Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Biophysical Chemistry 147 (2010) 35–41

Contents lists available at ScienceDirect

Biophysical Chemistry

journal homepage: http://www.elsevier.com/locate/biophyschem

Construction of an implicit membrane environment for the lattice Monte Carlo simulation of transmembrane protein

Yantao Chen *, Mingliang Wang, Qianling Zhang, Jianhong Liu

Shenzhen Key Laboratory of Functional Polymer, College of Chemistry and Chemical Engineering, Shenzhen University, Shenzhen 518060, China

ARTICLE INFO

Article history:
Received 30 October 2009
Received in revised form 17 December 2009
Accepted 18 December 2009
Available online 28 December 2009

Keywords:
Protein folding
Transmembrane protein
Implicit membrane environment
Monte Carlo simulation
Lattice chain

ABSTRACT

Due to the complexity of biological membrane, computer simulation of transmembrane protein's folding is challenging. In this paper, an implicit biological membrane environment has been constructed in lattice space, in which the lipid chains and water molecules were represented by the unoccupied lattice sites. The biological membrane was characterized with three features: stronger hydrogen bonding interaction, membrane lateral pressure, and lipophobicity index for the amino acid residues. In addition to the hydrocarbon core spanning region and the water solution, the lipid interface has also been represented in this implicit membrane environment, which was proved to be effective for the transmembrane protein's folding. The associated Monte Carlo simulations have been performed for SARS-CoV E protein and M2 protein segment (residues 18–60) of influenza A virus. It was found that the coil–helix transition of the transmembrane segment occurred earlier than the coil–globule transition of the two terminal domains. The folding process and final orientation of the amphipathic helical block in water solution are obviously influenced by its corresponding hydrophobicity/lipophobicity. Therefore, this implicit membrane environment, though in lattice space, can make an elaborate balance between different driving forces for the membrane protein's folding, thus offering a potential means for the simulation of transmembrane protein oligomers in feasible time.

1. Introduction

Membrane proteins account for approximately 25% of open reading frame in most genomes [1] and perform key functions such as energy conversion, solute transportation, secretion, and signal transduction. However, the principle of membrane protein structure remains one of the greatest challenges in structural biology [2–4]. Compared with the estimated 1700 structures which are needed to account for each structural family [5], only about 180 high resolution structures for membrane proteins have been determined [6]. The complexity of the biological membrane environment causes much difficulty for related experimental research.

Computer simulation has become a standard tool to explore the structure and dynamics of membrane proteins with increasing computer power and advanced simulation algorithms [7,8]. Current atomic simulations are limited to the timescale of about 100 ns for membrane protein-containing systems because of the complexity and large size of the lipid membrane. Alternatively, much of the attention has been paid on the treatment of lipid membrane in two fashions. Similar with the coarse-graining process on proteins [9–13], the lipid chains can be grouped into united particles, therefore greatly simplifying system freedom and accelerating the simulation [14,15]. Another approach is to employ the implicit membrane environment [16]. For example, a generalized Born solvation model has been extended to describe the membrane as a hydrophobic zone [17,18]. In a lattice simulation, Chen et al. [19] set one region of the simulation box as an implicit membrane with the characterization of hydrophobicity and lateral pressure. Though, the lipid interface and the lipid hydrocarbon core are not respectively considered in the existing implicit membrane models. It was noted that the generalized Born model may not correctly include the proper interfacial characteristics, thus significantly affecting the transient protein structure while crossing the membrane interface [20].

Due to the limited computing power, most of simulation researches were focused on the transmembrane domain and related proton channel function [15,18,21], and rather rare simulations attached importance to membrane protein with a full chain length [22]. However, the terminal domains are of much biological importance. For example, the cytoplasmic domain of Virus protein U (Vpu) of HIV-1 accelerates the degradation of the CD4 receptor in the infected cells [23]; the cytoplasmic domain of M2 protein of influenza A virus favors the tetrameric assembly and stabilizes the ion channel [24,25].

Thus, an effective simulation for membrane protein should not only embed the transmembrane segments into the membrane but should also include the two terminal domains. This is a big challenge in view of the computing time. In this work, we constructed a biological membrane environment with implicit lipid chains and implicit water molecules in lattice space. Transmembrane proteins with a full chain length were

* Corresponding author.
E-mail address: ytchen@szu.edu.cn (Y. Chen).

0301-4622/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.bpc.2009.12.008

embedded into this lattice model and simulated via the dynamic Monte Carlo method in feasible time. In the following section, the simulated materials and methods are presented. Computer simulation results are given in the third section, and a summary of this work is provided in the last section.

2. Materials and methods

2.1. Simulated materials

The folding processes of two short membrane proteins have been simulated, which are both small structural proteins with transmembrane sequences bearing helical structures.

E protein, a small envelope protein with 76 amino acids, executes many biological functions for severe acute respiratory syndrome (SARS) associated coronavirus [26]. Bioinformatics indicated that SARS-CoV E protein is likely to adopt a transmembrane conformation [27]. Based on *in vitro* experiments and simulation researches, Torres et al found that the transmembrane section of SARS-CoV E protein adopts a helical structure with N-terminal and C-terminal on the opposite sides of the membrane [21,28]. Unfortunately, this protein's high resolution structure has not been determined yet.

M2, a 97-residue homotetrameric transmembrane protein, forms a proton-selective channel vital to the function of influenza A virus [29]. Spatial structure of M2 protein's proton channel has been acquired [30,31]. The polypeptide segment from residue 18 to residue 60 of M2 protein [M2(18–60)] can form stable tetramer structure [31]. This M2 (18–60) segment was chosen as a simulating object, which contains an unstructured N terminus, a transmembrane helix (residues 25–46), and another C-terminal amphipathic helix (residues 51–59).

