

Does a Simple Lattice Protein Folding Model Exhibit Self-Organized Criticality?

Arun Bajracharya, Tyler Schiewe, Yura Sim; Faculty Advisor: Joelle Murray
Linfield College, McMinnville, OR

Special thanks to the Linfield College Physics Department and the Linfield College Undergraduate Research Fund

Introduction

Proteins are known to fold into tertiary structures that determine their functionality in living organisms. However, the way they consistently fold to the same structures is unknown. Our research sees if the folding process can be viewed computationally through the lens of self-organized criticality using a simple lattice-bound protein.

Self-Organized Criticality (SOC)

An SOC system is a dynamical system at a critical state that is characterized by power law behavior. It follows that these systems are both time and scale invariant.

The Sand Pile Model is the first and most well known example of a self-organizing system.

1. Grains of sand are individually dropped onto a flat plane (iterations).
2. The sand builds up into small sand piles which build up into a large sand pile. If the slope of the sand pile becomes too steep, avalanches will occur.
3. Eventually a critical slope is reached in which all sizes of avalanches can occur. The avalanche distribution follows a power law.
4. Once the critical slope is reached, the avalanche distribution will keep the system at the critical slope, thereby making it self-organizing.
5. Larger sand piles will have larger avalanche sizes but their distributions will have a similar slope (scale invariance) and the distribution will not be affected by the amount of iterations (time invariance).

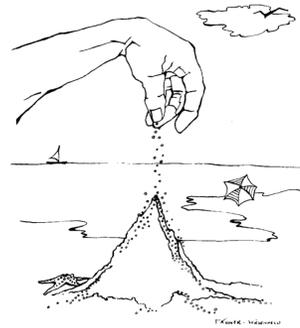


Figure (1) Illustration by Elaine Wiesenfeld depicting the Sand Pile Model.

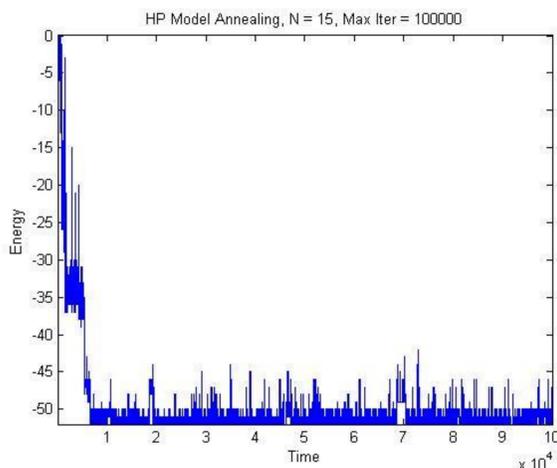


Figure (2) Typical energy plot of a folding lattice protein

For our lattice proteins, avalanches are associated with successive folds. An avalanche is said to start when an allowed fold occurs after a stopped fold. The avalanche size is equal to the number of successive allowed folds. Note some folds are allowed to increase the energy of the protein according to a probability function known as the Boltzmann Factor.

Model & Algorithm

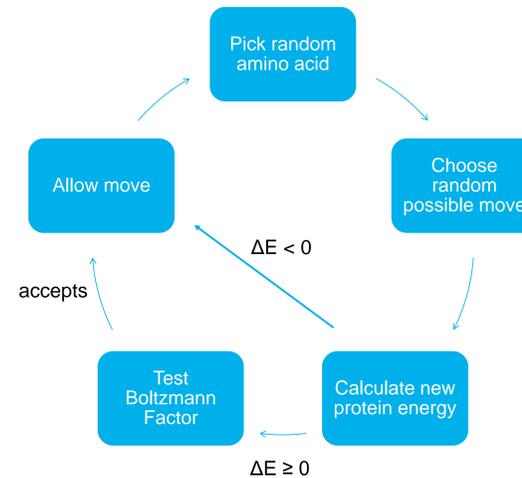


Figure (3) Algorithm Flow Diagram

- Amino acids are bound to a 3D grid (see Fig. 6). Bonds cannot stretch more than 1 unit. New bonds cannot be formed or destroyed.
- There are two types of amino acids in our protein: hydrophobic and hydrophilic. In a real folded protein, hydrophobic acids gather together in the center, away from the outside water molecules. We used an energy array to simulate this behavior to give our proteins their folding dynamics.
- A fixed $N \times N$ array provides the energy between any two amino acids. The energy value is subtracted from the total energy value if the two acids are unit distance from each other.

