J Comput Aided Mol Des (2012) 26:363–374
DOI 10.1007/s10822-012-9562-1

Replica exchange Monte-Carlo simulations of helix bundle membrane proteins: rotational parameters of helices

H.-H. Wu · C.-C. Chen · C.-M. Chen

Received: 14 September 2011 / Accepted: 18 March 2012 / Published online: 31 March 2012

Abstract We propose a united-residue model of membrane proteins to investigate the structures of helix bundle membrane proteins (HBMPs) using coarse-grained (CG) replica exchange Monte-Carlo (REMC) simulations. To demonstrate the method, it is used to identify the ground state of HBMPs in a CG model, including bacteriorhodopsin (BR), halorhodopsin (HR), and their subdomains. The rotational parameters of transmembrane helices (TMHs) are extracted directly from the simulations, which can be compared with their experimental measurements from site-directed dichroism. In particular, the effects of amphiphilic interaction among the surfaces of TMHs on the rotational angles of helices are discussed. The proposed CG model gives a reasonably good structure prediction of HBMPs, as well as a clear physical picture for the packing, tilting, orientation, and rotation of TMHs. The root mean square deviation (RMSD) in coordinates of C_α atoms of the ground state CG structure from the X-ray structure is 5.03 Å for BR and 6.70 Å for HR. The final structure of HBMPs is obtained from the all-atom molecular dynamics simulations by refining the predicted CG structure, whose RMSD is 4.38 Å for BR and 5.70 Å for HR.

Keywords Helix bundle membrane proteins · Monte-Carlo simulations · Molecular dynamics simulations · Structure prediction

Electronic supplementary material The online version of this article (doi:10.1007/s10822-012-9562-1) contains supplementary material, which is available to authorized users.

H.-H. Wu · C.-C. Chen · C.-M. Chen (✉)
Department of Physics, National Taiwan Normal University, 88 Sec. 4 Ting-Chou Rd., Taipei 116, Taiwan
e-mail: cchen@phy.ntnu.edu.tw

Introduction

With the availability of whole genome sequences, bioinformatic analyses show that more than a quarter of proteins coded in genomes are membrane proteins (MPs) [1–3]. Functionally normal MPs are vital to the survival of living cells; their functions include cell–cell contact, surface recognition, cytoskeleton contact, signaling, enzymatic activity, or transporting substances across the membrane [4]. Many known diseases result from the defects of MPs. The clinical importance of MPs is demonstrated by the fact that more than 50 % of known drugs in use today target MPs [5, 6], which are also responsible for the uptake, metabolism, and clearance of these pharmacologically active substances. Despite their biological and pharmaceutical importance, due to difficulties in crystallizing MPs, only about 200 unique structures have been derived so far [7]. As the attempts of using experimental methods to obtain membrane protein structures have encountered difficulties, great efforts have been directed at analyzing membrane proteins on a theoretical basis by model building, with the aid of low resolution structural data. Therefore, there exist great incentives for computational and theoretical studies of MPs [8–15].

Structures of transmembrane domains of MPs can be categorized into three classes: those with a β-barrel structure, those that cross the lipid membrane with a single α-helix of length between 17 and 25 amino acids, and those that transverse the membrane with an α-helix bundle (multiple α-helices). Among them, theoretical analysis has shown that helix bundles are much more abundant than β-barrels. The folding of helix bundle MPs (HBMPs) is usually understood by a two stage model [16–19]: Independently stable helices are formed in lipid membrane in the first stage, and the helices interact with others to form a functional MP in the second stage. Statistical analysis of available α-helical MPs

in the protein data bank (PDB) has revealed useful information related to the environmental preferences of amino acids within the membrane, which was first reported by Rees et al. in 1989 that membrane-exposed amino acids are usually more hydrophobic than buried amino acids in the photosynthetic reaction center structure [20]. More recent analysis of amino acid distributions in α-helical MPs has shown that large hydrophobic amino acids (e.g., leucine, isoleucine, valine, and phenylalanine) favor the lipid-exposed environment, while small side-chain amino acids (e.g., glycine, alanine, serine, and threonine) favor helix–helix interfaces [21–23]. Further analysis of residue–residue pairwise propensities in α-helical MPs suggested that polar and cation–pi interactions are more frequent in α-helical MPs than in water-soluble proteins [24]. In addition, recent experiments also showed the role of polar amino acids in mediating helix–helix association in the membrane [25, 26].

From the above observations, a comparative modeling of α-helical MPs using the Rosetta-Membrane method was proposed by considering residue–residue and residue–environment interactions [27]. The preliminary tests of this method yielded reasonable structure prediction for some small MPs (<4 Å root mean square deviation, RMSD) but poor prediction for large MPs (>9 Å RMSD). The failure in predicting the structure of large MPs seems to be consistent with the analysis of MPs by Stevens and Arkin [28], which suggests that maximization of van der Waal's (vdW) forces, in addition to a general segregation of hydrophobicities driven by lipid exclusion, also plays an important role in the folding of large MPs. Another threading method, TASSER (Threading ASSEmbly Refinement), threaded the query sequence on parts of solved protein structures, and then refined the resulting template by ab initio algorithms. Test of the method on a set of 38 MP structures led to 12 structures with RMSD to native less than 4 Å, but many others with RMSD to native greater than 6 Å. Two physics-based methods, MembStruk and PREDICT, were specifically designed to predict the structures of G protein-coupled receptors. These methods produced reasonably good structure of bovine rhodopsin with RMSD about 3 Å. However, the best predictions were selected based on structural similarity between models. More stringent energetic discrimination between models is desired. A comprehensive review of MP structure prediction can be found in a recent article by Barth [29].

Considerable efforts have been made on the structure prediction of proteins using purely physics-based protein models [30]. In 1997, Duan and Kollman folded the 36-residue villin headpiece by all-atom simulations in explicit solvent for 2 months on parallel supercomputers with structures up to 4.5 Å. More recently, this small protein was folded by Pande and coworkers to 1.7 Å with a total simulation time of 300 μs. However, the validity of these physics-based all-atom models has not been systematically tested, and it is rather impossible to use them for structure prediction of typical-size proteins consisting of 100–300 residues due to high computational cost. Instead, these all-atom models can be used to refine low-resolution protein structures, which can be derived from reduced physics-based models. For example, the UNRES (UNited-RESidue) method developed by Scheraga and coworkers represents a residue by two interacting united atoms, C_α and the side chain center. This method has been systematically applied to many CASP (Critical Assessment of techniques for protein Structure Prediction) targets since 1998. The most notable prediction by this approach is for T061 from CASP3, for which a predicted structure of 4.2 Å to the native for a 95-residue helical protein was generated with an accuracy gap from the rest of structures. In CASP6, this approach also folded a structure genomic target of 102-residue T0230 to a structure within 7.3 Å. Another physics-based method, ASTRO-FOLD, recently proposed by Floudas and coworkers for protein structure prediction has constructed a structure of 5.2 Å for a four-helical bundle protein of 102 residues in a double-blind prediction. Although purely physics-based methods represent the most promising approaches to solve the problem of protein structure prediction, they remain to be the most challenging approaches due to their best-but-still-low accuracy.

Based on the two-stage model, we have previously demonstrated the feasibility in predicting the native structure of HBMPs by a coarse-grained (CG) approach of computer simulations with very limited experimental information [14, 15]. By simulating the packing of sequence dependent helices, this approach leads to reasonable structure prediction of retinal proteins, including bacteriorhodopsin (BR), halorhodopsin (HR), sensory rhodopsin II (SRII), and (bovine) rhodopsin, with RMSD between 1.9 and 3.0 Å. The predicted HBMP structures have the lowest energy of the CG model, which includes most physical interactions. However, each transmembrane helix (TMH) in this CG model is represented by a heterogeneous polymer cylinder and a monomer represents a turn of helices. The rotational angle (Ω) along its long axis is thus ignored in the CG Monte-Carlo (MC) simulations of HBMPs. To restore the rotation angle of each helix in the all-atom (AA) refinement, we assumed that the most hydrophobic surface of helices are aligned to face the membrane core [31]. However, this policy for the rotational angles of helices based on hydrophobicity may not be universal for all MPs. Other interactions might also contribute to the rotational angles of helices [28].

