EFFECTS OF RIBOSOMAL EXIT TUNNEL ON PROTEIN'S COTRANSLATIONAL FOLDING

BUI PHUONG THUY

Nam Dinh University of Technology Education,
Phu Nghia, Loc Ha, Nam Dinh, Vietnam

TRINH XUAN HOANG

Institute of Physics, Vietnam Academy of Science and Technology,
10 Dao Tan, Ba Dinh, Hanoi, Vietnam

Email: hoang@iop.vast.ac.vn

Abstract. *In vivo, folding of many proteins occurs during their synthesis in the ribosome and continues after they have escaped from the ribosomal exit tunnel. In this research, we investigate the confinement effects of the ribosome on the cotranslational folding of three proteins, of PDB codes 1PGA, 1CRN and 2RJX, by using a coarse-grained model and molecular dynamics simulation. The exit tunnel is modeled as a hollow cylinder attached to a flat wall, whereas a Go-like model is adopted for the proteins. Our results show that the exit tunnel has a strong effect on the folding mechanism by setting an order by which the secondary and tertiary structures are formed. For protein 1PGA, the folding follows two different folding routes. The presence of the tunnel also improves the foldability of protein.*

I. INTRODUCTION

Significant insights on the protein folding problem [1] have been obtained in recent years through protein engineering experiments [2] and corroborated theoretical frameworks supported by computer simulations [3–5]. Yet, our current understanding of protein folding is mostly based on *in vitro* experiments [6]. Folding *in vivo* might be quite different from that in the test tube. For examples, folding *in vivo* could be affected by macromolecular crowding [7] in the intracellular environment. Folding of newborn or nascent proteins is under influence of the ribosome and other molecular machineries such as chaperones [8].

Accumulated experimental evidences have confirmed that protein folds cotranslationally, i.e. nascent polypeptide chain starts to fold during their synthesis in the ribosome [9–11]. Furthermore, the ribosome has been shown to promote efficient folding of nascent proteins [12]. Interestingly, the non-equilibrium character of the translational process also favors the formation of helices near the C-terminus [13]. Of particular importance for nascent protein folding is the ribosomal exit tunnel [14], a narrow channel in the ribosome large subunit, through which nascent proteins are secreted to cellular environment. The exit tunnel is about 80-100$\mathring{A}$ in length and 10-20$\mathring{A}$ in diameter [15]. While the tunnel

is too narrow for a full tertiary structure of protein, small modular structures such as α-helix and β-hairpin are able form inside the tunnel [16].

In this study we investigate the effect of the exit tunnel on folding of nascent proteins. In particular, we focus on how folding occurs inside the tunnel and whether the presence of the tunnel helps protein to fold more correctly. Such informations will be obtained by coarse-grained molecular dynamics simulations. We will study cotranslational folding for three short proteins with PDB codes 1PGA, 1CRN and 2JRX. The basic difference between these proteins resides on their native structures: 2JRX has only α-helices, whereas 1PGA and 1CRN have both α-helices and β-sheets.

II. MODEL AND METHODS

Go-like model with repulsive potential from ribosomal exit tunnel

Go-like models [17] have been very successful in studies of protein folding mechanism [18, 19]. The main idea of these models is to assign attractive interactions only to the native contacts. In this research we consider a version of off-lattice Go-like model which essentially follows Ref. [20]. The exit tunnel is modeled as a hollow cylinder of length of 100Åand of uniform diameter of 20Å(Fig. 1). The nascent protein is supposed to grow from the left end of the tunnel and escapes out of the tunnel through the right end. The latter is attached to a flat wall mimicking the ribosome body. The protein translation direction is from the N-terminus to the C-terminus.

