Modifying a Mini Protein with Two Conformational States to Instead Adopt Only One Conformation

Oliver Cho, Nicholas Wells, Colin Smith
Department of Chemistry
Wesleyan University, Middletown, CT 06459

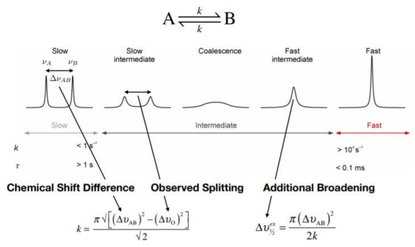
Wesleyan University

Introduction

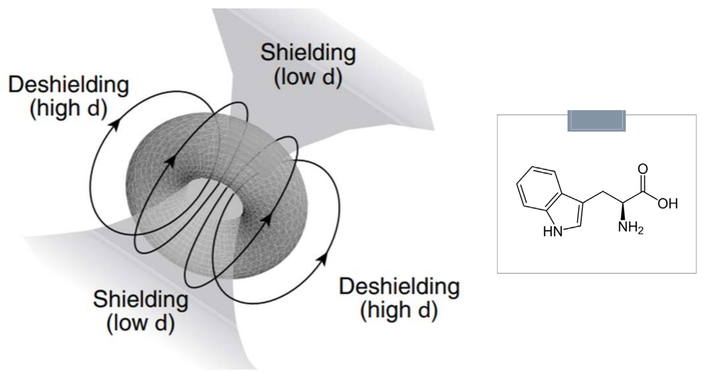
EHEE_rd2_0005 is a computationally designed mini protein composed of 40 residues. Exhibiting little structural change at 95°C or in solution of up to 4M GuHCl, EHEE_rd2_0005 is extremely thermodynamically stable.[1] While the structure of this protein was modeled easily in simulations, computational protein models are imprecise. The structure of these mini proteins must be further determined through analytical techniques after biological synthesis.

Our lab had previously attempted to analyze this mini protein using NMR spectroscopy, but our members were unable to assign several NMR peaks because of peak overlap. We suspected that this problematic peak overlap was due to the intermediate-rate dynamic exchange of our TRP-56 residue between two equally preferred conformations.

Dynamic Exchange Regimes

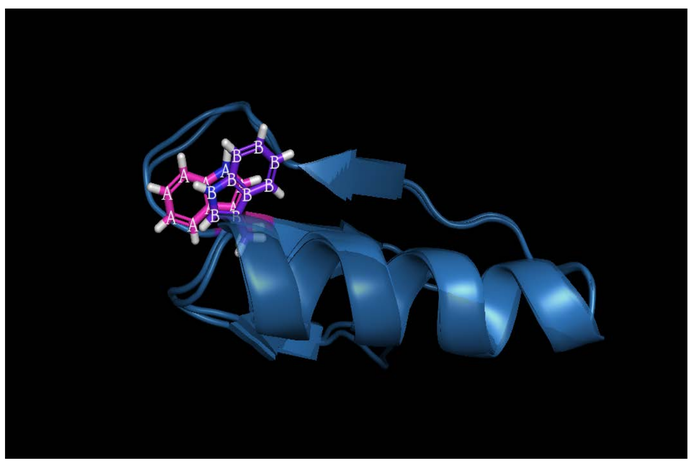


Dynamic exchange of tryptophan can be particularly problematic in NMR peak assignment because its aromatic ring current effects, combined with its large size, strongly influences the chemical shifts of neighboring nuclei.



Aromatic Ring Current Effects

In order to identify effective strategies for controlling conformational isomerism in proteins, we tested various single-point mutations that would not create drastic changes in the protein backbone structure.



