Protein escape at an atomistic model of the ribosomal exit tunnel

Journal of Physics: Conference Series **1506** (2020) 012022 doi:10.1088/1742-6596/1506/1/012022

Protein escape at an atomistic model of the ribosomal exit tunnel

Phuong Thuy Bui[1] **and Trinh Xuan Hoang**[2]

[1] Institute of Theoretical and Applied Research, Duy Tan University, Hanoi 100000, Vietnam
[2] Institute of Physics, Vietnam Academy of Science and Technology, 10 Dao Tan, Ba Dinh, Hanoi, Vietnam

E-mail: buiphuongthuy@duytan.edu.vn

Abstract. We study the post-translational escape of nascent proteins at the ribosomal exit tunnel by using the Go-like model for proteins and a real shape atomistic tunnel built on the protein data bank (PDB) structure of a ribosome of *Haloarcula marismortui*. The full translation and escape processes of the immunoglobulin binding B1 domain of protein G (GB1) at the tunnel were simulated by using Langevin dynamics. We show that at the simulation temperature corresponding to a physiological temperature, the escape time follows quite well the one-dimensional diffusion model proposed in our earlier works. The relationship between folding and escape obtained for the atomistic tunnel is similar to those obtained previously for the cylinder tunnel.

1. Introduction

In recent years, there have been much efforts to understand the role of the ribosome in protein folding inside cells (for a recent review, see [1]). During biosynthesis, a newly synthesized polypeptide chain emerges from the ribosomal exit tunnel and may begin to fold on the ribosome in a process described as co-translational folding [2]. It has been shown that folding can occur even inside the ribosome tunnel whose dimensions allow the formation of secondary structure elements and simple tertiary units [3, 4]. Co-translational folding however is limited to partially folded conformations. To become fully folded, a nascent protein must escape the exit tunnel after its translation is completed. The escape process should not be too quickly because this would increase the chance of aggregation [5] of the partially folded nascent protein. On the other hand, this process cannot be too slow because it would decrease the productivity of the ribosome. In recent works [6, 7], we have shown that the escape process is concomitant with the folding process at the tunnel and the two processes impact on each other. Folding accelerates escape while a gradual escape from the tunnel helps the protein to fold more correctly. Interestingly, the escape time of protein from the tunnel was shown to be consistent with a simple diffusion model corresponding to the first passage time of a Brownian particle in a linear potential field.

Our previous works however consider only a highly simplified model of the ribosomal exit tunnel, which takes the shape of a hollow cylinder with repulsive wall. Real exit tunnel is known to have irregular shape which also depends on the type of organism [8]. Thus, the aim of our present study is to work on a more realistic model of the ribosomal tunnel to check the validity of the previous results as well as to investigate the effect of tunnel's shape on the folding and escape of nascent proteins. For this purpose, we consider an atomistic tunnel that is based on the PDB

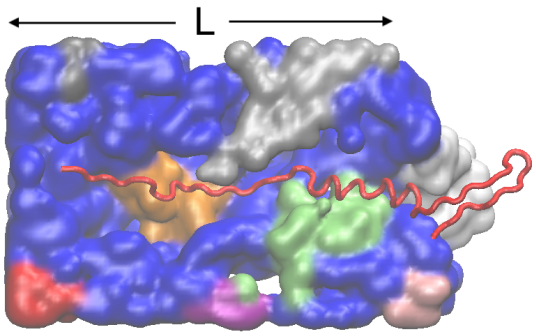

Figure 1. A nascent protein conformation inside the atomistic ribosomal exit tunnel considered in the present study.

structure of the large ribosome subunit of *Haloarcula marismortui*. In our model, only heavy atoms of the ribosomal RNAs are considered whereas the ribosomal proteins are represented only by their C_α atoms. The latter representation is applied also to nascent proteins within a standard Go-like model [9]. In order to accelerate the simulations the ribosome's atoms are kept fixed in all simulations. Thus the tunnel acts as a passive channel and only its shape is taken to the consideration.

