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.

© 2014 Elsevier Inc. All rights reserved.

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/© 2014 Elsevier Inc. All rights reserved.

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.

Here we parameterize the unit quaternions as

$$a = \cos\alpha\ \cos\gamma, \tag{4a}$$

$$r_x = \sin\alpha\ \cos(\beta+\gamma), \tag{4b}$$

$$r_y = \sin\alpha\ \sin(\beta+\gamma), \tag{4c}$$

$$r_z = \cos\alpha\ \sin\gamma, \tag{4d}$$

in which $\alpha \in [0, \pi/2]$, $\beta \in [0, 2\pi]$, and $\gamma \in [-\pi/2, \pi/2]$. Any unit quaternion with $a \geq 0$ can be uniquely represented by these three parameters as $Q(\alpha, \beta, \gamma)$, with the corresponding energy denoted by $E(z^c, Q(\alpha, \beta, \gamma))$. The Jacobian for the coordinate transformation in Eq. (4) is $\sin\alpha\cos\alpha$. Eq. (3) can therefore be evaluated as

$$G(z^c) = -k_BT \ln \frac{\int_0^{\pi/2} d\alpha \int_0^{2\pi} d\beta \int_{-\pi/2}^{\pi/2} d\gamma \exp[-E(z^c, Q(\alpha,\beta,\gamma))/k_BT] \sin\alpha\cos\alpha}{V_\Omega}. \tag{5}$$

With this parameterization, the volume of the orientation space is similarly given by

$$V_\Omega = \int_0^{\pi/2} d\alpha \int_0^{2\pi} d\beta \int_{-\pi/2}^{\pi/2} d\gamma \sin\alpha\cos\alpha = \pi^2, \tag{6}$$

which is indeed one half of the surface area of a unit 4D hypersphere.

According to the quaternion multiplication rule [23], $Q(\alpha, \beta, \gamma)$ defined in Eq. (4) can be decomposed into two factors:

$$Q(\alpha, \beta, \gamma) = Q_z(\gamma)Q_{xy}(\alpha, \beta), \tag{7}$$

in which

$$Q_{xy}(\alpha, \beta) = [\cos\alpha,\ \sin\alpha\cos\beta,\ \sin\alpha\sin\beta, 0], \tag{8}$$

$$Q_z(\gamma) = [\cos\gamma, 0, 0, \sin\gamma]. \tag{9}$$

Eq. (7) indicates that the rotation described by $Q(\alpha, \beta, \gamma)$ is equivalent to two sequential rotations: a rotation $Q_{xy}(\alpha, \beta)$ about an axis in the xy plane followed by a rotation $Q_z(\gamma)$ about the z-axis. Furthermore, the second rotation $Q_z(\gamma)$ does not change the z-coordinates and hence the energy. The energy can therefore be expressed as $E(z^c, Q(\alpha, \beta, \gamma)) = E(z^c, Q_{xy}(\alpha, \beta))$, independent of the parameter γ. The triple integral in Eq. (5) is thus reduced to a double integral:

$$G(z^c) = -k_BT \ln \frac{\int_0^{\pi/2} d\alpha \int_0^{2\pi} d\beta \exp[-E(z^c, Q_{xy}(\alpha,\beta))/k_BT] \sin\alpha\cos\alpha}{\pi}. \tag{10}$$

The free energy $G(z^c)$ can be obtained by numerically calculating the double integral in Eq. (10). In our calculations here we made an additional variable transformation $t \equiv \cos(2\alpha)$, with $t \in [-1, 1]$. With this transformation, we built a two-dimensional grid by evenly dividing the parameter space of t and β and evaluated the energy and the Boltzmann factor e^{-E/k_BT} at each grid point. The average Boltzmann factor over all the grid points then determines $G(z^c)$. We note that this is not the most efficient method to calculate the integral above. Nonetheless, because the energy evaluation is computationally cheap here, we can afford a large number of grid points and obtain accurate numerical solutions (see Section 3). If the computation speed becomes an issue, however, advanced integration scheme with better efficiency can be adopted [24].

