Biochimica et Biophysica Acta 1818 (2012) 252–262

Contents lists available at ScienceDirect

Biochimica et Biophysica Acta

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

An ensemble dynamics approach to decipher solid-state NMR observables of membrane proteins ☆

Wonpil Im *, Sunhwan Jo, Taehoon Kim

Department of Molecular Biosciences and Center for Bioinformatics, The University of Kansas, 2030 Becker Drive, Lawrence, KS 66047, USA

ARTICLE INFO

Article history:
Received 15 June 2011
Received in revised form 22 July 2011
Accepted 30 July 2011
Available online 8 August 2011

Keywords:
Chemical shift anisotropy
Dipolar coupling
Deuterium quadrupolar splitting
Semi-static fitting model
Dynamic fitting model
Potential of mean force

ABSTRACT

Solid-state NMR (SSNMR) is an invaluable tool for determining orientations of membrane proteins and peptides in lipid bilayers. Such orientational descriptions provide essential information about membrane protein functions. However, when a semi-static single conformer model is used to interpret various SSNMR observables, important dynamics information can be missing, and, sometimes, even orientational information can be misinterpreted. In addition, over the last decade, molecular dynamics (MD) simulation and semi-static SSNMR interpretation have shown certain levels of discrepancies in terms of transmembrane helix orientation and dynamics. Dynamic fitting models have recently been proposed to resolve these discrepancies by taking into account transmembrane helix whole body motions using additional parameters. As an alternative approach, we have developed SSNMR ensemble dynamics (SSNMR-ED) using multiple conformer models, which generates an ensemble of structures that satisfies the experimental observables without any fitting parameters. In this review, various computational methods for determining transmembrane helix orientations are discussed, and the distributions of VpuTM (from HIV-1) and WALP23 (a synthetic peptide) orientations from SSNMR-ED simulations are compared with those from MD simulations and semi-static/dynamic fitting models. Such comparisons illustrate that SSNMR-ED can be used as a general means to extract both membrane protein structure and dynamics from the SSNMR measurements. This article is part of a Special Issue entitled: Membrane protein structure and function.

1. Introduction

Unlike globular proteins, the orientation of each transmembrane (TM) helix in membrane proteins with respect to membrane bilayers is one of the most crucial pieces of information for membrane protein structure and function. Such a TM helix orientation is generally described by TM helix tilt (with respect to the membrane normal) and rotation (around the helical axis) (Fig. 1). There are increasing number of experimental observations that the change in TM helix orientation is directly associated with membrane protein function [1]. For instance, bacterial mechanosensitive channels regulate the osmotic pressure by releasing small osmolytes through the channel opening under extreme hypo-osmotic conditions [2,3]; the channel gating involves reorientation and tilting of the TM helices [4,5]. The state-of-the-art X-ray crystallography has shown impressive successes in determining various multi-TM membrane proteins using detergents [6]. However, it is still challenging to obtain the structural (orientational) information of membrane proteins with single-pass TM helix, smaller number of TM helices, and their oligomers in bilayer environments. Note that membrane proteins with single-pass TM helix are abundant [7], and their association and conformational changes are often involved in TM-induced signaling and regulation; receptors with single-pass TM helix make up about 30% of the entire human TM receptors [8]. As an alternative and complementary approach to X-ray crystallography, various spectroscopic methods such as solid-state nuclear magnetic resonance (SSNMR) spectroscopy, electron paramagnetic resonance (EPR) spectroscopy, attenuated total reflection-Fourier transform infrared (ATR-FTIR) spectroscopy, circular dichroism (CD) spectroscopy, and fluorescence spectroscopy are widely used to study the TM helix orientation and its dependence on the bilayer property, such as lipid composition or hydrophobic thickness. Among these techniques, SSNMR is particularly attractive because it can determine the (atomic resolution) structure and orientation of membrane proteins and peptides in their native membrane environment.

Recent developments in SSNMR, especially selective isotopic labeling techniques and novel pulse sequences, have made it possible to collect high-resolution NMR spectra from aligned samples [9–12].

Abbreviations: CSA, chemical shift anisotropy; DC, dipolar coupling; DQS, deuterium quadrupolar splitting; EPR, electron paramagnetic resonance; GALA, geometric analysis of labeled alanine; PISA, polar index slant angle; PISEMA, polarization inversion spin exchange at the magic angle; SSNMR, solid-state nuclear magnetic resonance; SSNMR-ED, SSNMR ensemble dynamics

☆ This article is part of a Special Issue entitled: Membrane protein structure and function.

* Corresponding author. Tel.: +1 785 864 1993; fax: +1 785 864 5558.
E-mail address: wonpil@ku.edu (W. Im).

0005-2736/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.bbamem.2011.07.048

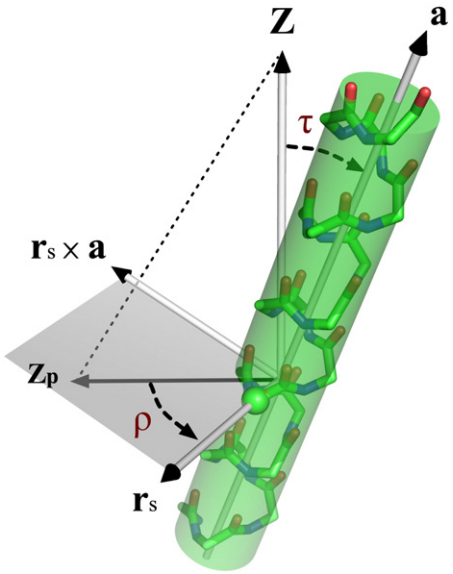

Fig. 1. Transmembrane helix orientation in terms of its tilt (τ) and rotation (ρ). τ is defined by the angle between the helical principal axis (**a**) and the unit vector along the *Z*-axis (**z**), which is parallel to the membrane normal by definition. ρ is defined by the angle between the projections of the *Z*-axis ($\mathbf{Z}_p = \mathbf{z} - (\mathbf{z}\cdot\mathbf{a})\mathbf{a}$) and the internal reference vector ($\mathbf{r}_s$) on the plane perpendicular to the helical principal axis. The detailed expressions can be found in our previous works [58,59]. The green sphere represents the internal reference atom. (VpuTM) The helical axis was defined by the Cα atoms of residue 7–25 of VpuTM and Cα atom of Ile^8 was used for the reference atom. To compare the rotation angle with the previous study (due to a different definition of ρ), 180° was added to ρ. (WALP23) Gly^1 has been widely used to define ρ (mostly when a rigid helix is used) [30,48,72], but it is not an appropriate choice for a dynamic system because of its flexibility at the terminal. Instead, we used Leu^{10} to avoid the flexibility problem and for facile comparison with other results based on Gly^1, because Leu^{10} is at the almost exact opposite position to Gly^1 on the helical wheel projection of WALP23.

In particular, polarization inversion spin exchange at the magic angle (PISEMA) experiment allows simultaneous measurement of pairwise ^{15}N chemical shift anisotropy (CSA) and ^1H–^{15}N dipolar coupling (DC) of isotopically labeled amide backbone of proteins [9]. The two-dimensional display of PISEMA spectrum shows a characteristic pattern, so-called polar index slant angle (PISA) wheels, which can be used to identify the helical segment and to determine its orientation [13,14]. SSNMR CSA and DC measurements have been applied to study the structure and the dynamics of various membrane proteins: gramicidin A (gA) [15–17], viral protein "u" (Vpu) [14,18], bacteriophage coat proteins [19,20], influenza M2 channel [21,22], mercury transporter [23], phospholamban [24], and antimicrobial peptides [25–27]. In addition, SSNMR deuterium quadrupolar splitting (DQS), another type of SSNMR observable, has been used to characterize the orientations and the dynamics of various synthetic single-pass TM helices, such as WALP and its variant peptides [28–31], gA [32], epidermal growth factor receptor (EGFR) [10], and antimicrobial peptide PGLa [33].

The most popular computational techniques to determine the TM helix orientation using the SSNMR observables (CSA, DC, and DQS) are based on geometric search approach or its variations [11,14,34,35]; i.e., experimentally derived structures are determined by searching a set of particular orientations (tilt and rotation) that have minimal root mean square deviations (RMSD) between the calculated and the experimental SSNMR observables using an ideal, rigid helix. The rationale behind these approaches is that TM helix segments are relatively rigid due to enhanced backbone hydrogen bonding in (low dielectric) bilayer environments [36,37]. The non-specified motional averaging is implicitly accounted by scaling the calculated observables using generalized order parameters, which makes these techniques quasi- or semi-static. These approaches are relatively straightforward, and one can quickly find a particular TM helix orientation that satisfies the experimental SSNMR observables. The extent of TM helix orientational variability is generally defined by the orientational space that yields the RMSD below the experimental error, e.g., less than 10 ppm for CSA, 1 kHz for DC, and 1 kHz for DQS [11,14–16,18,24,38]; the resulting structures generally suggest that the TM helices have very limited orientational freedom, e.g., ±1–2° in terms of the helix tilt angle.

However, increasing evidence from other experiments and molecular dynamics (MD) simulations indicates that orientational fluctuations of TM helices in bilayers are substantially larger, e.g., ±7–10° in terms of the helix tilt angle [38–46]. In addition, the TM helix tilt angle interpreted by a semi-static model using DQS is relatively small (~5° for WALP23 in DMPC) [30,31], and significantly different tilt angles (~30°) have been observed in a number of independent MD simulation studies [45,47–50]. Recently, several dynamic fitting models have been proposed to include TM helix orientational variability in the SSNMR structure determination [51–53]. In these approaches, additional fitting parameters are introduced to represent the orientational fluctuation of TM helix as a continuous distribution around the mean orientation; more fitting parameters could be introduced to capture complex motions, such as kink or peptide bond librations, other than the whole body motion [54]. However, addition of such fitting parameters and the underlying assumption of population distribution can be arbitrary and may not provide realistic dynamic information and orientational variability because TM helix orientations may deviate from a normal distribution. In general, extracting information about conformational variability directly from NMR and X-ray diffraction is not an easy task because the experimental observables represent time- and ensemble-averaged values [55–57].

Recently, we have developed an SSNMR ensemble dynamics (SSNMR-ED) technique with the purpose of extracting both TM helix orientation and intrinsic dynamics that are embedded in the SSNMR observables [58,59]. In SSNMR-ED, an ensemble of structures (i.e., replicas) is simulated in parallel MD simulations, and ensemble-averaged SSNMR observables across the replicas rather than a single individual structure are restrained to the experimental values. This approach is valid because the observables measured in SSNMR experiments are time- and ensemble-averaged properties. In addition, SSNMR-ED is advantageous because of its ability to generate an ensemble of structures (i.e., TM helix orientation distribution) that satisfies the experimental observables without additional fitting parameters. Similar techniques have been used in solution NMR structure determination [60,61] as well as in X-ray structure refinement [62], and the efficacy of refining structures based on the ensemble of property has been shown through the previous studies.

