

## Growth of homopolymer chain

A.Benyoussef<sup>1</sup>, Y.El Amraoui<sup>2</sup> and A.El kenz<sup>1</sup>

<sup>1</sup>) *Laboratoire de Magnétisme et de Physique des Hautes Energies,  
Département de Physique, Faculté des Sciences, B.P. 1014, Rabat, Morocco*

<sup>2</sup>) *Faculté des Sciences et Techniques, Errachidia, Morocco*

Thermodynamic and kinetic properties of an homopolymer chain, in two dimensional square lattice, are studied by the use of chain growth algorithm based on the self avoiding walks process. Which is in accord with what happen in nature. All monomers are considered to be hydrophobic. In the thermodynamical study we examine the critical behaviour of the chain by following the evolution of the chain geometry with temperature. This is done by studying the behaviour of the mean distance  $R$  chosen to describe the homopolymer structure. Moreover, we give the behaviour of thermodynamic quantities such as conformational energy and specific heat. The effect of a force applied, in a fixed direction, on the chain when it is growing is also investigated. Finally, in a kinetic study, we analyse the evolution with temperature of the mean time necessarily for a given conformation to reach its native state.

### I. INTRODUCTION

As it's shown in previous studies, a protein folds to its native conformation when the number of topological contacts between hydrophobic monomers in the chain increases in order to minimise the protein energy [1]. Hydrophobic energy is considered to be the dominant force which really leads to the folded state of the protein [2,3]. This energy is a function of temperature and the number of contacts in the chain decreases when temperature increases [4]. As a result of this hydrophobic driven force the energy landscape of the protein is represented by great number of minima each one corresponds to a compact structure and the smallest one of them corresponds to the native structure [5]. Because of the domination of this hydrophobic energy it's reasonable to approximate the folding of the protein by simple models in which hydrophobic interaction is the only driven force that leads to folding, whereas the other interactions are neglected [6,7]. indeed, In the HP model (H is hydrophobic monomer and P is polar monomer), one consider each lattice nearest neighbours HH contacts interaction has energy  $\epsilon < 0$  whereas all other contacts types (HP, PP) are neutral have zero energy [6,7]. In the HP<sup>+</sup> model, the energy is defined by  $\epsilon < 0$  for each native HP or PP contacts and  $-\epsilon$  for all non native contacts  $-\epsilon > 0$  [6,7]. Until now, the most of studies of protein folding by simulation treat the problem by Metropolis folding algorithms using the whole chain. In these algorithms a new configuration of the chain is sampled by re-growing only part of the chain, and so studying pathways that lead to the native conformation of the protein [8,9]. Unlike Metropolis algorithms using the whole chain, the chain growth algorithm generates configurations segment-by-segment, as what happen really in nature where the protein is formed by adding monomers to the chain until the full length of the protein is reached. This is done when

RNA is decoded to determine the protein sequence. A transition probability is used for growing the links, which is guided by Boltzmann statistics [10]. The configurations generated contribute significantly to the thermodynamic average at a given temperature. One of the most important advantages of the chain growth algorithm is that conformations generated are statistically independent, since the chain growth configurations are not dynamically connected they lead to enhanced exploration of configuration space.

Our aim is to study thermodynamic and kinetic properties of a homopolymer chain in two dimensional square lattice by a Monte Carlo simulation using a chain growth algorithm. For this, we use a simplification of the HP model in which the monomers of the chain are chosen to be hydrophobic. So, there is no polar monomers in the chain. Our study is divided on two parts. The first one is concerning thermodynamical study, in which we localise the collapse transition by following the evolution of a parameter  $R$  describing the chain geometry, and which behaves as the gyration radius. Then we study thermodynamic quantities as the conformational energy of the chain as a function of temperature and as a function of  $R$ , and the specific heat as a function of temperature. In order to complete this study we examine the effect, on the homopolymer collapse, of a force applied in a fixed direction on the chain when it is growing. In the second and last part of this paper we study the kinetic of the chain by the growth algorithm and so we analyse the evolution of the mean time necessarily for a given conformation to reach its native state.

