MOLECULAR DYNAMICS SIMULATION OF MEMBRANE PEPTIDES WITH AN
IMPLICIT MEMBRANE MODEL

By

Maryam Sayadi

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Chemistry- Doctor of Philosophy

2013

ABSTRACT

MOLECULAR DYNAMICS SIMULATION OF MEMBRANE PEPTIDES WITH AN
IMPLICIT MEMBRANE MODEL

By

Maryam Sayadi

Molecular dynamics simulation is a powerful technique to study structure and function of biological systems. Over the time, by developing more sophisticated models and faster computers, interest in simulating larger and more complicated systems has grown. With the current supercomputers, molecular dynamics simulation of huge systems, which we could only dream of several years ago, has become possible but yet computationally expensive. Implicit treatment of the solvent (and membrane) is one of the widely used methods to reduce the computational costs, which also can be used in combination with other enhanced sampling methods. However, there is still some room for improving both the accuracy and the speed of such methods.

Molecular dynamics simulations of Phospholamban (PLB) have been carried out with the Heterogeneous Dielectric Generalized Born method (HDG) with an implicit representation of the solvent and membrane. The conformational sampling of PLB in the phosphorylated and unphosphorylated states suggests that PLB has a steric effect on the E2-E1 conformational transition of Sarcoplasmic Reticulum Ca^{2+}ATPase (SERCA). This steric effect disrupts the Ca^{2+} ion uptake cycle and inhibits SERCA. Phosphorylation of PLB at Ser16 induces some structural changes in the PLB that diminishes the steric effect and allows SERCA to visit the E1 state and transports Ca^{+2} ion.

A Hamiltonian replica exchange has been also introduced to enhance sampling in the

implicit solvent simulations. In this technique, the solvent's dielectric constant exchanges between the replicas. It is proposed that visiting replicas with the lower dielectric constants alters the energy surface and improves the sampling speed.

I dedicate this work to my beloved husband, Borzoo Bonakdarpour and my parents.

ACKNOWLEDGMENTS

I wish to thank, first and foremost, my beloved husband, Borzoo Bonakdarpour, for all the support and love. I wish I were as strong, reliable, and hard working as he is.

It gives me great pleasure in acknowledging my advisor, Dr. Michael Feig. I am grateful for his patience, guidance and financial support. He was the best mentor I could possibly have. He taught me the way of the scientific and critical thinking, which I will always benefit from in my future carriers.

I wish to thank my parents for their strength, patience and kindness. It was very hard to be away from my loved ones. Their strength gave me courage to never give up.

I am indebted to Mr. Peter Briggs, the director of the MSU Office for International Students and Scholars, and Dr. Gary Blanchard, for their help and support during the three years that I was trying to obtain a student visa.

I would also like to acknowledge the help and guidance of my committee members, Dr. Robert I. Cukier, Dr. David Weliky and Dr. A. Daniel Jones.

I was greatly fortunate to be able to have one of my best friends, Dr. Afra Panahi, as a colleague and a dear friend in my group. I vastly enjoyed the times we spent together and the scientific discussions that we had over the time.

I am lucky to have had the honor of working with all the members of Feig group. For most, Dr. Seiichiro Tanizaki whom I am grateful for his help when I initially joined the group.

I greatly enjoyed the fun and the laughter I had with my friend and roommate Dr. Atefeh Garzan and I am thankful for the time that we had.

I am forever grateful for all the support and kindness from all my friends, colleagues and professors in Michigan State University.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1 Introduction

Thirty years ago, it was difficult to picture computational methods as being as effective and reliable as experimental methods in studying large biological macromolecules. The first molecular dynamics (MD) [1] simulation study reported in 1977 [2] was only a few ps long and involved basic pancreatic trypsin inhibitor or BPTI . Study of this small biological system suggested that proteins are dynamic systems, not completely rigid as the X-Ray crystallography [3, 4] pictured. Since then, because of better and faster computers and improvements to computational methods, larger biological molecules such as RNA polymerase [5] and DNA complexes [6, 7] have been studied. In addition to the increase in the system size, the accessible simulation time scale has been extended to several microseconds [8] and, in some cases, a millisecond [9, 10].

Most experimental techniques such as Nuclear Magnetic Resonance (NMR) [11] and Fluorescence Resonance Energy Transfer (FRET) [12] spectroscopy, only collect average structural properties of a system rather than information about single molecules. This is especially problematic when the system is very dynamic and can exist in several conformational states. Using a technique such as NMR, one can suggest an ensemble of structures where their average structural properties match with NMR data. X-Ray crystallography, on the other hand, suggests atomistic details about the most stable structure under crystallographic conditions. However, the X-Ray crystallography structure is not necessarily the most stable conformation under biological conditions [13]. By using MD, atomistic insight is provided for each individual structure in an ensemble of structures, and the weight for each conformation can be extracted through sampling. An example for a dynamic system is phospholamban, a 52 amino acid

membrane-bound protein, which is studied in this work by MD and will be discussed thoroughly in the next two chapters.

There are two main challenges in MD: realistic modeling of the system and a high computational cost. In MD simulations, modeling involves an appropriate particle-based representation, often at the atomistic level, and a description of all the interactions by mathematical equations. The latter is called the force field, which can vary significantly and is the key to a realistic description of a given biological system at the desired level of detail. In some force fields, very simple and crude models are used. But even more sophisticated force fields often do not describe all aspects of a given system with sufficient accuracy. For example, the CHARMM22 force field has been reported to over-stabilize π-helical structures relative to α-helices [14]. AMBER-99SB and AMBER-03 force fields, on the other hand, destabilize helical structures in peptides and proteins [15]. Therefore, one should be careful in interpreting the data obtained from each force field.

The second challenge in MD is the computational cost. Because of limitations in the speed and number of calculations that can be done even with today's supercomputers, some approximations and simplifications must be made to keep the computational cost affordable. Currently, using parallel computing [16-18] it is possible to simulate systems with as many as millions of atoms [17, 19]. Other than the growth in the system size, accessible time scales have been also expanded over the years. For example, on the highly specialized Anton computer, using 512 nodes, a millisecond simulation of a system including 23,558 atoms takes slightly more than two months [16]. Because this much computer power is not always accessible, alternative techniques to decrease the computational cost of MD simulations have been extensively sought out over recent years. One approach to meeting this   challenge   involves

simplification of the system so that the number of calculations per time step is reduced. For example, instead of using an atomistic model for the solvent, one can develop an implicit model that approximates all the environmental effects for any peptide or protein [20, 21].

Another advantage of implicit solvent is the lack of solvent relaxation. Relaxation times for lipid molecules are significantly longer than for water molecules (μs for the lipids [22, 23] vs. ps for the water molecules [24, 25]). Therefore, implicit representation of the membrane is especially appealing in MD simulations [26]. Further simplification of a given system is possible with coarse-grained models where the resolution of the system is reduced from standard atomistic representations [27, 28].

In order to extract meaningful thermodynamic properties from MD, the simulations should converge, which means that all of the relevant conformational space should be sampled with correct Boltzmann distribution weights. Other than using the implicit solvent/membrane models mentioned before, enhanced sampling methods can also help with achieving faster convergence in MD [29]. Replica exchange methods (REMD) [30] are among the most popular enhanced sampling techniques. In REMD, several copies (replicas) of the simulation are run in parallel, each with slightly different conditions. The details of REMD will be discussed later in chapter 4.

In the remainder of this chapter, the theory and methodologies of MD simulations and implicit solvent methods will be discussed. The implicit membrane method (HDGB) will be introduced in section 1.5, which was used for most of the simulations described in this work. MD simulations of phospholamban will be discussed in chapters 2 and 3. New developments of implicit solvent/membrane methods will be discussed in later chapters. In chapter 4, a new replica exchange technique is introduced and preliminary work on extending the current implicit

membrane model with a treatment of the nonpolar contribution of the solvation free energy is described in the appendix.

1.1 Molecular dynamics simulation

Molecular dynamics (MD) simulations explore the time-dependent behavior of a given system based on Newton's equation of motion. In all-atom MD simulations, biological systems are modeled in atomistic detail [1, 31]. Atomic forces (F_i) are calculated from a given expression for the potential energy (U) according to Equation 1-1

$$F_i = -\nabla_i U = m_i a_i \hspace{4cm} \text{1-1}$$

where m_i and a_i are the mass and acceleration of atom i respectively. At each MD step, new atomic positions and velocities are calculated by integrating the equation of motion. The potential energy function describes all of the interactions that are present in the system. Usually the potential energy function consists of a sum over different types of interactions in the system such as bonds, angles, dihedrals, and non-bonded interactions (see Equation 1-2). Parameters in the potential energy function, such as atomic partial charges and equilibrium bonding parameters, are optimized based on experimental and ab initio data. These parameters together with the potential energy function then make up the force field. Several force fields are available today for simulations of biomolecules. The most widely used force fields are CHARMM [32, 33], AMBER [34], OPLS [35] and GROMOS [36, 37].

In the following section, the CHARMM force field that has been used in all of the simulations of this work is described in more detail.

1.2　CHARMM force field

The potential energy function used in the CHARMM force field [32, 38] is shown in equation 1-2:

$$
\begin{aligned}
U(\vec{R}) = &\sum_{bonds} K_b(b-b_o)^2 + \sum_{angles} K_\theta(\theta-\theta_o)^2 + \sum_{Urey-Bradley} K_{UB}(S-S_o)^2 + \\
&\sum_{dihedrals} K_\varphi(1+\cos(n\varphi-\delta)) + \sum_{improper} K_w(w-w_o)^2 + \sum_{residues} U_{CMAP}(\varphi,\psi) + \\
&\sum_{\substack{non-bonded \\ pairs}} \left\{ \varepsilon_{ij}^{min} \left[\left(\frac{R_{ij}^{min}}{r_{ij}}\right)^{12} - \left(\frac{R_{ij}^{min}}{r_{ij}}\right)^6 \right] + \frac{q_i q_j}{4\pi\varepsilon_o \varepsilon r_{ij}} \right\}
\end{aligned}
\qquad 1\text{-}2
$$

The basic assumption is that different terms of the potential energy are additive. The individual terms can be divided into internal and non-bonded terms. The internal term includes the bond, valence angle, Urey-Bradley, dihedral angle, improper angle, and CMAP terms. The bond and valence angle terms, the first and second terms in Equation 1-2, are simple harmonic potentials and describe the potential caused by deviations from the equilibrium bond length and angle values (See Figure 1-1). In the CHARMM force field, a harmonic potential also restrains 1-3 distances with a so-called Urey-Bradley potential (See Equation 1-2). The dihedral term involves four consecutively connected atoms to describe rotations around the central bond in a periodic manner, as shown in Equation 1-2. The improper dihedral angle term is used to limit out-of-plane distortions of aromatic rings or sp^2 carbons attached to a carboxylate group (See Figure 1-2). Torsion angles are further modeled with a grid-based cross-correlation term, called CMAP, that is added to the CHARMM force field as a correction term for peptide backbone dihedral angles [32, 38]. Non-bonded interactions are described by a Lennard-Jones potential to

5

model hard-sphere repulsion and van der Waals attraction as well as Coulomb's law to represent electrostatic terms (see Equation 1-2).
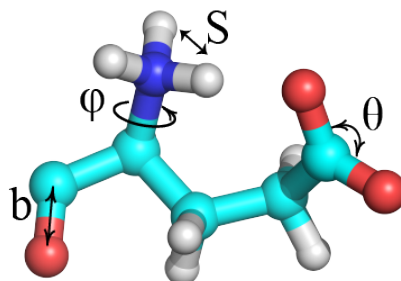
Figure 1-1 Illustration of bond (b), valence angle (θ), torsion angle (φ) and Urey-Bradley (S) contributions in the potential energy function. For interpretation of the references to color in this and all other figures, the reader is referred to the electronic version of this dissertation.
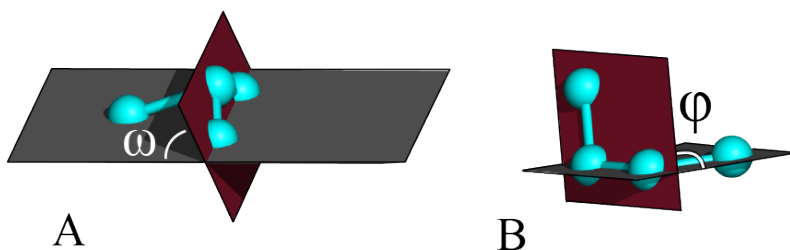
Figure 1-2 Illustration of A) improper, and B) dihedral angle.

1.3 Implicit solvent models

Simulating in vacuum is the easiest and least expensive method to study proteins and peptides. However, solvent can severely affect structure, dynamics and function of biological molecules [39, 40]. Water can form hydrogen bonds with the solute. In addition, dipole moments of water molecules shield charge-charge interactions between solute atoms. Explicit inclusion of water molecules is the obvious solution, but results in a dramatic increase in the system size. For example, in a recent MD simulation of the complete satellite tobacco mosaic

virus, 299,855 water molecules are necessary to fully solvate the 165,290 atoms of protein and nucleic acid. Expanding the system size affects the computational cost in two ways. First, the number of non-bonded interactions in a system increases with the order of N^2, where N is the number of atoms. And second, with a large number of water molecules in the system, most of the computational time is spent on sampling the solvent degrees of freedom in addition to the solute. An alternative technique to avoid these issues with explicit solvent is the use of implicit solvent that captures the mean field effects of the solvent environment [21]. There are several types of implicit solvent methods with various levels of approximation [20]. Some techniques approximate solvent effects with very simple models and keep the computational cost on the order of vacuum simulations. For example, the implicit model of Fraternali *et. al.* which is based on the solvent-accessible surface area and will be described later, is only 30 percent slower than vacuum simulations [41]. However, usually there is a trade-off between how fast and how accurately a given method can reproduce the explicit solvent results [42, 43]. More accurate implicit solvent methods are based on solutions to the Poisson equation [44] and are usually used as a benchmark for other implicit solvent methods.

In the following, some of the implicit solvent models are discussed starting with the most simplified methods.

1.3.1 Implicit solvent models based on dielectric screening functions

The main idea of the methods in this category is to express the screening of the charge-charge interactions by means of a linear [45-47], or non-linear [48, 49] dielectric screening function. The simplest models involve a distance-dependent dielectric constant ε in the

Coulomb expression that reaches bulk values when charges are further apart and more water molecules can screen the charges. The speed of these methods is in the order of the vacuum simulations but they cannot reproduce the solvation free energies in good agreement with explicit solvent simulations. The poor performance of such techniques is mainly the result of an inability to distinguish between solvent-exposed and buried atoms of the solute. In order to improve this method, Lazaridis *et al*. introduced EEF1 model [50] where the solvation free energy was decomposed into a self-energy component and screened charge-charge interactions. The self-energy component of the solvation free energy for atom *i* was defined as the solvation free energy of that atom when it was fully solvent exposed minus the effect of the surrounding groups or the solvent-excluded volume (V_j in equation 1-3).

$$\Delta G_i^{slv} = \Delta G_i^{ref} - \sum_{j \neq 1} f_i(r_{ij})V_j \qquad\qquad 1\text{-}3$$

Charge-charge interactions were screened with the distance-dependent dielectric function $\varepsilon(r)=f(r)$, where r_{ij} is the distance between solute atoms *i* and *j*. Some modifications to this method were later introduced by Mallik *et al*. allowing *f(r)* to depend on the distance of each atom from the surface [51]. In other methods, alternative screening functions were implemented, in which different measures of the solvent exposure of each atom were considered [52, 53].

1.3.2   Implicit solvent models based on the solvent accessible surface area

The main assumption in models based on the solvent accessible surface area (SASA) is that the solvent exposed surface of each atom is a measure of its interactions with the solvent

[54, 55]. All of the solute-solvent interactions, including the electrostatic and Lennard-Jones contribution, can be approximated with a linear function of SASA (see Equation 1-4).

$$G_{solvation} = \sum_i \gamma_i SASA_i$$
                                                                                        1-4

In Equation 1-4, γ_i is the atomic surface tension, usually obtained from fitting Equation 1-4 to experimental data. Wesson *et al.* implemented the idea in the CHARMM program [53] and parameterized their potential for five different atom types using Wolfenden's vacuum to water transfer free energies of the amino acid side chain analogs [56]. Schiffer *et al.* implemented the same method in the AMBER force filed [57].

In the implementation of this method in the GROMOS force field by Fraternali *et al.*, only two atom types, one for carbon and one for nitrogen or oxygen, were considered for representing the hydrophobic and hydrophilic properties, respectively [41]. The speed of their method, which is only 30 percent slower than vacuum, is due in part to a fast approximate calculation of SASA following a method developed by Hasel *et al.* [58]. Similar methods have been proposed by other groups [41, 58-60].

1.3.3 Implicit solvent models based on the hydration shell volume

The solvation free energy in the methods of this category is approximated to be proportional to the solvent exposed volume of the first solvation shell [61] as:

$$G_{hyd} = \sum_i \delta_i (VHS)_i$$
                                                                                        1-5

where δ_i is the empirical free energy of hydration density for atom i and VHS_i is the volume of the first solvation shell around atom i. Because of difficulties in calculating overlapping

hydration shell volumes, initial models have used approximations for estimating the hydration shell volume. Hopfinger *et al.* approximated the total hydration shell volume as a sum of subgroup hydration shell volumes [62]. Each subgroup had been minimized separately in the presence of solvent molecules to estimate the hydration shell volume. Kang *et al.* improved the method by mathematically calculating the exposed hydration shell volume for each group from their coordinates by calculating double and triple overlapping terms [63]. The main effort of such techniques has been focused on the accurate calculation of the excluded volume. However, because these methods are based on very simple physics, low computational cost is often accompanied by very rough approximations, which leads to deviations from the solvation free energy values calculated in explicit solvent simulations. More sophisticated methods that are based on stronger theory will be discussed in the next section.

1.3.4 Implicit solvent methods based on Poisson equation

In the Poisson methods, it is assumed that the polarization effects of the electric field that arise from the solute charges are linear and local. The electrostatic potential that is generated in response of the solvent to the charge density of the solute can be calculated through the Poisson equation as:

$$\nabla.(\varepsilon(r)\nabla\phi(r)) = -4\pi\rho(r) \qquad\qquad 1\text{-}6$$

In Equation 1-6, $\varepsilon(r)$ is the dielectric constant, $\varphi(r)$ is the electric potential, and $\rho(r)$ is the solute charge density. By calculating $\varphi(r)$ when the solute is placed in the aqueous solvent and in vacuum, and subtracting these values, the electrostatic contribution of the solvation free energy is calculated as:

$$\Delta G_{elec,solv} = \frac{1}{2}\sum_i q_i(\phi_i^{80} - \phi_i^1)$$

<div align="right">1-7</div>

Poisson methods are mainly different in the way they solve the Poisson equation. There are several algorithms for solving the PB equation among which the grid-based finite difference method is the most commonly used technique [64]. PB methods have been widely used in single point free energy [65-67] and pKa shift calculations [68, 69]. However, because of high computational cost, and the fact that smooth derivatives are difficult to obtain from the numerical solutions of the PB equation, their application in MD simulations has been limited [70, 71].

1.3.5 Generalized Born Methods

1.3.5.1   Main idea

In 1990, Still *et al.* introduced an empirical equation to approximate the PB-derived solvation free energies [72]. The solvation free energy of a single ion with the charge of q and radius of a is determined by the Born equation as:

$$\Delta G_{Born} = -\frac{q^2}{4\pi\varepsilon_o a}(1-\frac{1}{\varepsilon_w}) = -166\frac{q^2}{a}(1-\frac{1}{\varepsilon_w})$$

<div align="right">1-8</div>

In Equation 1-8, q is in the unit of electron charges, a is in the unit of Å and ΔG is in the unit of kcal/mol. The electrostatic contribution of the solvation free energy for a system composed of several charges, can be approximated as the difference between the screened and non-screened Coulomb interactions (first term in Equation 1-9) added to the sum of the Born self-energies (second term in Equation 1-9).

$$\Delta G_{elec} = -332 \sum_{i}^{N} \sum_{j \neq 1}^{N} \frac{q_i q_j}{r_{ij}} (1 - \frac{1}{\varepsilon_w}) - 166 \sum_{i}^{N} \frac{q_i^2}{a_i} (1 - \frac{1}{\varepsilon_w})$$ 1-9

Still *et al.* tried to generalize Equation 1-9 by finding a function (f_{GB}) that satisfied Equation 1-10. They called it the generalized Born (GB) equation:

$$\Delta G_{elec} = -166(1 - \frac{1}{\varepsilon_w}) \sum_{i}^{N} \sum_{j}^{N} \frac{q_i q_j}{f_{GB}}$$ 1-10

They introduced f_{GB} in the form of Equation 1-11, which is the most widely used form of f_{GB}:

1-11

$$f_{GB} = \sqrt{r_{ij}^2 + a_i a_j \exp(-r_{ij}^2 / F a_i a_j)}$$

In Equation 1-11, α_i is the effective Born radius for atom i, $r_{i,j}$ is the distance between atoms i and j, and F is an empirical factor set to 4 [72]. When $r_{i,j}$ is zero, Equation 1-11 is reduced to the simple Born equation. The effective Born radius introduced in Equation 1-11, describes approximately how far each atom is from the surface of the solute. When an atom is buried inside the solute molecule, large Born radius in Equation 1-11 leads to small contribution of that atom in the total free energy of solvation. If there is a single charge in the center of a spherical solute, the Born radius is the radius of that sphere. However, if the solute is not spherical or the charge is not at the center of the solute, the calculation of the Born radius is more complicated. One can determine the Born radius for atom i by turning all the other charges off and solve the PB equation for the new charge density. By submitting the calculated free energy in Born equation and solving for the radius and repeating this procedure, it is possible to calculate Born radii for all the atoms. This means that Poisson equation would be solved N times with N being the total number of the solute atoms. Therefore, the challenge of practically employing the GB

formalism is to determine Born radii by means other than solving the Poisson equation. It has

been shown that if the Born radii are calculated by solving the Poisson equation as mentioned

above, the solvation free energies calculated from Equation 1-10 are close to the solvation free

energies calculated by Poisson equation [73]. However, it has also been shown recently that

individual pairwise solvation free energies calculated from Equation 1-10 may show significant

deviation from the PB calculated solvation free energies [74]. It has been suggested that the

agreement between PB and GB methods (with correct Born radii) is the result of error

cancelation in the pairwise sum [74]. In order to reduce the individual pairwise errors, Onufriev

et al. proposed an improved method which unlike Still's formalism, not only depends on the

Born radii, but also is a function of the gradient of Born radii to mimic the deviations from

spherical shape of the solute [74].

1.3.5.2   Coulomb Field Approximation

In continuum solvent methods, at the charge density boundary where the solute

separates from the solvent, the dielectric constant varies sharply. The solute charge distribution

induces polarization on the solvent molecules and creates a reaction field on the solute

molecules in return. In GB methods, this induced reaction field and the related electric

displacement are ignored. In other words, it is assumed that the electric displacement is only

Coulombic in form and it does not change near the boundaries. This is called the Coulomb-field

approximation (CFA). As a result, the required work to bring a single charge, q_i, in a molecule

with the interior dielectric constant of ε_{int}, into the solvent with the dielectric constant of ε_{out}

can be written as Equation 1-12 where all other charges of the solute are turned off.

$$W_i = \frac{1}{8}\int(\frac{D}{\varepsilon}).D\,dV \approx \frac{1}{8\pi}\int_{in}\frac{q_i^2}{r^4\varepsilon_{in}}dV + \frac{1}{8\pi}\int_{out}\frac{q_i^2}{r^4\varepsilon_{out}}dV \qquad \text{1-12}$$

The electrostatic contribution of the solvation free energy for solvating this single charge is the difference between the work that is required to bring the charge from vacuum with the dielectric constant of 1 into the solvent with the dielectric constant of ε_{solv}:

$$\Delta G_{solvation,electrostatic} = W_{i(\varepsilon_{out}=1)} - W_{i(\varepsilon_{out}=\varepsilon_{solv})} = -\frac{1}{8\pi}(1-\frac{1}{\varepsilon_{solv}})\int_{out}\frac{q_i^2}{r^4}dV \qquad \text{1-13}$$

By comparing Equation 1-13 with the Born equation (shown in Equation 1-14)
1-14

$$\Delta G_{Born} = -\frac{q^2}{2\alpha}(1-\frac{1}{\varepsilon_w}) \qquad \text{1-14}$$

It is possible to formulate Born radius as Equation 1-15 [75]:

$$\frac{1}{\alpha_i^{-1}} = \frac{1}{4\pi}\int\frac{1}{r^4}dV = \frac{1}{R_i} - \frac{1}{4\pi}\int_{in,r>\alpha_i}\frac{1}{r^4}dV \qquad \text{1-15}$$

In Equation 1-15, α_i is the Born radius, R_i is the van der Waals radius of atom i, and the integration over the space outside of the solute is rewritten as the difference between the region where $\infty > r > \alpha_i$, and the interior region of the solute.

In an alternative approach, Ghosh $et\ al.$ converted the volume integral in Equation 1-15 to a surface integral by using the Green's theorem [76]. In order to improve the accuracy of their method, Ghosh $et\ al.$ also added empirical correction terms to the solvation free energy in their GB equation [76].

Grant $et\ al.$ tried to avoid CFA by introducing an exclusion function to scale the Born radii [77]. This function described how much an atom was masked by other atoms and how

much it was exposed to the solvent. The main reason for introducing such a function was to correct any polarization effects rising from the neighboring atoms that replaced the solvent molecules [77]. Haberthür *et al.*, introduced a different functional form for the Born radii [78]. In their method, called FACTS, Born radius was described as a function of the solvent volume around each atom and the symmetry of the surrounding solvent [78].

1.3.5.3   GBMV Method

GB methods are traditionally divided into two different classes based on the method they use to solve Equation 1-10. The first class takes advantage of surface/volume integration to solve the GB equation [72, 76, 77, 79-84]. The second class uses pairwise summation over all the atoms instead [75, 78, 85-96]. Here, only the first class of GB methods will be discussed.

In order to reduce the errors due to CFA, Lee *et al.* introduced a GB method based on molecular volume integration, called GBMV [79], with an empirical correction term in the definition of the Born radius:

$$A_5 = P \sqrt{\frac{1}{2R_i^2} - \frac{1}{4\pi} \int_{in,r>\alpha} \frac{1}{r^5} dV}$$ 1-16

$$\alpha_i = \frac{1}{-A_4 + A_5} + \alpha_O$$ 1-17

The term A_4 in Equation   1-17 is the CFA term from Equation   1-15. In Equation   1-16, P and α are adjustable parameters. In their method, Lee *et al.* introduced an analytical form to calculate the molecular volume that was based on superposition of exponential atomic functions (Equation   1-18) through a switching function (defined in Equation   1-19). This switching function guaranteed that the molecular volume decayed to zero at the surface of the molecule.

$$V_{MV}(r) = \frac{1}{1 + \exp(\beta[(\sum_j A_j(r)) - \lambda])}$$

1-18

$$A_j(r) = \exp(\frac{\gamma_o \ln(\lambda)}{R_j^2}|r_j|^4)$$

1-19

In the above equations, γ_0, λ and β are adjustable parameters. This method had been

implemented in the CHARMM force field and reported to be ten times slower that the vacuum

calculations for a protein with the size of about 60 residues [79]. In terms of accuracy, the

correlation coefficient of the calculated Born radii versus PB values increased from 0.77 to 0.95

for a test protein with 226 residues [97]. Single-point solvation free energy calculations on a

data set of single chain proteins also showed a significant improvement where the correlation

with PB calculated energies increased from 0.69 to 0.99 [79]. Later on, Lee *et al.* improved their

method by introducing a higher order correction term to CFA [80] as:

$$A_7 = 4\sqrt{\frac{1}{4R_i^2} - \frac{1}{4\pi} \int_{in, r > \alpha} \frac{1}{r^7} dV}$$

1-20

$$\alpha_i = \frac{S}{C_o A_4 + C_1 A_7} + D$$

1-21

where R_i was the van der Waals radius of atom i and C_0, C_1, S and D were adjustable

parameters.

The other improvement in the GBMV method by Lee *et al.* was a vector based method

for approximating the molecular volume instead of the van der Waals-based volume that results

simply from the overlapping atomic spheres [80]. The need for such a modification arose from

the inability of the previous techniques to include gaps in the molecular volume calculations. To

construct the standard molecular volume (SMV), a solvent-accessible volume is first

constructed by superposition of atomic spheres. The radius of each atomic sphere is the sum of

the radius of a water probe and the van der Waals radius of that atom. In the next step, all of the water probes that their centers are on the surface of the solvent-accessible volume are curved out. With this technique, the surface regions that are not occupied by the solute atoms, yet are not large enough to be occupied by water molecules, are not included in the high dielectric region of the solvent. Initially, Lee *et al.* approximated SMV with a sum of atom-centered quadratic exponential functions [79]. The analytical SMV technique could not properly approximate the regions on the surface where is not occupied by water molecules or any solute atoms. When such an area is inside of a cluster of atoms, it is called a gap region and when it is outside of a cluster of atoms, it is called an open region. GBMV tends to overestimate the molecular volume in the open regions and underestimate it in the gap regions [79]. In order to solve this problem, Lee *et al.* introduced a scaling factor based on the vector summation of the atomic vectors. These vectors connect each point of interest to the center of each atom. Because of the vector-based nature of this modification, differentiating between the open s and gap regions became possible. In a gap region, the vectors added up destructively while in an open region this addition is constructive. This modification reduced the molecular volume in the open regions and enhanced it in the gap regions. The form of the new molecular volume function was:

$$S_{MV} = S_O \left[\sum_j F_{MV}(\|r_j\|) \right] \frac{\sum_j \|r_j\|^2 F_{MV}(\|r_j\|)^2}{\left\| \sum_j r_j F_{MV}(\|r_j\|) \right\|^2} \qquad \text{1-22}$$

where $\|r_j\|$ was the norm of the vector connecting any point of interest to the center of atom j, S_O was an adjustable parameter and F_{MV} was an atomic function defined as :

$$F_{MV}(r_j) = \frac{C_j^2}{[C_j + \|r_j\|^2 - R_j^2]^2}, \quad C_j = P_1 R_j + P_2 \qquad \text{1-23}$$

with P_1 and P_2 as adjustable parameters. The new method (called GBMV2) has been reported to be two times slower than the original GBMV method [80].

1.3.6 Other implicit solvent methods

Dzubiella *et al*. introduced a different implicit solvent approach where the total solvation free energy is decomposed into four different terms each of which is a function of the solvent exposed volume of the solute: A term for generating the cavity against the solvent pressure, a term for solvent rearrangements, short and long range van der Waals interaction term between solute and solvent, and finally, an electrostatic term. The solvation free energy is then minimized for the correct solvent exposed volume of the solute [98]. Unlike other methods, where pre-definition of the solute-solvent boundary is required for the free energy calculations, in this method, by minimizing the functional form of the total solvation free energy, boundary geometry would be produced as an output [98-100].

1.4   Hybrid implicit / explicit solvation models

In order to combine the benefits of implicit and explicit solvent methods, hybrid implicit/explicit solvent methods have been introduced [101]. Primary hydration shell method (PHS) [102] is an example for such hybrid methods. In PHS, a thin layer of explicit water molecules is generated around the peptide. A restrained potential is used in this shell to prevent

the water molecules from escaping from the hydration shell. Computational costs of this method have been reported to be considerably lower than the fully solvated systems. However, because the hydration shell is too thin, properties of water may not be realistic.

A different approach has been recently suggested by Fennell *et al*. [103]. In their method, called semi-explicit assembly (SEA), Fennell *et al.* defined a shell around the solute cavity where they placed virtual water molecules based on a series of pre-calculated explicit solvent simulations. The orientation and distance of each water molecule from solute atoms were estimated from the pre-calculated simulations. After replacing the water molecules at each MD step, all the solvent-solvent and solvent-solute interactions were calculated explicitly. The solvent outside of the hydration shell has been treated through continuous Onsager equation [104]. This method has been shown to be able to accurately calculate solvation free energies faster than the explicit solvent simulations [103].

The main problem with the hybrid implicit/explicit solvent methods is that the behavior of the implicit solvent near the implicit/explicit boundary is not realistic. The presence of explicit water molecules, near the boundary region, affects the bulk behavior of the implicit solvent in these regions. On the other hand, the implicit solvent media may affect the configuration of water molecules in the explicit region near the boundary and result in some deviations from the bulk water properties. In order to overcome these problems, one can include more water molecules in the explicit shell around the solute. However, by including more and more water molecules in the explicit solvent shell, the hybrid method gets more expensive where finally looks more like an explicit solvent model with a cut off equal to the hydration shell. The required relaxation time for the explicit water molecules in the hydration shell is another disadvantage of the hybrid explicit/implicit solvent methods.

A different approach in solving the problems with the boundary region is to define a transition method that can successfully switch the implicit and explicit solvent properties in the interface regions [105].

1.5   Treatment of membrane; HDGB formalism

As mentioned previously, highly accurate solvation free energies can be calculated using the GBMV2 method. An implicit description of other biological environment such as membrane bilayers was encouraged by such achievements. The simplest description of the membrane is a low dielectric constant slab with a thickness matching the hydrophobic region of the bilayer. Using such a simplified description of the membrane, Spassov *et al*. adjusted their GB method for the membrane environment [86]. The main assumption in their method was that the membrane had the same dielectric constant as the solute. This assumption let them extend the integration region in Equation   1-15 to the membrane/water interface and then divide the integration region into the membrane and solute regions. Spassov *et al*. approximated the integral over the membrane region with an empirical formula that was only a function of z, the distance from the membrane center [86]. Im *et al*. [106] also treated the membrane as an extension of the solute cavity and modified the volume function in their GB method accordingly. The problem with such simplified models is that these methods are limited to two-dielectric systems. Sharp changes of the dielectric constant at the boundary region is another problem [107]. To overcome the second problem, Ulmschneider *et al*. tried to modify the molecular volume function to smoothly change at the interface of membrane and water [85].

