Computer simulation of knotted proteins unfold and translocation through nano-pores

M. A. Shahzad ^{1,2}

¹ *Department of Physics, Hazara University, Pakistan.*

² *Dipartimento di Fisica, Università di Camerino, Camerino, Italy*

(Dated: July 27, 2018)

We study the unfold and translocation of knotted protein, YibK and YbeA, through α -hemolysin nano-pore via a coarse grained computational model. We observe that knot of protein unfold in advance before the translocation take place. We also characterized the translocation mechanism by studying the thermodynamical and kinetic properties of the process. In particular, we study the average of translocation time, and the translocation probability as a function of pulling force F acting in the channel. In limit of low pulling inward constant force acting along the axis of the pore, the YibK knotted protein takes longer average translocation time as compare to YbeA knotted protein.

I. INTRODUCTION

The passage of a molecule through a biological pore, the so-called translocation, is a fundamental process occurring in a variety of biological processes [1], such as virus infection [2], and gene expression in eukaryotic cells [3]. The translocation process involves the exchange of proteins, ions, energy sources, genetic information and any particle or aggregate that plays a role in cell functioning and living. For instance, translocation of bio-polymer across membranes channel is universal in biology and includes the passage of RNA strands inside nuclear pores [4], absorption of oligonucleotides on suitable sites [5], transport of protein back and forth from the cells [6], mitochondrial import, protein degradation by Adenosine triphosphate (ATP)-dependent proteases and protein synthesis [7, 8].

The biological cell is covered by all types of pores and channels that control exchange of ions and molecules between sub-cellular compartment. These pathways are of vital importance to cellular function. Examples are: ion channels at the cell surface that control the flow of ions; pores inserted into cell membranes upon viral infection that serve as conduits for genome transfer; the nuclear pore complex (NPC) that regulates the transport of mRNA and proteins across the nuclear envelope of eukaryotic cells; and pore across the used for protein secretion into cell organelles.

Proteins play an important role in transport process across the cell membranes. The transport can be either passive or active. The passive transport involves the movement of biochemicals from areas of high concentration to areas of low concentration and requires no energy for transportation. The passive transport can take place by diffusion or by interposition of carriers to which the transported substances are bounded. On the other hand, active transport are biological processes that move oxygen, water and nutrients into cells and remove waste products. Active transport requires chemical energy because it is the movement of biochemicals from areas of lower concentration to areas of higher concentration. In active transport, trans-membrane proteins use the Adenosine-Tri-Phosphate (ATP) chemical energy

to drive a molecule against its concentration gradient. A mechanism of selection, based on the interaction between signal recognition particles and relative receptors, allows the trans-membrane proteins to detect the system to be transported.

Proteins are not only the building blocks of trans-membrane channels, but are also commonly transported between organelles and cells, in folded and unfolded form. Very often membrane apertures are not large enough to enter the proteins in their native (folded) conformations. For instance, the narrowest constriction in the proteasome organelle is only 13 Å in diameter. It is thus necessary for proteins to unfold in advance or during their passage across the pore [10, 11].

Here we used a coarse-grained computational model to simulate numerically the unfold and translocation process for the knotted protein through a α -hemolysin channel. We show that the knot in knotted protein unfold during the translocation process.

II. COMPUTER MODELING OF STRUCTURED POLYPEPTIDE CHAIN AND NANOPORE

We used a coarse-grained computational model to simulate numerically the translocation process for the protein through a channel. We consider the Gō-like model as a natural approach to access the impact of the molecules structural properties on translocation. Below we show in detail the importing mechanism, pore and particle interaction potential, numerical integration scheme and equation of motion for the material points constituting the system.