### II. MODEL AND METHOD

In our chain growth model, proteins are generated segment-by-segment on a two dimensional square lattice. Based on the self avoiding walks process (SAW), in order to prevent

the overlap of excluded volume sites on this lattice, the possible growth directions are given by the direction vector  $V = xi + yj$  where  $(x,y) \in \{(0,1),(0,-1),(1,0),(-1,0)\}$  and considering the distance between two adjacent points of the lattice to be equal to 1. The square lattice has three allowed moves from a given site by excluding the fourth one prevented by the SAW. All conformations generated by this method are of minimal energy. The model is described by the Hamiltonian defined as:

$$H = \sum_{ij} \epsilon \Delta_{ij}$$

where  $\epsilon = -2.3$  and  $\Delta_{ij}$  is given by

$$\Delta_{ij} = \begin{cases} 0 & \text{if } i \text{ and } j \text{ are not nearest neighbours} \\ 0 & \text{if } i \text{ and } j \text{ are nearest neighbours in the chain} \\ 0 & \text{if } i \text{ and } j \text{ are topological nearest neighbours} \\ -2.3 & \text{if } i \text{ and } j \text{ are topological nearest neighbours} \end{cases}$$

is then expressed as:

In order to introduce the temperature dependence of the collapse, a transition probability is used for regrowing links, which also allows the computation of different ensemble averaged quantities. Every direction chosen at a step of the chain growth is determined by the probability

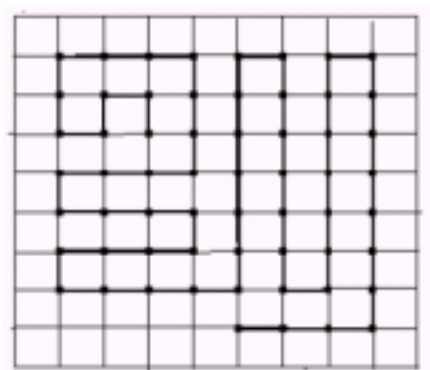
$$P_i \propto e^{-\beta E_{ij}}$$

Where  $\beta = 1/KT$ .

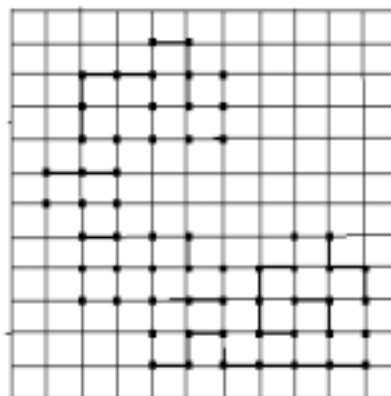
With this simulation we generate conformations where their geometries depend on temperature. The algorithm insures that all structures generated at a given temperature are of minimal energy.

### III. THERMODYNAMIC STUDY

The algorithm described above allows us to generate conformations where their geometries depend on temperature. Two examples are given to illustrate this situation. At low temperature, as it is shown in figure 1, the generated chain is of compact structure with maximum number of contacts, whereas at high temperatures (figure 2) generated chain is unfolded with less number of contacts.



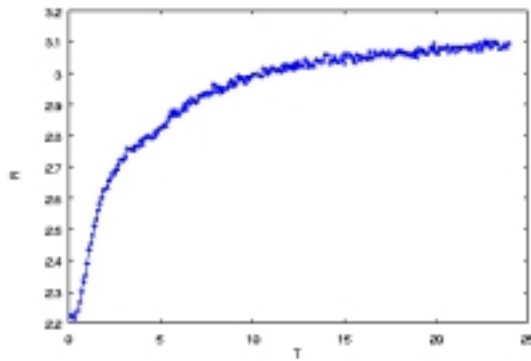
**FIG.1.** A homopolymer chain of 60 monomers generated in a bidimensional square lattice at low temperature.