Here, we construct a united-residue model of HBMPs in which the angular representation of a TMH in the membrane is described by a tilting angle (Θ), an orientation angle (Φ), and a rotational angle (Ω) as illustrated in Fig. 1. In this model, a polypeptide chain is represented as a sequence of α-

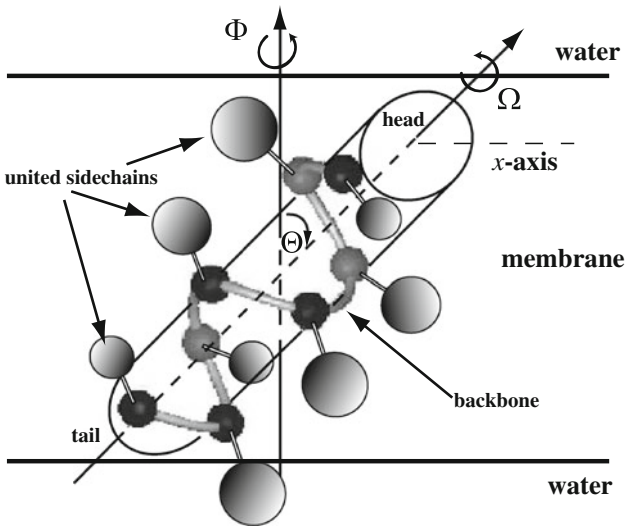

Fig. 1 Schematic representation of a TMH showing its tilting (Θ), orientation (Φ), and rotational (Ω) angles. The Φ angle is defined to be the included angle between the *x*-axis and the projection of the helix axis on the membrane, while the Ω angle is defined to be the included angle between the *x*-axis and the vector from the helix axis to the most hydrophobic amphiphilic face

carbon atoms (C_α's). The C_α's are linked together by rigid bonds, which constitute the backbone. United side chains are connected to the backbone by rigid bonds. This model contains a phenomenal energy term which models the helix–helix interaction in an inhomogeneous membrane environment to avoid simulating lipid membrane explicitly with a tremendous computational cost. These angular parameters (Θ, Φ, Ω) of TMHs calculated computationally can be directly compared with their experimental measurements from site-directed dichroism [32]. A detailed description of our CG model is given in Sect. 2. Two HBMPs, BR and HR, are tested without using their 3D structure information in PDB. They are retinal proteins found in the purple membrane of Halobacterium salinarium [33] and serve as excellent model systems for constructing a physical model to predict the structure of HBMPs and to study thermodynamics of MP folding, due to the relatively simple topology of their structure. The lowest-energy state structure of HBMPs in this model was obtained using the replica exchange Monte-Carlo (REMC) algorithm [34], as described in Sect. 3. In Sect. 4, we describe the folding dynamics of HBMPs using CG MC simulations and a refinement of the folded structure by AA molecular dynamics (MD) simulations using AMBER9 [35]. In Sect. 5, we discuss results from our computer simulations of BR, HR, and their subdomains. The predicted packing, tilting, orientation, and rotation of TMHs of HBMPs are reasonably consistent with experimental data, and their physical origins can be well understood from our CG model. AA refinement of our CG structures of HBMPs has a RMSD of 4–6 Å. Section 6 gives our conclusions.

Model

In the present study, we have adopted a united-residue model to simulate a protein chain, in which each residue of the chain is represented by a sphere with certain vdW radius and hydrophobicity. Based on the two-stage model, we assume that the initial structure of HBMPs contains several helices randomly residing in the membrane, which are constrained by flexible inter-helix loops. Each sequence dependent TMH was first constructed as a standard helix with the φ and ψ torsional angles of residues equal to -60 and -40 degrees. These helix templates were then subject to an energy minimization using AMBER9 with the force field frcmod.ff99SB to derive their lowest energy conformation, which maintains basic structural features of TMHs. After constructing all TMHs in the AA representation, the position of C_α atoms in the helix has been assigned as the location of the corresponding residues in the united-residue model, as shown in Fig. 1. Hydrogen bonding within each TMH was assumed in our CG simulations by fixing the helical structure of each TMH. The loop constraint is modeled by limiting the head-to-tail distance between two consecutive helices, which is proportional to the number of residues in the loops (each residue is about 3.8 Å). Here, head and tail of a TMH refer to the two end points on its axis as shown in Fig. 1. If the head-to-tail distance (r_{ht}) between two consecutive helices is smaller than the length of the corresponding loop, they can move without feeling the loop constraint. However, it is not permitted for r_{ht} to be larger than the maximal length. In other words, the potential for the loop constraint is a step function of r_{ht}, which is zero if r_{ht} is less than its maximal distance and is infinity if it is larger than the maximal distance. These helices are allowed to diffuse in the membrane, to tilt (Θ) and to rotate (Φ) along the membrane normal direction, as well as to rotate (Ω) along their long axis, as shown in Fig. 1.

Among various physical interactions, evidences show that the vdW interaction and side-chain packing among TMHs mostly determine the tertiary structure of MPs [4, 16, 17]. Although inter-helical hydrogen bonding, ion pairs, and disulfide bonds have been considered as alternative sources of stability, there are only few cases demonstrating the importance of these alternative interactions. In our model, the vdW interaction between helices is expressed as

$$E_{\mathrm{vdw}} = e_1 \sum_{\substack{i,j \\ i \neq j}} \sum_{m,n} \left\{ \left[\frac{r_0^{m_i} + r_0^{n_j}}{r(m_i, n_j)} \right]^{12} - \left[\frac{r_0^{m_i} + r_0^{n_j}}{r(m_i, n_j)} \right]^{6} \right\}, \qquad (1)$$

where e_1 is the strength of the vdW interaction, $r_0^{m_i}$ is the corresponding vdW radius of m-th residue in i-th helix, and $r(m_i, n_j)$ represents the distance between residues m_i and n_j. The adopted value of vdW radius (r_0) for each residue is

calculated from its estimated vdW volume (V_{vdW}) [36], i.e., $V_{\mathrm{vdW}} = (4\pi/3)r_0^3$. The vdW interaction in Eq. (1) is a sum of all vdW energy between residues, which is sequence dependent in our CG model.

The helix–water interaction E_{hw} is modeled by using a rescaled Kyte-Doolittle hydropathy index (HI), i.e., (Ala, Arg, Asn, Asp, Cys, Gln, Glu, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr, Val) = (0.4, −1, −0.78, −0.78, 0.56, −0.78, −0.78, −0.09, −0.71, 1, 0.84, −0.87, 0.42, 0.62, −0.36, −0.18, −0.16, −0.2, −0.29, 0.93), with strength e_2, which is mainly determined by the Gibbs free energy change for transferring amino acids from water into condensed vapor [37]. In other words, we have

$$E_{\mathrm{hw}} = e_2 \sum_k \mathrm{HI}_k, \tag{2}$$

where k runs over all residues in TMHs when they are in water region (the value of HI is zero for all residues when they are in the membrane region). We note that our choice of the Kyte-Doolittle scale is not unique and there exist a number of other hydrophobicity scales which are based on different methods used to measure hydrophobicity [38]. Since we have set $e_2 = 1$ in our simulations, one unit of energy is about 3 kcal/mol (the energy change for transferring Isoleucine from water into condensed vapor).

