Available online at www.sciencedirect.com
ScienceDirect
Journal of Structural Biology 162 (2008) 237–247
Journal of Structural Biology
www.elsevier.com/locate/yjsbi

Packing of transmembrane helices in bacteriorhodopsin folding: Structure and thermodynamics

C.-C. Chen [a], C.-C. Wei [b], Y.-C. Sun [b], C.-M. Chen [a,*]

[a] Department of Physics, National Taiwan Normal University, 88 sec. 4, Ting-Chou Road, Taipei, Taiwan
[b] Department of Chemistry, National Taiwan Normal University, Taipei, Taiwan

Received 15 August 2007; received in revised form 3 January 2008; accepted 4 January 2008
Available online 11 January 2008

Abstract

We propose a coarse-grained (CG) model to study the native structure and physical properties of helical membrane proteins (HMPs) using off-lattice computer simulations. Instead of considering sequence heterogeneity explicitly, we model its effect on the packing of helices by employing a mean packing parameter r_0, which is calculated from an all-atom (AA) model. Specifically, this CG model is applied to investigate the packing of helices in bacteriorhodopsin (BR), and predicts the seven helix bundle structure of BR with a root mean square deviation (RMSD) in coordinates of helix backbone atoms (N, C, C_α) of 3.99 Å from its crystal structure. This predicted structure is further refined in an AA model by Amber and the refined structure has a RMSD (in coordinates of helix backbone atoms) of 2.64 Å. The predicted packing position, tilting angle, and orientation angle of each helix in the refined structure are consistent with experimental data and their physical origins can be well understood in our model. Our results show that a reasonably good structure of BR can be predicted by using such a dual-scale approach, provided that its secondary structure is known. Starting from a random initial configuration, the folded structure can be obtained in days using a regular desktop computer. Various thermodynamic properties of helix packing of BR are also investigated in this CG model.

Keywords: Transmembrane helix packing; Structure prediction; Thermodynamics; Monte-Carlo simulations; Molecular dynamics simulations; Coarse-grained model

1. Introduction

Membrane proteins (MPs) perform important and diverse functions in living cells, such as regulation, communication, and assisting the folding of other MPs (White and Wimley, 1999). They are partially buried in the non-polar environment of a lipid membrane, where the hydrophobic effect is absent. Since lipid tails are unable to form hydrogen bonds with proteins, the intra-chain hydrogen bonding along the backbone of proteins plays a significant role to form their native structure in a membrane. Only two main structural motifs are observed for MPs: membrane-spanning α-helix bundles and β-barrels, the former being predominant. Although analyses show that more than a quarter of all proteins coded in genomes are MPs (Gerstein, 1998; Wallin and von Heijne, 1998; Krogh et al., 2001), due to difficulties in crystallizing MPs, less than 200 MPs have known crystallographic structures so far (data obtained from http://blanco.biomol.uci.edu/Membrane_Proteins_xtal.html). Therefore, there exist great incentives for computational and theoretical studies of MPs (Milik and Skolnick, 1992; Chen, 2000; Floriano et al., 2000; Chen and Chen, 2003). As the information technology advances, computer assisted structure predictions and dynamic studies of MPs might serve as a powerful tool to understand the biological functions of MPs.

The retinal protein bacteriorhodopsin (BR) is a MP found in the purple membrane of *Halobacterium salinarium* and serves as a light-driven proton pump. It converts the energy of green light (of wavelength 500–650 nm) into an

* Corresponding author.
E-mail address: cchen@phy.ntnu.edu.tw (C.-M. Chen).

1047-8477/$ - see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.jsb.2008.01.003

electrochemical proton gradient, which is used by the cell to produce ATP. BR is generally assumed to be structurally similar to G protein-coupled receptors like (GPCR-like) MPs, which are the largest protein family and involved in all types of stimulus-response pathways. It is the focus of much interest and has become a paradigm for MPs in general and transporters in particular (Oesterhelt and Stoeckenius, 1973; Oesterhelt, 1976; Henderson, 1977; Dobbs et al., 2002; Kokubo and Okamoto, 2004). Its structure and function have been analyzed in great detail using a variety of experimental techniques. Structurally, it has a topology of seven transmembrane (TM) helices arranged in two arcs, an inner one containing helices B, C, and D and an outer one comprising helices E, F, G, and A. Between helices B, C, F and G, there is a TM channel, which accommodates a retinal to separate the extracellular half channel from the cytoplasmic half channel. London and coworkers have demonstrated denaturation and renaturation of BR under a wide variety of conditions (Huang et al., 1981; London and Khorana, 1982). Understanding the structure and thermodynamics of this membrane residing retinal protein is crucial to further investigate its biological functions.

The unique structural topology of BR serves as a simple model for the study of computer assisted structure predictions of helical MPs (HMPs). Many investigators have used the structure of BR (or bovine rhodopsin recently) as a template to build models of GPCR-like MPs (Pardo et al., 1992; Davies et al., 1996; Baldwin, 1998; Herzyk and Hubbard, 1998). An advance in modeling GPCR was made by a study from Baldwin (Baldwin, 1998), in which knowledge from sequence and homology as well as important conjectures of MP structure were used to construct a model for rhodopsin. However, due to the low sequence identity (less than 30%) between most GPCR-like MPs and rhodopsin (or BR), it should be noted that their arrangement of TM helices could be very different, which leads to an incorrect prediction of their structures. In fact, the average sequence identity of 99% of human GPCRs to bovine rhodopsin is lower than 20%. Another template based method for structure prediction is the threading method (Zhang et al., 2006), whose success depends on the completeness of the library of solved structures in the protein data bank (PDB) (Berman et al., 2000). An alternative method for structure prediction of MPs without using homology has been developed by Goddard III and coworkers (Trabanino et al., 2004; Kalani et al., 2004), which requires an initial input of the experimental electron density map of MPs. The *ab initio* approach is based on the global minimization of a physical potential energy function, which thus far has had limited success for small proteins (Simons et al., 1997; Liwo et al., 1999; Zhang et al., 2003; Kokubo and Okamoto, 2004). Several recent attempts have tried to integrate knowledge-based potentials with physical interactions to study the packing of MP helices, which demonstrate how the existing experimental structure database can be utilized to help the structure prediction of MPs (Chen and Xu, 2006; Yarov-Yarovoy et al., 2006). We note that most previous studies of MP structures (such as homology modeling and threading) show little interest on the thermodynamic hypothesis of protein folding, which is the main focus of this article. In addition, most previous methods in studying the structure of large proteins often require extensive computation. Our study intends to obtain a reasonably good structure within limited computational time by a dual-scale approach. In the first step, we will identify important physical interactions that would determine the structure of BR and construct our coarse-grained (CG) model. A serious search for the ground state of the system is then carried out by both the Monte-Carlo (MC) simulations and the parallel tempering (PT) algorithm (Hansmann, 1997). In the second step, an all-atom (AA) representation of our predicted CG structure of BR will be refined to get the final structure using Amber (Case et al., 2002).

It has been suggested that folding of many integral HMPs can be understood by a two-stage model (Popot and Engelman, 1990; Booth and Curran, 1999; Pappu et al., 1999; Popot and Engelman, 2000): 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. In this paper, based on the two-stage model, we demonstrate the feasibility of predicting the native structure of BR by finding the lowest energy channel state of a CG protein model using MC simulations. Dominant physical interactions responsible for the folding of BR are discussed in detail in the Model section. In the Simulation Methods section, we describe the procedure in carrying out a CG MC simulation of MPs, which are used to get kinetics and thermodynamics of MP folding from a random structure to the ground state structure. The ground state structure has been identified by the PT algorithm. The procedure to refine the predicted CG structure is described in the all-atom refinement section. Our results on the predicted structure and thermodynamics of BR folding are discussed in the Section 5.