2. Models and methods

2.1. Simulation models

For nascent proteins, a Go-like model [9, 10, 11, 12], in which each amino acid residue is represented by a C_α atom, is considered. The intramolecular potential energy of a protein is given by

$$\begin{aligned} E &= \sum_{i=1}^{N-1} K_b(r_{i,i+1}-b)^2 + \sum_{i=2}^{N-1} K_\theta(\theta_i-\theta_i^*)^2 + \sum_{n=1,3}\sum_{i=2}^{N-2} K_\phi^{(n)}[1+\cos(n(\phi_i-\phi_i^*))] + \\ &+ \sum_{i+3<j} \epsilon\left[5\left(\frac{r_{ij}^*}{r_{ij}}\right)^{12} - 6\left(\frac{r_{ij}^*}{r_{ij}}\right)^{10}\right]\Delta_{ij} + \sum_{i+3<j}\epsilon\left(\frac{\sigma}{r_{ij}}\right)^{12}(1-\Delta_{ij}) \ , \end{aligned} \tag{1}$$

where N is the number of residues; r_{ij} is the distance between residues i and j; θ and ϕ are the bond angle and dihedral angle; n is equal to either 1 or 3; the star superscript corresponds to the native state; Δ is the native contact map with Δ_{ij} equal to 1 if there is a native contact between i and j and equal to 0 otherwise. The native contact map is defined based on an all-atom consideration of the protein PDB structure and a contact cut-off distance dependent on the atomic van der Waals (vdW) radii (the C3 map as described in our previous work [7]). Energy is given in units of ϵ, which corresponds to the depth of the 10-12 Lennard-Jones potential given in the fourth term of Eq. (1). The parameters used in the Go-like model are $b = 3.8$ Å, $\sigma = 5$ Å, $K_b = 100\ \epsilon\text{Å}^{-2}$, $K_\theta = 20\ \epsilon(\text{rad})^{-2}$, $K_\phi^{(1)} = -\epsilon$, $K_\phi^{(3)} = -0.5\ \epsilon$.

To build up the atomistic model of the ribosomal exit tunnel, we used the fully refined crystal structure of the *Haloarcula Marismortui* large ribosomal subunit with the PDB code 1jj2. Only a part of the large subunit surrounding the tunnel was taken to be included in the model (Fig. 1). In particular, we excluded the ribosome's atoms that are further than 30 Å from a chosen axis that goes roughly through the middle of the tunnel. We kept only heavy atoms for the ribosomal RNAs and only C_α atoms for the ribosomal proteins forming the tunnel. The interaction potential between the C_α atoms of nascent protein and the C_α atoms of ribosomal

protein is assumed to be repulsive and takes the form of $\epsilon(\sigma/r_{ij})^{12}$ like in the fifth term of Eq. (1). The interaction between a nascent protein's C_α (a) and a ribosomal RNA heavy atom (b) is also repulsive and given by the potential

$$V(r) = \epsilon \left(\frac{R_a + R_b + R_+}{r_{ab}} \right)^{12} , \tag{2}$$

where r_{ab} is the distance between a and b; $R_a = 2.5$ Å is an effective radius of amino acid; R_b is the atomic vdW radius of the ribosomal atom. R_+ is an additive radius accounting for the fact that hydrogen atoms are not considered in the model of the tunnel. R_+ should have a value close but smaller than the vdW radius of hydrogen due to the covalent bonds between hydrogen and the heavy atoms.