In this review, we aim to provide an overview on various computational techniques to determine TM helix orientations using the SSNMR observables (CSA, DC, and DQS) with an emphasis on SSNMR-ED. In the next section, the SSNMR observables are introduced in terms of their relationship with TM helix orientation, and various computational approaches for SSNMR structure determination are discussed together with possible effects of TM orientational (motional) averaging. Backgrounds and formalisms of SSNMR-ED are then given together with applications to the TM domain of Vpu (VpuTM) and WALP23. With the purpose of cross-validation and critical examination of each computational approach's efficacy, the SSNMR-ED results are compared with those from MD simulations and semi-static/dynamic fitting methods. The paper is then concluded with perspectives and future directions on computational studies of membrane protein structure, dynamics, and function in combination with experimental SSNMR observables.

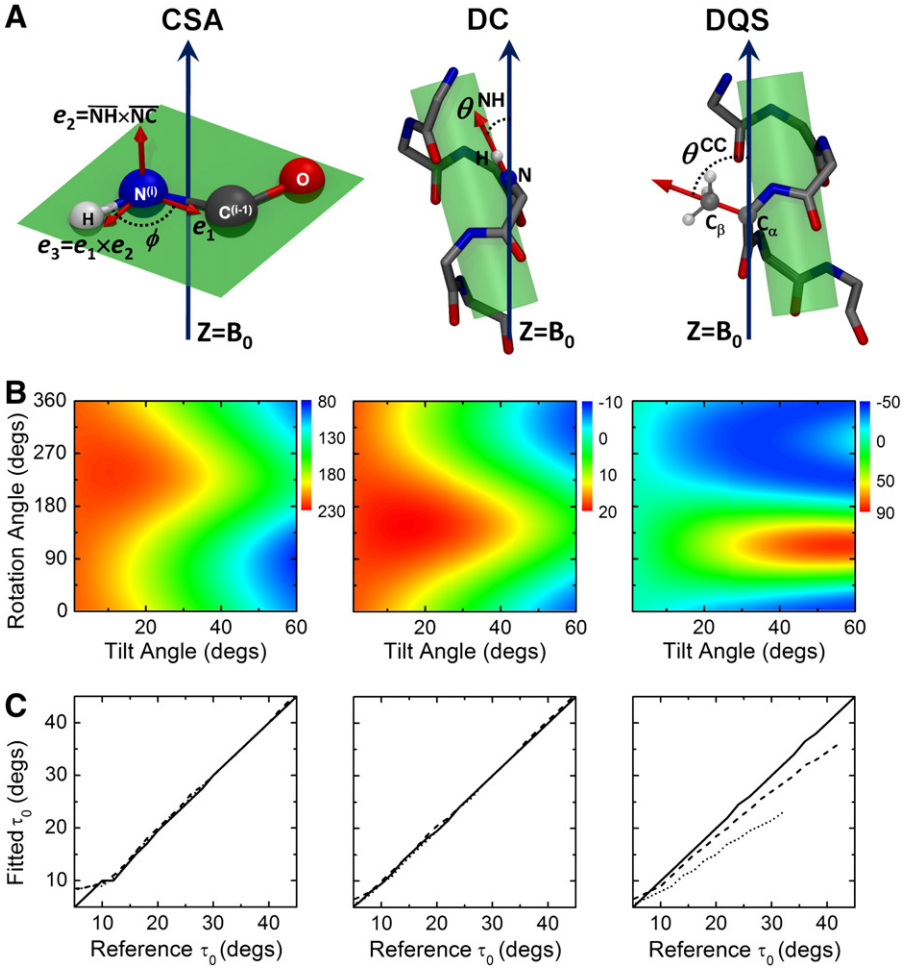

Fig. 2. Relationship between SSNMR observables and TM helix orientation. (A) Schematic representation of internuclear vectors that are used to calculate instantaneous CSA, DC, and DQS values. (B) CSA, DC, and DQS values (from left to right) of Ala^{10} residue with respect to τ and ρ of the model peptide ($(Ala)_{23}$). (C) Semi-static fitting was performed using synthetic CSA, DC, and DQS data, which were generated by assuming three different amplitude of motional average; i) $\delta_\tau = 0°$ and $\delta_\rho = 0°$ (solid line), ii) $\delta_\tau = 10°$ and $\delta_\rho = 30°$ (dashed line), and iii) $\delta_\tau = 10°$ and $\delta_\rho = 60°$ (dotted line). We excluded the τ in which the semi-static fitting did not yield the RMSD below the thresholds, e.g., 10 ppm for CSA, 1 kHz for DC, and 2 kHz for DQS.

2. SSNMR observables and TM helix orientation

2.1. Relationship between SSNMR observables and TM helix orientation

In this work, the direction of the external magnetic field ($\mathbf{B}_0$) is defined as the *Z*-axis. Fig. 2A shows the key variables that relate each of the SSNMR observables to TM helix orientation; the peptide-plane orientation with respect to the *Z*-axis for CSA (σ), the angle (θ^{NH}) between the ^{15}N–^{1}H vector and the *Z*-axis for DC (ν_{DC}), and the angle (θ^{CC}) between the $C_\alpha - C_\beta$ bond of labeled alanine (virtually the three C_β–D vectors due to their fast rotation) and the *Z*-axis for DQS (ν_{DQS}). These variables can be optimized to determine TM helix orientations by mapping a set of the SSNMR observables to a TM helix. In other words, TM helix orientation can be determined by searching the orientational space (i.e., tilt (τ) and rotation (ρ)) that minimizes the RMSD between the experimental (χ^{exp}) and the calculated (χ^{calc}) values;

$$\mathrm{RMSD}_\chi = \sqrt{\frac{1}{N_\chi}\sum_{i=1}^{N_\chi}\left(|\chi_i^{calc}| - \chi_i^{exp}\right)^2} \quad (1)$$

where χ is either σ (CSA), ν_{DC} (DC), or ν_{DQS} (DQS); n.b., the absolute value of χ^{calc} is used because the positive and the negative DC and DQS values are indistinguishable in SSNMR experiment. Each property is calculated using the key variables in Fig. 2A and the following formulas.

$$\sigma^{calc} = \sigma_{11}\hat{e}_{1,z}^2 + \sigma_{22}\hat{e}_{2,z}^2 + \sigma_{33}\hat{e}_{3,z}^2 \quad (2)$$

where σ_{nn} and $\hat{e}_{n,z}$ ($n = 1, 2, 3$) are the instantaneous magnitude and unit vector *Z*-component of chemical shift tensors. According to the rigid tensor approximation, $\hat{e}_2$ is defined by the cross product of $\mathbf{r}_{NC}$ and $\mathbf{r}_{NH}$. Then, $\hat{e}_1$ is defined by a rotation angle ϕ from $\mathbf{r}_{NH}$ on the peptide plane defined by N, C, and H atoms (Fig. 2A: CSA). Finally, $\hat{e}_3$ is defined by the cross product of $\hat{e}_1$ and $\hat{e}_2$. For the case of VpuTM in this work, σ_{11}, σ_{22}, and σ_{33} were set to 64, 77, and 222 ppm, respectively, and 108.5° was used for ϕ [63].

$$\nu_{DC}^{calc} = \frac{K_{DC}}{2}\left(3\cos^2\theta^{NH} - 1\right) \quad (3)$$

where the dipolar coupling constant (K_{DC}) is set to 21.016 kHz with $|\mathbf{r}_{NH}| = 1.05$ Å.

$$\nu_{DQS}^{calc} = \frac{3K_{DQS}}{4}\left(3\cos^2\theta^{CC} - 1\right) \quad (4)$$

where the quadrupolar coupling constant (K_{DQS}) is set to 56 kHz, which corresponds to one-third of a typical quadrupolar coupling constant due to the fast rotation of the methyl group [64].

To illustrate how TM helix orientation affects the calculated SSNMR observables, we simulated a simple polyalanine peptide with 23 residues, $(Ala)_{23}$, that has an ideal α-helical geometry ($\Phi = -57.8°$, $\psi = -47.0°$). CSA, DC, and DQS for each residue were calculated with respect to the TM helix's τ and ρ. While Fig. 2B shows the calculated CSA, DC, and DQS of a single residue (Ala^{10}), the distribution shows all possible values that are accessible for each residue because different residues have the values that are shifted along ρ at a given τ. Fig. 2B also illustrates that there are multiple solutions (i.e., TM helix orientation) for given SSNMR observables. The SSNMR observables have at least 2-fold degeneracy along ρ at a given τ; i.e., there are at least two ρ that yield the same CSA, DC, and DQS at a given τ. In particular, DC and DQS can have up to 8-fold degeneracy because of the following reasons. First, there are regions with negative values and such regions increase as τ increases; since only the absolute value is considered as in SSNMR measurement, there will be up to 4-fold degeneracy along ρ. Second, due to $\cos^2\theta$ in Eqs. (3)–(4), there will be additional 2-fold degeneracy along τ when DC and DQS values are less than 1/2 K_{DC} and 3/4 K_{DQS}. Given that DC and DQS experiments only measure the absolute values, this feature implies the possibility of a certain level of sign averaging when the TM helix dynamics (motional averaging) is considered (see Section 2.3). The spatial orientational distribution (so-called mosaic spread) due to sample alignments can also affect the experimental SSNMR observables [65,66]. In this work, however, such spatial orientational distributions are not considered, as the original experimental studies of WALP23 [30,31] and VpuTM [63] did not consider them.

2.2. Computational approaches to determine TM helix orientations

Various computational methodologies have been developed to efficiently translate the SSNMR observables into TM helix orientations. These approaches are based on TM helix orientational search (mostly τ and ρ) to minimize the RMSD between the experimental data and the calculated values. However, they differ in the following two aspects; (i) how to calculate the ensemble average of the SSNMR observables calculated by Eqs. (2)–(4), and (ii) how to perform the orientational/conformational search of TM helices. The ensemble average of the SSNMR observables can be written in the following general formula,

$$\left\langle \chi^{calc} \right\rangle_{ens} = \int d\tau d\rho \; \chi^{calc}(\tau,\rho) P(\tau,\rho) \tag{5}$$

where $\chi^{calc}(\tau,\rho)$ is either σ^{calc}, ν_{DC}^{calc}, or ν_{DQS}^{calc} at given τ and ρ, and $P(\tau,\rho)$ is the probability of a TM helix to be in certain τ and ρ.