2.2. The implicit biological membrane environment

A simple model has been constructed to simulate membrane protein's environment. As shown in Fig. 1, there are three regions in this model, i.e. the regions for aqueous solution, for the membrane interface, and the lipid hydrocarbon core spanning region. The unoccupied lattice sites implicitly represent lipid segments in the membrane region or water molecules in the aqueous region. Both of them can freely exchange with polypeptide chain's residues. Therefore, it can be called as an implicit membrane model.

The biological membrane was characterized in three aspects. An insertion of membrane protein would disturb the integrity of the compact membrane, consequently increasing the internal energy of the membrane [2]. As a reflection, an effective lateral pressure (P) [19] is introduced to minimize the polypeptide chain's lateral area (A) when inserted into the implicit membrane. Secondly, the strength of hydrogen bonding interaction was set as 3.6 kcal/mol in lipid medium and 1.6 kcal/mol in aqueous environment according to Sheu et al's simulation results [32].

Hydrophobicity and lipophobicity are also considered to distinguish the membrane region from the aqueous environment. New hydrophobicity scales (HYDS) and lipophobicity scales (LIPS) have been established based on a statistical analysis of amino acid residue's environmental preference by Mokrab et al. [33]. The lipophobicity scales are further divided into two classes for the lipid bilayer interface (LIPS-I) and the hydrocarbon core regions (LIPS-C). Mokrab et al's results were correlated very well with the Kyte–Doolittle index [34] for hydrophobicity (Fig. S1 in Supplementary materials) and the White–Wimley index [2] for lipophobicity (Fig. S2). According to the fitting line's slope in Figs. S1 and S2, the Kyte–Doolittle index is about 4.9 times of HYDS and the White–Wimley index is about 4.3 times of LIPS-C, showing that the hydrophobicity scales and the lipophobicity scales obtained from different environments can be reasonably related.

In our simulation, the hydrophobic index (I_{water}), the lipophobic indexes (I_{inter} for membrane interface and I_{core} for hydrocarbon core region) are multiples of HYDS, LIPS-I and LIPS-C with a coefficient ($\varepsilon_{\text{phobic}}$). $\varepsilon_{\text{phobic}}$ was set as 4.8 kcal/mol, just three times of the hydrogen bonding energy in aqueous solution.

2.3. Lattice chain and associated potentials

A highly coarse-grained lattice model was employed in this work. Each residue is taken as a basic unit and represented by a minimal cube ($a\times a\times a$) with 8 neighboring lattice sites at its apexes [13]. The bond length between two consecutive residues may fluctuate within a defined range. This bond fluctuation model contains 5 bond lengths and 87 permitted bond orientations [35,36], much more relaxable than the conventional single-site lattice model [37,38], thus allowing the construction of a more “realistic” helix with little anisotropy in lattice space. In addition, as a branching point, disulfide bond can be easily embedded into this lattice model.

The energy function for our lattice model can be simply written as

$$E = E_l + E_\theta + E_{\text{press}} + E_{\text{chiral}} + E_{\text{hbond}} + E_{\text{phobic}}. \quad (1)$$

The potentials associated with the fluctuations of the bond length and bond angle between consecutive residues are, respectively, given by

$$E_l = \sum_{i=1}^{N-1} u_{l,i} = \sum_{i=1}^{N-1} \frac{1}{2} k_l (l_i - l_o)^2 \quad (2)$$

and

$$E_\theta = \sum_{i=2}^{N-1} u_{\theta,i} = \sum_{i=2}^{N-1} \frac{1}{2} k_\theta (\cos\theta_i - \cos\theta_o)^2. \quad (3)$$

Here, N is the polypeptide chain length; k_l and k_θ are the elastic constants respectively; l_i is the module of the bond vector $\boldsymbol{l}_i$ which indicates the vector connecting residue i to residue $i+1$; θ_i is the angle between $\boldsymbol{l}_{i-1}$ and $\boldsymbol{l}_i$; θ_o and l_o are the underlying expectation values. θ_o is chosen as 96°, according to the real value of polypeptide chain [39]. l_o is set to be $\sqrt{6}$, the distance between (0,0,0) and (2,1,1), which is the most favorable relaxation mode in this lattice model.

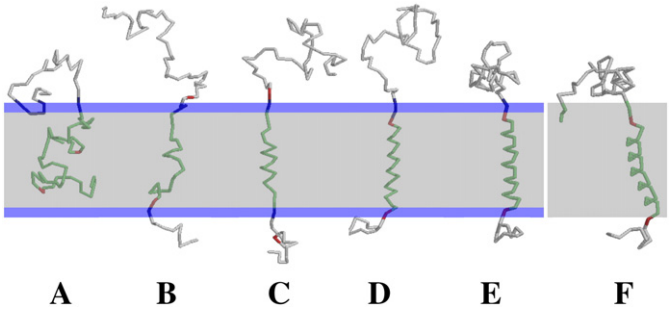

Fig. 1. Sketch map for the biological membrane environment and some typical snapshots of SARS-CoV E protein during the associated folding process. Gray zone and blue zone represent the membrane's hydrocarbon core and lipid interface respectively. The polypeptide segments in hydrocarbon core and lipid interface and water solution are colored in green, blue, and gray correspondingly. The two ends of the helical segment are labeled in red color. (A) $1/T^* = 0$; (B) $1/T^* = 0.64$; (C) $1/T^* = 1.28$; (D) $1/T^* = 2.08$; (E) $1/T^* = 8.16$; (F) $1/T^* = 8.16$.

E_{press} is employed to account the energy variation of membrane resulting from the insertion of polypeptide segments:

$$E_{\text{press}} = PA = \sum_{i=1}^{N-1} Pl_{p,i}^2. \tag{4}$$

Here, P is the effective lateral pressure, A is the chain's lateral area, and $l_{p,i}$ is the projected length of bond vector $\boldsymbol{l}_i$ on the membrane plane. Other than staying in water solution, the polypeptide segment that immerged in the membrane has to maintain an orientation perpendicular to the membrane plane so as to minimize the damage on the integrity of the membrane.