In addition, detailed studies of model hydrophobic helices in phospholipid bilayers have shown that lipids in the immediate neighborhood of a helix are perturbed due to the helix–lipid interaction [39, 40]. For example, the hydrophobic matching could influence the lipids on alkyl chain flexibility and order. We thus model the helix–lipid interaction E_{hl} by a tilting energy of the helices in the membrane

$$E_{hl} = e_3 \sum_i (1 - \cos\Theta_i), \tag{3}$$

where Θ_i is the tilting angle of the i-th helix and the energy coefficient e_3 is mainly determined by the resistance of the alkyl chains to changes in their length [41]. The tilting energy increases if a helix is tilted from the membrane normal, due to the increase in the contact between lipids and helix. Since the secondary structure of MPs is assumed to be unchanged in our simulations, helix tilting will reduce the unfavorable contact between helices and water. Therefore the competition between E_{hw} and E_{hl} will determine the tilting angle of helices.

Our energy terms E_{hw} and$d$ E_{hl} are modeled assuming that the environment of helices in the membrane region is uniform, which is not precise when a MP is buried in the membrane. Residues at the inner surface of a MP are expected to experience a more hydrophilic environment than those at the outer surface. From previous studies [42, 43], it is known that the configuration of a helix inside the membrane depends mainly on (1) the hydrophobic association between its non-polar faces and the lipid environment as well as (2) the hydrophilic association between its polar faces with those of other helices. In other words, in the absence of other interactions, a TMH would tend to arrange its most hydrophobic face toward membrane and its most hydrophilic face toward the MP center. To model this effect, we first characterize surface hydrophobicity of TMHs. Each TMH was divided into four amphiphilic faces and its hydrophilic residues located at both ends were ignored due to their exposure to water. As a result, each amphiphilic face contains 5–6 residues and its hydrophobicity is represented by a HI vector of these residues [44]. Here we denote the k-th element of j-th amphiphilic face HI vector in i-th helix by $\varepsilon^{\mathrm{amp}}_{i,j,k}$, whose value is given by the rescaled Kyte-Doolittle HI. The HI value of the lipid environment ($\varepsilon^{\mathrm{amp}}_{\mathrm{lipid}}$) is defined to be 1. In our CG model, the amphiphilic interaction between helices and their inhomogeneous environment is expressed as

$$E_{\mathrm{amp}} = -e_4 \left\{ \sum_{\substack{i_1,i_2,j_1,j_2 \\ i_1 \neq i_2}} f^{j_1,j_2}_{i_1,i_2} \cdot \left(\sum_k \varepsilon^{\mathrm{amp}}_{i_1,j_1,k} \cdot \varepsilon^{\mathrm{amp}}_{i_2,j_2,k} + \text{“i.n.n.r.”} \right) + \left(\sum_{i,j,k} \chi_{i,j} \cdot \varepsilon^{\mathrm{amp}}_{i,j,k} \cdot \varepsilon^{\mathrm{amp}}_{\mathrm{lipid}} \right) \right\} \tag{4}$$

Here the first two terms are helix–helix interactions and the third term is helix–lipid interaction. The value of $f^{j_1,j_2}_{i_1,i_2}$ is 1 if the j_1-th amphiphilic face of helix i_1 and the j_2-th amphiphilic face of helix i_2 can have proper contact, and is 0 otherwise [45]. The first term is a direct scalar product of HI vectors of corresponding amphiphilic faces. The second term “i.n.n.r.”, however, considers the interaction of nearest neighbor residues with a strength factor $\lambda < 1$. More specifically, we express “i.n.n.r.” as $\lambda\left(\sum_{k=1}^{4} \varepsilon^{\mathrm{amp}}_{i_1,j_1,k} \cdot \varepsilon^{\mathrm{amp}}_{i_2,j_2,k+1} + \sum_{k=2}^{5} \varepsilon^{\mathrm{amp}}_{i_1,j_1,k} \cdot \varepsilon^{\mathrm{amp}}_{i_2,j_2,k-1}\right)$. The value of $\chi_{i,j}$ is 1 if the j-th amphiphilic face of helix i does not have proper contact with any amphiphilic face of other helices, and is 0 otherwise. The proposed energy term E_{amp} in Eq. (4) mainly determines the rotational angles of TMHs. Its value is minimized if the most hydrophobic face of TMHs is in contact with the lipid environment. We note that a number of structures of different rotational angles could have the lowest value of E_{amp} in our CG model, due to the range of defined proper contact between two amphiphilic faces in Ref. [45]. This simple expression of the amphiphilic interaction is proposed to give the first order approximation of the rotational angles of helices.

Finally we discuss the effect of the retinal in stabilizing the channel state over a hexagonal packing state which is expected to have lower vdW energy. For such a hexagonal packing state, the retinal will be outside the HBMP and has

unfavorable contacts with lipids. Conversely, the retinal will form hydrogen bonding with water molecules and have favored contacts with helices, if it is in the channel state [46, 47]. In our CG model, the retinal molecule is simplified as a rod of four spheres with a separation distance about 3.7 Å, representing the carbon clusters of C_1–C_6, C_7–C_9, C_{10}–C_{12}, and C_{13}–C_{15} whose vdW radius is 3.22, 2.58, 2.35, and 2.58 Å, respectively. We thus model the non-covalent interaction between the retinal and its environment by a contact energy between retinal and helices,

$$E_{\text{contact}} = -e_5 \sum_{i=1}^{7} \varepsilon(\Delta r_i), \tag{5}$$

where Δr_i is the shortest distance between the axes of retinal and i-th helix, and $\varepsilon(\Delta r_i)$ is 1 if 6 Å $< \Delta r_i <$ 11 Å and 0 otherwise. We note that hydrogen bonding within each TMH was assumed in our CG model by fixing the helical structure of each TMH during MC simulations. Therefore, to find the ground state structure of HBMPs, the relevant physical quantity to be minimized in our model is the total energy $E_{\text{total}} = E_{\text{vdw}} + E_{\text{hw}} + E_{\text{hl}} + E_{\text{amp}} + E_{\text{contact}}$.

Searching algorithm for the ground state structure

The REMC method has been shown to have good search properties in protein studies, whose basic idea is a conditional temperature swapping of two configurations of the protein, each in a regular canonical simulation at different temperatures [34, 48]. This approach effectively enhances the probability for the protein to get out of local energy minima. Given two configurations, each with energies and temperatures E_1, T_1 and E_2, T_2, respectively, the probability of swapping their simulation temperatures is given by $P = \min\left\{1, \exp\left[-(E_2 - E_1)\left(\frac{1}{kT_1} - \frac{1}{kT_2}\right)\right]\right\}$. Our simulation data are gathered as follows. For 4-helix subdomains, MC simulations of 10 replicas of a protein with random initial states are executed independently at different temperatures (T = 0.08, 0.1, 0.13, 0.17, 0.21, 0.28, 0.36, 0.46, 0.59, and 0.77) (in units of 3 kcal/mol). In our distributed computing system consisting of a cluster of quad core 1.8 GHz AMD Opteron CPUs, MC simulation of each replica is carried out at a single core. After a run time of 200 MC steps, all information of replicas will be collected by a server computer. The server computer will then attempt to swap temperature between replicas based on the above swapping probability. For 7-helix HBMPs, since the required number of replicas in REMC grows with the square root of the degree of freedom of the systems, we have simulated the folding of MPs by using 15 replicas with a different temperature set (0.100, 0.125, 0.156, 0.195, 0.244, 0.305, 0.381, 0.477, 0.596, 0.745, 0.931, 1.164, 1.455, 1.819, 2.274) and a swapping frequency of one swap per 10^3 MC steps. Our simulations have shown that, under the above conditions, the swapping probability between temperatures is about 25–53 %, which usually leads to an efficient search of the lowest energy state structure in available replicas among various temperatures [49].