The simulations were carried out using molecular dynamics method with the Langevin equations of motion and a Verlet algorithm [6]. The amino acids are assumed to have an uniform mass, m. Temperature is given in units of ϵ/k_B, whereas time is measured in units of $\tau = \sqrt{m\sigma^2/\epsilon}$. We used the same value for the friction coefficient in the Langevin equation, $\zeta = 1\ m\tau^{-1}$. In each simulation, first the nascent polypeptide chain was grown inside the tunnel from the position of the petidyl transferate center (PTC) at a constant speed with the growth time $t_g = 100\ \tau$ per amino acid. It has been shown that this growth time is slow enough to produce a converged behavior of translated nascent protein conformations [6]. After the protein chain is completely translated, the simulation was run until the protein fully escapes from the tunnel. The escape time was measured from the moment of complete translation. We also used an umbrella sampling technique [13] with restraint potentials as given in Ref. [6] to sample low probability's protein conformations along the escape pathways.

2.2. Diffusion model

In the diffusion model [6, 7], the protein escape process is modelled as one-dimensional diffusion of a Brownian particle in a linear potential field with a diffusion constant D and an absorbing boundary condition at a distance L. Given the potential of the form of $U(x) = -kx$ where x is the coordinate of the particle, the distribution of the escape time was found [7] as

$$g(t) = \frac{L}{\sqrt{4\pi D t^3}} \exp\left[-\frac{(L - D\beta k t)^2}{4Dt} \right], \tag{3}$$

where L is the tunnel length and $\beta = (k_B T)^{-1}$ is the inverse temperature. Using the distribution in Eq. (3) one obtains the mean escape time

$$\mu_t \equiv \langle t \rangle = \int_0^\infty t\, g(t)\, dt = \frac{L}{D\beta k} , \tag{4}$$

and the standard deviation

$$\sigma_t \equiv (\langle t^2 \rangle - \langle t \rangle^2)^{\frac{1}{2}} = \frac{\sqrt{2\beta k L}}{D(\beta k)^2} . \tag{5}$$

Note that both μ_t and σ_t diverges when $k = 0$, for which $g(t)$ becomes a heavy-tailed Lévy distribution. It has been shown that D and βk depends on L, on the escaping protein and on other conditions such as the crowders' volume fraction outside the tunnel [7].

3. Results and discussion

As a nascent protein we consider the immunoglobulin binding B1 domain of protein G (GB1) with the PDB code 1pga, which has been considered in our previous studies. This protein has