2. Model

Previous studies using lattice MC simulations have shown the feasibility in correctly predicting the number of TM α-helices of a HMP and their location in sequence (Chen and Chen, 2003). TM helices will insert into the membrane in vivo either spontaneously or, more probably, via a translocon. In the latter case, our previous computer simulations suggest that the formation of TM helices is much faster than the packing of TM segments since only local interactions are involved in helix formation (Chen and Chen, 2003). As proposed in the two-stage model, we assume that the initial structure of BR contains seven rigid cylinders randomly residing in the membrane, which are constrained by flexible inter-helix loops. Each cylinder consists of 5 monomers and the size of each monomer varies from 4.5 to 7.5 Å depending on the number of residues

in each monomer. Excluded volume constraint is imposed on these cylinders, which means that no crossing of cylinders is allowed. These TM helices can be identified on the basis of hydrophobicity as described in our previous work (Chen and Chen, 2003) and their sequences adopted here are (WIWL) (ALGT) (ALMGL) (GTLY) (FLVK) (helix A: 10–30), (PDAKK) (FYAIT) (TLVPA) (IAFTM) (YLSML) (helix B: 37–61), (RYAD) (WLFT) (TPL) (LLLD) (LALL) (helix C: 82–100), (QGTIL) (ALVG) (ADGI) (MIGT) (GLVGA) (helix D: 105–126), (RFVW) (WAIS) (TAA) (MLYI) (LYVL) (helix E: 134–152), (EVAST) (FKVLR) (NVTV) (VLWSA) (YPVVW) (helix F: 166–189), and (NIET) (LLFM) (VLD) (VSAK) (VGFG) (helix G: 202–220). Here amino acids within the same group are considered to belong to the same monomer in the CG model. 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 3.8 Å). If the head-to-tail distance 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 the head-to-tail distance to be larger than the maximal length. These helices are allowed to diffuse in the membrane, as well as to tilt (Θ) and rotate (Φ) along the membrane normal direction, as shown in Fig. 1. The rotation of the helices around their main axis is irrelevant in this CG model due to the simplification of TM helices. Among various physical interactions, evidences show that the van der Waals (vdW) interaction and side-chain packing among TM helices mostly determine the tertiary structure of MPs (Popot and Engelman, 1990, 2000). Although interhelical 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 packing interaction between helices is expressed as

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

where e_1 is the strength of the packing interaction and r_0 determines the minimum of E_{packing}. In this representation, each helix is considered to be a rigid polymer cylinder and a monomer represents 3–5 residues. The distance between m-th monomer in helix i and n-th monomer in helix j is denoted by $r(m_i,n_j)$. The packing interaction in Eq. (1) is a sum of all vdW energy between monomers, which is approximated without sequence dependence in our CG model. A mean parameter r_0 is employed instead of considering sequence heterogeneity explicitly, which is some sort of average value evaluated using AA model of the helices. To estimate its value, we first construct those seven helices of BR individually as standard helices with the φ and ψ torsional angles of residues equal to -60 and -40 degree, based on the secondary structure of BR. Each standard helix is subject to an energy minimization using Amber7. The lowest vdW energy of each pair of neighboring helices is calculated as a function of their separation distance by varying their tilting (Θ), orientation (Φ), and rotational (Ω) angles at each distance, as shown in Fig. 2. The packing parameter r_0 (7.8 Å) in Eq. (1) is obtained by averaging its values predicted from these curves, which is consistent with its experimental measurement (7.9 Å) from the PDB structure of BR. We note that the value of r_0 is in general sequence dependent since different proteins would have different packing size. For example, it is calculated to be 8.3 Å for HR (experimental measurement is 8.24 Å) and 8.4 Å for SRII (experimental measurement is 8.5 Å). This packing interaction has determined the relative helix posi-

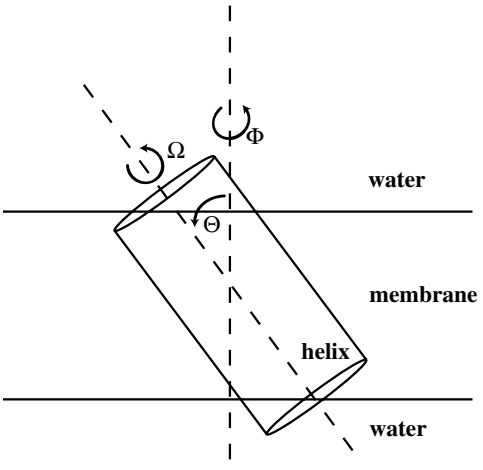

Fig. 1. A schematic representation of a transmembrane helix showing its tilting (Θ), orientation (Φ), and rotational (Ω) angles.

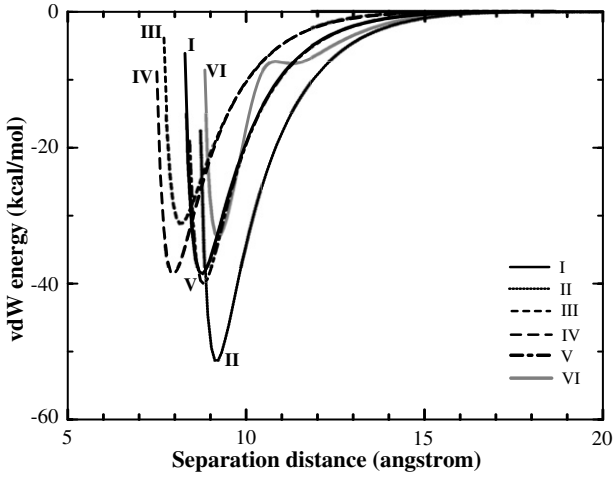

Fig. 2. Energy curves of vdW interaction for each pair of neighboring helices in BR as a function of their separation distance. The vdW energy curves are of the lowest energy at each separation distance by varying tilting, orientational, and rotational angles. Curve I is for the pair of helices A and B, curve II is for the pair of helices B and C, curve III is for the pair of helices C and D, curve IV is for the pair of helices D and E, curve V is for the pair of helices E and F, and curve VI is for the pair of helices F and G.

tions of several HMPs, provided that a suitable r_0 is obtained from an AA model calculation. We further note that the determination of r_0 relies on AA calculations for pairs of neighboring helices, which requires knowledge on the topology of helix packing. Since a slight change in the value of r_0 only changes the packing size but not the packing topology, a rough estimation of r_0 can be used to obtain the packing topology first. This packing topology will then allow a more accurate estimation of r_0.

The helix–water interaction $E_{\rm hw}$ can be modeled by using a rescaled Kyte-Doolittle hydropathy (HP) index (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 the water phase into their compressed gas phase (Kyte and Doolittle, 1982). Here, the HP index of residues is rescaled to be values between −1 and 1. In our CG model, the hydropathy index of each monomer corresponds to its cumulative hydropathy (CHP) index of associated residues, i.e., $\mathrm{CHP}_j = \begin{cases} \sum_i \mathrm{HP}_i, & \text{if monomer } j \text{ is in water} \\ 0, & \text{if monomer } j \text{ is in membrane} \end{cases}$, where i is the residue index in monomer j of a TM helix. Thus one can express $E_{hw} = e_2 \sum_j \mathrm{CHP}_j$, where j runs over all monomers in TM helices. 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 (Huschilt et al., 1985; Subczynski et al., 1998). We thus model the helix–lipid interaction by a tilting energy $E_{\rm hl} = e_3 \Sigma_i (1-\cos\Theta_i)$ of the helices in the membrane, where Θ_i is the tilting angle of the i-th helix. The tilting energy increases if a helix is tilted from the membrane normal, due to the increase in the contact between lipids and helix. In principle, the competition of the helix–water and helix–lipid interactions would determine the tilting angles of helices. A proper estimation of the value of e_3/e_2 would allow us to predict the tilting angles of helices correctly. To estimate its best value, we have calculated the tilting angles of helices for 12 HMPs in the PDB (1E12, 1F88, 1H68, 1JGJ, 1BL8, 1FX8, 1J4N, 1OED, 1PW4, 2B2F, 2GFP, 2IC8) by varying e_3/e_2 using MC simulations. The predicted tilting angles of helices are then compared with their values acquired from the PDB. As shown in Fig. 3, the value of e_3/e_2 is estimated to be about 0.7 by minimizing the RMSD of the helix tilting angles, $\Theta_{\rm rmsd} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\Theta_i - \Theta_i^0)^2}$, where Θ_i and Θ_i^0 are the predicted and acquired tilting angles of helices. In other words, if one uses $e_3/e_2 = 0.7$ to predict the helix tilting of the above HMPs, it is expected to obtain the best results. Among these 12 HMPs, the first four are GPCR-like MPs, 1BL8 and 1FX8 are ion channels, 1FX8 and 1J4N are aquaporin-like MPs, 1PW4, 2B2F, and 2GFP are transporter MPs, and 2IC8 belongs

