Using Programming to Simulate and Visualize Proteins

Ariyana Abando and John M. Finke
Division of Biomedical Sciences , SIAS, University of Washington Tacoma
Tacoma, WA 98406

Background

Significance

The use of programming was to conduct simulations and structural visualizations that can help understand the functions of biomolecules at the molecular level. This approach can be used to study structures from any organism. This understanding can lead to advances in medicine, healthcare, and bio-engineering

Objective:

1. Make protein models where the folded state is the lowest energy (native topology model).
2. Find temperature for transition state.
3. Compare the results of transition state with experimental assessments.

Project Background

Chymotrypsin Inhibitor 2 (CI2) is a special model of protein folding because of its simplicity. It has only three important conformation states: (1) Folded states; (2) Unfolded states; and (3) transition states between the Folded and Unfolded states. The transition states are the least understood and most difficult to study.

Questions:

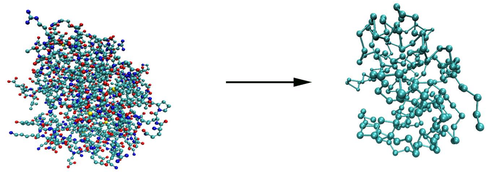
What is the structure of the transition state simulations in a native topology model?

How does this simulated structure compare to the experimental assessments?

The Approach

All-atom structure
(closer to reality)

Simple C_α model
(easier to simulate biological processes)

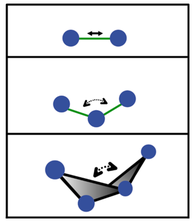


All-atom structure is stripped down to a C_α atom “skeleton model” enabling simulation of biologically-relevant processes such as protein folding .

Methods

Define energy Terms “Force Field” or “Rules of the Unviverse”

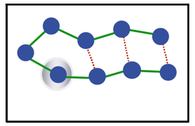
$$E_{total} = E_{bond} + E_{angle} + E_{dihedral} + E_{LJ} + E_{rep}$$



$$E_{bonds} = \sum_{bonds} \frac{1}{2} \varepsilon_r (r - r_0)^2$$

$$E_{angle} = \sum_{angles} \frac{1}{2} \varepsilon_\theta (\theta - \theta_0)^2$$

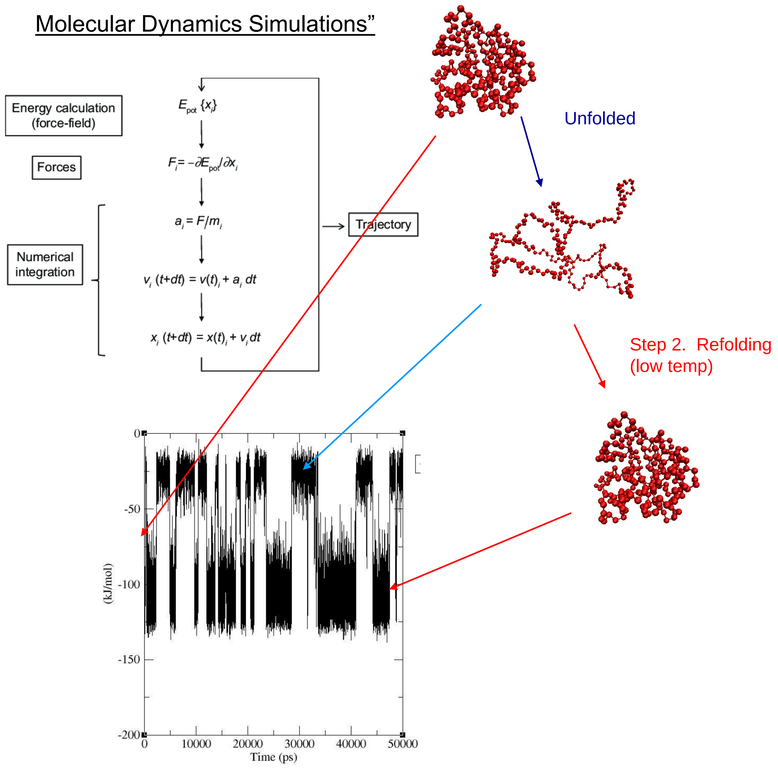
$$E_{dihedral} = \sum_{dihedrals} \left[\varepsilon_\phi^1 \left[1 - \cos(\phi - \phi_0) \right] + \varepsilon_\phi^2 \left[1 - \cos(3(\phi - \phi_0)) \right] \right]$$



$$E_{LJ} = \sum_{|i-j|>3} \varepsilon_{LJ} \left[5 \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - 6 \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{10} \right]$$

$$E_{rep} = \sum_{i,j} \varepsilon_{rep} \left(\frac{\sigma_{12}}{r_{12}} \right)^{12}$$

Molecular Dynamics Simulations”



Results

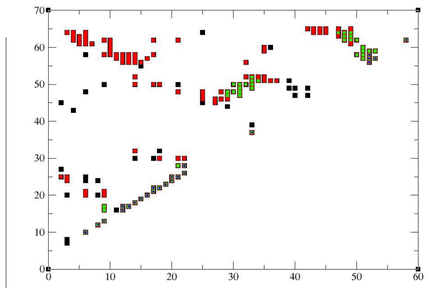


Figure 1.

Black squares represent the contacts between residues belonging to the crystal structure of CI2. Red squares represent the contacts of CI2 for the folded structure. Green squares are the contacts of CI2 in transition state. Blue represents the unfolded contacts. Important to note that squares shown are for contacts that are present more than 50% of the time. Overlapping squares are contacts that each structure has in common.

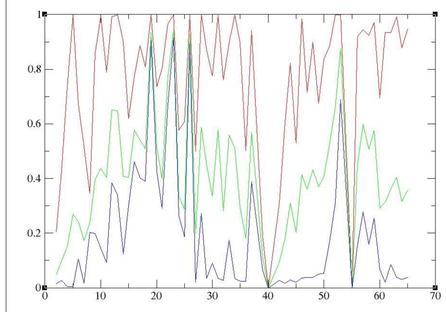


Figure 2.

Graph colors follow the same representation as in Figure 1. Points that are shown are averages of contact points within protein. It is easier to see what residues within CI2 have in common between conformation states.

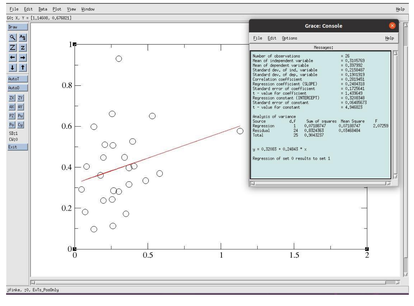


Figure 3.

Correlation of Simulated Data in Figure 2 (green lines) with experimental assessment of transition state structure (phi-value analysis).2

Some agreement (slope = 0.25; $r^2 = 0.28$). But definitely room for improvement.

Conclusion

Hypothesis is partially supported (experimental transition state structure is partially captured). There is still the need for improvement to get closer results. Simulations and visualizations improved my understanding of the function of protein folding and improved programming skills.

References:

1. Rao, M.K., Chapman, T.R. Finke J.M., (2008) Crystallographic B-factors highlight energetic frustration in aldolase folding. J. Phys. Chem. B., Vol. 112, pp. 10417-10431.
2. Itzhaki and Fersht (1995) The structure of the transition state for folding of chymotrypsin inhibitor 2 analyzed by protein engineering methods: evidence for a nucleation-condensation mechanism for protein folding. J. Mol. Biol., Vol. 254, pp. 260-288.