including S-S bridges)
- Torsion and non-equilibrium sampling

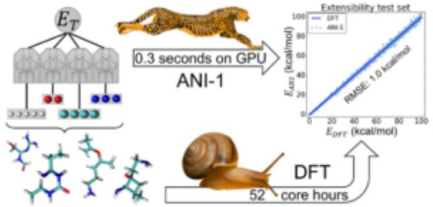
Large Dataset

DFT calculations

ML model

Calculation time:

About a week on one of big national computing resources



Smith, Justin S.; Isayev, Olexandr; Roitberg, Adrian E.

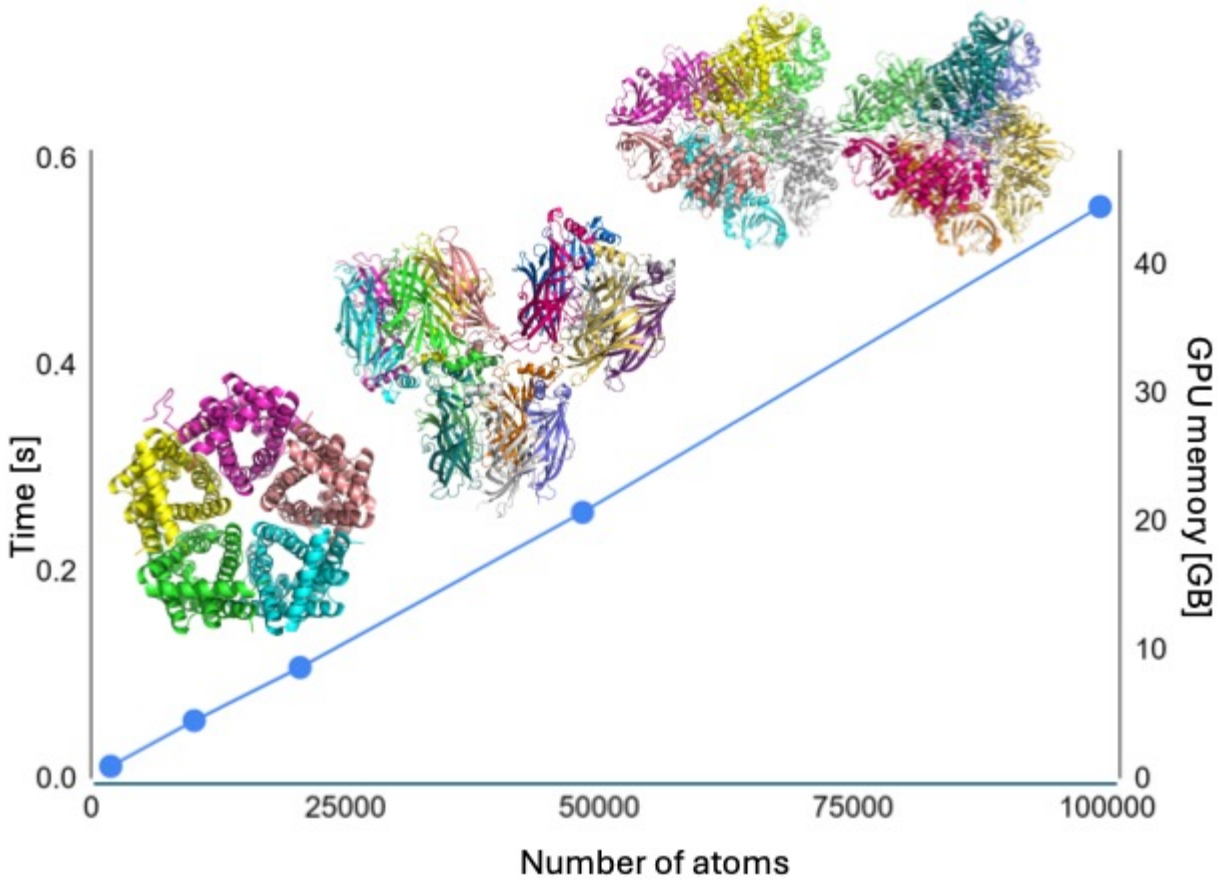
ANI-1: an extensible neural network potential with DFT accuracy at force field computational cost

Journal Article

In: Chemical Science, iss. 8, pp. 3192-3203, 2017.