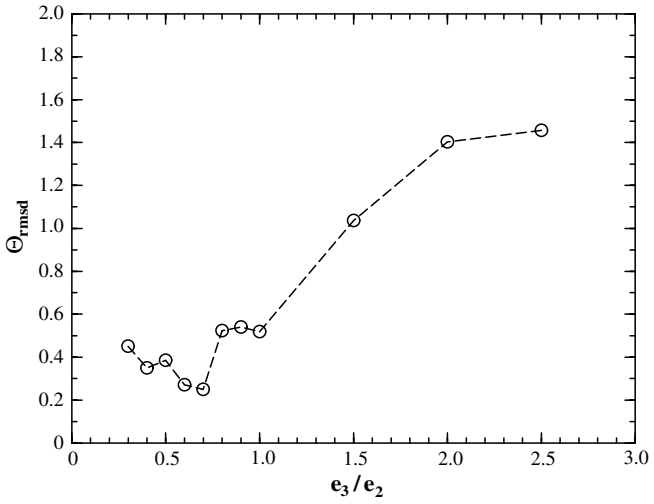

Fig. 3. RMSD of the helix tilting angles ($\Theta_{\rm rmsd} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\Theta_i - \Theta_i^0)^2}$) as a function of e_3/e_2. Here, Θ_i and Θ_i^0 are the predicted and acquired tilting angles of helices, for the following 12 helix-bundle MPs in the PDB: 1E12, 1F88, 1H68, 1JGJ, 1BL8, 1FX8, 1J4N, 1OED, 1PW4, 2B2F, 2GFP, 2IC8.

to intramembrane proteases. The reason that we include 4 GPCR-like MPs in this estimation is because they are the largest protein family known. We note that, to calculate the value of Θ_i with the lowest energy, MC simulations of each helix have been carried out for various values of e_3/e_2 in our CG model.

Finally we discuss the effect of the retinal in stabilizing the channel state over a hexagonal packing state. The hexagonal packing state, in which one helix of BR is surrounded by the other six helices, is expected to have the lowest packing energy since there are more contacts between helices (Kokubo and Okamoto, 2004). However, for such a hexagonal packing state, the retinal will be outside the helix bundle 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 (Nina et al., 1995; Baudry et al., 1999). In our CG model, the retinal is represented by a rod of length 12 Å and radius 1.6 Å, which is covalently bound to Lys-216 of the G-helix and allowed to move in the membrane. We thus model the non-covalent interaction between the retinal and its environment by a contact energy between retinal and helices, $E_{\rm contact} = e_4 \sum_{i=1}^{7} \varepsilon(\Delta r_i)$, where Δr_i is the shortest distance between the axes of retinal and i-th helix, and $\varepsilon(\Delta r_i)$ is 1 if Δr_i is between 6 and 9 Å and 0 otherwise. We note that $E_{\rm contact}$ is residue independent since it is used to model mainly the energy difference between retinal–water interaction and retinal–lipid interaction.

According to the thermodynamic hypothesis of protein folding, the native state of proteins is the global minimum of free energy (Anfinsen, 1973). To find the ground state structure of BR, the relevant physical quantity to be minimized in our model is the total energy $E = E_{\rm packing} + E_{\rm hw} + E_{\rm hl} + E_{\rm contact}$. In principle, several properties in our

model would be different for different MPs, including identification of each monomer in helices, the cumulative hydropathy index CHP_j, and the packing parameter r_0. We note that our model is a minimum model for MP folding, which has only been tested for 7TM receptors. A more comprehensive model and a more realistic representation of proteins are still under our investigation, in which case all residues of TM helices are explicitly considered.

3. Simulation methods

To study the structure and helix packing dynamics of BR, the simulation box is divided into three regions: a hydrophobic lipid layer (of thickness 26 Å) sandwiched by two water regions. The protein chain is represented by seven rigid cylinders (TM helices) located in the membrane phase and constrained by flexible inter-helix loops. Each cylinder consists of 5 monomers and the retinal is represented by a rod, which is covalently bound to Lys-216 of the G-helix and allowed to move in the membrane. The presence of this retinal molecule in the structure formation of BR will prevent the protein from forming a hexagonal packing structure. The helix packing of BR is simulated by the Metropolis MC algorithm in a continuum space at a constant temperature T (Chen and Higgs, 1998). At each instant, a cylinder is picked up at random and attempts to diffuse in an arbitrary direction, to tilt away from the z-axis, or to rotate along the z-axis. The rotation of the helices around their main axis is irrelevant in our CG model due to the simplification of TM helices. If any attempted move of cylinders satisfies the constraints of excluded volume and inter-helical loops, the move is accepted with probability $w = \min[1, \exp(-\Delta E/kT)]$ (the Metropolis criteria), where ΔE is the energy change of the system. In these simulations, we choose $e_1 = 0.25$, $e_2 = 1$, $e_3 = 0.7$, $e_4 = -0.5$, and $kT = 0.1$. The choice of these parameters is not unique. The RMSD of the ground state structure is unchanged if we slightly change the value of e_1 or e_4. However, for $e_1 > 0.4$ ($e_4 = -0.5$) or $e_4 > -0.3$ ($e_1 = 0.25$), the structure of the ground state has a RMSD (in coordinates of helix backbone atoms) of 5.67 Å from the native structure of BR. The thermal energy has no effect on the predicted structure as long as enough simulation time is used.

4. All-atom refinement

In addition to examining folded structure of BR using a CG model, an AA calculation is desired to see if one can obtain BR folded structure at atomic level. Amber was used to carry out this task. In particular, we will examine the ability of Amber in refining our predicted CG structures. To construct the AA representation of the predicted CG structure of BR from our MC simulations, we first build the seven helices of BR individually as standard helices with the φ and ψ torsional angles of residues equal to −60 and −40 degree, based on the secondary structure of BR. Each standard helix is subject to an energy minimization using Amber7. The seven energy-minimized helices constructed are used to replace the rigid cylinders of BR in the CG model by fitting the center of mass and the axis of helices. Moreover the rotation angle along the long axis (Ω) of each helix is chosen to align the most hydrophobic surface of helices to face the membrane core (Trabanino et al., 2004). Inter-helix loops are added arbitrarily by hand to connect consecutive helices without changing the CG structure of BR in the AA representation. We note that, however, this policy for the rotational angles of helices by considering hydrophobicity is not universal for all MPs. Other interactions might also contribute to the rotational angles of helices. The retinal is covalently bound to Lys-216 of the G-helix and allowed to move in the membrane. The atomic charges of the retinal and Lys-216 are taken from the work of Tajkhorshid and coworkers (Tajkhorshid et al., 1999). This structure is then refined by an energy minimization, which is proceeded with 5000 steps of steep descent method and 10,000 steps of conjugate gradient method. Here both water and membrane are implicit in our model. The hydrophobic core of membrane is treated as a dielectric medium of dielectric constant $\kappa = 2.5$ [its value is between 2 and 4 (Tsong, 1990)] and the dielectric constant of water is 80. As a first order approximation, we treat the environment of HMPs as a uniform dielectric, which screens out charges by a factor $1/\kappa$. More sophisticated model of the environment of HMPs can be adopted to improve the predicted structures of HMPs. For example, if a HMP contains a sub-structure in water, a better model of lipid and water would be required to get a good prediction. Other values of κ (2.0 and 3.0) are also used, but no substantial differences in the folded structure are observed. 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 are applied to the backbone atoms of all helices, including the torsional angles (φ and ψ) and the distance between N and O atoms of hydrogen bonds in the helices, as well as the residue position on the lipid mid-plane (LMP). 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. We note that typical values of temperature coupling parameter are between 0.5 and 5 ps for protein simulations, and too small values of this parameter might cause unrealistic fluctuations. For the purpose of structure refinement, any value in this range should make no significant difference.