$Q_{xy}(\alpha, \beta)$ (see Eq. (8)) represents a rotation of 2α about an axis (with an angle β from the x-axis) in the xy plane. If we define $\tau \equiv 2\alpha$ and $\rho \equiv \beta$ with $\tau \in [0, \pi]$ and $\rho \in [0, 2\pi]$, the parameters τ and ρ are actually the tilt angle and the rotational angle, respectively, that are commonly used to denote the orientations of membrane proteins [16,25]. Furthermore, τ and ρ can also be taken as the polar and azimuthal angles in a spherical coordinate system, such that each rotation $Q_{xy}(\alpha, \beta)$ is mapped to a unique point (specified by coordinates τ and ρ) on the surface of a unit 3D sphere. The north pole of the sphere (with $\tau = 0$) represents the identity rotation, thus corresponding to the orientation of the reference structure. It can be shown that the sphere surface is a proper representation of the orientations, with the area of any region on the surface proportional to the volume of the corresponding orientation space. This property can also be easily seen if one fixes the orientation of the protein and instead rotates the membrane normal direction, as described in Ref. [13]. In light of this, Eq. (10) is essentially an integration of the Boltzmann factor e^{-E/k_BT} on the sphere.

2.4. Computational details

We calculated the free energy profile $G(z^c)$ for twelve proteins of similar size, as listed below. (1) M2 proton channel of influenza A virus, with the reference conformation (152 residues) of the homotetramer taken from the first structure in PDB 2RLF [26]. (2) Outer membrane protein A (OMPA), with the reference conformation (176 residues) taken from the first structure in PDB 1G90 [27]. (3) Adenylate kinase (AdK), with the reference conformations (214 residues) taken from chain A in PDBs 1AKE [28] and 4AKE [29] for the closed and open structures, respectively. (4) Ubiquitin (Ubq), with the reference conformation (76 residues) taken from PDB 1UBQ [30]. (5) Fibroblast growth factor-1 (FGF1), with the reference conformation (127 residues) taken from chain A in PDB 1AFC [31]. (6) Endothelial monocyte-activating protein II (EMAPII) in the mature form, with the reference conformation (166 residues) taken from PDB 1E7Z after removing the Leu-Glu-His_6 tag [9] in the C-terminal. (7) Human growth hormone (hGH), with the reference conformation (186 residues) taken from PDB 1HGU [32]. (8) Myoglobin (Myo), with the reference conformation (154 residues) taken from PDB 2MGE [33]. (9) Catalytic fragment of human fibroblast collagenase (MMP-1), with the reference conformation (170 residues) taken from the first structure in PDB 1AYK [34]. (10) Superoxide dismutase (SOD), with the reference monomer conformation (151 residues) taken from chain O in PDB 2SOD [35]. (11) Amicyanin (AC), with the reference conformation (105 residues) taken from PDB 1AAJ [36]. (12) Trypsin inhibitor DE-3, with the reference conformation (166 residues) taken from PDB 1TIE [37].

As discussed earlier, the energy of the protein is a function of one translational (z^c) and two rotational (t and β) degrees of freedom. In this study z^c is sampled from −70 Å to 70 Å. We evenly divided the sampled range in each degree of freedom into 1000 grid points, thus resulting in a total of 10^9 grid points in the combined three-dimensional parameter space. For each grid point we accordingly translated and rotated the reference structure and then evaluated the energy. The computations were implemented in C++, and took 5–8 h for the 10^9 energy evaluations on a single AMD Opteron processor at 2.6 GHz.

3. Results

As described in Methods, we numerically integrated the Boltzmann factor in the orientation space to obtain the free energy $G(z^c)$. Although the description of the protein orientation involves the use of a reference structure, the results of the integration should not depend on the particular orientation of the reference. We performed additional calculations to verify this property and to validate our integration scheme. Specifically, we generated some new reference orientations by random rigid-body rotations on the original reference, and repeated the integrations using the new references. These calculations show that all such integrations indeed give essentially the same results, with the differences in the free energy well below 10^{-4} k_BT. Therefore, as expected, our calculated

