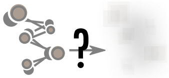
All-atom refinement to cryo-em densities

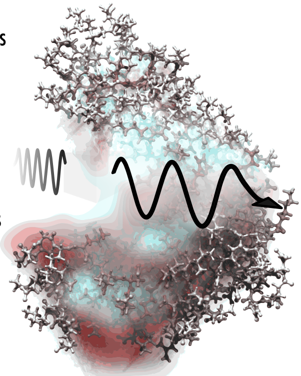
Christian Blau, Erik Lindahl, Stockholms Univeristet
christian.blau@scilifelab.se

Deriving and assessing MD potentials for refinement against experimental data

-Deriving-a-potential ——— the-Bayes-way

Find an ensemble of atom configurations {R} that best matches a given cryo-EM density.





Bayes Model: EM-densities (turqouis and red) report how incident electrons (black wave) interact with atoms (white ball-and stick) with volume elements (gray).

Goodness-of-fit measure defines Refinement potential function

- Find a potential function V that describes {R}

$V = k_B T \log p(R|\text{density})$

- Given a density, how probable is configuration R?

$p(R|\text{density}) = p(R)\ p(\text{density}|R)\ /\ p(\text{density})$

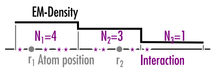
$V = V_{forcefield} + V_{fit} - \text{const.}$

- Given configuration R, how likely is a density?

$p(\text{density}|R)$

Model assumptions for goodness-of-fit

- EM-densities report a number of interactions with atoms



$p(\text{density}|R) = p(\{N_{interaction}\}|R)$

- Voxels are statistically independent

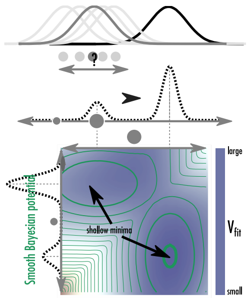
$p(\{N_{interaction}\}|R) = p(N_1|R)\ p(N_2|R)\ ...$

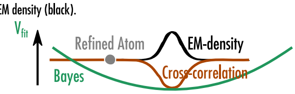
- Atoms have thermal motion and scatter independently

$p(\text{interaction in voxel}|R) = p_{cross\ section} \star p_{thermal} + ...$

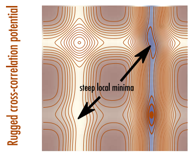
Bayes goodness-of-fit reduces local minima in 1d model system

Find atom position (gray) that best matches given cryo-EM density (black).



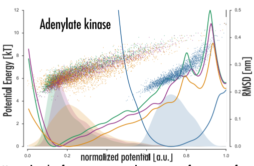


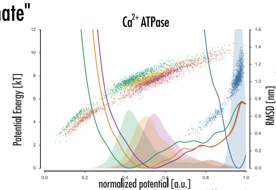
Bayesian potential avoids local minima where single atoms get "stuck" in density



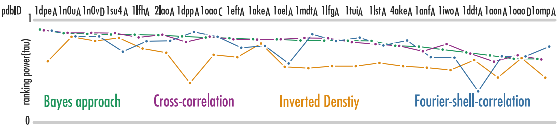
-Testing-potential-quality

Goodness-of-fit as a "reaction coordinate"

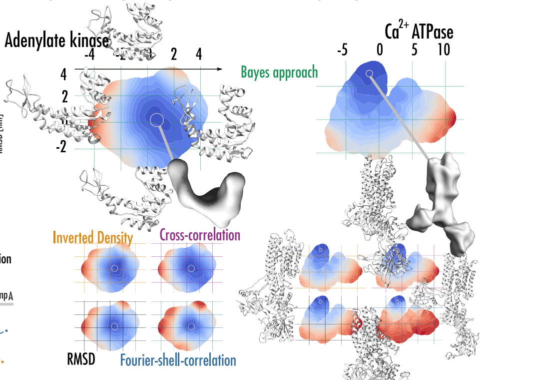




Normalized refinement potentials as scoring function in free MD simulations. To generate a large number of very different decoys, simulations were run at 400K. Pseudo-free energies estimate the difficulty to force the system to a good fit, RMSD vs. Potential point clouds report the estimated quality of the fit. Kendalls tau rank correlation coefficient (below) measures the structure prediction quality of goodness-of-fit measures



Refining = biasing energy landscapes; reweighting reveals bias



Shape of bias to protein motion depends on the refinement potenial. Bayesian refinement (top) is smoother than inverted density, cross-correlation and integrated fourier-shell correlation based potentials

-Appling-the-potential

The biasing force from a density

$\mathbf{F(R)} = -\text{grad}\ V(\mathbf{R})$

For density refinement applications:

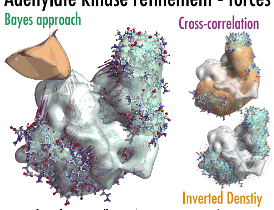
$\mathbf{F(R)} = -\text{grad}\ V(\rho^{exp}, \rho^{sim}(\mathbf{R}))$

Naive implementations expensive and error prone. Use instead

$\mathbf{F(R)} = -\sum_{voxel} \partial V(\rho^{exp}, \rho^{sim}) * \text{grad}(\rho(\mathbf{R}))$

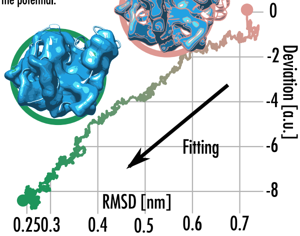
Differential density gives fat and easy force calculation for density based potentials thorugh fourier transform.

Adenylate kinase refinement - forces



Forces for refining an all atom structure against a density(grey). Simulated density in green, differental density in ochre.

Bayesian refinement against density generated from closed state crystal structure (blue). Biased MD simulation finds target structures unkown to the potential.



Acknowledgements

I thank the Knut and Alice Wallenberg Foundation and the Carl-Trygger Stiftelse for funding.

I further thank Helmut Grubmüller, for intial support with the project, Carsten Kutzner for implementation of cross-correlation potential to gromacs and Andrea Vaiana and Lars Bock for for initial feed-back.