5. Results and discussion

As suggested by Anfinsen, the native state of proteins is the global minimum of free energy (Anfinsen, 1973). Fig. 4 shows a typical run of helix packing of BR starting from a

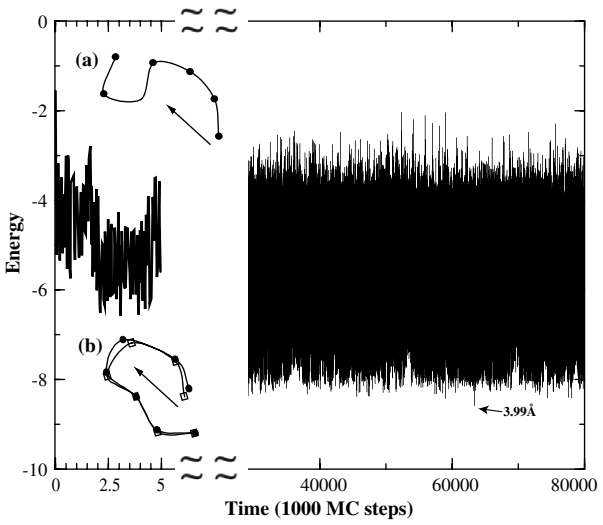

Fig. 4. Potential energy of BR as a function of time calculated from MC simulations. The energy of the chain as a function of simulation time is plotted for every 10 MC steps initially and for every 1000 steps otherwise. Only the highest and lowest energies during this time window are shown. The insets show the average helix positions of the initial structure (a) as well as the overlap of the predicted and crystal structures (b). Here, 1 MC step is defined to be the time period within which each helix attempts to move 50 times. The lowest energy of the energy curve is also noted with its RMSD (in coordinates of helix backbone atoms).

random initial configuration as shown in the inset (a), in which the energy of BR drops rapidly from −1.5 to −6.3 during the first 2500 MC steps and the lowest energy (−8.6) is observed at about 63 million MC steps. The ground state structure has a RMSD (in coordinates of helix backbone atoms: N, C, C_α) of 3.99 Å from the PDB structure of BR (Pebay-Peyroula et al., 1997). Here, 1 MC step is defined to be the time period within which each helix attempts to move 50 times. The lowest energy state is noted with its RMSD. The energy of the chain as a function of simulation time is plotted for every 10 MC steps initially and for every 1000 steps otherwise. Only the highest and the lowest energies within this time window are shown. Approximately, starting from the random initial configuration of MPs, it takes only a few days to find the ground state of our CG model by using a single core AMD Opteron CPU with a clock speed of 1.8 GHz. This ground state is consistent with the lowest energy state obtained using the parallel tempering (PT) algorithm, in which 10 replicas of temperature 0.08, 0.10, 0.13, 0.17, 0.21, 0.28, 0.36, 0.46, 0.59, and 0.77 are adopted for parallel simulations (Hansmann, 1997). The PT simulations consist of 10^7 MC steps and an exchange of conformations between pairs of replicas at neighboring temperatures is attempted simultaneously after 10^3 MC steps for each replica. In our MC simulations, 10 arbitrary initial states have been adopted in searching for the ground state structure of BR. Only 5 of them can find the ground state identified by PT within 10^8 MC steps. In the inset (b) of Fig. 4, the average helix positions of this ground state structure are compared to those of the crystal structure of BR on the LMP, which shows a remarkable consistency. For this ground state, the RMSD of helix positions at LMP (RMSD-LMP) is 1.23 Å and the RMSD of tilting angles (Θ_{rmsd}) is 8.28 degrees. Fig. 5 shows the relationship between energy and RMSD (in coordinates of helix backbone atoms) for 100 structures observed in our simulation, including 30 lowest energy structures and 70 randomly chosen high energy structures. It is clear that these low energy structures resemble the native structure and their RMSD (in coordinates of helix backbone atoms) is smaller than 5 Å. It is also clear that, from Fig. 4, the folded state corresponds to a basin of attraction on the energy surface since states of energy less than −8 (or RMSD in coordinates of helix backbone atoms less than 5 Å) appear repeatedly throughout our simulations. It is safe for us to note that the correspondence between our ground state structure and the native structure is consistent with the thermodynamic hypothesis, instead of a coincidence. In Fig. 6, curve 1 shows the values of Θ_{rmsd} as a function of time, which starts at the value of 19 degrees and fluctuates between 5 and 40°. Curve 2 shows the value of RMSD-LMP as a function of time, which starts at the value of 18 Å and fluctuates between 1 and 25 Å.

The 3D structures of BR predicted from our simulations are compared to its X-ray structure as shown in Fig. 7, in which the seven-helix structures of BR are depicted for the X-ray structure (black lines), MC prediction (light grey ribbons), and MD refinement (dark grey ribbons). The similarity among these three structures is visible. It can be seen, from Fig. 7, that the refinement of most helices in BR is apparent in the AA model. The overall RMSD in coordinates of helix backbone atoms of the lowest energy structure from the X-ray structure is 3.99 Å from the CG model prediction. This CG structure is then energy minimized in the AA representation using Amber7, which leads

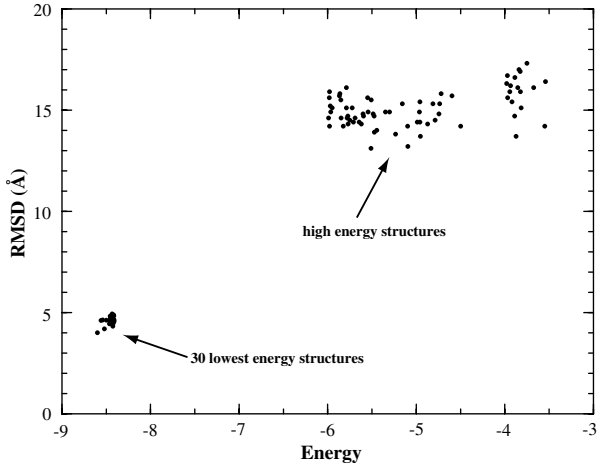

Fig. 5. A scattered diagram in the energy-RMSD (in coordinates of helix backbone atoms) plane for 100 observed structures, including 30 lowest energy structures and 70 randomly chosen high energy structures. The low energy conformations have a RMSD less than 5 Å, while the high energy states have a large RMSD about 15 Å.

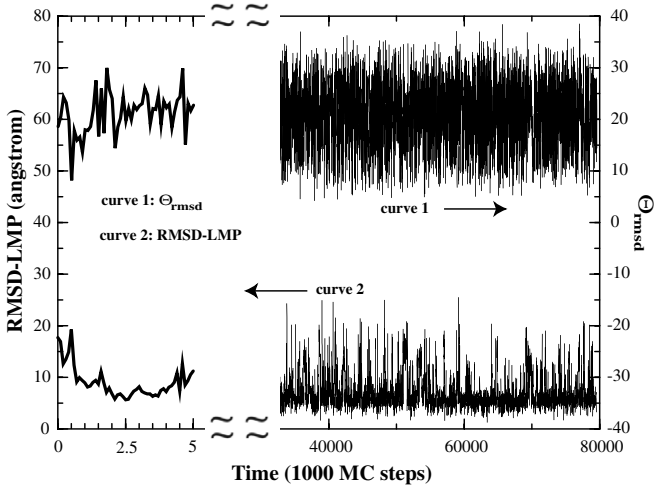

Fig. 6. Θ_{rmsd} and RMSD-LMP of BR as a function of time.