A. Protein numerical model: The Gō-like approach

Protein folding is a complex process involving large number of degree of freedom. A linear polypeptide protein chain is organized into a space-filling, compact, and well-defined 3-dimensional structure. The monomeric unit in a polypeptide chain is the peptide group. The C, O, N, and H atoms in the amino acid lie in the same

FIG. 1. Left upper panel: Full atomic representation of YibK knotted protein. Right upper panel : Secondary structure of YibK revealed from its protein data bank file, PDB-ID: IJ85. Lower left panel : YibK protein coarse grained description. In Gō model only the C_{α} carbon atoms along the main backbone chain of the protein are depicted. Lower right panel: The rectangle represent the knotted region .

plane, and successive planes defined angles ϕ and ψ . The conformation of a chain of n amino acids can be defined by $2n$ parameters. Therefore, performing traditional molecular dynamics simulation with atomistic detail poses a challenge if one wished to exhaustively explore the folding landscape of a moderately sized protein with finite computation resources. For this reason coarse-grained representation of the polypeptide chain have been employed in folding simulation. The phenomenological off-grid model proposed by Nobuhiro-Gō and Harold A. Scheraga [12, 14] is a coarse-grained model that identifies all the amino acids with their C_{α} carbon atoms. Rather than representing each atom in the protein explicitly groups of atoms can be treated as a single coarse-grained site. This way, each residue corresponds to a single coarse-grained site placed at the C_{α} carbon atom. Through this way, protein

are reduced to a chain of material points coinciding spatially with the C_{α} carbon atoms of the protein backbone and no side chain characteristic are retained. The atomic mass of the material points are the same and equal to 1 in the so-called code measure units or internal units. Different refinements of this model are available in literature [13, 15–19], and are generally referred as Gō-Like models. The approach presented in our numerical work is adopted from reference [13, 20, 22, 23].

The native structure of a protein is taken from the protein data bank (PDB) file and the interaction between the amino acids are grouped into native and non-native state. The protein data bank (PDB) format file contains information about atomic coordinates for various types of proteins, small molecules, ions and water. It provides a standard representation for macromolecular structure

FIG. 2. Structure of YbeA from *Escherichia*. Left upper panel: Full atomic representation of YbeA knotted protein. Right upper panel: Secondary structure of YbeA revealed from its protein data bank file, PDB-ID: 1NS5. Left lower panel: YbeA protein coarse grained description. In Gō model only the C_{α} carbon atoms along the main backbone chain of the protein are depicted. Right lower panel: The rectangle represent the knotted region .

data which is derived from X-ray diffraction and nuclear magnetic resonance (NMR) studies.

In Gō-like model, energy minima are assigned to a particular reference configuration, usually coincident with the native structure of the considered protein. This is why Gō-models, where only native contacts contribute to the free energy, can be properly applied to small globular protein whose crystalline structure is already known and the folding dynamics are mainly determined by the topology of the native conformation. The Gō-like model generates almost ideal folding pathway. Also, the native conformation so obtained are stable and closed to the reference one, which can be extracted from PDB. The Gō-like model is ideal for two-state folding dynamics simulation, but it has been used also to predict intermediate configuration [13].

The main limitation of the Gō-model is that its topo-

logical nature does not account for chemical factors, determining a considerable disagreement between real and simulated time scale. However, these discrepancies do not alter the folding thermodynamics and the correct sequence of events. In most cases the model allows the correct assessment of the free energy and of folding pathways.

Using the Gō-like approach for the numerical simulation, we used the following potentials acting on the material point of the protein: (1) Peptide potential (or bond potential) V_p , (2) Bending angle potential V_{θ} , (3) Twisted angle potential V_{ϕ} , and (4) Non-bonded interactions (Lennard-Jones potential or barrier) V_{nb} . These interactions are the simplest potential energy function that can reproduce the basic feature of protein energy landscapes at a mesoscopic level of detail, and it has proved to give insight into a remarkably broad range of prop-

FIG. 3. Schematic view of knotted protein pulled across a α -hemolysin nano-pore from CIS to TRANS-side. The confinement effect inside the pore is model by a step-like $[0, L]$ cylindrical soft-core repulsive potential. A constant force F_x in the direction of x -axis is used to drive the protein inside the channel. The force is always applied to the foremost residue inside the pore .

erties. The combinations of a potential energy function and all the parameters that go into it constitutes a so-called force field.