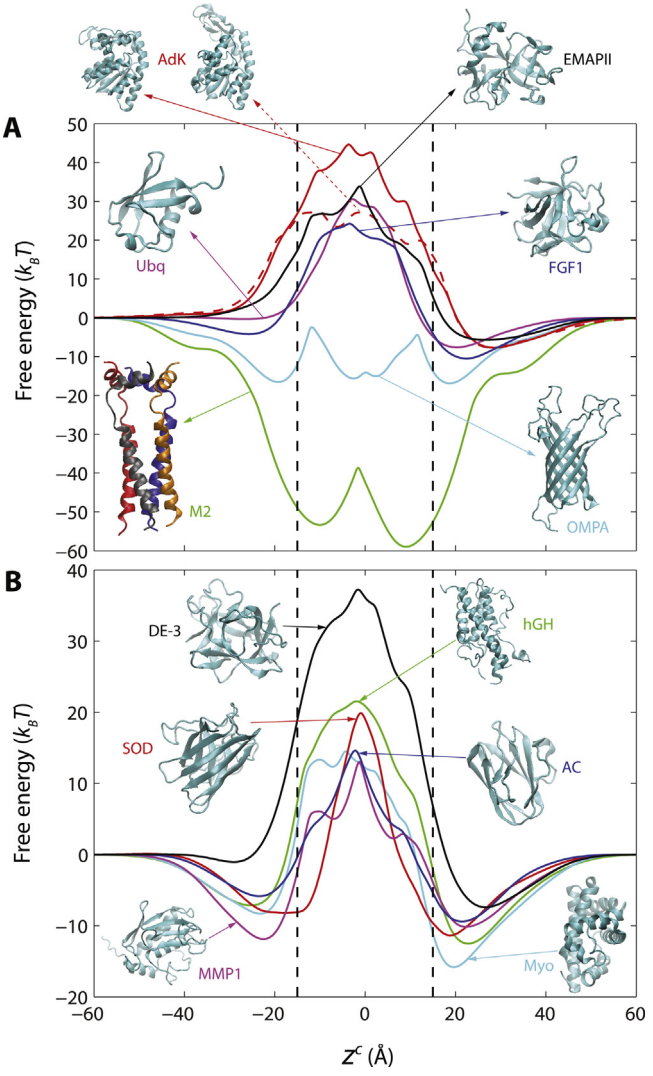

Fig. 1. Calculated free energy profiles for twelve proteins ((A) M2, OMPA, Ubq, AdK, FGF1, EMAPII; (B) hGH, Myo, MMP-1, SOD, AC, DE-3) after integrating out the orientation degrees of freedom. The free energy is a function of z^c, the z coordinate of the center of the C_α atoms. Results for the closed and open conformations of AdK are shown in *solid* and *dashed* curves, respectively. The two vertical *dashed* lines indicate the 30 Å-thick hydrophobic region of the membrane. The molecular images were rendered in VMD [45].

free energy depends only on the intrinsic conformation of the reference structure, but not on its specific orientation state.

Due to the coarse-grained nature of the membrane potential adopted here, our calculated free energy is not expected to be sufficiently accurate for determining the exact binding affinity of the protein to the membrane. Nonetheless, the estimated free energy profiles here could distinguish membrane-bound proteins from those not residing in the membrane. We therefore carried out the calculations for multiple proteins in each category, as shown in Fig. 1. Because the membrane potential [15] was developed and previously tested on a group of integral transmembrane proteins, its applicability on water-soluble proteins is the focus of our test here. In particular, we included a number of known water-soluble proteins in the calculations to see whether the potential could correctly predict an unfavorable interaction of these proteins with the membrane. Our chosen water-soluble proteins cover the three major structural classes, i.e., the α domain, the α/β domain, and the β domain. We also note that this potential is somewhat asymmetric with respect to the center of the membrane, reflecting different properties of the two leaflets of the lipid bilayer. This asymmetry is thus also present in the integrated free energy profiles here.

Among the studied proteins, M2 (in the homotetramer form) and OMPA are known to form channels in the membrane. Indeed, their free energy profiles (Fig. 1A) exhibit attractive wells in the interior of the membrane with lower free energy than in the bulk region. The profiles for both proteins exhibit multiple minima, which correspond to different orientations, as will be further discussed later. Overall, the low energies for M2 and OMPA inside the membrane are consistent with their identities as membrane proteins.

For water-soluble proteins (AdK, Ubq, hGH, Myo, MMP-1, SOD, AC, DE-3), in contrast, the free energy profiles (Fig. 1) exhibit repulsive barriers in the membrane region with higher free energy than in the bulk. Because AdK may adopt multiple conformations [28,29], we performed the calculations for both the closed and the open structures of the protein. Our results (Fig. 1A) show major energetic barriers for both conformations, with the barrier for the open conformation somewhat lower than that of the closed conformation.