2.2.1. Semi-static fitting approach

This is a simple, popular approach in which experimentally derived TM helix orientations are determined by searching a particular orientation (τ and ρ) of an ideal, rigid helix to minimize the RMSD between the calculated and the experimental SSNMR observables. As mentioned in Introduction, the term "semi-static" is used to broadly include some non-specified motion that is taken into account by simply multiplying a global order parameter (often less than 1) to Eqs. (2)–(4). In this approach, the ensemble average is neglected and each τ and ρ that yields the RMSD below the experimental error (e.g., less than 10 ppm for CSA, 1 kHz for DC, and 1 kHz for DQS [11,14–16,18,24,38]) is considered to be accessible. Previous studies showed that the accessible orientations determined by the semi-static fitting approach happen to be very narrow or delta function-like. The most probable reason is that the semi-static fitting methods do not consider any motional averaging.

2.2.2. Dynamic fitting approach

This method considers TM helix motional averaging by taking various forms of $P(\tau,\rho)$ in Eq. (5), such as the uniform or Gaussian distribution.

$$Uniform: P(\tau,\rho) = \begin{cases} \dfrac{1}{4\delta_\tau\delta_\rho} & \text{when } \tau_0-\delta_\tau < \tau < \tau_0+\delta_\tau \text{ and} \\ & \rho_0-\delta_\rho < \rho < \rho_0\delta_\rho \\ 0 & \text{otherwise} \end{cases} \tag{6}$$

$$Gaussian: P(\tau,\rho) = N_0 \exp\left(-\frac{(\tau-\tau_0)^2}{2\delta_\tau^2}\right)\exp\left(-\frac{(\rho-\rho_0)^2}{2\delta_\rho^2}\right) \tag{7}$$

where τ_0 and ρ_0 are the distribution center, δ_τ and δ_ρ control the width of the distribution, and N_0 is a normalization constant. This approach requires additional parameters to define the extent of TM helix orientational variability, i.e., δ_τ and δ_ρ in Eqs. (6)–(7). All these parameters are often adjusted to minimize the RMSD between the calculated and the experimental SSNMR observables. This approach is based on the fact that the correlation time of axial-motions (related to helix rotation) and off-axial motions (related to helix tilting) are significantly faster than the NMR timescale [17,32] and the MD simulation studies with various single-pass TM helices yielded more significant motions (fluctuations) in terms of τ and ρ than the semi-static approaches [38–46]. It should be noted that the ideal peptide geometry is assumed in general fitting methods; otherwise more fitting parameters are required to model a TM helix in different geometries such as kink or peptide bond librations [54].

2.2.3. Ensemble dynamics approach

This method is different from the semi-static and dynamic fitting approaches in that the orientational/conformational search of TM helices is performed by restrained MD simulation. In addition, the ensemble properties are calculated as an average of multiple conformers (i.e., replicas), and thus the motional averaging, if any, is explicitly considered by multiple conformers. It is implicitly assumed that the conformational ensemble average is the same as the time average, and the chemical exchange between different conformations are much faster than the NMR time scale. For the simplest approach, Eq. (5) becomes

$$\left\langle \chi^{calc} \right\rangle_{ens} = \frac{1}{N_{REP}} \sum_i^{N_{REP}} \chi_i^{calc}. \tag{8}$$

where N_{REP} is the number of replicas used in SSNMR-ED. The advantage of this method is that the assumption about the TM helix geometry or the underlying TM helix distribution, $P(\tau,\rho)$, is not required, but can be acquired after the structure determination process. Because of such explicit consideration of helix conformational flexibility in this approach, a global order parameter (often less than 1) used in the semi-static and dynamic fitting approaches is set to 1. In Section 3, we will explain SSNMR-ED in detail.

2.2.4. Free energy based approach

In principle, one can rigorously calculate $P(\tau,\rho)$ in Eq. (5) by calculating the two-dimensional (2D) potential of mean force (PMF) as a function of τ and ρ, $\mathcal{W}(\tau,\rho)$.

$$P(\tau,\rho) = \frac{\exp(-\mathcal{W}(\tau,\rho)/k_BT)}{\int d\tau' d\rho' \exp(-\mathcal{W}(\tau',\rho')/k_BT)} \tag{9}$$

where k_B is the Boltzmann constant. This approach is different from the aforementioned ones in that the TM helix orientation and dynamics are not obtained by the minimization of the RMSD between the experimental and the calculated SSNMR observables (by Eq. (5)), but simply become a natural product when the RMSD is reasonable. Although the 2D-PMF calculation in explicit membranes itself is still a

challenging task, we have shown that such an ensemble average based on Eqs. (5) and (9) can be done in an implicit membrane model; the calculated ensemble-averaged CSA of protegrin-1 (PG-1) antimicrobial peptide are shown to be comparable to experimental values, but the resulting distribution of PG-1 orientation is strikingly different from the semi-static interpretation [26,46]. In addition, Esteban-Martín et al. used empirical partition free energies of individual amino acid residues from aqueous solution to a membrane bilayer, and calculated the relative population of WALP23 orientations. They showed that the back-calculated DQS values had good correlation with experimental data only when the ensemble average was considered [67]. Alternatively, one can estimate $\mathcal{W}(\tau,\rho)$ from MD simulation trajectories, but this approach generally suffers from insufficient conformational sampling. Despite these difficulties, the PMF-based approach could be potentially powerful and useful in the near future with increasing computational resources and accuracy of molecular mechanics force fields.

2.3. Possible averaging of SSNMR observables due to TM helix orientational motions

The measured SSNMR observables can be affected by various TM helix motions, such as peptide-bond librations (on the timescale of 10^{-9} s), whole-body axial diffusion (10^{-8}–10^{-7} s), and whole-body off-axial reorientations (10^{-6}–10^{-5} s) [17,32]; whole-body axial diffusion collectively represents the motions related to helix rotation and whole-body off-axial reorientations are related to helix tilting. These timescales were estimated by fitting the parameters that represent distinct motional relaxation models. Our recent all-atom MD simulations of VpuTM in explicit lipid bilayers suggested faster timescale (~10^{-9} s) of τ and ρ fluctuations around its average τ and ρ [58]. Considering the limited amount of sampling and the absence of large scale off-axis fluctuation due to bilayer undulation, the faster timescale reported in the MD simulation likely suggests the upper bound of the fluctuation time scale. Nonetheless, such fast motions below the NMR time scale (10^{-3} s) indicate that the orientational averaging in SSNMR experiment is feasible and the TM helix dynamics needs to be considered when interpreting SSNMR observables. Although Strandberg et al. [52] and Esteban-Martín et al. [67] have done very careful analyses on the impact of possible TM orientational average, we have considered a simple case here, which is instructive and useful for discussion in the rest of the paper.

Let us consider the simple TM model of the $(Ala)_{23}$ ideal helix. For each τ (0° to 45° by 1°) and ρ (−180° to 180° by 2°), we first generated synthetic CSA, DC, and DQS data of each residue using Eqs. (2)–(4). Then, assuming a uniform distribution with various sets of δ_τ and δ_ρ in Eq. (6), we calculated the ensemble-averaged CSA, DC, and DQS of each residue at specific τ_0 (2° to 45° for every 2°) and ρ_0 (fixed at 180°) using Eq. (5). The resulting ensemble-averaged values were used as target (experimental) data in the semi-static model. Fig. 2C shows the relationship between the target τ_0 (for synthetic ensemble-averaged data) and the average τ within acceptable RMSD (based on the semi-static model) for various amplitudes of fluctuation ($\delta_\tau=\delta_\rho=0°$; $\delta_\tau=10°$ and $\delta_\rho=30°$; $\delta_\tau=10°$ and $\delta_\rho=60°$); a similar result was obtained when the target data were made by the Gaussian distribution instead of the uniform distribution. Clearly, when the amplitude of fluctuations is small, e.g., $\delta_\tau<10°$ and $\delta_\rho<30°$, the semi-static approach predicted τ near the target τ_0. However, such correspondence gets worse as the amplitude of rotational fluctuations increases in DQS. For example, the semi-static model yielded the lowest RMSD at $\tau=20°$ with the synthetic ensemble-averaged DQS at $\tau_0=30°$ with $\delta_\tau=10°$ and $\delta_\rho=60°$. This clearly indicates that if experimentally measured DQS includes a large motional averaging, the semi-static model may yield an incorrect TM orientation; the motional averaging in DQS at large τ also includes a certain level of sign averaging (Fig. 2B). For CSA and DC, the correspondence between the target τ_0 and the lowest RMSD τ is well maintained within reasonable TM motional fluctuations ($\delta_\tau<10°$ and $\delta_\rho<60°$). However, it was shown by Esteban-Martín et al. that CSA and DC are also susceptible to motional averaging and the target τ_0 will not be accurately predicted when a large amplitude of motional fluctuations is present ($\delta_\tau>10°$) [51]. More importantly, the semi-static model was not able to reproduce the orientational variability imposed in the synthetic data, suggesting that if measured CSA and DC include a certain level of motional averaging, the semi-static model may not capture such orientational variability despite correct average orientation.

These illustrations clearly indicate the impact of possible TM dynamics (motional averaging) on measured SSNMR observables. In this context, the main questions are what is the nature of TM dynamics and how one can explore such dynamics based on the SSNMR observables, which were the motivations for the development of the SSNMR-ED technique.

3. SSNMR ensemble dynamics

3.1. Backgrounds and formalisms

The ensemble dynamics simulation is designed to perform the parallel MD simulation with a certain number of replicas (N_{REP}) using a set of biased potentials (U_χ) that restrains the ensemble-averaged property ($\langle\chi^{calc}\rangle_{ens}$) to the experimental value (χ^{exp}); U_χ is given by

$$U_\chi = N_{REP}\sum_{i=1}^{N_\chi} k_\chi\left(\left|\left\langle\chi_i^{calc}\right\rangle_{ens}\right|-\chi_i^{exp}\right)^2 \quad (10)$$

where N_χ is the number of target experimental observables and k_χ is the force constant.

