Q|R

Quantum Refinement Project

Holger Kruse PSICON 2020

Q|R iterative refinement of biomolecular models

- **Refinement - restrained optimization of *T* w.r.t. model parameters:**

$$T = T_{\mathrm{DATA}} + wT_{\mathrm{RESTRAINTS}}$$

Atomic model to experimental data fit	External “a priori” information about the problem
Primary	**Secondary**

- **Fit model to exp. data as best as possible, without violating available “a priori” knowledge about the molecule (too much).**
- **Restraints are needed due to lack of data quality and amount**

Q|R iterative refinement of biomolecular models

- **Refinement - restrained optimization of *T* w.r.t. model parameters:**

$$T = T_{DATA} + wT_{RESTRAINTS}$$

$$T_{RESTRAINTS} = T_{BOND} + T_{ANGLE} + T_{DIHEDRAL} + T_{PLANARITY} + T_{NONBONDED} + T_{CHIRALITY}$$

$$T_{BOND} = \Sigma_{\text{all bonded pairs}} w(d_{ideal} - d_{model})^2 \text{ , } \ldots$$

standard geometry restraints (force-field) are too simple to describe reality

Q|R iterative refinement of biomolecular models

- **Refinement - restrained optimization of *T* w.r.t. model parameters:**

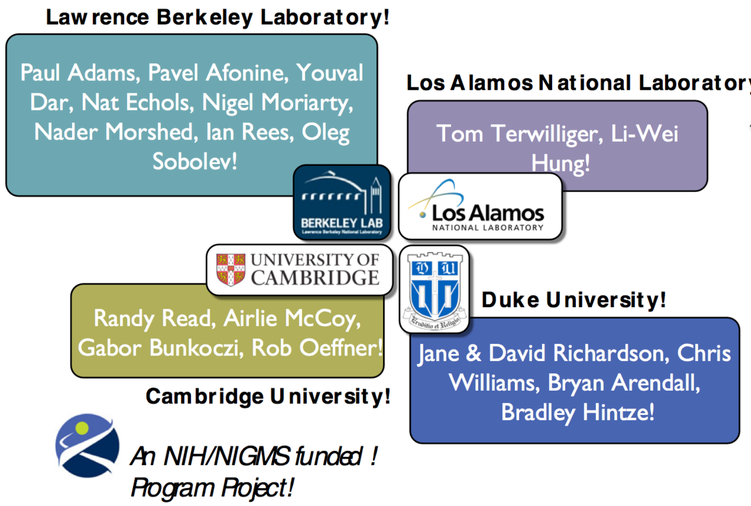
$$T = T_{\mathrm{DATA}} + wT_{\mathrm{RESTRAINTS}}$$

Information arising from QM calculations

Replace standard geometry restrains with restraints from QM (e.g. DFT-D3, GFN-xTB,.. calculations)

Q|R The Team(s)

PHENIX/LBNL



ICQSM - Shanghai University

VibESLab
vibes-lab.org

Q|R

Crystallographers + Quantum Chemists

+ Research Software Engineers

Malgorzata Biczysko (ICQMS)
Pavel Afonine (LBNL, ICQMS)
Nigel Moriarty (LBNL)
Mark Waller (ICQMS, now Pending.AI)
Holger Kruse (IBP)
+ students!

Institute of Biophysics - CAS

AKADEMIE VĚD ČESKÉ REPUBLIKY

Project Challenges and Milestones

Proof of concept and basic functionality

[#0] inception

Q|R: quantum based refinement: Acta Cryst. (2017). D73, 45-52

[#1] 'solve' QM scaling issue -> automated fragmentation approach

Solving the scalability issue in quantum-based refinement: Q|R#1, Acta Cryst. (2017). D73, 1020-1028

[#2] full crystal point-group symmetry support

Including crystallographic symmetry in quantum-based refinement: Q|*R*#2,Acta Cryst. (2020). D76, 41-50

[#3] Cryo-EM structures (real-space refinement)

Real-space quantum-based refinement for cryo-EM: Q|*R*#3,Acta Cryst. (2020). D76, 1184-1191.

[#4] treatment of alternative locations

(in preparation)

[#?] treatment of transition metal systems

[#?] treatment of nucleic acids

[#?] ...

Advanced concepts

[#?] showcase QM refinement on 'interesting' structures

[#?] explore and recommend QM methods (GPU codes or SE)

[#?] ...