In order to present a more realistic picture for the implicit membrane, Tanizaki *et al*. [84] developed an extension to the GBMV2 method which was called the heterogeneous dielectric

20

generalized Born formalism (HDGB). Previously, Feig *et al.* showed that the GB solvation free

energies calculated with the GBMV2 method were in very good agreement with the PB derived

solvation free energies in water. However, nonlinear errors raised for other solvents with lower

dielectric constants [83]. Motivated by Kirkwood expression for the off-center charge reaction

field potential, Feig *et al.* suggested that the Born radius in Equation   1-21 should be a function

of ε_{water} as well. The modified Born radius equation would read as:

$$\alpha_i = \frac{1}{C_O A_4 + C_1 \left(\frac{3\varepsilon_w}{3\varepsilon_w + 2}\right) A_7} + D + \frac{E}{\varepsilon_w + 1} \qquad\qquad 1\text{-}24$$

with the adjustable parameters of C_O =0.3255, C_1=1.085, D=-0.14 and E=-0.15 [83]. Motivated

by Equation 1-24, Tanizaki *et al.* rewrote the GB equation for a membrane as:

$$\Delta G_{elec} = -166 \sum_i^N \sum_j^N (1 - \frac{1}{\varepsilon_{ij}}) \frac{q_i q_j}{\sqrt{r_{ij}^2 + \alpha_i(\varepsilon_i)\alpha_j(\varepsilon_j)\exp\left[\frac{r_{ij}^2}{F\alpha_i(\varepsilon_i)\alpha_j(\varepsilon_j)}\right]}} \qquad 1\text{-}25$$

In Equation   1-25, the dielectric constant is local and varies based on the location of each atom

through an apparent dielectric constant, ε':

$$\varepsilon_{ij} = \frac{1}{2}\left[\varepsilon'(z_i) + \varepsilon'(z_j)\right], \quad \varepsilon_i = \varepsilon'(z_i) \qquad\qquad 1\text{-}26$$

In order to determine ε', Tanizaki *et al.* considered several infinite slabs with different dielectric

constants that were stacked along the z direction to mimic the membrane. However, ε' at each

distance z from the membrane center was not defined simply by the dielectric constant of the

slab were the point of interest was located but as an effective dielectric constant. The effective

dielectric constant incorporates the effect of different dielectric slabs at a distance that causes

polarization at each slab boundary. The apparent dielectric profile was constructed by solving the

Born equation for ε, with a charged spherical probe with electrostatic solvation free energies

obtained from solutions of the PE for the probe in the multilayer dielectric environment [83]. The apparent dielectric profile introduced in HDGB method is sensitive to the number of membrane slabs and also to the radius and charge of the probe. The profile introduced by Tanizaki *et al.* was generated with a 1 Å radius probe and with one positive charge. They used a three layer system with dielectric constants of 2, 7 and 80 to specifically mimic DPPC bilayer [83] (See Figure 1-3).
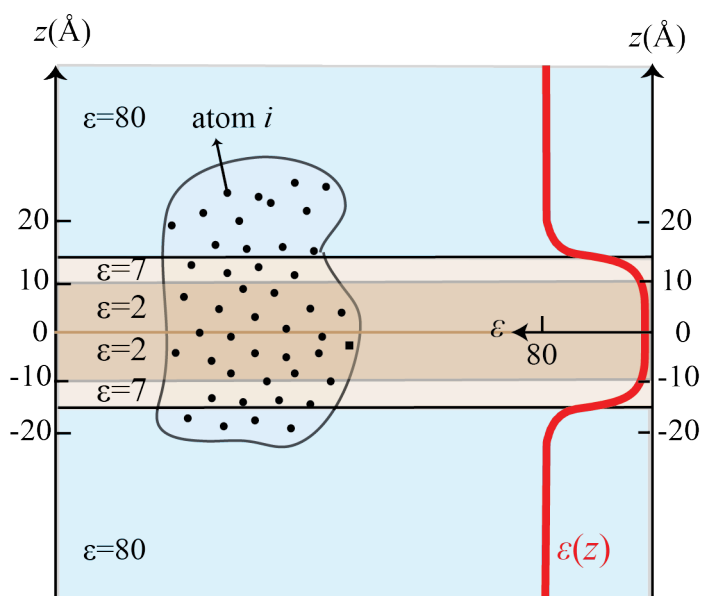
Figure 1-3 Heterogeneous model of a membrane bilayer in the HDGB method [84]. The solute keeps its atomistic details in the HDGB method. The dielectric profile is plotted in red along the z axis. The dielectric profile is generated from solving the Poisson equation for a system with three slabs. Each slab is represented in a box with a different color or shade. The blue color represents water and brown represents the membrane. The dielectric constant of each slab is shown in the figure.

An implicit treatment of membrane bilayers is not limited to the GB formalism. For example, in an extension of the EEF1 method, the distant-dependent dielectric constant function has been modified to change as a function of the distance from the membrane center [108].

Through out this dissertation, a re-optimized version of the HDGB formalism was used to study membrane proteins. The modification from the original HDGB formalism is described in chapter 2.

1.6   Non-polar contribution of the solvation free energy

As mentioned before, a solvent-accessible surface area term has been used to approximate implicit solvent environment in SASA-based methods [41, 53-55, 57, 59, 109]. This approach however, was later used to approximate the non-polar contribution of the solvation free energy, where electrostatic contribution was modeled by other means [72, 78, 85, 86]. Estimating the non-polar contribution of the solvation free energy with a functional form that is proportional to the solvent accessible surface area of the solute has been one of the most widely used methods to study solvation processes [72, 78, 85, 86]. In 1999, Lum *et al*. proposed that for hydrophobic solutes, this approximation is only correct if the solutes is large [110]. They proposed that for the small solutes, the solvation mechanism is different and the solvation free energy is proportional to the solute volume instead. In addition to the size effect, it was suggested that the solvation free energy is not a linear function of the solvent accessible surface area for alkanes [111-113]. A decomposition of the hydration process into enthalpic and entropic terms, suggested that while the entropy loss of solvent was linearly proportional to the solvent accessible surface area of the solute, the enthalpic gain of the solute-solvent interactions, or the van der Waals dispersion term, could not be approximated by a simple SASA model [114]. Gallichio *et al*. tried to model the

23

enthalpic and entropic terms, by fitting the solvation free energies of alkanes to the functional form of:

$$\Delta G_{np} = \sum_{i=1}^{N} \left[\gamma(t_i) A_i + \alpha(t_i) \right]$$ 1-27

where the first term is a SASA model term and $\alpha(t_i)$ is an atomic adjustable parameter [114]. It is assumed that the van der Waals dispersion term, which is independent of the solute shape, can be estimated through integration of attractive Lennard-Jones potential between solvent and solute, weighted with the number density of the solvent (see Equation 1-28) [115].

$$U_{vdW}(i) = \int \rho_w u_{vdW}^{(i)} (|r - r_i|) d^3 r$$ 1-28

This idea was further developed in the AGBNP method [87] by approximating the volume integral Equation 1-28 by the definition of Born radius. The result was the van der Waals dispersion term shown in Equation 1-29:

$$\Delta G_{vdW} = \sum_{i=1} a_i \frac{v_i}{(\alpha_i + R_w)^3}$$ 1-29

where α_i is the Born radius of atom i, v_i is an adjustable parameter and R_w is the probe radius or the shortest distance from the solute surface that is accessible for the solvent. R_w was set to 1.4 Å, the radius of a water probe. In Equation 1-29, a_i is the attractive Lennard Jones potential between all the solvent atoms and solute atom i. However, the effects of the solvent atoms are not explicitly defined in Equation 1-30. The mean attractive van der Waals term is determined by introducing ρ_w in the integration in Equation 1-28.

$$a_i = -\frac{16}{3} \pi \rho_w \varepsilon_w \sigma_{iw}^6$$ 1-30

24

In Equation   1-30, ρ_w is the number density of water molecules, and ε_{iw} and σ_{iw} were the

Lennard-Jones parameters describing the attractive van der Waals term between solute atom i

and water:

$$\varepsilon_{iw} = \sqrt{\varepsilon_i \varepsilon_w}, \quad \sigma_{iw} = \sqrt{\sigma_i \sigma_w}$$

1-31

Fennel *et al*. used a completely different approach to approximate $\Delta G_{solv,np}$ [116]. In their

method, they pre-calculate Lennard-Jones energies for different ε and σ values via an explicit

water simulation where all the charges were turned off. They defined a new solvent accessible

surface by adding the van der Waals radius of each solute atom to the average separation

distance between the solvent and that specific atom. These separation values were estimated

from radial distribution functions from the explicit solvent simulations. Fennel *et al*. defined a

vector (v), connecting each point on the new solvent accessible surface (point A) and the center

of the associated surface solute atom. They averaged the Lennard-Jones terms between a water

molecule at point A and the solute molecules along vector v, which were in a distant cut off from

the surface solute atom. The average Lennard-Jones term for each solute atom was then fitted to

a Lennard-Jones potential function to find the effective ε and σ. These effective Lennard-Jones

parameters were then used to calculate the non-polar contribution of solvation free energy.

Fennel *et al*. suggested that for poly aromatic hydrocarbons, where the SASA model fails to

approximate the non-polar contribution of the solvation free energies, their model successfully

captures the attractive Lennard-Jones contribution [116].

All the above-mentioned methods are used to approximate the non-polar contribution of

solvation free energy in the solvent. However, because most of the solvent exposed residues on a

peptide are hydrophilic, the electrostatic contribution of solvation free energy is usually the

dominant contribution to the total solvation free energy. In case of the membrane proteins, the weight of different contributions in the solvation free energy shifts significantly. Membrane-exposed side-chains are usually hydrophobic. This means that the electrostatic interactions between the membrane and solute are not as dominant as before and the contribution of the non-polar terms in the solvation free energy becomes more important in the membrane environment. Therefore, a good non-polar model for estimating the non-polar contribution of the solvation free energy is much more important for implicit membrane methods. In the implicit membrane methods mentioned before, the non-polar contribution of the solvation free energy is approximated by a SASA model. In the Im *et al.* method, membrane effects are captured by introducing a z-dependent membrane exclusion function in definition of atomic SASA [106]. Tanizaki *et al.* added a switching function to the SASA term in the HDGB method, which guarantees that the SASA term goes to zero at the center of the membrane and equals to one in the bulk water. SASA models cannot capture the total non-polar contribution of the non-polar interactions between solvent/membrane and solute. Therefore, in order to improve the HDGB method, modeling of the non-polar contribution of the solvation free energy should be modified. In the appendix, such a modification in the HDGB formalism is suggested by defining a second term in the non-polar contribution of the solvation free energy to approximate the attractive van der Waals interaction between the solute and solvent. A detailed description of this new term will be discussed thoroughly in the appendix.

1.7   Limitations of the implicit solvent methods

One of the problems with the GB formalism and implicit solvent methods in general is that solvent responds in the same way to positive and negative charges. However, in reality,

water molecules orient differently around positive and negative charges. The Born radius calculation is not sensitive to the charge sign. Therefore asymmetric behavior of water around charges with different signs must be captures by other means in the GB methods. One attempt to address this problem is to consider a few explicit water molecules at the places where the charge density is very high [117]. In a different approach, Born radii were adjusted based on the induced charge density on the surface [118]. Introducing dipoles in the first hydration shell has been also reported to be effective in differentiating between the solute's negative and positive charges [103].

An implicit representation of the environment may also be problematic in the context of membrane when insertion of explicit water molecules into the membrane has been shown to facilitate the existence of charged residue side chains in the low dielectric environment of membrane interior [119]. In the explicit solvent/membrane simulation, lipid chains can interact with the surface of the peptide explicitly. If there is a hydrophobic residue at the surface of the peptide near the membrane/water interface, lipid chains can move to some extent and shield the hydrophobic side-chain from the high dielectric solvent. Shielding a hydrophobic reside by explicit lipid chains or accompanying a charged side-chain by water molecules into the membrane can be approximated by a membrane model with flexible thickness.

The above mentioned problems are on top of the more general problem of uncertainty about the force fields which has been reported to sometimes over/under stabilize a specific conformation or a secondary structure for proteins [120]. In addition, for studying specific interactions between the solvent or membrane and the solute, the explicit representation of solvent/membrane is required and clearly implicit solvent/membrane techniques cannot be used.

In spite of all these caveats, implicit membrane methods have been used successfully to study a wide range of systems [106, 121-125].

1.8   Phospholamban; A case study for the HDGB method

It has been known for many years that stimulation of the intact heart by β-adrenergic agonists increases the rate and intensity of the cardiac muscle contraction and relaxation processes [126]. It is believed that β-adrenergic stimulation affects contraction-relaxation of heart muscles by changing the calcium ion flux in and out of the sarcoplasmic reticulum (SR) [127] where calcium ions are stored. These effects are made via activation of cyclic adenosine monophosphate (cAMP) with the biological role of a second messenger. Through the β-adrenergic stimulation, sarcoplasmic reticulum Ca^{2+}-ATPase isoform 2a, (SERCA) [128] a calcium ion pump on SR membrane, is being phosphorylated. In the Ca^{2+} uptake cycle, SERCA exchanges conformation between the E2 and E1 states, with low and high affinity for Ca^{2+} ion respectively. SERCA's activity is regulated by a transmembrane protein named phospholamban (PLB) [129] (see Figure 1-4). The name "phospholamban", which means "phosphate receptor", refers to the protein undergoing phosphorylation at Ser16 by cAMP-dependent protein kinase (PKA) [130]. Phosphorylation of PLB at Ser16 has been found to be linked to the level of Ca^{2+} ion uptake stimulated by cAMP [127, 131]. In the unphosphorylated form, PLB inhibits SERCA and phosphorylation of PLB releases the inhibitory effect. PLB has been reported to be the substrate for another protein kinase in SR, calcium/calmodulin-dependent protein kinase (CaMKII). Although CaMKII phosphorylation of PLB at Thr17 is reported to show additive effects on Ser16 phosphorylation [132, 133], some studies proposed that Ser16 phosphorylation

is sufficient for maximal stimulation of SERCA activity [134, 135]. In spite that in vitro studies suggested independent mechanisms for Ser16 and Thr17 phosphorylation of PLB [132], all the attempts to phosphorylate PLB by CaMKII in intact heart cells failed unless in cases where the cAMP level of the cell was high [136-138]. It is suggested that Thr17 phosphorylation of PLB with CaMKII occurs only in two cases: When the Ca^{2+} ion concentration is high and the phosphatase that deactivates CaMKII is inhibited, or when the cAMP level in the cell is high which increases the Ca^{2+} ion concentration in the cell [139]. Knowing more about the mechanism of releasing the inhibitory effect of PLB and the physiological role of Ser16 and/or Thr17 phosphorylation requires more detailed structural knowledge about PLB, phosphorylated PLB and also PLB-SERCA interaction site. Cross-linking studies have suggested close interactions between SERCA and PLB. Solid state NMR studies of the transmembrane domain of PLB suggested that the mobility of PLB is significantly reduced in the presence of SERCA [140]. These findings suggested a direct interaction between SERCA and PLB. However, details about the interaction site of these two transmembrane proteins are not clear yet. Though the crystal structure of SERCA in both E1 and E2 from is available [141, 142], there are still some questions about the structure of PLB. Solution and solid-state NMR studies of the monomeric PLB mutant reported an L-shaped structure for PLB, which consists of a transmembrane (TM) helix that is attached to a cytoplasmic helix through a hinge [143-145]. The transmembrane domain of PLB is proposed to be able to fit into the transmembrane groove of SERCA in the E2 state [146]. This groove gets smaller in size by the E2 to E1 transition of SERCA. It has also been shown that while the cytoplasmic domain of PLB is not sufficient to inhibit SERCA [147], TM domain of PLB exhibits the same inhibitory function as the full length PLB [148]. Cytoplasmic domain of PLB, where the phosphorylate-able residues Ser16 and Thr17 are placed,

has shown significant dynamics. It is suggested that the cytoplasmic domain exists in equilibrium between the extended and bent conformations. In the presence of SERCA, this equilibrium shifts toward the extended structures [149].
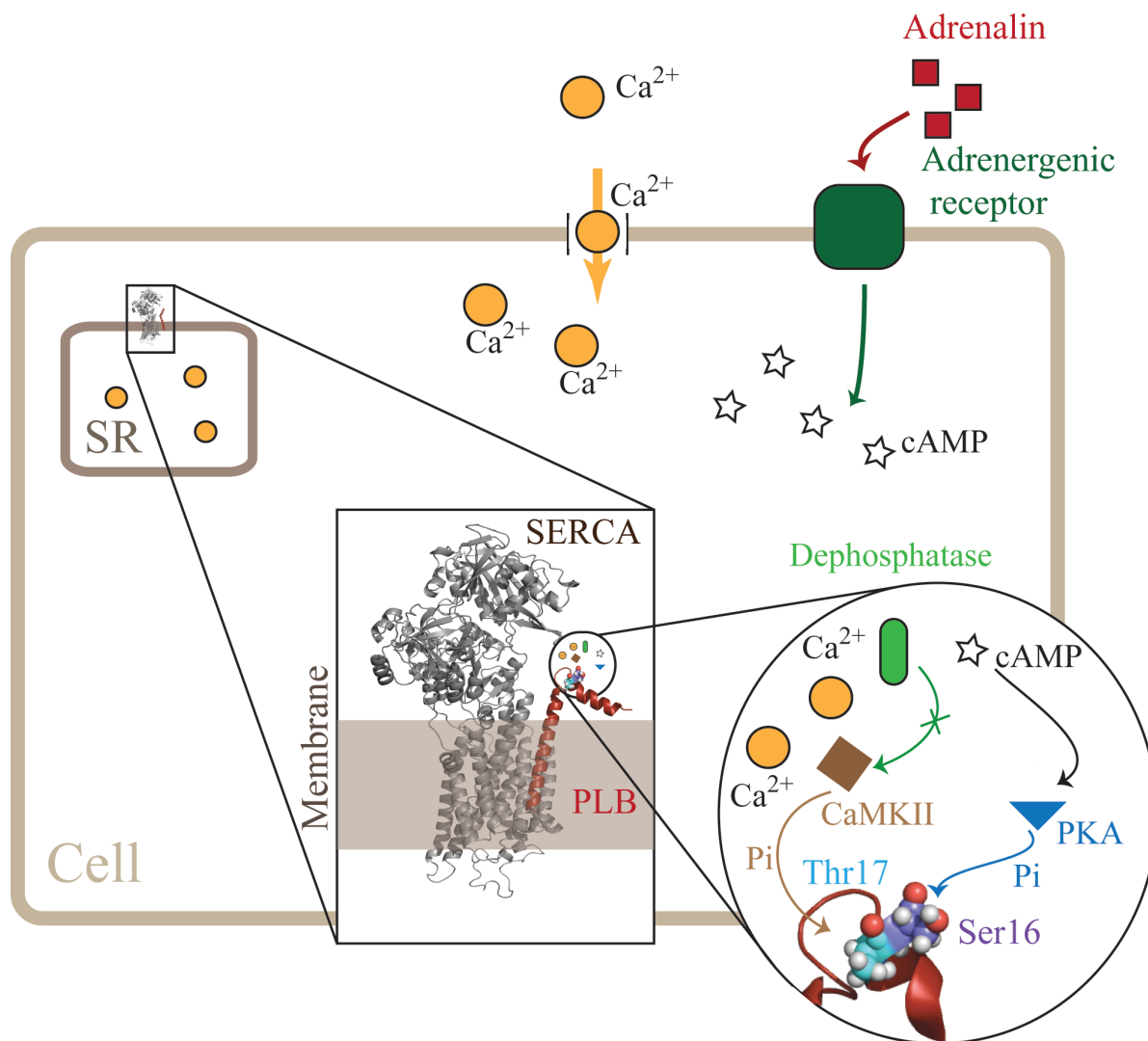
Figure 1-4 Schematic representation of PLB phosphorylation process. Different mechanisms of PLB phosphorylation at Ser16 and Thr17 are also shown in the figure.

PLB has been detected in both pentameric and monomeric forms [150, 151]. Because the monomeric mutant of PLB exhibits super inhibitory effects, is has been suggested that the active oligomeric form of PLB is the monomeric conformation [129, 152]. Other experiments have also shown that in the presence of SERCA, PLB is depolymerized into its monomeric state [152, 153].

MD simulations are useful for studying dynamic structures. In case of PLB, the reported conformation consists of an ensemble of 20 structures from solution NMR data in DPC micelles [143]. The flexibility of the loop and the lack of enough solution NMR data for the residues in this region are the main reasons that solution NMR studies are not sufficient to provide tertiary structural information for PLB. NMR orientational restrains obtained from the solid state NMR studies of PLB [154] were used in a refinement protocol [155] to introduce a new ensemble of structures for PLB. However, as mentioned before, the weight of each structure in the whole ensemble is not clear. Moreover, the NMR restrains are average properties. This means that it is possible to have a new ensemble of structures, with average structural properties that match the NMR data while the individual structures do not satisfy the structural restrains. In such a case, MD simulations could be very helpful in generating a weighted ensemble of structures where the average structural properties are in agreement with the NMR data. In chapters 2 and 3, MD simulations are used to study conformational sampling of PLB and phosphorylated PLB. Because of the implicit representation of the membrane in the HDGB method, it is easy to increase or reduce the membrane thickness by scaling the ε and non-polar profiles. Chapter 2 is devoted to the study of the effects of changing the membrane thickness on PLB conformational sampling and comparing the average structural restrains with the NMR data. In chapter 3, the conformational sampling of Ser16 and Thr17 phosphorylated PLB is being discussed.

2 Effect of membrane thickness on conformational sampling of phospholamban from computer simulations [156]

2.1   Abstract

The conformational sampling of monomeric, membrane-bound phospholamban is described from computer simulations. PLB plays a key role as a regulator of sarcoplasmic reticulum calcium ATPase (SERCA). An implicit membrane model is used in conjunction with replica exchange molecular dynamics simulations to reach μs-ms time scales. The implicit membrane model was also used to study the effect of different membrane thicknesses by scaling the low-dielectric region. The conformational sampling with the membrane model mimicking DPPC bilayers is overall in good agreement with experimental measurements, but consists of a wide variety of different conformations including structures not previously described. The conformational ensemble shifts significantly in the presence of thinner or thicker membranes, which has implications for the structure and dynamics of PLB in physiological membranes and offers a new interpretation of previous experimental measurements of PLB in detergents and microsomal membrane.

2.2   Introduction

Calcium plays an essential role in muscle contractions. During muscle relaxation, sarcoplasmic/endoplasmic reticulum (SR/ER) calcium ATPase (SERCA) pumps Ca^{2+} from the

cytosol to the lumen across the SR/ER membrane [129]. The activity of the cardiac isoform of Ca^{2+}-ATPase, SERCA2a, is regulated by phospholamban (PLB), a 52-amino acid transmembrane (TM) protein [129]. The interaction with unphosphorylated PLB is believed to lock SERCA into the reduced calcium affinity E2 form thereby inhibiting Ca^{2+} transport [157]. Inhibition is released when PLB is phosphorylated at S16 and/or T17 [132]. Phosphorylation of PLB is proposed to either result in dissociation of PLB from SERCA or in an altered interaction between PLB and SERCA so that SERCA can switch from the inactive E2 form to the active E1 form with high Ca^{2+} affinity. However, details of this regulatory mechanism remain unclear [158]. The importance of PLB in heart muscle function is underscored by a link between PLB mutations and heart failure in humans[159]

As a first step towards understanding the inhibitory mechanism of PLB, many studies have focused on the structural analysis of PLB without SERCA. PLB shows a strong tendency to oligomerize and in particular to form pentamers [152, 160]. However, most experimental evidence based on electron paramagnetic resonance (EPR) [161] fluorescence [162], and mutagenesis [152, 157] studies suggest that the monomeric form of PLB is primarily responsible for inhibition of SERCA. Therefore, most studies of PLB are concerned with the monomeric form.

The C41F mutant of PLB has a low propensity to form pentamers but retains biological activity. NMR studies of this mutant in CHCl3/MeOH indicate the presence of two α-helices connected by a β-turn suggesting an L-shaped structure [163]. Solid-state NMR (SSNMR) studies of the triple mutant in a mixed bilayer [145] suggest an interhelical angle of $80 \pm 20°$ consistent with L- or T-shaped conformations. A slightly different value of ~66° was obtained

using multidimensional SSNMR[144]. Furthermore, solution NMR-derived structures of C36A/C41F/C46A mutants in dodecylphosphocholine (DPC) micelles show three distinct structural domains for PLB [149]: a short cytoplasmic (CP) helix (domain Ia, residues 1-16), a hinge with a β-turn type III conformation (residues 17-22), and a long hydrophobic transmembrane (TM) helix (domains Ib, 23-30, and II, 31-52) also with an interhelical angle of 80° [143]. Further studies of PLB with multidimensional SSNMR and hybrid solution NMR and SSNMR suggest that the CP helix adopts angles of 93-102° with respect to the membrane normal while the TM helix is tilted with an angle of 22-24º [144, 154].

The static models of PLB described so far are contrasted by evidence of significant conformational dynamics. There is evidence of two topologies for the TM domain from solution [164] and SSNMR [144]. Two well-resolved conformational states of the CP domain have also been observed in solution NMR and EPR spectroscopy [165, 166] One conformation, called the T state, is ordered and in direct contact with the membrane surface while the other conformation, the R state, is dynamically disordered and detached from the membrane. Furthermore, fluorescence resonance energy transfer (FRET) measurements suggest that the CP domain of PLB in the absence of SERCA assumes a wide range of structures relative to the TM domain [167].

Molecular dynamics (MD) simulations provide atomistic insight into the structural information and dynamics of PLB. Several MD simulations of partial or full length PLB in different environments have been reported so far [168-172]. One of the first MD simulations of full length PLB in DPPC bilayer, water and methanol maintained the two well-defined TM and CP helical domains over 5-10 ns, and showed large-amplitude rigid-body motions of these domains, in particular for the CP domain[172].

Replica exchange MD simulations of the CP domain of PLB and phosphorylated PLB (pPLB) as well as the R9C mutant in explicit water showed the importance of P21 in maintaining the local helical structure of the CP domain of PLB [169]. Constant-temperature MD simulations of full-length PLB and pPLB for 30 ns, suggested that the TM helix and the lipid bilayer have only a minor effect on the conformational sampling of the CP domain of PLB [168]. Furthermore, a 15-ns MD simulation of PLB in POPC bilayer confirmed the presence of dynamic motions involving the CP domain of PLB and preservation of mostly helical structures for the CP and TM domains [170]. Finally, a recent over 15 ns MD simulation study of PLB in explicit POPC bilayer confirmed the helix-hinge-helix structure of PLB but found that the interhelical angle and the position of the CP helix relative to the membrane are highly dependent on the starting structure indicating that simulations of PLB over 10-ns time scales are not sufficient to fully describe the conformational dynamics of PLB [171].

Here, we describe replica exchange simulations of monomeric PLB with an implicit membrane model that effectively cover at least μs time scales and provide a more comprehensive view of PLB dynamics than previously reported simulations. Implicit membrane simulations have been used before to cover long-time dynamics in the interaction of peptides with membranes [106, 121-125]. The use of an implicit membrane also offers the unique advantage that the membrane thickness can be varied easily. For the first time, our simulation results suggest that the conformational sampling of PLB may indeed differ for very thin and very thick membranes, which has important consequences for the interpretation of conflicting experimental data and offers new insights into the mechanism by which PLB regulates SERCA.

2.3 Methods

2.3.1 Molecular Dynamics Simulations of Phospholamban

Replica exchange MD simulations [30] of monomeric PLB with the wild-type sequence were started from model 1 of the NMR ensemble for the C36A/C41F/C46A PLB mutant in DPC micelles (PDB entry 1N7L [143]). Standard protonation states (pH=7) and zwitterionic termini were used. The initial structure was oriented in the implicit membrane with the TM helix (domain II) parallel to the membrane normal and the CP helix (domain I) parallel to and above the membrane surface. The CHARMM22 all-atom force field [32] in combination with the CMAP correction term [38] was used to describe intramolecular interactions along with the heterogeneous dielectric generalized Born (HDGB) implicit membrane model to reflect the membrane environment. No cutoffs were applied to non-bonded interactions. Temperature replica exchange MD simulations were started from the energy-minimized initial structure with eight replicas spanning a temperature range of 300-400K in an exponential fashion. Exchanges were attempted every 500 MD steps. The acceptance ratio between adjacent replicas was between 22% and 33%. A time step of 1.5 fs was used to maintain stable simulations with the implicit membrane model [173]. Langevin dynamics [174] was used with a friction coefficient of 10 ps^{-1} [175] for all non-hydrogen atoms to control the temperature of the system. The SHAKE algorithm was used to constrain bond lengths involving hydrogen atoms [176]. All of the simulations were carried out with CHARMM [177], version 34a2 in combination with the MMTSB Tool Set [178] to facilitate the replica exchange simulations and analyze the simulation results.

2.3.2 Implicit Membrane Model

An implicit model of the phospholipid bilayer is applied here to reach long time scales that are otherwise inaccessible with conventional explicit lipid simulations. Implicit models reduce the computational time per time step but the more significant advantage comes from instant relaxation of the environment, which would otherwise incur kinetic barriers with an explicit representation, especially in lipid bilayer environments. In order to obtain a realistic representation of the physical characteristics of membrane environments, the HDGB model [179] consisting of a variable dielectric continuum along the membrane normal is applied. This model is based on the Generalized Born (GB) formalism and involves a modified description of the electrostatic component of $\Delta G_{solvation}$ [80]:

$$\Delta G_{elec} = -166 \sum_{i=1}^{n} \sum_{j=1}^{n} \left[1 - \frac{1}{(1/2)(\varepsilon_i - \varepsilon_j)} \right] \frac{q_i q_j}{\sqrt{r_{ij}^2 + \alpha_i(\varepsilon_i)\alpha_j(\varepsilon_j)\exp[-r_{ij}^2 / 8\alpha_i(\varepsilon_i)\alpha_j(\varepsilon_j)]}} \quad \text{2-1}$$

where q_i are atomic charges, n is the number of atoms, r_{ij} are interatomic distances, and α_i are the effective atomic Born radii. The dielectric constant (ε) for each atom in Equation 2-1 varies as a function of z, along the membrane normal, and reflects the effective dielectric constant at a certain membrane insertion depth. The ε profile was initially generated by solving the Poisson equation for a spherical probe at different locations in a system composed of dielectric slabs [179] and subsequently optimized to match free energies of insertion from experiment and explicit membrane simulations as described below (see Figure 2-1). Other parameters of the GB model were set as described previously [80, 83, 173, 179].
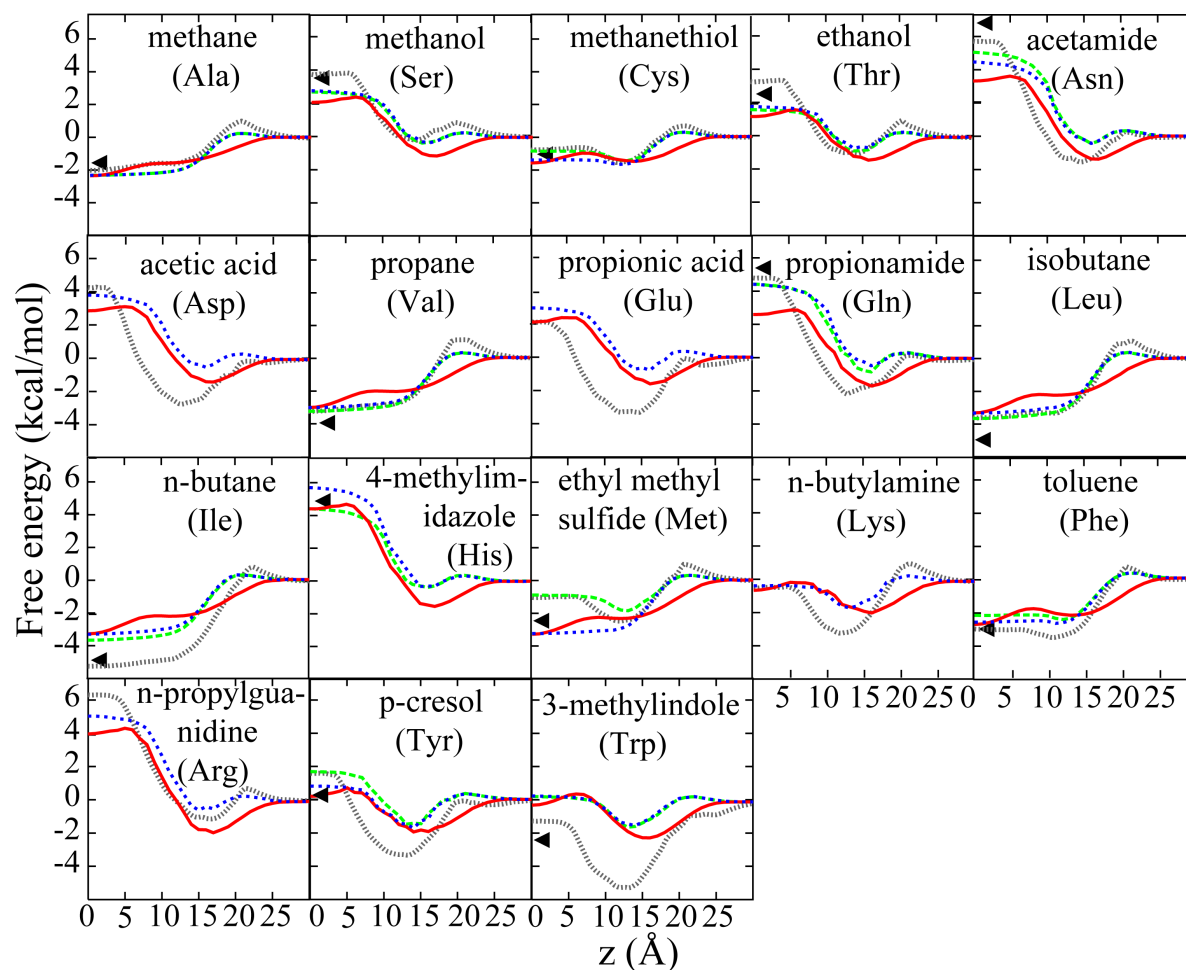
Figure 2-1 Free energy insertion profiles of amino acid side chain analogs calculated with explicit solvent simulations with OPLS charges [35](⋯⋯), original HDGB method (——), modified HDGB method with CHARMM charges [32](⋯⋯⋯), and modified HDGB method with OPLS charges (----). Experimental transfer free energies from cyclohexane to water are marked by ◄ [197].

The non-polar component of the solvation free energy is described by the solvent accessible area (SASA) model:

$$\Delta G_{non-polar} = \gamma \sum_{i=1}^{n} S(z_i)A_i \qquad\qquad 2\text{-}2$$

where A_i is the solvent accessible surface area of the i*th* atom, z_i is the distance of atom i from the membrane center along the membrane normal, γ is the empirical surface tension parameter set to 0.015 kcal/mol [179] and *S(z)* is a switching function to reflect the change of the surface tension along membrane normal. The *S(z)* function was initially determined from explicit lipid simulations for the insertion of O_2 into lipid bilayers [179, 180] and subsequently optimized along with the ε profile to match insertion free energies (see Figure 2-1).