Based on the CSA and DC restraint potentials that Lee et al. developed for single-conformer SSNMR structure determination [68], we have developed the ensemble CSA, DC, and DQS restraint potentials by calculating the ensemble-averaged property and distributing the ensemble-averaged forces over N_{REP} replicas at each simulation time step. Based on Eqs. (2)–(4) and Eq. (8), the ensemble-averaged CSA, DC, and DQS are calculated by the following formalisms; for simplicity, we dropped the index i for ith observable in Eq. (10).

$$\begin{aligned}\left\langle\sigma^{calc}\right\rangle_{ens} &= \sigma_{11}\left\langle\hat{e}_{1,z}^2\right\rangle+\sigma_{22}\left\langle\hat{e}_{2,z}^2\right\rangle+\sigma_{33}\left\langle\hat{e}_{3,z}^2\right\rangle \quad (11)\\ &= \frac{\sigma_{11}}{N_{REP}}\sum_{m=1}^{N_{REP}}\hat{e}_{1,z,m}^2+\frac{\sigma_{22}}{N_{REP}}\sum_{m=1}^{N_{REP}}\hat{e}_{2,z,m}^2+\frac{\sigma_{33}}{N_{REP}}\sum_{m=1}^{N_{REP}}\hat{e}_{3,z,m}^2\end{aligned}$$

$$\left\langle\nu_{DC}^{calc}\right\rangle_{ens} = \frac{K_{DC}}{2}\left\langle 3\cos^2\theta^{NH}-1\right\rangle = \frac{1}{N_{REP}}\frac{K_{DC}}{2}\sum_{m=1}^{N_{REP}}\left(3\cos^2\theta_m^{NH}-1\right) \quad (12)$$

$$\left\langle\nu_{DQS}^{calc}\right\rangle_{ens} = \frac{3K_{DQS}}{4}\left\langle 3\cos^2\theta^{CC}-1\right\rangle = \frac{1}{N_{REP}}\frac{3K_{DQS}}{4}\sum_{m=1}^{N_{REP}}\left(3\cos^2\theta_m^{CC}-1\right). \quad (13)$$

The total potential energy (U_{TOTAL}) of the ensemble system is then expressed as

$$U_{TOTAL} = U_{CHARMM}+U_{CSA}+U_{DC}+U_{DQS} \quad (14)$$

where U_{CHARMM} is the standard CHARMM potential energy of the ensemble system, based on our implementation in CHARMM [69]. We have checked the numerical accuracy of the CSA, DC, and DQS ensemble restraint potentials; by definition, the total energy of all the replicas (not the total energy of each replica) was well converged during the NVE dynamics (Fig. 3). In summary, the ensemble-averaged SSNMR observables are calculated by taking averages across all replicas in SSNMR-ED (Eqs. (11)–(13)). Therefore, the motional/sign averaging, if any, is explicitly considered and included in the resulting TM helix orientation distribution that satisfies the experimental SSNMR observables.

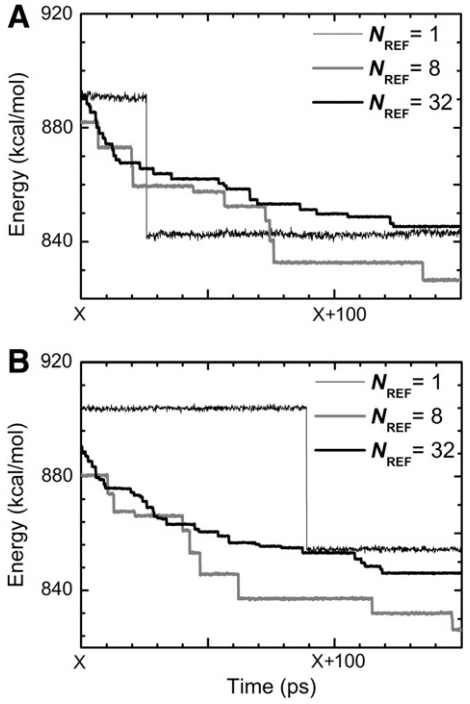

Fig. 3. Total energy profiles of all the replicas (when $N_{REP} = 1, 8$, and 32) with (A) CSA and (B) DC restraint potentials during NVE dynamics. *X* represents arbitrary time point during 1-ns simulation. A sudden energy changes are due to velocity reassignment during the dynamics.

The SSNMR-ED simulation protocols in implicit membrane models are well described in References [58,59] and have been successfully applied to VpuTM (with CSA and DC restraints) and WALP23 (with DQS restraint). Therefore, in the following subsections, we highlight the main findings from the SSNMR-ED simulations to illustrate the efficacy of SSNMR-ED in extracting the TM helix orientation and dynamics possibly embedded in the experimental SSNMR observables.

3.2. VpuTM

Vpu is a small membrane protein of 81 residues, encoded in the HIV-1 virus. The native protein is composed of a hydrophobic TM and amphipathic C-terminal domains. Park and Opella have measured CSA and DC of the TM domain (VpuTM) in three different lipid environments, and showed that the TM helix's τ increases as the bilayer thickness decreases [63]; 18° in DOPC/DOPG, 27° in DMPC/DMPG, and 35° in DLPC/DLPG 9:1 mixed bilayers (hereafter, for simplicity, DOPC/DOPG, DMPC/DMPG, and DLPC/DLPG are referred to as DOPC, DMPC, and DLPC, respectively).

Fig. 4A and B shows the RMSD of CSA and DC of the structure ensembles determined by the SSNMR-ED simulation in DMPC. Regardless of N_{REP}, the RMSD of CSA and DC are well below the acceptable range (i.e., less than 10 ppm for CSA and 1 kHz for DC), indicating that the generated structure ensemble agrees well with the SSNMR experimental observables. Interestingly, as N_{REP} increases, the RMSD of CSA and DC decreases, suggesting that the resulting structure ensembles with more replicas better represent the experimental observables. We have cross-validated that such observations are not the consequence of reduction in a data-to-parameter ratio (see Section 4.1). While the RMSD indicate good agreement with experimental data, the calculated ensemble structures show striking differences in orientational variability with different N_{REP}. Fig. 4C and D shows the τ and ρ of the ensemble structures. When $N_{REP} = 1$, the average τ and ρ are $29 \pm 1°$ and $309 \pm 9°$, respectively. These values agree well with the results from the semi-static method ($\tau = 27°$ and $\rho = 345°$) [63], considering the fact that conformational flexibility is explicitly included in our calculations even with $N_{REP} = 1$. When $N_{REP} \geq 2$, the ensemble-averaged τ and ρ remain similar values, but the variations in τ and ρ are increased to about 8–9° and 50° in terms of standard deviation of the structures from 250 independent SSNMR-ED simulations. This clearly implies that the experimental CSA and DC may have already included such extents of motional averaging, discussed in Section 2.3. The MD simulation in explicit bilayers shows a similar level of orientational variability, supporting that the resulting TM orientation from SSNMR-ED is physically relevant (see Section 4.2). Such increases in orientational variability are also observed in other structure ensembles from different lipid bilayer systems, as shown in Fig. 4E and F; the ensemble-averaged τ and ρ in DLPC and DOPC also match well with the values derived from the semi-static method [63]. Clearly, in accordance with the hydrophobic mismatch concept [45,70,71], VpuTM's τ increases as the bilayer hydrophobic thickness decreases in order to maximize the match between the TM helix's hydrophobic length and the bilayer hydrophobic thickness.

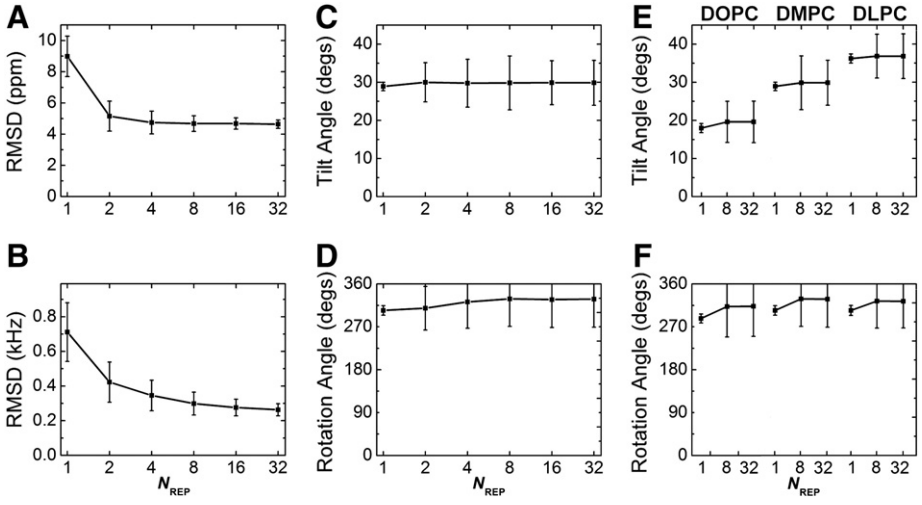

Fig. 4. Validation of VpuTM structure ensemble determined by SSNMR-ED. (A and B) CSA and DC RMSD from the experimental observables as a function of N_{REP} in a DMPC bilayer environment. (C and D) τ and ρ of VpuTM in DMPC. (E and F) VpuTM orientation determined by SSNMR-ED in DOPC, DMPC, and DLPC bilayers. The average τ from the SSNMR-ED are $17.7 \pm 1.3°$ (DOPC), $28.9 \pm 1.1°$ (DMPC), and $36.2 \pm 1.2°$ (DLPC), and the average ρ from the SSNMR-ED are $299 \pm 10°$ (DOPC), $309 \pm 10°$ (DMPC), and $304 \pm 10°$ (DLPC) when $N_{REP} = 1$. The average τ from the SSNMR-ED are $18.9 \pm 9.0°$ (DOPC), $29.5 \pm 8.7°$ (DMPC), and $36.8 \pm 5.8°$ (DLPC), and the average ρ from the SSNMR-ED are $336 \pm 64°$ (DOPC), $330 \pm 55°$ (DMPC), and $323 \pm 56°$ (DLPC) when $N_{REP} = 8$.

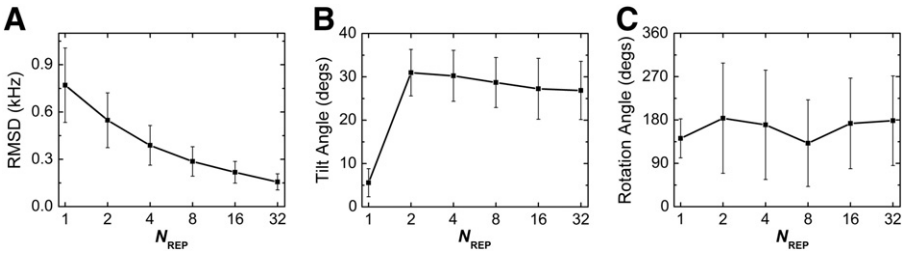

Fig. 5. (A) RMSD of DQS and (B and C) τ and ρ of WALP23 as a function of N_{REP}.