Simulation methods

The dynamic CG simulation of HBMP folding is performed in a simulation box, which is divided into three regions: a membrane core (of thickness 26–32 Å) sandwiched by two water regions [17]. Our simulations of HBMP folding start from a random initial arrangement of TMHs in the membrane region with proper loop constraints. The retinal is permanently linked to the G-helix on one end while the other end is allowed to move in the membrane. The presence of this retinal molecule in the structure formation of HBMPs will block helices from entering the pore region of the helix-bundle. The folding of HBMPs is simulated by the REMC algorithm in a continuum space. In our simulations, seven rigid helices are allowed to move in the simulation box by changing their Θ, Φ, and Ω angles, as well as the position of their center of mass. Here the Ω angle of a helix is defined to be the rotational angle of its first amphiphilic face. The transformation of residue positions from (x, y, z) to (x', y', z'), i.e., $\begin{pmatrix} x\prime \\ y\prime \\ z\prime \end{pmatrix} = J \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ is derived from imposing an angular change to a helix from (Θ, Φ, Ω) to $(\Theta', \Phi', \Omega') = (\Theta + \Delta\Theta, \Phi + \Delta\Phi, \Omega + \Delta\Omega)$ Here the transformation matrix corresponding to this angular change is

$$J = \begin{pmatrix} \cos\Phi' & -\sin\Phi' & 0 \\ \sin\Phi' & \cos\Phi' & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\Theta' & \sin\Theta' \\ 0 & -\sin\Theta' & \cos\Theta' \end{pmatrix} \begin{pmatrix} \cos\Phi' & \sin\Phi' & 0 \\ -\sin\Phi' & \cos\Phi' & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos(\Delta\Omega + \Phi) & -\sin(\Delta\Omega + \Phi) & 0 \\ \sin(\Delta\Omega + \Phi) & \cos(\Delta\Omega + \Phi) & 0 \\ 0 & 0 & 1 \end{pmatrix} \times \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\Theta & -\sin\Theta \\ 0 & \sin\Theta & \cos\Theta \end{pmatrix} \begin{pmatrix} \cos\Phi & \sin\Phi & 0 \\ -\sin\Phi & \cos\Phi & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

If any attempted move of cylinders satisfies the constraints of excluded volume and inter-helical loops, the move is accepted with probability $w = \min[1, \exp(-\Delta E/kT)]$, where ΔE is the energy change of the system and T is the temperature of a replica. Since the REMC method effectively avoids the trapping of the protein chain at local energy minima, typically it would take about 10^6 MC steps to find the ground state structure of the chain using our computer cluster.

In addition to examining folded structure of HBMPs using a CG model, an AA calculation is implemented to improve the folded structures of HBMPs at the atomic level. To construct the AA representation of the lowest energy structure of HBMPs in our CG model, those energy minimized helices in the united-residue model as described in Sect. 2 are replaced by their corresponding helices in the AA representation. Inter-helix loops are added to connect consecutive helices using MC simulations with a spring potential between corresponding ends of loops and helices, in which case helices of HBMPs are frozen as a template to be added with loops. The retinal is covalently bound to Lys-216 of the G-helix. It is placed randomly and allowed to move in the simulation. The atomic charges of the retinal and Lys-216 are taken from Tajkhorshid et al. [50] This structure is then refined by an energy minimization with 5,000 steps of steep descent method and 10,000 steps of conjugate gradient method. Here the hydrophobic core of membrane is treated as a dielectric medium of dielectric constant $\kappa = 2.5$ (its value is between 2 and 4) [51]. As a first order approximation, we treat the environment of MPs as a uniform dielectric, which screens out charges by a factor $1/\kappa$. More sophisticated model of the environment of MPs can be adopted to improve the predicted structures of MPs. Starting from the energy-minimized structure, we carry out restrained MD simulations to further refine the folded structure by allowing both helices and loop segments to move. The restraints include the torsional angles ($|\Delta\phi| \leq 1°$ and $|\Delta\psi| \leq 1°$) and the distance between N and O atoms of hydrogen bonds in the helices ($|\Delta r| \leq 0.1$ Å). The time step is 2 fs. The bonds associated with hydrogen atoms were fixed at their equilibrium bond lengths. The cutoff distance for non-bonded interactions is 100 Å to include all atom–atom non-bonded interactions. The temperature coupling parameter for a constant temperature simulation is set to be 5 ps.

Results and discussion

According to the thermodynamic hypothesis of protein folding, the native state of proteins is the global minimum of free energy [52]. This hypothesis is consistent with the experimental results of London and coworkers, in which they have demonstrated denaturation and renaturation of BR under a wide variety of experimental conditions [53, 54]. To investigate the validity of our CG model of HBMPs, we have simulated the folding of BR (PDB code 1PY6) and HR (PDB code 1E12), as well as their 4-helix subdomains (TMHs 3–6), starting from random initial configurations. The lowest energy structure of HBMPs is identified by the REMC method within 10^7 MC steps. The typical energy parameters used in our simulations are $e_1 = 0.25$, $e_2 = 1.0$, $e_3 = 0.7$, $e_4 = 0.6$, $e_5 = 2$, and $\lambda = \{0.0, 0.2\}$, but the choice of these parameters is not unique. The RMSD in coordinates of C_α atoms of our predicted CG structures from their PDB structures is in the range of 3–7 Å.

4-helix subdomain of BR (BRD4) and HR (HRD4)

To test whether our method is able to predict the packing of TMHs and to investigate the appropriate range of various parameters in our CG model and in the REMC method, we first simulate the packing of 4-helix subdomain of BR (TMHs 3–6), which has a relatively stable structure [27]. Since we exclude helix 7 of BR, which is chemically bonded to the retinal, we do not consider the energy term in Eq. (5). As demonstrated in the supporting information, the RMSD in coordinates of C_α atoms of the ground state CG structure from the PDB structure is 4.44 Å for $\lambda = 0.0$ and is 3.96 Å for $\lambda = 0.2$. The RMSD of rotational angles of our predicted structures is defined by $\Omega_{\mathrm{RMSD}} = \sqrt{n^{-1}\sum_{i=1}^{n}\left(\Omega_i - \Omega_i^0\right)^2}$, where Ω_i is the rotational angle of i-th helix, and Ω_i^0 is its corresponding value in the PDB. It is found that Ω_{RMSD} is 42.39° for $\lambda = 0.0$ and is 40.10° for $\lambda = 0.2$. Similar simulations have been carried out to predict the packing of 4-helix subdomain of HR (TMHs 3–6). The RMSD in coordinates of C_α atoms of the ground state CG structure from the PDB structure is 4.49 Å for $\lambda = 0.0$ and is 4.37 Å for $\lambda = 0.2$. The value of Ω_{RMSD} of our predicted structures is 50.99° for $\lambda = 0.0$ and is 49.85° for $\lambda = 0.2$. The proposed energy term E_{amp} in Eq. (4) for the rotational angles of TMHs is validated by the high correlation between E_{amp} and Ω_{RMSD} as shown in the supplementary information. However, since many structures of different rotational angles in our CG model could have the lowest value of E_{amp}, it is expected that the RMSD of the rotational angles of the predicted structure would not be small ($40° < \Omega_{\mathrm{RMSD}} \leq 50°$). In general, our CG model has a better prediction of BRD4 and HRD4 in the packing, tilting, orientation, and rotation of helices for the model with $\lambda = 0.2$. However, the differences among our predictions using different values of λ are not significant.