2.3.3 Analysis of simulation results

The weighted histogram analysis method (WHAM) [181] was used to generate potential of mean force (PMF) maps from the replica exchange trajectories. The angle between the CP and the TM helices (interhelical angle, θ) and the distance between the center of mass of the CP helix and the projection onto the TM helical axis (d-COM) were chosen as the reaction coordinates for generating the PMFs (see Figure 2-2). The helices were described by residues 4-12 and 25-40 for the CP and TM helices, respectively. Only part of the complete TM helix was used in this definition in order to reduce the effect of local bending within the TM helix on the calculated θ. Other structural properties were calculated using slightly different definitions to match experimental data as described in the caption of Table 2-1. Clustering of PLB conformations was carried out with the K-means method implemented in the MMTSB Tool Set [178]. Pymol was used to visualize the simulation results and generate molecular graphics [182].
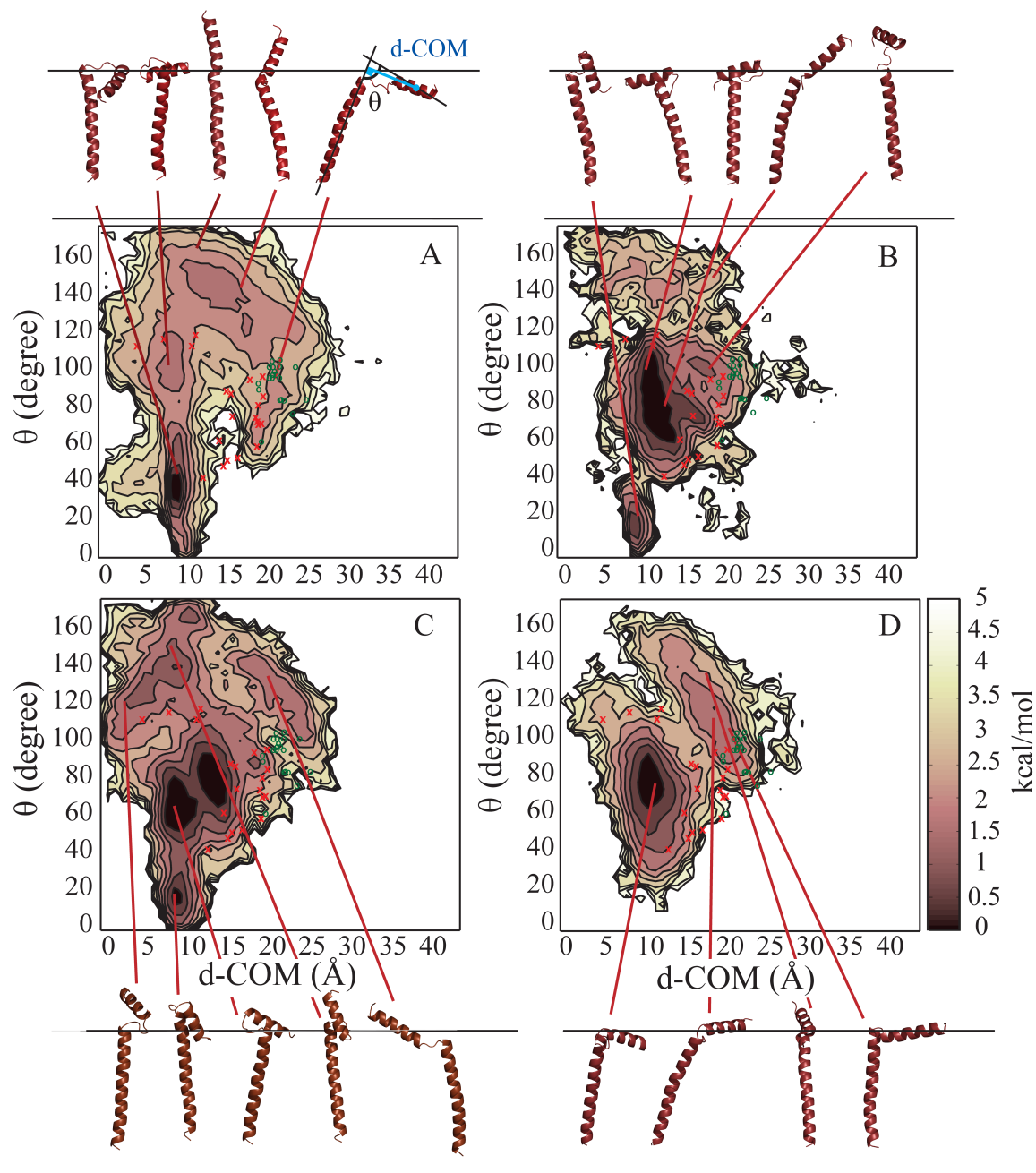
Figure 2-2 Energy landscape for PLB in different membrane models; A) MM1, B) MM2, C) MM3, and D) MM4. Representative structures for the clusters corresponding to the minima are shown in cartoon representation. NMR structures with PDB codes 1N7L and 2KB7 are mapped and shown as red x's and green o's respectively [143, 154].

Membrane model	MM1	MM2	MM3	MM4	Experiment
Interhelical angle, θ (degrees)	81±13	87 ±11	83 ±12	82 ±10	68 ±23[143] 80 ±20[145]
TM tilt, α_{TM} (degrees)	19.5±2.7	13.3 ±1.9	13.6 ±1.8	13.0 ±1.7	21.2 ±2[144] 24±2[154]
CP tilt, α_{CP} (degrees)	96.4±14	96.4 ±12	101.2 ±14	98.6 ±18	93 ±6[144] 102±2[154]
Interdomain distance (Å)	16.1±1.7	18.4±1.4	15.1±1.4	13.4±1.4	~20 [158]
TM insertion (Å)	7.3±0.7	5.6±0.6	5.8±0.6	5.7±0.6	5.5±0.5 [154][†]
CP insertion (Å)	27.8±1.6	30.1±1.2	31.3±1.2	29.5±1.3	16.2±0.8[154][†]
TM rotation, ρ_{TM} (degrees)	213±14	196±10	215±8	208±7	~198 [154]
CP rotation, ρ_{CP} (degrees)	187±16	209±16	136±8	229±8	92±3 [154][†]

Table 2-1 Average structural properties of PLB calculated from simulations with different membrane models. Experimental values in DPC micelles [143], mixed DOPC/DOPE or DOPC lipid bilayers [144, 145, 154], and liposomes consisting of SR lipids [158] are given for comparison. Values reported in [154] as a result of refinement instead of direct experimental measurements are marked by †. Tilt angles α_{TM} and α_{CP} are with respect to membrane normal, the interdomain distance is the distance between C_α atoms of residues Y6 and C24 [158], TM and

Table 2-1 (Cont'd)

CP insertion depths are based on the center of mass for residues 23-52 (TM helix) and residues 1-16 (CP helix) relative to the membrane center (z=0), and rotation angles ρ_{CP} and ρ_{TM} are defined as previously described [154] with helical axes obtained by averaging backbone N(i)-O(i+4) vectors for residues 23-38 (TM helix) and residues 3-10 (CP helix). Only the upper part of the TM helix was considered to avoid artifacts due to TM helix bending. Errors for simulation results are obtained from comparing block averages.

2.4   Results

2.4.1 Optimization of implicit membrane model

The HDGB implicit membrane model was initially parameterized based on limited reference data from experiments and explicit lipid insertion profiles. Since the original publication of the HDGB model, additional high-quality amino acid side chain analog insertion profiles from explicit lipid bilayer simulations have become available [183]. This data was used to optimize the previously published ε and non-polar profiles [179]. The resulting optimized profiles are depicted in Figure 2-3 and a comparison of implicit and explicit membrane insertion profiles is shown in Figure 2-1. Overall, the agreement between the implicit and explicit insertion profiles is remarkably good, especially when the OPLS force field [35] is used with HDGB method to match the explicit lipid profiles that also used the OPLS force field. At the same time, the free energies of insertion predicted with the implicit model agree well with experimental water-cyclohexane transfer free energies [184]. There are some discrepancies for polar residues that exhibit interfacial minima near 15 Å due to the development of water defects

42

in explicit lipid simulations [183]. While water defects are not possible in the HDGB model this is expected to be less of a problem for peptides and proteins interacting with the lipid bilayer where the development of water defects around a polar residue is likely to be discouraged by neighboring hydrophobic residues.
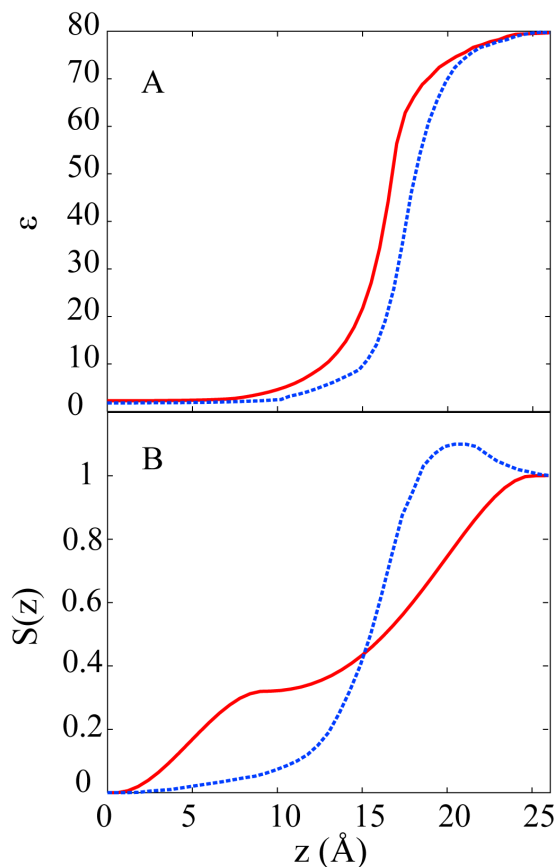
Figure 2-3 Dielectric and nonpolar profiles. The optimized profiles are in blue and the original profiles are in red.

The optimized profiles shown in Figure 2-3 are appropriate for DPPC bilayer, but stretched or compressed profiles can be generated through simple scaling to examine the effect of different membrane thicknesses. The profiles in Figure 2-3 have a thickness of 26 Å for the ε

profile (the value of z where the bulk value of ε is reached), and 26 Å for the non-polar profile (the value of z where the non-polar function, $S(z)$, reaches 1). Three additional profiles were considered with dielectric/non-polar profile thicknesses of 24.5/24 Å, 25.5/24 Å, and 26.5/27 Å to model membranes that are thinner or thicker than DPPC bilayers. MM2 may be representative of phospholipid bilayers with shorter fatty acid tails than DPPC. MM1 is even thinner and may represent bilayers consisting of extremely short fatty acid tails like detergent molecules while MM4 may represent phospholipid bilayers with longer fatty acid tails as often present in physiological membranes [185]. The resulting ε and non-polar profiles are shown in Figure 2-4. Amino acid insertion profiles changed only moderately as a result of the shifted ε and non-polar profiles (see example in Figure 2-5).
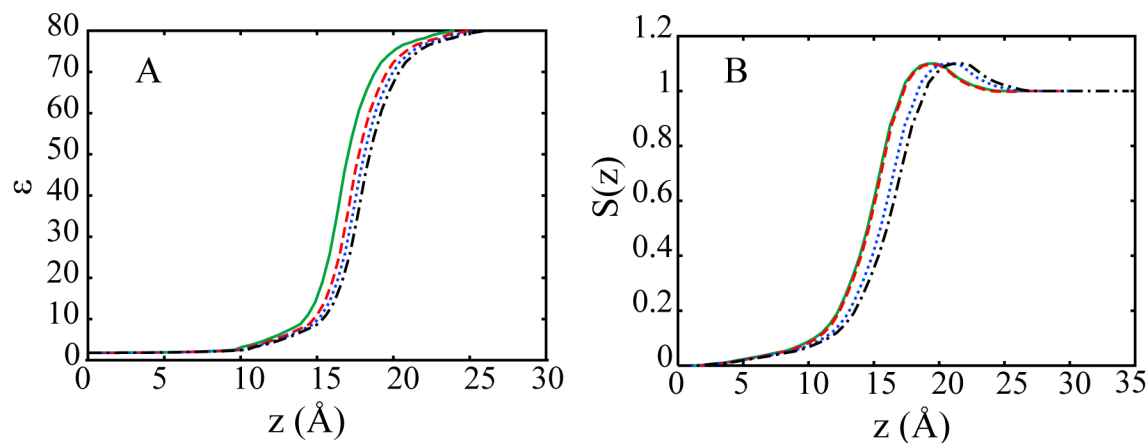
Figure 2-4 Dielectric profiles scaled to 24.5 Å (MM1,━━), 25.5 Å (MM2,▬ ▬ ▬), and 26.5 Å (MM4,▬·▬··▬) compared with the dielectric profile optimized for DPPC bilayer (MM3, thickness of 26 Å (⸳⸳⸳⸳⸳⸳⸳)). B) Non-polar profiles scaled to 24 Å (models 1 and 2,▬ ▬ ▬ and━━), and 27 Å (MM4,▬·▬··▬). The non-polar profile optimized for DPPC bilayer (MM3) has a thickness of 26 Å (⸳⸳⸳⸳⸳⸳⸳).
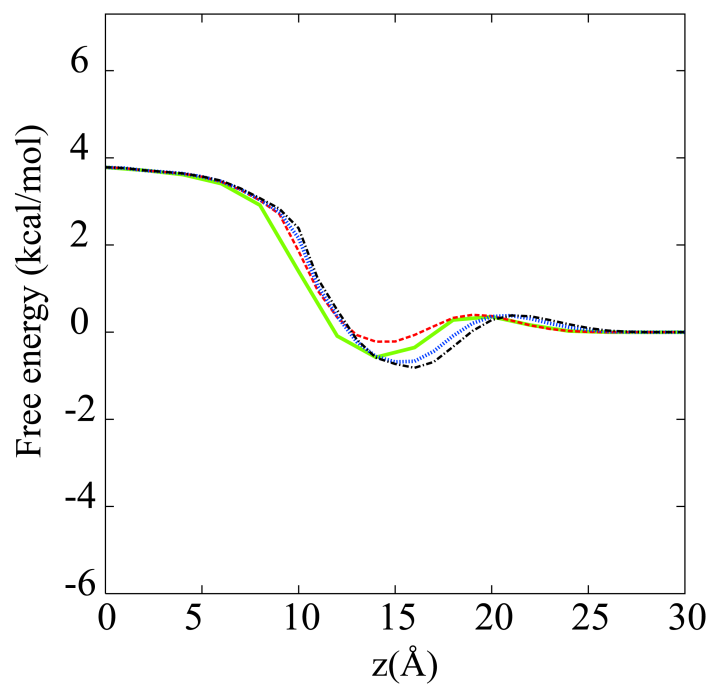
Figure 2-5 Free energy insertion profiles for asparagine amino acid analog (acetamide) in different membrane models: MM1 (——), MM2 (– – –), MM3 (॥॥॥॥) and MM4 (–·–·–).
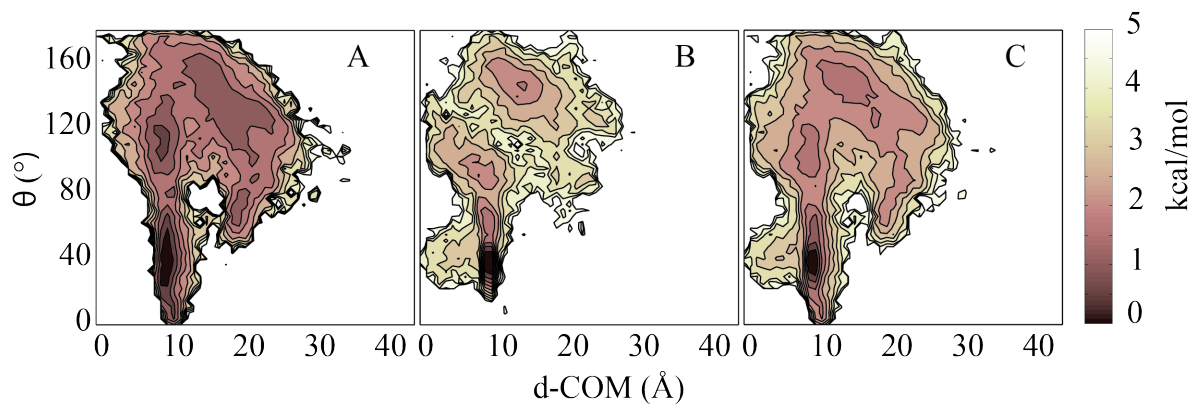
Figure 2-6 Energy landscapes for PLB sampled in membrane MM1. A) PMF generated from first, B) second, and C) combined replica exchange simulations for reaction coordinates according to Figure 2-1. Total production simulation times are 22.5 and 45 ns for the individual and combined simulations, respectively.

2.4.2 Replica exchange simulations of PLB with implicit membrane

Replica exchange simulations of PLB with different implicit membrane profiles were carried out for 30,000 cycles (22.5 ns/replica) each. The first 10,000 cycles were discarded as equilibration and the last 20,000 cycles (15 ns) of each simulation were used for analysis. In order to improve statistical convergence further, two separate simulations were run for each system rather than extending the time for a single simulation. Figure 2-6 shows a comparison of PMFs generated from two separate replica exchange simulations for MM1. The PMFs are qualitatively similar, but the quantitative differences between the two PMFs suggest an error of up to 1 kcal/mol in the less populated regions, which translates into a maximum error of 0.7 kcal/mol when combining two PMFs according to standard error analysis. Consequently, all of the subsequent PMFs are averaged from two independent replica exchange simulations.

Figure 2-2C shows the sampling of PLB as a function of θ, the interhelical angle, and d-COM, the TM axis-CP helix distance, for the DPPC membrane model (MM3). A broad range of conformations is sampled, including the structures from the NMR ensembles in DPC micelles (PDB id: 1N7L) [143] and recently refined models based on SSNMR data in DOPC bilayer (PDB id: 2KB7) [154] that have been mapped onto the PMFs based on their reaction coordinate values. The simulated PLB structures were clustered and the conformations closest to the cluster centers corresponding to minima in the PMF are shown in cartoon representation. The predominant minima correspond to L- and T-shaped conformations with T-shaped conformations being more prominent while the NMR ensembles emphasize L-shaped conformations more. The combination of solution NMR and lipid bilayer SSNMR in the 2KB7 ensemble has led to a shift

in populations towards larger d-COM values, still within the range of structures sampled in the simulations but clearly distinct from the main minima of the simulated structures.
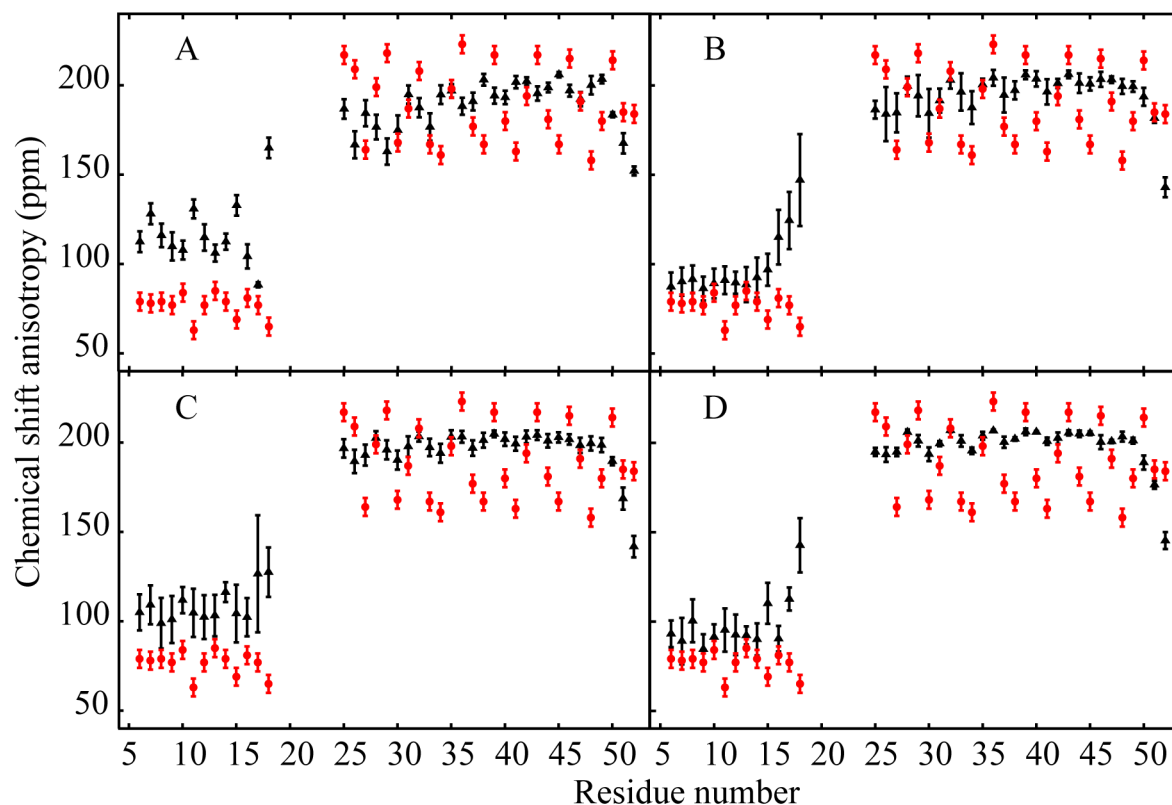
Figure 2-7 Average chemical shift Anisotropy (CSA) for different residues for A) MM1, B) MM2, C) MM3, and D) MM4 with error bars shown with ⊢⊖⊣. Experimental Values and their error bars are shown with ⊨◀⊣.

In order to compare further with the primary experimental data, average chemical shift anisotropy (CSA) and ^{1}H-^{15}N dipolar couplings (DC) were calculated from the simulations and compared against the experimental values (see Figures 2-7 and 2-8). Overall, the simulation results with the DPPC model are in good agreement with the SSNMR data. Average CSA values for the CP domain are slightly higher than experimental values as a result of sampling some

structures with large θ (see Figure 2-2). Deviations in the TM region also include a shift to slightly larger values (corresponding to slightly less tilted TM helices in the simulations). We also find that the averages from the simulations underestimate the amplitude of the periodic variations of both CSA and DC values. However, this is not an indication of significantly different structural ensembles since the amplitude is highly sensitive to the helical tilt angle and the average orientation of the N-H vector with respect to the helical axis [186].
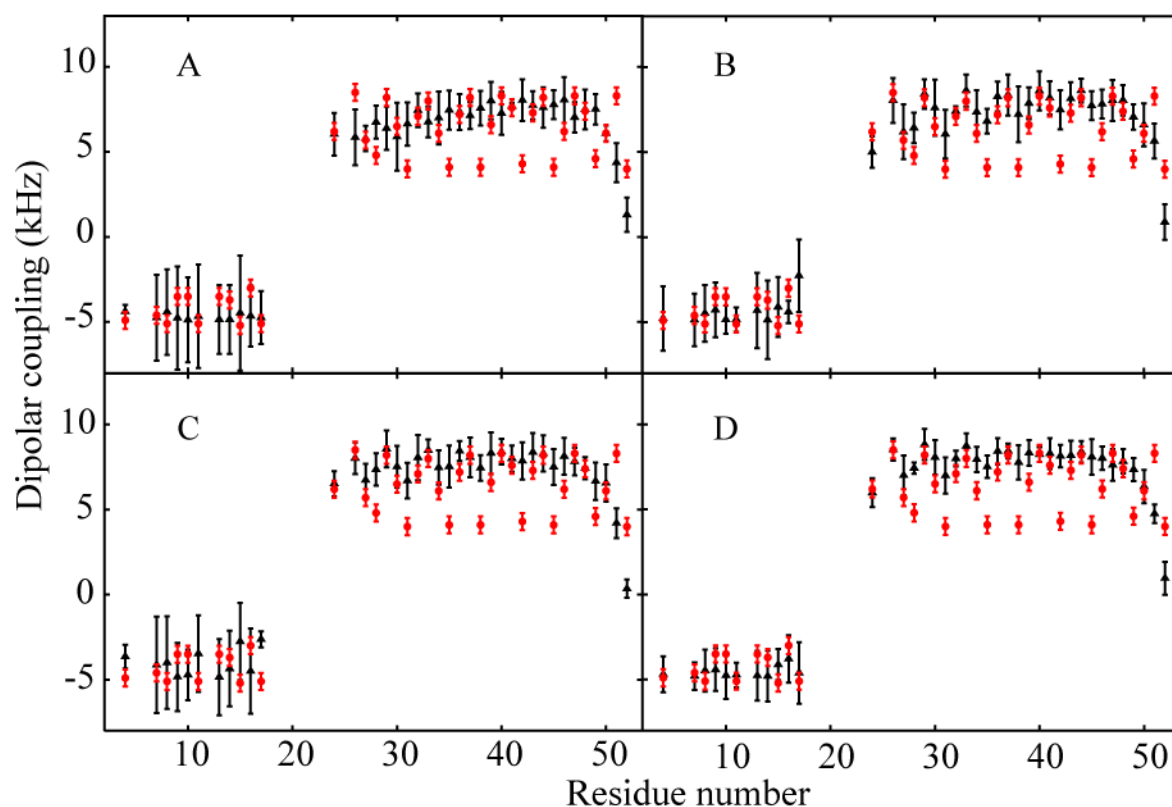
Figure 2-8 Average dipolar coupling (DC) for different residues for A) MM1, B) MM2, C) MM3, and D) MM4 with error bars shown with ⊢⊙⊣. Experimental Values and their error bars are shown with ⊢◀⊣.

In other conformations seen in the simulations the CP helix interacts less strongly with the membrane interface because of a more extended configuration or it inserts partially into the head-group region of the membrane in an almost parallel orientation relative to the TM helix. The latter conformation has not been described previously by experiments or any of the simulation studies. Based on our simulations, it is estimated that this compact conformation would be populated about 8% of the time and may therefore not be easily detectable in experimental studies or in short simulations. However, in the absence of clear experimental evidence, we cannot rule out that this conformation is an artifact of the simulation methodology employed here.

2.4.3   Effect of membrane thickness on PLB sampling

Implicit membrane models with stretched or compressed profiles were used to examine the effect of membrane thickness on PLB sampling. Figure 2-2 shows the PMFs generated with four different membrane models. It can be seen that the structures with small interhelical angles (θ) of 0-50° are increasingly populated at reduced membrane thickness model (MM1) and vanish in the model with the thickest membrane (MM4). At the same time, the pronounced broad minimum with θ=60-120° is present in all but MM1. Structures with large values of θ where the CP helix points away from the membrane surface are present in all profiles, but in the thickest membrane these structures are less populated and there are more partially bent structures with θ near 120° rather than fully extended structures with θ near 180°. The structures from the NMR ensembles overlap reasonably well with the simulated distributions for MM2-4 but less so for the thinnest membrane model, MM1, where L-shaped conformations are largely missing. The average DC values from the simulations agree well with the SSNMR restraints for all models,

49

but CSA values deviate significantly for the CP part in MM1 and for residues 15 to 18 in MM2 and MM4.

Average values of θ, tilt and rotation angles and insertion depths for the CP and TM helices, and the interdomain distance (between Y6 and C24) are reported in Table 1. The simulation results can be compared with experimental values for θ, the TM tilt and rotation angles, the CP tilt angle, and the interdomain distance [143-145, 154, 158] The other properties are compared against the reported values for the structural ensemble that resulted from optimization in a simple membrane potential [187] in the context of experimental restraints [154]. Overall, the agreement is quite good and there are only modest differences between the four membrane models despite significant differences in the conformational sampling suggesting that the averages are not very sensitive to the actual structural ensembles. We note that the TM tilt angle is reduced in the simulations as evident already from the DC and CSA data described above. Furthermore, the interdomain distance appears to be reduced in the simulated structures compared to the experimental data. However, there are uncertainties in the exact interpretation of the FRET experiments that were carried out to obtain these data [158]. The largest discrepancy between the simulated structures and the ensemble derived from experiments is in the CP insertion depth. In the simulations, the center of the CP helix is located at the membrane-water interface at about 30 Å from the membrane center while the structures from the recent NMR ensemble (2KB7) have a much more deeply inserted CP helix. However, while the exact insertion depth was not measured experimentally, the reported value of about 16 Å appears to be at least in part a result of the simple E_z membrane potential that was used during structure optimization. We found that optimization of representative structures extracted from our simulations with the E_z potential generally led to orientations with more deeply inserted CP

50

helices suggesting that deep insertion of CP helices into the lipid bilayer is more favorable with

the E_z potential than with our energy function.
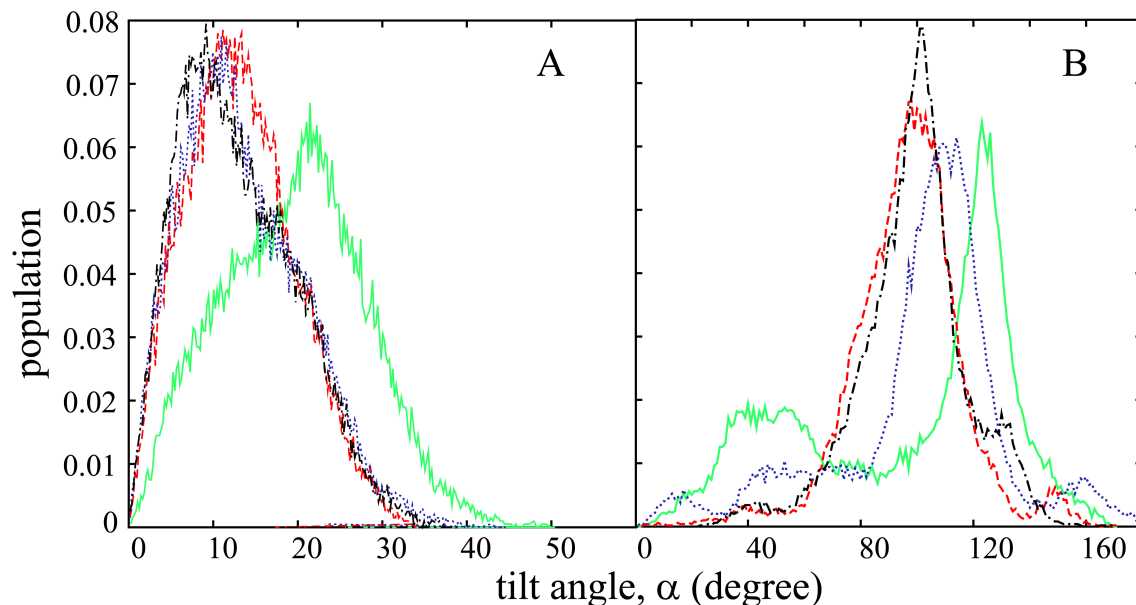
Figure 2-9 Distribution of the tilt angle (α) of A) TM helix defined by the angle between
the TM helical axis and membrane normal B) CP helix defined by the angle between the CP
helical axis and membrane normal for different membrane models: MM1 (——), MM2 (‒ ‒ ‒),
MM3 (.......) and MM4 (–..–..–).

An analysis of the distributions of the TM and CP helix tilt angles (see Figure 2-9), the

interdomain distances (see Figure 2-10), and the CP and TM rotation angles (see Figure 2-11)

indicates that the average values from the simulations are in fact a result of broad conformational

ensembles that consist of multiple states. For example, the average CP tilt angle does not vary

much between the different membrane models. However, the actual distribution is quite different

for MM1 compared to MM2 and MM4. In MM2 and MM4 there is a single peak near the

average value while MM1 has two separate peaks corresponding to the extended and compact

states with tilt angles of 120° and 40°, respectively. The distribution for MM3 appears to be

intermediate between MM1 and MM2/MM4. The distributions of the interdomain distances and helix rotation angles are more complex as a result of the broad ensemble of structures that are sampled in the simulations highlighting again that simple averages as measured experimentally may not fully capture the structural diversity of the actual structural ensembles of monomeric PLB. In the case of the TM helix rotation we find multiple states, a major state with a rotation angle of ~220° and minor states at ~120° and ~50°. In addition, there is another state with a rotation angle of ~270° that is predominantly populated with the MM1 model. The CP helix rotation angle appears to be very sensitive to the membrane model. With the MM3 model, the CP helix is predominantly rotated such that the hydrophobic residues A11, V4, and A15 are facing the membrane (ρ_{CP}<180) while those residues mostly face the solvent with the MM2 and MM4 models (ρ_{CP}>180).
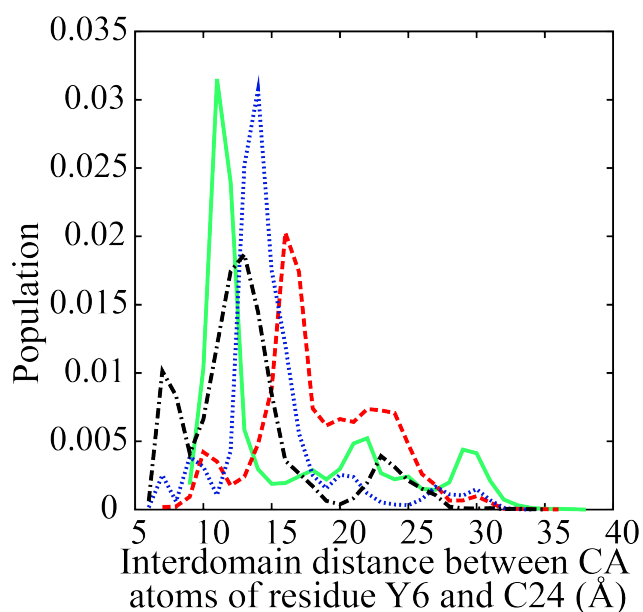
Figure 2-10 Interdomain distance distribution for different membrane models: MM1 (———), MM2 (- - - -), MM3 (.......) and MM4 (–·–·–). Interdomain distance is defined as the distance between CA atoms of the residues Y6 and C24.

52
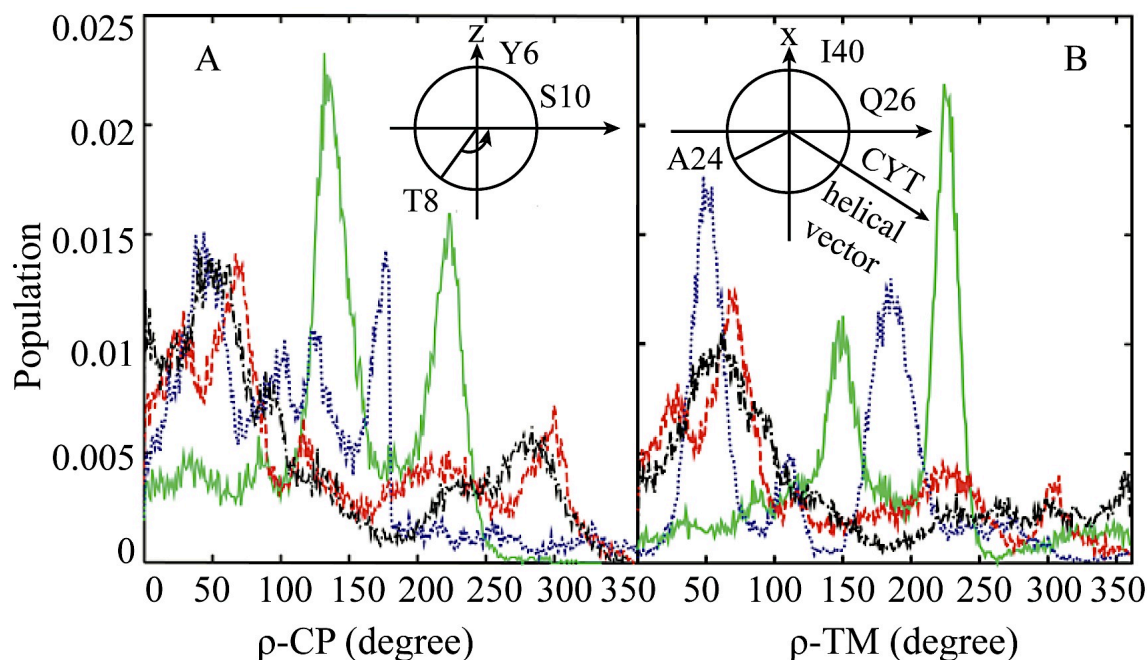
Figure 2-11 A) Distribution of the rotation angle (ρ) of the CP helix defined by the angle between the y axis and the vector connecting the amide nitrogen atom of T8 and the center of the CYT helical axis in yz plane in agreement with reference [154] B) Distribution of the rotation angle (ρ) of the TM helix defined by the angle between the CP helical axis and the vector connecting the amide nitrogen atom of A24 and the center of the TM helical axis in xy plane for different membrane models: MM1 (——), MM2 (▬ ▬ ▬), MM3 (·······) and MM4 (▬·▬·▬).