3.3. WALP23

Among various SSNMR observables, DQS measurement has been used to characterize the orientations of synthetic single-pass TM helices, such as WALP and its variant peptides [30,72]. The TM helix's τ interpreted by the semi-static geometric analysis of labeled alanine (GALA) method is relatively small (~5° for WALP23 in DMPC) [30,31], which differs significantly from much larger τ (~30°) observed in MD simulations [45,47–50]. Fig. 5B and C shows the orientation of WALP23 in terms of τ and ρ as a function of N_{REP}. While the ensemble-averaged DQS are in close agreement with experimental values in terms of RMSD (Fig. 5A), the orientation distribution of the ensemble structures shows a striking difference between $N_{REP}=1$ and $N_{REP}\geq 2$. The τ and ρ for $N_{REP}=1$ are $5.6\pm3.2°$ and $141.8\pm40.6°$ with the DQS RMSD of ~0.8 kHz, which agrees very well with the result ($\tau=5.5°$, $\rho=153.5°$, and RMSD = 0.9 kHz) based on the GALA method [30]. Interestingly, for $N_{REP}\geq 2$, much larger τ are observed, reaching at ~27° on average when $N_{REP}=32$, while ρ remains similar average values with increased fluctuations. Similar to the VpuTM case, the ensemble orientation appears to converge when $N_{REP}\geq 8$, demonstrating that it is possible to find a well-converged solution to the applied DQS restraints [73].

How does the structure ensemble show such a dramatic increase in τ with smaller DQS RMSD when $N_{REP}\geq 2$? Fig. 6 shows the DQS distribution of each alanine in the ensemble structures for $N_{REP}=1$, 8, and 32. Surprisingly, each distribution is very broad except when $N_{REP}=1$; when $N_{REP}\geq 2$, DQS of each alanine cover the whole DQS range (−42 to 84 kHz in Fig. 2B). It is notable that each distribution has the highest population on the negative DQS side and a broad distribution on the positive one. It becomes clear that such a broad DQS distribution results from the various TM helix orientations in the structure ensemble. Thus, the absolute value of the ensemble-averaged DQS becomes smaller (0–10 kHz), suggesting motional/sign averaging in the ^{2}H-NMR experiment, discussed in Section 2.3. It should be stressed that, when $N_{REP}\geq 2$, the calculated DQS of individual structures are different from the experimental one, but their ensemble-averaged DQS (regardless of their signs) show excellent agreement with experimental observables.

4. Comparison of SSNMR-ED with MD and other interpretation methods

By design, the structure ensemble determined by the (biased) SSNMR-ED simulation agrees well with the experimental data. However, the applied restraints could force the generated structures to be trapped in physically irrelevant orientations that only satisfy the experimental SSNMR observables. Cross-validation is one way of examining if the orientation distributions of such ensemble structures are realistic [74]. Another way is to check if the ensemble orientations are compatible with those from (unbiased) standard MD simulation or the 2D-PMF as a function of τ and ρ, assuming that such unbiased simulation and PMF calculation produce a reasonable and physically relevant TM helix orientation distribution.

4.1. Cross-validation

To validate if the orientation distributions of VpuTM ensemble structures are realistic and physically relevant, we performed cross-validation, which is an unbiased measure of fit [74]. For cross-validation, SSNMR-ED simulations without either CSA or DC restraint set were performed, and the RMSD of NMR observables that are not incorporated in SSNMR-ED were calculated. For example, the CSA RMSD is calculated from a structure ensemble determined by SSNMR-ED only using the DC restraint set, and vice versa. As shown in Fig. 7, unbiased CSA and DC

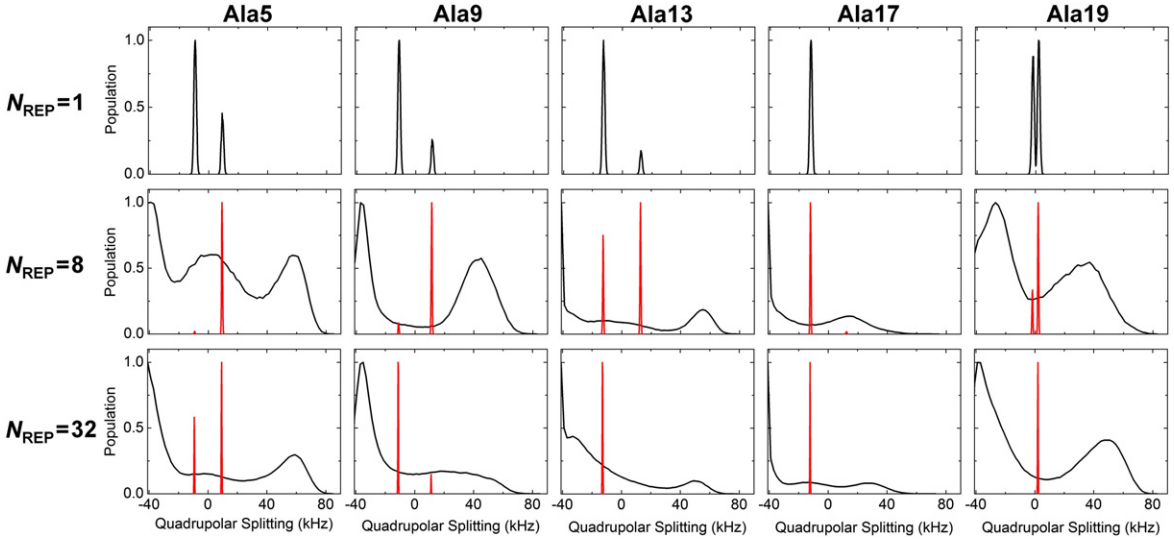

Fig. 6. Distribution of calculated (black) and ensemble averaged (red) DQS of each alanine as a function of N_{REP}. The population is normalized by setting the highest value to 1 for easy comparison for each system.

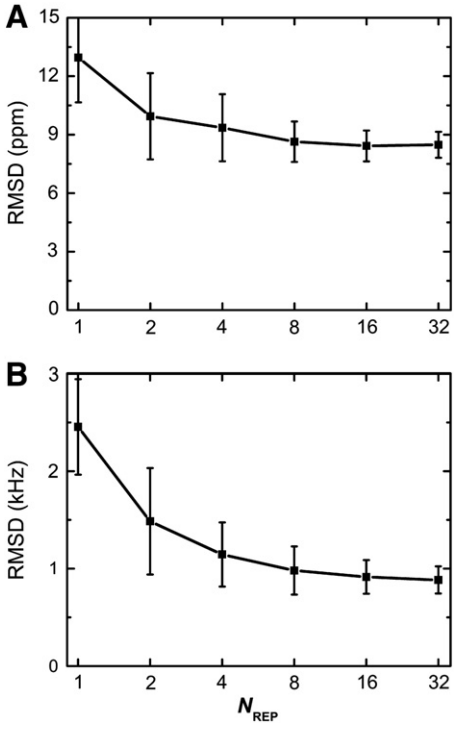

Fig. 7. (A) CSA and (B) DC RMSD from cross-validation SSNMR-ED simulation in which either the CSA or DC restraint was excluded.

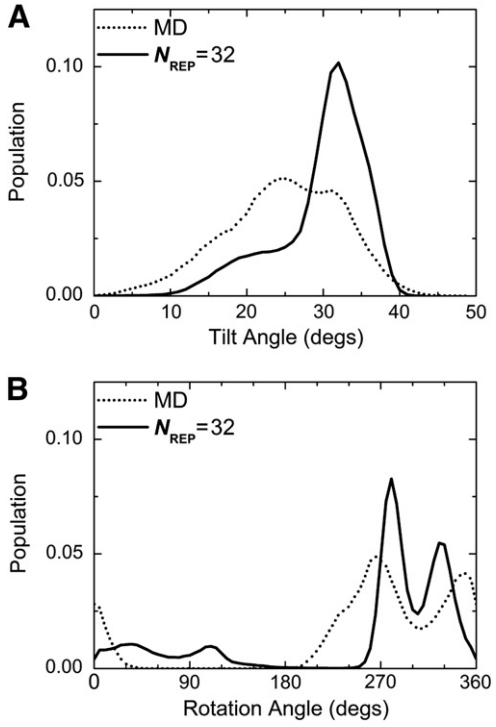

Fig. 8. VpuTM's (A) τ and (B) ρ distributions of structure ensembles (with $N_{REP}=32$) in DMPC. Dotted lines represent the distributions from standard MD simulation. The average τ and ρ from the standard MD simulations are 24.7 ± 7.5° and 295 ± 47°, respectively; the error bars are the standard deviation from the five independent trajectories. (see Reference [58] for the results in DOPC and DLPC).

RMSD become lower when more replicas are used, supporting the idea that using more replicas indeed generates structure ensembles that are more representative than the semi-static method with single conformer. This cross-validation also demonstrates that improved quality of fit is not because of imposing unphysical bias to the structure or the reduction of a data-to-parameter ratio,

Cross-validation of SSNMR-ED with DQS for WALP23 is not an easy task. Since the number of restraints (5 DQS) in the SSNMR-ED simulation is not enough to perform reliable cross-validation, the orientational variability of the SSNMR-ED structures is validated with the 2D-PMF calculation as a function of WALP23's τ and ρ in the next subsection.

4.2. SSNMR-ED versus MD/2D-PMF

To examine if the VpuTM orientation distributions of SSNMR-ED ensemble structures are compatible with those from standard MD simulation, we have performed a total of 1.5 μs comparative MD simulations of VpuTM in explicit lipid bilayers, starting from τ = 0° (see Reference [58] for details). Fig. 8 shows the comparisons between SSNMR-ED (with $N_{REP}=32$) and MD for τ and ρ distributions. As mentioned in Section 3.2, the structures determined with $N_{REP}=1$ produces a much narrower distribution than those from MD simulation and significantly underestimates the variability of TM helix orientation. When $N_{REP}\geq 8$, the structures determined by SSNMR-ED show similar orientational variability to the ones from MD simulation. However, detailed examination of the τ and ρ distributions revealed that the overall shape of the distributions does not exactly match in the structure ensembles from MD and SSNMR-ED. Such differences could arise from the limited MD simulation time and the different representation of membranes in MD (explicit) and SSNMR-ED (implicit).