α-helix associated potentials are included according to our previous work [13] with minor modification. The hydrogen bonding potential E_{hbond} is expressed as

$$E_{\text{hbond}} = \sum_{i=1}^{N_h} \sum_{j=h_{N,i}}^{h_{C,i}-4} u_{\text{hbond},i,j}, \tag{5}$$

where

$$u_{\text{hbond},j} = \begin{cases} 0 & (d_{j,j+4} < d_1) \\ -\varepsilon_{\text{hbond}} & (d_1 \le d_{j,j+4} \le d_2) \\ 0 & (d_{j,j+4} > d_2) \end{cases}. \tag{6}$$

Here, N_h is the total number of helical blocks for the simulated protein, and h_N and h_C are the two ends of the helical blocks. The hydrogen bonding occurs when the distance $d_{i,i+4}$ between residue i and residue $i+4$ is among the range from $d_1=\sqrt{10}$ to $d_2=\sqrt{12}$. For SARS-CoV E protein, $H=1$, $h_{N,1}=11$ and $h_{C,1}=34$; for M2(18–60) segment, $H=2$, $h_{N,1}=25$, $h_{C,1}=46$, $h_{N,2}=51$ and $h_{C,2}=59$. $\varepsilon_{\text{hbond}}=\varepsilon_0=1.6$ kcal/mol in aqueous solution and the lipid interface. $\varepsilon_{\text{hbond}}=9\varepsilon_0/4$ in the hydrocarbon core spanning regions of membrane.

In our coarse-grained model, the side chains are not reflected explicitly. Alternatively, the following potential is employed to have the chirality effect included

$$E_{\text{chiral}} = \sum_{i=1}^{N_h} \sum_{j=h_{N,i}}^{h_{C,i}-3} u_{\text{chiral},i,j}, \tag{7}$$

where

$$u_{\text{chiral},j} = \begin{cases} \varepsilon_{\text{chiral}} & (\mathbf{l}_j \times \mathbf{l}_{j+1} \cdot \mathbf{l}_{j+2} \le 0) \\ 0 & (\mathbf{l}_j \times \mathbf{l}_{j+1} \cdot \mathbf{l}_{j+2} > 0) \end{cases}. \tag{8}$$

Accordingly, a non-right-handed orientation for three consecutive bonds is penalized with a positive energy ε_0. This simple definition is just for the formation of α-helix, and the construction of other complex protein structures in a coarse-grained model may consult some delicate representations for side chain's orientation [40].

The energy function for hydrophobic interaction with water and lipophobic interaction with membrane is represented by

$$E_{\text{phobic}} = \sum_{i=1}^{N} u_{\text{phobic},i} = \sum_{i=1}^{N} \left(\frac{m_{\text{water},i} I_{\text{water}}}{M} + \frac{m_{\text{inter},i} I_{\text{inter}}}{M} + \frac{m_{\text{core},i} I_{\text{core}}}{M} \right). \tag{9}$$

Here, m_{water}, m_{inter}, and m_{core} are the number of unoccupied lattice sites around the polypeptide chain in the environments of the water solution, the lipid interface and the hydrocarbon core respectively. Only the nearest neighboring unoccupied lattice site are considered, thus the maximum is $M=24$. The hydrophobic index (I_{water}) and the lipophobic index (I_{inter} and I_{core}) have been determined as described in Sub-section 2.2, and their values for the twenty natural amino acid residues can be found in Table S1 of the Supplementary materials.

The other parameters of our energy function are listed as following: $\varepsilon_{\text{hbond}}=\varepsilon_0$, $\varepsilon_{\text{chiral}}=\varepsilon_0/2$, $k_l=\varepsilon_0/2$, $k_\theta=\varepsilon_0/2$, $P=\varepsilon_0/4$.

2.4. Simulation details

The lattice box is composed of $150\times150\times150$ cubic lattice sites, and the periodic boundary condition is set along in X and Y directions. In the Z direction, the N-terminal ectodomain can penetrate into the membrane, but is not allowed to enter the cytoplasmic side, and vice versa for the C-terminal domain, thus the membrane protein's insertion process is not taken into account in this work just for simplicity. In this lattice space, the thicknesses of the implicit membrane's lipid interface and hydrocarbon core are set as $3a$ and $18a$ respectively according to the structure of a dioleoylphosphocholine (DOPC) bilayer [2].

The dynamic Monte Carlo simulation was employed to perform chain relaxation. Each Monte Carlo step (MCS) was composed of N attempts. For each attempt, a residue was first selected randomly, then one of the nearest neighbor "site groups" along six principal directions was selected randomly [35,36]. Only when all the sites in the selected "site group" remain unoccupied, the movement can be attempted. Metropolis importance sampling [41] was employed as the ultimate criterion for the acceptance or rejection in every attempting movement.

According to the idea of simulated annealing, a series of temperatures from the athermal state to the lowest simulated temperature were sequentially selected for each trajectory. At each temperature, w MCS's were used for relaxation to achieve its thermodynamic equilibration, and then another w MCS's were used for the collection of one thousand statistics. Longer relaxation times were implemented for lower temperatures. In this simulation, $w=1\ 000\ 000$ MCS's at the athermal state, and $w=9\ 000\ 000$ MCS's at the lowest temperature. Each trajectory started from an independent random coiled conformation. For both proteins, it took no more than 600 h to run all of the 50 trajectories on a PC with 2.33 GHz CPU.

2.5. Measured quantities

Generally, the thermodynamic transition can be estimated from the peak of specific heat. The reduced specific heat at a certain temperature can be examined by the following form [13]

$$c_v^* = \frac{C_V}{Nk_B} = \frac{\langle E^2\rangle - \langle E\rangle^2}{N(k_BT)^2}, \tag{10}$$

where '< >' represents the ensemble average.

