# Homework #1-Physics 498Bio

by Soon Yong, Chang – Oct/15/01

*Review of paper "A thermodynamic model for prebiotic RNA-protein co-evolution" by Erzan Tuzel and Ayse Erzan, cond-mat/0107315 v1 Jul/2001*

## INTRODUCTION & BACKGROUND

The proteins consist of building blocks known as amino acids. These amino acids form polymer chains with well-defined sequences given by RNA. The basic sequence of the polymer chain is known as the primary structure. However, the functional characteristics of the proteins are also defined by structures obtained from folding the primary structure on itself. These are the secondary and tertiary structures. The folding process has to proceed in an accurate and fast manner in order for the protein structure to be attributed with useful functions.

In this study, the situation that precedes the existence of the cells (prebiotic) is supposed. At the same time, the RNA is allowed to have an independent existence from the DNA acting as the synthesizer of the protein. There is supposed to be a primeval "soup" rich in different amino acid molecules where the RNAs as well as the protein molecules float freely. The evolutionary advantages that favor those proteins (and RNAs that encode them) that can act as "heat pumps" folding and unfolding is studied. Some kind of heat dissipation mechanism is necessary in order to permit the existence of the low entropy bio-molecular structures (RNA, proteins). A favorable scenario for the cooperative evolution of RNA-protein is the one where this prebiotic "soup" is contained within the inter-connected cavities of the porous rock. The surface to volume ratio is basically ~ 1/r, so that smaller pores maximize the probabilities of the contact with the surface for dissipation. The heat is absorbed by the protein molecules during the unfolding process and it is dissipated through the rock surface in the fast folding. This paper explores the feasibility of such a refrigeration scheme by proposing a simple model Hamiltonian and studying the efficiency of the folding-unfolding cycle of the proteins.

Although the proteins that exist today escape this simple argument, it is still an interesting quest that of trying to formulate the physical reasons that made possible the selection of particular proteins structures along with RNAs that can synthesize them. This is, in my view, a question worthy enough for undertaking in this homework assignment.

## METODS

RNA is injecting information during the formation of the protein chains from the basic amino acid components. The entropy is being lowered. A simple way of lowering entropy is to lose heat. Thus when the protein molecule is detached from RNA, it is in the "unfolded state" carrying away heat. The existence of hydrophobic parts (residues) in the protein facilitates its diffusion toward the rock surface. After the exchange of heat is made, the protein molecule folds on itself in such a way to avoid the contact of the hydrophobic residues with water. This is important because now the low entropy protein is free to wander around gathering heat from the

"soup" as well as from other RNA molecules. This is the cycle that is repeated over and over, with an increasing number of participating proteins.

The physics of the protein synthesis and folding (and unfolding) is captured by the following model:

+There is an energy gap ($J$) between the correct sequence of the amino acids and the incorrect one. There is only one correct order of the monomers that lowers the energy. Thus other combinations (errors) are statistically disfavored.

+Even if the protein as a chain of monomers has a very large number of degrees of freedom with correspondingly large number of possible configurations (folded states), there is a preferential "pathway" of successive folding steps that connect the folded state with the denatured state. In this way, the protein molecule does not have to try and probe many available configurations "wasting time".

+A comparison with an inverted tree structure can be made. From the unfolded state (root), the protein structure follows a particular pathway through different branched until it gets to the folded state (destination node).

+The pathway is probabilistically favored by being associated with the lowering of energy in the amount ($\varepsilon$).

+The existence of the folding pathway can be turned "on" and "off" by the "guiding parameter" ($\lambda$).$\lambda$ =0 is equivalent to having the protein in one of the only two states; "folded" and "unfolded". $\lambda$ = 1 allows the protein to be in any intermediate states along the pathway, thus having more possible configurations. This consideration can modify the partition function with slightly different outcomes.

Once the energetic model is defined, the usual partition function can be calculates (at the constant temperature ($T$) as the rock is the heat reservoir for the system). Thermodynamic entropy ($S$) can easily be calculated from the Helmholtz free energy.

Also the quantity called "order parameter" ($\psi$) that measures the averaged degree of folding, can be obtained. $\psi = 0$ means that all the proteins of the ensamble are unfolded and $\psi = 1$ means that all the proteins of the ensamble are folded.

This model is quite simple with only a small number of parameters ($J$, $\varepsilon$, $\lambda$, $T$, $N$ and $q$). $q$ is the number of possible branches for each node of the "inverted tree", of which only one leads through the "pathway". The input of these quantities is quite arbitrary with some experimental justifications. This could be one of the weaknesses of the model.

**RESULTS & DISCUSSION**

First of all summarizing briefly the result:

+An abrupt change of the order parameter ($\psi$) is observed as the temperature is raised (at *Tc*). The value goes down from 1 (fully folded) to 0 (fully unfolded) at the temperature *Tc*. Being guided ($\lambda$ = 1) or unguided ($\lambda$ = 0) affects how fast the protein can be unfolded ($\lambda$ = 0 the fastest).

+The existence of a large entropic gap is detected between the two states of protein (folded/unfolded). The entropy is abruptly lowered below *Tc*.

+A "heat engine" working between two isothermals (soup and rock temperatures) and two isoentropic states ($\lambda$ = 1 and $\lambda$ = 0) can be established.

In conclusion, this simple model proved the feasibility of the heat dissipation mechanism described in the introduction. Supportive evidence of this mechanism can be found in the nature; the existence of very fast folding proteins under conditions of extremely hot temperature. We can reason that the RNA molecules that synthesize fast folding molecules might have replicated more and became winners in the evolutionary struggle.

As the authors have pointed out, there is an important pitfall. If we assume that the heat transfer through the rock surface is not complete and the protein molecules are to be found in partially folded states, the exposed hydrophobic core would favor the aggregated state of multiple polymeric structure. In this case, the dissipation cycle is disrupted. However, this difficulty is circumvented by assuming that there are "molecular helpers" present in the soup that avoid such aggregation.

I think that the objective of this "toy" model has been achieved, with some obvious shortcomings because of its extreme simplicity. For example, it does not consider the right-handed chirality of the amino acid molecules. This could be explained by the arguments of selection that would bias favorably the energetics. The parameters such as *J* and $\varepsilon$ ignore completely the specificity. But, in the end there is no guarantee that the life began in the porous rock and we cannot discard other possible scenarios where some kind of adversity and competing tendencies would make a winner out of many possible candidates.