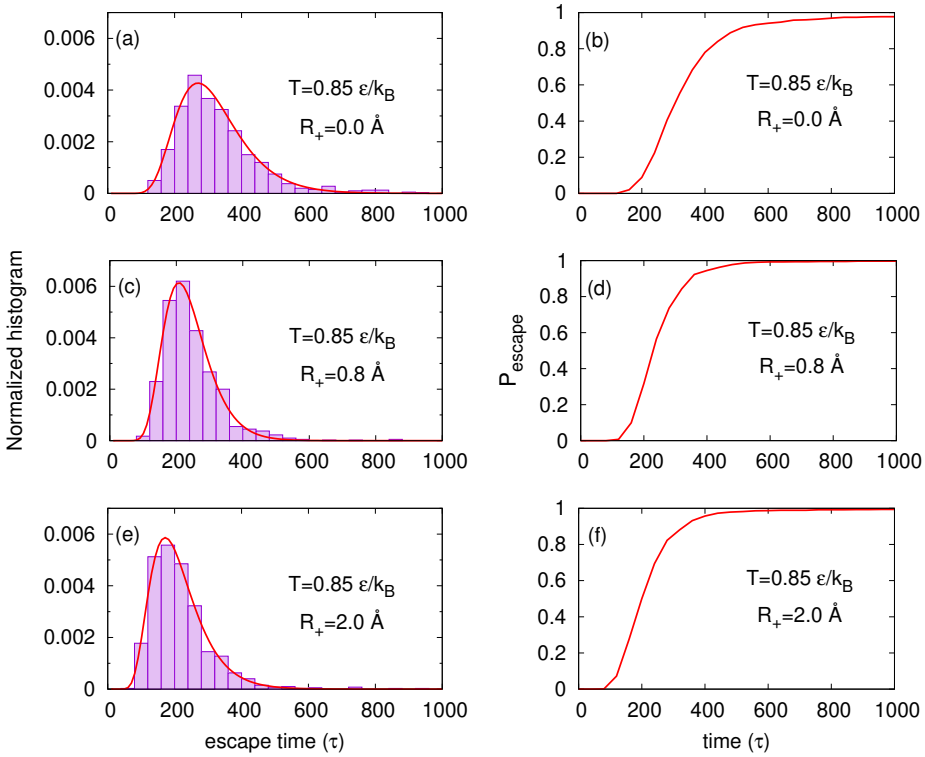

Figure 2. Normalized histogram of the escape time (left) and the time dependence of the escape probability P_{escape} (right) for protein GB1 at temperature $T = 0.85\,\epsilon/k_B$ for $R_+ = 0$ (a,b), 0.8 Å (c,d) and 2 Å (e,f). The data were obtained from 500 independent escape trajectories. The histograms are fitted to the escape time distribution predicted by the diffusion model [7] (solid line).

a length of $N = 56$ residues. An example of a fully translated conformation of GB1 inside the tunnel is shown in Fig. 1. In the present study, we choose the simulation temperature to be $T = 0.85\ \epsilon/k_B$. Note that the folding transition temperature for GB1 in the given Go-like model is $T_f = 1.004\ \epsilon/k_B$ [7]. The melting temperature for wild-type GB1 at pH 5.5 reported by experiment is 80.5°C [14]. Thus, the simulation temperature $T = 0.85\ \epsilon/k_B$ corresponds to a physiological temperature of 26.3°C. Note that T denotes the absolute temperature.

To examine the ribosomal tunnel model, we have tested several values of R_+ for the interaction potential between the nascent protein's amino acids and the ribosomal RNA heavy atoms. Figure 2(a,c,e) shows the escape time distributions obtained for GB1 at temperature $T = 0.85\ \epsilon/k_B$ for R_+ equal to 0, 0.8 and 2 Å. It shows that as R_+ increases, the escape time distribution tends to shrink and move towards smaller values of the escape time, indicating that protein escapes faster as the tunnel becomes narrower. Note that for all cases the escape time distributions are well fitted to the diffusion model proposed in our earlier works [6, 7]. Thus, the diffusion model remains valid for a real shape tunnel and its applicability does not depend on the potential parameters such as R_+.

Figure 2(b,d,f) shows the escape probability (P_{escape}) of GB1 as a function of time for the three considered values of R_+. It can be seen that for $R_+ = 0$, P_{escape} does not reach the value of 1 for the time up to 1000 τ, while this is not the case for $R_+ = 0.8$ and 2 Å. Thus the escape process is sensitive to the width of the tunnel. For $R_+ = 0$, we found that a small but significant fraction of escape trajectories found that the protein is trapped inside the tunnel for a long

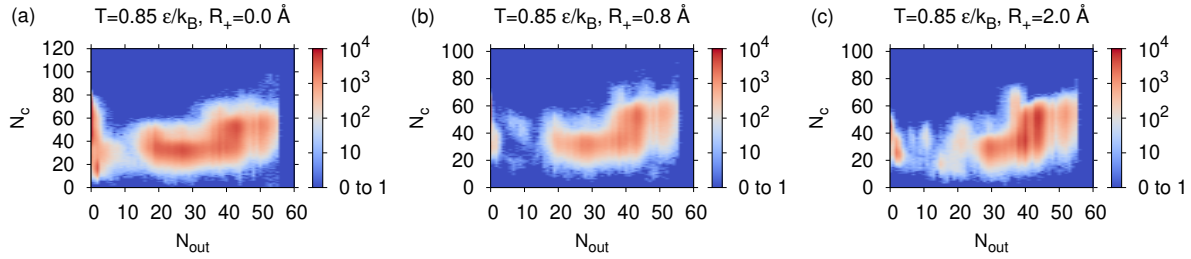

Figure 3. Histogram of protein conformations sampled during the escape process as function of the number or residues outside the tunnel ($N_{\rm out}$) and the number of native contacts (N_c) obtained at $T = 0.85\,\epsilon/k_B$ for the tunnel with $R_+ = 0$ (a), 0.8 Å (b) and 2 Å (c). The data in each histogram were collected from 500 independent escape trajectories.