The mean square radius of gyration $<S^2>$ has been employed to describe the size of chain conformation, which can be calculated by the following equation

$$S^2 = \frac{1}{N}\left\langle \sum_{i=1}^{N} (\mathbf{r}_i - \mathbf{r}_{\text{cm}})^2 \right\rangle. \tag{11}$$

Here, $\boldsymbol{r}_i$ is the position vector for the ith residue, and $\boldsymbol{r}_{\text{cm}}$ is the center of mass of the polypeptide chain.

Further, the radius-of-gyration tensor $\boldsymbol{S}$ can be defined as

$$\mathbf{S} = \frac{1}{N}\mathbf{R}\mathbf{R}^T, \tag{12}$$

where

$$\mathbf{R} = (\mathbf{r}_1, \mathbf{r}_2, \cdots\cdots, \mathbf{r}_N). \tag{13}$$

Here, $\boldsymbol{R}$ is the position matrix for a chain's specific conformation and $\boldsymbol{R}^T$ is the corresponding transposed matrix. Via the diagonalization of $\boldsymbol{S}$, three eigenvalues and related eigenvectors can be computed [42]. Tilt angle (ω) was used to describe helical block's relative orientation, which

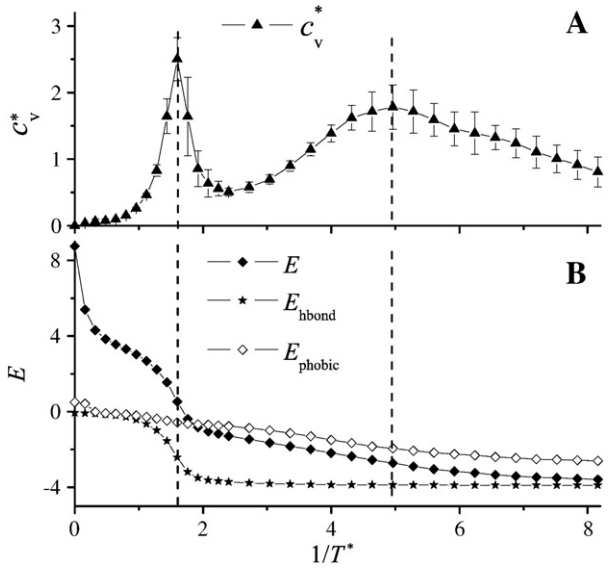

Fig. 2. Reduced specific heat (c_v^*) and chain energy (E) as functions of inversed temperature ($1/T^*$) for SARS-CoV E protein. The dashed lines indicate the peak positions of specific heat corresponding to the coil–helix transition point and the coil–globule transition point. Error bars come from the standard deviations of all trajectories.

is defined as the angle of the conformation's longest principle with respect to the membrane normal:

$$\omega = \arccos\left(\frac{\lambda_{max,z}}{\lambda_{max}}\right), \tag{14}$$

where λ_{max} is the eigenvector's module in the longest principle axis, and $\lambda_{max,z}$ is the component in the direction of membrane normal, i.e., Z direction.

Pattern recognition algorithm [43] is employed to determine the so-called α-helical block. Two consecutive α-helical H-bonds (from i to $i+4$) can shape a minimal α-helical block, in which all residues along the sequence belong to this helical block. Two minimal α-helical blocks can build a long α-helical block with no less than two shared consecutive residues. The maximal helical block can reach the chain length of a polypeptide segment. Based on this algorithm, the helical ratio and the length of helix blocks, i.e. θ and L_h, can be determined.

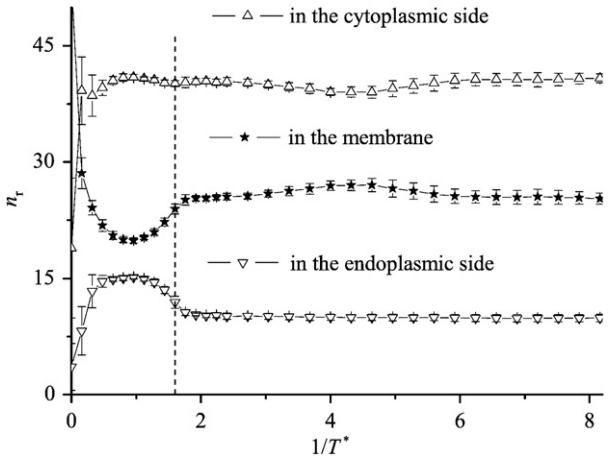

Fig. 3. Number of residues (n_r) in specific regions of the membrane as functions of inversed temperature ($1/T^*$) for SARS-CoV E protein. The dashed line indicates the peak position of specific heat corresponding to the coil–helix transition point.

3. Results and discussions

3.1. SARS-CoV E protein

3.1.1. Two separated thermodynamic transitions

Fig. 1 presents several typical snapshots for the thermodynamic folding progress of SARS-CoV E protein. The polypeptide chain started folding from a coil conformation (Fig. 1A), and finally evolved into a transmembrane protein (Fig. 1E). Fig. 2 shows the variations of reduced specific heat and chain energy during thermal annealing. Here, a dimensionless temperature T^* ($\equiv k_B T/\varepsilon_0$) is defined, where k_B is Boltzmann constant. Clearly, two separated peaks occurred for the reduced specific heat curve at $1/T^* = 1.60$ and 4.96, which should correspond to two thermodynamic transitions. The hydrogen bonding energy (E_{hbond}) and the hydrophobic and lipophobic energy (E_{phobic}) decreased rapidly around $1/T^* = 1.60$ and 4.96 respectively, indicating that those two transitions are the coil–helix transition and the coil–globule transition.