**FIG.2.** A homopolymer chain of 60 monomers generated in a bidimensional square lattice at high temperature.

#### A. THE BEHAVIOUR OF THE RADIUS OF GYRATION R

Then the radius of gyration  $R$  characterise the conformation geometry and the native state is determined by the compactness of the chain. It's reasonable to study the collapse transition by analysing the behaviour of  $R$  with temperature  $T$ . So we construct the chain a great number of times at a fixed value of temperature in order to calculate the mean distance  $R$  as a function of temperature. Figure 3 represent the evolution of the mean distance  $R$  of a chain of 20 monomers. The curve shows that under a critical temperature characterising the collapse transition  $T_c = 0.5$ ,  $R$  is minimal and approximately constant  $R = R_b$  (for  $N=20$  we can have  $R_b = 2.22$ ). This means that in this range of temperatures all conformations generated by the chain growth algorithm are compact. As the temperature increases from  $T_c$ ,  $R$  begin to increase. In this situation the chain trays to occupy large surface because of contacts breaking by heat effect. The same behaviour had been observed using the gyration radius [9].



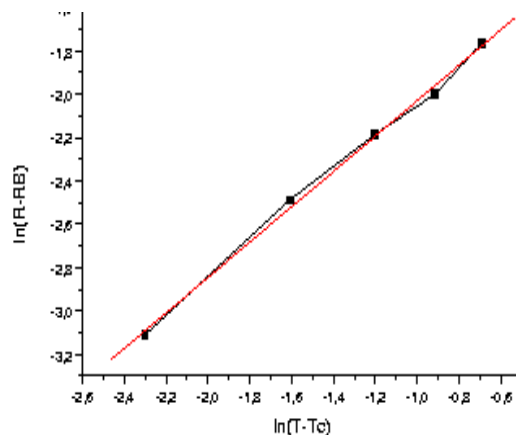
**FIG.3.** The evolution of the mean distance  $R$  as a function of temperature of a chain of 20 monomers.

### B. CRITICAL EXPONENTS

The critical exponent of the parameter  $R$  can be calculated through the relation :

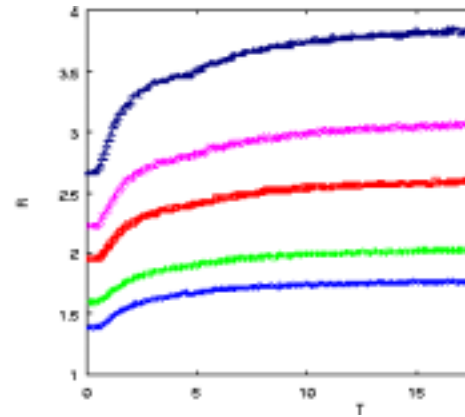
$$R - R_b = (T - T_c)^\beta$$

Numerically we find  $\beta$  by a linear fit of the curve representing the function defined by  $\ln(R - R_b) = f(\ln(T - T_c))$  (figure 4) which gives  $\beta = 0.817$

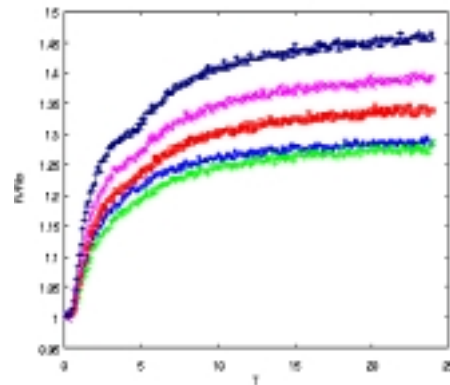


**FIG.4.** The curve representing  $\log(R - R_b) = f(T - T_c)$  for a chain of 20 monomers.

In order to study the size effect of the chain on collapse, we represent the distance  $R$  for different sizes (figure 5), we see that the transition point for short chains doesn't change with the size changes. We also draw the normalised curves  $R/R_b = f(T)$  for different values of the chain length (figure 6). This figure shows that near the transition point all curves are approximately confounded which mean that the critical exponent  $\beta$  doesn't change with size variation.



**FIG.5.** Curves representing the mean distance  $R = f(T)$  for different sizes ,from the top to the bottom respectively  $N=30, N=20, N=15, N=10$  and  $N=8$ .