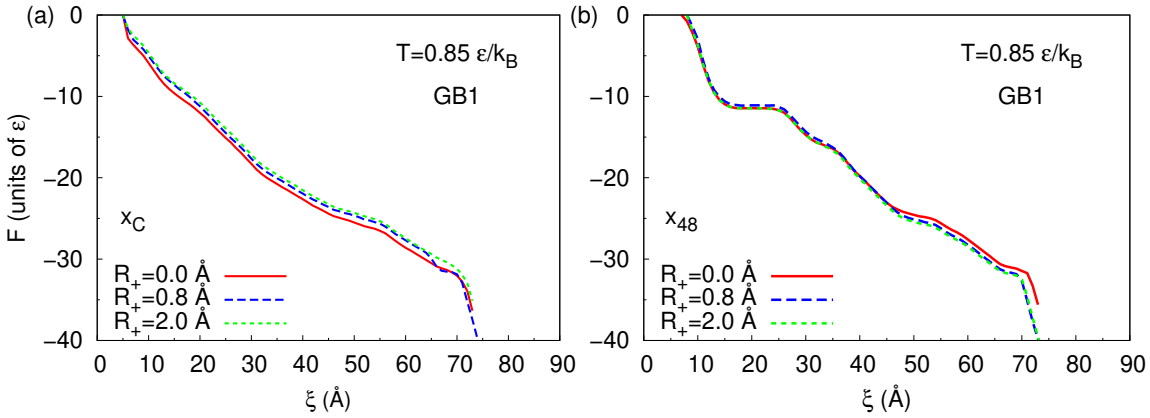

Figure 4. Dependence of the free energy, F, on the coordinate ξ being the x-coordinate of the C-terminal residue (a) and the 48th residue (b) for the protein GB1 escaping the tunnel at $T = 0.85\ \epsilon/k_B$ for $R_+ = 0$, 0.8 and 2 Å, as indicated.

time. Note that the van der Waals radius of hydrogen is 1.09 Å. Thus, physically, both values of $R_+ = 0$ and $R_+ = 2$ Å are not adequate (too small or too big) to account for the presence of hydrogen atoms in the tunnel's wall. The value of $R_+ = 0.8$ Å is compatible with hydrogen vdW radius, and at the same time results in a good escape behavior, i.e. without trapping inside the tunnel, of GB1 as shown in Fig. 2.

Figure 3 shows the histogram of protein conformations sampled during the escape process as function of the number or residues outside the tunnel ($N_{\rm out}$) and the number of native contacts (N_c) obtained for the three considered values of R_+. They look similarly but one can see that as R_+ increases, the cloud of high density points moves towards higher values of $N_{\rm out}$. The appearance of a cloud in the middle ranges of N_c and $N_{\rm out}$ indicates a concomittancy of the escape process with the folding process, even though the broadness of the cloud indicates the stochastic nature of the escape process.

We have calculated also the free energy of escape, F, at the atomistic tunnel for protein GB1 by using the umbrella sampling technique [13] and the weighted histogram method [15]. Following the previous work [6], we have chosen the x-coordinates of the C-terminal residue

and the 48th residue as coordinates for the free energy. These residues correspond to the latest residues in the two typical pathways of the escape process of GB1 at the cylinder tunnel at low temperatures (see Ref. [6]). Figure 4(a) shows that F is a decreasing function of the x-coordinate of the protein's C-terminus. This dependence is not perfectly but closely consistent with the linear potential field of the diffusion model. Figure 4(b) shows that with the 48th residue, the free energy profile has a plateau around $\xi = 20$ Å. This plateau corresponds to a slow escape when the C-terminal β-hairpin is formed inside the tunnel. We have found that for the atomistic tunnel and at $T = 0.85\ \epsilon/k_B$ the probability that the C-terminal β-hairpin is formed inside the tunnel is only a few percent, meaning that the pathway corresponding to Fig. 4(a) is dominating. Figure 4 also shows that the free energies for different values of R_+ differ slightly from each other.