Due to the membrane's lateral pressure, chain energy (E) decreased rapidly from the athermal state. As shown in Fig. 3, at high temperature, the number of residues (n_r) in the membrane decreased quickly when being extruded out of membrane (Fig. 1B), and then mounted up around the coil–helix transition point. The hydrogen bond tends to form in the hydrogen core region, therefore stretching chain segment back into the membrane after overcoming the large lateral-pressure potential (Fig. 1C). After the coil–helix transition, the partition of residues has been fixed, i.e., about 10 residues in the endoplasmic side, about 25 residues in the membrane, about 41 residues in the cytoplasmic side. The two ends of the transmembrane helix are just near the lipid interface of membrane (Figs. 1D and E).

The variation of chain conformation's size has been explored in Fig. 4. The transmembrane segment (11–34) was first elongated at high temperature, and then shrank quickly around the coil–helix transition point. In contrast, the two terminal domains (1–10 and 35–76) decreased around the coil–globule transition point.

Thus, the thermodynamic folding process of SARS-CoV E protein may be imagined as follows. In the athermal state, the polypeptide chain can

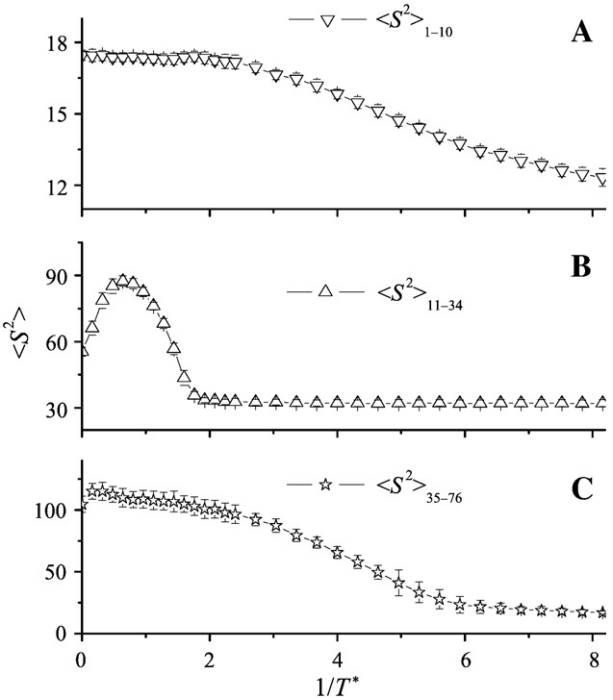

Fig. 4. Mean square radius of gyration $<S^2>$ as a function of inversed temperature ($1/T^*$) for SARS-CoV E protein's three chain segments.

relax freely through the membrane environment, adopting a conformation similar to the random coil (Fig. 1A); during thermal annealing, the chain segment was extruded outside of the membrane due to the lateral pressure, shaping an expanded conformation (Fig. 1B); then hydrogen bonding interaction dominated the chain's relaxation, and the inserted chain segment shaped into a regular helical structure in the membrane (Fig. 1C and D); at lower temperature, the hydrophobic interaction prevailed in the folding process, and the two terminal domains collapsed into compact globule conformations (Fig. 1E). This folding process indicated that the coil–helix transition and the coil–globule transition were apart from each other for the SARS-CoV E protein.

3.1.2. The role of lipid interface

The lipid interface is a necessary portion for the biological membrane [2]. A comparison has been made between the cases with and without the lipid interface in our implicit membrane model. Though the formation of transmembrane helix was not affected by the deletion of lipid interface (Fig. 5A), there were clear differences for the C-terminal domain. For the case without lipid interface, the last several residues in the C-terminal with strong hydrophobicity prefer to penetrate into the hydrocarbon core of the membrane as shown (Fig. 1F), decreasing the number of residues (n_r) in the cytoplasmic side (Fig. 5B). The increased size of the C-terminal domain's conformation ($<S^2>_{35-76}$ in Fig. 5C) indicated that the C-terminal domain would bend over the membrane interface (Fig. 1F) rather than collapse into a compact globular conformation.

The hydrophobic force at the lipid interface is only one half of that in water solution [33], but holds equal lateral pressure with the hydrocarbon core. As a result, the C-terminal domain would prefer to stay outside and then collapsed into a compact globule at low temperature. The role of the lipid interface can also be confirmed in the kinetic process. Baumgaertner et al. found that the lipid–water interface may represent a barrier against the membrane protein's spontaneous insertion [20].

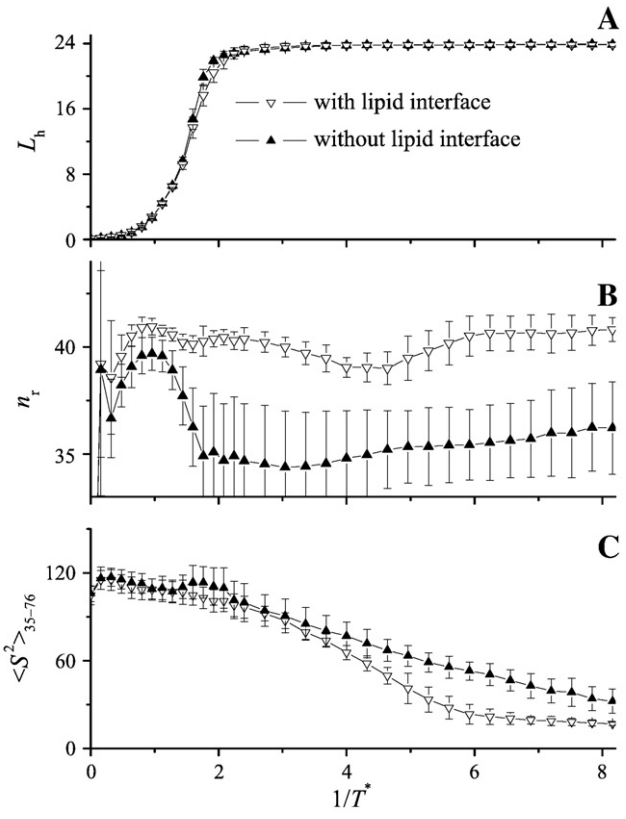

Fig. 5. Comparison of the length of helix block (L_h), the number of residues (n_r) in the cytoplasmic side and the mean square radius of gyration $<S^2>_{35-76}$ for SARS-CoV E protein's segment between the cases with and without membrane's lipid interface.

