Protein Structure Refinement with Molecular Dynamics Simulations

Michael Feig
Vahid Mirjalili
Keenan Noyes

Biochemistry & Molecular Biology; Chemistry
Michigan State University

NIH GM084953 Thank you Janna Wehrle!
NSF TeraGrid TG-MCB090003
FEIG group CASP11 Protocol

- Run 40 x 30 ns (1.2 μs / target) molecular dynamics simulation
- Explicit solvent
- Latest CHARMM force field (c36)
- Weak harmonic restraints on C-alpha atoms

~100,000 CPU core hours PER TARGET!
Simple scoring does not work!
Why Averaging?
Tolerance to scoring function noise
Why Averaging?

Replication of experimental data

Crystal structure
Avg. RMSF (~0.5 Å)

Actual dynamics (w/respect to X-ray)
Avg. RMSF (~0.5 Å)
Probably there are substates ...

... but most likely none of those look exactly like the X-ray average.
FEIG group CASP11 Protocol – Part 2

• **Average structure** from subset of conformations
• Run **additional MD with strong restraints** on C-alpha to fix stereochemistry (➔ good MolProbity scores)

➔ **MODEL 1**

We also tried to sample selectively certain regions to improve the structure further but that was not especially successful

➔ **MODELS 2-5**
CASPR11 performance

**Avg. ΔGDT-HA:** 3.4

**CASP10:** 2.6

**19 GDT-HA > 60**

**11 GDT-HA > 60**
CASP11 vs. CASP10

Why did we do better in CASP11 vs. CASP10?

• More sampling
• Weaker restraints
• Optimized subset selection
Best Sampling vs. Best Model

![Graph showing the comparison between best sampling and best model, with the sampling range and submitted model 1 highlighted.](image-url)
What is improved?

Blue: initial model
Pink: refined model
Red: exp. structure

Successful adjustment of secondary structure elements.
Flexible loops remain challenging.