The interaction of the helices in the membrane was further analyzed by examining the distribution of the insertion depths of the TM and CP helices. The corresponding data for the four membrane models are shown in Figure 2-12. The insertion depths of the CP helix with MM2-4 follow similar Gaussian-like distributions with a peak near 30 Å. This means that the CP helix for the most part remains close to the membrane-water interface despite a significant fraction of structures with large interhelical angles. This is accomplished in part by tilting the

TM helix and by the formation of compact linker structures that allow the CP helix to stay close to the membrane as evident from the conformations shown in Figure 2-2 In contrast, the insertion depth distribution for the CP helix exhibits a much broader distribution with MM1 where one peak is located near 25 Å and an additional broad feature extends from 35-48Å. The first peak corresponds to compact structures with small values of θ, while the second feature reflects extended structures with large values of θ. Thus, only with the thinnest membrane we observe extended conformations where the CP helix is located far above the membrane.
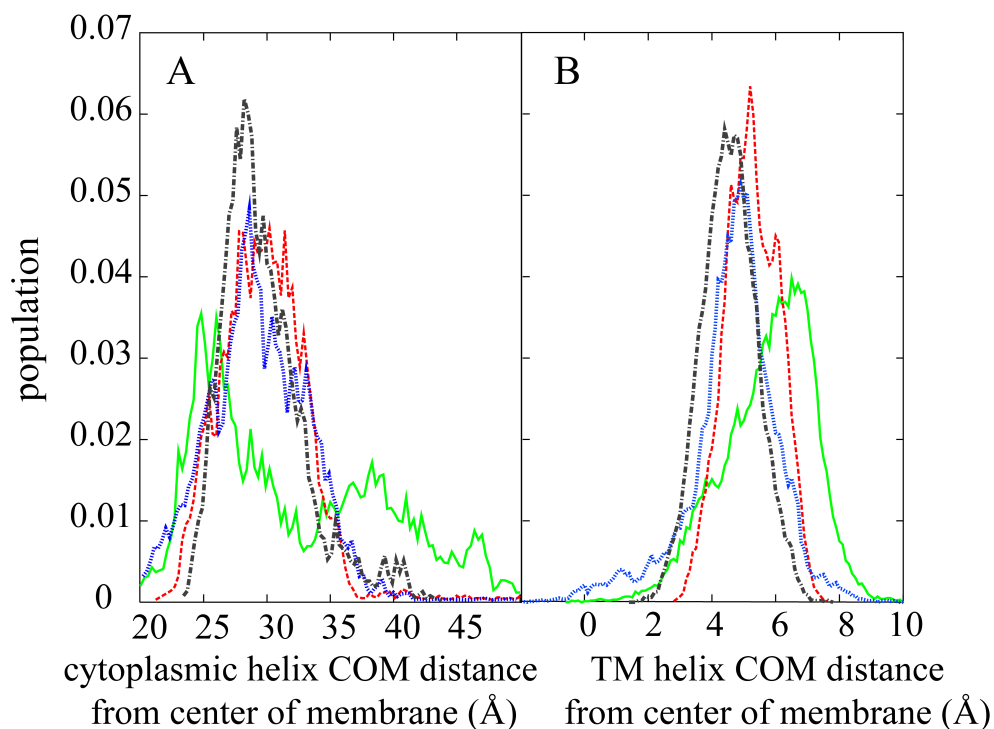
Figure 2-12 A) Distribution of CP helix insertion. B) Distribution of the TM helix (residues 25-52) insertion for different membrane models: MM1 (──), MM2 (----), MM3 (……) and MM4 (─··─··─).

The TM helix is generally shifted towards the CP side of the membrane (see Figure 2-12B). Again, MM2-4 show similar distributions of the distances from the center of the membrane, while MM1 stands out. The center of mass of the TM helix is on average 5 Å away from the membrane center with models 2-4 but in the thinnest membrane its COM is displaced further to about 7 Å from the membrane center. The larger displacement of the TM helix with the thinnest membrane is possible because of the hydrophobic mismatch and results in the hinge region being pushed further above the membrane-water interface. With the hinge region away from the membrane, the CP helix can form compact structures with small values of θ where the polar N-terminus points towards the membrane but is still not significantly inserted into the hydrophobic core of the membrane. At the same time, the sampling of more extended structures with the CP helix far from the membrane interface as described above is facilitated by the raised hinge region, and L- and T-shaped structures are not especially favorable with the CP helix far from the membrane-water interface in those conformations.
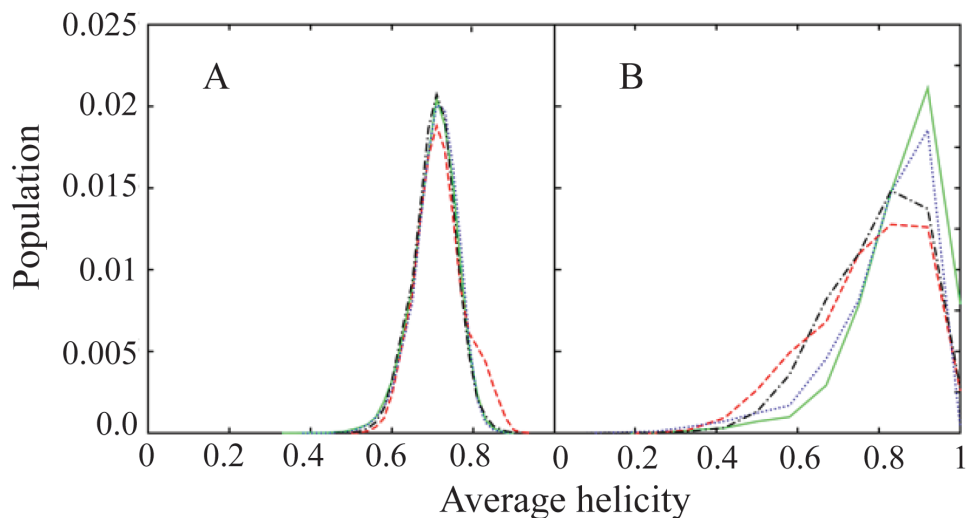
Figure 2-13 Distribution of the average helicity calculated for A) the full length PLB and B) domain Ia (residue 1-16) for different membrane models: MM1 (——), MM2 (- - - -), MM3 (·······) and MM4 (–·–·–).

Finally, distributions of the average helicity for the CP domain and full length PLB are given in Figure 2-13. The CP domain is largely helical in our simulations with a small subset of partially helical conformations. These results are in agreement with most experimental studies except for a study by the Baldus group who interpreted dynamics in the CP domain as sampling of non-helical conformations [188].

2.5   Discussion

The conformational analysis of membrane-bound peptides has been challenging with both experimental and computational methods. Experimental methods often do not provide full atomistic resolution and/or have to compromise on the systems that can be studied (e.g. mutated PLB in DPC micelles vs. wild-type PLB in phospholipid bilayers). At the same time computational methods are limited by compromises between force field accuracy and the extent of sampling that can be achieved. Here, the new implicit membrane model, HDGB, is employed to overcome sampling limitations in explicit lipid simulations. The implicit model offers significantly improved conformational sampling, in particular because the kinetic barriers due to lipid rearrangements in response to conformational changes of a given solute are absent. Because of the nature of implicit solvent, the increased computational efficiency comes at the cost of reduced accuracy. In particular, specific lipid-peptide interactions cannot be fully represented with an implicit approach. However, the implicit model has been optimized carefully to reflect the mean-field energetics of explicit all-atom lipid bilayer environments while previous simulations of peptides and proteins with implicit membranes [106, 121-125] have resulted in good agreement with experimental data.

56

It is estimated that μs-ms time scales can be accessed with our simulations, which is much longer than the time scales covered in previous simulation studies. The use of implicit solvent with Langevin dynamics and a small friction coefficient by itself has been shown to accelerate the crossing of barriers in aqueous solvent by a factor 2-10 [175] while temperature replica exchange sampling is expected to accelerate barrier crossings further by at least a factor of 10. This means that our replica exchange implicit solvent simulations with ~20 ns/replica should be equivalent to at least μs time scale sampling. Furthermore, the mean-field energetics provided by the implicit membrane model avoids kinetic barriers due to lipid relaxation processes that may occur on time scales ranging from ns to seconds [189]. Therefore, microsecond time scale implicit membrane simulations are probably equivalent to millisecond or longer explicit membrane simulations. Indeed, we found significant conformational sampling of a wide range of structures that far exceeds the conformational variety observed in previous simulation studies and agrees with the experimental finding of significant conformational dynamics in PLB [166].

The use of an implicit model also offers the unique advantage of being able to easily change the membrane thickness. Here, four different thicknesses were considered in order to examine the effect of membrane thickness on PLB sampling. It was found that sampling of PLB differs significantly between the thinnest and thickest membranes. With the thickest membranes PLB predominantly samples L- and T-shaped conformations as well as semi-extended structures, but in all cases the CP helix closely interacts with the membrane surface. With the thinnest membrane, most of the sampling involves compact structures with small interhelical angles where the CP and TM helices are nearly parallel. In addition, fully extended structures are sampled where the CP helix is located far above the membrane interface.

The intermediate model MM3 was parameterized to reflect DPPC bilayer. From a comparison of MD simulations of DPC micelles [190] and DPPC bilayers [191], the overall membrane thickness is slightly reduced with DPC over DPPC as in the MM2 model. The thinnest model, MM1, may then be representative of detergents such as the $C_{12}E_8$ detergent with even shorter hydrocarbon tails. On the other hand, the physiological cardiac sarcoplasmic reticulum membrane (CSR) with an average fatty acid chain length of 18.1 carbon atoms [192] compared to DPPC with 16 carbon atom tails corresponds best to the thickest membrane model (MM4). It should be emphasized, however, that only MM3 was optimized to model a specific membrane, DPPC, and that the other models are simply introduced to study the general effect of reduced or increased membrane thickness on PLB sampling. In order to model other specific types of lipid environments, further optimizations would be required. In this study, we are also neglecting the possible impact of membrane curvature that may be important to accurately model micellar systems.

The conformational sampling with MM2 (Figure 2-2B) is indeed in good agreement with the structural NMR ensemble in DPC micelles but the results for MM2 and MM3 are similar, suggesting that PLB sampling in DPC micelles and DPPC bilayers may not be significantly different and that DPC micelles may be good mimics for DOPC/DPPC bilayers in experimental or computational studies. In contrast, there are more significant differences between MM2/MM3 and MM4 suggesting that experimental results obtained in DPC detergents or DOPC/DPPC bilayers may not be fully reflective of the behavior of PLB in native SR membranes.

The original NMR ensemble in DPC micelles was recently re-refined by applying an geometric restraints from new SSNMR measurements of PLB in DOPC (which is similar to DPPC) [154] along with an empirical membrane potential [187]. The resulting ensemble still

consists of mostly L-shaped structures but the CP domain is generally further extended and more inserted into the lipid bilayer compared to the originally reported ensemble. It may seem that the shift in the structural ensemble as a result of the SSNMR data in DOPC bilayers suggests genuine differences in the conformational sampling between DPC micelles and DOPC bilayers. However, neither set of data (NOE restraints from solution NMR in DPC micelles and SSNMR data in DOPC bilayers) are sufficient to unambiguously describe the conformation of PLB and even when both data sets are combined significant uncertainties remain, especially with respect to the insertion of the CP domain into the bilayer. Furthermore, it should be kept in mind, that structures consistent with the time- and ensemble-averaged NMR data are not necessarily expected to agree with the actual conformational distribution on sub-ms time scales.

	θ	d-COM	α-TM	α-CP	Helicity	z-TM	z-CP	ρ-TM	ρ-CP
Basin 1	67±2	9.1±0.8	10±1	111±2	0.7±0.1	6.8±0.1	29.6±0.9	127±4	99±5
Basin 2	79±2	13.9±0.7	17±1	109±2	0.7±0.1	6.0±0.1	29.3±0.9	59±5	103±6

Table 2-2 Average properties for conformations in basin 1 and 2 in Fig. 3*C*; θ is the interhelical angle, d-COM is the distance between the center of mass of the CP helix from the projection onto the TM helical axis, α-TM and α-CP are the tilt angles of the TM and CP helices respect to membrane normal respectively, helicity is the averaged helicity of the full length PLB, *z*-TM and *z*-CP are the distance of the center of the membrane from the COM of the TM and CP helices respectively, ρ-TM is the TM rotation angle defined by the angle between the CP helical axis and the vector connecting the amide nitrogen of A24 and the center of the TM helical axis in

Table 2-2 (Con't)

xy plane, ρ-CP is the CP rotation angle defined by the angle between y axis and the vector connecting the amide nitrogen of T8 and the center of the CP helical axis in yz plane. ρ-CP is defined in agreement with reference [154].

An interesting observation is that fully extended conformations with the CP helix located far above the membrane surface are only sampled with the thinnest membrane. PLB has been previously suggested to interact with SERCA in an extended form. This idea is based primarily on the observation of a cross-link between K3 of PLB and K400 of SERCA in $C_{12}E_8$ detergent [193]. Such an interaction is only possible when PLB is fully extended. However, this cross-link could not be reproduced by other groups, especially the Jones group, who used insect cell microsomes instead of detergent [194]. Our simulations offer an explanation for this apparent controversy. $C_{12}E_8$ detergent has a very short lipid tail consisting of only 12 carbon atoms and is best modeled with a thin membrane similar to MM1. At the same time, microsomal membranes are thicker than DPPC bilayers according to electron microscopy measurements [185, 191] and are best modeled by MM4. Our simulations show that PLB conformations with a fully extended CP helix are not observed in thicker membranes but are a feature of very thin membranes. Therefore, it appears that the observation of the PLB-K400 cross-link [193] may be a result of the particular detergent that was used. Because native SR membranes are also thicker, this finding has implications for the role of extended conformations of PLB in interactions with SERCA. Based on our result (see Figure 2-2D) PLB assumes primarily T- and L-shaped conformations in the thickest membrane with two main basins, one consisting of more compact T-shaped structures, the other one consisting of more extended, but still mostly L-shaped

60

conformations (See Table 2-2). Therefore, one may conclude that in SR membranes PLB interacts with SERCA primarily in one of those conformations instead of a fully extended form as suggested earlier. At the same time, it has been known for a long time that membrane thickness affects SERCA activity. In particular, long chain, unsaturated fatty acids maximize calcium transport across SR membranes [195, 196]. While this may be a direct consequence of lipid-protein interactions, it is also possible, based on our findings, that SERCA activity is diminished in thinner membranes because of increased inhibition by PLB under the assumption that PLB does in fact inhibit SERCA more readily in the extended form. In order to address these possibilities in more detail, future studies will need to focus on the actual interactions between PLB and SERCA. Furthermore, the effect of phosphorylation on PLB structure in the context of different membranes and in interactions with SERCA has not been considered here and will also need to be addressed in more detail in future studies to arrive at a complete understanding of SERCA regulation by PLB.

2.6   Conclusion

The conformational sampling of PLB on ns-μs time scale was investigated with an implicit membrane model. Significant conformational dynamics were observed that are in overall good agreement with experimental measurements. Furthermore, different membrane thicknesses were modeled by appropriate scaling of the implicit membrane model. Membrane model mimicking DPPC bilayers results in conformational ensembles with a high population of T- and L-shaped structures as well as more extended structures with larger interhelical angles. In addition, a small, but significant population of compact, highly bent structures was found that have not been described previously. The sampling of compact and fully extended structures is

enhanced with the thinnest membrane model while those conformations are suppressed with the thickest membrane model that most closely resembles the actual physiological SR membranes. These findings suggest that partially extended conformations rather than the fully extended ones may play a significant role in PLB-SERCA interactions under physiological conditions.

2.7   Acknowledgments

3 Role of Conformational Sampling of Ser16 and Thr17-phosphorylated Phospholamban in Interactions with SERCA[198]

3.1 Abstract

Phosphorylation of phospholamban (PLB) at Ser16 and/ or Thr17 is believed to release its inhibitory effect on sarcoplasmic reticulum calcium ATPase. Ser16 phosphorylation of PLB has been suggested to cause a conformational change that alters the interaction between the enzyme and protein. Using computer simulations, the conformational sampling of Ser16 phosphorylated PLB in implicit membrane environment is compared here with the unphosphorylated PLB system to investigate these conformational changes. The results suggest that conformational changes in the cytoplasmic domain of PLB upon phosphorylation at Ser16 increase the likelihood of unfavorable interactions with SERCA in the E2 state prompting a conformational switch of SERCA from E2 to E1. Phosphorylation of PLB at Thr17 on the other hand does not appear to affect interactions with SERCA significantly suggesting that the mechanism of releasing the inhibitory effect is different between Thr17 phosphorylated and Ser16 phosphorylated PLB.

3.2 Introduction

Phospholamban (PLB), a 52 residue transmembrane (TM) protein, plays an essential role in regulating sarco-/endoplasmic reticulum calcium ATPase (SERCA), a calcium pump in heart muscle [129] by reducing the enzyme affinity for calcium [157]. PLB can be phosphorylated at Ser16 and/or Thr17 by cAMP-dependent protein kinase A (PKA) and Ca^{2+}-calmodulin

63

dependent protein kinase (CAM kinase), respectively, in response to β-adrenergic stimulation [199]. PLB phosphorylation increases the Ca^{2+} affinity of SERCA and releases the inhibitory effect [200]. Several *in vivo* studies have shown that phosphorylation of Ser16 and Thr17 is sequential [137, 201]. *In vitro* studies in sarcoplasmic reticulum (SR) membranes on the other hand have shown that Thr17 phosphorylation can be stimulated by electrical pulse without the prerequisite of Ser16 phosphorylation [202-204]. It has been suggested that the sequential phosphorylation *in vivo* is a result of interactions between PKA and CaMKII pathways [205]. Ser16 phosphorylation of PLB appears to be sufficient to release the inhibitory effect for the maximum cardiac response to β-adrenergic stimulation [135]. Therefore, the physiological role of Thr17 phosphorylation is not fully understood yet. Interestingly, aerobic interval training preferentially increases Thr17 phosphorylation [206]. It has been suggested that Thr17 and Ser16 phosphorylation may have an additive effect in releasing the inhibitory function of PLB both *in vitro* [132, 207] and *in vivo* [139]. The possibly different physiological role of Ser16 vs. Thr17 phosphorylation may also point at mechanistic differences in SERCA-PLB interactions with PLB phosphorylation at either site.

To better understand the mechanism by which Ser16 and/or Thr17 phosphorylation relieves inhibition of SERCA by PLB, detailed knowledge of the structural changes of PLB upon phosphorylation is essential. Most experimental and computational studies have been focused on unphosphorylated PLB and relatively little detailed structural information is available for phosphorylated PLB. NMR studies of unphosphorylated PLB in DPC micelles [143] and lipid membranes [144, 154] suggest an ensemble of L-shaped structures where a long TM helix is connected to a dynamic cytoplasmic (CP) helix via a flexible linker. Based on NMR data, average interhelical angle of 80 ± 20° has been reported in different solutions [208], and micellar

and membrane environments [143, 144], while other experimental data indicate that the CP domain of PLB is in fact in equilibrium between dynamically ordered and disordered conformations [165, 166]. The average structural properties and conformational dynamics of PLB have been reproduced by molecular dynamics (MD) simulations from our group [156] and others [168-172].

Phosphorylation of PLB appears to shift the equilibrium of the cytoplasmic helix towards a "disordered" state [209, 210], although the exact molecular interpretation of this finding is unclear. Furthermore, a reduction of helical content in the CP helix upon Ser16 phosphorylation has been suggested from CD [209, 210], NMR [211] and attenuated total reflection FTIR (ATR-FTIR) experiments [212]. NMR studies on a 36-residue long N-terminal fragment of Ser16 and Thr17 phosphorylated PLB also report a shorter CP helix from residues 2 to 12 [213] while the interhelical angle in double phosphorylated PLB appears to be increased to $100 \pm 35°$ [213]. Results from fluorescence resonance energy transfer (FRET) studies indicate a reduced interdomain distance between residues Tyr6 and Cys24 after Ser16 phosphorylation [158]. This is consistent with fluorescence quenching studies of Tyr6 that suggest that the phosphorylation reduces the solvent accessibility of Tyr6 is a result of a conformational change [214].

Several experimental studies are consistent with the idea that PLB is still attached to SERCA after phosphorylation [158, 210, 215-217]. On the other hand, an earlier cross-link study reported the abolishment of cross-links between SERCA and PLB upon phosphorylation by PKA [146], which was interpreted as complete dissociation of PLB from SERCA. However, a more recent cross-link study is more consistent with data supporting the hypothesis that PLB does not dissociate from SERCA upon phosphorylation [218]. In that study, phosphorylation of PLB on Ser16 increased the distances between certain residues of PLB and SERCA while other cross-

links remained the same [218]. An early EPR study also suggests the complete dissociation of PLB from SERCA [218]. There, the shift to the pentameric state of PLB upon phosphorylation was observed in the pentamer-monomer PLB equilibrium. This shift was interpreted as an evidence for dissociation of PLB from SERCA, which binds to the monomeric state of PLB [219]. Finally, fluorescence measurements of PLB SERCA co-crystals have suggested an alternative hypothesis that phosphorylation of PLB may relieve the inhibitory effect by facilitating the structural coupling between two interacting SERCA units that is otherwise disrupted by unphosphorylated PLB [220].

A few MD simulations of phosphorylated PLB (pPLB) have been reported to date and resulted in the following findings: A decrease in helicity was observed upon Ser16 phosphorylation in simulations of the cytoplasmic PLB domain (residues 1-25) [169]. MD simulations of full length pPLB in an explicit POPC membrane bilayer also showed a decrease in the helical content while interactions of phosphoserine with the membrane head group region appeared to result in lower mobility of the CP domain [170]. Another MD simulation study using both replica exchange simulations of the CP domain of PLB and a constant temperature MD simulation of full-length pPLB has provided further insights. The results from this study suggest a decrease in the interdomain distance, a decreased interhelical angle, and again a loss of helical contents in the CP helix upon phosphorylation as a result of interactions between the Ser16 phosphate group and residues Arg9, Arg13 and/or Arg14 [168]. These simulations of pPLB provide important insight into the structure and dynamics of PLB upon phosphorylation but the covered time scales remain relatively short compared to the experimentally observed dynamics, reaching at best tens of nanoseconds in the replica exchange runs.

In order to describe the dynamics of phosphorylated PLB over much longer time scales, we are reporting here results from replica exchange implicit membrane simulations of PLB that is phosphorylated either at Ser16 or at Thr17. We used here the HDGB implicit membrane model that has been successfully applied in previous simulations of unphosphorylated PLB [156], influenza fusion peptide [221], and integral membrane proteins [222]. An implicit description of membrane and solvent drastically accelerates sampling because lipid relaxation processes are avoided. We estimate that the combination of using the implicit membrane model with replica exchange sampling provides access to dynamics that would otherwise be observed on μs to ms time scales. The drawback of using an implicit approach is of course the neglect of explicit lipid-peptide interactions, however, previous studies from our group suggest that the HDGB model used here is able to capture the structure and dynamics of membrane-bound peptides, including PLB, in a highly realistic fashion [156, 221].

In the following, the methods used here are briefly summarized before the results are presented and discussed.

3.3   Methods

3.3.1 Simulations

Replica exchange MD simulations [30] of monomeric PLB phosphorylated at Ser16 (p16-PLB) or Thr17 (p17-PLB) were carried out and compared with a previous simulation of unphosphorylated PLB [156]. Model 1 of the NMR ensemble for the C36A/C41F/C46A PLB mutant (PDB entry 1N7L [143]) was used as the initial structure. Zwitterionic termini and standard protonation states (pH=7) were applied. The phosphoserine and phosphothreonine side

chains were assumed to be fully deprotonated with a -2 charge. Starting conformations were oriented in the implicit membrane with the transmembrane (TM) helix parallel to the membrane normal and the cytoplasmic (CP) helix above and parallel to the membrane surface. The CHARMM22 all-atom force-field [32] along with the CMAP correction term [38] was used to describe peptide interactions. No cutoffs were applied to the non-bonded interactions. The implicit description of the membrane involved the heterogeneous dielectric generalized Born model (HDGB) [179]. Optimized parameters for DPPC bilayer [156] were used in all simulations. Other implicit solvent parameters were set as described previously [223]. The SHAKE algorithm [176] was applied to constrain bonds involving hydrogen atoms. To control the temperature, Langevin dynamics [174] was used with a friction coefficient of 10 ps^{-1} applied to heavy atoms [175]. A time step of 1.5 fs was used to maintain stable simulations with the implicit membrane model as described previously [173]. All simulations were carried out with version 34a2 of CHARMM [177] in combination with the MMTSB Tool Set [178].

Replica exchange simulations consisted of eight replicas over a temperature range of 300-400K, spaced exponentially. All of the replicas were started with energy-minimized structures. Exchanges were attempted after every 500 MD steps (0.75 ps). The resulting exchange acceptance ratio was 21-30% between adjacent replicas. Four separate replica exchange simulations were carried out for each system and sampling data was combined to improve statistical significance [156]. Each simulation was carried out for 50,000 cycles (37.5 ns/replica) from which the first 30,000 cycles were considered as equilibration and not included in the analysis. The total length of the production runs is 67.5 ns (22.5 ns for each simulation) for the non-phosphorylated PLB and 90 ns for the p16 and p17 phosphorylated PLB systems (22.5 ns for each simulation).

The weighted histogram analysis method (WHAM) [181] was used to generate potentials of mean force (PMFs) at 300K to include sampling data at higher temperatures as well. The interhelical angle and the distance between the center of mass of the CP helix and its projection onto the TM helix (d-COM) were chosen as the primary reaction coordinates for generating the PMFs (see Figure 3-1). The interhelical angle (θ) is defined as the angle between the CP (residue 4-12) and TM (residue 23-52) helical axis. The center of mass of the CP helix was calculated based on residues 1 to 16. Note, that the definitions of the reaction coordinates here are slightly different from our previous definition [156] to better describe conformations with less well-defined helical structure. Further analysis involved clustering with the K-means method as implemented in the MMTSB Tool Set [178]. PyMOL was used to generate molecular graphics [182]. MATLAB version 7.9.0.529 (R2009b) was used to generate PMFs and for data mining analysis.
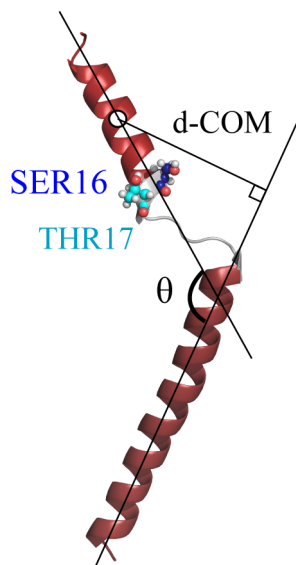
Figure 3-1 Reaction coordinates used in generating the PMF plots; interhelical angle θ, and the distance between the center of mass of the CP helix from its projection onto the TM

Figure 3-1 (cont'd)

helical axis, d-COM. Shorter CP helical domain that is used to define θ is shown in red as well as the TM domain.

3.3.2 SERCA-PLB docking

In order to test whether different PLB conformations are likely to interact favorably with SERCA the following procedure was applied: The TM domain of PLB composed of residues 21-52 (called T-PLB) from the NMR structure (PDB: 1N7L) was manually docked to the structure of SERCA in the E2 conformation (PDB: 1IWO [224]). To obtain the docked complex, cross-linking information according to Table 3-1 was used [194, 225-227]. The final distances in the resulting SERCA-PLB complex were close to the cross-linking distances (see Table 3-1). TM domain of all different conformations of PLB and phosphorylated PLB were then aligned the docked T-PLB. From all of the aligned PLB structures, residues 1-22 (called cpPLB) were selected and added to the PDB structure of SERCA in the E2 conformation. The complexes were minimized for only 100 steps to correct any possible clashes between the side chains of the two proteins. In order to obtain a docking score, minimized structures of the complexes were then reduced to a coarse-grained (CG) model with a single site per residue. The CG model was used to focus on the overall shape compatibility and avoid excessive sensitivity to individual atomic contacts that could be relieved through side chain reorientation. The TM domain of PLB in all of the structures is well preserved in the simulations. Therefore, only the CP domain of PLB is included in the complexes and calculation of the docking score. Each CG particle is placed at the average position of all atoms in a given residue. Only Lennard-Jones interactions were considered in the docking score with the parameters reported in Table 3-2. Energies were then

calculated for the CG models of the SERCA-cpPLB complexes to determine whether CP domain

of PLB structures clash with SERCA or not. An energy cutoff of 20 kcal/mol was found to

roughly match a more subjective analysis based on visual inspection. Following this protocol, the

structural ensembles from the simulations were evaluated to determine the fraction of

conformations that are likely to interact favorably with SERCA.

	V89C-V49C	K328-N27C	T317C-L31C	C318-N30C	K328-N30C
Experiment (no PKA) (Å)	<5[227]	<10[194]	<10[226]	<10[225]	<15[194]
Experimental changes in the cross-link after adding PKA	none[218]	none[218]	-80%[218]	-50%[218]	none[218]
PLB:SERCA-E2	4.9	10.5	8.1	10.2	13.0
P16-PLB :SERCA-E2	3.8	10.2	11.9	11.7	10.7

Table 3-1 Cross-linking distances and corresponding distances after manual docking of

PLB with SERCA-E2 (in Å). All distances are measured between side chain S atom of (mutated)

Cysteines and side chain N atom of Lysine.

Residue	ε (kcal/mol)	R/2 (Å)
Ala	-0.07	2.732104
Val	-0.07	2.587710
Ile	-0.07	2.587306
Leu	-0.07	2.735415
Pro	-0.07	2.923360
Met	-0.07	2.764147
Phe	-0.07	3.329596
Trp	-0.07	3.132287
Gly	-0.07	2.603561
Ser	-0.07	2.738570
Thr	-0.07	2.590725
Cys	-0.07	2.734203
Asn	-0.07	2.740384
Gln	-0.07	3.039055
Tyr	-0.07	3.320570
Lys	-0.07	3.114759
Arg	-0.07	3.384245
His	-0.07	2.984095
Asp	-0.07	2.739218
Glu	-0.07	2.590725

Table 3-2 Lennard-Jones parameters used during docking with the standard Lennard-Jones functional form of $V=[\varepsilon_{i,j} (R_{i,j}/r_{i,j})^{12} - 2(R_{i,j}/r_{i,j})^{6}]$ where $\varepsilon_{i,j}=(\varepsilon_i . \varepsilon_j)^{1/2}$ and $R_{i,j}=R_i/2 + R_j/2$.

3.4 Results

Replica exchange simulations of phosphorylated PLB are compared here with previous simulations of unphosphorylated PLB [156]. In order to characterize the dynamics of the cytoplasmic (CP) helix, the simulation data was first analyzed in terms of the angle between the TM and CP helices, θ, and the distance of the center of mass of the CP helix from its projection

on the TM helical axis, d-COM (see method section for detailed definitions). The resulting PMFs and representative structures corresponding to the minima are shown in Figures 3-2A-C and Figure 3-3, respectively. A wide range of structures was sampled with either of the two phosphorylation states. Generally, the ensembles include compact structures, where the CP helix interacts with the membrane interface, and extended structures where the CP helix points away from the membrane. Compact structures include L- and T-shaped states with θ values between 50-90$^{\text{o}}$ as well as conformations with small values of θ of less than 30$^{\text{o}}$, where the CP helix is packed in an antiparallel fashion against the TM helix.

In unphosphorylated PLB, most of the sampling involves L- and T-shaped structures with a small additional population of extended states (see Figures 3-2A and 3-3). Upon phosphorylation at Ser16, additional local minima regions appear in the free energy landscapes. The most dominant structures are still L- and T-shaped states (cf. B3, B4, and B5 in Figure 3-3) with the L-shaped structures more populated relative to the unphosphorylated state. Looking at Figure 3-3 it can be seen that the structures in the most populated cluster B5 as well as B6 have a break in the CP domain. These structures are not observed in unphosphorylated PLB and are apparently stabilized by extra hydrogen bonds that involve the phosphorylated Ser16. On the other hand, structures with small θ values below 30$^{\text{o}}$ are less populated and structures with $\theta<10^{\text{o}}$ are absent altogether in the Ser16 phosphorylated PLB ensemble. This can also be explained by extra hydrogen bonds formed in the CP domain. These extra hydrogen bonds cause the CP helix to break and the N-terminal part comes upward and points away from the membrane center instead if inserting deeply into the membrane. Reduced helicity in the CP helix upon Ser16 phosphorylation is also apparent in the other structures shown in Figure 3-3 (cf. B3 and B7). Thr17 phosphorylation, on the other hand, leads to an ensemble that is more similar to

unphosphorylated PLB. However, there is more significant sampling of extended structures
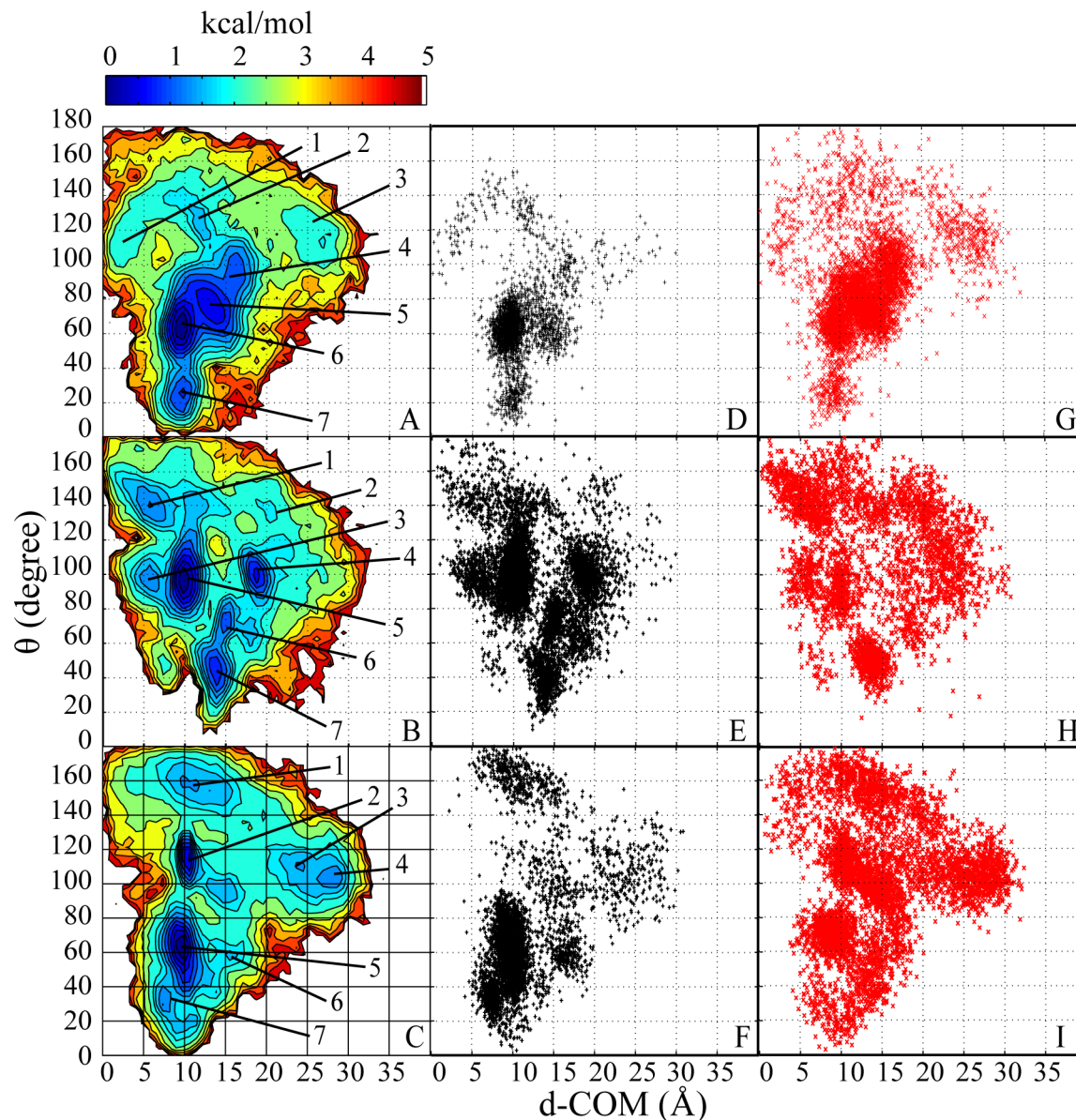
Figure 3-2 PMFs as a function of θ and d-COM (see Figure 3-1), for (A) PLB, (B) p16-PLB, (C) p17-PLB. Representative structures from cluster analysis that correspond to minima are shown. Mapping structures with (shown with ✖) and without clash (shown with ✚) in the θ and d-COM coordination space for (D, G) PLB, (E, H) p16-PLB, and (F, I) p17-PLB. Conformations sterically incompatible with complex formation are shown in the middle column in black, compatible conformations are shown in red in the right column.
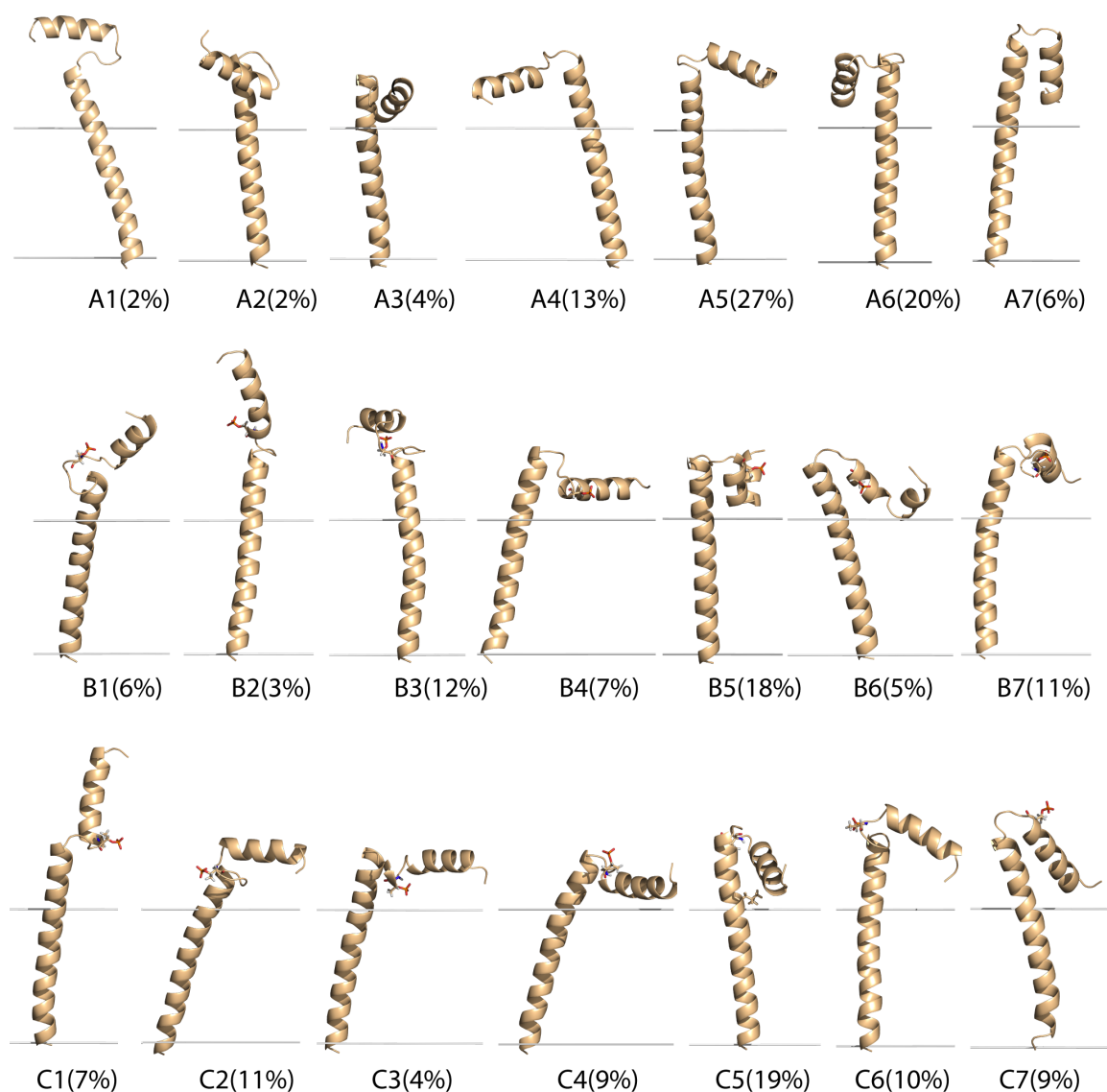
Figure 3-3 Representative structures from cluster analysis for (A1-7) PLB, (B1-7) p16-PLB, and (C1-7) p17-PLB that correspond to minima on Figure 3-2A-C. In all structures, the orientation with respect to the membrane was preserved. Lines represent the membrane hydrophobic core and are placed at 15Å distance above and below the membrane center (z=0). Ser16 in B1-7 and Thr17 in C1-7 are shown with stick representation. Cluster population percentages are given in parentheses.