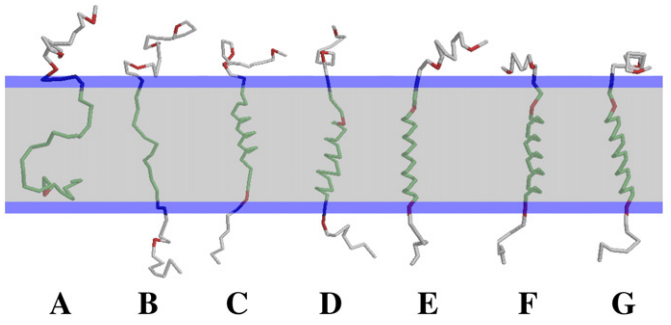

Fig. 6. Typical snapshots of the M2(18–60) segment during the folding process. Gray zone and blue zone represent the membrane's hydrocarbon core and lipid interface respectively. The polypeptide segments in hydrocarbon core and lipid interface and water solution are colored in green, blue, and gray correspondingly. The two ends of the helical segment are labeled in red color. (A) $1/T^*=0$; (B) $1/T^*=0.48$; (C) $1/T^*=1.6$; (D) $1/T^*=2.56$; (E) $1/T^*=7.36$; (F) $1/T^*=9.6$; (G) $1/T^*=9.6$. (F) and (G) are the same conformation but from different angles of view.

3.2. The M2(18–60) segment

3.2.1. Helix formation depending on hydrophobicity

Different from SARS-CoV E protein, the M2(18–60) segment can form two helices. From an initial coil conformation (Fig. 6A), the chain was elongated (Fig. 6B). Then, the helix in the membrane first shaped (Fig. 6C), and later the helix outside of the membrane (Fig. 6D and E). A folded conformation for the M2(18–60) segment formed lastly at low temperature(Figs. 6F and G).

The folding process of the helices is also illustrated based on the variation of the helical ratio (θ) in Fig. 7. The half ratio of the helical property was chosen to quantificationally describe the folding process for the helix(25–46) segments and the helix(51–59) segments, which are at $1/T^*=1.44$ and 2.64 respectively. The transmembrane helix (25–46) formed at higher temperature because of stronger hydrogen bonding in the hydrocarbon core of the membrane. Due to the big energy gap for hydrogen bonding between in hydrocarbon core and in water solution, the hydrogen bonded residues preferred to immerge into the membrane. In contrast, the amphipathic helix(51–59) shaped more slowly in water solution for lack of enough hydrophobicity. Therefore, the helix formation process was affected by its corresponding hydrophobicity or lipophobicity, indicating a certain degree of coupling between helix formation and chain collapse.

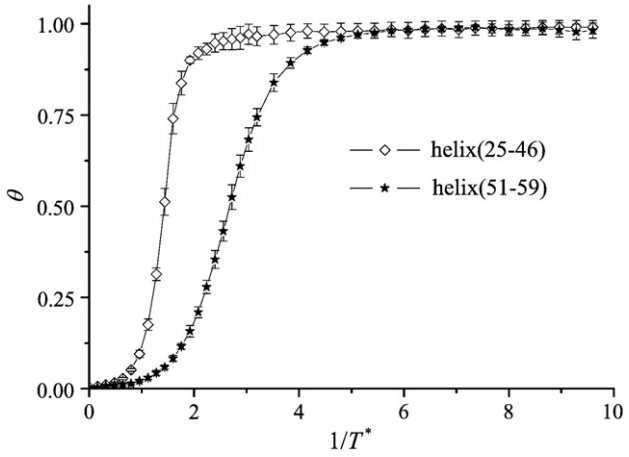

Fig. 7. Helical ratio (θ) as a function of inversed temperature ($1/T^*$) for the M2(18–60) segment's two helices.

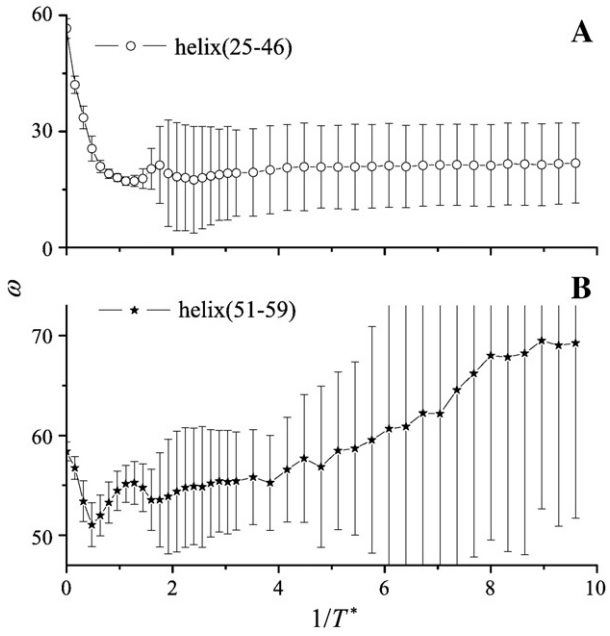

Fig. 8. Tilt angle (ω) as a function of inversed temperature ($1/T^*$) for the M2(18–60) segment's two helices.