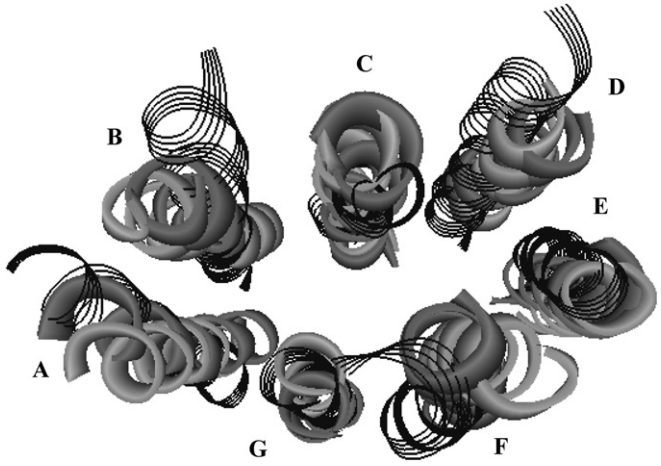

Fig. 7. Comparison of BR structures from MC prediction (light grey ribbons) and MD refinement (dark grey ribbons) with its X-ray structure (black lines). Significant improvement in the MD refinement is observed for helices A, B, E, and F.

to an energy-minimized structure of RMSD (in coordinates of helix backbone atoms) of 2.9 Å from the X-ray structure. With the above mentioned restraint sets, as shown in Fig. 8, the 5 ns MD simulation gives a RMSD (in coordinates of helix backbone atoms) curve (curve 1) ranged between 2.4 and 3.0 Å and the potential energy (curve 2) of BR decreases systematically with time from 1930 kcal/mol to below 1800 kcal/mol. The lowest energy structure observed in our MD simulation has a RMSD (in coordinates of helix backbone atoms) of 2.64 Å. For this refined structure of BR, histograms of the deviation of heavy atoms in BR from their native positions show a peak at around 2.5 Å, as demonstrated in Fig. 9. Here, Fig. 9(A) shows the distribution in the position deviation of atoms N, C, and C_α in the backbone while Fig. 9(B) shows the distribution in the position deviation of C_β atoms in the side chains. It is clearly seen from Fig. 9 (A) that most heavy atoms in the backbone have a deviation between 1 and 4 Å from their native position. On the other hand, C_β

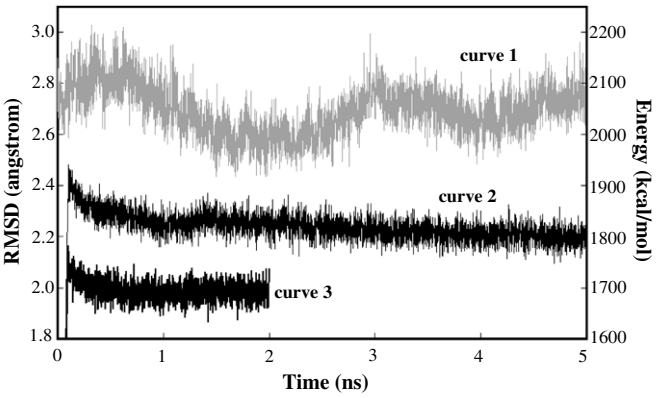

Fig. 8. RMSD (in coordinates of helix backbone atoms) and potential energy of BR calculated for MD simulations. Curve 1 is the RMSD of backbone atoms of the seven helices in the MD trajectory, curve 2 is the potential energy curve of BR obtained from the restrained MD simulation starting from MC predicted structure, and curve 3 is the potential energy curve of a restrained MD simulation starting from the X-ray structure.

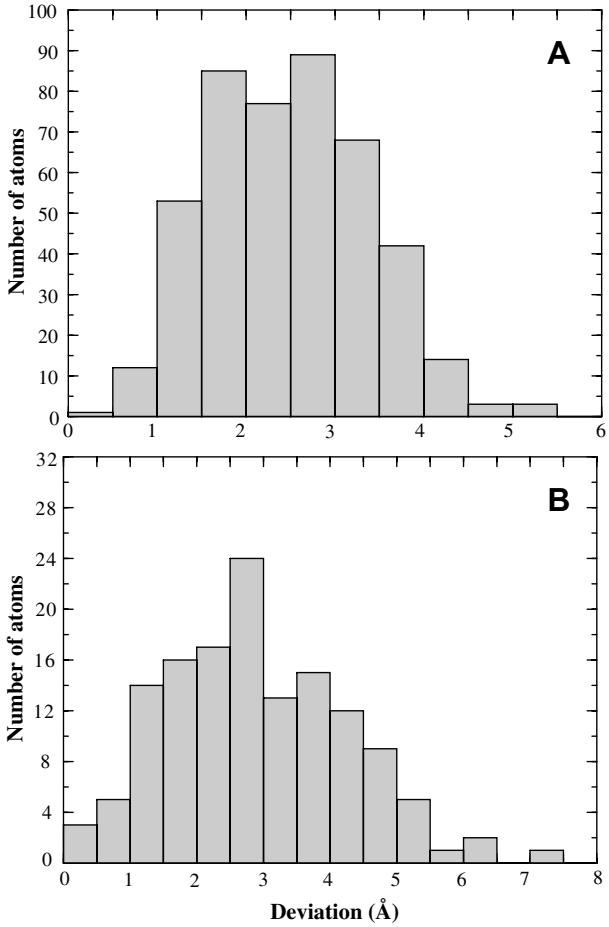

Fig. 9. Histograms of the deviation of heavy atoms in BR from their native positions: histogram (A) shows the distribution in the deviation of atoms N, C, and C_α in the backbone while histogram (B) shows the distribution in the deviation of C_β atoms in the side chains.

atoms in the side chains have a broader distribution in their deviation from the native positions, as shown in Fig. 9 (B). Roughly speaking, half of heavy atoms in the seven helices

have a deviation from their native position less than 2.6 Å. Our results are also consistent with Table 1 that most heavy atoms in helices D and F have smaller deviation from their native positions. We have also applied our dual-scale approach to predict structures of other HMPs. In the low resolution prediction, the RMSD value of helix backbone atoms from the crystal structure is 2.59 Å for HR and 3.12 Å for SRII. This RMSD (in coordinates of helix backbone atoms) value is improved to be 1.89 Å for HR and 1.92 Å for SRII in the high resolution prediction.

Table 1
A comparison of the predicted values of Θ and Φ of the seven helices of BR with their values calculated from the X-ray structure (1AP9)

BR	Θ			Φ		
Helix	PDB	MC	MD	PDB	MC	MD
A	27.76	17.53	20.86	176.78	355.49	169.52
B	21.04	14.49	14.61	129.42	256.88	141.61
C	14.81	30.78	10.63	79.75	200.38	111.74
D	22.40	16.13	18.05	133.29	260.10	122.28
E	0.00	4.38	10.40	Arbitrary	256.79	158.16
F	11.58	15.87	12.54	198.38	256.44	204.36
G	19.77	19.84	17.13	174.86	152.20	145.56

The Φ value of helix E is arbitrary since its Θ value is set to be 0.