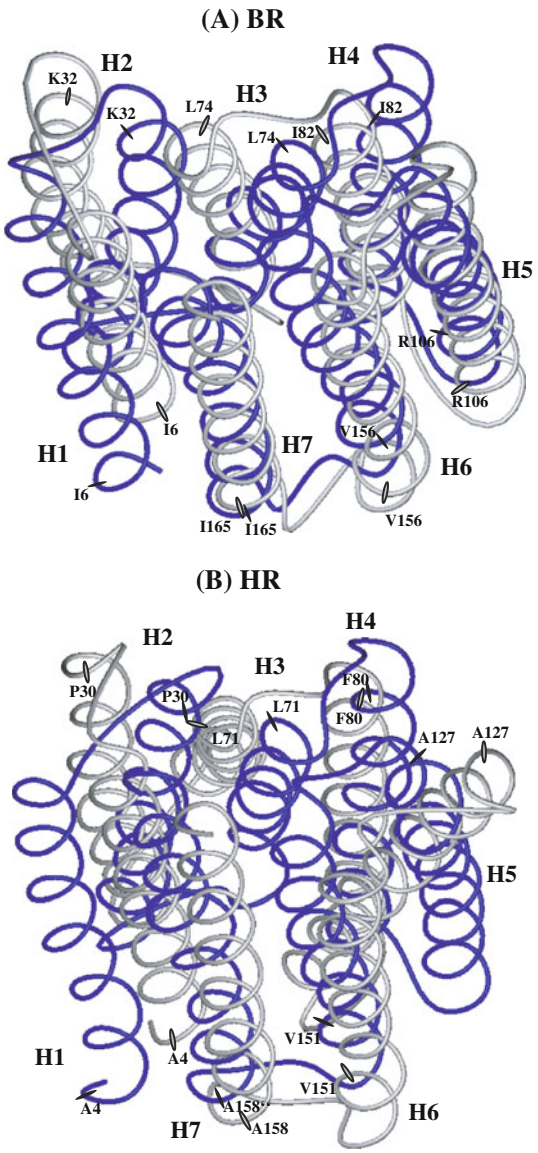

Fig. 2 Comparison of the PDB structure (*dark ribbons*) from our CG predictions (*grey ribbons*) for **a** BR and **b** HR. Here we have shown the location of one residue for each helix to visualize the prediction error in the rotational angle

Bacteriorhodopsin (7 helices)

After testing our CG model on subdomains of BR and HR for different values of λ, we conclude that the CG model with $\lambda = 0.2$ leads to better structure predictions and thus apply it to study the folding of HBMPs. Bacteriorhodopsin (PDB code 1PY6) is a representative of a family of bacterial rhodopsin structures, the largest known family of MP structures available in the PDB. It is a MP found in the purple membrane of Halobacterium salinarium and serves as a light-driven proton pump. In Fig. 2a, we show the predicted CG structure of BR (grey ribbons, retinal not shown) from our model with a RMSD in coordinates of C_α atoms of 5.03 Å. The RMSD values of tilting and orientation angles of the predicted BR structure are calculated to be 10.82° and 108.11°, respectively. From Table 1, it is found that the Θ angle of most helices is much smaller than its value from the PDB structure, which is due to the large value of membrane thickness (32 Å) used in our simulation. For a smaller value of membrane thickness (26 Å), the predicted structure has $\Theta_{RMSD} \approx 4.2°$. The prediction error in the rotational angles of helices is mostly between 0° and 60°, except for helix 1 whose prediction error is 112.29°. The Ω_{RMSD} value of our CG structure of BR is 59.32°. A comparison of our predictions and experimental data for the three angular parameters of TMHs for BR is shown in Table 1. The general structure of our prediction for BR is reasonably consistent with its PDB structure (dark ribbons), and the major discrepancy between our predicted structure and the PDB structure comes from helices 1–3. Helices 1 and 2 are shifted away from their packing position, while the orientation angle of helix 3 deviates significantly from its orientation in the PDB structure with an error of 139.22°. Due to our representation of amino acids with the united-residue model, the poor packing of these helices is the main source of prediction error. This error could be reduced by using more appropriate values of vdW radii in the model calculated from an AA model.

To investigate if our CG model does provide a sound physical basis for the structure prediction of HBMPs, the scattered diagram in Fig. 3a shows the relationship between energy and RMSD for 1,700 saved CG structures. Protein structures with high energy usually have a very large RMSD and are not shown in this scattered diagram. In general, structures with lower energy in our CG model tend to have smaller RMSD, which is consistent with the thermodynamic hypothesis of protein folding. The calculated correlation coefficient between energy and RMSD is 0.63 for BR. We note that, in our CG model, the energy of the predicted BR structure is −60.6, while it is −50.3 (−17.2) for the PDB structure with (without) local energy minimization. The relatively high energy for the PDB structure is due to unfavorable residue overlaps.

Halorhodopsin (7 helices)

To further test our CG model, similar REMC simulations have been carried out to obtain the structure of HR (PDB code 2JAF). HR is a light driven chloride pump, which converts the energy of green light into an electrochemical chloride gradient. A comparison of our predicted structure of HR (grey ribbons) with its PDB structure (dark ribbons) is shown in Fig. 2b. The predicted structure of HR has an energy of −70.1 in the CG model, while the model energy of the PDB structure is −58.0 with local energy

Table 1 A comparison of our predictions and experimental data for the three angular parameters (Θ, Φ, Ω) of TMHs for BR

Angle	Helix						
	H1	H2	H3	H4	H5	H6	H7
Θ							
PDB	21.14	1.98	10.57	4.96	14.51	12.01	12.63
MC ($L = 26$)	2.40 (17.79)	0.15 (5.93)	2.91 (5.22)	1.09 (7.08)	2.98 (19.52)	0.10 (17.02)	1.82 (8.34)
MD	13.53	10.66	4.53	2.9	6.5	9.55	7.65
Φ							
PDB	308.25	17.19	123.19	71.62	40.11	189.65	336.90
MC	236.63	46.98	262.41	56.15	242.93	44.12	235.49
MD	257.83	28.65	220.59	34.38	286.74	137.51	303.67
Ω							
PDB	335.75	299.66	315.13	253.25	190.79	112.30	73.91
MC	88.24	249.24	262.99	194.23	228.04	69.33	74.48
MD	79.64	241.18	293.93	190.22	209.13	78.51	43.54

The values of Θ predicted from MC simulations are for the case of $L = 32$, while its values in parentheses are results from MC simulations using $L = 26$

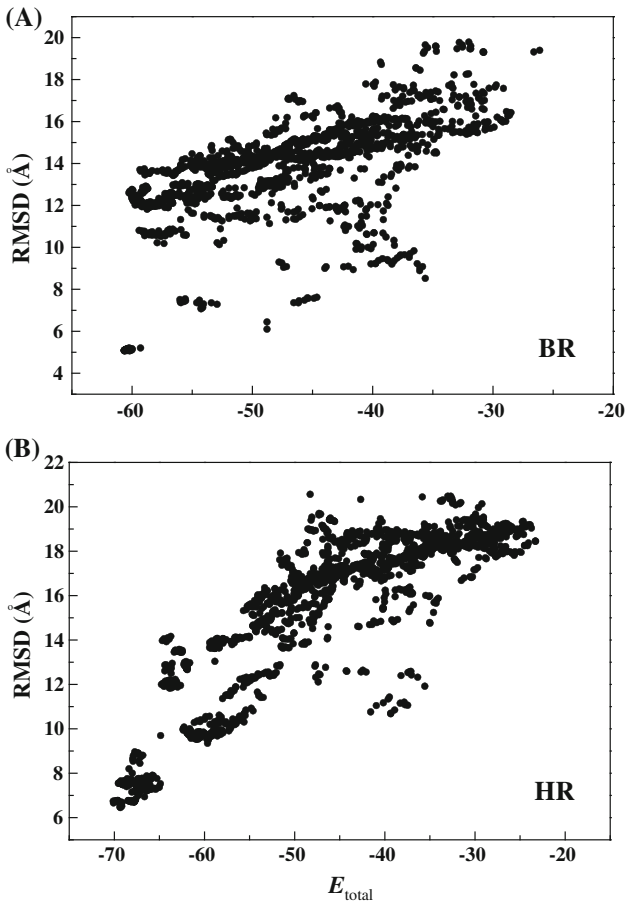

Fig. 3 Scatter plots of **a** BR and **b** HR in the energy-RMSD (from the PDB structure) plane for 1,700–1,800 observed structures. The value of RMSD is calculated only for C_α atoms in TMHs

minimization and is −26.3 without energy minimization. The relatively high energy for the PDB structure is due to unfavorable residue overlaps. The overall RMSD in coordinates of C_α atoms of our CG predicted structure of HR is 6.70 Å. The RMSD values of tilting and orientation angles of the predicted HR structure are 13.00° and 103.95°, respectively. The Ω_{RMSD} value of the predicted CG structure of HR is calculated to be 46.03°. A comparison of our predictions and experimental data for the three angular parameters of TMHs for HR is given in Table 2. Figure 3b shows the scattered diagram of energy and RMSD for 1,800 saved structures of HR. The correlation coefficient between energy and RMSD is found to be 0.86, which provides consistent evidence in verifying the physical hypothesis of protein folding.