($\theta > 150^\circ$) as well as structures with small values of θ ($\theta < 30^\circ$). T-shaped conformations with small d-COM values (like B4) appear to be less prominent upon Thr17 phosphorylation (see Figure 3-2), while highly extended conformations and completely L-shaped conformations such as C1 and C4 in Figure 3-3 are sampled more extensively.

To understand the effect of phosphorylation further, a number of other structural properties were analyzed: First, the average helicity was calculated as a function of residue number. The results, presented in Figure 3-4, show that phosphorylation at Ser16 significantly reduces the helicity in the CP domain of PLB. This is in agreement with the experimental data [211, 212] and confirms results from previous simulation studies [30, 169, 170]. NMR studies of the PLB (1-25) peptide suggest that residues 12 through 17 unwind upon phosphorylation at Ser16 [211]. According to NMR studies of the full-length monomeric PLB, residues 14 through 17 from the C terminus and residue 3 from the N terminus of the CP helix in the Ia domain of PLB unfold upon phosphorylation at Ser16 [209]. As it is shown in Figure 3-4, on average, residues 7 through 11 from the CP helix lose their high helical content upon phosphorylation at Ser16. The number of residues that lose their helical property is more than what is suggested by NMR study of the full-length PLB [209] but is in good agreement with the NMR study of the PLB (1-25) peptide [211]. The helical content of the Ib domain is not influenced by the Ser16 phosphorylation of PLB in agreement with the NMR data [209]. Interestingly, phosphorylation at Thr17 does not appear to affect helicity compared to unphosphorylated PLB.
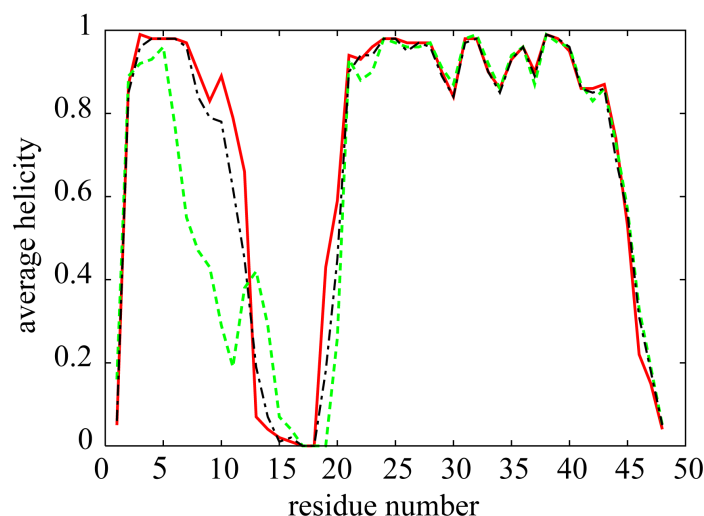
Figure 3-4 Average helicity distribution for different residues for PLB (——), p16-PLB(‑ ‑ ‑) and p17-PLB(•• ••).

Second, interdomain distances between the oxygen atom of the hydroxyl group on Tyr6 and the Cβ atom of Ala24 were calculated (see Figure 3-5). The corresponding distance distributions differ significantly as a function of phosphorylation state. In its unphosphorylated state, PLB exhibits interdomain distances that lie mostly between 12 and 22 Å. Upon phosphorylation, a diverse set of structures gives rise to a much broader distribution with distances between 3 to 30 Å. Phosphorylation at Thr17 results in a slightly wider interdomain distances relative to Ser16 phosphorylated conformations.. FRET experiments have suggested that phosphorylation of PLB at Ser16 decreases the interdomain distance by 3 Å [158]. From the simulations, we found a 2 Å decrease in the overall average interdomain distance, which is considered to be in excellent agreement with experimental data (see Table 3-3) considering the uncertainties in interpreting the FRET data when sampling of multiple conformations with a wide distribution of distance is involved.
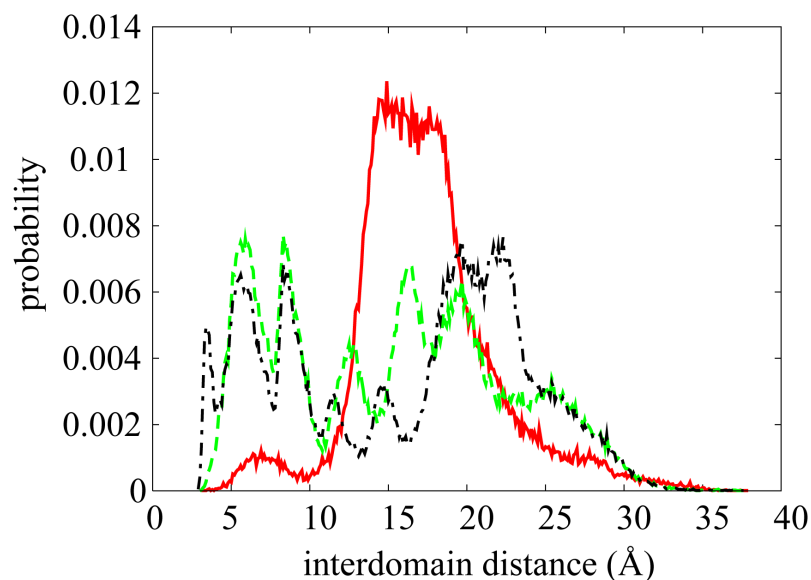
Figure 3-5 Distribution of the interdomain distance between the oxygen atom of the hydroxyl group on Tyr 6 and the Cβ atom of Ala24 for PLB (——), p16-PLB (– – –), and p17-PLB(·· ··).

The solvent accessibility of Tyr6 was further analyzed by calculating its solvent accessible surface area (SASA). We found that phosphorylation at Ser16 increases sampling of conformations where Tyr6 is partially or completely buried. In unphosphorylated PLB, Tyr6 is solvent-exposed in almost all of the structures (see Figure 3-6). The average solvent-accessible surface area (see Table 3-3) is reduced by 52% upon Ser16 phosphorylation, but the margin of error is relatively large. This result agrees at least qualitatively with the 31% decrease in solvent-accessibility measured by fluorescence studies [214].
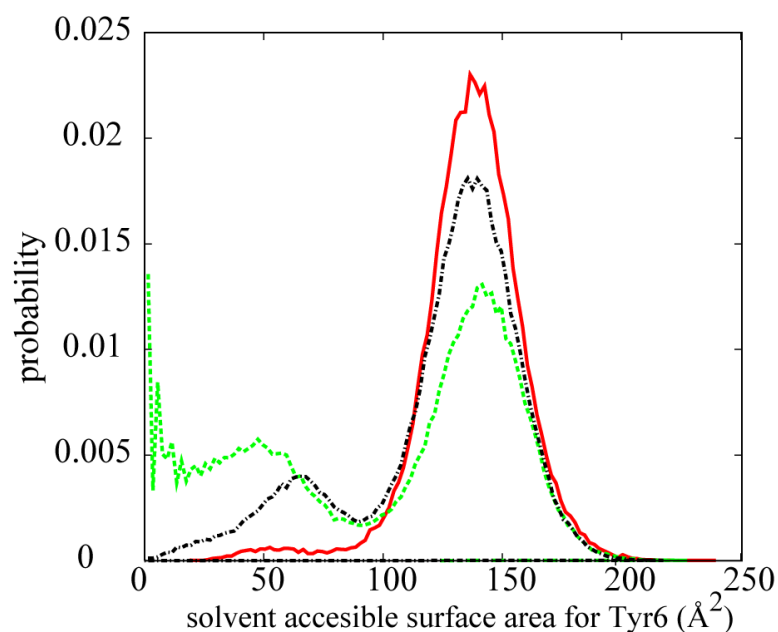
Figure 3-6 Distribution of Tyr6 solvent accessible surface area for PLB (──), p16-PLB (▬ ▬), and p17-PLB(**.. ..**).

	Simulation			Experiment	
	PLB	**p16-PLB**	**p17-PLB**	**PLB**	**p16-PLB**
Interhelical angle (º)[1]	77.4±7.1 (26.8)	98.0±8.2 (31.0)	89.7±6.2 (36.4)	~80±20 [143, 208] ~66 [144]	N/A
Interdomain distance (Å)[2]	17.3±1.0 (4.4)	15.3±2.6 (7.1)	17.9 ±1.2 (7.5)	21.1±0.9 [158]	18.2±0.6 [158]
Solvent exposed area for Tyr6 (Å)	139.9±2.3 (21.6)	66.4±30.1 (54.2)	129.5±4.6 (35.6)	31 % decrease upon phosphorylation [214]	

Table 3-3 Comparison between the calculated properties from the simulation and experimental values for the PLB and phosphorylated PLB systems. Uncertainties are obtained from comparing values between different simulations. Standard deviations are given in parentheses.

Finally, the key question is how the conformational changes in PLB upon phosphorylation affect interactions with SERCA and how Ca^{2+} flux is consequently regulated by PLB. While experimental structures of SERCA-PLB complexes are not available so far, EPR data suggests that by binding to SERCA, the equilibrium between the R and T states of PLB shifts toward the R state without introducing new conformational states for PLB [217]. Under the assumption that PLB in the PLB-SERCA complex resembles conformations seen also in the monomeric form, we performed docking of the PLB structures sampled in our simulations (in the absence of SERCA) to the inactive E2 conformation of SERCA. This was achieved by docking the TM helix of PLB into SERCA according to cross-linking data first and then, superimposing each of the sampled PLB structures at the TM helix region (see Methods section for more details). We then analyzed clashes between the CP helix of PLB and SERCA to determine whether the complex was viable. The results of this analysis are given in Table 3-4. It is found that about 60% of the conformations of unphosphorylated PLB but only 31% of the conformations of Ser16-phosphorylated sampled in the absence of SERCA are structurally compatible with complex formation. We note, that the quality of the intermolecular interactions plays an additional role in determining SERCA-PLB interactions. However, without reliable structural information, such an analysis is likely not meaningful and was therefore not carried out here. Based on the differences in the amount of steric clashes, the simulation results predict that Ser16 phosphorylation would have a destabilizing effect on the PLB-SERCA-E2 complex and likely require a conformational rearrangement of PLB with respect to SERCA if Ser16 phosphorylated PLB stays bound to SERCA as suggested by experiments [158, 210, 215-217]. Presumably, such a change would allow SERCA to switch from E2 to the active E1 form. Interestingly, Thr17 phosphorylation seems to have a different effect. About half of the

conformations sampled in the simulation are still compatible with complex formation. Therefore, at least the interaction of the PLB's CP helix with SERCA may not be completely different from the unphosphorylated complex.

Phosphorylation	None	Ser16	Thr17
Percentage of structures that clash with SERCA	38	69	48

Table 3-4 Percentage of the PLB structures that clash with SERCA in E2 state as a function of phosphorylation.

To understand the conformational characteristics of the structures that clash with SERCA better, the sampled conformations were divided into two groups based on whether they clash with SERCA or not. In Figure 3-2D-I all these structures have been projected onto θ and d-COM for the PLB and phosphorylated PLB ensembles. Comparing Figures 3-2E, 3-2F, 3-2H and 3-2I shows that even though most of the L- and T-shaped conformation with θ values between 60 and 100º potentially clash with SERCA in the Ser16 phosphorylated state, L- and T- shaped Thr17 phosphorylated PLB structures are compatible with SERCA. This suggests that the sampled L- and T- shaped structures, which are the most sampled conformations in phosphorylated states, are structurally different in Ser16 and Thr17 phosphorylated PLB.

3.5   Discussion

In this paper, the effects of Ser16 and Thr17 phosphorylation on the conformational sampling of PLB were studied and compared with a previous study of PLB in DPPC bilayer by means of the HDGB implicit membrane model [156]. To the best of our knowledge, this work is

the first simulation reported for single site phosphorylated PLB at Thr17. This work also covers conformational sampling of p16-PLB over time scales that are much longer than previously reported simulations of Ser16 PLB. Previous short explicit simulations did not reach the time scales needed for lipid relaxation or focused only on the CP domain of PLB in water. Using an implicit description of the membrane with the HDGB model allowed us to cover much longer time scales with moderate computational resources. Furthermore, the use of replica exchange simulation methodology and averaging over multiple separate simulations further increases the effective time scales much beyond of what could be reached with conventional constant-temperature explicit lipid simulations. We estimate that although our simulations are only on the scale of tens of nanoseconds, the dynamics described here occurs in reality on μs to ms timescales as a result of the combined acceleration from replica exchange sampling and avoidance of slow lipid relaxation kinetics.

The key finding of the present work is that phosphorylation of PLB at Ser16 significantly affects conformational sampling of PLB in the CP domain in such a way that it increases the chances of a clash with SERCA in the E2 state. Upon phosphorylation at Ser16, the helical content of the C terminus of the CP domain reduces and new conformations with a break in the CP domain are sampled. Also structures with completely buried Tyr6 residue are observed for p16-PLB. These types of conformations are not observed for p17-PLB. Our results suggest that if p16-PLB remains in the same binding site as PLB, the number of clashes between SERCA and
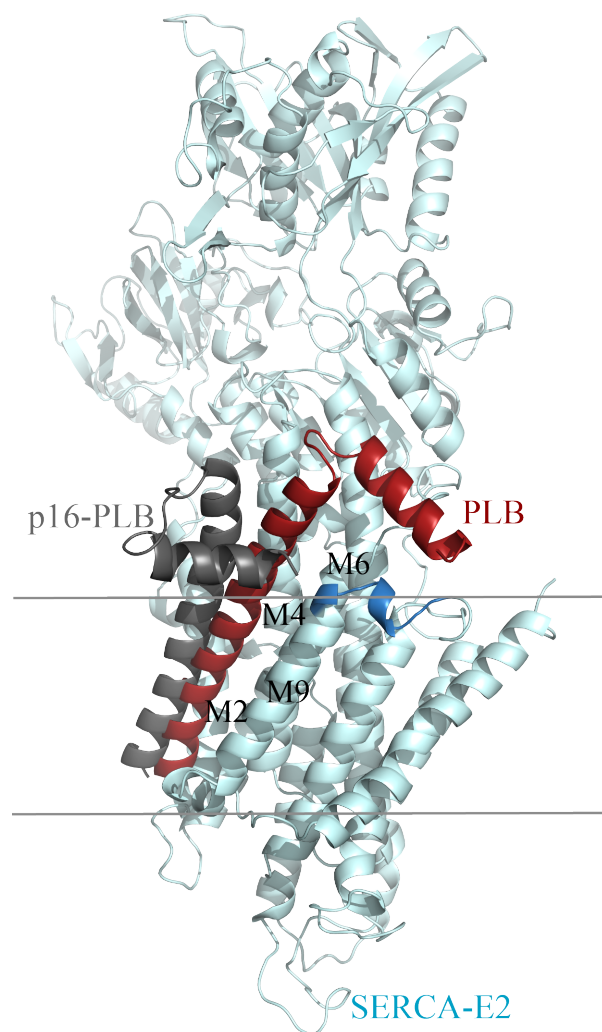
Figure 3-7 Manually docked model of PLB and p16-PLB into SERCA in the E2 conformation. PLB and p16-PLB structures correspond to the minima in the PMFs and SERCA structure is from the crystal structure with the PDB entry code of 1IWO. Upper part of the TM helix number M9 of SERCA and part of the loop connecting helices M8 and M9 are highlighted.

PLB increases significantly. To remain bound to SERCA, we therefore hypothesis that p16-PLB changes its interaction site to prevent such clashes. A model constructed from crystal structure of SERCA in the E2 form and one of the most sampled p16-PLB structures in our simulation is

shown in Figure 3-7. This complex model takes into account available cross-link data for p16-PLB, which is summarized in Table 3-3 (see Methods section). For comparison, a model constructed with one of the most highly populated structures in our previous simulation of unphosphorylated PLB simulation [228] and SERCA-E2 complex is shown in Figure 3-7. The use of the cross-link data introduces uncertainty into the proposed complex model, but while individual cross-links provide only loose restraints, combined data from all available cross-links limits the possibility for alternative conformations. As can be seen, the TM domain of p16-PLB is more detached from SERCA relative to the TM domain of PLB in order to prevent clashes between the membrane-inserted CP domain of p16-PLB and the loop connecting helices M8 and M9 in SERCA (highlighted in Figure 3-7). Such a partial detachment is consistent with the cross-links reported for the lower part of helix M4 in SERCA after PLB phosphorylation at Ser16 [218]. Previous modeling studies of the PLB-SERCA complex have suggested that the TM domain of PLB closely interacts with a groove constructed by helices M2, M4, M6 and M9 of SERCA in the E2 state in the absence of calcium ions thereby preventing SERCA from undergoing a conformational switch to E1 [146]. According to our model, p16-PLB becomes sufficiently detached from this groove to allow the transition from E2 to E1 and subsequently enable enzyme activity [158, 210].

The case of p17-PLB appears to involve a different mechanism. According to our results, the conformations sampled for p17-PLB are still largely compatible with the tight complex proposed for unphosphorylated PLB. A possible explanation for how p17-PLB may nevertheless regulate SERCA activity would involve a shift in binding equilibrium towards PLB dissociation upon p17 phosphorylation. The two different proposed mechanisms for Ser16 and Thr17 phosphorylation are illustrated in Figure 3-8. We further speculate that Ser16 phosphorylation

may be the typical regulatory mechanism for normal cycling of SERCA activity while Thr17 phosphorylation may become involved under special circumstances where complete PLB dissociation is more desirable.
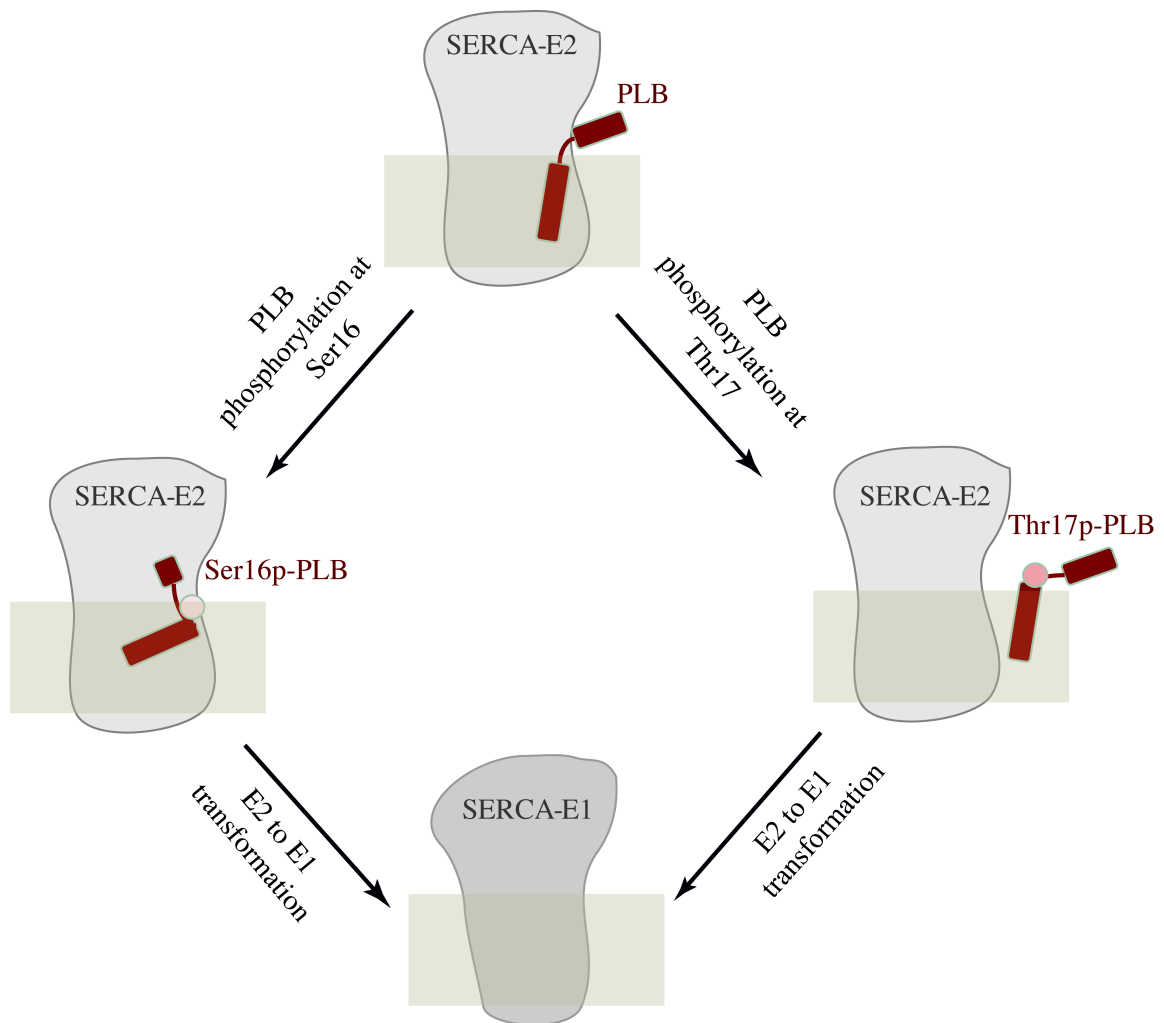
Figure 3-8 A hypothetical diagram for the interaction of PLB and SERCA. Phosphorylation of PLB at Ser16 changes the binding site of SERCA and p16-PLB while PLB is still partially bound to SERCA. The p17PLB hypothetically detaches from SERCA. Either phosphorylation pathway induces the E2 to E1 conformational change in SERCA and activates the pump.

Ser16 phosphorylation alone is sufficient for the maximal cardiac response to β-adrenergic stimulation [135]. However, differential effects have been reported for Thr17 phosphorylation with respect to the frequency dependence during electrical stimulation [203, 204], in increased heart cell contractions, in calcium transients of aerobic interval trained mice [206], and during mechanical recovery after stresses like acidosis and stunning [229]. Based on these observations, Thr17 phosphorylation of PLB may play a role when fast changes in the heart rate are needed or when a fast response is required for heart functioning such as an electric shock. Complete dissociation of PLB as proposed for p17-PLB may facilitate a more rapid activation of SERCA compared to p16-PLB, which is proposed to remain bound to SERCA. Clearly, further studies are needed to understand these questions more completely. In particular, there is a need for more detailed structural data with respect to PLB-SERCA interactions.

Several *in vivo* studies suggest a sequential nature of phosphorylation [138, 201, 230] on the two phosphorylation sites and indicate that Thr17 phosphorylation may typically depend on PLB pre-phosphorylation at Ser16 [231]. Furthermore, experimental studies have suggested an additive effect of double-phosphorylation in terms of activating SERCA [132, 207, 229]. This raises the question of how the double-phosphorylation state affects the conformational sampling of PLB compared to the single-phosphorylated PLB variants. Simulations of the double-phosphorylated state are difficult because of significant pKa shifts expected to result from two neighboring phosphate groups that are likely to partial charge neutralization of one of the phosphates. Because we expect that the pKa shifts are strongly coupled to conformational changes of PLB, we believe that pH dynamics simulation techniques are needed to adequately study this system. Such simulations are outside the scope of this study but are being considered in future work.

3.6   Conclusion

Conformational changes introduced by phosphorylation of PLB at Ser16 and Thr17 have been studied through molecular dynamics simulations. Taking advantage of an implicit membrane environment and replica exchange sampling phosphorylated PLB conformations were sampled on estimated μs-ms timescales. Phosphorylation on either site introduced different changes in the conformational sampling and the resulting energy landscapes. Phosphorylation at Ser16 increases the chances of clash with SERCA and suggests that steric effects cause the reorientation of PLB respect to SERCA. Thr17 phosphorylation on the other hand, does not seem to increase the number of clashed structures with SERCA significantly. This suggests that instead of reorientation, PLB may dissociate from SERCA after phosphorylation on Thr17. Based on the model we present here Ser16 and Thr17 phosphorylation of PLB appear to have different mechanisms, which may reflect different physiological role of each individual phosphorylation.

3.7   Acknowledgements

4 Eps-REMD; A Hamiltonian replica exchange method with an exchangeable dielectric constant

4.1 Introduction

In the straight MD simulations, studying a system with more than one million atoms for timescales larger than tens of μs has remained an ambitious goal. In a rugged energy landscape, the simulation may be trapped in the local minimum free energy states. Therefore, a large extent of the simulation time should be spent to cross the energy barriers. As a result, developing enhanced sampling techniques has become very popular in the recent years. Replica exchange molecular dynamics methods (REMD) are among the most effective techniques to avoid oversampling local minima in the MD simulations. [30].

In REMD, several non-interacting copies (replicas) of a system are simulated simultaneously and independently, each with different conditions. In the most common implementation, the condition to be varied between the replicas is the temperature. At specific intervals, replicas are subject to exchange according to the Metropolis criterion, shown in Equation 4-1:

$$T_{mn} = \begin{cases} 1 & \text{if } \Delta_{mn} \leq 0 \\ \exp(-\Delta_{mn}) & \text{if } \Delta_{mn} > 0 \end{cases}$$

4-1

where

$$\Delta_{mn} = [\beta_n - \beta_m][E(q_m) - E(q_n)]$$
$$\beta_m = \frac{1}{k_B T_m}$$

4-2

In Equation 4-2, k_B is the Boltzmann constant, $E(q_m)$ and $E(q_n)$ are the potential energies of the

conformations in replicas m and n. If an exchange is accepted, the state in the low temperature replica gains more kinetic energy to escape the energy traps by transferring to the higher temperature replica. In a more general approach, replicas with different Hamiltonians are subject to exchange [232]. This technique is called HREM, the Hamiltonian replica exchange method [232]. HREM could be performed through scaling a component of the total Hamiltonian [233] or by adding different harmonic potential functions to different replicas. One of the examples for the latter case, is the distance replica exchange method (DREM) [234].

Experimental data has shown that using solvents with different dielectric constants can alter the structural properties of the peptides [235-238]. Computational studies suggest that changing the solvent's dielectric constant (ε), could affect the strength of the intermolecular hydrogen bonds via changing the partial charges shielding on the proteins [239]. Therefore, the stability and conformational sampling of a system could be affected by the choice of the solvent dielectric [239-243]. In a low ε environment, charges are less shielded. Therefore all the electrostatic interactions are stronger and hydrogen bonds are more stable. The cellular dielectric environment is different from the bulk water. For example in the membrane hydrophobic core, dielectric constant reduces from the bulk water value [107]. High concentration of the proteins can also reduce the effective dielectric constant of the solvent by simply replacing the solvent volume [244]. However, most of the MD simulation studies are performed in a very diluted condition where the dielectric constant of the solvent is close to the bulk water value. In addition, in systems with hydrophilic side chains, low dielectric environment may lead to the formation of the side chain-backbone hydrogen bonds, which in turn partially unfolds the peptides. Other than the electrostatics effect, solvent has a hydrophobic effect on the solute. Hydrophobic side chains

have the tendency to escape from water and form a hydrophobic core. Forming the hydrophobic core is an important driving force in the protein folding process [245].

Changing the solvent's dielectric constant has complicated effects on the stability and conformation of the proteins and peptides. What is clear is that, changing the solvent's dielectric constant alters the energy landscape. In the implicit solvent simulations, the dielectric constant of the solvent is included explicitly in the Hamiltonian. Therefore, in a technique such as the HREM, by exchanging the solvent's ε between the replicas, one can explore the conformational space faster. Unlike the T-REMD method, in which higher temperature replicas mostly unfold the structures, the new technique (called Eps-REMD) focuses on sampling conformations with different folds and can be an alternative for studying such systems.

In this work we have introduced and applied the Eps-REMD techniques to study the conformational space of Met-enkephalin and the synthetic polypeptide (AAQAA)$_3$. The alanine based polypeptide (AAQAA)$_3$ has been a widely studied system to examine the helical propensity of the peptides [246]. (AAQAA)$_3$ has been studies thoroughly by both experiments [247-249] and computational simulations [14, 250]. (AAQAA)$_3$ has been reported to form helical structures. However, the helical content and the relative weight of the α- to π-helical structures depend on the force field used in the simulation [14].

Met-enkephalin is an inhibitory neurotransmitter that binds to the opioid receptor [251]. Experimental [251] and computational studies [234, 252-265] have suggested a wide range of conformations for Met-enkephalin that are sensitive to the environment. Because of the wide variety of the sampled conformations, this peptide has been used as a test system in several enhanced sampling methods [234, 252, 260, 264]. In general Met-enkephalin forms open and

closed conformations, where a salt-bridge is formed between the charged termini in the closed form. Because of the presence of the conformations with a salt-bridge present in the conformational space of Met-enkephalin, this system is a good test case for our new method, which focuses on changing the stability of charged-charged interactions. Sampling of Met-enkephalin is reported to be sensitive to the force field and method of use [256-263, 265]. However, with the use of the same force field and the same implicit solvent method, in this work, we try to reproduce the conformational space generated in the regular T-REMD with the new Eps-REMD technique.

4.2   Methods

4.2.1   The Eps-REMD Method

In the dielectric constant replica exchange method, M independent copies of the system are simulated simultaneously with different dielectric constants, ε_m (m =1,..,M). The equilibrium probability for each replica is

$$P_m = \frac{1}{Z_m} \exp(-\beta E_m(X_m, \varepsilon_m))$$

4-3

where $\beta = 1/(k_B T)$, Z_m is the partition function of the mth replica and X_m is the set of coordinates of the system. In the GBMV implicit solvent method, E_m in Equation 4-3, is a function of ε and X through Equation 4-4, which is a part of the total Hamiltonian and estimates the electrostatic contribution of the solvation free energy.

$$\Delta G_{elec} = -166(1 - \frac{1}{\varepsilon_w}) \sum_i^N \sum_j^N \frac{q_i q_j}{\sqrt{r_{ij}^2 + \alpha_i \alpha_j \exp(-r_{ij}^2 / 4\alpha_i \alpha_j)}} \qquad \text{4-4}$$

In the equilibrium condition, the detailed balance condition:

$$P_m(X_m,\varepsilon_m)P_n(X_n,\varepsilon_n)T_{m,n} = P_n(X_m,\varepsilon_n)P_m(X_n,\varepsilon_m)T_{n,m} \qquad \text{4-5}$$

should hold, where $T_{m,n}$ is the transition probability of exchanging condition m to n. Using Equation 4-5, the transition probability ratio is determined as it shown in Equation 4-6:

$$\frac{T_{m,n}}{T_{n,m}} = \exp(-\Delta_{nm}) \qquad \text{4-6}$$

where

$$\Delta_{nm} = -\frac{1}{k_B T}\{[E_m(q_n,\varepsilon_m) - E_m(q_m,\varepsilon_m)] - [E_n(q_m,\varepsilon_n) - E_n(q_n,\varepsilon_n)]\} \qquad \text{4-7}$$

In Equation 4-7 the temperature, T, is constant in different replicas. An exchange would be accepted based on the Metropolis criteria (see Equation 4-1).

4.2.2   Molecular dynamics simulations

Two peptides were simulated with both the regular temperature replica exchange (T-REMD) and dielectric constant replica exchange (Eps-REMD) techniques, (AAQAA)$_3$ and Met-enkephalin. All the simulations were carried out with an extension to the implicit solvent method GBMV [79, 266] that allows for the simulation of different dielectric environments [83]. All of the GBMV parameters were set to the values described in the original GBMV papers [79, 266] except for β = -12 and S_o =0.65 which were reported to improve the numerical stability of the simulations [173]. Surface tension parameter, γ, was set to 0.015 (cal/molÅ^2).