A protein conformation is given by the positions of the C_α atoms. The energy of a conformation Γ in presence of the tunnel is given by:

$$
\begin{aligned}
E(\Gamma) &= \underbrace{\sum_{bonds} K_b(b-b_0)^2}_{\text{bondstretching}} + \underbrace{\sum_{angles} K_\theta(\theta-\theta_0)^2}_{\text{angle bending}} + \underbrace{\sum_{dihedral} K_\phi^{(n)}(1+cos(n(\phi-\phi_0)))}_{\text{dihedral angle}} \\
&+ \underbrace{\sum_{j>i+3} 4\epsilon\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12}-\left(\frac{\sigma_{ij}}{r_{ij}}\right)^6\right]c_{ij}}_{\text{Lennard-Jones}} + \underbrace{\sum_{j>i+3}\epsilon\left(\frac{\sigma_0}{r_{ij}}\right)^{12}(1-c_{ij})}_{repulsion} \\
&+ \underbrace{\sum_i\left\{4\epsilon\left[\left(\frac{\sigma_1}{\rho_i}\right)^{12}-\left(\frac{\sigma_1}{\rho_i}\right)^6\right]+\epsilon\right\}}_{\text{repulsion of tunnel wall and attached wall}},
\end{aligned}
\tag{1}
$$

where the subscript 0 is a notation for native state; b is the distance between two consecutive residues along the chain; θ is the bond angle formed by three subsequent residues and ϕ is the dihedral angle defined by four subsequent residues along the chain. r_{ij} is the distance between i and j residues. ρ_i is the distance from residue i either to the tunnel wall or the attached wall depending whether it is inside the tunnel or not. $c_{ij}=1$ if i and j are a pair of residues which forms a contact in the native state (a native contact) and $c_{ij}=0$ if they do not. A native contact is defined if the distance between the residues in the native state is smaller than 7.5$\mathring{A}$ and the residues are separated by at least 3 other residues along
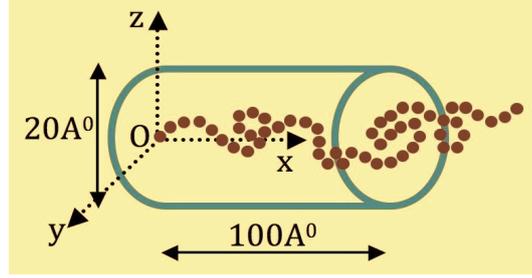
Fig. 1. Sketch of confinement geometry of ribosomal exit tunnel modeled as an uniform cylinder and protein cotranslational folding within the tunnel.

the chain. The chosen coefficients are $K_b = 100\epsilon\mathring{A}^{-2}, K_\theta = 20\epsilon(rad)^{-2}, K_\phi^{(1)} = -\epsilon, K_\phi^{(3)} = -0.5\epsilon, \sigma_0 = \sigma_1 = 5\mathring{A}$ and $\sigma_{ij} = 2^{-1/6}(r_{ij})_0$. The last term in Eq. (1) corresponds to the repulsion from the wall of the cylinder.

Molecular Dynamics method

The time evolution of folding for each protein was simulated by using Molecular Dynamics (MD) method with Langevin equation:

$$m\ddot{\vec{r}} = \vec{f} - \gamma\dot{\vec{r}} + \vec{\Gamma}, \tag{2}$$

where $\vec{f}$ is the force due to molecular interaction potentials in the model, γ is the friction constant. $\vec{\Gamma}$ is a random force in form of a white noise of zero mean and time correlation $\langle\Gamma_\alpha(t)\Gamma_\beta(t + t')\rangle = 2\gamma mk_BT\delta(t')\delta_{\alpha\beta}$, where m is the mass of the amino acid, k_B is the Boltzmann constant, T is the temperature, α and β denote any of the x, y, or z components. If $\gamma = 0$, the Langevin equation turns into the deterministic Newton equation. Distance is measured in units of $\mathring{A}$, and the unit of time is $\tau = \mathring{A}\sqrt{m/\epsilon}$. Temperature is measured in units of ϵ/k_B. The friction coefficient is chosen to be $\gamma = 0.5 \ m\tau^{-1}$. For simplicity, we use a reduced unit system in which $m = 1$, $\epsilon = 1$, and $k_B = 1$. Eq. (2) is integrated using the Verlet algorithm with time step $\Delta t = 0.01\tau$.