All atom refinement of MP structures

The predicted HBMP structures from our CG model have been refined using AMBER9. In Fig. 4, the 10 ns MD simulation gives a RMSD curve (in coordinates of helix backbone atoms and from the PDB structure) ranged 4–5 Å for BR and 5–7 Å for HR. The deviation of our predicted structures of HBMPs could be significantly smaller (3–4 Å) if the parametric values in our CG model are tuned. The potential energy (model) curves describe the change of potential energy of MPs in refining the predicted CG structures, which decreases systematically with time from 2,200 to below 1,800 kcal/mol for BR and from 2,300 to below 1,800 kcal/mol for HR. The potential energy (native) curves describe the time evolution of MPs' potential energy, in which case the initial structure in the

Table 2 A comparison of our predictions and experimental data for the three angular parameters (Θ, Φ, Ω) of TMHs for HR

Angle	Helix						
	H1	H2	H3	H4	H5	H6	H7
Θ							
PDB	19.52	1.19	6.26	8.23	17.73	13.89	16.18
MC ($L = 26$)	4.31 (22.17)	2.19 (4.01)	27.52 (5.37)	2.44 (14.32)	26.87 (21.05)	2.38 (10.96)	0.38 (24.38)
Φ							
PDB	315.70	76.20	77.35	107.72	13.18	170.17	332.32
MC	2.29	134.65	91.67	335.18	148.40	319.71	80.21
Ω							
PDB	349.50	243.51	292.21	269.29	171.31	91.67	81.36
MC	70.29	243.50	243.51	256.11	238.35	56.15	82.51

The values of Θ predicted from MC simulations are for the case of $L = 32$, while its values in parentheses are results from MC simulations using $L = 26$

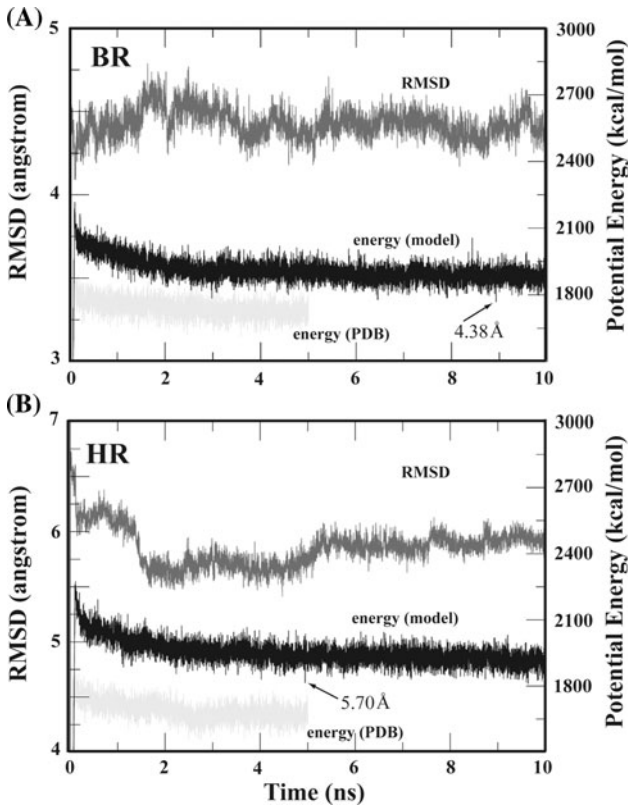

Fig. 4 RMSD (from the PDB structure) and potential energy of BR and HR calculated for MD simulations. Curves labeled RMSD show the RMSD of C_α atoms of the seven helices in the MD trajectory, curves labeled energy (model) are the potential energy curve obtained from the restrained MD simulation starting from MC predicted structure, and curves labeled energy (native) are the potential energy curve of a restrained MD simulation starting from the PDB structure

MD simulation of MPs is their PDB structure. The potential energy of our refined structures is about 150–200 kcal/mol higher than that of the corresponding PDB structures, which is slightly greater than the thermal fluctuation of the potential energy. However, the relaxation of MPs' structure in MD simulation is not fast enough to find the native structure in 10 ns. The lowest energy structure obtained during our MD refinement has a RMSD in coordinates of C_α atoms of 4.38 Å from the PDB structure for BR and 5.70 Å from the PDB structure for HR. These results are not as good as those predicted from our earlier model [14], but are comparable with predictions from other models. For example, Yarov-Yarovoy et al. used the Rosetta-Membrane method to build up the conformation of BR with a RMSD of 8.70 Å [27], while Kokubo and Okamoto [13] combined the CHARMM force field and REMC to obtain the lowest-energy structure of BR with a RMSD of 10.06 Å. In the present model, it is found that packing of TMHs is affected by the adopted values of vdW radii of amino acids. However, measured values of vdW radii of atoms (or residues) could differ significantly, depending on the measuring techniques and the chemical environment of atoms. This factor seems to be a major source of errors in predicting the packing of TMHs for BR and HR. Figure 5 shows the comparison of the MD refined structures (side view) of BR (a) and HR (b) with their PDB structures. Comparing Fig. 5 with Fig. 2, it is seen that the MD refinement of BR and HR improves our predictions in the packing, tilting, and orientation of TMHs. The RMSD of tilting angles decreases from 10.82° (CG structure) to 6.20° (MD refinement) for BR, and from 13.00° (CG structure) to 9.81° (MD refinement) for HR. The RMSD of orientation angles decreases from 108.11° (CG structure) to 65.72° (MD refinement) for BR, and from 103.95° (CG structure) to 43.28° (MD refinement) for HR. The RMSD of rotational angles is also improved after MD refinement, which slightly decreases from 59.32° (CG structure) to 54.82° (MD refinement) for BR and from 46.03° (CG structure) to 41.94° (MD refinement) for HR. The difficulty to refine the

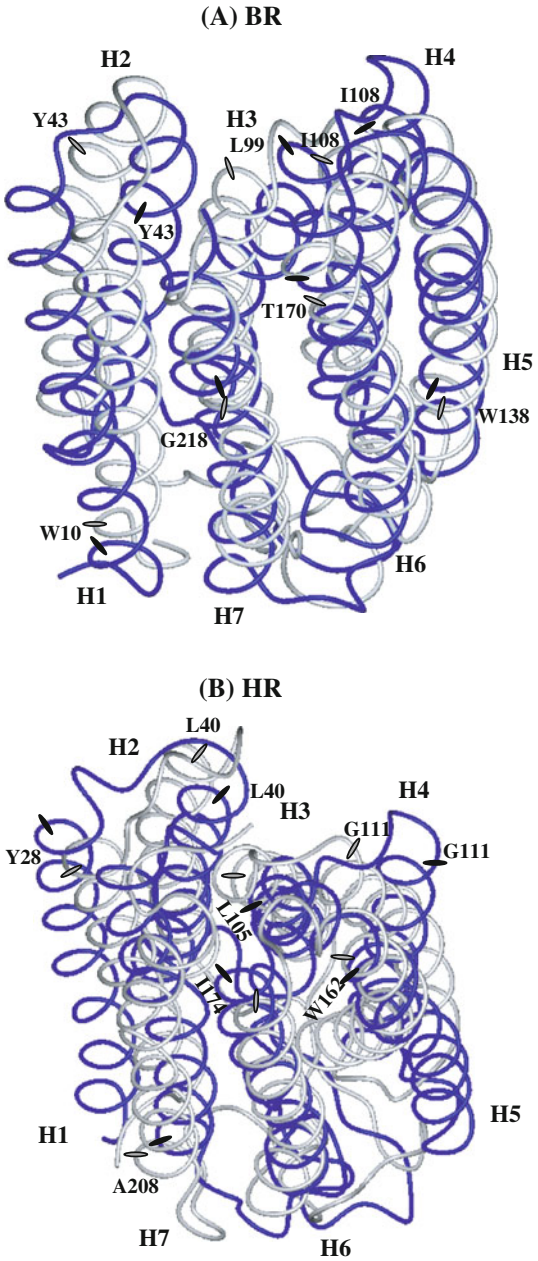

Fig. 5 Comparison of the PDB structure (*dark ribbons*) from our AA refined predictions (*grey ribbons*) for **a** BR and **b** HR. Here we have shown the location of one residue for each helix to visualize the prediction error in the rotational angle