**FIG.6.** Curves representing the mean distance  $R/R_b = f(T)$  for different sizes ,from the top to the bottom respectively  $N=30, N=20, N=15, N=10$  and  $N=8$ .

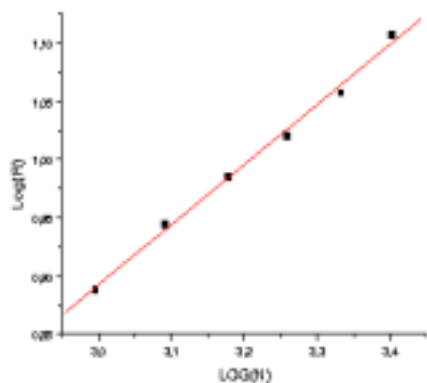
For  $T = T_c$  we also calculate the critical exponent defined by the relation :

$$R = N^\nu$$

where  $N$  is the chain length.  $\nu$  is determined by a linear fit of the curve representing  $\ln(R) = \ln(N)$  (figure 7) which gives the value  $\nu = 0.515$ . This value can be compared with the exponent  $\nu = 4/7$  found in the case of the two dimensional square lattice of a self avoiding walk dynamic using the gyration radius with  $N \rightarrow \infty$  [11]. We can then calculate the critical exponent  $\nu_1$  by analogy with magnetic systems considering  $R$  as an order parameter is defined by the relation the relation:

$$R = N^{\beta/\nu_1} f$$

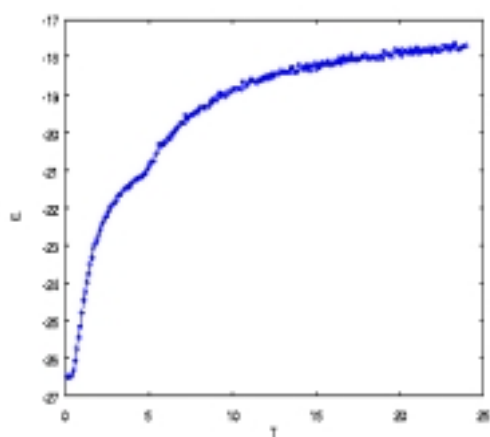
where  $f$  is a homogenous function. So  $\nu = \beta/\nu_1$  which gives  $\nu_1 = 1.06$ .



**FIG.7.** Curve representing  $\text{Log}(R)=f(\text{Log}(N))$  for a chain of 20 monomers

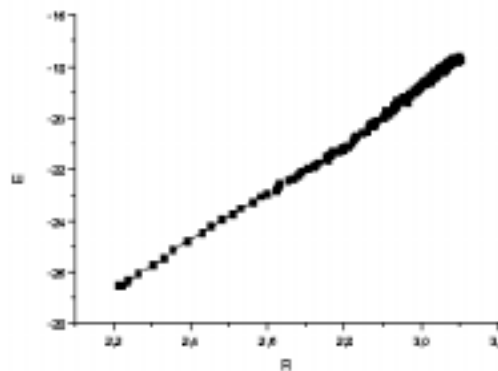
### C. CONFORMATIONAL ENERGY OF THE CHAIN

The mean conformational energy of the chain is calculated by constructing the chain big number of times at a given temperature, and averaging by the number of constructions. Figure 8 gives the shape of the mean energy as a function of temperature. From this curve we observe for  $T < T_c$  that the energy is approximately constant and minimal which geometrically corresponds to compact structures, and as the temperature increases from the value  $T=T_c$  the energy increases and so the conformation becomes unfolded by minimising the number of contacts between its monomers. The same behaviour is observed using the chain relaxation of a heteropolymer protein [12].



**FIG.8.** The conformational energy as a function of temperature for a chain of 20 monomers.

We also represented energy as a function of the parameter  $R$  (figure 9), we find that this curve is linear which means that the conformational energy is proportional to the distance  $R$ . And so the energy can be calculated as a function of the distance  $R$  by the relation  $E=AR+B$ , a linear fit of this curve gives us the values  $A=10,434$ ,  $B=-50,05$ .