Let $\mathbf{r}_i (i = 1, \dots, m)$ be the position vector of the m residues identified by their C_α Carbon atom in the reference (native) configuration, and $\mathbf{r}_i^0 (i = 1, \dots, m)$ be the position vector of m residue in the current configuration. The peptide potential, responsible for the covalent bonds between the beads of the polymer chain, has the following expression:

$$V_p(\mathbf{r}_{i,i+1}) = \frac{k_p}{2} (\mathbf{r}_{i,i+1} - \mathbf{r}_{i,i+1}^0)^2, \quad (1)$$

where $\mathbf{r}_{i,i+1} = \mathbf{r}_i - \mathbf{r}_{i+1}$, $\mathbf{r}_{i,i+1}^0 = \mathbf{r}_i^0 - \mathbf{r}_{i+1}^0$ are the position vector between the bonded residues i and $i + 1$ in the instantaneous and native configuration respectively. The norm of position vector is the bond length. The empirical constant $k_p = \frac{100\epsilon}{d_0^2}$ is in term of the equilibrium length parameter d_0 , and is in the unit of energy parameter ϵ . In our simulation, $d_0 = 3.8$ Å is the average distance between two adjacent amino acids and ϵ sets the energy scale of the model. The bond potential is nothing more than a simple harmonic potential with spring constant k_p .

The angular potential V_θ is used to recover the secondary structure of protein in reference native conformation. Mathematically, it is equivalent to peptide potential V_p by replacing the relative displacement by angular

difference, that is

$$V_\theta(\theta_i) = \frac{1}{2} k_\theta (\theta_i - \theta_i^0)^2, \quad (2)$$

where $k_\theta = 20\epsilon \text{ rad}^{-2}$ is the elastic constant expressed in term of the energy computational unit ϵ , and θ_i, θ_i^0 are bond angles formed by three adjacent beads in the simulated (time-dependent) and native conformation, respectively.

The dihedral potential (torsion) is 1-4 interaction, and are expressed as a function of the twisted angles ϕ_i and ϕ_i^0 , again refereed respectively to the actual and crystal configuration. The dihedral potential is important for the recovery of the correct protein secondary structure. The twisted angle is the angle formed between the two planes determined by four consecutive amino acids along the chain. The definition of twisted angle potential V_ϕ is

$$V_\phi(\phi_i) = k_\phi^{(1)} [1 - \cos(\phi_i - \phi_i^0)] + k_\phi^{(3)} [1 - \cos 3(\phi_i - \phi_i^0)], \quad (3)$$

where $k_\phi^{(1)} = \epsilon$ and $k_\phi^{(3)} = \epsilon/2$ are dihedral constants expressed in term of energy unit ϵ .

Non-bonded (nb) interactions between nonconsecutive amino acids are modeled with Lennard-John 12-10 potential. In Gō-like model a distinction is made among the pairs of residues that interact following a potential that has also an attractive part, in addition to a repulsive part. The criteria for this distinction is made on the

FIG. 4. Snapshots of unfold mechanism. Upper left Panel: YibK protein with knot region marked by rectangle. Upper Right Panel: Unfold of the knot during the first simulation time. Lower left Panel: YbeA protein with knot describe by rectangle. Lower Right Panel: Unfold of the knot obtained after the first simulation run time. .

basis of the native distance with respect to a parameter of the model, the so-called cut-off radius, R_c . Native contact [24] are identified by the cut-off distance R_c to be chosen such that two residues i and j form a native interaction if their distance r_{ij} in the native state is less than cut-off distance R_c . This criterion selects a certain number of native contacts on the Proteins crystallographic structure. On the other hand, when two residues are not in native contact $r_{ij} > R_c$, they interact only through the Lennard-Jones repulsive tail $(\sigma/r_{ij})^{12}$, where $\sigma = 4.5 \text{ \AA}$

is a free length parameter correlated with the extension of the excluded volume (self-avoiding polymer chain). In other words, such residues in the protein chain will interact attractively if brought to a distance greater than the native one and repulsive otherwise. The expression for Lennard-Jones potential is:

$$V_{native}(r_{ij}) = \epsilon \left[5 \left(\frac{r_{ij}^0}{r_{ij}} \right)^{12} - 6 \left(\frac{r_{ij}^0}{r_{ij}} \right)^{10} \right], \quad (4)$$

where all the symbols have been already defined. When

FIG. 5. Ubiquitin translocation probability P_{Tr} as a function of force. The error-bar are calculated using the statistical error formula .