Discussion of Results

The avalanche distributions for figures 4 and 5 were made from an average of 300 folding proteins' distributions. Our proteins' avalanche distributions were not as linear as we anticipated. This would mean that it is not obvious that there is a power law. However, the distribution looks almost completely linear above zero on the y axis. Apparently, the large avalanches (represented in the "tail" of the distribution) were only being sampled at a very small frequency. This leads us to believe that with more statistics--possibly much greater than 300 runs--we would see a definite power law.

The fact that the distribution is mostly linear for the entirety of the folding process suggests to us that our proteins are always in a critical state. This characteristic differs from other SOC models like the Sand Pile Model. It is difficult to know what this means; it could mean that our system is not self-organizing, or it could mean that the nature and dynamics of protein folding are closely linked to SOC behavior.

Moving forward, we realize the need to improve computational efficiency. This would allow us to perform more runs and thus provide better statistics from which we could find a power law. Already our efforts have more than doubled computational time while adding more measurement features.

Since our approach to understanding protein folding dynamics is fairly new, our definition of lattice protein avalanches and rules for lattice protein dynamics are still largely experimental and subject to change.

Data

N = 15

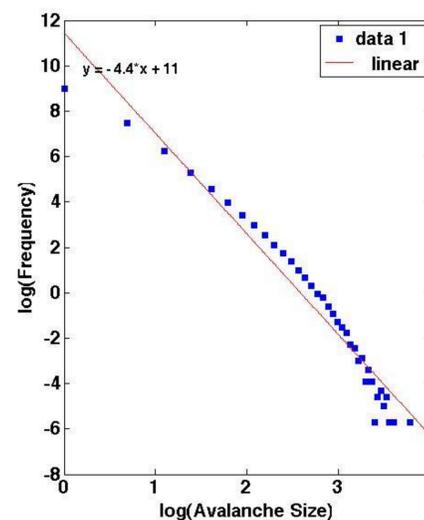


Figure (4) Avalanche Distribution [N=15]

N = 75

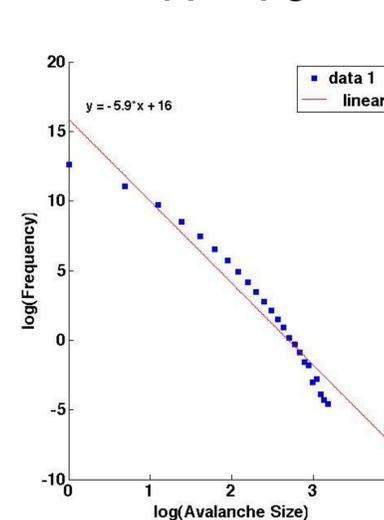


Figure (5) Avalanche Distribution [N=75]

Since SOC systems are supposed to follow a power law, an avalanche distribution on a log plot should show a straight line. Also, because SOC systems are scale invariant, the slope for the linear fit for the 15 acid protein should be similar to that of the 75 acid protein.

Future Research

- Artificial Avalanche Stopping
 - See if setting an avalanche size limit inhibits protein folding
- Multiple Acid Folding
 - Fold two or more acids at the same time
- Non-lattice Proteins
 - Allow acids to move freely through 3D space and allow bonds to stretch or possibly break
- Contact Maps
 - Use contact maps to measure different things about the folding process including fold patterns and folded state consistency.

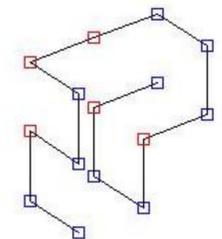


Figure (6) Folded N=15 protein viewed in perspective. Red proteins are hydrophobic, blue are hydrophilic

References:

- "Computational Physics", Nicholas J. Giordano, Hisao Nakanishi, Pearson/Prentice Hall (2006)
- "How Nature Works: The Science of Self-Organized Criticality", P. Bak, Copernicus New York, NY (1996)
- "Theoretical Models of SOC Systems", Markus J. Aschwanden, Self-Organized Criticality Systems (2013)