For WALP23, to validate the orientational variability of the SSNMR-ED structures, we have carried out the 2D-PMF calculation as a function of WALP23's τ and ρ in an implicit membrane model (see Reference [59] for details). Such a validation is important for two reasons. First, the number of restraints (5 DQS) in the SSNMR-ED simulation is not enough to perform reliable cross-validation. Second, a regular MD simulation often suffers from insufficient sampling to obtain a reliable and converged TM helix orientation distribution. Fig. 9 compares the 2D-PMF with the SSNMR-ED ensemble structure distribution with $N_{REP}=32$. The high population in the ensemble structures well matches with the thermally-accessible region in the 2D-PMF; the thermally-accessible orientations are defined by the regions that have free energies within 0.6 kcal/mol from the minimum-PMF orientation. Most structures from the SSNMR-ED simulations are within 3.0 kcal/mol from the minimum-PMF orientation. This comparison again supports the notion that the orientation distribution of the ensemble structures from the SSNMR-ED simulation is physically relevant.

Despite some differences, these independent comparisons between SSNMR-ED and MD (or 2D-PMF) for VpuTM and WALP23 clearly demonstrate that SSNMR-ED has the ability to generate an ensemble of

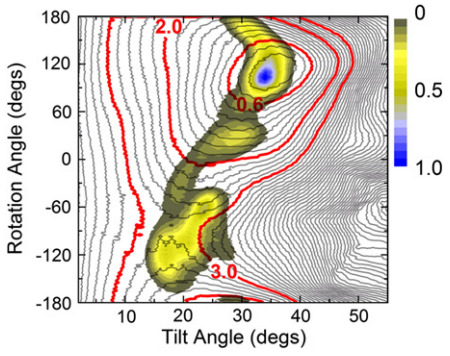

Fig. 9. Overlap of the ensemble structure distribution (with $N_{REP}=32$) and 2D-PMF. The contour lines (gray) are drawn every 0.2 kcal/mol and red line indicates 0.6, 2.0, and 3.0 kcal/mol from the PMF minimum. The population is normalized by setting the highest value to 1.

structures (i.e., TM helix orientation distribution) that satisfies the experimental observables within a reasonable physical (force field) model. Therefore, SSNMR-ED can be used to extract both TM helix orientation and dynamics simultaneously from the SSNMR observables.

4.3. SSNMR-ED versus semi-static/dynamic fitting methods

While we already mentioned the semi-static results (published by others) for VpuTM and WALP23 in Sections 3.2 and 3.3, their orientations determined by the semi-static and Gaussian-based dynamic fitting methods are compared with the SSNMR-ED results in this subsection. In the Gaussian-based dynamic fitting calculation, τ_0, δ_τ, ρ_0, and δ_ρ in Eq. (7) were considered as fitting parameters and independently varied from 0° to 55° by 1° (τ_0), 0° to 20° by 2° (δ_τ), −180° to 180° by 1° (ρ_0), and 0° to 120° by 5° (δ_ρ), respectively; n.b., when $\delta_\tau = \delta_\rho = 0°$, it becomes the semi-static model in Eq. (6). For each set of parameters (τ_0, δ_τ, ρ_0, and δ_ρ), the ensemble-averaged values were calculated using Eq. (5) and used to compute the RMSD from experimental data. Order parameters were set to 1.0 for VpuTM (CSA and DC) and 0.88 for WALP23 (DQS) in DMPC, following the original References [11,63]. Each set of parameters that yielded the RMSD below 10 ppm (CSA), 1 kHz (DC), and 2 kHz (DQS) were accepted. In practice, there are many sets of parameters that are acceptable and it may not be reasonable to choose only one set simply based on the lowest RMSD. Instead, the accessible orientational distribution can be calculated by averaging the Gaussian distributions with acceptable parameter sets.

$$\langle P(\tau,\rho)\rangle = \frac{1}{N_{fit}}\sum_{i}^{N_{fit}} P_i(\tau,\rho) \quad (15)$$

where N_{fit} is the number of {τ_0, δ_τ, ρ_0, and δ_ρ} below the RMSD thresholds. The resulting $\langle P(\tau,\rho)\rangle$ is plotted in Fig. 10A and B (gray to red scale) and compared with the SSNMR-ED result with $N_{REP} = 32$ (gray to blue scale). Although one could use different weights for different parameter sets in Eq. (15), the equal weight was applied because it is not possible to determine which sets are more important than the others in the dynamic fitting approach. Nevertheless, the resulting population distribution $\langle P(\tau,\rho)\rangle$ is useful to visualize the resulting parameter space and to compare the result with different methods.

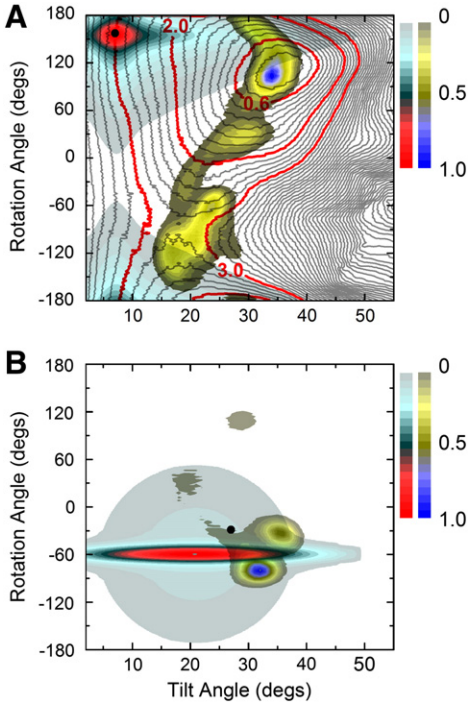

Fig. 10. (A) Overlap of the ensemble structure distribution (with $N_{REP} = 32$, gray to blue scale), 2D-PMF (gray contour line), semi-static fitting (black spot: the lowest RMSD; 1.42 kHz), and the probability of Gaussian distributions from dynamic fitting (gray to red scale) of WALP23. (B) Overlap of the ensemble structure distribution (with $N_{REP} = 32$, gray to blue scale), semi-static fitting (black spot: the lowest RMSD), and probability of Gaussian distribution from dynamic fitting (gray to red scale) of VpuTM. The population is normalized by setting the highest value to 1 for easy comparison.

For WALP23, the Gaussian-based dynamic fitting method yields a slightly larger τ and a broader orientational distribution than the semi-static method (Fig. 10A). However, the most populations determined by the fitting method are still restricted in much smaller τ (<10°), compared to the previous MD (~30°), SSNMR-ED, and 2D-PMF results. In the case of VpuTM, the Gaussian-based dynamic fitting method yields a slightly smaller τ and a broader orientational distribution than the semi-static method (Fig. 10B); the δ_τ and δ_ρ are about 1.0° and 7.9° (semi-static) and 6.2° and 3.6° (dynamic fitting; average over accessible parameters). Note that there are too many δ_τ for narrow ranges of τ_0, ρ_0, and δ_ρ in the acceptable parameter sets, so that the resulting population appears to be much broader along τ in Fig. 10B. The Gaussian-based fitting method still yielded restricted TM helix dynamics, compared to the MD (7.5° and 47° for τ and ρ) and SSNMR-ED (8–9° and 50°) results in Fig. 8.

These comparisons for both VpuTM and WALP23 illustrate that the dynamic fitting method naturally adds more orientational fluctuations in the determined TM orientations (compared to the semi-empirical method), but the resulting populations are still under strong influence of the SSNMR experimental values similar to the semi-empirical method. Consequently, the dynamic fitting method yields smaller TM helix dynamics in VpuTM and much smaller τ in WALP23, compared to the MD (or 2D-PMF) and SSNMR-ED results.

5. Summary and perspectives

We have described various computational techniques (semi-static/dynamic fitting/SSNMR-ED) to determine TM helix orientations using the SSNMR observables (CSA, DC, and DQS) with an emphasis on newly developed SSNMR-ED. With VpuTM (with CSA and DC) and WALP23 (with DQS) as model systems, we have validated the SSNMR-ED technique and compared the TM orientation distributions of resulting ensemble structures with those from MD simulations (including the PMF calculations) and semi-static/dynamic fitting models. By allowing individual replicas more orientational freedom, the resulting SSNMR-ED structure ensembles show better agreement with experimental observables including motional/sign averaging and much greater orientational variability that overall matches well with the MD simulation.

Compared to the semi-static and dynamic fitting approaches, the primary advantage of SSNMR-ED is the ability to generate an ensemble of structures (i.e., TM helix orientation distribution) that satisfies the experimental observables within a reasonable physical (force field) model without prior knowledge about the underlying distribution or motional modes. This allows one to extract both TM helix orientation and dynamics from SSNMR observables. Such an advantage is indeed rooted in the development of sophisticated SSNMR restraint potentials [68] that can be used in MD simulations with other general potential energy function, which makes SSNMR-ED general and unique. In this context, our perspectives and future directions on computational studies of membrane protein structure and dynamics in combination with experimental SSNMR observables are discussed in the following.

5.1. GWALP23

This is a synthetic peptide (acetyl-GGALW(LA)$_6$LWLAGA-ethanolamide), which is different from WALP23 in terms of the numbers of Trp anchors and their relative positions [75]. Compared to WALP23, there are experimental CSA, DC, and DQS available for GWALP23, which were measured by Koeppe and co-workers [75,76]. Interestingly, the TM helix

orientation determined by the semi-static model using CSA and DC yielded similar orientation determined using DQS. Therefore, it will be interesting to use SSNMR-ED to examine the TM helix orientations with multiple independent SSNMR measurements, and elucidate the role of Trp anchors in different TM orientation distributions and dynamics.

5.2. SSNMR-ED with more complex topologies and TM oligomers

So far, we have focused on simple single-pass TM helices, and thus it is a natural extension to use SSNMR-ED for membrane proteins with more complex topologies and TM hetero-/homo-oligomers. In this context, the major coat protein of Pf1 bacteriophage [20] and phospholamban (PLN) [24,77], a regulator of SERCA Ca^{2+} pump, are interesting systems to explore; both have a single-pass TM with a periplasmic or cytoplasmic helical domain. In particular, PLN is known to form a homo-pentamer. In addition, MerF, a mercuric ion transporter, can also be an excellent model system because it has two TM helices [23]. Both PLN homo-pentamer and MerF are useful model systems to understand the extent of TM orientational variability depending on the TM assembly. Therefore, the applications of SSNMR-ED to these systems can provide insights into complex membrane protein structure, dynamics, and function because intrinsic dynamics and/or distinct configurations of different domains can be directly related to membrane protein function. In addition, to examine the physical relevance of the resulting structure ensembles, various PMF calculations based on the restraint potentials for helix tilt, rotation, and crossing angles can be compared with the SSNMR-ED simulations.