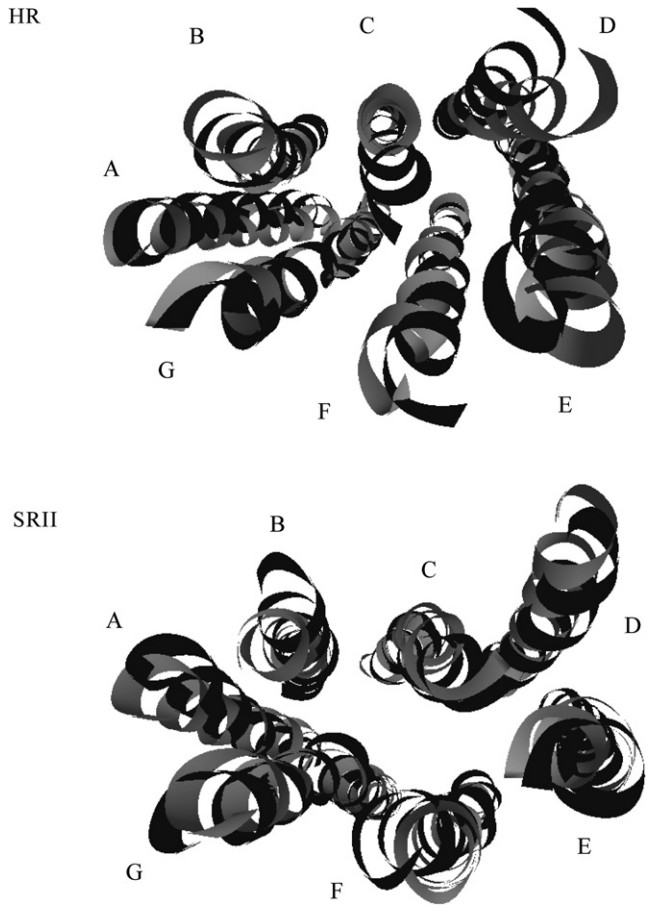

Fig. 10. Comparison of the PDB structure with our predicted structure for HR and SRII. Black ribbons represent helices in the PDB structure, and grey ribbons represent helices in our predicted structure obtained from the dual-scale approach.

Fig. 10 shows the overlap of the crystal structure and our predicted structure for HR and SRII. Since a detail analysis of other MPs is not the main focus here, it will be discussed elsewhere.

To examine if the structural deviation mainly arises from force field employed, we carry out simulation with the same protocol except that the starting structure is the crystal structure of BR. The MD simulation for 2 ns gives a low RMSD (in coordinates of helix backbone atoms) curve at about 1 Å. The potential energy curve versus time is shown as the curve 3 in Fig. 8, which is fluctuating around 1700 kcal/mol. This energy is lower than that of curve 2 by about 100 kcal/mol, due to lower vdW and electrostatic energies. The three curves in Fig. 8 show that the present computational protocol is able to obtain reasonable good structure, which is close to the tertiary structure of BR in PDB. Starting from the predicted CG structure, the lowest energy structure in our MD simulations of BR has a RMSD of 2.64 Å for helix backbone atoms and a RMSD of 3.19 Å for C_β atoms in the side chains. However, in a 5 ns simulation, sampling was not sufficient enough to refine folded structure to a RMSD (in coordinates of helix backbone atoms) value of 1 Å. Nevertheless, structure refinement was observed in helices A, B, E, and, F, as illustrated in Fig. 7. Future studies in enhancing sampling with larger computer power should allow us to study this slow motion dynamics.

The predicted tilting (Θ) and orientation (Φ) angles of helices in BR are compared to their values calculated from the crystal structure of BR in Table 1. As expected, the predicted tilting angles are consistent with their values acquired from the X-ray structure. For the predicted CG structure from MC simulations, the calculated value of Θ_{rmsd} is 8.28 degrees. On the other hand, the predicted orientation angles deviate from their experimental values substantially, due to the lack of information of the side-chain packing in our CG model. The calculated value of $\Phi_{\mathrm{rmsd}} = \sqrt{\frac{1}{7}\sum_{i=1}^{7}(\Theta_i - \Theta_i^0)^2}$ is about 100 degrees. This deviation in the predicted values of orientation angles can be greatly improved by including side-chain information into the CG model, or by refining the predicted structure in AA models. Here, a refinement from a 5 ns restrained MD simulation leads to a structure of $\Theta_{\mathrm{rmsd}} \approx 5.87$ and $\Phi_{\mathrm{rmsd}} \approx 20°$.

After predicting a reasonably good folded structure of BR without using 3D structure information of BR in the PDB, we believe that those interactions in our model should dominate the helix packing process of HMPs and that our model is suitable for studying their thermodynamic properties. Here, we use the multiple histogram method, as described in the Appendix A, to calculate various thermodynamic quantities of BR folding (Ferrenberg and Swendsen, 1989; Newman and Barkema, 1999). The relative entropy as a function of conformational energy is demonstrated in Fig. 11. The maximum of entropy observed in Fig. 11 can be understood straightforwardly.

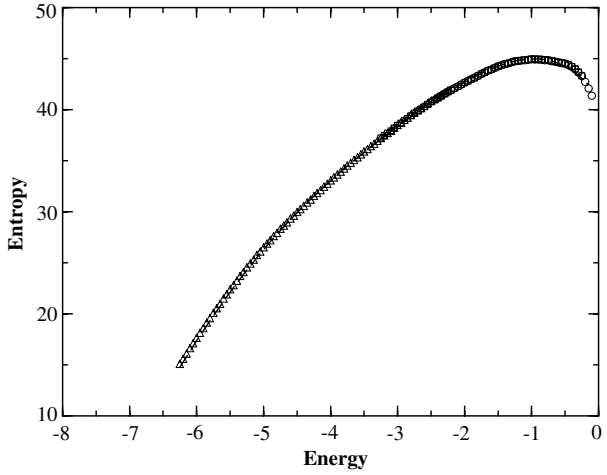

Fig. 11. Relative entropy as a function of conformational energy computed from MC simulations at three selected temperatures: $kT = 0.15$ (triangles), $kT = 0.5$ (squares) and $kT = 1000$ (circles).

At the limit of high energies, the protein has an extended alignment of helices and the associated entropy is small due to limited number of possible configurations. On the other hand, the protein has a unique alignment of helices toward the ground state energy, in which case the entropy of the protein is also small. It is rather natural to expect a maximum of entropy between these two extremes. As shown in Fig. 12, the total energy of BR as a function of temperature is calculated from the multiple histogram method using our MC simulations at five different temperatures ($kT = 0.1$, 0.25, 0.5, 0.75 and 1.0) and the specific heat of BR is shown in the inset. A pronounced single peak at $kT = 0.42$ is observed in the specific heat curve. Similar specific heat curve with a pronounced single peak has also been observed in experimental measurement of BR folding/unfolding using differential scanning calorimetry (Kahn et al., 1992). A direct comparison of our results with the experimental data would require a careful calibration of

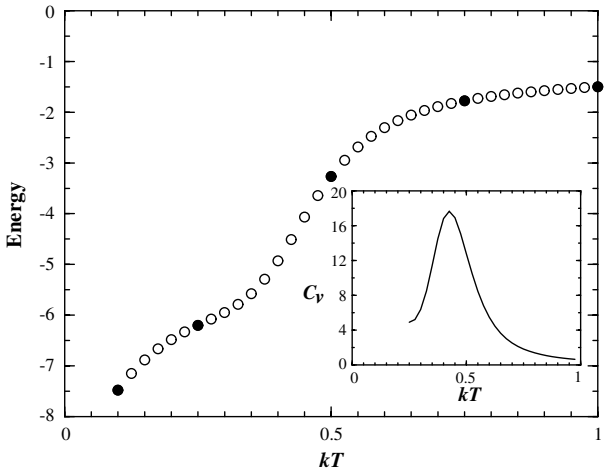

Fig. 12. Total energy of BR as a function of temperature calculated by the multiple histogram method. The inset displays the specific heat C_v, which has a pronounced peak at $kT = 0.42$.

our parameters using AA models, which is still under our investigation. It is interesting to note that our results are consistent with the experimental finding of Kahn et al. (Kahn et al., 1992), who have investigated three possible factors stabilizing helix–helix association in BR, including the extra-membranous loops connecting the helices, the binding interactions between retinal and helices, and the vdW packing of helices. They concluded that BR does not dissociate into separate fragments upon cutting inter-helical connections or removing retinal by bleaching, and the vdW packing effect will bring the helices together in a non-specific way. It is clear from our model that the inter-helical connections only serve as a constraining factor, instead of playing an important role in the stability of BR. Moreover the observed stable helix packing after bleaching implies the existence of the hexagonal packing state found in our simulation of BR without a retinal.