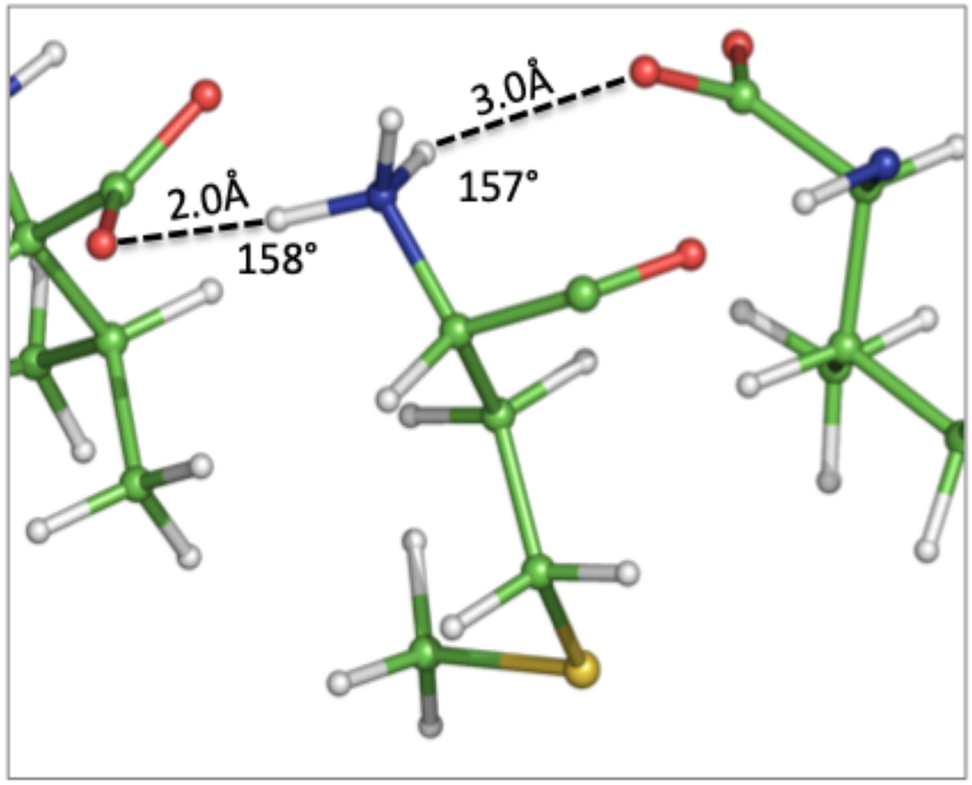
GOAL: Being able to refine complete PDB database (semi) automatically

Q|R showcase

QM refinement of whole (small) protein:

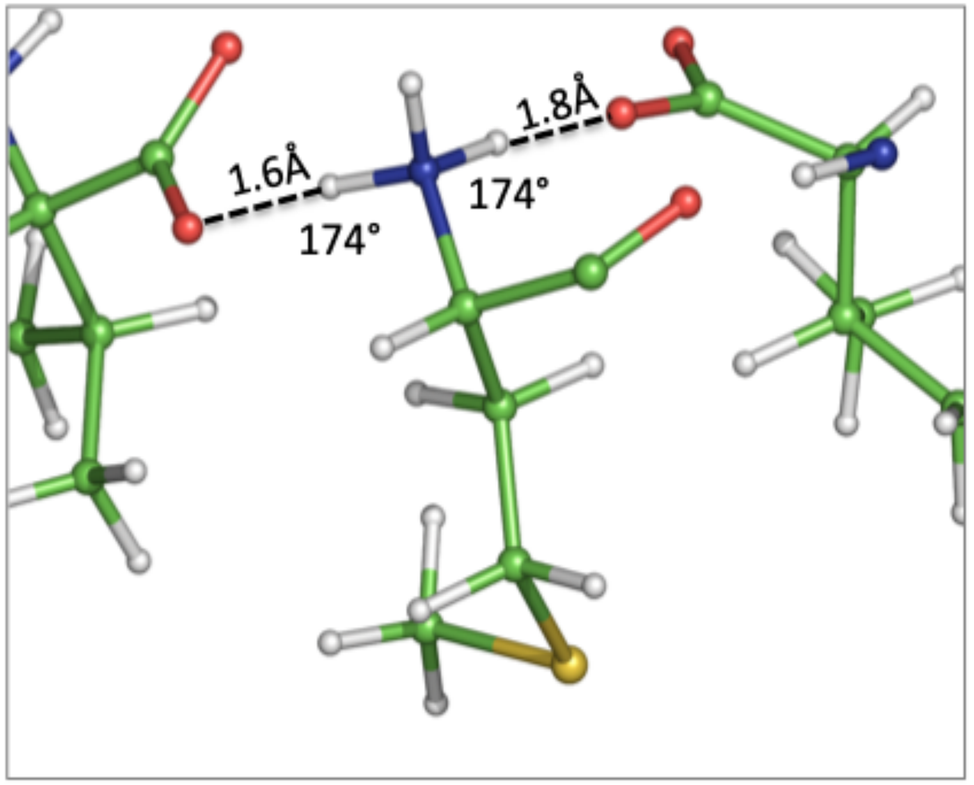
-> the amyloid cross-beta spine (PDB code: 2oNA), HF-D3/6-31G (TeraChem)

Standard refinement



Poor H-bonding

QM refinement



Improved H-bonding

M. Zheng, N.W. Moriarty, Y. Xu, J. R. Reimers, P. V. Afonine, M. P. Waller Acta Cryst. D (2017)

Q|R Standout Features

- global QM refinement
- no size limitations (fragmentation)
- no point group limitations
- (soon™) alternative locations
- for Xray and Cryo-EM structures
- open-source (qrefine.com)

- new python code from scratch
  - established libraries:
    - CCTBX-> crystallography toolbox
    - ASE -> interfaces QM programs
- modern software practices
  - extensive unit + regression tests (CI)
  - cross-team code review
- naturally integrates into PHENIX workflow

short-term Software Engineering Goals

- make switch from python2.7 to python3 (CCTBX/PHENIX dependent)
- custom ASE engines available for Gaussian, ORCA, Turbomole, Terachem, MOPAC, XTB and ANI. But various limitations.
- Switch to QCEngine as QM program backend (after py3 switch)
- Add PSI4+BrianQC program as second GPU-driven program

Q|R Overview

- Replace standard geometry restraints with QM data
- Treat entire molecule using QM
- integrate into existing refinement tools CCTBX/PHENIX

qrefine.com
github.com/qrefine/qrefine/

Thank you for listening! :-)

Scope

- improved non-covalent interactions, unusual regions, ligands,...
- Eliminate bias of man-made restraints (restraints libraries)
- analysis of fine structural details and chemistry