Overall, our calculated free energy profiles correctly describe membrane proteins with favorable interactions and higher probabilities in the membrane, and water-soluble proteins with unfavorable interactions and low probabilities in the membrane. By comparing the free energy of EMAPII (Fig. 1A) to other proteins, it is clear that EMAPII is much more similar to a typical water-soluble protein than to a membrane protein. The free energy of EMAPII features a high barrier in the interior of the membrane, indicating that it is energetically highly unfavorable for this protein to enter the membrane.

As discussed in Methods, with the degeneracy of in-plane rotations, the entire orientation space can be mapped onto the surface of a 3D sphere [13]. Here we use this sphere to visualize the energy variation in the orientation space, using the six proteins in Fig. 1A as examples. Fig. 2 shows the orientational dependence for these proteins, in which the lowest energy along the translational degree of freedom (z^c) is displayed for each orientation. M2 is a homotetramer with an approximate four-fold symmetry in the crystal structure [26], which gives rise to an approximate rotational symmetry in its orientation sphere (Fig. 2A). Two major low-energy regions on the sphere are located near the north and south poles, respectively. The spatial arrangements represented by these two regions are roughly inversions of each other, with the C-terminal of the protein on the extra- and intracellular sides, respectively. In these inverted orientations, the center of the entire protein has some offset from the midpoint ($z = 0$) of the membrane, thus giving rise to two separate minima in the 1D free energy profile (Fig. 1A). Notably, the spatial arrangements with the lowest energies are not at the north or south pole of the sphere, but instead correspond to tilted orientations as shown in Fig. 2A. The angles between the symmetry axis and the membrane normal are 39° and 33° for the two orientations. These values are in good agreement with previous calculations on a single M2 helix [21], and with NMR studies [38]. For OMPA (Fig. 2D), in addition to the transmembrane orientations as in M2, there also exist low-energy configurations in which the protein was bound to either surface of the membrane. These orientations correspond to the multiple minima in the free energy profile of OMPA (Fig. 1A).

In the orientation spheres of FGF1, EMAPII, Ubq and AdK, in contrast, the spatial arrangement with the global minimum energy corresponds to the protein binding on the extracellular surface of the membrane (Fig. 2B, C, E and F). In fact, for all orientations of these proteins, the lowest energy along z^c corresponds to surface-bound structures. Other water-soluble proteins also have similar preference to the membrane surfaces, as indicated by the shallow free-energy valleys in those regions (Fig. 1B). Such attraction