In all of the simulations, version c34a2 of the CHARMM force field [32] was used with the CMAP correction [38]. Eps-REMD method was implemented in the Multiscale Modeling Tools for Structural Biology (MMTSB) tool set [178] to facilitate using the CHARMM program [177] for the replica exchange simulations. T-REMD simulations and data analysis were taking care of by means of MMTSB tool set as well [178]. Pymol was used to generate molecular graphics [182]. An 18 Å cutoff was applied for calculating non-bonded interaction energies. An integration time step of 1.5 fs was used for all the systems. The bonds that involved hydrogen atoms were constrained with the SHAKE algorithm [176]. Langevin dynamics [174] was used to control the temperature with a uniformed friction coefficient of 10 ps^{-1} for all the non-hydrogen atoms.

All of the simulations were started from the fully extended structures. For each system, one Eps-REMD simulation was performed with eight replicas with dielectric constants of 5, 7, 9, 12, 15, 20, 40 and 80. These ε values were chosen from a series of test runs to maintain the exchange rates of almost 50% between all the neighboring replicas. Each replica exchange cycle consists of 500 steps (0.75 ps) of MD simulations. Temperature was controlled by means of Langevin dynamics [174] and was fixed at 300 K for all the Eps-REMD simulations.

For comparison, one T-REMD simulation was performed for each system with eight replicas. The dielectric constant of the surrounding medium was set to 80 for all the T-REMD simulations. The temperature was spaced exponentially from 300 to 500K for (AAQAA)$_3$ and from 300 to 400K for Met-enkephalin in the T-REMD simulations. Combined (two-dimensional) T-Eps-REMD1 simulations were also performed on both of the peptides with the conditions that are mentioned in Table 4-1. For the Met-enkephalin peptide, a second set of condition was tested for the combined REMD technique, called T-Eps-REMD2. The conditions

for the T-Eps-REMD2 are also reported in Table 4-1. The exchange rates for all of the REMD simulations are reported in Table 4-2 and Table 4-3. Each replica was simulated for 120,000 cycles (90 ns), of which the last 60 ns was used for analysis. The $(AAQAA)_3$ system was blocked at the C-terminal with an amide group. The standard N-terminus (NH_3^+) was added to the N-terminal of $(AAQAA)_3$. Met-enkephalin was simulated with charged termini.

T-Eps1-REMD			T-Eps2-REMD		
Replica	T	ε	Replica	T	ε
1	300	5	1	300	5
2	300	15	2	300	15
3	300	40	3	300	40
4	300	80	4	300	80
5	312.6	80	5	322.4	80
6	325.7	80	6	346.4	80
7	339.4	80	7	372.3	80
8	353.6	80	8	400	80

Table 4-1 Temperature and dielectric constants for different replicas in the T-Eps-REMD1 and T-Eps-REMD2 methods.

	Exchange (%)		
Replicas	T-REMD	Eps-REMD	T-Eps-REMD1
1,2	34	64	92
2,3	30	51	50
3,4	28	51	52
4,5	32	51	47
5,6	40	51	57
6,7	42	52	55
7,8	43	32	53

Table 4-2 Exchange rates for the (AAQAA)$_3$ peptide in simulations with different REMD

techniques.

	Exchange (%)			
Replicas	T- REMD	Eps-REMD	T-Eps-REMD1	T-Eps-REMD2
1,2	73	62	78	82
2,3	74	51	50	51
3,4	74	51	52	52
4,5	74	51	63	49
5,6	75	51	75	58
6,7	75	52	75	57
7,8	76	47	74	58

Table 4-3 Exchange rates for Met-enkephalin peptide in simulations with different

Table 4-3 (cont'd)

REMD techniques.

4.3   Results

In order to test the new technique, a dielectric constant replica exchange (Eps-REMD) simulation of $(AAQAA)_3$ has been carried out and compared with the regular T-REMD simulation. In Figure 4-1, potential of mean force (PMF) for the two replica exchange methods are shown. In the T-REMD simulation, most of the sampled structures have a high helical content from 0.68 to 0.92 (see Figure 4-1A). The population of the other structures decreases significantly and regions with low helical content are sampled less. No conformational sampling is observed in the regions with zero average helicity values. In Figure 4-1B, energy landscape for the same system generated from Eps-REMD simulation is shown. There are significant differences between the two energy landscapes. A minimum region with large values of the average helicity is observed in the T-REMD energy map. In the Eps-REMD simulation, conformations with a wide range of helical content have been sampled. An additional low energy region has been appeared in the Eps-REMD energy map with small average helicity values. In addition, a significant population of structures with the average helicity value of 0.5 is present in the Eps-REMD energy landscape. Population of similar conformations in the T-REMD energy map is much lower.
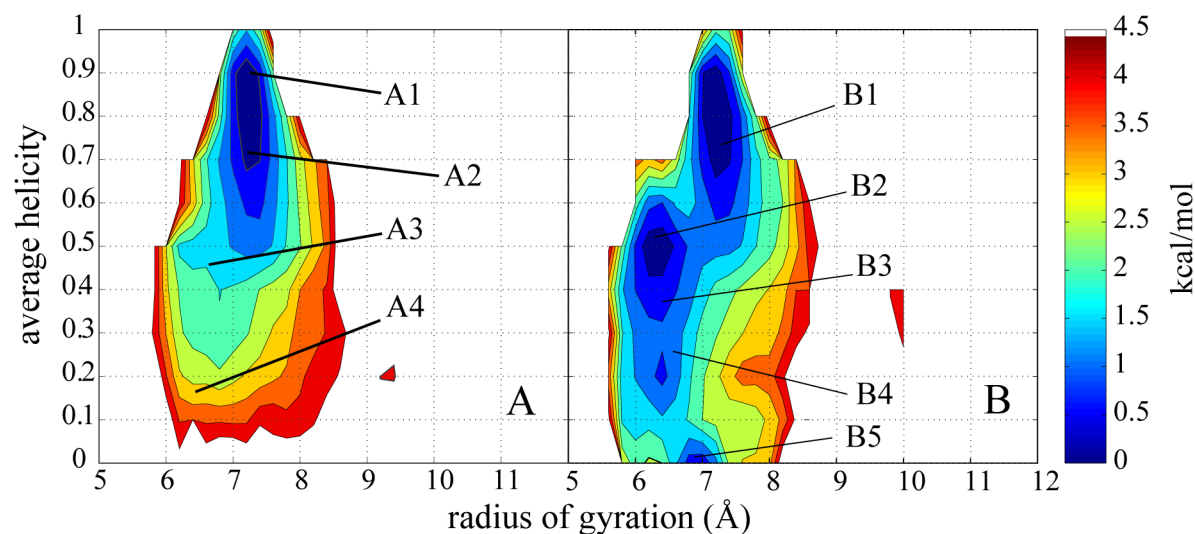
Figure 4-1 Energy landscape for A) T-REMD and, B) Eps-REMD simulations of

(AAQAA)$_3$. The representative structures chosen from clustering the conformations are labeled

and mapped on the energy surface with A1-A4 for the T-REMD and B1-B5 for the Eps-REMD

simulations. The representative structures are shown in Figure 4-2.

In order to gain additional insight into the structural properties of the sampled structures,

all the conformations were clustered and the structures closest to the center of each cluster were

chosen as the representative conformations for that cluster. In Figure 4-2, these representative

conformations are shown. The structures with very small average helicity (such as conformations

B3 and B4 in Figure 4-2) as well as some conformations that were not observed in the T-REMD

simulation (such as B5 in Figure 4-2) were sampled in the Eps-REMD simulation. A closer look

at the conformations sampled in the Eps-REMD simulation is useful for a better understanding of

the discrepancy between the energy landscapes generated with the two different REMD methods.

The hydrogen bonds that are observed in the low helical content structures in the Eps-REMD

simulation are not between the i and $i+4$ residues. Such hydrogen bonds are responsible for the formation of the non-helical structures in the Eps-REMD simulation (Figure 4-3B).
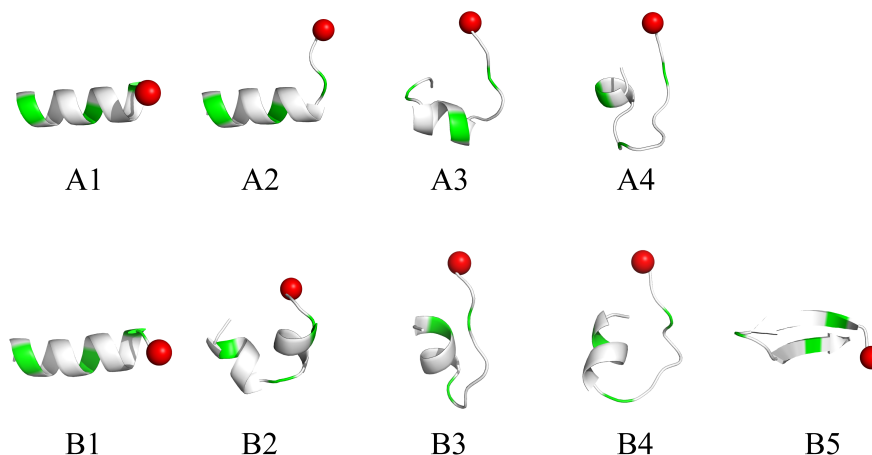
Figure 4-2 Representative structures for (AAQAA)$_3$ generated with the A1-A4) T-REMD and, B1-B5) Eps-REMD simulation. After clustering all of the conformations, the structures closest to the center of the clusters were chosen and mapped on the PMFs (see Figure 4-1). In the cartoon representation, Ala and Gln residues are color-coded with white and green respectively. N-termini are marked with red spheres.
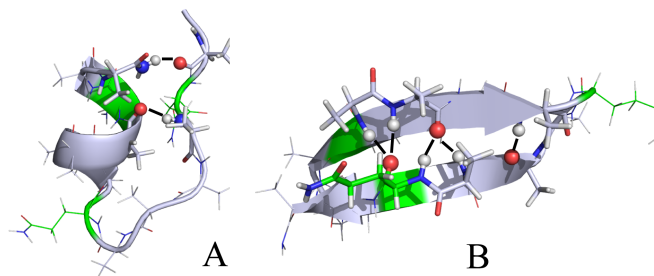
Figure 4-3 Hydrogen bonds that lead to misfolded conformations. Structures A and B correspond to the two minima in the low helicity region on the energy landscape generated by Eps-REMD (B3 and B5 in Figure 4-2B respectively). Side chains are presented with lines and

Figure 4-3 (cont'd)

the residues involved in hydrogen bonding, other than the i, i+4 hydrogen bonds, are shown by the sticks representation. The atoms involved in hydrogen bonds are shown with spheres. Residues are color-coded with white for the Ala and green for the Gln residues. Atomic color-coding is as following: Red for oxygen, white for hydrogen and blue for the nitrogen atoms. The hydrogen bonds are shown with black solid lines.

Differences in the PMFs always might be the result of insufficient sampling. So it is worthwhile to have a look at the effect of simulation time on the generated PMFs. In Figure 4-4A and B, three PMFs have been shown for each of the T-REMD and Eps-REMD simulations of $(AAQAA)_3$ that are generated at different time scales. In the Eps-REMD method generated PMFs, the minimum at the region with the average helicity values between 0.7 and 0.9 is visited less often from the 52.5 ns to the 67.5 ns time scale. The region with the average helicity values between 0.1 to 0.3 and the radius of gyration of 6.5 Å has also been changed gradually over the time. In the T-REMD generated PMFs, the energy surface at low average helicity regions undergoes slight changes.

Behavior of the salt bridges in the Eps-REMD simulation has also been studied by comparing Met-enkephalin conformational space sampled with the T-REMD and Eps-REMD methods. The results are presented in Figure 4-4. Comparing the PMFs generated with the T-REMD and Eps-REMD methods suggests that in the Eps-REMD simulation, only structures with very small end-to-end distance values are being sampled. Both of the termini in Met-enkephalin are charged. In Figure 4-6, representative structures for Met-enkephalin from the T-REMD and Eps-REMD simulations are presented. Figure 4-6 suggests that in the T-REMD simulation, wide

99

range of end-to-end distance values are observed between the sampled structures, while in the Eps-REMD simulation mainly structures with small end-to-end distances are being sampled.
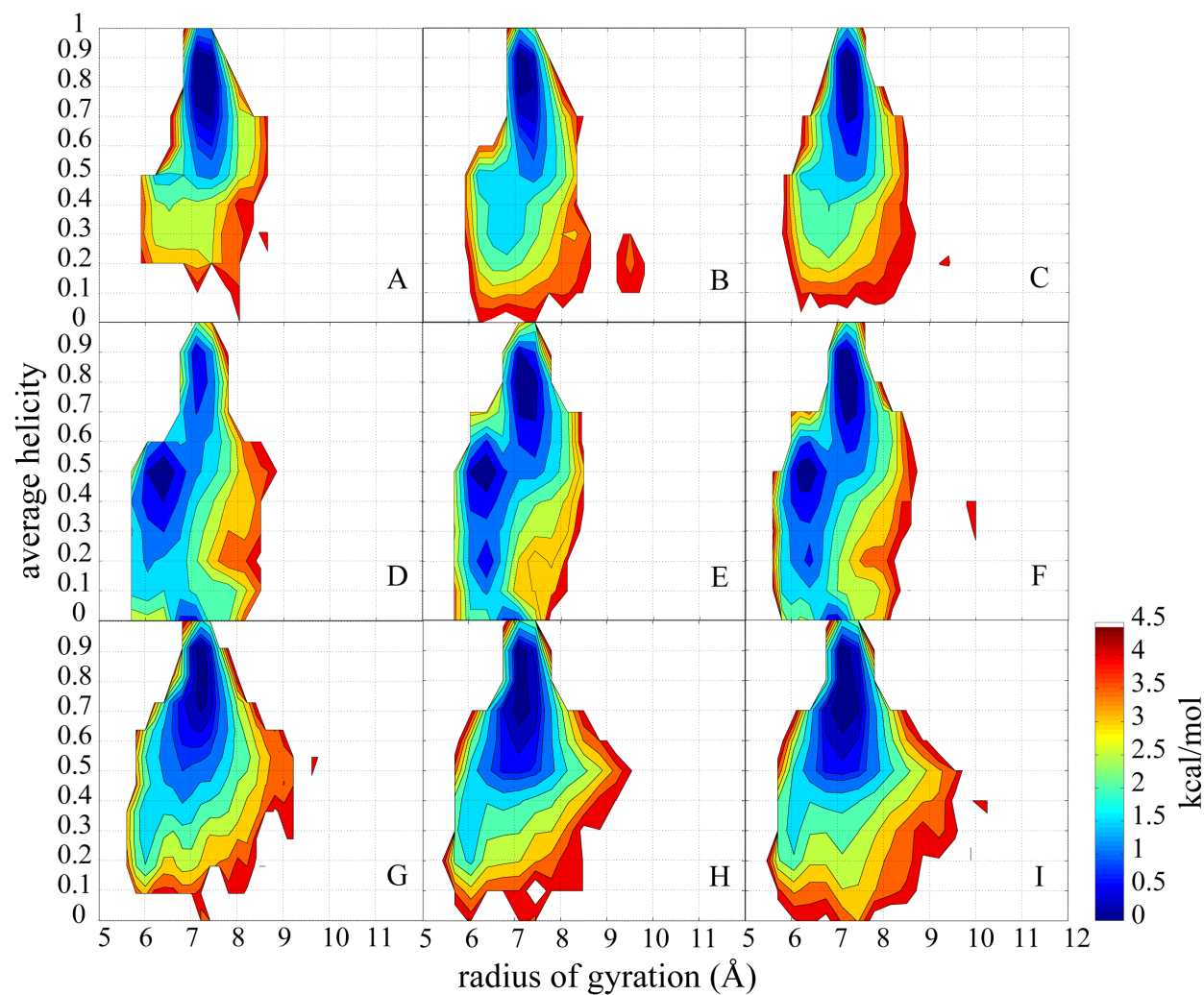
Figure 4-4 Energy landscape for (AAQAA)$_3$ system generated with A-C) T-REMD, D-E) Eps-REMD and G-I) T-Eps-REMD1 methods. The PMFs are generated from 30 ns to: A, D and G) 52.5 ns, B, E and H) 67.5 ns, and C, F and I) 90 ns. First 30 ns of the simulations have been considered as equilibration and discarded in the analysis.

This observation is consistent with the energy landscapes generated with the two REMD methods. Sampling is much broader in the T-REMD simulation and even a small population of structures with end- to-end distance values of almost 16 Å has been sampled (see Figure 4-5). The significant differences between the T-REMD and Eps-REMD simulation energy maps suggest that the Eps-REMD method is far from convergence with the assumption that the T-REMD simulation has been already converged.
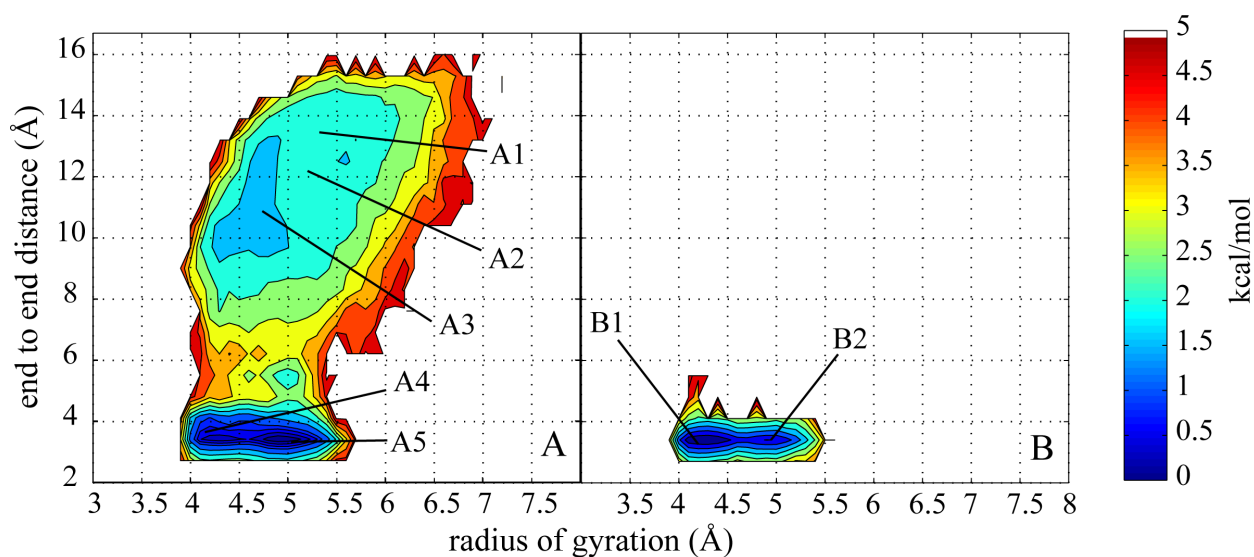
Figure 4-5 Energy landscapes for Met-enkephalin generated by the A) T-REMD and B) Eps-REMD methods. Representative structures from Figure 4-6 are mapped on the PMFs.
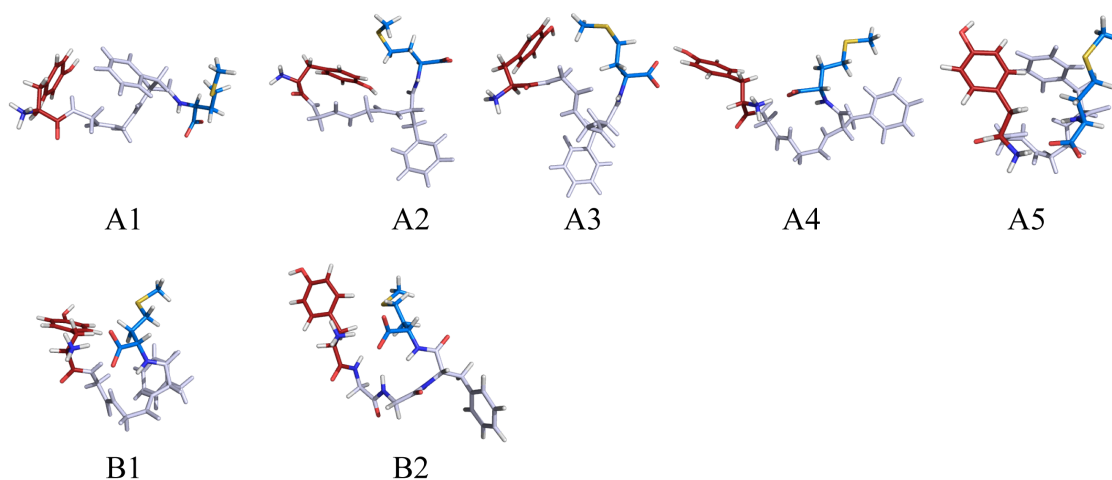
Figure 4-6 Representative structures for Met-enkephalin generated by A1-A5) T-REMD and B1- B2) Eps-REMD methods. The termini are color coded with red and blue.

T-Eps1-REMD			T-Eps2-REMD		
Replica	T	ε	Replica	T	ε
1	300	5	1	300	5
2	300	15	2	300	15
3	300	40	3	300	40
4	300	80	4	300	80
5	312.6	80	5	322.4	80
6	325.7	80	6	346.4	80
7	339.4	80	7	372.3	80
8	353.6	80	8	400	80

Table 4-4 Temperature and dielectric constants for different replicas in the T-Eps-REMD1 and T-Eps-REMD2 methods.

In an attempt to improve the Eps-REMD method's sampling, a combined T and ε replica exchange method, called T-Eps-REMD1, is introduced. The ε and T values for each of the replicas are reported in Table 4-1. In the first attempt to combine T- and Eps-REMD, four replicas out of the eight were spanning temperatures from 312.6 to 353.6 K while in the other four replicas, ε was changing from 5 to 80 at T=300K. The combined T-Eps-REMD1 method was applied to the $(AAQAA)_3$ system and the generated PMF is shown in Figure 4-7. By exchanging both T and ε in the T-Eps-REMD1 method, unrealistic conformations that were sampled with the Eps-REMD method (such as structures B4 and B5 in Figure 4-2) are not observed in the conformational space of the $(AAQAA)_3$ system anymore. Studying the energy landscape generated with the T-Eps-REMD1 technique at different time scales suggests that the energy surface does not change significantly over the last 40ns of the simulation (compare Figures 4-4G, H and I). The PMFs generated by the T-REMD and T-Eps-REMD1 methods are much more similar compare to the PMF generated by the Eps-REMD technique. However, T-Eps-REMD1 method could not reproduce the exact PMF generated by the T-REMD and there are still some slight differences between the PMFs generated by the two techniques (compare Figure 4-1A and Figure 4-7).
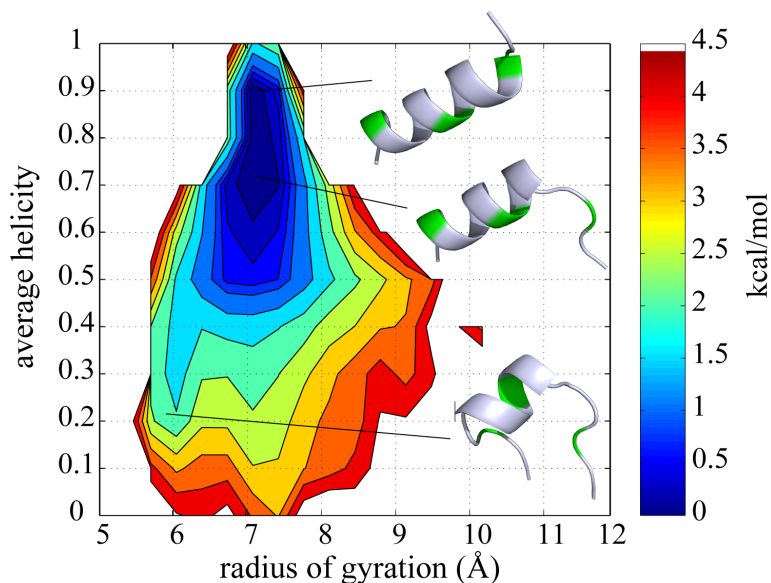
Figure 4-7 Energy landscape for (AAQAA)$_3$ generated with T-Eps-REMD1 method with the condition set mentioned in Table 4-1.

The T-Eps-REMD1 method is also tested for the Met-enkephalin system and the generated PMF is presented in Figure 4-8A. Comparing with Figure 4-5, it is clear that the PMF generated by the T-Eps-REMD1 method is in better agreement with the T-REMD plot (Figure 4-5A). Areas with the end to end distance values higher than 4 Å, that were clearly missing in the PMF generated by the Eps-REMD method in Figure 4-5B, are sampled now with the T-Eps-REMD1 technique (see Figure 4-8A). Again, sampling is not exactly similar between the T-REMD and T-Eps-REMD1 methods (compare Figure 4-5A and 4-8A). The regions with end to end distance values between 8 and 14 Å are sampled more often in the T-REMD method. The highest temperature visited by each replica in the T-Eps-REMD1 simulation is 353.6K, which might not be high enough for crossing the energy barriers and reproducing the T-REMD simulation energy map. Therefore, in the next test case, T was distributed between replicas in

such a way that each replica could visit higher temperatures up to 400K. T and ε values for this REMD condition set (called T-Eps-REMD2) are reported in Table 4-1. These new conditions were used to simulate Met-enkephalin and the generated PMF was compared with the energy landscapes from the T-Eps-REMD1 and T-REMD simulations (see Figures 4-4A, 4-8A and 4-8B). Visiting higher temperatures, with the T-Eps-REMD2 conditions, changes the conformational sampling slightly relative to the T-Eps-REMD1 simulation. These minor differences are more pronounced in the large end-to-end distance value regions, covering end-to-end distance values between 9 to 12 Å, which are visited more often with the T-Eps-REMD2 condition set. These results suggest that longer simulation time might be required for convergence of the T-Eps-REMD simulations.
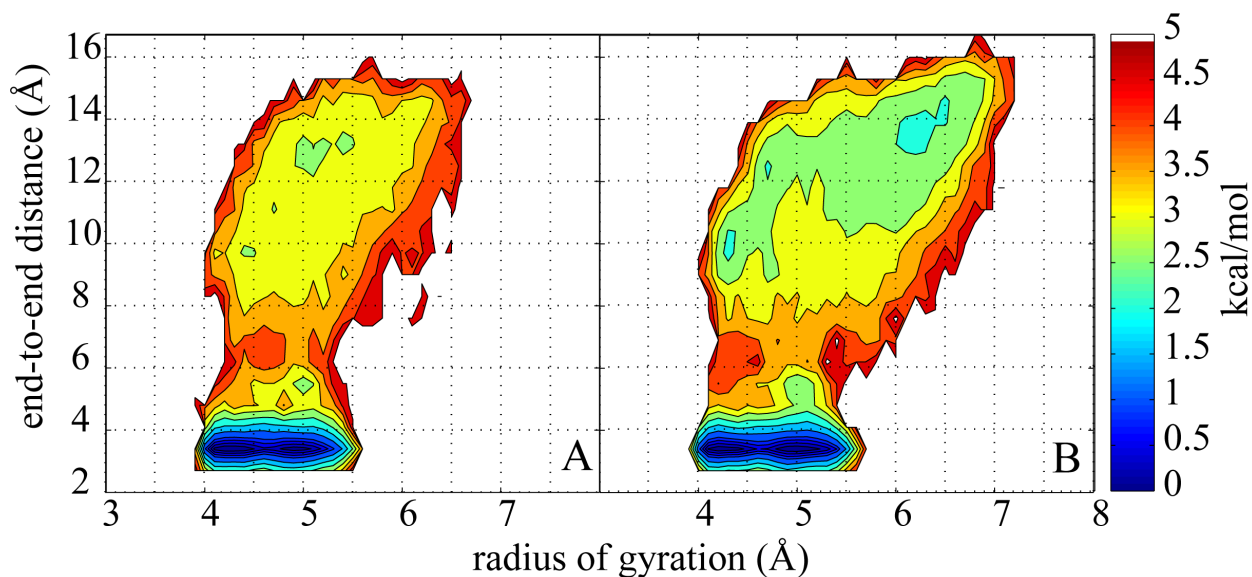
Figure 4-8 Energy landscape for Met-enkephalin generated by A) T-Eps1-REMD and, B) T-Eps2-REMD methods.
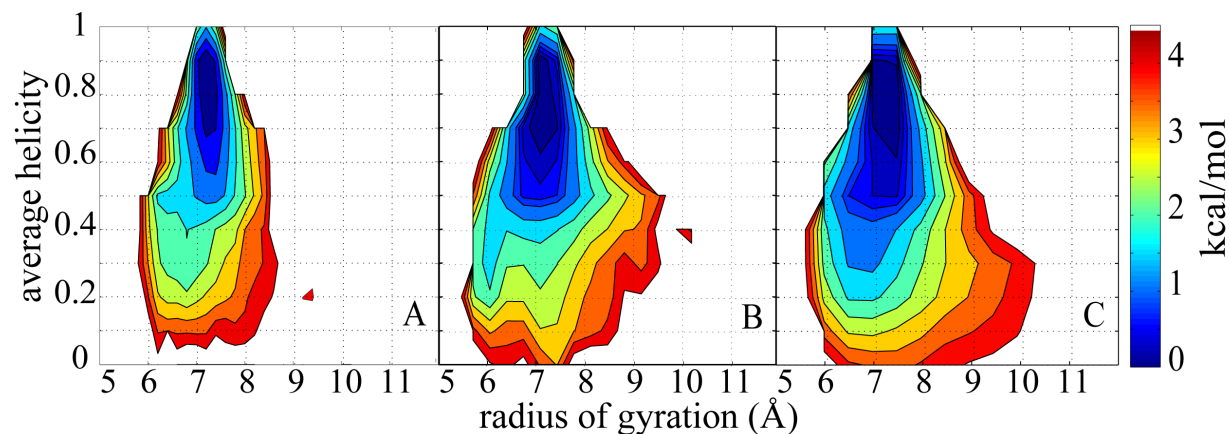
Figure 4-9 Energy landscape for the (AAQAA)$_3$ peptide generated with A and C) the T-REMD and B) T-Eps-REMD1 methods. In generating the PMFs in A and B only the sampling from one replica (the replica with ε=80 and T=300K) has been used. WHAM has been used to generate the PMF in panel C from all the 8 replicas in the T-REMD simulations.

In order to study the effects of the weighted histogram analysis (WHAM) [181] technique on generating the PMFs, the energy landscape of the (AAQAA)$_3$ peptide from T-REMD simulation is generated without (called T-REMD, Figure 4-9A) and with (called T-WHAM, Figure 4-9C) WHAM technique and is compared with the PMF from the T-Eps-REMS1 (Figure 4-9B) simulation that is generated without WHAM. If WHAM is not used, only sampling form one of the replicas with the condition that represent the biological condition (in this case, T=300K and ε=80) is used. By using WHAM, the biased conformational samplings from other replicas are reweighted and contributed to the PMF. This extra contribution results in a more converged energy landscape. Overall, the energy maps are quite similar with slight differences, mostly in the high-energy regions. The deep minimum area is more localized in the T-REMD simulation map compare to the T-WHAM (compare Figure 4-9A and Figure 4-9C). By using the

sampling only from one replica in the T-REMD (such as Figure 4-9A), there is no sampling in the areas with the radius of gyration values higher than 9Å. In contrast, in both T-Eps-REMD1 (Figure 4-9B) and T-Wham (Figure 4-9C) PMFs, sampling is extended to the radius of gyration values higher than 9 Å. The T-Eps-REMD1 map is very similar to the T-WHAM energy map with one exception of a more extended minimum region with the average helicity value of 0.3 and the radius of gyration of 6 Å.
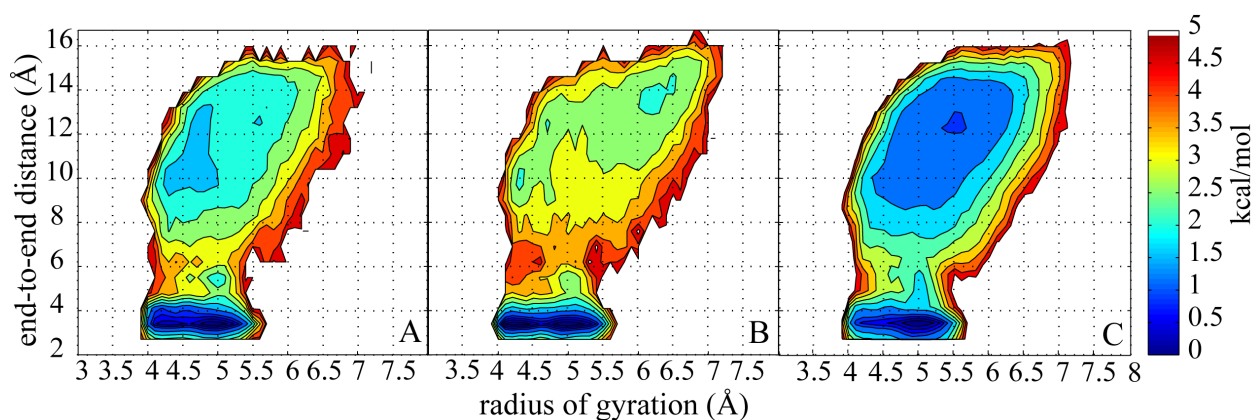
Figure 4-10 Energy landscapes generated for the Met-enkephalin peptide from A and C) the T-REMD and B) T-Eps-REMD2 simulations. Panel C presents the PMF generated with the reweighted sampling from all of the 8 replicas in the T-REMD simulation, using WHAM. For the PMFs presented in the panels A and B, only conformations from one replica (the replica with T=300K and ε=80) is used.

In Figure 4-10, effects of using WHAM in generating the PMFs from the T-REMD simulation of the Met-enkephalin system is shown. The differences between the T-REMD energy landscapes generated without (Figure 4-10A) and with (Figure 4-10C) WHAM are more

pronounced in the Met-enkephalin system, especially for the regions with the end-to-end
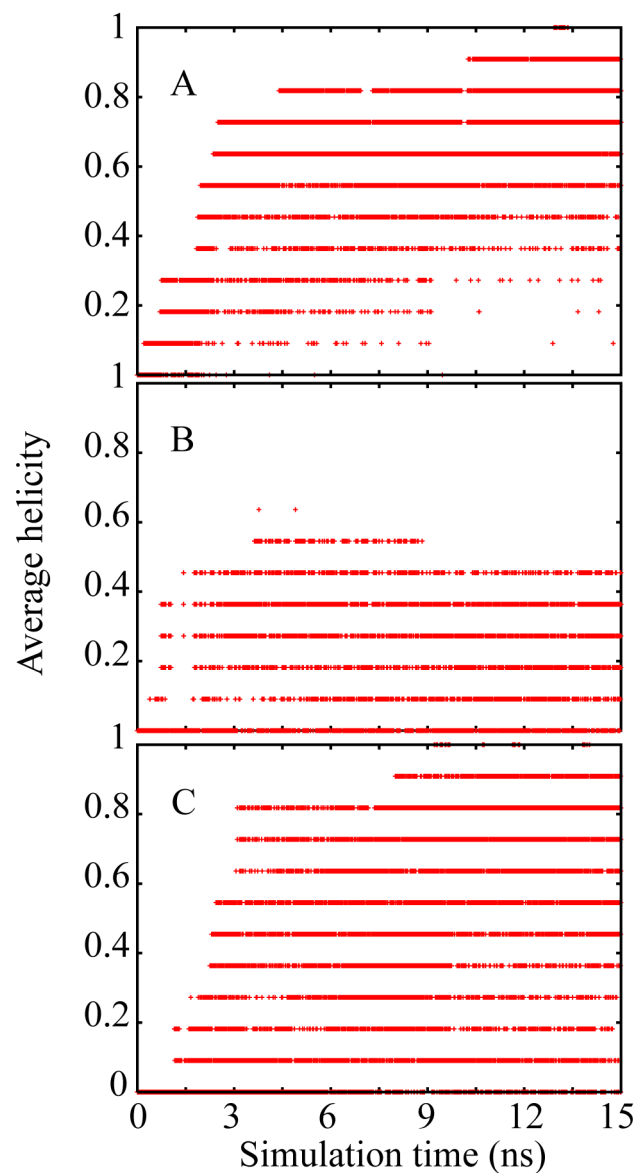
Figure 4-11 Monitoring average helicity values over the simulation time for A) the T-REMD B) Eps-REMD and, C) T-Eps-REMD1 methods for the (AAQAA)$_3$ peptide.

distance values between 9 and 14 Å. Sampling in these regions are higher when the PMF is generated with WHAM (compare Figure 4-10A and 4-10C). These differences suggest that

ignoring the reweighted sampling form other replicas may result in an un-converged energy landscape. Even though Met-enkephalin is a very small system with only 5 residues, comparing the energy landscapes from the T-REMD and T-Eps-REMD1 simulations with the T-WHAM map suggests that if sampling from only one replica is being used, more simulation time should be extended so that all of the conformational space is sampled with one single replica.