III. RESULTS AND DISCUSSION

Effect of ribosomal exit tunnel on the formation of secondary and tertiary structures

The first issue we investigate is the effect of ribosomal exit tunnel on the formation of secondary structures for 1PGA protein. Our simulations show that the secondary structures of 1PGA including the α-helix and the β-hairpin are able to formed within the tunnel. The simulation temperature was chosen to be $T = 0.5\epsilon/k_B$, which is below the folding transition temperature (about $0.9\epsilon/k_B$ in our model). Fig. 2 shows a two-dimensional histogram of conformation as function of the number of native contacts, (Q), and the number of amino acids emerged from the exit port, (N_{out}), in 100 folding trajectories. It clearly shows that the formation of the tertiary structures follows two different folding routes.
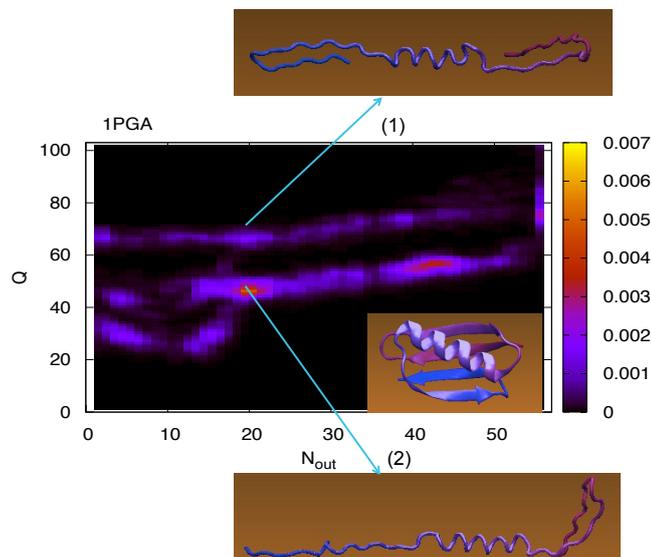
Fig. 2. Histogram of 100 folding trajectories as function of number of native contacts (Q) and number of amino acids emerged from exit port (N_{out}) for 1PGA protein.

When the first amino acid emerges from the exit tunnel, the nascent chain is partially folded with $25 - 65$ native contacts formed. These contacts are due to the presence of the secondary structures, the α-helix and the β-hairpins. It can be seen in Fig. 2 that in both folding routes the first β-hairpin (associated with the N-terminus) and the α-helix are formed, but the second β-hairpin (associated with the C-terminus) is formed only in the upper folding route (with higher Q). For the latter, the number of contacts gradually increases, whereas for the lower folding route (with lower Q), there is a jump in Q when the last amino acid is released from the exit port.

Previous experimental studies have showed that the nascent polypeptide chain can form the α-helix and the β-hairpin within the exit tunnel [16]. It has been also shown that near the exit port protein can also form the β-sheets and a partial tertiary structure corresponding to a sub-domain folding [21,22]. Our above simulation results are consistent with these experimental findings. A new insight from our study is that the folding of 1PGA in the presence of the exit tunnel proceeds following two different routes.