3.2.2. Helix orientation

Helix segment's orientation with respect to the membrane plane has been described by tilt angle (ω) as shown in Fig. 8. After being in shape, the transmembrane helix(25–46) adopted an orientation with $\omega \approx 20°$, consistent with the experimental results [44,45]. The tilt angle for the amphipathic helix(51–59) was about 55° at high temperature when in coil conformation, and then mounted up to about 70° at low temperature because of lipophilic interaction with the membrane interface, which means that the helix(51–59) is nearly parallel to the membrane. The orientation fluctuation as displayed by the error bar in Fig. 8B was somewhat considerable. This may resulted from the weak hydrophobicity of the helix(51–59), indicating that oligomer may be a more stable form for M2 protein. As illustrated in the NMR experiment [31], the M2(18–60) tetramer forms a C-terminal base via the right-handed packing of four amphipathic helices.

The presence of the membrane region breaks the spatial symmetry, therefore local anisotropy should be carefully dealt with for lattice model [9]. In this bond fluctuation model, α-carbon bond vector holds 87 orientations and the square of hydrogen bond length can also fluctuate between 10 and 12, thus allowing the formed regular helix with little anisotropic orientation. Supporting evidence can be found in Figs. 6 and 8. Helix(25–46) and helix(51–59) adopted clearly different orientations in lattice space, and the orientation swayed in a big range.

4. Summary

In this work, a biological membrane model with implicit lipid segments and implicit water molecules has been constructed in lattice space. The polypeptide chain was coarse-grained into a strand of minimal cubes connected by fluctuating α-carbon bonds, and the unoccupied lattice sites were regarded as the implicit lipid segments and implicit water molecules. The biological membrane was characterized with three features: strong hydrogen bonding, membrane lateral pressure, and lipophobicity. Moreover, another feature of this implicit membrane model is the representation of lipid interface, which was proved to be very resultful for the folding of transmembrane protein.

The associated Monte Carlo simulation has been performed. In the folding process of SARS-CoV E protein, the coil–helix transition occurred obviously earlier than the coil–globule transition, demonstrating that the coil–helix transition and the coil–globule transition were apart from each other in the thermodynamic folding process of the transmembrane protein. For the M2(18–60) segment, the transmembrane helix(25–46) formed earlier than the amphipathic helix(51–59), and the latter finally adopted an orientation nearly parallel to the membrane surface. Thus, the coil–helix transition process for the membrane protein segment and final layout in the membrane environment are obviously influenced by its corresponding hydrophobicity or lipophobicity. The different folding processes for helix(25–46) and helix(51–59) resulted from the competition between membrane lateral pressure, hydrogen bonding strength, and related hydrophobicity or lipophobicity. Such phenomena bear evidence of our implicit membrane model on the elaborate balance between different driving forces for membrane protein.

Lastly, most of transmembrane proteins form oligomers *in vivo* [28,31]. Schnell's experiment [31] indicated that the C-terminal domain was of much significance for the assembling or oligomerization of membrane proteins, thereby reinforcing the worthy of simulating transmembrane protein with two terminal domains. In this work, the implicit membrane environment and associated coarse-grained model can efficiently reduce the workload and therefore provide ensemble averaged results in feasible time. Thus our model can be a potential means for simulating the folding and dynamics of transmembrane protein oligomers.

Acknowledgements

The authors are grateful to Prof. Bingshi Li for the invaluable discussions on this manuscript. This work was supported by NSF of China (Grant Nos. 20804023, 20673073, and 20601019) and the Cultivation Project of Excellent Young Innovative Talents of Higher Education of Guangdong Province (Grant No. LYM08086).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.bpc.2009.12.008.

References

[1] E. Wallin, G. von Heijne, Genome-wide analysis of integral membrane proteins from eubacterial, archaean, and eukaryotic organisms, Protein Sci. 7 (1998) 1029–1038.
[2] S.H. White, W.C. Wimley, Membrane protein folding and stability: physical principles, Annu. Rev. Biophys. Biomolec. Struct. 28 (1999) 319–365.
[3] J.U. Bowie, Solving the membrane protein folding problem, Nature 438 (2005) 581–589.
[4] A. Elofsson, G. von Heijne, Membrane protein structure: prediction versus reality, Annu. Rev. Biochem. 76 (2007) 125–140.
[5] A. Oberai, Y. Ihm, S. Kim, J.U. Bowie, A limited universe of membrane protein families and folds, Protein Sci. 15 (2006) 1723–1734.
[6] S.H. White, Biophysical dissection of membrane proteins, Nature 459 (2009) 344–346.
[7] W.L. Ash, M.R. Zlomislic, E.O. Oloo, D.P. Tieleman, Computer simulations of membrane proteins, Biochim. Biophys. Acta-Biomembr. 1666 (2004) 158–189.
[8] G.S. Ayton, G.A. Voth, Systematic multiscale simulation of membrane protein systems, Curr. Opin. Struct. Biol. 19 (2009) 138–144.
[9] A. Kolinski, J. Skolnick, Reduced models of proteins and their applications, Polymer 45 (2004) 511–524.
[10] V. Tozzini, Coarse-grained models for proteins, Curr. Opin. Struct. Biol. 15 (2005) 144–150.
[11] S.Y. Kim, J. Lee, Folding simulations of small proteins, Biophys. Chem. 115 (2005) 195–200.
[12] L. Zhang, D.N. Lu, Z. Liu, How native proteins aggregate in solution: a dynamic Monte Carlo simulation, Biophys. Chem. 133 (2008) 71–80.
[13] Y.T. Chen, Y.Q. Zhou, J.D. Ding, The helix-coil transition re-visited, Proteins 69 (2007) 58–68.
[14] S.J. Marrink, H.J. Risselada, S. Yefimov, D.P. Tieleman, A.H. de Vries, The MARTINI force field: coarse grained model for biomolecular simulations, J. Phys. Chem. B 111 (2007) 7812–7824.
[15] P.J. Bond, J. Holyoake, A. Ivetac, S. Khalid, M.S.P. Sansom, Coarse-grained molecular dynamics simulations of membrane proteins and peptides, J. Struct. Biol. 157 (2007) 593–605.
[16] M. Feig, Implicit membrane models for membrane protein simulation, in: A. Kukol (Ed.), Methods in Molecular Biology [M], Humana Press Inc., 2008, pp. 181–196.
[17] W. Im, M. Feig, C.L. Brooks, An implicit membrane generalized Born theory for the study of structure, stability, and interactions of membrane proteins, Biophys. J. 85 (2003) 2900–2918.