Generating the correct PMF is not the only objective of the new T-Eps-REMD method. It is also essential to show that the new method is potentially faster in sampling the conformational space and folding an extended structure. Although the (AAQAA)$_3$ system is a very small peptide which folds very fast with the regular T-REMD method, it is worthwhile to look at the folding speed in the different methods introduced here and compare them with the T-REMD technique. We have calculated the average helicity for this peptide from the replica with the biological condition (T=300K and ε=80). Starting from the very beginning of the simulation, when the peptide is completely unfolded, the average helicity has been monitored over the first 15 ns of the simulation. The results are shown in Figure 4-11. As expected from the Eps-REMD technique, structures with small average helicity values are overestimated and in the first 15 ns of the simulation, no conformation with average helicity values higher than 0.6 has been sampled (see Figure 4-11B). Both the T-REMD and T-Eps-REMD1 methods are successful in folding the peptide within the first few nanoseconds of the simulation. However with the T-Eps-REMD1 technique, conformations with more than 80% helical content are sampled almost 1 ns faster. It should be mentioned that sampling conformations with very high average helicity values does not necessary mean that the simulation converges faster. As it is shown in Figure 4-11C, over the first 15 ns of the T-Eps-REMD1 simulation, many non-helical structures are also sampled, which suggests that the simulation is far from convergence. However, in a large protein with several

Figure 4-12 Occurrence of different conditions in each replica for A) T-REMD, B) Eps-REMD

and C) T-Eps-REMD methods for the (AAQAA)$_3$ system.

folds, faster sampling of helical conformations may suggest a promising technique for accelerating the process of protein folding.

An interesting feature of the Eps-REMD and T-Eps-REMD is the reduction of the phase separation in the REMD simulations. As it is suggested in Figure 4-12, phase separation can occur in the T-REMD simulations when some replicas do not visit all of the temperatures. Phase separation over time would cause participation of less number of the replicas in the sampling of the condition of interest (T=300) and loosing the effectiveness of the REMD method. In such a case, for long periods of time, no exchange would be accepted between the replicas. By introducing ε, as a new exchangeable condition, there would be enough overlap between the replicas and exchanges occur all along the simulation time. In Figure 4-12B, it is shown that the occurrence of each condition in different replicas has been increased and all replicas visit all the conditions (different ε in this case) frequently. In the T-Eps-REMD (Figure 4-12C) frequency of occurrence has been slightly reduced relative to the Eps-REMD but it is still more often than the T-REMD simulation.

4.4   Discussion

As a new approach to the replica exchange methods, in this work, dielectric constant of the solvent has been exchanged between the neighboring replicas. Forming the hydrophobic core is the driving force in the protein folding process. Therefore, one might expect that high ε in the solvents such as water, facilitates the folding process. However, folding is more complicated that just forming the hydrophobic core. Inter and intra molecular hydrogen bonds play an essential role in the folding process. In a very low ε solvent, charges are not shielded and

hydrogen bonds and salt bridges are more stable. If a peptide consists of residues with nonpolar side chains, lowering ε would lead to forming high helical content structures [243]. In the presence of the polar and charged residues, this picture becomes more complicated, because not only backbone hydrogen bonding can occur, but also side chain-side chain and side chain-backbone hydrogen bonds can form. What is clear though is that, changing ε results in sampling different structures and changes the stability of the sampled conformations. At Low ε environment, the hydrophobic effect weakens. Therefore, conformations lose their compactness. This effect could help the conformations to visit a different tertiary structure without the necessity of getting completely unfold. In real biological systems, chaperones are known to stabilize such intermediate states and prevent aggregation [267]. All these reasons suggest that the solvent ε can be used as an exchangeable term between the replicas in the HREM regime. However, as it is shown, HREM with just different ε is slower in sampling the conformational space relative to the T-REMD method. By lowering $\varepsilon,$ all charge-charge interactions get stronger and unrealistic hydrogen bonds could form (see Figure 4-3). These hydrogen bonds are stable especially in the low ε environment. Therefore, it is necessary to introduce T as the second exchangeable parameter in the HREM method in order to break the unrealistic hydrogen bonds.

It is worthwhile to mention that, a two-dimensional HREM was tested for T and $\varepsilon.$ However, this technique was very slow to converge because these two parameters have the opposite effects on stabilizing the secondary and tertiary structure of the peptides (data not shown). The high T replicas are necessary to destabilize the wrong charge-charge interactions while the low T replicas are required to stabilize the correct secondary structure of the peptides. The condition that the sampled structures are being analyzed is 80 for ε and 300K for T.

Therefore any combination of T higher than 300K with ε values lower than 80 would cancel the effectiveness of HREM. By adding T as an exchangeable parameter in the T-Eps-REMD method introduced here, the sampling space is much more similar to the T-REMD simulation. Small differences between the maps could be the result of not enough sampling. However, applying a higher temperature in the second version of the T-Eps-REMD method did slightly improve the agreement between the energy landscapes for the Met-enkephalin peptide (see Figure 4-8).

4.5 Conclusion

In this chapter, T-Eps-REMD method has been introduced as an alternative for the T-REMD method. Exchanging only the solvent ε does not reproduce the low energy regions in the T-REMD energy landscapes and results in forming unrealistic conformations with wrong hydrogen bonds. By adding temperature, as an exchangeable condition, to the Eps-REMD, the agreement between the generated energy landscapes improves to a large extent. By lowering ε it is possible to decrease the energy barriers and sample a wider range of conformational space. Visiting higher T replicas would break the wrong hydrogen bonds and salt-bridges more easily. Here, the new method is tested on two small peptides. The initial results are promising but longer simulations are needed to confirm the convergence in both T-REMD and T-Eps-REMD simulations. However, the current results suggest that this technique is at least as effective as the T-REMD method and it has the potential to work well for the large systems where several secondary structural domains are present.

5 Conclusions and the future direction

Heterogeneous Dielectric Generalized Born method, HDGB, is one of the most accurate implicit solvent/membrane methods to study the behavior of the biological systems. The main focus of the current work is to show that this method could generate accurate results comparable with the experimental data and could suggest new hypothesis about the function and mechanism of the biological systems.

Molecular dynamics simulation of Phospholamban, PLB, using the HDGB method, successfully suggested that the conformational space of this peptide was broader than what the NMR experiment was reporting at the time. Structural properties of PLB, averaged over the structures from the conformational sampling of PLB were in a very good agreement with the experimental data. More over, studying the conformational sampling of PLB in different phosphorylation states suggested a mechanism for the inhibitory function of PLB. The data suggested that PLB binds to SERCA in the E2 state and locks the conformational state of SERCA. By disrupting the E2 to E1 conformational change in SERCA, the Ca^{2+} cycle is interrupted and Ca^{+2} flux is reduced in the muscle cell.

Phosphorylation of PLB at Ser16 reduces the number of structures that can be fitted into the SERCA binding groove by introducing some structural changes in PLB. Therefore, by phosphorylating the Ser16 residue on PLB, population of the structures that are forced to leave the groove, increases which improves the chance of transformation of SERCA from E2 into the E1 state and proceeding the Ca^{2+} uptake cycle.

Such a structural change was not observed after Thr17 phosphorylation of PLB. Therefore, it is possible to assume that the effect of Thr17 phosphorylation on the inhibitory mechanism of PLB is not structural. An alternate hypothesis is that phosphorylation of PLB at Thr17 affects the energetics between SERCA and PLB. However, the preliminary work on studying the effects of Thr17 phosphorylation of PLB on the SERCA-PLB binding free energy did not come to a clear conclusion. More detailed studies are required to confirm this hypothesis. An example for such a case study would be a MD simulation of the PLB-SERCA complex where SERCA is in the E2 state and PLB is phosphorylated at Thr17. PLB conformations are structurally compatible with the PLB-SERCA complex formation. Therefore, it is possible to construct a model from the Thr17 phosphorylated PLB and SERCA. Studying the stability of such a system would be useful to enlighten the influence of Thr17 phosphorylation on the PLB-SERCA binding.

A further step in studying the PLB inhibitory function could be an MD simulation of the PLB-SERCA complex where PLB is in the unphosphorylated state. Studying such a system can provide a clear picture for the PLB and SERCA binding site and answer many questions about PLB-SERCA interaction, which with the lack of any crystal structure for the complex, are currently difficult to answer. It is also worthwhile to investigate the changes in the PLB conformation and interaction with SERCA along the E2 to E1 transition of SERCA. In such a biased MD simulation, conformational state of SERCA can be forced to change from E2, in the SECA-PLB complex, to E1. Such studies require a balanced description of the peptide-peptide and peptide-solvent interactions. As will be discussed in the Appendix section, the current HDGB method does not include the favorable dispersion term. Therefore modifying the HDGB method to include this term might result in a better balance in the energetics of any system,

which consists of more than one peptide in the membrane. In addition, because SERCA has a large size compare to PLB, seeking more innovative enhanced sampling methods to study such a system would be of an interest. A technique such as the one described in section 4 might be useful to investigate conformational sampling of SERCA.

Current work suggests that PLB phosphorylation has two modes. One is when Ser16 is phosphorylated and the PLB-SERCA binding site is changed but PLB is still in close interaction with SERCA. The second mode is when PLB is phosphorylated at Thr17 and dissociates from SERCA completely. According to this scheme, Thr17 phosphorylation has a stronger effect on removing the inhibitory effect of PLB than the Ser16 phosphorylation. The reason is that when PLB is away from SERCA in the Thr17 phosphorylated mode, it is more difficult for PLB to associate with SERCA after dephosphorylation. Therefore, interruption of the E1-E2 transition of SERCA is more difficult and the Ca^{2+} ion flux is stronger.

However, it seems that single phosphorylation of Thr17 is not the way that the nature inhibits SERCA in the regular conditions. In the *in vivo* studies, Thr17 phosphorylation has been reported to follow Ser16 phosphorylation. The only condition under which Thr17 is phosphorylated in the *in vivo* studies without Ser16 being phosphorylated is when the concentration of Ca^{2+} ion in the cell is very high and when the phosphatase that dephosphorylates PLB is inhibited [139]. Experimental studies also suggest that Ser16 phosphorylation of PLB is sufficient for the maximal cardiac response to the β-adrenergic stimulation in the *in vivo* studies [135]. These findings suggest that in the regular conditions, single site phosphorylation of PLB at Thr17 does not occur. A hypothesis for the reason why single site phosphorylation of PLB at Thr17 is not favorable in the regular condition is that, because of the complete dissociation of Thr17 phosphorylated PLB from SERCA, it is more

difficult for PLB to associate SERCA again for another cycle of the Ca^{2+} uptake. Another possibility is that after dissociation of PLB from SERCA, it aggregates and forms a pentamer, which is not the active form of PLB. Studying the pentameric form of PLB associated with SERCA would also be a very interesting research to conduct. The electric stimulation of the heart cells has been reported to induce singe site phosphorylation of PLB at Thr17 in a frequency-dependent manner [268]. This finding together with the possibility of the single-site Thr17 phosphorylation of PLB in the high concentration of Ca^{2+} ion, suggests that this mechanism of PLB phosphorylation is used when a sudden change in the contraction-relaxation cycle of the muscle cell is required. In such a case a signal different from the β-adrenergic stimulation induce a large flux of Ca^{2+} ion into the SR.

Finally, the effect of the single site Thr17 phosphorylation could be different from the effect of the dual site Thr17 and Ser16 phosphorylated PLB. *In vivo* studies suggest that Ser16 phosphorylated PLB exhibits the same inhibitory effect as the dual site phosphorylated PLB [205]. Studying the dual site phosphorylated PLB with the MD simulations and comparing it with the case where PLB is phosphorylated at only one site, could help to answer sever questions about the function and role of the dual site phosphorylated PLB.

APPENDIX

Implementation of a van der Waals dispersion term in the HDGB method

A-1 Introduction

Solvation plays an essential role in different biological processes [269, 270]. Therefore it is crucial to model the solvation effects accurately when simulating the biological systems. The most detailed model of the solvent is described by representing all of the solvent molecules explicitly. In this case, all of the solvent-solvent and solvent-solute interactions are calculated during the simulation. However, if studying some specific interactions between the water molecules and the solute is not required, it is possible to describe the solvent effects implicitly and reduce the computational costs [50, 54, 55, 61, 65-69, 72, 83, 94].

Most of the recent implicit solvent methods divide solvation free energy into the polar and non-polar terms. The contribution of the non-polar term to the total solvation free energy in water is small, relative to the polar term. However, in the hydrophobic environment of membrane bilayers, the non-polar contribution of the solvation free energy is significantly higher. Therefore, an accurate description of the non-polar term in the solvation free energy, in particular the van der Waals dispersion term, is especially important for the implicit membrane models.

The nonpolar contribution of the solvation free energy, ΔG_{np}, is usually treated with the solvent accessible surface area (SASA) models [72, 78, 84-86, 94, 271-275]. In the SASA

model, ΔG_{np} is approximated to be linearly proportional to the solvent accessible surface area of the solute as:

$$\Delta G_{np} = \sum_i \gamma_i SASA_i + b$$

A-1

where $SASA_i$ is the solvent accessible surface area for the atom type i of the solute, and γ_i is an adjustable parameter corresponding to the surface tension coefficient. In Equation A-1, b is set to zero in some methods [79, 94, 273] or it is another adjustable parameter to obtain a better fit to the experimental data [115]. The experimental data, which is used for parameterization of Equation A-1, varies from the transfer free energies of alkanes from pure liquid hydrocarbons to water [276], to the solubility of the amino acid analogs in the organic solvents [277]. A wide range of γ values, from very small values, such as 5 cal/molÅ^2 [271], to values as large as 138 cal/molÅ^2 [113], has been reported in different studies based on the method of parameterization of Equation A-1.

However, it is suggested that for small organic molecules, solvation effects could be better described with the solvent accessible volume [63]. The solvent accessible volume models have shown to work well for both the non-charged [278, 279] and charged [279, 280] organic molecules. Further studies suggested that the solute size affects the dependency of the solvation free energy to either the solvent accessible surface or the solvent accessible volume. If the solute is small, the solvation free energy scales better with the solute accessible volume [110, 281-283]. The non-polar contribution of the solvation free energy can be divided into the cavity and van der Waals dispersion terms according to Equation A-2:

$$\Delta G_{np} = \Delta G_{cav} + \Delta G_{vdw-dis}$$ A-2

The cavity term is an unfavorable term, which describes the cost of cavity formation to accommodate the solute in the solvent. The dispersion term arises from the attractive van der Waals interactions between the solute and solvent molecules. When separating the cost of cavity formation from the van der Waals interactions, the SASA model is commonly used to capture only the cavity term in the non-polar contribution of the solvation free energy [115, 284-286]. The van der Waals term is then described with a different formalism. The necessity of defining a non-SASA term to describe the attractive van der Waals interactions has been further confirmed by Gallicchio *et al*. when they tried to fit the solvation free energies of the small organic molecules to a model with both the SASA and non-SASA terms [287]. In addition, studies on the association free energies of the organic molecules [288] and small peptides [286] failed to approximate the non-polar contribution of the solvation free energies with a SASA-only model. In a very crude approximation, the van der Waals dispersion term can be described with a very simple term that only depends on the solute atom type [287] (the second term in Equation A-3).

$$\Delta G_{np} = \sum_i \gamma_i(t_i) SASA_i + \alpha(t_i)$$ A-3

In Equation A-3, t_i specifies the solute atom types. For a more sophisticated description of the dispersion term, other methods decomposed the Lennard-Jones potential into the attractive and repulsive terms using the Weeks-Chandler-Andersen (WCA) [289] scheme. The repulsive term, which describes the cavity formation, is approximated with a SASA model [290]. The dispersion term is then approximated by integrating the attractive part of the Lennard-Jones potential over the solvent space, according to Equation A-4 [284, 291, 292]:

$$\Delta G_{disp} \cong \sum_i^N \rho_w C_i \int_{solvent} dr(r - r_i)^{-6}$$

A-4

where N is the total number of the solute atoms, C_i is an atomic constant, r_i is the location of atom i, and ρ_w is the bulk number density of water. The water number density could be assumed to be homogenous such as Equation A-4 [87] or it could be weighted by the radial distribution function of water [291].

In the AGBNP method, introduced by Gallicchio *et al.*, the volume integral in Equation A-4 is approximated with the definition of the Born radii from the Generalized Born formalism [87]:

$$\Delta G_{vdw-dis} \cong \sum_i a_i \frac{A_i}{(\alpha_i + R_w)^3}$$

A-5

where α_i is the Born radius of atom I, R_w is the radius of a water probe or the solvent offset, a_i is a dimensionless parameter. In Equation A-5, A_i is defined as:

$$A_i = -\frac{16}{3} \pi \rho_w \varepsilon_{iw} \sigma_{iw}^6$$

A-6

where ρ_w is the number density of the water molecules in the standard condition with the value of

0.33428 Å^{-3}, and ε_{iw} and σ_{iw} are the water-solute Lennard-Jones (LJ) parameters for the solute atom i. LJ parameters in Equation A-6 are calculated as:

$$\varepsilon_{iw} = \sqrt{\varepsilon_i \varepsilon_w}, \quad \sigma_{iw} = (\frac{\sigma_i + \sigma_w}{2})$$

A-7

where ε_i, ε_w, σ_i, and σ_w are from the CHARMM force field as well as the LJ equation:

$$U_{LJ} = \varepsilon_{ij}[(\frac{\sigma_{ij}}{r_{ij}})^{12} - 2(\frac{\sigma_{ij}}{r_{ij}})^6]$$

A-8

In the current version of the implicit membrane method used in the most of this dissertation, the heterogeneous dielectric generalized Born (HDGB) [84] formalism, the total non-polar term of the solvation free energy is approximated by a SASA model which does not include the attractive LJ interactions between the membrane and the solute atoms. This leads to the underestimation of the solvation free energy in the membrane, especially for large proteins with large SASA values. Here, a dispersion term is added to the HDGB implicit membrane model. We will show that for a membrane protein test case, introducing the dispersion term in the HDGB method and balancing the cavity term, leads to a more energetically favorable environment for the membrane protein in the membrane relative to the solvent.

A-2 Methods

A-2-1 The dispersion term in the HDGB method

In the modified HDGB formalism, the non-polar contribution of the solvation free energy is calculated according to Equation A-4 with the cavity term:

$$\Delta G_{cavity} = \sum_{i=1}^{N} SASA_i \gamma S(z)$$

A-9

where $SASA_i$ is the solvent accessible surface area for atom i, γ is the surface tension parameter and $S(z)$ is a switching function called the non-polar profile (for more details see the introduction chapter). In order to include the dispersion term in the HDGB method, we have modified the dispersion term introduced in the AGBNP method [87] by Gallicchio *et al*. (Equations A-5 and

A-6). In the membrane, the solute is exposed to a heterogeneous environment that varies based on the distance from the membrane center (z). The effect of the heterogeneous environment is the presence of different types of 'solute' atoms with different densities as a function of z. Therefore, the number density, ρ_w, in Equation A-6 should change as a function of z and the solute atom types. For example, the number density of the oxygen atoms in the DPPC bilayer decreases from the head group region toward the membrane center. In contrast, the number density of the hydrogen atoms increases toward the center of the membrane. Therefore, it is necessary to define separate density profiles for different atom types of the membrane bilayer. Equations A-5 and A-6 are then modified for the membrane as:

$$\Delta G_{vdw-dis} = \sum_{i}^{N} \sum_{j}^{M} a_i \frac{A_{i,j}}{(\alpha_i + R_j)^3}$$

A-10

and

$$A_{i,j}(z) = -\frac{16}{3}\pi\rho_j(z)\varepsilon_{ij}\sigma_{ij}^6$$

A-11

In the modified expression for $\Delta G_{vdw-dis}$, there is an additional sum over membrane atom types and oxygen of water molecules, where N is the total number of the solute atoms and M is the number of the bilayer atom types plus one for water. Different atom types in the DPPC bilayer are shown in Figure A-1. From the atom types shown in Figure A-1, only four atom types are used in the dispersion model: H, C, C2 and O atom types. Contributions of the membrane atom types N and P in the total solute-membrane van der Waals term are negligible and therefore these membrane atom types are neglected in the dispersion model. In the xy plane in the Cartesian coordinate system, the distribution of the membrane atom types is assumed to be homogenous.

Therefore, the number densities would be constant in the x and y directions. However, the
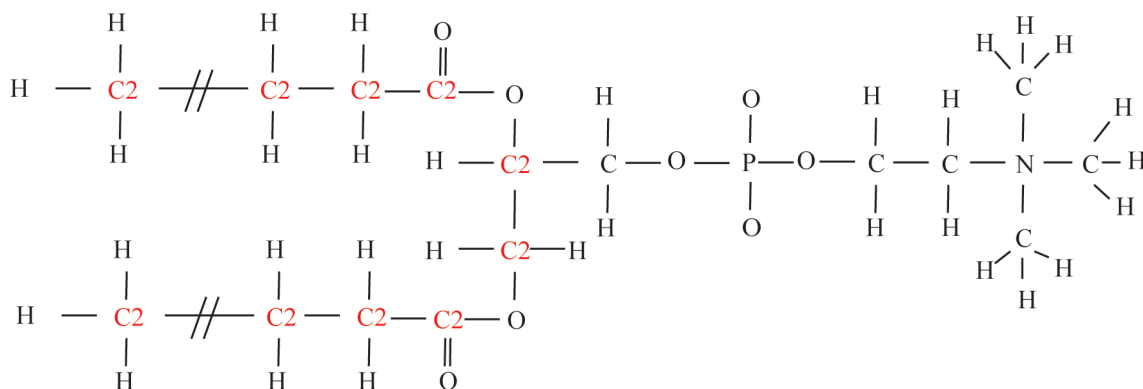
Figure A-1 Presentation of different atom types in a DPPC molecule. Four different atom types have been used in the HDGB-vdW method: O, H, C, C2. The carbon atoms are divided into to groups, one for the head group and one for the tail, shown with C2. The color representation is only to differentiate between the carbon atom types.

number density of different bilayer atom types varies along the membrane normal. In order to determine the number density profile for each atom type, explicit solvent simulation of DPPC bilayer was used. In addition to the explicit solvent simulation of the DPPC system, a second simulation with a membrane protein inserted in the lipid bilayer was studied to study the effect of the presence of a protein in the number density of the water and bilayer atom types. The details for both of the MD simulations have been reported in the method section. After the systems were equilibrated, each system was divided into different slabs with the thickness of 1 Å each. As mentioned above, in the presence of a peptide, the distribution of the bilayer atom types and water may change. Therefore, It is worthwhile to compare the density profiles for different atom types in the presence and absence of the peptide. Radial distribution functions have been used to determine the density profiles for different atom types. In the presence of the peptide, atoms from the surface of the peptide have been used as the probes to calculate the radial distribution

functions. In the system where no peptide is present, water molecules have been used as the probes. In the presence of the peptide, radial distribution function of different atom types will be underestimated because of the exclusion volume of the peptide. Therefore, it is necessary to scale the radial distribution functions before comparing them with the bilayer-only systems according to Equation A-12. As it will be discussed later, the lipid-only system has been used to scale the radial distribution functions of the atom types in the bilayer-peptide system.

$$g(r)_{bilayer-peptide} = \frac{N_{bilayer-only}}{N_{bilayer-peptide}} g(r)_{bilayer-only} \qquad \text{A-12}$$

In Equation A-12, g(r) is the radial distribution function for each water/bilayer atom type and N is the total number density of bilayer and water atom types.

For the bilayer only system, radial distribution functions of all the water/bilayer atom types were calculated. The results are presented in Figure A-2. Each panel in Figure A-2 corresponds to a different slab in the bilayer. In each panel, radial distribution function of six atom types are plotted, two for water and four for the bilayer. From each panel, the average radial distribution of each atom type in each slab is calculated and plotted against z, the distance of each slab from the membrane center, in Figure A-3. The total radial distribution of all the atom types has been calculated and plotted in Figure A-3 as well. The total distribution curve suggests that the total number density in the lipid bilayer is approximately constant along the membrane normal with the average value of 1.1 atoms/$Å^3$. This average value has been used to scale the density profiles generated from the bilayer-peptide system. First, the radial distribution functions of the four bilayer atom types plus water oxygen have been generated with the same

technique that mentioned before as shown in Figure A-4 (dispersion term for the hydrogen atoms
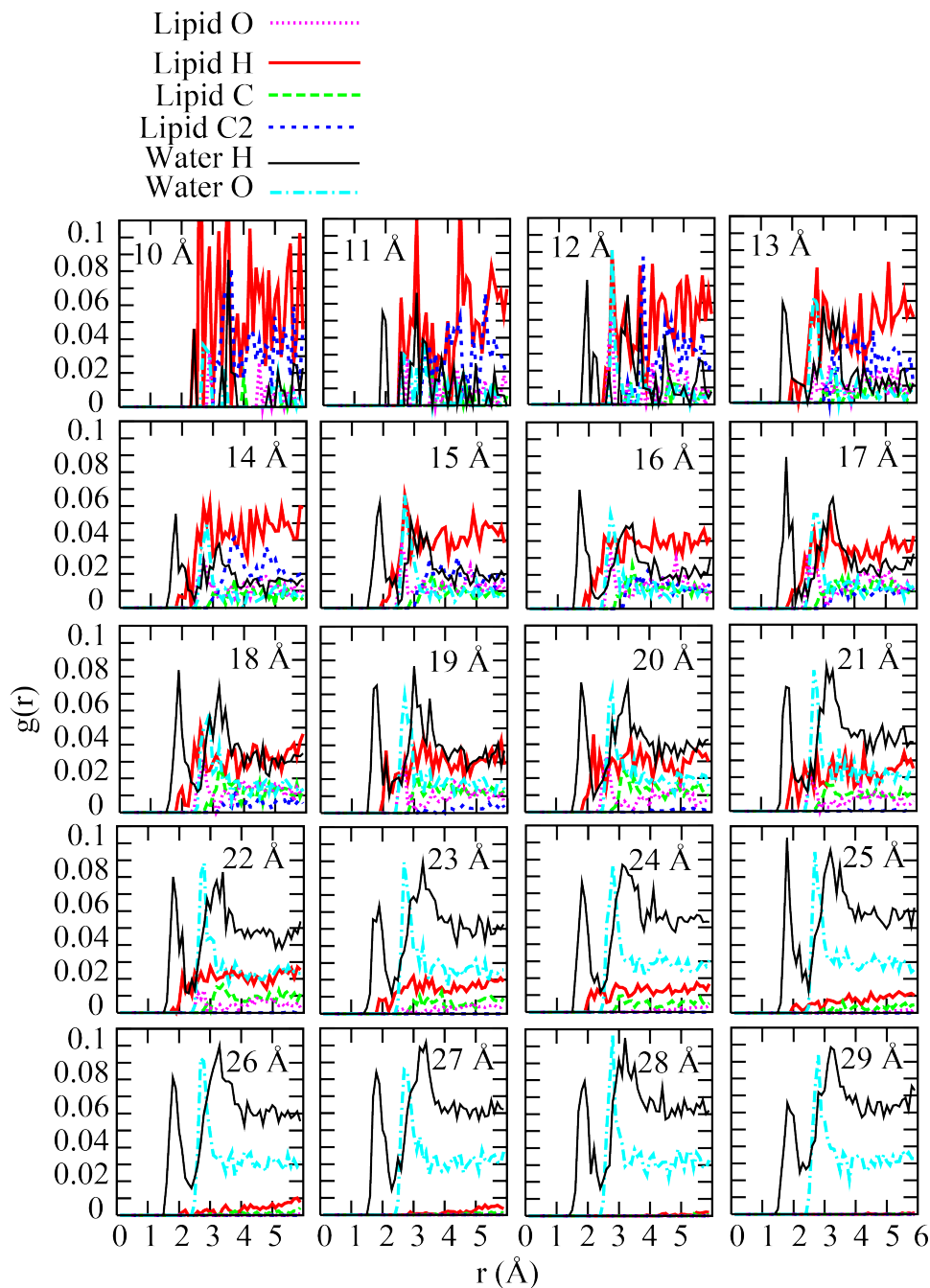
Figure A-2 Radial distribution functions for different slabs of the explicit DPPC system. Each panel presents the radial distribution functions calculated in one slab. The distance of each slab from the membrane center is mentioned on each pannel. At each slab, six radial distribution functions have been calculated, two for water and four for the bilayer.
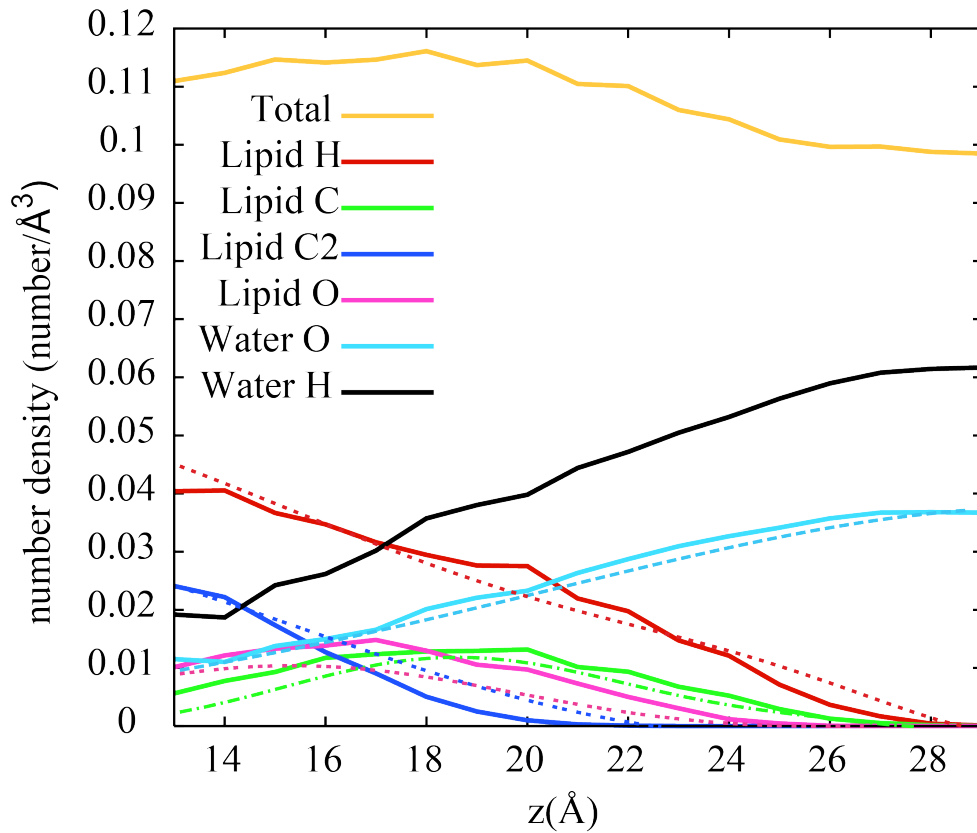
Figure A-3 Average number density of different atom types are plotted az a function of z, distance from the membrane center. Each data point is from averaging the radial distribution function from a pannel in Figure A-2. The same color coding has been used in this Figures A-3 and A-2. The solid lines are form the bilayer-only system while the dotted lines are from the bilayer-protein system calculations. The total number density has been calculated from the bilayer-only system and plottted in yellow.
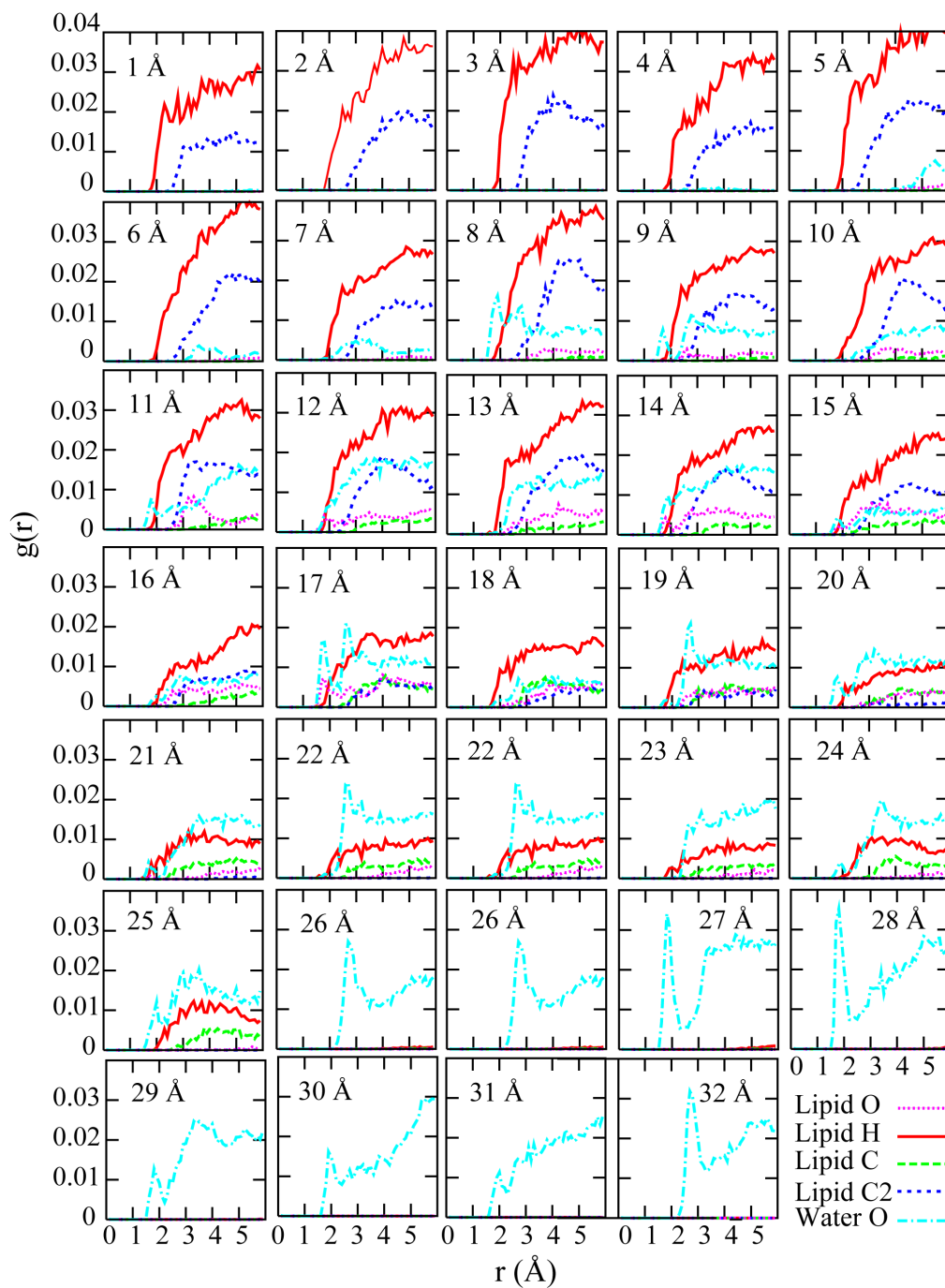
Figure A-4 Scaled radial distribution functions for the lipid atom types and water oxygen for different slabs of a system including DPPC bilayer and a membrane protein. Color coding is shown in the figure. Each panel describes the radial distribution functions for one specific slab with the distance from the center of membrane that is reported on the panel.
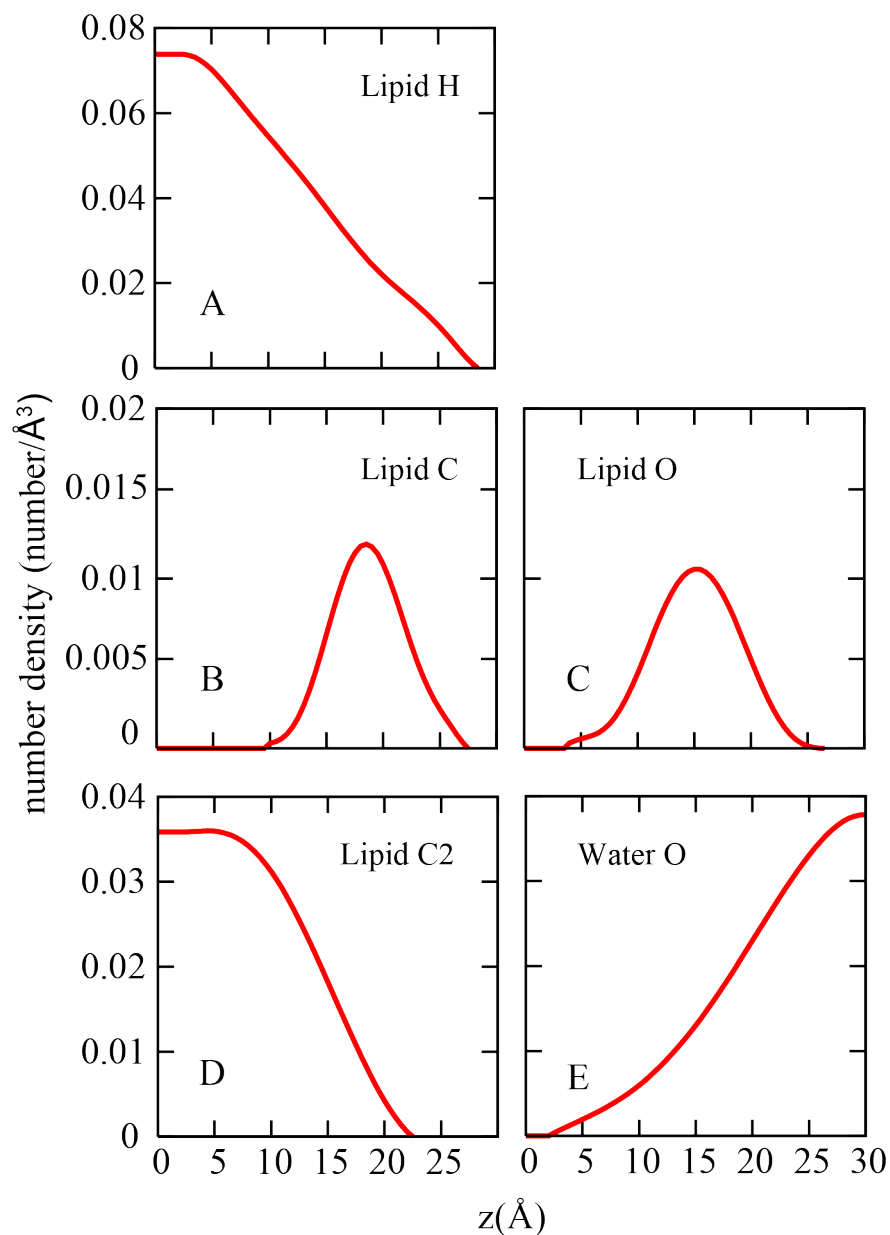
Figure A-5 Final density profiles for the four bilayer atom types and water. The density profiles are extracted from the scaled density profiles generated from the bilayer-protein system (for more details see the methods section).

of water are not considered here because Lennard-Jones interactions between water hydrogen atoms and other atoms are neglected in the CHARMM force field). Then the average radial

distribution for each atom type at each slab was calculated and plotted along z to generate a density profile. These profiles then scaled based on the average total number density value, 1.1 atom/ $\mathring{A}^3$ ($N_{bilayer-only}$ in Equation A-12), to generate the corrected density profiles (dotted lines in Figure A-3). In Figure A-3, the density profiles from the bilayer-peptide system are plotted only in the range from 13 to 29 $\mathring{A}$ to be comparable with the profiles generated from the bilayer-only system. In the bilayer-only system, no data is available for the z values less than 13 $\mathring{A}$ because there is no penetrated water for those slabs. Comparing the density profiles generated from the two systems (with and without the peptide) suggest that the distribution of the bilayer atom types and water is almost the same, at least in the regions that are more than 13 $\mathring{A}$ away from the membrane center. Finally, using Matlab, a polynomial (with the order between 5 and 7) was fitted to each density profile generated from the bilayer-peptide system. The final density distribution profiles are presented in Figure A-5.