6. Conclusions

In summary, we have shown that the proposed simple model for the helix packing of HMPs is able to predict the structure of BR consisting of more than 250 amino acids and to calculate various equilibrium thermodynamic quantities. The structure of BR is efficiently predicted with a RMSD (in coordinates of helix backbone atoms) less than 4 Å using our CG model. The packing position and tilting angle of each helix of BR can be calculated and well understood in the CG model. It is found that the vdW interaction among helices and the retinal–environment interaction determine their relative position in membrane, while the competition between the helix–water and helix–lipid interactions determines the tilting angle of each helix. In the absence of the retinal, the lowest energy state is found to have a non-channel hexagonal structure, which is also found previously from an AA calculation (Kokubo and Okamoto, 2004). Further refinement of this predicted structure by AA models can substantially reduce the value of RMSD (in coordinates of helix backbone atoms) to 2.64 Å. Side-chain packing is found to be responsible for the orientation of helices. We note that, while the secondary structure information of BR can be derived by computer simulations based on the hydropathy index approach, it is considered to be known in this study. In addition, 3D structure information of BR in PDB is not required in this approach to predict helix packing of BR. Further computer simulations show a pronounced peak of the specific heat curve, which is consistent with the experimental observation. It should be noted that, although this approach obtains a reasonably good structure for several HMPs, our model is not designed to cover all MPs, particularly for beta barrel MPs. A more comprehensive model and a more realistic representation of proteins are still under our investigation. It is also important for an accurate prediction of the secondary structures of MPs in order to obtain a good prediction on the 3D structure of MPs. The construction of a general model for MP

structure prediction is so far unrealistic and remains to be one of the greatest challenges (Bowie, 2005).

Acknowledgments

The authors thank Dr. Tajkhorshid for providing the partial charge force field parameters for the retinal. This work is supported by the National Science Council of Taiwan under Grant No. NSC 96-2112-M-003-002. Y.-C. Sun acknowledges the partial support by the National Science Council of Taiwan under Grant No. NSC 93-2113-M-003-014. The authors thank the computer center at College of Science of NTNU and the NCHC for providing partial CPU time for the MD simulations.

Appendix A. Multiple histogram method

For a MC simulation with normal Boltzmann importance sampling, the probability $p(E)$ of generating a state with energy E on any one time-step is $p(E) \equiv N(E)/n = \rho(E)e^{-\beta E}/Z$, where the histogram $N(E)$ is the number of times out of n that the system is measured to have energy E, $\rho(E)$ is the density of states, $\beta=1/(kT)$ is the inverse temperature, and $Z = \sum_E \rho(E)e^{-\beta E}$ is the partition function. After performing a number of different simulations at a number of inverse temperatures β_i, we obtain a number of different estimates of the density of states, i.e., $\rho_i(E) = N_i(E)Z_i e^{\beta_i E} n_i^{-1}$. Since $\rho(E)$ depends only on the system we are studying, and not on the temperature, each of these estimates is an estimate of the same function. The best estimate of $\rho(E)$ can be obtained by performing the weighted average over the estimates $\rho_i(E)$, which gives more weight to individual estimates in regions where the corresponding histograms have more samples. As shown in the Appendix B, we express the estimate of $\rho(E)$ as $\rho(E) = \sum_i N_i(E)/\sum_j n_j Z_j^{-1} e^{-\beta_j E}$ (Newman and Barkema, 1999). The relative entropy as a function of conformational energy is obtained using $S(E) = \ln\rho(E)$. One can also obtain the conformational energy and the specific heat as a function of temperature from the partition function using $E = -\partial \ln Z/\partial\beta$ and $C_v = \partial E/\partial T$.

Appendix B. Estimate of the density of state

From the MC computer simulations of HMP helix packing, it is found that the errors $\Delta N_i(E)$ on the number of samples in each bin of a histogram are Poissonian since the measurements of the energy of the system can be considered to be independent. It is known that the error $\Delta N_i(E)$ of a Poissonian is equivalent to the square root of the average histogram $\overline{N_i(E)}$, i.e., $\Delta N_i(E) = \sqrt{\overline{N_i(E)}}$. Assuming that one makes a very large number of simulations all at the inverse temperature β_i and forms a histogram of each run, it is reasonable to approximate Poissonian errors as Gaussian. For a number of estimated density of states $\rho_i(E)$, the best estimate of $\rho(E)$ in the multiple histogram method can be expressed as $\rho(E) = \sum_i \rho_i(E)\sigma_i^{-2}/\sum_j \sigma_j^{-2}$, where $\sigma_i^2 = \sigma_i^2(E)$ is the variance of our estimates $\rho_i(E)$ of the density of states. In other words, the best estimate gives more weight to individual estimates in regions where the corresponding histograms have more samples. Since the Poissonian variation in $N_i(E)$ is the only source of error in $\rho_i(E)$, the error σ_i on $\rho_i(E)$ can be expressed as $\sigma_i = \Delta N_i(E) Z_i e^{\beta_i E} n_i^{-1} = \sqrt{\overline{N_i(E)}} Z_i e^{\beta_i E} n_i^{-1}$. Furthermore, if we were to do an infinite number of simulations at β_i, in principle the true density of states $\rho(E)$ can also be expressed as $\rho(E) = \overline{N_i(E)} Z_i e^{\beta_i E} n_i^{-1}$. Therefore, we have $\sigma_i^2 = \overline{N_i(E)}(Z_i e^{\beta_i E} n_i^{-1})^2 = \frac{\rho^2(E)}{\overline{N_i(E)}}$. The best estimate of $\rho(E)$ can then be calculated by performing the weighted average as $\rho(E) = \sum_i \rho_i(E)\rho^{-2}(E)\overline{N_i(E)}/\sum_j \rho^{-2}(E)\overline{N_j(E)} = \sum_i N_i(E)[Z_i e^{\beta_i E} n_i^{-1}\overline{N_i(E)}]/\sum_j \overline{N_j(E)}$. It is easy to verify that $\rho(E) = \sum_i N_i(E)/\sum_j n_j Z_j^{-1} e^{-\beta_j E}$ by substituting the expression for $\overline{N_j(E)}$.

References

Anfinsen, C., 1973. Principles that govern the folding of protein chains. Science 181, 223–230.

Baldwin, J.M., 1998. The probable arrangement of the helices in G protein-coupled receptors. EMBO J. 12, 1693–1703.

Baudry, J., Crouzy, S., Roux, B., Smith, J.C., 1999. Simulation analysis of the retinal conformational equilibrium in dark-adapted bacteriorhodopsin. Biophys. J. 76, 1909–1917.

Berman, H.M. et al., 2000. The protein data bank. Nucleic Acids Res. 28, 235–242.

Booth, P.J., Curran, A.R., 1999. Membrane protein folding. Curr. Opin. Struct. Biol. 9, 115–121.

Bowie, J.U., 2005. Solving the membrane protein folding problem. Nature 438, 581–589.

Case, D.A. et al., 2002. Amber 7.

Chen, C.-M., Higgs, P.G., 1998. Monte-Carlo simulations of polymer crystallisation in dilute solution. J. Chem. Phys. 108, 4305–4314.