Atomistic ribosomal tunnels have been considered in a number of simulation studies of nascent proteins, but not in the context of protein escape. Elcock [16] has considered the ribosomal tunnel also from Haloarcula marismortui, for which he used an uniform repulsive potential of $(\sigma_{ij}/r_{ij})^{12}$ with σ_{ij} changing from 4 Å to 6 Å for the interaction between the protein C_α atoms and the heavy atoms of the ribosome. In our case, $\sigma_{ij} = R_a + R_b + R_+$ is not uniform and depends on the atom type. Particularly, for carbon of $R_b = 1.7$ Å, one obtains $\sigma_{ij} = 5$ Å given $R_+ = 0.8$ Å. For nitrogen one gets $\sigma_{ij} = 4.85$ Å, for oxygen $\sigma_{ij} = 4.82$ Å, and for phosphorus $\sigma_{ij} = 5.1$ Å.

4. Conclusion

We have worked out an atomistic model of the ribosomal exit tunnel with suitable potentials for studying the escape of nascent proteins. We have shown that a non-zero additive value R_+ is needed for the repulsive potential between the nascent protein and the heavy atoms of RNAs to compensate the missing of hydrogen atoms in the tunnel model in order to get a good escape behavior of the GB1 protein, and that $R_+ = 0.8$ Å is a reasonable value. It is shown that protein escape at a real shape ribosomal exit tunnel is similar to what found for a cylinder tunnel considered in our previous works. The escape time distributions for the atomistic tunnel are consistent with the earlier proposed diffusion model, confirming the strength and validness of this model in studying the escape of nascent proteins from the ribosomal tunnel.

Acknowledgments

One of the authors (TXH) acknowledges the support of VAST for high-level researchers under grant No. NVCC05.09/19-19.

References

[1] Javed A, Christodoulou J, Cabrita L D and Orlova E V 2017 *Acta Crystallogr. D* **73** 509–521
[2] Netzer W J and Hartl F U 1997 *Nature* **388** 343–349
[3] Lu J and Deutsch C 2005 *Nat. Struct. Mol. Biol.* **12** 1123–1129
[4] Kosolapov A and Deutsch C 2009 *Nat. Struct. Mol. Biol.* **16** 405–411
[5] Dobson C M 2003 *Nature* **426** 884–890
[6] Bui P T and Hoang T X 2016 *J. Chem. Phys.* **144** 095102
[7] Bui P T and Hoang T X 2018 *J. Chem. Phys.* **149** 045102
[8] Dao Duc K, Batra S S, Bhattacharya N, Cate J H and Song Y S 2019 *Nucleic Acids Research* **47** 4198–4210
[9] Go N 1983 *Ann. Rev. Biophys. Bioeng.* **12** 183–210
[10] Hoang T X and Cieplak M 2000 *J. Chem. Phys.* **112** 6851–6862
[11] Hoang T X and Cieplak M 2000 *J. Chem. Phys.* **113** 8319–8328
[12] Clementi C, Nymeyer H and Onuchic J N 2000 *J. Mol. Biol.* **298** 937–953
[13] Torrie G M and Valleau J P 1977 *J. Comput. Phys.* **23** 187–199
[14] Campos-Olivas R, Aziz R, Helms G L, Evans J N and Gronenborn A M 2002 *FEBS Lett.* **517** 55–60
[15] Ferrenberg A M and Swendsen R H 1989 *Phys. Rev. Lett.* **63** 1195–1198
[16] Elcock A H 2006 *PLoS Comp. Biol.* **2** 0824–0841