$r_{ij}^0 > R_c$, the purely repulsive contribution $V_{nonnative}$ is assigned to the pair of amino acids considered. This term enhances the cooperativity in the folding process and takes the form of a Lennard-Jones barrier

$$V_{nonnative}(r_{ij}) = \frac{10}{3}\epsilon\left(\frac{\sigma}{r_{ij}}\right)^{12}. \quad (5)$$

The non-bonded potential V_{nb} summarized the possible long range interaction just described above and reads as

$$V_{nb}(r_{ij}) = \begin{cases} \epsilon \left[5 \left(\frac{r_{ij}^0}{r_{ij}} \right)^{12} - 6 \left(\frac{r_{ij}^0}{r_{ij}} \right)^{10} \right] & r_{ij}^0 \leq R_c \\ \frac{10\epsilon}{3} \left(\frac{\sigma}{r_{ij}} \right)^{12} & r_{ij}^0 > R_c. \end{cases} \quad (6)$$

The total potential acting on all the residues of the proteins is then :

$$V_{Go}(r_{ij}) = \sum_{i=1}^{m-2} V_p(r_{i,i+1}) + \sum_{i=1}^{m-2} V_\theta(\theta_i) + \sum_{i=1}^{m-3} V_\phi(\phi_i) + \sum_{i,j>i+3} V_{nb}(r_{ij}). \quad (7)$$

Figure (1, left upper Panel) illustrate the basic atomic structure of YibK knotted protein. The secondary structure with beta-sheets and α -helices is shown in fig (1, right upper Panel). The coarse-grained representation of knotted protein is described in fig (1, left lower Panel) with color that vary according to its residues type. Figure (1, right lower Panel) illustrate the Yibk knotted protein with trefoil knot region marked my rectangular. The detail structure of YbeA knotted protein are shown in Figure (2 (Figures produced by VMD-software).

B. Pore Model

The confinement effect on protein dynamics can be represented by a step-like soft-core repulsive cylindrical potential. The cylinder axis of symmetry is set adjacent with the x -axis of the frame of reference used for protein translocation simulation. The same x -axis direction is used to develop the mechanical pulling of the protein by using a constant force F_x applied to the foremost beads inside the confinement. The constant force is used in analogy with the electrical potential in voltage driven translocation experiment. The schematic representation of the system, the potential and the pulling mechanism are shown in Fig. (3).

The expression of the pore potential is given by:

$$V_c(x, y, z) = V_0 \left(\frac{\rho}{R_p} \right)^{2q} \Theta[x(L - x)], \quad (8)$$

where $V_0 = 2\epsilon$ and $\Theta(s) = [1 + \tanh(\alpha s)]/2$ is a smooth step-like function limiting the action of the pore potential in the effective region $[0, L]$. L and R_p are pore length and radius respectively. Also, $\rho = \sqrt{(y^2 - z^2)}$ is the radial coordinate. The parameter q tunes the potential (soft-wall) stiffness, and α modulates the soft step-like profile in the x -direction; the larger the α , the steeper the step. In our simulation, we consider $q = 1$ and $\alpha = 2 \text{ \AA}^2$. The driving force F_x acts only in the region in front of the pore mouth $x \in [-2, 0]$, and inside the channel $[0, L]$. Pore length $L = 100 \text{ \AA}$ and radius $R_p = 10 \text{ \AA}$ are taken from α HL structure data [25].