Abstract | **Links** | BibTeX | Tags: ANI, Machine learning potential

Time & Memory Scaling: single energy calculation

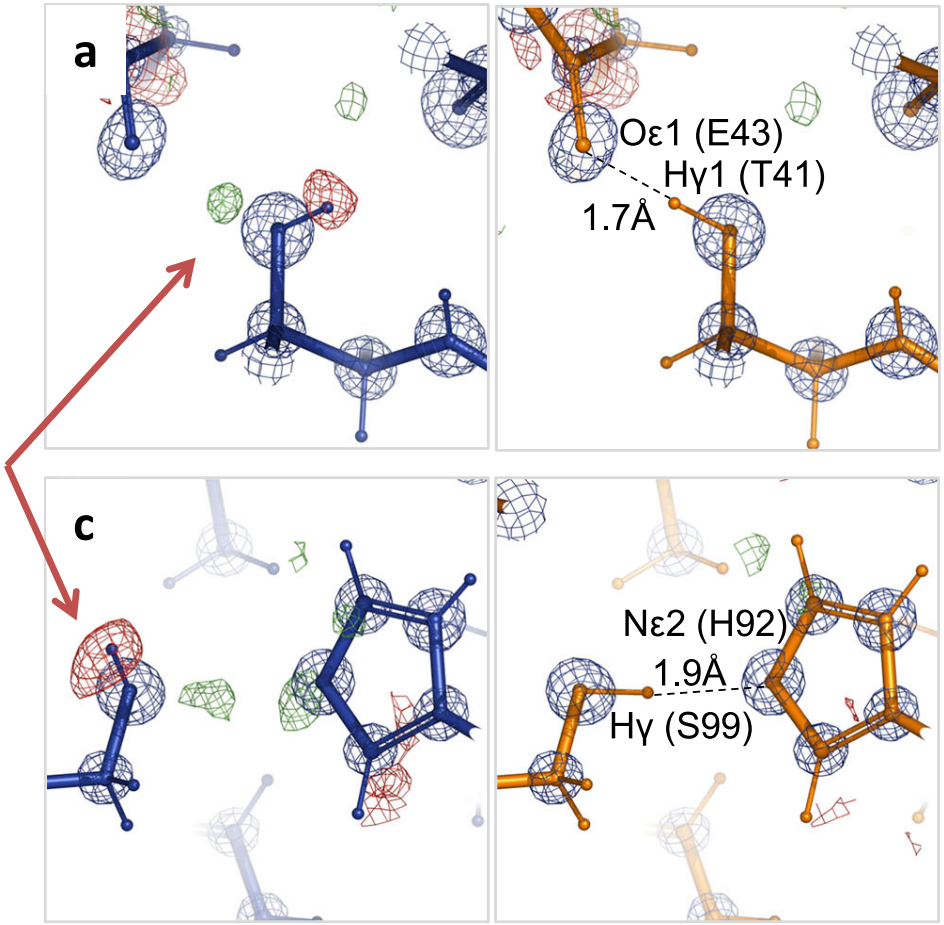


Scope of AQuaRef

- High resolution:
  - Standard library based restraints can hide specific low-signal chemistry in your structure
- Low resolution:
  - Data alone cannot meaningful geometry. Need a lot of restraints. Some of them hard to set correctly
- Any resolution:
  - Ligands. Ligands can be parameterized (”CIF files”). Ligand-protein interactions are governed by NCIs that are not parameterized in any refinement

AQuaRef: high resolution

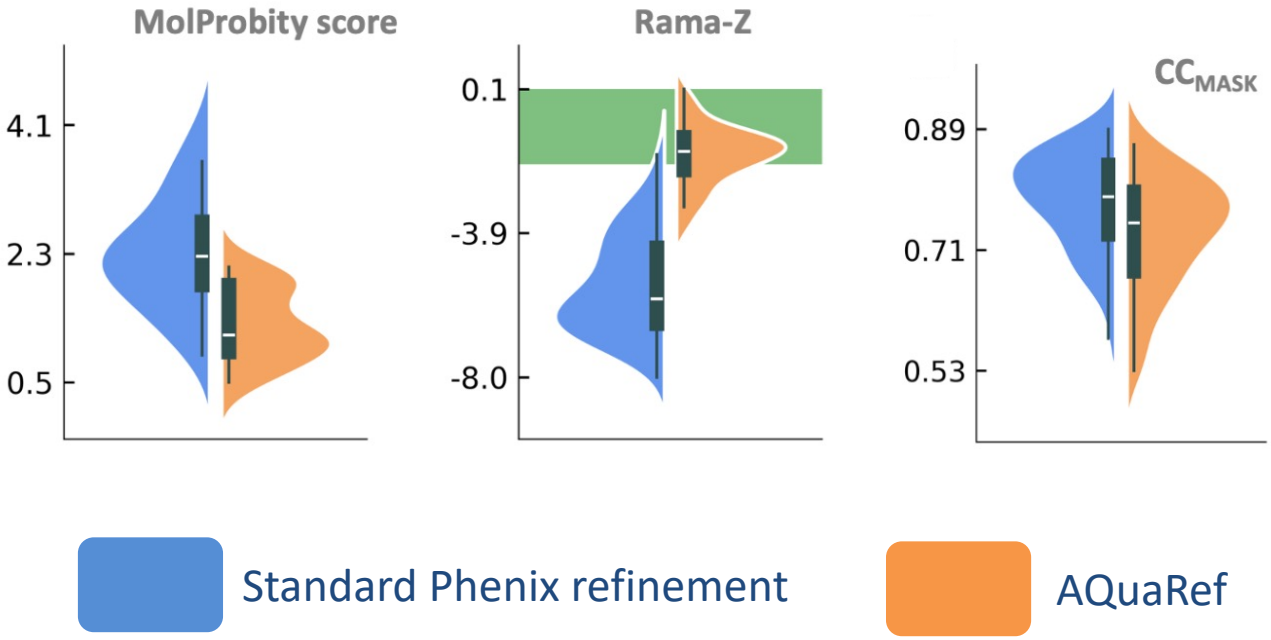
Initial position of H is far from data support (green blob), so it rotates thanks QM restraints awareness of a better orientation