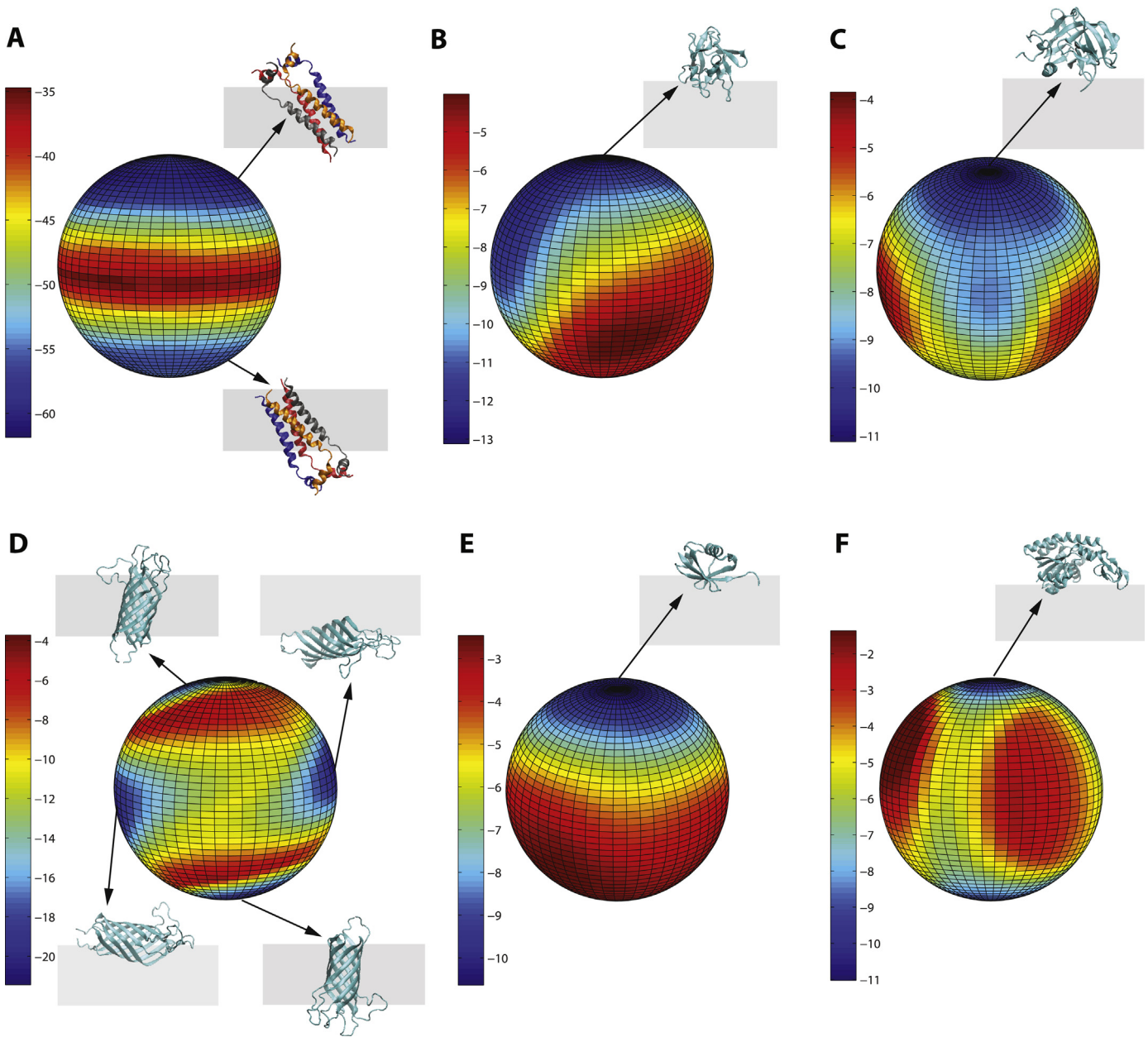

Fig. 2. Minimum energies (in k_BT) for protein orientations mapped onto a sphere surface for (A) M2, (B) FGF1, (C) EMAPII, (D) OMPA, (E) Ubq, and (F) AdK (open conformation). Every point on the sphere corresponds to a unique orientation. For each orientation, the lowest energy along the translational (z^c) degree of freedom is identified and displayed by color. Spatial arrangements corresponding to the energy minima in this orientation space are shown by the molecular images, with the membrane displayed as a 30 Å-thick *gray* slab. For M2 and OMPA, the chosen reference orientation (represented by the north pole) has the pore axis coincide with the *z*-axis. For other proteins, the spatial arrangement with the global energy minimum is taken as the reference orientation.

to the membrane surface may be partly attributed to the charged residues on the surface of the water-soluble proteins. Nonetheless, these proteins all encounter major energetic barriers when further penetrating into the membrane.

4. Discussion

Sampling and search in the orientation space of membrane proteins have been performed in a number of studies [13–21,25]. Here we further generalized this approach and carried out integrations of protein orientations to obtain a 1D free energy profile (Fig. 1) that characterizes the interaction of the protein with the membrane. Importantly, the entire orientation space is uniformly sampled in the integration, and the result does not depend on the specific rotational state of the reference structure used to define the orientations. The 1D free energy provides an estimate of the relative propensity for the protein to be in the membrane vis-à-vis in the bulk water, thereby distinguishing membrane proteins from water-soluble ones.