C. Equation of Motion: Langevin dynamics

The equation of motion which governs the motion of material point is the well-know Langevin equation which is usually employed in coarse-grained molecular dynamics simulation. Consequently, numerical investigations are performed at constant temperature. The over-damped limit is actually implemented ($\ddot{r} = 0$) and a standard Verlet algorithm is used as numerical scheme for time integration [26]. The Langevin equation is given by

$$M\ddot{\mathbf{r}}_i = -\gamma M\dot{\mathbf{r}}_i + \mathbf{F}_i^{conf} + \mathbf{W}_i \quad (9)$$

where $\mathbf{F}_i^{conf}$ is the sum of all the internal and external forces acting on residue i . Here γ is the friction coefficient used to keep the temperature constant (also referred as Langevin thermostat). The random force $\mathbf{W}_i$ accounts for thermal fluctuation, being a delta-correlated stationary and standard Gaussian process (white noise) with variance $\langle W_\alpha(t)W_\alpha(t') \rangle = 2\gamma MRT\delta_{\alpha,\alpha'}\delta(t - t')$. The random force satisfies the fluctuation-dissipation theory; the mean-square of $\mathbf{W}$ is proportional to the corresponding friction coefficient γ .

FIG. 6. Trajectory of YibK knotted protein in collective coordinate Q as a function of time using a constant external pulling force of $F = 0.75$ at the C-terminus. YibK: A trace of the number of peptide count at the CIS-side (filled blue circle), and TRANS end (filled brown triangle).

III. RESULTS AND DISCUSSION

We used Langevin molecular dynamics simulation (Verlet integration scheme) to transport knotted protein YbeA (PDB ID 1NS5) and YibK (PDB ID 1J85) across a static α -hemolysin pore. The folding temperature occurred at $T^* = 0.77$ in reduced temperature units R/ϵ , corresponding to the experimental denaturation temperature $T = 338$ K. This defined the energy scale to the value $\epsilon \simeq 0.88$ kcal mol $^{-1}$. Using the time scale $t_u = \sigma(M/120\epsilon)^{1/2}$ we can obtained the physical time unit from the simulation time. With $\epsilon \simeq 0.88$ kcal mol $^{-1}$, $\sigma = 4.5$ Å, and assuming the average mass of amino acid is $M \sim 136$ Da, we get $t_u \sim 0.25ps$. In computer simulation the time step and friction coefficient used in the equation of motion (Langevin dynamics) are $h = 0.001t_u$ and $\gamma = 0.25t_u^{-1}$, respectively. Further, the unit of force is defined as $f_u = \epsilon \text{ Å}^{-1} \sim 6$ pN.

A. Knotted protein Unfold

The protein in its native conformation is placed initially at the CIS-side of the pore. As shown in fig (4, upper left and right panel) the knotted region of Yibk knotted protein are marked by rectangle with residues color follow by red, blue, green, yellow, black and white. During the first simulation run time the knot unfold as describe in figure () with the same residue colors. In contrast to Piotz Szymczak [27] finding where an on/off pulling force untie the knot of protein, we observed that a simple constant pulling force can able to unfold the knot of knotted protein. Figure (4, lower left and right panel) show the initial fold conformation of YbeA knotted protein at the begiing of simulation. The lower right panel of figure () decrbe its unfold conformation. The mechanism of unfold of the trefoil knot in our simulation

is similar to fold mechanism as observed by Stefan Wallin *et al.*[28].

B. Knotted protein translocation

The simulation was run until the protein was fully expelled out from the cis to tans-side, and almost complete refolding occurred. However, the translocation may fail at low external force F within a large assigned waiting time t_w ; in this case, the run stops, discards its statistics, and restarts a new trajectory. Figure (5) illustrates the plot of translocation probability P_{Tr} as a function of force. The probability of translocation P_{Tr} can be defined as the number of translocation successes over the number of total runs, within the time $t_w \simeq 15^3 t_u$. The translocation becomes ($P_{Tr} > 1/2$) when the force exceeds a critical value $F_c \simeq 0.75f_u$ for YbeA and $F_c \simeq 0.5f_u$ for YibK knotted protein, while below these thresholds the probability goes rapidly to 0. The red square data represent the translocation probability for YibK protein while the filled blue circle describe the YbeA knotted protein translocation probability data.