Atoms or groups of atoms orient to make plausible NCIs if that does not require crossing energy barriers

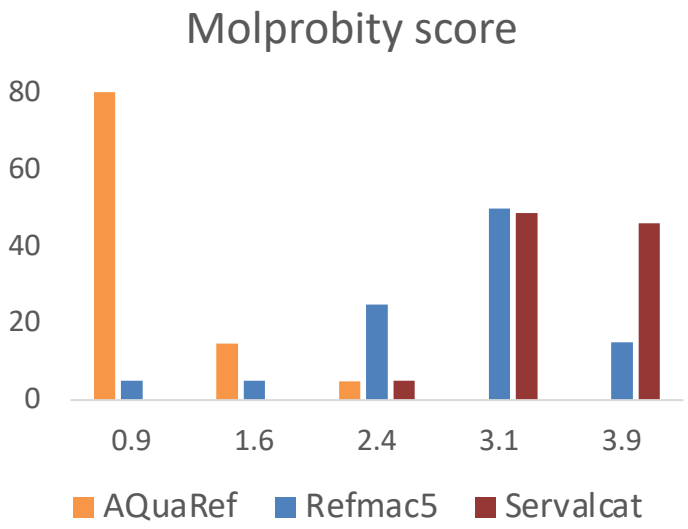
AQuaRef: low resolution

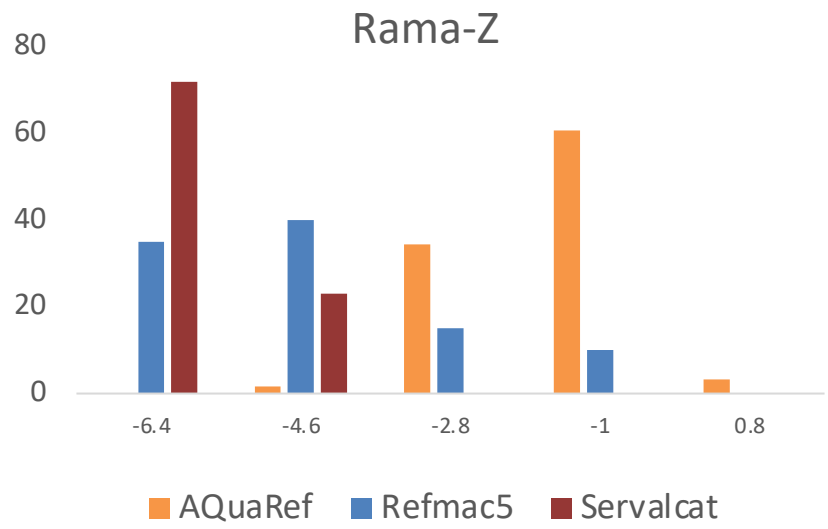
40 cryo-EM low resolution models (3Å or worse)



Testing: AQuaRef *vs* others

40 cryo-EM low resolution models (3Å or worse)





AQuaRef: model requirements

- Atom-complete models with all hydrogens
  - No bare bonds
  - One missing in whole model = Failed refinement
- Meaningful starting geometry
  - No severe geometry violations
  - Reactive potential: a pair of clashing atoms can be mistakenly considered as forming covalent bonds

AQuaRef: model requirements

- Linux + GPU + CUDA 12
  - No Mac / Windows
- No alternative conformations
- Proteins and protein-like ligands + water

AQuaRef: coming up next

- Later this year:
  - Support of Ligands
  - Handling alternative conformations
- Longer term:
  - Combine with aspherical form-factors (best for cryoEM and high-res MX)

AQuaRef

nature communications

Article

https://doi.org/10.1038/s41467-025-64313-1

AQuaRef: machine learning accelerated quantum refinement of protein structures

Received: 29 July 2024

Accepted: 12 September 2025

Published online: 17 October 2025

Roman Zubatyuk[1,7], Malgorzata Biczysko[2,7], Kavindri Ranasinghe[3,7], Nigel W. Moriarty[4], Hatice Gokcan[1], Holger Kruse[5], Billy K. Poon[4], Paul D. Adams[4,6], Mark P. Waller[5], Adrian E. Roitberg[3], Olexandr Isayev[1] ✉ & Pavel V. Afonine[4] ✉