The free energy profile for EMAPII exhibits typical features of a water-soluble protein, with a high barrier in the interior of the membrane. Similar to other water-soluble proteins, therefore, EMAPII would not prefer to enter the membrane. Based on our calculations here, a "naïve" release mechanism, in which EMAPII independently penetrates and crosses the membrane by simple passive diffusion, thus appears to be unfavorable. Hypothetically, the EMAPII release might involve cofactors, similar to the case of FGF1, which has been characterized as a non-classical released protein. Here FGF1 has a free energy profile similar to EMAPII, and experimentally its export is dependent upon cofactors such

as S100A13 [39]. To investigate the effect of cofactors, we also calculated the free energy profiles for two protein complexes: a heterotetramer consisting of two FGF1 and two S100A13 monomers [40], and a heterohexamer consisting of two FGF1, two S100A13, and two C2A domains of synaptotagmin-1 [40]. Our results show that the free energy barriers for the FGF1-S100A13 heterotetramer and the FGF1-S100A13-C2A heterohexamer are even higher than that of an FGF1 monomer, thus indicating that these protein complexes as described by the crystal structures [40] would still not spontaneously enter or cross the lipid membrane through simple diffusion. Indeed, it was suggested that the protein complex needs to undergo conformational changes such as partial denaturation before the release [39,40]. Furthermore, the FGF1 release may involve destabilization of the lipid bilayer [8,41], which cannot be described by our model here.

We also acknowledge two limitations in our approach. First, the EMAPII crystal structure [9] used here is in the mature form, whereas the release of premature EMAPII was also observed in experiments. Due to lack of crystal structure, currently we cannot perform similar calculations on the premature EMAPII. Moreover, a rigid EMAPII conformation is assumed in our calculations. Although no experimental evidence has suggested that EMAPII would undergo any major conformational changes [42] when crossing the lipid bilayer, we cannot completely rule out such a possibility [39,40]. Whereas this study provides an evaluation of a naive mechanism for EMAPII release, further investigations, especially with experimental input, are still needed to elucidate the exact pathway and mechanism.

We show in this study that the entire orientation space can be projected onto the surface of a 3D sphere [13], after removing the degeneracy due to in-plane rotations about the membrane normal. The area of any given region in this sphere surface is always proportional to the volume of its represented orientation space, regardless of the reference structure. This rigorous representation is not only convenient for visualizing the orientation-dependent energies here, but could also facilitate the efficient sampling or search of the protein orientations [13–21,25].

We also note that the knowledge-based membrane potential [15] adopted here is of approximate nature only. One particular concern is that the potential does not treat buried and exposed residues in the protein differently. Nonetheless, this model is shown to be in good agreement with some more recent membrane potentials [18]. Furthermore, the integration method here can be readily combined with other membrane energy models [13,16,18–20,43,44] (some of which may have better accuracy and higher resolution), and should therefore be useful for the general analysis of proteins.

Acknowledgment

This research is supported by the Indiana University Collaborative Research Grant (IUCRG) fund of the Office of the Vice President for Research. The calculations were performed on a Linux cluster at School of Science, IUPUI.

