Journal of Molecular Graphics and Modelling 54 (2014) 141–147

Contents lists available at ScienceDirect

Journal of Molecular Graphics and Modelling

journal homepage: www.elsevier.com/locate/JMGM

Evaluating membrane affinity by integrating protein orientations

Fangqiang Zhu [a,*], Matthias Clauss [b]

[a] Department of Physics, Indiana University – Purdue University Indianapolis, United States
[b] Department of Cellular and Integrative Physiology, Indiana University School of Medicine, Indianapolis, IN, United States

ARTICLE INFO

Article history:
Accepted 15 October 2014
Available online 29 October 2014

Keywords:
Protein orientation
Free energy
Coarse-grained modeling
Non-classical release
EMAPII
FGF1

ABSTRACT

Energetic interactions of a protein with lipid bilayers determine its propensity to reside in the membrane. Here we seek to evaluate the membrane interactions for EMAPII, a protein found to be released from the cell by unknown mechanisms, as well as several other proteins. Using a knowledge-based coarse-grained membrane potential, we calculate the free energy profiles for these proteins by integrating out the orientation degrees of freedom. Due to the invariance of energy under in-plane rotations about the membrane normal, the orientation space can be reduced to two dimensions and mapped onto the surface of a unit sphere, thus making visualization, sampling and integration more convenient. The integrated free energy profiles determine the relative probabilities along the membrane normal for the proteins regardless of their orientations, and display distinctive characteristics for membrane proteins and water-soluble proteins. The membrane interactions for EMAPII exhibit typical features of a water-soluble protein, with a high energetic barrier to enter or cross the membrane. Our results thus suggest that similar to the non-classical export of FGF1, the release of EMAPII would involve more complicated mechanisms than simple passive diffusion across the membrane.

1. Introduction

Extracellular proteins are either derived from transmembrane proteins by proteolytic cleavage or secreted as soluble non-membrane-binding proteins. Commonly the secreted proteins require a leader sequence and travel through the Golgi. Recently, however, a new class of secreted proteins from intracellular cytosolic origin have been identified, including fibroblast growth factor-1 (FGF1), interleukin-1alpha (IL-1α) and endothelial monocyte-activating protein II (EMAPII), which all lack a leader sequence and the release is not through the classical exocytosis pathway involving the Golgi apparatus [1–3]. In the cytosol, EMAPII is a non-enzymatic part of a multi-enzyme aminoacyl-tRNA synthetase (ARS) complex [4] and as such also known as p43 protein and ARS-interacting multifunctional protein 1 (AIMP1). EMAPII can be released by stress including hypoxia, cigarette smoke, HIV envelope protein gp120 and apoptosis [3–6]. The release mechanism of EMAPII is currently unclear, and may involve the translocation across membranes through hydrophobic interactions with the lipid bilayer. Alternatively, similar non-classical released proteins FGF1 and IL-1α involve clustering with other proteins leading to membrane destabilization [7,8].

Given that EMAPII was originally considered a water-soluble protein based on its cellular location and crystallography [9], its ability to cross the cell membrane and reach the extracellular space, as mentioned above, appears highly unusual. It is thus helpful to explicitly assess the interaction between the protein and the lipid membrane. Hypothetically, if the interaction energy of a protein with the membrane is similar in magnitude to that with the bulk water, the protein would experience a flat energetic profile across the lipid bilayer and would be able to penetrate the membrane through simple passive diffusion without encountering significant barriers. Here we computationally evaluate the protein-membrane interactions to evaluate the feasibility of this hypothetical release mechanism for EMAPII.

The interaction energy between a protein and the lipid membrane plays a major role in determining whether the protein is membrane-bound. Naturally, membrane proteins and water-soluble proteins have favorable and unfavorable interactions with the lipid bilayer, respectively. Computationally, first-principle calculations of protein-membrane interaction energy at all-atom level are expensive and currently still challenging [10–12]. Alternatively, a number of approximate membrane potentials have been developed over the past decade, enabling convenient and much faster estimations of the membrane interaction energy. Just to name a few, Tusnády et al. proposed the TMDET algorithm [13], based on the hydrophobicity scale and some structural factors, to provide a quantitative measure for the protein-membrane interaction. This

* Corresponding author.
E-mail address: fzhu0@iupui.edu (F. Zhu).

http://dx.doi.org/10.1016/j.jmgm.2014.10.009
1093-3263/

algorithm was used to build the PDBTM database [14] for all membrane proteins with known crystal structures. Ulmschneider et al. developed an implicit membrane potential based on a comprehensive survey of high-resolution crystal structures of membrane proteins [15]. Lomize et al. proposed a physics-based protocol [16] to compute the transfer energies of proteins from water to the lipid bilayer, and applied this method to analyze the membrane proteins in the OPM database [17]. DeGrado and coworkers developed knowledge-based potentials [18,19] for determining the position and orientation of proteins relative to the membrane. Recently, Nugent and Jones proposed another knowledge-based potential and search algorithms for the positioning and refinement of membrane proteins [20].