5.3. Ensemble dynamics with RDC

Although it is not discussed in this review, residual dipolar coupling (RDC) measured in solution NMR provides orientational information and has been proven to be useful in membrane protein structure determination. For example, both Pf1 coat protein [20] and PLN have experimental RDC available [78,79]. We have separately developed a RDC restraint potential that does not require the alignment tensor information (or reasonable starting structures) as an input [80]. Extension of the RDC restraint potential for ensemble dynamics and its applications together with SSNMR observables can provide in-depth understanding of membrane protein structure, dynamics, and function.

5.4. Membrane protein structure refinement in explicit lipids

What is currently lacking in SSNMR structure determination is the detailed information of protein interactions with bilayers. In this context, the membrane protein structure refinement in explicit bilayers, i.e., restrained MD simulation of a membrane protein with the SSNMR observables (with/without RDCs), can provide such information at the atomic resolution. The restrained MD simulation with SSNMR observables is an attractive approach because experimental SSNMR data do not provide sufficient number of restraints to define sidechain–sidechain interactions and sidechain–lipid interactions in the bilayer environments. The rationale behind this approach is to maximally use the available experimental observables with additional benefits from the state-of-the-art MD simulation of membrane systems.

5.5. SSNMR-ED in explicit bilayers

With increasing computational resources, it will be feasible to perform SSNMR-ED simulations in explicit bilayers. So far, our SSNMR-ED simulations have been performed in implicit membrane models mainly due to limited resources and illustration purpose of the SSNMR-ED's efficacy. Although such implicit membrane models reasonably capture the bilayer environments, explicit bilayers are more realistic, including lipid adaptation due to TM helix orientational changes. Therefore, it would be necessary to compare the SSNMR-ED results from implicit and explicit bilayers. Such comparisons can be also beneficial for better understanding of TM helix dynamics in different membrane models.

In conclusion, SSNMR-ED can be used as a general strategy to extract both TM helix orientation and dynamics simultaneously from the SSNMR measurements, and one may apply this knowledge to investigating the influence of the TM helix orientation and its variability on the structure and function of biologically important systems.

Acknowledgment

This work was supported by the NSF MCB-0918374 and TeraGrid resources provided by Purdue University (NSF OCI-0503992).

References

[1] O.S. Andersen, R.E. Koeppe II, Bilayer thickness and membrane protein function: an energetic perspective, Annu. Rev. Biophys. Biomol. Struct. 36 (2007) 107–130.
[2] S. Steinbacher, R. Bass, P. Strop, D.C. Rees, Structures of the prokaryotic mechanosensitive channels MscL and MscS, Curr. Top. Membr. 58 (2007) 1–24.
[3] B. Corry, B. Martinac, Bacterial mechanosensitive channels: experiment and theory, Biochim. Biophys. Acta 1778 (2008) 1859–1870.
[4] M. Betanzos, C.-S. Chiang, H.R. Guy, S. Sukharev, A large iris-like expansion of a mechanosensitive channel protein induced by membrane tension, Nat. Struct. Biol. 9 (2002) 704–710.
[5] Z. Liu, C.S. Gandhi, D.C. Rees, Structure of a tetrameric MscL in an expanded intermediate state, Nature 461 (2009) 120–124.
[6] S. White, Membrane Proteins of Known 3D Structures, http://blanco.biomol.uci.edu/mpstruc/.
[7] I.T. Arkin, A.T. Brunger, D.M. Engelman, Are there dominant membrane protein families with a given number of helices? Proteins 28 (1997) 465–466.
[8] M.S. Almen, K.J. Nordstrom, R. Fredriksson, H.B. Schioth, Mapping the human membrane proteome: a majority of the human membrane proteins can be classified according to function and evolutionary origin, BMC Biol. 7 (2009) 50.
[9] C. Wu, A. Ramamoorthy, S. Opella, High-resolution heteronuclear dipolar solid-state NMR-spectroscopy, J. Magn. Reson. Ser. A 109 (1994) 270–272.
[10] D.H. Jones, K.R. Barber, E.W. VanDerLoo, C.W. Grant, Epidermal growth factor receptor transmembrane domain; ^{2}H NMR implications for orientation and motion in a bilayer environment, Biochemistry 37 (1998) 16780–16787.
[11] P.C.A. van der Wel, E. Strandberg, J.A. Killian, R.E. Koeppe II, Geometry and intrinsic tilt of a tryptophan-anchored transmembrane alpha-helix determined by ^{2}H NMR, Biophys. J. 83 (2002) 1479–1488.
[12] A.A. Nevzorov, S.J. Opella, A "magic sandwich" pulse sequence with reduced offset dependence for high-resolution separated local field spectroscopy, J. Magn. Reson. 164 (2003) 182–186.
[13] T.A. Cross, J.R. Quine, Protein structure in anisotropic environments: development of orientational constraints, Concepts Magn. Reson. 12 (2000) 55–70.
[14] F.M. Marassi, S.J. Opella, A solid-state NMR index of helical membrane protein structure and topology, J. Magn. Reson. 144 (2000) 150–155.
[15] R. Smith, D.E. Thomas, F. Separovic, A.R. Atkins, B.A. Cornell, Determination of the structure of a membrane-incorporated ion channel. Solid-state nuclear magnetic resonance studies of gramicidin A, Biophys. J. 56 (1989) 307–314.
[16] R.R. Ketchem, W. Hu, T.A. Cross, High-resolution conformation of gramicidin A in a lipid bilayer by solid-state NMR, Science 261 (1993) 1457–1460.
[17] C. Fares, J. Qian, J. Davis, Magic angle spinning and static oriented sample NMR studies of the relaxation in the rotating frame of membrane peptides, J. Chem. Phys. 122 (2005).
[18] S.H. Park, A.A. De Angelis, A.A. Nevzorov, C.H. Wu, S.J. Opella, Three-dimensional structure of the transmembrane domain of Vpu from HIV-1 in aligned phospholipid bicelles, Biophys. J. 91 (2006) 3032–3042.
[19] S.J. Opella, A.C. Zeri, S.H. Park, Structure, dynamics, and assembly of filamentous bacteriophages by nuclear magnetic resonance spectroscopy, Annu. Rev. Phys. Chem. 59 (2008) 635–657.
[20] S.H. Park, F.M. Marassi, D. Black, S.J. Opella, Structure and dynamics of the membrane-bound form of Pf1 coat protein: implications of structural rearrangement for virus assembly, Biophys. J. 99 (2010) 1465–1474.
[21] F.A. Kovacs, J.K. Denny, Z. Song, J.R. Quine, T.A. Cross, Helix tilt of the M2 transmembrane peptide from influenza A virus: an intrinsic property, J. Mol. Biol. 295 (2000) 117–125.
[22] C. Tian, K. Tobler, R.A. Lamb, L.H. Pinto, T.A. Cross, Expression and initial structural insights from solid-state NMR of the M2 proton channel from influenza A virus, Biochemistry 41 (2002) 11294–11300.
[23] A.A. De Angelis, S.C. Howell, A.A. Nevzorov, S.J. Opella, Structure determination of a membrane protein with two trans-membrane helices in aligned phospholipid bicelles by solid-state NMR spectroscopy, J. Am. Chem. Soc. 128 (2006) 12256–12267.
[24] N.J. Traaseth, J.J. Buffy, J. Zamoon, G. Veglia, Structural dynamics and topology of phospholamban in oriented lipid bilayers using multidimensional solid-state NMR, Biochemistry 45 (2006) 13827–13834.