We continue to study two other proteins: 1CRN and 2RJX. The results are shown in Fig. 3 and Fig. 4. For 1CRN, folding clearly follows a single route (Fig. 3) and proceeds in three distinct steps. For $N_{out} < 30$ only α-helices are formed, and no tertiary structure is found. For N_{out} between 35 and 45, there is a jump in Q which indicates that an essential portion of tertiary structure has been established. The last jump in Q happens at $N_{out} \approx 46$, i.e. equal to the full length of the protein ($N = 46$). This jump corresponds to a release of the C-terminal tail of the protein from the exit port and the formation of the last tertiary contacts with the rest of the protein.
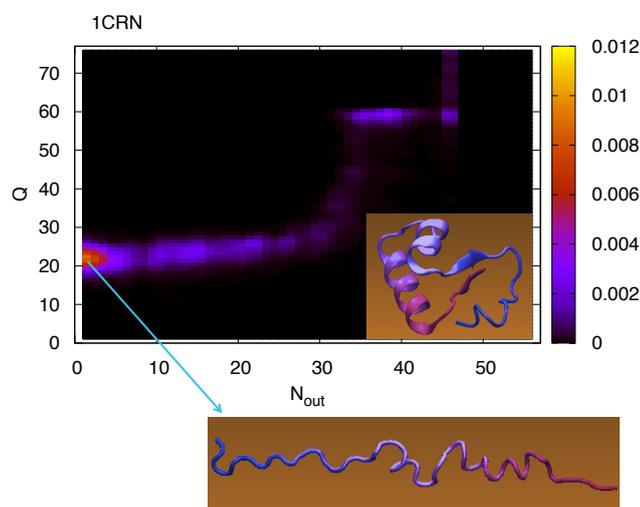
Fig. 3. Histogram of 100 folding trajectories as function of number of native contacts (Q) and number of amino acids emerged from exit port (N_{out}) for 1CRN protein.
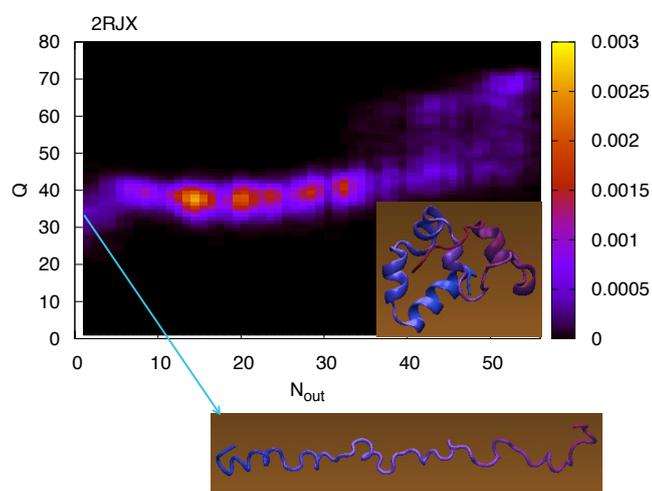

A

Fig. 4. Histogram of 100 folding trajectories as function of number of native contacts (Q) and number of amino acids emerged from exit port (N_{out}) for 2RJX protein.

The folding of 2RJX also follows a single route for $N_{out} < 35$, when only the α-helices are able to form. The protein starts to form tertiary contacts when more than 35 residues are released from the exit port. Fig. 4 shows that for $N_{out} > 35$ the number of

native contacts, Q, varies to a large extent, which indicates that there is no clear ordering in tertiary contact formation for this protein.

Our results show different folding scenarios for three different proteins. The only common thing is that the presence of exit tunnel helps the establishment of the secondary structures before the formation of tertiary structure. The formation of β-hairpin in side the tunnel is less probable than that of the α-helix which leads to two different folding routes in the case of 1PGA. Depending on the protein, the tertiary structure can form gradually (1PGA, upper folding route), in steps (1CRN) or without ordering (2RJX).

Effect of ribosomal exit tunnel on the successful folding probability

We study also the effect of ribosomal exit tunnel of the folding efficiency of 1PGA protein. We compare the probability of successful folding, P_{fold} for three cases: a) refolding without translation, b) cotranslational folding without the exit tunnel, and c) cotranslational folding with the exit tunnel. The results are shown in Fig. 5 where P_{fold} is plotted against time. It is interesting to note that the presence of the exit tunnel leads to an increase of about 4% in P_{fold} at large t. Thus, the exit tunnel has a favorable effect on folding efficiency. This finding is consistent with recent experimental result [12] on the folding of T4 lysozyme with and without the ribosome.
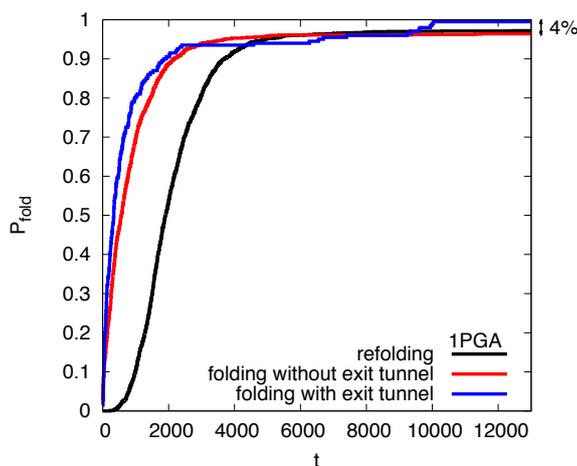