[18] M.B. Ulmschneider, J.P. Ulmschneider, M.S.P. Sansom, A. Di Nola, A generalized Born implicit-membrane representation compared to experimental insertion free energies, Biophys. J. 92 (2007) 2338–2349.
[19] C.M. Chen, C.C. Chen, Computer simulations of membrane protein folding: structure and dynamics, Biophys. J. 84 (2003) 1902–1908.
[20] M.M. Sperotto, S. May, A. Baumgaertner, Modelling of proteins in membranes, Chem. Phys. Lipids 141 (2006) 2–29.
[21] J. Torres, J.F. Wang, K. Parthasarathy, D.X. Liu, The transmembrane oligomers of coronavirus protein E, Biophys. J. 88 (2005) 1283–1290.
[22] V. Lemaitre, D. Willbold, A. Watts, W.B. Fischer, Full length Vpu from HIV-1: combining molecular dynamics simulations with NMR spectroscopy, J. Biomol. Struct. Dyn. 23 (2006) 485–496.
[23] K. Levesque, A. Finzi, J. Binette, E.A. Cohen, Role of CD4 receptor down-regulation during HIV-1 infection, Curr. HIV Res. 2 (2004) 51–59.
[24] G.G. Kochendoerfer, D. Salom, J.D. Lear, R. Wilk-Orescan, S.B.H. Kent, W.F. DeGrado, Total chemical synthesis of the integral membrane protein influenza A virus M2: role of its C-terminal domain in tetramer assembly, Biochemistry 38 (1999) 11905–11913.
[25] K. Tobler, M.L. Kelly, L.H. Pinto, R.A. Lamb, Effect of cytoplasmic tail truncations on the activity of the M-2 ion channel of influenza A virus, J. Virol. 73 (1999) 9695–9701.
[26] D.X. Liu, Q. Yuan, Y. Liao, Coronavirus envelope protein: a small membrane protein with multiple functions, Cell. Mol. Life Sci. 64 (2007) 2043–2048.
[27] P.A. Rota, M.S. Oberste, et al., Characterization of a novel coronavirus associated with severe acute respiratory syndrome, Science 300 (2003) 1394–1399.
[28] J. Torres, K. Parthasarathy, X. Lin, R. Saravanan, A. Kukol, D.X. Liu, Model of a putative pore: the pentameric alpha-helical bundle of SARS coronavirus E protein in lipid bilayers, Biophys. J. 91 (2006) 938–947.
[29] L.H. Pinto, R.A. Lamb, The M2 proton channels of influenza A and B viruses, J. Biol. Chem. 281 (2006) 8997–9000.
[30] A.L. Stouffer, R. Acharya, D. Salom, A.S. Levine, L. Di Costanzo, C.S. Soto, V. Tereshko, V. Nanda, S. Stayrook, W.F. DeGrado, Structural basis for the function and inhibition of an influenza virus proton channel, Nature 451 (2008) 596-U513.
[31] J.R. Schnell, J.J. Chou, Structure and mechanism of the M2 proton channel of influenza A virus, Nature 451 (2008) 591-U512.
[32] S.Y. Sheu, E.W. Schlag, H.L. Selzle, D.Y. Yang, Hydrogen bonds in membrane proteins, J. Phys. Chem. B 113 (2009) 5318–5326.
[33] Y. Mokrab, T.J. Stevens, K. Mizuguchi, Lipophobicity and the residue environments of the transmembrane alpha-helical bundle, Proteins 74 (2009) 32–49.
[34] J. Kyte, R.F. Doolittle, A simple method for displaying the hydropathic character of a protein, J. Mol. Biol. 157 (1982) 105–132.
[35] I. Carmesin, K. Kremer, The bond fluctuation method: a new effective algorithm for the dynamics of polymers in all spatial dimensions, Macromolecules 21 (1988) 2819–2823.
[36] H.P. Deutsch, K. Binder, Inter-diffusion and self-diffusion in polymer mixtures: a Monte Carlo study, J. Chem. Phys. 94 (1991) 2294–2304.
[37] P.H. Verdier, W.H. Stockmayer, Monte Carlo calculations on the dynamics of polymers in dilute solution, J. Chem. Phys. 36 (1962) 227–235.
[38] H.J. Hilhorst, J.M. Deutch, Analysis of Monte Carlo results on the kinetics of lattice polymer chains with excluded volume, J. Chem. Phys. 63 (1975) 5153–5161.
[39] A. Mukherjee, B. Bagchi, Correlation between rate of folding, energy landscape, and topology in the folding of a model protein HP-36, J. Chem. Phys. 118 (2003) 4733–4747.
[40] N.V. Buchete, J.E. Straub, D. Thirumalai, Continuous anisotropic representation of coarse-grained potentials for proteins by spherical harmonics synthesis, J. Molec. Graph. Model. 22 (2004) 441–450.
[41] N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, E. Teller, Equation of state calculations by fast computing machines, J. Chem. Phys. 21 (1953) 1087–1092.
[42] B.N. Parlett, The Symmetric Eigenvalue Problem, Prentice Hall, Englewood Cliffs, New Jersey, 1980.
[43] W. Kabsch, C. Sander, Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features, Biopolymers 22 (1983) 2577–2637.
[44] J.U. Bowie, Helix packing in membrane proteins, J. Mol. Biol. 272 (1997) 780–789.
[45] M.B. Ulmschneider, M.S.P. Sansom, A. Di Nola, Evaluating tilt angles of membrane-associated helices: comparison of computational and NMR techniques, Biophys. J. 90 (2006) 1650–1660.