# Impact of Insertion and Deletion Mutations on Protein Thermodynamics

**Ryan Wilkins and Stefan Wallin**
**Department of Physics and Physical Oceanography**
**Memorial University, St. John's, NL**

## Introduction

Insertion and deletion mutations are common events in protein evolution but occur much less frequently than point mutations. Due to this, the effects of inserting and deleting amino acids on the biophysical properties of proteins are much less studied and understood. Here, we use a coarse-grained protein model to assess the effects of these mutations on protein stability.

## Coarse-Grained Model

This study used a coarse-grained model [1,4] for protein folding with 7 atoms per amino acid and a reduced alphabet of 3 amino acids: polar (p), hydrophobic (h), and glycine (g). The energy function is given by:

$$E(C) = E_{hb} + E_{hp} + E_{local} + E_{exvol}$$

The stable native state for this model occurs at the minimum energy state. Model degrees of freedom are dihedral angles $\varphi$ and $\psi$, while other angles and bond lengths are held fixed.

## Computational Methods

To study the thermodynamic behaviour, simulated-tempering Monte Carlo (STMC) methods were used. Simulation parameters were as follows:

- $10^9$ steps per simulation
- 8 model temperatures from $k_BT = 0.47$ to $k_BT = 0.70$
- MC updates change either $T$ or chain geometry by rotating $\varphi$ and $\psi$

## Stability of the Native State

A thermodynamic observable for stability $S$ was defined. $S$ can be considered the population of the native state and takes a value between 0 and 1. $\Delta S$ can then be considered a change in stability between two sequences.

## Sequence $II_{\alpha\beta}$

Using the coarse-grained model, a 35 amino acid sequence labeled $II_{\alpha\beta}$ was designed to fold spontaneously into a protein exhibiting both an α-helix and β-hairpin structural elements. The initial sequence used was:

pphpphhpphpphhpggggphphphpggphphphp

This sequence was obtained using basic protein design principles:

- Glycine in turn regions
- p and h patterns such that the helix
- and sheet regions are amphipathic

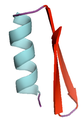

Figure 1: The fold of Sequence $II_{\alpha\beta}$.

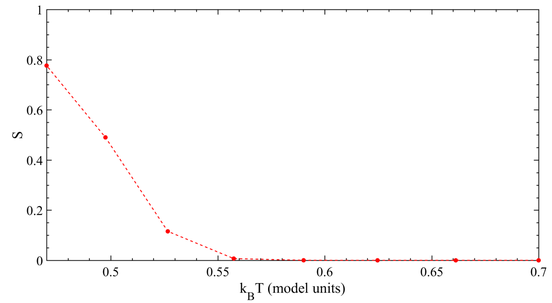


Figure 2: A melting curve for the initial sequence $II_{\alpha\beta}$.

As seen in the melting curve for the protein, the stability $S$ of the native state increases as $T$ decreases. For this reason, all $\Delta S$ measures between sequence $II_{\alpha\beta}$ and its mutants were made at $k_BT = 0.47$.

## Criterion for the Native State

To be considered for a native fold, each protein must exhibit:

- 75% of residues exhibit appropriate native state configuration
- 4 or more inter-strand hydrogen bonds in the β-hairpin
- 5 or more hydrophobic contacts between α-helix and β-hairpin

## How do Indels Impact Stability?

From the initial sequence $II_{\alpha\beta}$, 73 unique insertions and 25 unique deletions were found. Insertions were obtained by inserting each of the amino acids (p, h, and g) into each position in the sequence. Deletions were obtained by deleting each residue. Each sequence was simulated using STMC methods to find the stability of each sequence.

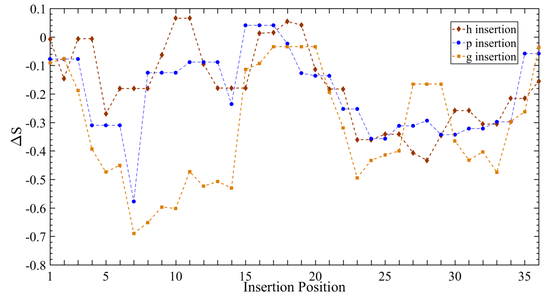


Figure 3: Stability changes resulting from insertion mutations to sequence $II_{\alpha\beta}$.

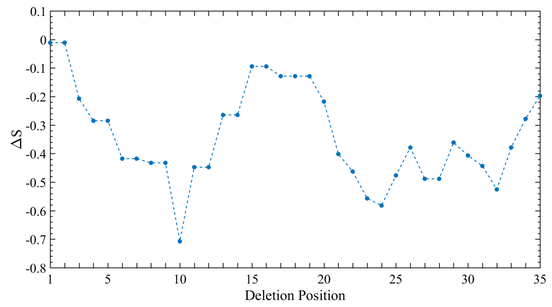


Figure 4: Stability changes resulting from deletion mutations to sequence $II_{\alpha\beta}$

## Conclusions

From these results, the following observations can be made:

- Insertions are generally more tolerated than deletions
- Inserting glycine is often more deleterious than p or h due to its helix-breaking tendencies
- Indels are better tolerated in loop and terminal regions than in a helix or sheet
- Some hydrophobic insertions can help stabilize protein structure
- Indels (other than glycine insertions) in the α-helix are less destructive than in the β-sheet

These results appear to agree with experimental observations on real proteins [2,3].

An important sequence pattern of "pphhpp" was identified to be important in helix formation. Indels which disrupted this pattern were less stable. Hydrophobic insertions extending this pattern were found to promote stability.

## Future Directions

These results can serve as a baseline for research in the following directions:

- Isolation of the two secondary structure elements to assess the impact of indels in isolation in comparison to a protein with tertiary structure
- Examining specific mutations with either stabilizing or extreme destabilizing effects to find out why this happens
- Exploring if and how indels can aid proteins in switching between stable folds in protein evolution
- Exploring similar proteins using more detailed all-atom protein models

## References

1) C. Holzgräfe and S. Wallin. Local versus global fold switching in protein evolution: insight from a three-letter continuous model. Physical Biology, 12(23), February 2015.

2) N. V. Grishin. Fold change in evolution of protein structures. J. Struct Biol., 134(2-3), 2001.

3) David Shortle and John Sondek. The emerging role of insertions and deletions in protein engineering. Current Opinion in Biotechnology, 6, 1995.

4) A. Bhattacherjee and S. Wallin. Coupled folding-binding in a hydrophobic/polar protein model: Impact of synergistic folding and disordered flanks. Biophysical Journal, 12(8):569-578, February 2012.

## Acknowledgments

The authors would like to thank Memorial University of Newfoundland, SciNet and Compute Canada for making this project possible.

## Further information

For further information, please visit:

http://www.physics.mun.ca/~stefan/