**FIG.9.** The conformational energy as a function of the mean distance  $R$  for a chain of 20 monomers.

### D. SPECIFIC HEAT

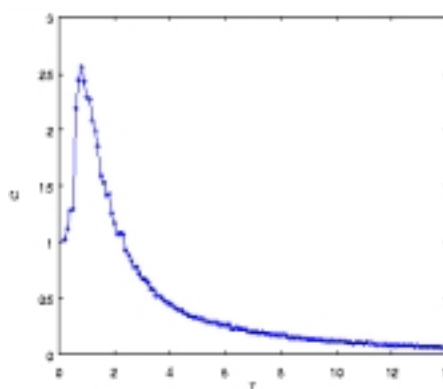
In this simulation the specific heat of the hydrophobic is defined by the relation :

$$C=dE/dT$$

and by the energy fluctuations:

$$C=1/(KT)^2(\langle E^2 \rangle - \langle E \rangle^2)$$

Where  $E$  is the conformational energy and the average is over all chain constructions at a fixed temperature. Figure 10 represents the specific heat as a function of temperature, the curve shows that  $C$  presents a peak at a temperature  $T=0.8$ , this behaviour is observed in heteropolymers using the relaxation of the chain[12].

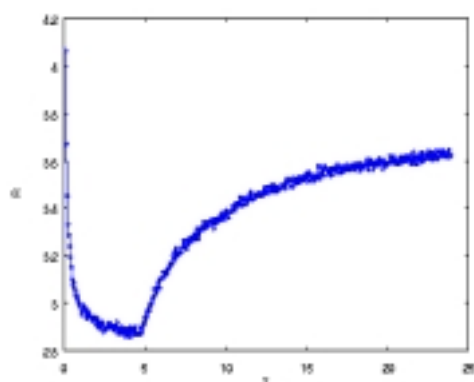


**FIG.10.** The specific heat as a function of temperature  $C=f(T)$  for a chain of 15 monomers.

### E. APPLICATION OF A FORCE ON THE PROTEIN

While growing the chain, we apply a force  $F$  in a fixed direction. The role of this force is to push the chain to be unfolded. And so we will have a competition between the power of the

hydrophobic interaction which has the tendency to collapse the conformation and the force  $F=0.08$ . Figure 11 shows the evolution of the chain grown by the algorithm in presence of this force. The behaviour of the distance  $R$  shows that at low temperature the protein is opened and as the temperature increases  $R$  decreases and so the conformation becomes collapsed. For a temperature  $T=5$ ,  $R$  increases and so the chain begins to become unfolded. From this result we can say that the application of the force on the hydrophobic chain, when it is growing leads to the appearance of two transitions: one at low temperature and the second at high temperature. This seems to be the same as results in previous studies showing the existence of two transitions called cold denaturation and warm denaturation [13,14]. We see also that the critical temperature has changed in the presence of the force, this had been observed when a force is applied on secondary structures proteins [15]. As a conclusion we can say that the force can be considered as a potential that has to be added to the Hamiltonian of the model in order to show the cold denaturation which disappears when the force isn't applied.

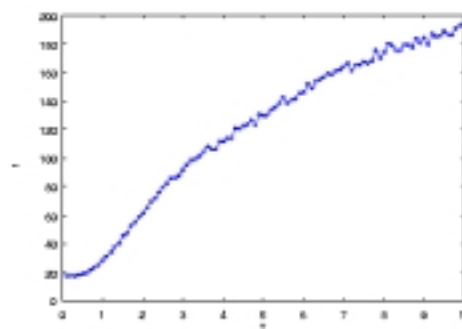


**FIG.11.** Curve representing  $R=f(T)$  with a force  $F=0.08$  applied on the chain in a fixed direction for a chain of 20 monomers.