The Two Conformations of TRP-56: "State A" and "State B"

Methods

A) Exhaustive yet Imprecise Single-Point Mutation Scan

Rosetta Commons

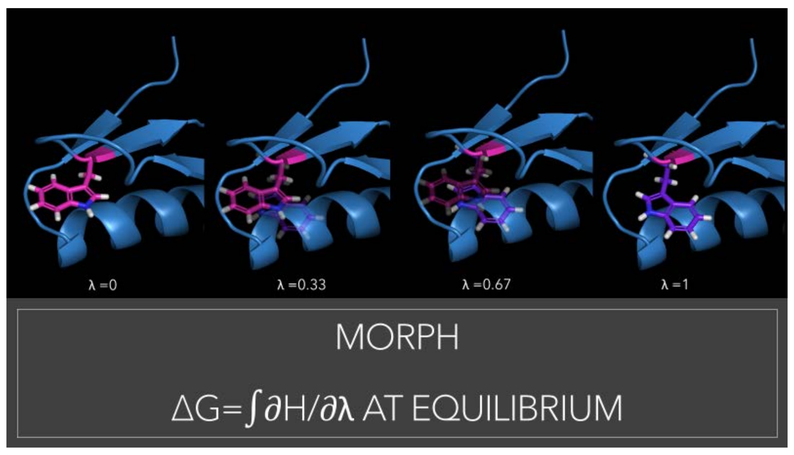
Because of the computing time required for molecular dynamics simulations, it was not feasible to run an exhaustive scan of all possible simulations using more rigorous alchemical methods. While we first chose mutations based only on the hypothetical attractive and repulsive forces in which the new side chains may participate, we later began to use Rosetta to help us choose mutations.

Given the possibility that we could have missed a mutation we had never considered, we ran an exhaustive scan of all possible single-point mutations using Rosetta, which uses a much less resource-intensive energy scoring system. We also did not allow for backbone changes in these Rosetta simulations, which sacrificed more accuracy for speed.

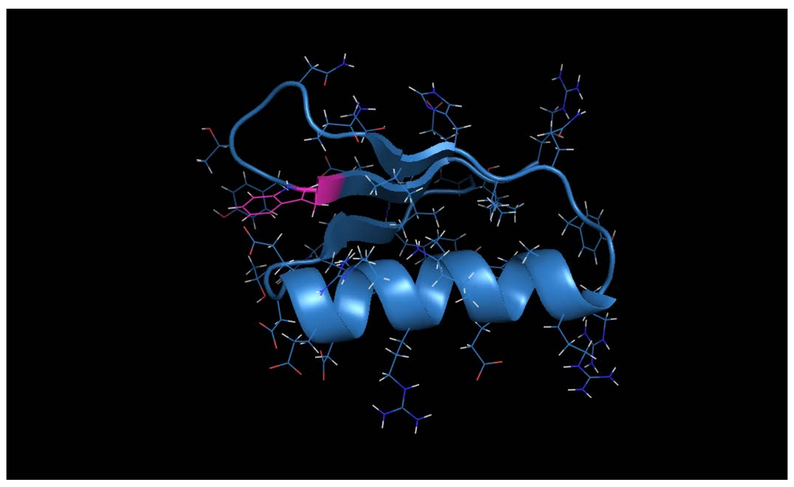
B) Slower but More Accurate Molecular Dynamics Simulations

GROMACS FAST. FLEXIBLE. FREE.

Using Gromacs for our molecular dynamics simulations, we measured the effectiveness of each of our mutations by measuring the free energy differences of the two conformations with alchemical morph simulations.



We used fast-growth simulations that morphed our residues from one state to the other over the course of 50 ps, which is much too fast to be at equilibrium. Instead of integrating over the change in energy, we used pmx to analyze our data. Depending on the data, pmx chooses one of three methods: the Bennett Acceptance Ratio, the Crooks Gaussian Intersection, or the Jarzynski Equality.