Whereas the membrane potentials above were mainly designed to find the optimal position and orientation of membrane proteins in the lipid bilayer, here we aim to evaluate the possibility for EMAPII to spontaneously enter the membrane. In general, the TMDET algorithm [13] provides a Q-value that could distinguish membrane and water-soluble proteins. In this study, alternatively, we seek to calculate a free-energy profile for the protein-membrane interactions at different vertical positions, thus offering an intuitive measure for the membrane affinity of the protein.

As in many coarse-grained calculations [15,21,22], here a folded protein domain is represented as a rigid body without internal conformational variation. In such cases the interaction energy depends on the spatial arrangement of the protein, i.e., its orientation and position relative to the membrane. Because the in-plane translation and rotation of the protein do not change the membrane energy, such energy only depends on three degrees of freedom, thus making a grid-based exhaustive search computationally feasible and affordable, as implemented in many studies using various membrane potentials [13–21]. Instead of searching for the optimal spatial arrangement as in those studies, here we integrate the contributions from the entire set of orientations to obtain a one-dimensional (1D) free energy profile that directly determines the equilibrium probabilities of the protein along the membrane normal and thus provides a quantitative measure for its propensity to access the membrane. We also show that the orientations can be properly mapped onto the surface of a sphere [13], which offers an intuitive and convenient way to visualize and sample the orientation space.

Among the available membrane potentials [13–20] discussed earlier, in this study we adopt the knowledge-based energy function developed by Ulmschneider et al. [15] due to its conceptual simplicity. This implicit potential has been adopted in a number of studies, e.g., to identify the orientation of transmembrane helices [21] and to simulate protein-protein binding near the membrane [22]. Here we use this energy model along with our integration method to evaluate the membrane interaction for EMAPII and several other proteins. As mentioned earlier, our calculations aim to assess the feasibility of a simple release mechanism, in which EMAPII enters and crosses the membrane through passive diffusion without undergoing large-scale conformational changes.

2. Methods

As mentioned earlier, we adopt the implicit membrane potential developed by Ulmschneider et al. [15] for calculating the interaction energy. In this section, we first introduce this energy function, and then define a 1D free energy that integrates out the orientation degrees of freedom. To numerically calculate this free energy, we describe our method to properly sample and integrate the orientation space, which utilizes the invariance of the potential energy under in-plane rotations to reduce the dimensionality of the sampling. Finally, we provide computational details for the proteins studied here. Throughout this article the membrane normal is denoted as the z-axis, with the membrane midplane at $z = 0$ and the extra- and intracellular spaces at the $+z$ and $-z$ sides, respectively.

2.1. Energy function

Here we adopt the residue-based implicit membrane potential developed in Ref. [15]. In this empirical representation of the membrane interaction energy, a potential of mean force along the membrane normal is provided for each type of amino acid [15]. The coordinates of the C_α atom represent the position of each amino acid. The interaction energy of a single residue with the membrane is a function of its z coordinate, and is given by [15]

$$e(z) = -k_B T \ln\{a_0 + a_1 \exp[-a_2(z-a_3)^2] + a_4 \exp[-a_5(z-a_6)^2]\}, \quad (1)$$

in which the parameters $a_0, \ldots, a_6$ depend on the type of the amino acid, and were taken from Ref. [15]. Given the coordinates of an entire protein, the total membrane energy is simply the addition of the individual energies from each residue:

$$E = \sum_i e_i(z_i), \quad (2)$$

where the summation is over all residues in the protein.

2.2. Free energy

In this study the protein structure is treated as a rigid body without internal conformational change. Consequently, the protein coordinates are completely determined by the three translational and three rotational degrees of freedom. Moreover, translations in the xy plane do not affect the interaction energy with the membrane. The energy $E(z^c, \Omega)$ thus depends on the center z^c of the C_α atoms and the orientation Ω of the protein. We may further integrate out the rotational degrees of freedom (Ω) and obtain a free energy as a function of z^c alone:

$$G(z^c) = -k_B T \ln \frac{\int d\Omega \exp[-E(z^c, \Omega)/k_B T]}{V_\Omega}, \quad (3)$$

where $V_\Omega \equiv \int d\Omega$ is the volume of the orientation space. Here the integrant is the Boltzmann factor which is proportional to the probability density. The free energy $G(z^c)$ thus determines the equilibrium distribution of the protein at different positions of the membrane normal, regardless of its orientation.

2.3. Integration of orientation space

To evaluate the integral in Eq. (3), one needs to properly sample the orientation space. To mathematically represent orientations, typically a reference state of the protein is specified, and any orientation can then be described by the rotation required to bring the reference to the current state. Any three-dimensional (3D) rigid-body rotation can be described by a unit quaternion [23], consisting of four elements denoted here as $Q \equiv [a, r_x, r_y, r_z]$ with $a^2 + r_x^2 + r_y^2 + r_z^2 = 1$. The quaternion represents a rotation about the axis $\vec{r} \equiv (r_x, r_y, r_z)$ by an amount θ, with $\cos(\theta/2) = a$. All such unit quaternions lie on the surface of a four-dimensional (4D) hypersphere. In fact this hypersphere surface is a proper representation of the entire rotational space, except for a redundancy due to the fact that Q and $-Q$ correspond to the same rotation [23]. We thus only need to sample one half of the surface representing, e.g., quaternions with $a \geq 0$, which then have a one-to-one correspondence to the 3D rotations.