Fig. 5. The successful folding probability, P_{fold}, of 1PGA protein in three cases: a) refolding (without translation), b) cotranslational folding without the exit tunnel and c) cotranslational folding with exit tunnel. $t = 0$ corresponds to the moment when the protein is fully unfolded (a), fully translated (b), and fully translated and released from the exit port (c). The probability P_{fold} for each case was calculated based on 500 independent folding trajectories.

In summary, we have shown that the folding mechanism of nascent proteins is significantly influenced by the ribosomal exit tunnel. The latter defines an order by which the secondary and tertiary structure are formed. It also improves the folding efficiency

of nascent proteins. The effect of ribosomal exit tunnel together with other cell mechanisms (such as the works of chaperones) help nascent proteins to avoid misfolding and aggregation.

REFERENCES

[1] C. B. Anfinsen, *Science* **181** (1973) 223-230.

[2] A. Matouschek, J. T. Kellis, L. Serrano and A. R. Fersht, *Nature* **340** (1989) 122-126.

[3] K. A. Dill, H. S. Chan, *Nat. Struct. Biol.* **4** (1997) 10-19.

[4] J. N. Onuchic, P. G. Wolynes, *Curr. Opin. Struct. Biol.* **14** (2004) 70-75.

[5] T. X. Hoang, A. Trovato, F. Seno, J. R. Banavar, and A. Maritan, *Proc. Natl. Acad. Sci. USA.* **101** (2004) 7960-7964.

[6] R. D. Schaeffer, A. Fersht, V. Daggett, *Curr. Opin. Struct. Biol.* **18** (2008) 4-9.

[7] M. S. Cheung, D. Klimov, and D. Thirmalai, *Proc. Natl. Acad. Sci. USA* **102** (2005) 4753-4758.

[8] F. Ulrich Hartl, and M. Hayer-Hartl, *Nat. Struct. Mol. Biol.* **16** (2009) 574-581.

[9] A. A. Komar, *Trends Biochem. Sci.* **34** (2009) 16-24.

[10] G. Kramer, D. Boehringer, N. Ban, B. Bukau, *Nat. Struct. Mol. Biol.* **16** (2009) 589-597.

[11] L. D. Cabrita, C. M. Dobson, J. Christodoulou, *Curr. Opin. Struct. Biol.* **20** (2010) 33-45.

[12] C. M. Kaiser, D. H. Goldman, J. D. Chodera, I. Tinoco Jr., C. Bustamante, *Science* **334** (2011) 1723.

[13] D. Marenduzzo, T. X. Hoang, F. Seno, M. Vendruscolo and A. Maritan, *Phys. Rev. Lett.* **95** (2005) 098103.

[14] D. N. Wilson, R. Beckmann, *Curr. Opin. Struct. Biol.* **21** (2011) 274-282.

[15] N. Ban, P. Nissen, J. Hansen, P. B. Moore, T. A. Steitz, *Science* **289** (2000) 905-920.

[16] A. Kosolapov, C. Deutsch, *Nat. Struct. Mol. Biol.* **16** (2009) 405-411.

[17] N. Go and H. Abe, *Biopolymers*, **20** (1981) 991.

[18] T. X. Hoang and M. Cieplak, *J. Chem. Phys.*, **113** (2000) 8319.

[19] D. Baker, *Nature* (London), **405** (2002) 39.

[20] C. Clementi, H. Nymeyer, J. N. Onuchic, *J. Mol. Biol.* **298** (2000) 937-953.

[21] R.J. Gilbert, P. Fucini, S. Connell, S. D. Fuller, K. H. Nierhaus, C. V. Robinson, C. M. Dobson, D. I. Stuart, *Mol. Cell* **14** (2004) 57-66.

[22] S. Fulle, H. Gohlke, *J. Mol. Biol.* **387** (2009) 502-517.