rotational angle of TMHs comes from the fact that TMHs pack closely for low energy structures and the barrier to change the rotational angle of TMHs from one local minimum to another is rather high.

In our CG model, the energy of the predicted BR structure is −60.6, while it is −50.3 for the PDB structure with local energy minimization. The model energy is −70.1 for the predicted HR structure, and is −58.0 for the PDB structure of HR with local minimization. These results indicate that various parameters in the CG model are not optimized and the proposed united-residue model for HBMP folding only predicts reasonable HBMP structure with RMSD ranged 5–7 Å. To optimize the model parameters, it is necessary to minimize the energy difference between the lowest energy structure and the PDB structure for a set of HBMPs. However, many HBMPs, such as bovine rhodopsin, contain kinked helices and the present CG model needs to be modified to include kinked helices. Such a prediction of kink position and kink angle of a TMH can be performed using clustering algorithms from a large population of decoy structures of the TMH generated during computer simulations, and is under our current investigation.

Our physics-based CG model can be used to study the packing of HBMPs in the membrane without relying on extensive computation and statistic data from PDB. Although the structures of HBMPs have been of significant experimental interest, the resolution of our predicted HBMP structures is not high enough for practical biomedical applications. For BR and HR, the RMSD values of Θ, Φ, Ω of their predicted structures are about 10°, 100°, and 50°, respectively. The accuracy of Θ can be improved by choosing proper value of membrane thickness since the tilting angle is determined by the competition of E_{hw} and E_{hl}. Large membrane thickness tends to leave the helices parallel to the membrane normal, while small membrane thickness tends to tilt the helices away from the membrane normal. In the case of BR, the value of Θ_{RMSD} is reduced to 4° if a smaller membrane thickness ($L = 26$) is used. Nevertheless, there is no quick remedy to improve the accuracy in the prediction of Φ and Ω in the present model. The orientation angle Φ is mainly determined by the packing of helices (E_{vdw}) in our CG model, whose poor prediction results from the simplification of the side chain of amino acids by a sphere. The rotational angle Ω is mainly determined by the energy term E_{amp}, but the definition of proper contact between amphiphilic faces in the present model leaves a large uncertainty of the structure with the lowest value of E_{amp}. Refinement of the predicted CG structures using AA simulations can significantly improve the tilting and orientation angles, but not the rotational angle. We also note that our simple model does not consider helical kinks, which are present in many HBMPs. In principle the position and angle of helical kinks can be predicted using an AA model and the present CG model can be readily modified to include kinked helices.

Conclusions

To conclude, we have proposed a united-residue model for HBMP folding by including most dominant physical interactions. These physical interactions can be used to

understand intuitively the packing, tilting, orientation, and rotation of TMHs in the folding of HBMPs. These angular parameters (Θ, Φ, Ω) of TMHs derived from simulating the folding of HBMPs can be compared with results from site-directed dichroism experiments, which measure single amide I vibration modes corresponding to ^{13}C-labeled sites within HBMPs. Here, we have investigated the folding of BR, HR, and their subdomains with limited experimental input (secondary structure information only). The ground state structure of HBMPs in our CG model can be identified efficiently by REMC simulations using 10 replicas for 4-helix MPs and 15 replicas for 7-helix MPs. It is shown that the energy of MP structures in our protein model strongly correlates with their RMSD from the native structure. The predicted CG structures of HBMPs are reasonably consistent with their PDB structures. The RMSD in coordinates of C_α atoms of the ground state CG structure is 5.03 Å for BR and 6.70 Å for HR. Refinement of the predicted CG structures by AA MD simulations reduce the value of RMSD to 4.38 Å for BR and 5.70 Å for HR. Our MD simulations suggest that near native structures cannot be obtained by refining the predicted CG structures within 10 ns using AMBER9. We note that the deviation of our predicted structures of HBMPs could be significantly smaller (3–4 Å) if the parametric values in our CG model are tuned. However, with a priori knowledge of MP's folded structures, the expected RMSD from this CG model would be between 4 and 6 Å.

Acknowledgments This work is supported by the National Science Council of Taiwan under grant of no. NSC 97-2112-M-003-005-MY2.

References

1. Gerstein M (1998) Patterns of protein-fold usage in fight microbial genomes: a comprehensive structural census. Proteins 33(4): 518–534
2. Wallin E, von Heijne G (1998) Genome-wide analysis of integral membrane proteins from eubacterial, archaean, and eukaryotic organisms. Protein Sci 7(4):1029–1038
3. Krogh A, Larsson B, von Heijne G, Sonnhammer ELL (2001) Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J Mol Biol 305(3): 567–580. doi:10.1006/jmbi.2000.4315
4. White SH, Wimley WC (1999) Membrane protein folding and stability: physical principles. Annu Rev Biophys Biomol Struct 28:319–365
5. Drews J (2000) Drug discovery: a historical perspective. Science 287(5460):1960–1964
6. Filmore D (2004) It's a GPCR world. Mod Drug Discov 7:24–26
7. Bowie JU (2005) Solving the membrane protein folding problem. Nature 438(7068):581–589. doi:10.1038/nature04395
8. Milik M, Skolnick J (1992) Spontaneous insertion of polypeptide-chains into membranes—a Monte-Carlo model. Proc Natl Acad Sci USA 89(20):9391–9395
9. Chen CM (2001) Lattice model of transmembrane polypeptide folding. Phys Rev E 63(1):010901. doi:10.1103/PhysRevE.63.010901
10. Floriano WB, Vaidehi N, Goddard WA, Singer MS, Shepherd GM (2000) Molecular mechanisms underlying differential odor responses of a mouse olfactory receptor. Proc Natl Acad Sci USA 97(20):10712–10716
11. Dobbs H, Orlandini E, Bonaccini R, Seno F (2002) Optimal potentials for predicting inter-helical packing in transmembrane proteins. Proteins 49(3):342–349. doi:10.1002/prot.10229
12. Chen CM, Chen CC (2003) Computer Simulations of membrane protein folding: structure and dynamics. Biophys J 84(3):1902–1908
13. Kokubo H, Okamoto Y (2004) Self-assembly of transmembrane helices of bacteriorhodopsin by a replica-exchange Monte Carlo simulation. Chem Phys Lett 392(1–3):168–175. doi:10.1016/j.cplett.2004.04.112
14. Chen CC, Chen CM (2009) A dual-scale approach toward structure prediction of retinal proteins. J Struct Biol 165(1):37–46. doi:10.1016/j.jsb.2008.10.001
15. Chen CC, Wei CC, Sun YC, Chen CM (2008) Packing of transmembrane helices in bacteriorhodopsin folding: structure and thermodynamics. J Struct Biol 162(2):237–247. doi:10.1016/j.jsb.2008.01.003
16. Popot JL, Engelman DM (1990) Membrane protein folding and oligomerization: the two-stage model. Biochemistry 29(17): 4031–4037
17. Popot JL, Engelman DM (2000) Helical membrane protein folding, stability, and evolution. Annu Rev Biochem 69:881–922
18. Booth P, Curran A (1999) Membrane protein folding. Curr Opin Struct Biol 9(1):115–121
19. Pappu RV, Marshall GR, Ponder JW (1999) A potential smoothing algorithm accurately predicts transmembrane helix packing. Nat Struct Biol 6(1):50–55
20. Rees D, DeAntonio L, Eisenberg D (1989) Hydrophobic organization of membrane proteins. Science 245(4917):510–513
21. Ulmschneider MB, Sansom MSP, Di Nola A (2005) Properties of integral membrane protein structures: derivation of an implicit membrane potential. Proteins 59(2):252–265. doi:10.1002/prot.20334
22. Pilpel Y, Ben-Tal N, Lancet D (1999) kPROT: a knowledge-based scale for the propensity of residue orientation in transmembrane segments. Application to membrane protein structure prediction. J Mol Biol 294(4):921–935. doi:10.1006/jmbi.1999.3257
23. Adamian L, Nanda V, DeGrado WF, Liang J (2005) Empirical lipid propensities of amino acid residues in multispan alpha helical membrane proteins. Proteins 59(3):496–509. doi:10.1002/prot.20456
24. Adamian L, Liang J (2001) Helix–helix packing and interfacial pairwise interactions of residues in membrane proteins. J Mol Biol 311(4):891–907. doi:10.1006/jmbi.2001.4908
25. Choma C, Gratkowski H, Lear JD, DeGrado WF (2000) Asparagine-mediated self-association of a model transmembrane helix. Nat Struct Biol 7(2):161–166. doi:10.1038/72440
26. Gratkowski H, Lear JD, DeGrado WF (2001) Polar side chains drive the association of model transmembrane peptides. Proc Natl Acad Sci USA 98(3):880–885. doi:10.1073/pnas.98.3.880
27. Yarov-Yarovoy V, Schonbrun J, Baker D (2006) Multipass membrane protein structure prediction using Rosetta. Proteins 62(4):1010–1025. doi:10.1002/prot.20817
28. Stevens TJ, Arkin IT (1999) Are membrane proteins "inside–out" proteins? Proteins 36(1):135–143. doi:10.1002/(SICI)1097-0134(19990701)36:1<135::AID-PROT11>3.0.CO;2-I
29. Barth P (2010) Prediction of three-dimensional transmembrane helical protein structures. In: Frishman D (ed) Structural bioinformatics of membrane proteins. SpringerWienNewYork, New York