Defining the collective variable coordinate Q ,

$$Q = \frac{1}{N} \sum_{i=1}^N q(x_i) \quad (10)$$

where x_i is the x -coordinate of the i th bead and $q(x)$ is the piecewise function defined as

$$q(x) = \begin{cases} 0, & \text{if } x < 0 \\ \frac{x}{L}, & \text{if } 0 \leq x < L \\ 1, & \text{if } x \geq L. \end{cases} \quad (11)$$

The value of collective variable coordinate $Q = 0$ corresponds to the case when the folded protein is completely on the Cis-side of the nano-pore, and $Q = 1$

FIG. 7. Trajectory of YbeA knotted protein in collective coordinate Q as a function of time using a constant external pulling force of $F = 0.75$ at the C-terminus. YbeA: A trace of the number of peptide count at the CIS-side (filled blue circle), and TRANS end (filled brown triangle) .

FIG. 8. Average translocation time τ of ubiquitin as a function of the force F . The red square circle represent YibK average translocation time while the blue circle show the YbeA translocation time data .

when the protein is completely translocate on Trans-side. The value $0 < Q < 1$ represent the coordinate when the beads is inside the nano-pore. Figure (6) and (6) illustrate the few trajectory of knotted protein in collective variable coordinate Q as a function of time during the unfolding and translocation through a α -hemolysin pore.

The translocation process can be further characterized by studying the statistics of translocation times. The time statistics of translocation events is accessible to experiment [11] in which current loop drops signal the occupation of the channel by the passing molecule ; for this reason, these times are also called the blockage times. In the simulation, the translocation time can be measured as first arrival time t at the channel end $x = L$ of the protein center of mass. The average translocation time

is shown in Fig. (8) as a function of force F .

IV. CONCLUSION

We have investigated the mechanical unfold and translocation of knotted protein through a biological α -hemolysin nano-pore using coarse-grained computational model. In this simplified model translocation process we import a knotted protein via a finite size channel by a uniform pulling force acting inside of the channel only. In particular, we present in detail the mechanical unfold and translocation of trefoil knotted protein Yibk and YieA. Our finding shows that the knot of the protein unfold in advance before or during the translocation process. We examine the kinetic and thermodynamical characteristic of the process by studying the statistics of blockage times and the translocation probability as a function of the pulling force F acting in the pore. As shown in the Fig. (5) of probability of the translocation as a function of force, the translocation dynamics occurs when the importing force exceeds a threshold value F_c depending on the free-energy barrier that protein has to overcome in order to slide along the channel. Such a free-energy barrier results from competition of the unfolding energy and the entropy associated with the confinement effects of the pore. Mathematically, the molecular dynamics simulations at the coarse-grained level of protein translocation can be described by a simple drift-diffusion Smoluchowski equation exploiting the first passage theory and radiation boundary condition.

Supporting Information

Movie S1 This movie file (produced by VMD software) represents the unfold and translocation of **YbiaK** through α HL nanopore having length $L = 100$ Å and radius $R_p = 10$ Å.

Movie S2 This movie file (produced by VMD software)

describes the unfold and translocation of **YbeA** through

α HL nanopore having length $L = 100 \text{ \AA}$ and radius $R_p = 10 \text{ \AA}$.

