11th International Conference
on Advanced Materials
Rio de Janeiro Brazil
September 20 - 25
ICAM2009

Simulation of Protein Folding in the square lattice: number of monomer-monomer contacts.

P. M. Silva[1], A. R. de Souza[1], J.R. Sambrano[1], A. Caliri[2]
[1] Grupo de Modelagem e Simulação Molecular–DQ/DM/FC/UNESP, 17033-360, Bauru - SP, Brazil
[2] Departamento de Física e Química – FCFRP – USP, 14040-903, Ribeirão Preto – SP, Brazil

Abstract – We study the folding of an amino acid sequence using the two dimensional Hydrophobic-Hydrophilic model (2D HP model) for protein structure formation. The hydrophobic-hydrophilic model proposed by Dill [2], is a free energy model of the native conformation of a protein. The folding is due to interactions between hydrophobic amino acids that tend to form a core in the spatial structure shielded from the surrounding water molecules by hydrophilic amino acids. The model is usually referred to as the HP model where H stands for hydrophobic and P stands for polar. The energy of configuration is calculated by looking at neighboring water sites around the hydrophobic amino acids.

In these work we present a computer simulation designed to predict the bi-dimensional structure of proteins given their amino acid sequence. How proteins fold is an important topic in the field of molecular biology, were we used a simple model that captures the essence of the important components of protein folding [1]. The linear sequence of the protein is composed of amino acids of only two types [2]: hydrophobic and Hydrophilic. This sequence is folded on a two-dimensional square lattice were each point of the chain can turn 90 degrees left, right, back or continue ahead.

Each protein is represented by a structure containing two arrays sorted according to X and Y-coordinates. We also store the energy of each protein structure configuration. The population consists of 2.155.667 such configurations. The energy function is simple: -1 for each direct contact occupying neighboring non-diagonal lattice points. Despite the simplicity of the HP model, the folding process has many behavioral similarities with the folding process in the real system, and used to formulate new hypothesis of protein structure formation. The success of the HP model as a computational tool came from the fact that the discrete set of possible conformations makes it possible to enumerate and to consider all conformations of small proteins. In table 1 we present the topological monomers neighbour in the square lattice. We can see that for a chain with 4 monomers there are 5 possible conformations, and only 1 with topological monomer neighbour, represented by the two monomers: 1 and 4. In the case of a chain with 6 monomers we can obtain 36 possible conformations and 18 that present topological neighbour. The energy of the conformation can be predicted from the possible choice of H or P monomers along the chain. For example the chain with 4 monomers will present a conformation most stable when the H monomers will be present at positions 1 and 4 along the chain.

Chain	N.T.C	N.C.T.N	Topological Neighbours
4	5	1	[1,4]
5	13	4	$[2,5]^2$,$[1,4]^2$
6	36	18	$[3,6]^5$, $[1,6]^3$, $[2,5]^3$, $[1,4]^2$,[3,4]
7	98	59	$[4,7]^{13}$, $[2,7]^4$,$[3.6]^{10}$, $[1,6]^2$, $[2,5]^9$, $[1,4]^{15}$
8	272	222	$[5,8]^{44}$,$[3,8]^{14}$,$[4,7]^{29}$,$[1,8]^{13}$,$[2,7]^5$,$[3,6]^{29}$
9	740	802	$[1,4]^{123}$,$[1,6]^{54}$,$[1,8]^{30}$,$[2,5]^{82}$,$[2,7]^{20}$,$[2,9]^{21}$,$[3,6]^{80}$,$[3,8]^{20}$,$[4,7]^{80}$,$[4,9]^{48}$,$[5,8]^{87}$,$[6,9]^{145}$
			N.C.T.N = Number of Configurations with Topological Neighbours
			N.T.C = Number of Total Configurations
			Chain = Number of Momomers in the Chain

Figure 1: Topological contacts between monomers in the chain.

References

[1] Rios, P. L. Caldarelli , G., Puttin proteins back into water, **Physical Review E**, 62, 8449 (1999).
[2] Lau, F.K. and Dill, A.K., A Lattice Statistical Mechanics Model of the Conformational and Sequence Spaces of Proteins. . **Macromolecules**,22,3986-3997. (1989)