30. Lee J, Wu S, Zhang Y (2009) Ab initio protein structure prediction. In: Rigden DJ (ed) From protein structure to function with bioinformatics. Springer, New York
31. Trabanino RJ, Hall SE, Vaidehi N, Floriano WB, Kam VW, Goddard WA III (2004) First principles predictions of the structure and function of G-protein-coupled receptors: validation for bovine rhodopsin. Biophys J 86(4):1904–1921. doi:10.1016/S0006-3495(04)74256-3
32. Arkin IT, MacKenzie KR, Brunger AT (1997) Site-directed dichroism as a method for obtaining rotational and orientational constraints for oriented polymers. J Am Chem Soc 119(38):8973–8980
33. Zheng L, Herzfeld J (1992) NMR studies of retinal proteins. J Bioenerg Biomembr 24(2):139–146
34. Hansmann UHE (1997) Parallel tempering algorithm for conformational studies of biological molecules. Chem Phys Lett 281:140–150
35. Case DA, Cheatham TE III, Darden T, Gohlke H, Luo R, Merz KM Jr, Onufriev A, Simmerling C, Wang B, Woods RJ (2005) The Amber biomolecular simulation programs. J Comput Chem 26(16):1668–1688. doi:10.1002/jcc.20290
36. Nolting B (2005) Protein folding kinetics: biophysical methods. Springer, New York
37. Kyte J, Doolittle RF (1982) A simple method for displaying the hydropathic character of a protein. J Mol Biol 157(1):105–132. doi:10.1016/0022-2836(82)90515-0
38. Rose GD, Wolfenden R (1993) Hydrogen bonding, hydrophobicity, packing, and protein folding. Annu Rev Biophys Biomol Struct 22:381–415. doi:10.1146/annurev.bb.22.060193.002121
39. Huschilt J, Hodges R, Davis J (1985) Phase equilibria in an amphiphilic peptide-phospholipid model membrane by deuterium nuclear magnetic resonance difference spectroscopy. Biochemistry 24(6):10
40. Subczynski WK, Lewis RN, McElhaney RN, Hodges RS, Hyde JS, Kusumi A (1998) Molecular organization and dynamics of 1-palmitoyl-2-oleoylphosphatidylcholine bilayers containing a transmembrane alpha-helical peptide. Biochemistry 37(9):3156–3164. doi:10.1021/bi972148+
41. May S, Ben-Shaul A (1999) Molecular theory of lipid-protein interaction and the L-alpha-H-II transition. Biophys J 76(2):751–767
42. McLean LR, Hagaman KA, Owen TJ, Krstenansky JL (1991) Minimal peptide length for interaction of amphipathic alpha-helical peptides with phosphatidylcholine liposomes. Biochemistry 30(1):31–37
43. Kiyota T, Lee S, Sugihara G (1996) Design and synthesis of amphiphilic alpha-helical model peptides with systematically varied hydrophobic-hydrophilic balance and their interaction with lipid- and bio-membranes. Biochemistry 35(40):13196–13204. doi:10.1021/bi961289t
44. For simplicity, the HI vector of AFs only contains 5 elements. For an AF with 6 residues, its HI vector would be simplified to be <(HI1+HI2) ×0.5, (HI2+HI3) ×0.5, (HI3+HI4) ×0.5, (HI4+HI5) ×0.5, (HI5+HI6) ×0.5>
45. There are three conditions for the j1-th AF of helix i1 and the j2-th AF of helix i2 to have proper contact: (1) the distance between these two helices is less than 30 Å, (2) there is no other helix in between them, and (3) the distance between j1-th AF and j2-th AF are the shortest in all 16 AF pairs between these two helices
46. Nina M, Roux B, Smith JC (1995) Functional interactions in bacteriorhodopsin: a theoretical analysis of retinal hydrogen bonding with water. Biophys J 68(1):25–39. doi:10.1016/S0006-3495(95)80184-0
47. Baudry J, Crouzy S, Roux B, Smith JC (1999) Simulation analysis of the retinal conformational equilibrium in dark-adapted bacteriorhodopsin. Biophys J 76(4):1909–1917
48. Swendsen RH, Wang JS (1986) Replica Monte Carlo simulation of spin glasses. Phys Rev Lett 57(21):2607–2609
49. Kofke DA (2002) On the acceptance probability of replica-exchange Monte Carlo trials. J Chem Phys 117(15):6911–6914. doi:10.1063/11507776
50. Tajkhorshid E, Paizs B, Suhai S (1999) Role of isomerization barriers in the pK(a) control of the retinal Schiff base: a density functional study. J Phys Chem B 103(21):4518–4527
51. Tsong TY (1990) Electrical modulation of membrane proteins: enforced conformational oscillations and biological energy and signal transductions. Annu Rev Biophys Biophys Chem 19:83–106. doi:10.1146/annurev.bb.19.060190.000503
52. Anfinsen CB (1973) Principles that govern the folding of protein chains. Science 181(96):223–230
53. Huang KS, Bayley H, Liao MJ, London E, Khorana HG (1981) Refolding of an integral membrane protein. Denaturation, renaturation, and reconstitution of intact bacteriorhodopsin and two proteolytic fragments. J Biol Chem 256(8):3802–3809
54. London E, Khorana HG (1982) Denaturation and renaturation of bacteriorhodopsin in detergents and lipid-detergent mixtures. J Biol Chem 257(12):7003–7011