[25] B. Bechinger, M. Zasloff, S.J. Opella, Structure and dynamics of the antibiotic peptide PGLa in membranes by solution and solid-state nuclear magnetic resonance spectroscopy, Biophys. J. 74 (1998) 981–987.
[26] S. Yamaguchi, T. Hong, A. Waring, R.I. Lehrer, M. Hong, Solid-state NMR investigations of peptide–lipid interaction and orientation of a beta-sheet antimicrobial peptide, protegrin, Biochemistry 41 (2002) 9852–9862.
[27] A. Ramamoorthy, S. Thennarasu, D.-K. Lee, A. Tan, L. Maloy, Solid-state NMR investigation of the membrane-disrupting mechanism of antimicrobial peptides MSI-78 and MSI-594 derived from magainin 2 and melittin, Biophys. J. 91 (2006) 206–216.
[28] M.R.R. de Planque, J.-W.P. Boots, D.T.S. Rijkers, R.M.J. Liskamp, D.V. Greathouse, J.A. Killian, The effects of hydrophobic mismatch between phosphatidylcholine bilayers and transmembrane alpha-helical peptides depend on the nature of interfacially exposed aromatic and charged residues, Biochemistry 41 (2002) 8396–8404.
[29] E. Strandberg, S. Morein, D.T.S. Rijkers, R.M.J. Liskamp, P.C.A. van der Wel, J.A. Killian, Lipid dependence of membrane anchoring properties and snorkeling behavior of aromatic and charged residues in transmembrane peptides, Biochemistry 41 (2002) 7190–7198.
[30] E. Strandberg, S. Ozdirekcan, D.T.S. Rijkers, P.C.A. van der Wel, R.E. Koeppe II, R.M.J. Liskamp, J.A. Killian, Tilt angles of transmembrane model peptides in oriented and non-oriented lipid bilayers as determined by 2H solid-state NMR, Biophys. J. 86 (2004) 3709–3721.
[31] S. Ozdirekcan, D.T.S. Rijkers, R.M.J. Liskamp, J.A. Killian, Influence of flanking residues on tilt and rotation angles of transmembrane peptides in lipid bilayers. A solid-state ^{2}H NMR study, Biochemistry 44 (2005) 1004–1012.
[32] R.S. Prosser, J.H. Davis, Dynamics of an integral membrane peptide: a deuterium NMR relaxation study of gramicidin, Biophys. J. 66 (1994) 1429–1440.
[33] E. Strandberg, P. Wadhwani, P. Tremouilhac, U.H.N. Dürr, A.S. Ulrich, Solid-state NMR analysis of the PGLa peptide orientation in DMPC bilayers: structural fidelity of ^{2}H-labels versus high sensitivity of 19F-NMR, Biophys. J. 90 (2006) 1676–1686.
[34] J. Wang, J. Denny, C. Tian, S. Kim, Y. Mo, F. Kovacs, Z. Song, K. Nishimura, Z. Gan, R. Fu, J.R. Quine, T.A. Cross, Imaging membrane protein helical wheels, J. Magn. Reson. 144 (2000) 162–167.
[35] A.A. Nevzorov, S.J. Opella, Structural fitting of PISEMA spectra of aligned proteins, J. Magn. Reson. 160 (2003) 33–39.
[36] S. White, W. Wimley, Membrane protein folding and stability: physical principles, Annu. Rev. Biophys. Biomol. Struct. 28 (1999) 319–365.
[37] J.L. Popot, D.M. Engelman, Helical membrane protein folding, stability, and evolution, Annu. Rev. Biochem. 69 (2000) 881–922.
[38] A. Ramamoorthy, S.K. Kandasamy, D.-K. Lee, S. Kidambi, R.G. Larson, Structure, topology, and tilt of cell-signaling peptides containing nuclear localization sequences in membrane bilayers determined by solid-state NMR and molecular dynamics simulation studies, Biochemistry 46 (2007) 965–975.
[39] P.J. Bond, J. Holyoake, A. Ivetac, S. Khalid, M.S.P. Sansom, Coarse-grained molecular dynamics simulations of membrane proteins and peptides, J. Struct. Biol. 157 (2007) 593–605.
[40] D.J. Goodyear, S. Sharpe, C.W.M. Grant, M.R. Morrow, Molecular dynamics simulation of transmembrane polypeptide orientational fluctuations, Biophys. J. 88 (2005) 105–117.
[41] J. Lee, W. Im, Transmembrane helix tilting: insights from calculating the potential of mean force, Phys. Rev. Lett. 100 (2008) 018103.
[42] L. Shi, A. Cembran, J. Gao, G. Veglia, Tilt and azimuthal angles of a transmembrane peptide: a comparison between molecular dynamics calculations and solid-state NMR data of sarcolipin in lipid membranes, Biophys. J. 96 (2009) 3648–3662.
[43] T. Kim, J. Lee, W. Im, Molecular dynamics studies on structure and dynamics of phospholamban monomer and pentamer in membranes, Proteins 76 (2009) 86–98.
[44] V.V. Vostrikov, B.A. Hall, D.V. Greathouse, R.E. Koeppe II, M.S.P. Sansom, Changes in transmembrane helix alignment by arginine residues revealed by solid-state NMR experiments and coarse-grained MD simulations, J. Am. Chem. Soc. 132 (2010) 5803–5811.
[45] T. Kim, W. Im, Revisiting hydrophobic mismatch with free energy simulation studies of transmembrane helix tilt and rotation, Biophys. J. 99 (2010) 175–183.
[46] H. Rui, W. Im, Protegrin-1 orientation and physicochemical properties in membrane bilayers studied by potential of mean force calculations, J. Comput. Chem. 31 (2010) 2859–2867.
[47] W. Im, C.L. Brooks III, Interfacial folding and membrane insertion of designed peptides studied by molecular dynamics simulations, Proc. Natl. Acad. Sci. U.S.A. 102 (2005) 6771–6776.
[48] S. Ozdirekcan, C. Etchebest, J.A. Killian, P.F.J. Fuchs, On the orientation of a designed transmembrane peptide: toward the right tilt angle? J. Am. Chem. Soc. 129 (2007) 15174–15181.
[49] S. Esteban-Mart, J. Salgado, The dynamic orientation of membrane-bound peptides: bridging simulations and experiments, Biophys. J. 93 (2007) 4278–4288.
[50] L. Monticelli, D.P. Tieleman, P.F.J. Fuchs, Interpretation of ^{2}H-NMR experiments on the orientation of the transmembrane helix WALP23 by computer simulations, Biophys. J. 99 (2010) 1455–1464.
[51] S. Esteban-Martin, E. Strandberg, G. Fuertes, A.S. Ulrich, J. Salgado, Influence of whole-body dynamics on ^{15}N PISEMA NMR spectra of membrane proteins: a theoretical analysis, Biophys. J. 96 (2009) 3233–3241.
[52] E. Strandberg, S. Esteban-Martin, J. Salgado, A.S. Ulrich, Orientation and dynamics of peptides in membranes calculated from ^{2}H-NMR data, Biophys. J. 96 (2009) 3223–3232.
[53] A. Holt, L. Rougier, V. Reat, F. Jolibois, O. Saurel, J. Czaplicki, J.A. Killian, A. Milon, Order parameters of a transmembrane helix in a fluid bilayer: case study of a WALP peptide, Biophys. J. 98 (2010) 1864–1872.
[54] D.T. Murray, Y. Lu, T.A. Cross, J.R. Quine, Geometry of kinked protein helices from NMR data, J. Magn. Reson. 210 (2011) 82–89.
[55] A.M.J.J. Bonvin, A.T. Brünger, Conformational variability of solution nuclear magnetic resonance structures, JMBio 250 (1995) 80–93.
[56] O.F. Lange, N. Lakomek, C. Farès, G.F. Schröder, K.F.A. Walter, S. Becker, J. Meiler, H. Grubmüller, C. Griesinger, B.L. de Groot, Recognition dynamics up to microseconds revealed from an RDC-derived ubiquitin ensemble in solution, Science 320 (2008) 1471–1475.
[57] B. Richter, J. Gsponer, P. Várnai, X. Salvatella, M. Vendruscolo, The MUMO (minimal under-restraining minimal over-restraining) method for the determination of native state ensembles of proteins, J. Biomol. NMR 37 (2007) 117–135.
[58] S. Jo, W. Im, Transmembrane helix orientation and dynamics: insights from ensemble dynamics with solid-state NMR observables, Biophys. J. 100 (2011) 2913–2921.
[59] T. Kim, S. Jo, W. Im, Solid-state NMR ensemble dynamics as a mediator between experiment and simulation, Biophys. J. 100 (2011) 2922–2928.
[60] A.M. Bonvin, A.T. Brunger, Conformational variability of solution nuclear magnetic resonance structures, J. Mol. Biol. 250 (1995) 80–93.
[61] K. Lindorff-Larsen, R.B. Best, M.A. Depristo, C.M. Dobson, M. Vendruscolo, Simultaneous determination of protein structure and dynamics, Nature 433 (2005) 128–132.
[62] E.J. Levin, D.A. Kondrashov, G.E. Wesenberg, G.N. Phillips, Ensemble refinement of protein crystal structures: validation and application, Structure 15 (2007) 1040–1052.
[63] S.H. Park, S.J. Opella, Tilt angle of a trans-membrane helix is determined by hydrophobic mismatch, J. Mol. Biol. 350 (2005) 310–318.
[64] J.H. Davis, The description of membrane lipid conformation, order and dynamics by ^{2}H-NMR, Biochim. Biophys. Acta 737 (1983) 117–171.
[65] A. Nevzorov, S. Moltke, M. Heyn, M.F. Brown, Solid-state NMR line shapes of uniaxially oriented immobile systems, J. Am. Chem. Soc. 121 (1999) 7636–7643.
[66] A.V. Struts, G.F.J. Salgado, K. Tanaka, S. Krane, K. Nakanishi, M.F. Brown, Structural analysis and dynamics of retinal chromophore in dark and meta I states of rhodopsin from ^{2}H NMR of aligned membranes, J. Mol. Biol. 372 (2007) 50–66.
[67] S. Esteban-Martín, D. Giménez, G. Fuertes, J.s. Salgado, Orientational landscapes of peptides in membranes: prediction of ^{2}H NMR couplings in a dynamic context, Biochemistry 48 (2009) 11441–11448.
[68] J. Lee, J. Chen, C.L. Brooks III, W. Im, Application of solid-state NMR restraint potentials in membrane protein modeling, J. Magn. Reson. 193 (2008) 68–76.
[69] B.R. Brooks, C.L. Brooks III, A.D. Mackerell Jr., L. Nilsson, R.J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, A. Caflisch, L. Caves, Q. Cui, A.R. Dinner, M. Feig, S. Fischer, J. Gao, M. Hodoscek, W. Im, K. Kuczera, T. Lazaridis, J. Ma, V. Ovchinnikov, E. Paci, R.W. Pastor, C.B. Post, J.Z. Pu, M. Schaefer, B. Tidor, R.M. Venable, H.L. Woodcock, X. Wu, W. Yang, D.M. York, M. Karplus, CHARMM: the biomolecular simulation program, J. Comput. Chem. 30 (2009) 1545–1614.
[70] A. Holt, J.A. Killian, Orientation and dynamics of transmembrane peptides: the power of simple models, Eur. Biophys. J. 39 (2010) 609–621.
[71] J.A. Lundbaek, S.A. Collingwood, H.I. Ingolfsson, R. Kapoor, O.S. Andersen, Lipid bilayer regulation of membrane protein function: gramicidin channels as molecular force probes, J. R. Soc. Interface 7 (2010) 373–395.
[72] P.C.A. van der Wel, E. Strandberg, J.A. Killian, R.E. Koeppe II, Geometry and intrinsic tilt of a tryptophan-anchored transmembrane a-helix determined by ^{2}H NMR, Biophys. J. 83 (2002) 1479–1488.
[73] R.B. Best, M. Vendruscolo, Determination of protein structures consistent with NMR order parameters, J. Am. Chem. Soc. 126 (2004) 8090–8091.
[74] A.T. Brünger, G.M. Clore, A.M. Gronenborn, R. Saffrich, M. Nilges, Assessing the quality of solution nuclear magnetic resonance structures by complete cross-validation, Science 261 (1993) 328–331.
[75] V.V. Vostrikov, C.V. Grant, A.E. Daily, S.J. Opella, R.E. Koeppe II, Comparison of "Polarization Inversion with Spin Exchange at Magic Angle" and "Geometric Analysis of Labeled Alanines" methods for transmembrane helix alignment, J. Am. Chem. Soc. 130 (2008) 12584–12585.
[76] V.V. Vostrikov, A.E. Daily, D.V. Greathouse, R.E. Koeppe II, Charged or aromatic anchor residue dependence of transmembrane peptide tilt, J. Biol. Chem. 285 (2010) 31723–31730.
[77] K. Oxenoid, J.J. Chou, The structure of phospholamban pentamer reveals a channel-like architecture in membranes, Proc. Natl. Acad. Sci. U.S.A. 102 (2005) 10870–10875.
[78] J. Chou, S. Gaemers, B. Howder, J. Louis, A. Bax, A simple apparatus for generating stretched polyacrylamide gels, yielding uniform alignment of proteins and detergent micelles, J. Biomol. NMR 21 (2001) 377–382.
[79] R. Tycko, F.J. Blanco, Y. Ishii, Alignment of biopolymers in strained gels: a new way to create detectable dipole–dipole couplings in high-resolution biomolecular NMR, J. Am. Chem. Soc. 122 (2000) 9340–9341.
[80] T. Rathinavelan, W. Im, Explicit treatment of force contribution from alignment tensor using overdetermined linear equations and its application in NMR structure determination, J. Comput. Chem. 28 (2007) 1858–1864.