-
- [1] W. Wickner and R. Schekman. Protein translocation across biological membranes. *Science*, 310(5753), 1452:1456, 2005.
 - [2] M. T. Madigan, J. M. Matinko, and J. Parker. *Biology of Micro-organisms*. Prentice, Hall, Englewood Cliffs, NJ, 1997.
 - [3] B. Alberts, K. Roberts, D. Bray, J. Lewis, M. Raff, and J. D. Watson. *Molecular Biology of the Cell*, Garland Publishing, New York and London, 1994.
 - [4] J.O. Bustamante, J.A. Hanover, and A. Liepins. The ion channel behavior of the nuclear pore complex. *Journal of Membrane Biology*, 146(3), 239:251, 1995.
 - [5] B. Hanss, E. Leal-Pinto, L.A. Bruggeman, T.D. Copeland, and P.E. Klotman. Identification and characterization of a cell membrane nucleic acid channel. *Proceedings of the National Academy of Sciences*, 95(4), 1921:1926, 1998.
 - [6] M.S. Sanford and Blobel G. A protein-conducting channel in the endoplasmic reticulum. *Cell*, 65(3), 371:380, 1991.
 - [7] A. Matouschek. Protein unfolding an important process in vivo? *Opinion in Structural Biology*, 13(1), 98:109, 2003.
 - [8] S. Prakash and A. Matouschek. Protein unfolding in the cell. *Trends in Bio- chemical Sciences*, 29(11), 593:600, 2004.
 - [9] M.S. Sanford and Blobel G. A protein-conducting channel in the endoplasmic reticulum. *Cell*, 65(3), 371:380, 1991.
 - [10] John J. Kasianowicz, Eric Brandin, Daniel Branton, and David W. Deamer. Characterization of individual polynucleotide molecules using a membrane channel. *Proc. Natl. Acad. Sci. USA*, 93, 13770:13773, 1996.
 - [11] Amit Meller, Lucas Nivon, and Daniel Branton. Voltage-Driven DNA Translocations through a Nanopore. *Phys. Rev. Lett.*, 86, 3435, 2001.
 - [12] N. Gō and H. A. Scheraga. On the use of classical statistical mechanics in the treatment of polymer chain conformations. *Macromolecules*, 9, 535:542, 1976.
 - [13] C. Clementi, H. Nymeyer, and J.N. Onuchic. Topological and energetic factors: what determines the structural details of the transition state ensemble and on- route intermediates for protein folding? An investigation for small globular proteins. *J. Mol. Biol.*, 298, 937:953, 2000.
 - [14] N. Gō. Theoretical studies of protein folding. *Ann. Rev. of Biophys. and Bioeng.*, 12, 183, 1983.
 - [15] E. Shakhnovich. Theoretical studies of protein-folding thermodynamics and kinetics. *Current Opinion in Structural Biology*, 7, 29:40, 1997.
 - [16] D. Baker. A surprising simplicity to protein folding. *Nature*, 405, 39:42, 2000.
 - [17] J.K. Karanicolas and C.L. Brooks. Improved gō-like models demonstrate the robustness of protein folding mechanisms towards non-native interactions. *J. Mol. Bio.*, 334, 309:325, 2003.
 - [18] J.N. Onuchic and P.G. Wolynes. Theory of protein unfolding. *Curr. Opin. Struct. Bio.*, 14, 70:75, 2004.
 - [19] R.D. Jr. Hills and C.L. Brooks. Insights from coarse-grained gō-models for protein folding and dynamics. *Int. J. of Mol. Sci.*, 10, 889:905, 2009.
 - [20] C. Guardiani, M. Cencini, F. Cecconi. Coarse-grained modeling of protein unspecifically bound to DNA. *Phys. Bio.*, 11, 026003 2014.
 - [21] F. Cecconi, M. Bacci, M. Chinappi. Protein transport across nanopores: a statistical mechanical perspective from coarse-grained modeling and approaches. *Prot. Pept. Lett.*, 21, 227:234 2014.
 - [22] M. Chinappi, F. Cecconi, C.M. Casciola. Computational analysis of Maltose Binding Protein translocation. *Philos. Mag.*, 91, 2034:2048 2011.
 - [23] A. Ammenti, F. Cecconi, U. Marini-Bettolo-Marconi and A. Vulpiani. A statistical model for translocation of structured polypeptide chains through nanopores. *J. Phys. Chem. B*, 113, 10348 2009.
 - [24] S. Miyazawa, Jernigan R. L. Residue-residue potentials with a favorable contact pair term and an unfavorable high packing density term, for simulation and threading. *J Mol Biol.*, 256, 623:44, 1996.
 - [25] M. Chinappi, F. Cecconi, and C.M. Casciola. Computational analysis of maltose binding protein translocation. *Philos. Mag.*, 91, 2034:2048, 2011.
 - [26] L. Verlet. Computer experiments on classical fluids. i. thermodynamical properties of Lennard-Jones molecules. *Phys. Rev.*, 159, 98:103, 1967.
 - [27] P. Szymczak. Periodic forces trigger knot untying during translocation of knotted proteins. *Sci. Rep.* 6, 21702, (2016).
 - [28] S. Wallin, K. B. Zeldovich, and E. I. Shakhnovich. The folding mechanics of a knotted protein. *J Mol Biol.* 368(3): 884-893, 2007.