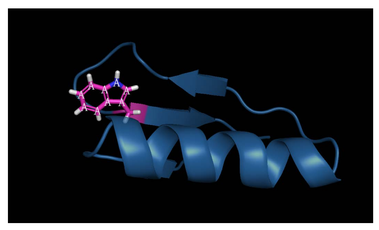


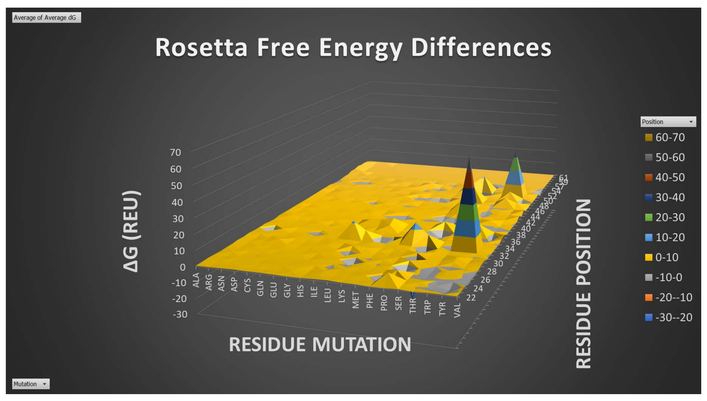
References

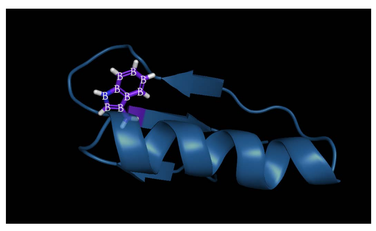
1. Rocklin, G. J., Chidyausiku, T. M., Goreshnik, I., Ford, A., Houliston, S., Lemak, A., . . . Baker, D. (2017). Global analysis of protein folding using massively parallel design, synthesis, and testing. *Science, 357*(6347), 168-175. doi:10.1126/science.aan0693

Results

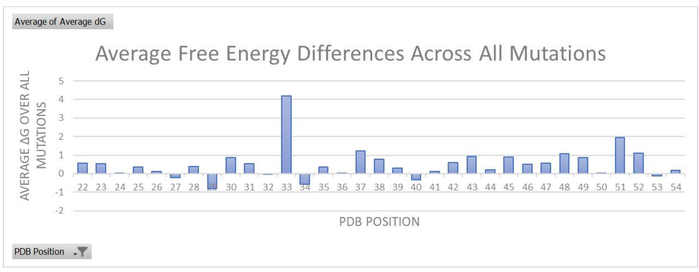
A) Rosetta





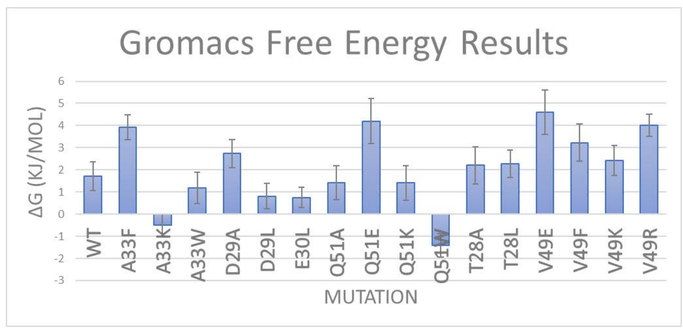


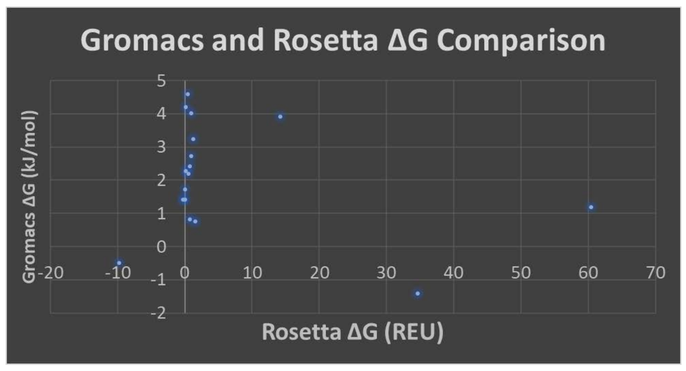
The analysis of our Rosetta data suggested that it is easier to mutate our protein to push the tryptophan ring into the "state A" conformation. This strategy's effectiveness is likely due to the strong secondary structure's ability to hold sterically large sidechains in place, which can push the TRP-56 sidechain outward.



Rosetta also helped us determine which residues would be better potential targets for mutation.

B) Gromacs





Despite its success in guiding our general strategies, Rosetta was a poor predictor of specific mutations' free energy values.

Acknowledgements

- Josh Dudley for struggling with me through the countless mysterious errors
- Professor Smith for the position in his lab and his guidance throughout the way