References

[1] I. Prudovsky, A. Mandinova, R. Soldi, C. Bagala, I. Graziani, et al., The non-classical export routes: FGF1 and IL-1alpha point the way, J. Cell Sci. 116 (2003) 4871–4881.
[2] I. Prudovsky, F. Tarantini, M. Landriscina, D. Neivandt, R. Soldi, et al., Secretion without Golgi, J. Cell. Biochem. 103 (2008) 1327–1343.
[3] U.E. Knies, H.A. Behrensdorf, C.A. Mitchell, U. Deutsch, W. Risau, et al., Regulation of endothelial monocyte-activating polypeptide II release by apoptosis, Proc. Natl. Acad. Sci. U.S.A. 95 (1998) 12322–12327.
[4] M.A. Schwarz, J. Kandel, J. Brett, J. Li, J. Hayward, et al., Endothelial-monocyte activating polypeptide II, a novel antitumor cytokine that suppresses primary and metastatic tumor growth and induces apoptosis in growing endothelial cells, J. Exp. Med. 190 (1999) 341–354.
[5] G. Barnett, A.M. Jakobsen, M. Tas, K. Rice, J. Carmichael, et al., Prostate adenocarcinoma cells release the novel proinflammatory polypeptide EMAP-II in response to stress, Cancer Res. 60 (2000) 2850–2857.
[6] L.A. Green, R. Yi, D. Petrusca, T. Wang, A. Elghouche, et al., HIV envelope protein gp120-induced apoptosis in lung microvascular endothelial cells by concerted upregulation of EMAP II and its receptor, CXCR3, Am. J. Physiol. Lung. Cell. Mol. Physiol. 306 (2014) L372–L382.
[7] C. Eder, Mechanisms of interleukin-1beta release, Immunobiology 214 (2009) 543–553.
[8] I. Graziani, C. Bagala, M. Duarte, R. Soldi, V. Kolev, et al., Release of FGF1 and p40 synaptotagmin 1 correlates with their membrane destabilizing ability, Biochem. Biophys. Res. Commun. 349 (2006) 192–199.
[9] L. Renault, P. Kerjan, S. Pasqualato, J. Menetrey, J.C. Robinson, et al., Structure of the EMAPII domain of human aminoacyl-tRNA synthetase complex reveals evolutionary dimer mimicry, EMBO J. 20 (2001) 570–578.
[10] A.P. Chetwynd, K.A. Scott, Y. Mokrab, M.S. Sansom, CGDB: a database of membrane protein/lipid interactions by coarse-grained molecular dynamics simulations, Mol. Membr. Biol. 25 (2008) 662–669.
[11] J.P. Ulmschneider, J.C. Smith, S.H. White, M.B. Ulmschneider, In silico partitioning and transmembrane insertion of hydrophobic peptides under equilibrium conditions, J. Am. Chem. Soc. 133 (2011) 15487–15495.
[12] J. Gumbart, B. Roux, Determination of membrane-insertion free energies by molecular dynamics simulations, Biophys. J. 102 (2012) 795–801.
[13] G.E. Tusnády, Z. Dosztányi, I. Simon, Transmembrane proteins in the Protein Data Bank: identification and classification, Bioinformatics 20 (2004) 2964–2972.
[14] D. Kozma, I. Simon, G.E. Tusnády, PDBTM: Protein Data Bank of transmembrane proteins after 8 years, Nucleic Acids Res. 41 (2013) D524–D529.
[15] M.B. Ulmschneider, M.S. Sansom, A. Di Nola, Properties of integral membrane protein structures: derivation of an implicit membrane potential, Proteins 59 (2005) 252–265.
[16] A.L. Lomize, I.D. Pogozheva, M.A. Lomize, H.I. Mosberg, Positioning of proteins in membranes: a computational approach, Protein Sci. 15 (2006) 1318–1333.
[17] M.A. Lomize, I.D. Pogozheva, H. Joo, H.I. Mosberg, A.L. Lomize, OPM database and PPM web server: resources for positioning of proteins in membranes, Nucleic Acids Res. 40 (2012) D370–D376.
[18] A. Senes, D.C. Chadi, P.B. Law, R.F. Walters, V. Nanda, et al., E(z), a depth-dependent potential for assessing the energies of insertion of amino acid side-chains into membranes: derivation and applications to determining the orientation of transmembrane and interfacial helices, J. Mol. Biol. 366 (2007) 436–448.
[19] C.A. Schramm, B.T. Hannigan, J.E. Donald, C. Keasar, J.G. Saven, et al., Knowledge-based potential for positioning membrane-associated structures and assessing residue-specific energetic contributions, Structure 20 (2012) 924–935.
[20] T. Nugent, D.T. Jones, Membrane protein orientation and refinement using a knowledge-based statistical potential, BMC Bioinform. 14 (2013) 276.
[21] M.B. Ulmschneider, M.S. Sansom, A. Di Nola, Evaluating tilt angles of membrane-associated helices: comparison of computational and NMR techniques, Biophys. J. 90 (2006) 1650–1660.
[22] Y.C. Kim, G. Hummer, Coarse-grained models for simulations of multiprotein complexes: application to ubiquitin binding, J. Mol. Biol. 375 (2008) 1416–1433.
[23] C.F. Karney, Quaternions in molecular modeling, J. Mol. Graph. Model. 25 (2007) 595–604.
[24] A.D. McLaren, Optimal numerical integration on a sphere, Math. Comput. 17 (1963) 361–383.
[25] S. Esteban-Martin, D. Gimenez, G. Fuertes, J. Salgado, Orientational landscapes of peptides in membranes: prediction of ^{2}H NMR couplings in a dynamic context, Biochemistry 48 (2009) 11441–11448.
[26] J.R. Schnell, J.J. Chou, Structure and mechanism of the M2 proton channel of influenza A virus, Nature 451 (2008) 591–595.
[27] A. Arora, F. Abildgaard, J.H. Bushweller, L.K. Tamm, Structure of outer membrane protein A transmembrane domain by NMR spectroscopy, Nat. Struct. Biol. 8 (2001) 334–338.
[28] C.W. Muller, G.E. Schulz, Structure of the complex between adenylate kinase from *Escherichia coli* and the inhibitor Ap5A refined at 1.9 A resolution. A model for a catalytic transition state, J. Mol. Biol. 224 (1992) 159–177.
[29] C.W. Muller, G.J. Schlauderer, J. Reinstein, G.E. Schulz, Adenylate kinase motions during catalysis: an energetic counterweight balancing substrate binding, Structure 4 (1996) 147–156.
[30] S. Vijay-Kumar, C.E. Bugg, W.J. Cook, Structure of ubiquitin refined at 1.8 A resolution, J. Mol. Biol. 194 (1987) 531–544.
[31] X. Zhu, B.T. Hsu, D.C. Rees, Structural studies of the binding of the anti-ulcer drug sucrose octasulfate to acidic fibroblast growth factor, Structure 1 (1993) 27–34.
[32] L. Chantalat, N.D. Jones, F. Korber, J. Navaza, A.G. Pavlovsky, The crystal-structure of wild-type growth-hormone at 2.5 angstrom resolution, Protein Pept. Lett. 2 (1995) 333–340.
[33] M.L. Quillin, R.M. Arduini, J.S. Olson, G.N. Phillips Jr., High-resolution crystal structures of distal histidine mutants of sperm whale myoglobin, J. Mol. Biol. 234 (1993) 140–155.