#### IV. KINETIC STUDY

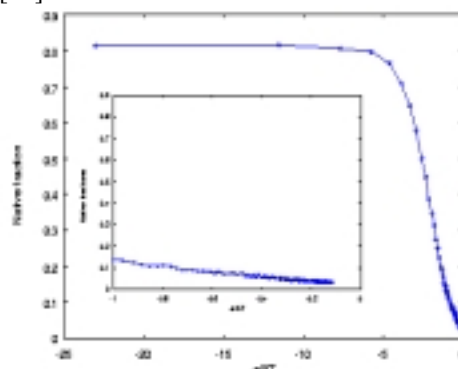
In order to introduce the kinetic in our chain growth algorithm, we give the chain while it is growing more chances to reach the native structure. We realise this kinetic by the following procedure: we grow the chain until its full length and we calculate its energy. If the conformation grown isn't native we break randomly a certain number of links and we regrow the chain until its full length is reached. We repeat this each time until reaching the native state and then we stop breaking links. The time is then calculated as the number of links broken to reach the native conformation at a fixed temperature. Figure 12 shows the evolution of the mean time of a given conformation to reach its native state as a function of temperature. We see

that the time increases as the temperature increases, which is in accord with results obtained for homopolymer's studies [16]. We also observe that the time values found by this model are smaller than those found by the chain relaxation models, it can be explained by the fact that in the growth model the chain can reach easily its native state than in other models.



**FIG.12.** The mean time for a chain to reach its native conformation  $t=f(T)$

We also represent the native fractions as a function of  $-e/KT$  where  $-e=-2.3$  is the contact energy in figure 13. The shape of this curve shows the same behaviour as observed in previous studies [16].



**FIG.13.** Curve representing the native fractions as a function of  $-e/KT$  with  $e=2.3$  is the contact energy

#### V. CONCLUSION

We have studied a homopolymer chain by a chain growth model as a first step for understanding heteropolymer protein folding. Our results show that the mean distance  $R$  is a good parameter for studying the critical behaviour of the chain. We have also calculated and represented thermodynamical functions such as the conformational energy and specific heat, we have examined the effect of a force applied on the chain while it is growing which also leads to the appearance of two transitions one at low temperature and the second at high temperature. In the kinetic study we have represented the mean

time for a given conformation to reach the native structure as a function of temperature, and represented at a second time the native fractions as a function of temperature. From all these results we can conclude that the chain growth model is a good approximation to understand the protein folding mechanism because of the reason that the process of forming proteins in nature is like the chain growth model.

### References

- [1] K. A. Dill and D. Stigter, *adv protein chem* 46, 59 (1995) .
- [2] K. A. Dill, *Biochemistry* 29, 7133 (1990).
- [3] H. S. Chan and K. A. Dill, *physics today* Vol. 46, No. 2, 24 (1993).
- [4] K. A. Dill, o. V. A. Darwin and H. Karen, *Biochemistry*, 28, 5439 (1989).
- [5] K. A. Dill, *Protein Science* 8, 1116 (1999).
- [6] H. S. Chan and K. A. Dill, *Proteins: structure, function and genetics* 30, 2 (1998).
- [7] J. C. Camacho, D. Thirumalai, *Proc.Natl, Sci.USA.* 90, 6369 (1993).
- [8] R. Miller,C. A. D. Michael, J. Fasolka, and A. C. Balazs *chem physics* 96, 768 (1991).
- [9] H. S. Chan and K. A. Dill, *J chem.phys* 100, 9238 (1994).
- [10] H. H. Gan, A. Tropsha, and T. Schlick, *J of chem phys* 13, 5511 (2000).
- [11] B. Duplatier and H. Saleur, *phys.Rev.lett* 59, 539 (1987).
- [12] K. A. Dill, S. Bromberg, K. Yue, K. M. Fiebig, D. P. Yee, P. D. Thomas and H. S. Chan, *Protein Science* 4, 561 (1995) .
- [13]P. L. Rios, and G. Caldarelli (2000)
- [14]K. A. Dill and D. Shortle, *Annu Rev Biochem* 60, 795 (1991).
- [15] H. Zhou, *Cond-Mat*(2001) 0112090 v1.
- [16] H. S. Chan and K. A. Dill, *J.Chem .Phys* 99, 2116 (1993).