Chen, C.-M., 2000. Lattice model of transmembrane polypeptide folding. Phys. Rev. E 63, 010901.

Chen, C.-M., Chen, C.-C., 2003. Computer simulations of membrane protein folding: structure and dynamics. Biophys. J. 84, 1902–1908.

Chen, Z., Xu, Y., 2006. Energetics and stability of transmembrane helix packing: a replica-exchange simulation with a knowledge-based membrane potential. Proteins 62, 539–552.

Davies, A., Schertler, G.F., Gowen, B.E., Saibil, H.R., 1996. Projection structure of an invertebrate rhodopsin. J. Struct. Biol. 117, 36–44.

Dobbs, H., Orlandini, E., Bonaccini, R., Seno, F., 2002. Optimal potentials for predicting inter-helical packing in transmembrane proteins. Proteins 49, 342–349.

Ferrenberg, A.M., Swendsen, R.H., 1989. Optimized Monte Carlo data analysis. Phys. Rev. Lett. 63, 1195–1198.

Floriano, W.B., Vaidehi, N., Goddard III, W.A., Singer, M.S., Shepherd, G.M., 2000. Molecular mechanisms underlying differential odor responses of a mouse olfactory receptor. Proc. Natl. Acad. Sci. USA 97, 10712–10716.

Gerstein, M., 1998. Patterns of protein-fold usage in eight microbial genomes: a comprehensive structural census. Proteins 33, 518–534.

Hansmann, U.H.E., 1997. Parallel tempering algorithm for conformational studies of biological molecules. Chem. Phys. Lett. 281, 140–150.

Henderson, R., 1977. The purple membrane from *Halobacterium halobium*. Annu. Rev. Biophys. Bioeng. 6, 87–109.

Herzyk, P., Hubbard, R.E., 1998. Combined biophysical and biochemical information confirms arrangement of transmembrane helices visible from the three-dimensional map of frog rhodopsin. J. Mol. Biol. 281, 741–754.

Huang, K.-S., Bayley, H., Liao, M.-J., London, E., Khorana, H.G., 1981. Refolding of an integral membrane protein. Denaturation, renaturation and reconstitution of intact bacteriorhodopsin and two proteolytic fragments. J. Biol. Chem. 256, 3802–3809.

Huschilt, J.C., Hodges, R.S., Davis, J.H., 1985. Phase equilibria in an amphiphilic peptide-phospholipid model membrane by deuterium nuclear magnetic resonance difference spectroscopy. Biochemistry 24, 1377–1386.

Kahn, T.W., Sturtevant, J.M., Engelman, D.M., 1992. Thermodynamic measurements of the contributions of helix-connecting loops and of retinal to the stability of bacteriorhodopsin. Biochemistry 31, 8829–8839.

Kalani, M.Y., Vaidehi, N., Hall, S.E., Trabanino, R.J., Freddolino, P.L., et al., 2004. The predicted 3D structure of the human D2 dopamine receptor and the binding site and binding affinities for agonist and antagonists. Proc. Natl. Acad. Sci. USA 101, 3815–3820.

Kokubo, H., Okamoto, Y., 2004. Self-assembly of transmembrane helices of bacteriorhodopsin by a replica-exchange monte carlo simulation. Chem. Phys. Lett. 392, 168–175.

Krogh, A., Larsson, B., von Heijne, G., Sonnhammer, E.L., 2001. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J. Mol. Biol. 305, 567–580.

Kyte, J., Doolittle, R.F., 1982. A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. 157, 105–132.

Liwo, A., Lee, J., Ripoll, D.R., Pillardy, J., Scheraga, H.A., 1999. Protein structure prediction by global optimization of a potential energy function. Proc. Natl. Acad. Sci. USA 96, 5482–5485.

London, E., Khorana, H.G., 1982. Denaturation and renaturation of bacteriorhodopsin in detergents and lipid-detergent mixtures. J. Biol. Chem. 257, 7003–7011.

Milik, M., Skolnick, J., 1992. Spontaneous insertion of polypeptide-chains into membranes—a Monte-Carlo model. Proc. Natl. Acad. Sci. USA 89, 9391–9395.

Newman, M.E.J., Barkema, G.T., 1999. Monte Carlo Methods in Statistical Physics. Clarendon Press, Oxford.

Nina, M., Roux, B., Smith, J.C., 1995. Functional interactions in bacteriorhodopsin: a theoretical analysis of retinal hydrogen bonding with water. Biophys. J. 68, 25–39.

Oesterhelt, D., Stoeckenius, W., 1973. Functions of a new photoreceptor membrane. Proc. Natl. Acad. Sci. USA 70, 2853–2857.

Oesterhelt, D., 1976. Bacteriorhodopsin as an example of a light driven proton pump. Angew. Chem., Int. Ed. Engl. 15, 17–24.

Pappu, R.V., Marshall, G.R., Ponder, J.W., 1999. A potential smoothing algorithm accurately predicts transmembrane helix packing. Nat. Struct. Biol. 6, 50–55.

Pardo, L., Ballesteros, J.A., Osman, R., Weinstein, H., 1992. On the use of the transmembrane domain of bacteriorhodopsin as a template for modeling the three-dimensional structure of guanine nucleotide-binding regulatory protein-coupled receptors. Proc. Natl. Acad. Sci. USA 89, 4009–4012.

Pebay-Peyroula, E., Rummel, G., Rosenbusch, J.P., Landau, E.M., 1997. X-ray structure of bacteriorhodopsin at 2. 5 angstroms from microcrystals grown in lipidic cubic phases. Science 277, 1676–1681.

Popot, J.-L., Engelman, D.M., 1990. Membrane protein folding and oligomerization: the two-stage model. Biochemistry 29, 4031–4037.

Popot, J.-L., Engelman, D.M., 2000. Helical membrane protein folding, stability, and evolution. Annu. Rev. Biochem. 69, 881–922.

Simons, K.T., Kooperberg, C., Huang, E., Baker, D., 1997. Assembly of protein tertiary structures from fragments with similar local sequences using simulated annealing and Bayesian scoring functions. J. Mol. Biol. 268, 209–225.

Subczynski, W.K., Lewis, R.N.A.H., McElhaney, R.N., Hodges, R.S., Hyde, J.S., Kusumi, A., 1998. Molecular organization and dynamics of 1-palmitoyl-2-oleoylphosphatidylcholine bilayers containing a transmembrane α-helical peptide. Biochemistry 37, 3156–3164.

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, 4518–4527.

Trabanino, R.J., Hall, S.E., Vaidehi, N., Floriano, W.B., Kam, V.W., et al., 2004. First principles predictions of the structure and function of G-protein-coupled receptors: validation for bovine rhodopsin. Biophys. J. 86, 1904–1921.

Tsong, T.Y., 1990. Electrical modulation of membrane proteins. Annu. Rev. Biophys. Biophys. Chem. 19, 83–106.

Wallin, E., von Heijne, G., 1998. Genome-wide analysis of integral membrane proteins from eubacterial, archaean, and eukaryotic organisms. Protein Sci. 7, 1029–1038.

White, S.H., Wimley, W.C., 1999. Membrane protein folding and stability: physical principles. Annu. Rev. Biophys. Biomol. Struct. 28, 319–365, For a review, see.

Yarov-Yarovoy, V., Schonbrun, J., Baker, D., 2006. Multipass membrane protein structure prediction using Rosetta. Proteins 62, 1010–1025.

Zhang, Y., Kolinski, A., Skolnick, J., 2003. TOUCHSTONE II: a new approach to ab initio protein structure prediction. Biophys. J. 85, 1145–1164.

Zhang, Y., DeVries, M.E., Skolnick, J., 2006. Structure modeling of all identified G protein-coupled receptors in the human genome. PLoS Comput. Biol. 2, 88–99.