[34] F.J. Moy, P.K. Chanda, S. Cosmi, M.R. Pisano, C. Urbano, et al., High-resolution solution structure of the inhibitor-free catalytic fragment of human fibroblast collagenase determined by multidimensional NMR, Biochemistry 37 (1998) 1495–1504.
[35] J.A. Tainer, E.D. Getzoff, K.M. Beem, J.S. Richardson, D.C. Richardson, Determination and analysis of the 2 A-structure of copper, zinc superoxide dismutase, J. Mol. Biol. 160 (1982) 181–217.
[36] R. Durley, L. Chen, L.W. Lim, F.S. Mathews, V.L. Davidson, Crystal structure analysis of amicyanin and apoamicyanin from Paracoccus denitrificans at 2.0 A and 1.8 A resolution, Protein Sci. 2 (1993) 739–752.
[37] S. Onesti, P. Brick, D.M. Blow, Crystal structure of a Kunitz-type trypsin inhibitor from *Erythrina caffra* seeds, J. Mol. Biol. 217 (1991) 153–176.
[38] J. Wang, S. Kim, F. Kovacs, T.A. Cross, Structure of the transmembrane region of the M2 protein H(+) channel, Protein Sci. 10 (2001) 2241–2250.
[39] K.M. Kathir, K. Ibrahim, D. Rajalingam, I. Prudovsky, C. Yu, et al., S100A13-lipid interactions-role in the non-classical release of the acidic fibroblast growth factor, Biochim. Biophys. Acta 1768 (2007) 3080–3089.
[40] S.K. Mohan, S.G. Rani, C. Yu, The heterohexameric complex structure, a component in the non-classical pathway for fibroblast growth factor 1 (FGF1) secretion, J. Biol. Chem. 285 (2010) 15464–15475.
[41] A. Kirov, H. Al-Hashimi, P. Solomon, C. Mazur, P.E. Thorpe, et al., Phosphatidylserine externalization and membrane blebbing are involved in the nonclassical export of FGF1, J. Cell. Biochem. 113 (2012) 956–966.
[42] D. Shental-Bechor, T. Haliloglu, N. Ben-Tal, Interactions of cationic-hydrophobic peptides with lipid bilayers: a Monte Carlo simulation method, Biophys. J. 93 (2007) 1858–1871.
[43] T.H. Schmidt, C. Kandt, LAMBADA and InflateGRO2: efficient membrane alignment and insertion of membrane proteins for molecular dynamics simulations, J. Chem. Inf. Model. 52 (2012) 2657–2669.
[44] M. Punta, A. Maritan, A knowledge-based scale for amino acid membrane propensity, Proteins 50 (2003) 114–121.
[45] W. Humphrey, A. Dalke, K. Schulten, VMD: visual molecular dynamics, J. Mol. Graph. 14 (1996) 33–38.