So far, we have estimated the $\rho_j(z)$ values in Equation A-11. It is also necessary to determine σ_j and ε_j in Equation A-11 and a_j and R_j in Equation A-10 where j is the water/bilayer atom type index. σ_j and ε_j values are extracted from the CHARMM force field using:

$$\varepsilon_j = \sum_{i=1}^{N_j} \varepsilon_i / N_j \quad , \quad \sigma_j = \sum_{i=1}^{N_j} \sigma_i / N_j \qquad \text{A-13}$$

where N_j is the number of bilayer atoms that belong to the atom type j. The σ and ε values for the DPPC bilayer with the above mentioned atom types are reported in Table A-1. Next, we need to

determine the offset values, R_j in Equation A-10. The offset value determines how close a lipid

atom can get to the peptide. The offset value for the water molecules is set to the radius of a

water molecule, which is 1.4 Å. This offset value is independent of the position of the water

Lipid atom type	ε	σ
H	-0.0307	1.2200
C	-0.0680	2.0350
C2	-0.0570	2.0190
O	-0.1100	1.6750

Table A-1 Lennard-Jones parameters for different DPPC atom types. The values are calculated by averaging the Lennard-Jones parameters of the atoms that belong to each atom type in a DPPC molecule from the CHARMM force field.

molecule relative to the membrane center. In the case of lipid atoms, determination of the offset

values is more complicated because of the non-spherical shape of the lipid chains. We have used

the radial distribution functions in Figures A-2 and A-4 to estimate the offset values for the

bilayer atom types. The offset value for each atom type is chosen as the shortest distance from

the probe, where the radial distribution function is non-zero. The offset values that are extracted

from the data presented in Figures A-2 and A-4 are reported in Table A-2.

```
                      ┌─────────────────────────────────┐
                      │  Determine the bilayer atom types │◄──────────────────┐
                      └─────────────────────────────────┘                     │
                                      │                                        │
                                      ▼                                        │
                      ┌─────────────────────────────────┐                     │
                      │   Determine the density profiles  │                     │
                      └─────────────────────────────────┘                     │
                                      │                                        │
                                      ▼                                        │
                      ┌─────────────────────┐                                 │
                      │  Change the ε profile │◄──────────────────────────────┤
                      └─────────────────────┘                                 │
                                      │                                        │
                                      ▼                                        │
                      ┌─────────────────────┐                                 │
                      │  Change the "a" values │◄─────────────────────────────┤
                      └─────────────────────┘                                 │
```
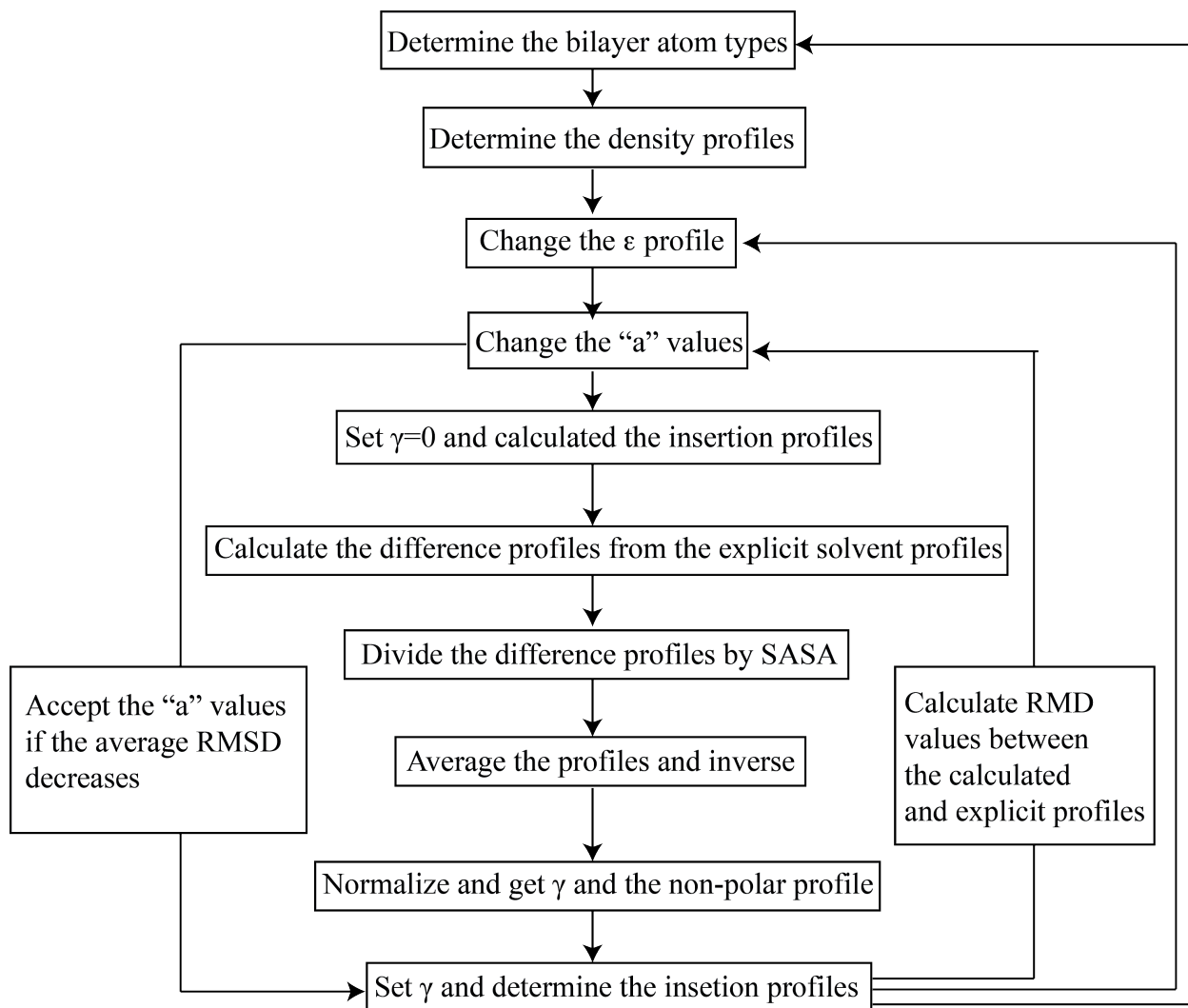
Figure A-6 The protocol for parametrization of the modified HDGB method.

A type	Offset (Å)
Lipid H	1.5
Lipid C	2.2
Lipid C2	2.2
Lipid O	1.5
Water O	1.4

Table A-2 Offset values (a_j) for bilayer atom types and water oxygen generated from the radial distribution functions as described in the methods section.

Finally, the fitting parameter, a_j in Equation A-6 needs to be determined. In order to find

the a_j values, the protocol presented in Figure A-6 has been used. The solvation free energies of

the amino acid side chain analogs were calculated at different z values with the new method and

fitted to the solvation free energy profiles of the amino acid side chain analogs from the explicit

solvent simulation. The fitting protocol is as follows: First, all of the a_j values were set to 1 and γ

was set to zero and insertion profiles for the amino acid side chain analogs were calculated. In

the HDGB formalism with the dispersion term modification, the solvation free energy would be

the sum of the electrostatic term and the non-polar term from Equation A-2. Therefore, by setting

γ=0, the solvation free energy would be calculated from the GB term and the dispersion term

with a missing cavity term. Then, the calculated insertion profiles were subtracted from the

explicit solvent profiles. These difference profiles indicate the missing contribution of the cavity

term in the total solvation free energy. The cavity term in the HDGB method is then calculated

through Equation A-5 (for more details see the introduction section). By dividing each difference

profile by the *SASA* value of that side chain analog, the missing term should then be the profile

for the cavity term, scaled by the γ value. These difference profiles that are scaled with the

1/*SASA* values have been plotted in Figure A-7. These profiles were then averaged for all of the

analogs. At each point on the average profile, the free energy value of that data point was

subtracted from the free energy value at z=0 value (see Figure A-8). The result would be referred

to as "the inversed average profile". The γ value was determined from the inversed average

profile at z=30 Å, where the profile reaches the plateau. The inversed average profile was then

normalized to obtain the cavity profile ($S(z)$ in Equation A-9). Finally the γ value was set to the
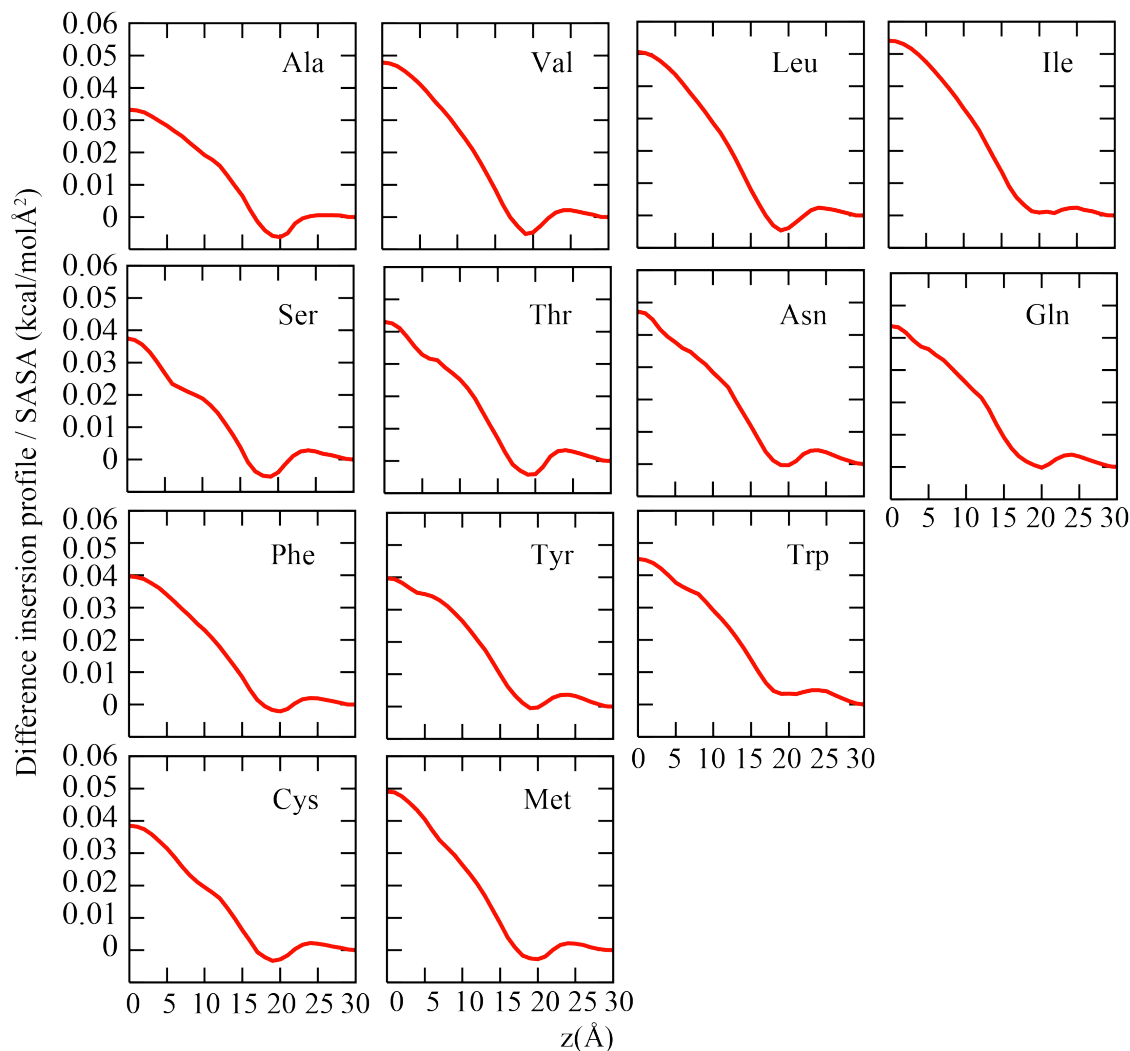
Figure A-7 The diference profiles generated from subtracting the insertion profiles in the HDGB-vdW method while *γ* is set to zero, from the reference insertion profiles from the explicit solvent simulations. The profiles are scaled with the inverse of the SASA value for each amino acid side chain analog. The reference profiles are calculated from the explicit solvent simulations [183, 293] with the OPLS charges [35, 294].
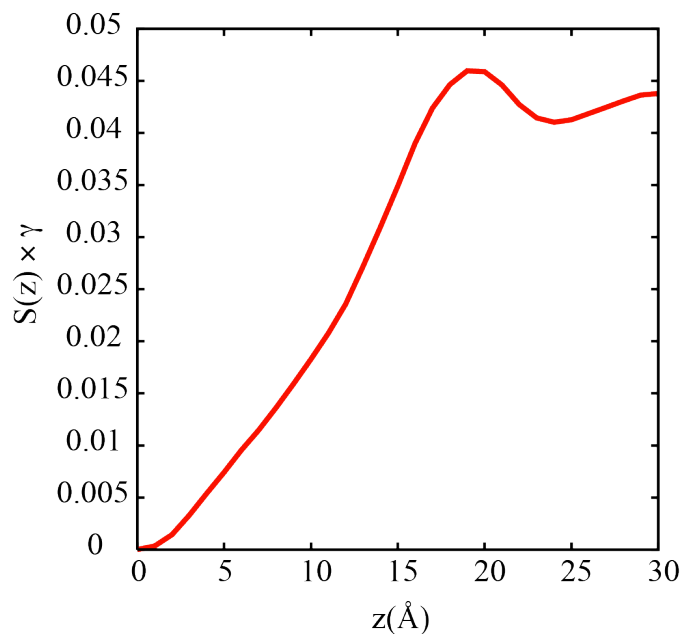
Figure A-8 The scaled non-polar profile generated from averaging the difference profiles and then reverse the average profile. The scaling factor is the γ value. The non-scaled profile reached to one in the bulk water, when z is larger than 30 Å.
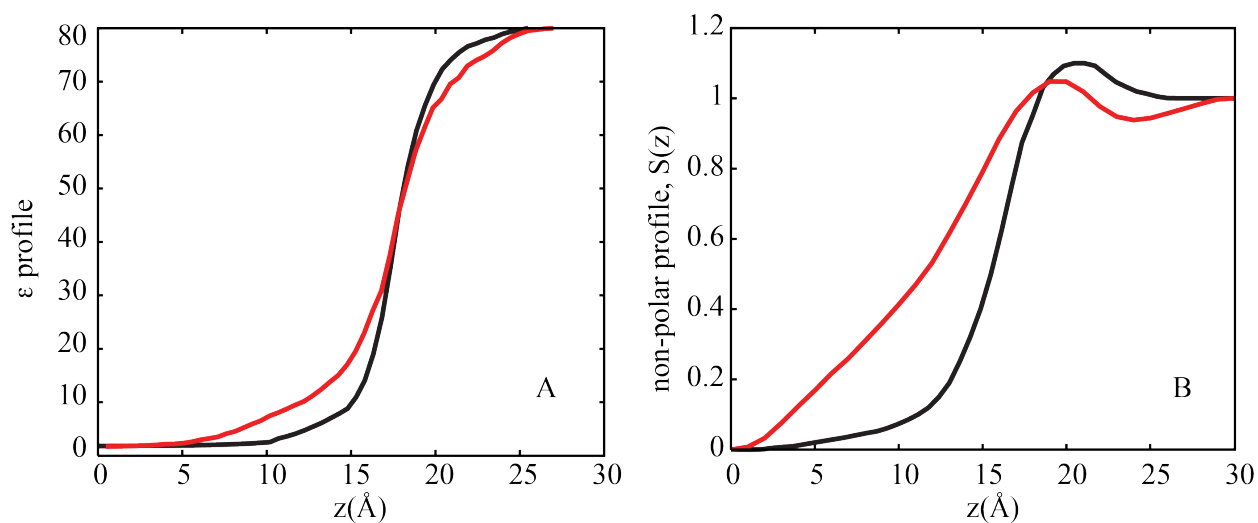
Figure A-9 The ε and non-polar profiles used in the HDGB methods. The original profiles are shown in black. The red curves describe the modified profiles after adding the dispersion term to the HDGB method and parametrization of the model.

determined value in the previous step and using the new cavity profile, the total solvation free energy of the amino acid side chain analogs were determined. The whole process was repeated until the desired fit was obtained. After each setting of the cavity profile, at least 100 steps of the downhill Monte Carlo (MC) were carried out for a better fit. The details of the MC technique are mentioned in section A-2-2. The final a_j values are presented in Table A-3. In Figure A-9, the final ε and non-polar profiles for the modified method (called HDGB-vdW) are presented and compared with the HDGB method with no implemented dispersion term. The γ value extracted from the last non-polar profile fitting was 0.044 cal/molÅ^2. However, during the MC steps, this value was reduced to 0.038 cal/molÅ^2.

Peptide atom type	a_j value
CT3 carbon	0.5
CT2 carbon	0.9
CT1 carbon	1.1
Aromatic carbon on the phenyl ring	0.4
Aromatic carbons in His	0.8
Other Aromatic carbons	0.4
Carbon in the C=O group	0.8
Carbon of methyl group (Ala analog)	1.4
Oxygen of the hydroxyl group on Tyr	1.45
Other hydroxyl oxygens	1.5
Nitrogen in Trp	0.3
Other nitrogens	0.45
Hydrogen in the hydroxyl group	1.5
Aromatic hydrogens	0.3
Hydrogen in the H-S bond	1.0
Hydrogen of methyl group (Ala analog)	1.35
Other hydrogens	0.95
Sulfur in Met	0.35
Sulfur in Cys	0.85
Phosphorus	0.75

Table A-3 The final a_j values determined from the parameterization of the HDGB-vdW method with the protocol described in Figure A-6.

A-2-2 MD Simulations

The dispersion term has been implemented in the CHARMM force field [33] version c36a4. CHARMM program [177] has been used in all of this work. The HDGB formalism [84] has been modified to include the new dispersion term. All the HDGB parameters are set to the values mentioned in the original paper [84] other than the γ value, which is determined during the fitting. Insertion profiles for the amino acid analogs have been calculated in this work. Amino acid analogs were placed at the center of the membrane initially. For each free energy calculation, the side chain analog was rotated around the x- and y-axis. The Boltzmann average of the solvation free energy was calculated for all the rotated analogs at a specific distance from the center of the membrane. The analogs were rotated 15° at a time. The average solvation free energies were calculated at the 1 Å intervals along the membrane normal up to the z value equal to 30 Å. Reference insertion profiles are from the explicit solvent simulations [183, 293] with the OPLS charges [35, 294].

In the fitting protocol, a downhill Monte Carlo (MC) sampling technique has been used for the γ and a_j values. At each MC step, γ or one of the a_j values were changed according to Equation A-14:

$$a_{j,new} = a_{j,old} \pm 0.05 \quad , \quad 0.3 \le a_{j,new} \le 2.5$$
$$\gamma_{new} = \gamma_{old} \pm 0.001 \quad , \quad 0.035 \le \gamma_{new} \le 0.1$$

A-14

After each MC step, RMSD between the reference and calculated profiles were determined and averaged. The MC step was accepted with the following condition:

$$\Delta = AveRMSD_{new} - AveRMSD_{old}$$

$$T = \begin{cases} 1 & if \; \Delta < 0 \\ 0 & if \; \Delta \geq 0 \end{cases}$$

A-15

where *AveRMSD* is the average RMSD values for all the insertion profiles.

The explicit solvent simulations of a DPPC bilayer with and without a membrane protein have been conducted with the CHARMM force field [33] version c36a1. The initial structures were built and equilibrated using the CHARMM-GUI website [295] which is a web-based graphical interface for the CHARMM users. The membrane-only system was constructed with 54 DPPC lipid molecules solvated with 2372 water molecules. The membrane-protein system was constructed with 52 DPPC lipid molecules in each leaflet and the total number of 6185 water molecules. Periodic boundary condition was used for both systems. The box dimensions were 41.2×41.2×80 Å and 64.5×64.5×83.6 Å for the membrane-only and membrane-protein systems respectively. The membrane protein which is used in the membrane-peptide system is DsbB (PDB code 2ZUQ) [296]. The systems were minimized for 1500 steps with the steepest descent and adapted-bases Newton−Raphson methods each. In the first 25000 MD steps of the equilibration runs, the time step of 1 fs was used to gain more stability for the constructed systems. The time step was then increased to 2 fs for the rest of the simulation. Temperature was set to 323.15 K and maintained constant with the use of Langevin dynamics [174] with the friction coefficient of 3 ps^{-1} for all the heavy atoms. SHAKE algorithm [297] was used to constrain the bonds that involve hydrogen. The water molecules were initially restrained with a harmonic potential to stay away from the hydrophobic core with a force constant of 2.5 kcal/mol. The lipids were also harmonically restrained with a 2.5 kcal/mol force constant. The restrains were gradually removed from the systems during four cycles of equilibration. The first two cycles consist of 25000 MD steps each (5 ns per cycle). The last two cycles include 50000 MD

steps (10 ns each). The equilibration protocol was obtained form the CHARMM-GUI website [295]. The production trajectory consists of 5 ns of the MD simulation and the frames were written into a dcd file after each 500 steps.

For a model dimer peptide called pVNVV (with the monomer sequence of LLLLV LLLLL LNLLL LLLVL LLLLL VL) the dissociation free energy profile has been calculated through the umbrella sampling simulation of the dimeric peptide with the HDGB method. The peptide was blocked with the acetyl and amine groups in the termini. For the initial structure of the dimer peptide, its crystal structure with the PDB code of 2ZTA was used [268]. Each monomer consists of a helical structure with a helix-helix distance of 8.7 Å. Umbrella sampling simulation of the dimer was performed with the Multiscale Modeling Tools for Structural Biology (MMTSB) tool set [178] to facilitate using the CHARMM program [177]. CHARMM force field [33] version c36a1 was used for the simulation. All the HDGB parameters were set to the values reported in the original HDGB paper [84]. γ was set to 0.015 cal/molÅ^2 for the HDGB and 0.038 cal/molÅ^2 for the HDGB-vdW method. Langevin dynamics [174] was used to keep the temperature at 298K with the friction coefficient of 10 ps^{-1} for all the heavy atoms. The cut-off value of 18 Å was used for the non-bonded interactions. SHAKE algorithm [176] was used to restrain all the bonds that involved hydrogen. A helix-helix distance restrain potential was used with the force constant of 5 kcal/molÅ^2 to restrain the dimer [298]. The initial structure was first equilibrated for 5 ns with the restrain potential to keep the helix-helix distance at 9 Å. The helix distance was then reduced to 7 Å or increased to 17 Å with the 0.5 Å increments at each 50 ps to generate the initial structure for each window. A 10 ns implicit solvent simulation was performed for each window. A force constant of 2.5 kcal/molÅ^2 was used to restrain the distance between

the two monomers. The last 5 ns simulation of each window was used to generate the dissociation free energy profile.
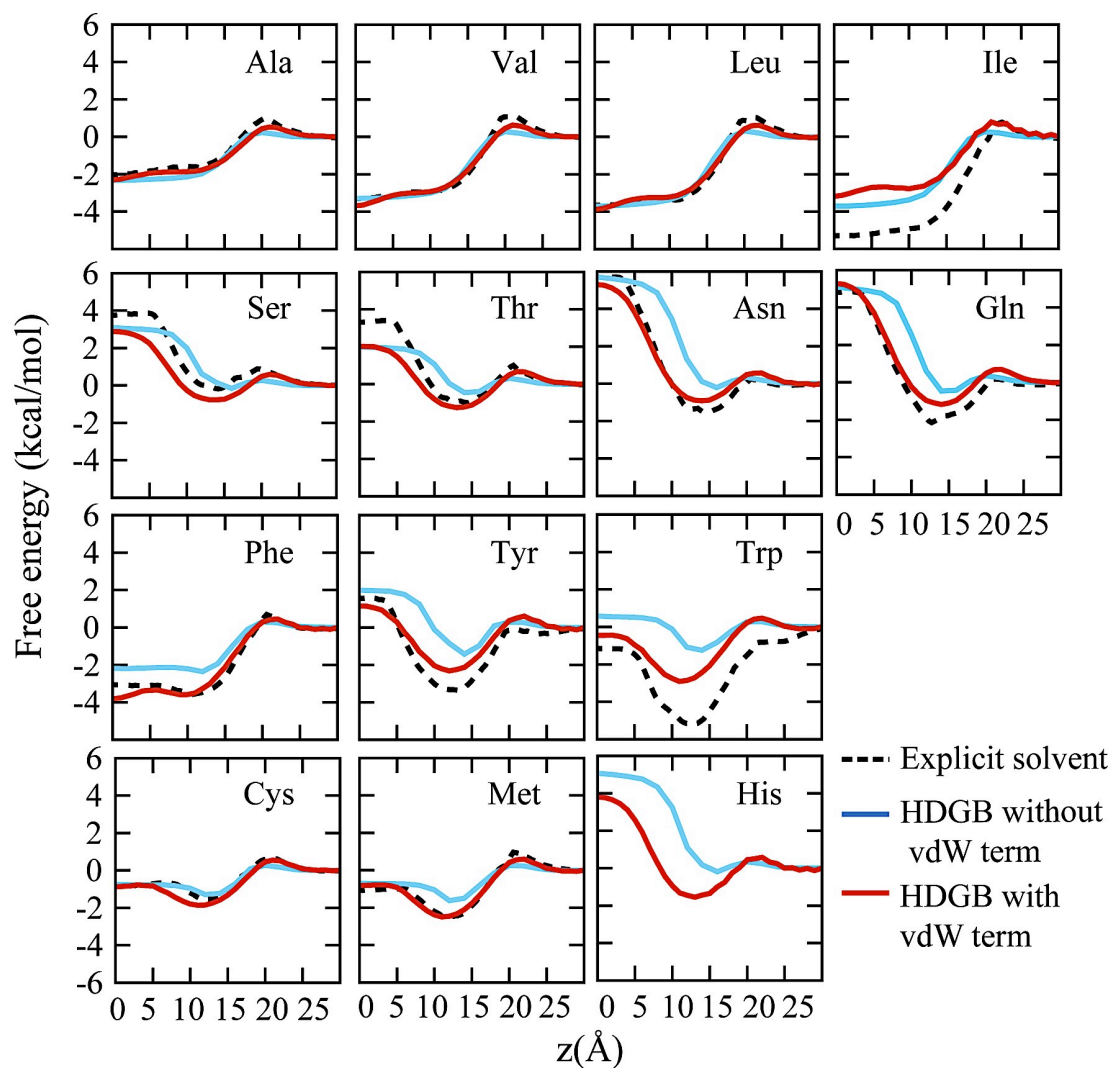
Figure A-10 The final amino acid anoalog insertion profiles calculated with the HDGB-vdW method. The profiles generated with the explicit solvent simulations [183, 293] and the HDGB method are also presented for comparision with the OPLS charges [35, 294].

A-3 Methods validation and discussion

For the validation of the new method, HDGB-vdW, the final amino acid analog insertion profiles have been plotted and compared with the profiles generated from the HDGB method with no dispersion term and the explicit solvent simulation in Figure A-10. The insertion profiles calculated from the HDGB method (with no vdW term) are plotted as well for comparison. In general, the fitted profiles are in very good agreement with the explicit solvent profiles. In some cases, the generated profiles have been improved from the HDGB method. The most improvements are for the Tyr and Trp analogs, where the deep minimum at the hydrophobic core region is better described with the HDGB-vdW method. Gln, Asn, Phe and Met profiles have also been improved.
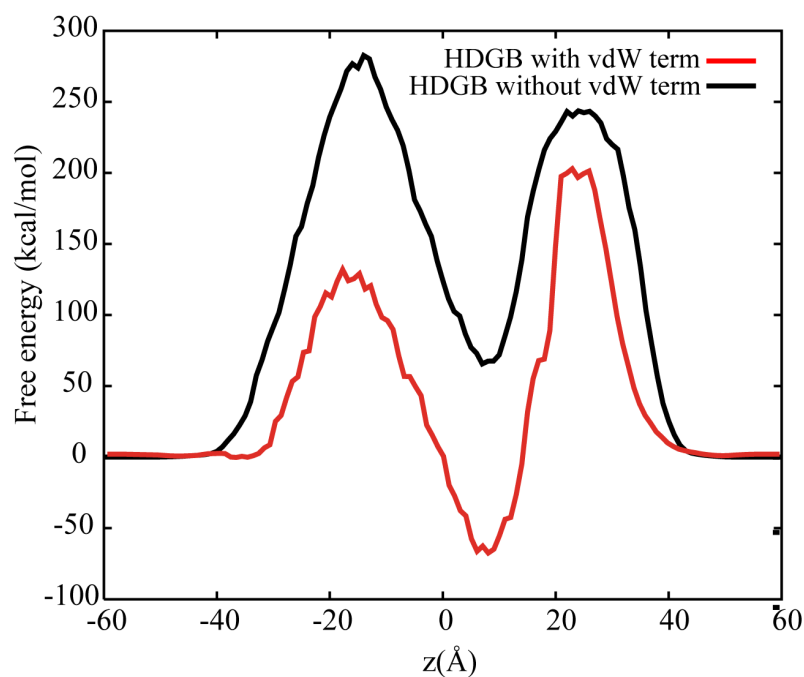
Figure A-11 Insertion profile for the DsbB membrane protein generated with the HDGB (red curve) and HDGB-vdW (black curve) methods. The insertion profiles are generated with the protocol described in the methods section. Free energy of solvation is calculated at differentpositions of the center of mass of the protein relative to the membrane center (z).

We have also looked at the insertion profile of a membrane protein (PDB code 2ZUQ), which we have chosen from the membrane protein data set of our collaborators in the Y. Sugita group. The insertion profiles for the membrane protein, calculated with the HDGB and HDGB-vdW methods are shown in Figure A-11.The HDGB method, with no dispersion term, cannot capture the balance between the various contributions of the solvation free energy in the membrane and water. This deficiency leads to the favorability of the membrane protein in the water relative to the membrane. As it is shown in Figure A-11, by introducing the vdW dispersion term in the HDGB method, the solvation free energy of the membrane protein is correctly estimated to be more favorable in the membrane. Addition of the new term does not seem to change the most favorable insertion depth for the membrane protein relative to the membrane center, which was determined to be almost 8 Å from the membrane center in the HDGB method. This suggests that in case of studying one single protein in the membrane, the HDGB method probably would give a good estimate for the energetics of the system. However, if the protein leaves the membrane during the simulation, the HDGB method with no vdW term is not suitable for simulating such a system. More over, because the attractive vdW interactions between the membrane and protein are missing in the HDGB method, any protein-protein interaction would be overestimated in the membrane. This overestimation was tested with calculating the dissociation profile of the pVNVV dimer with the HDGB method with no dispersion term and comparing it with the explicit solvent simulations. The generated profile is
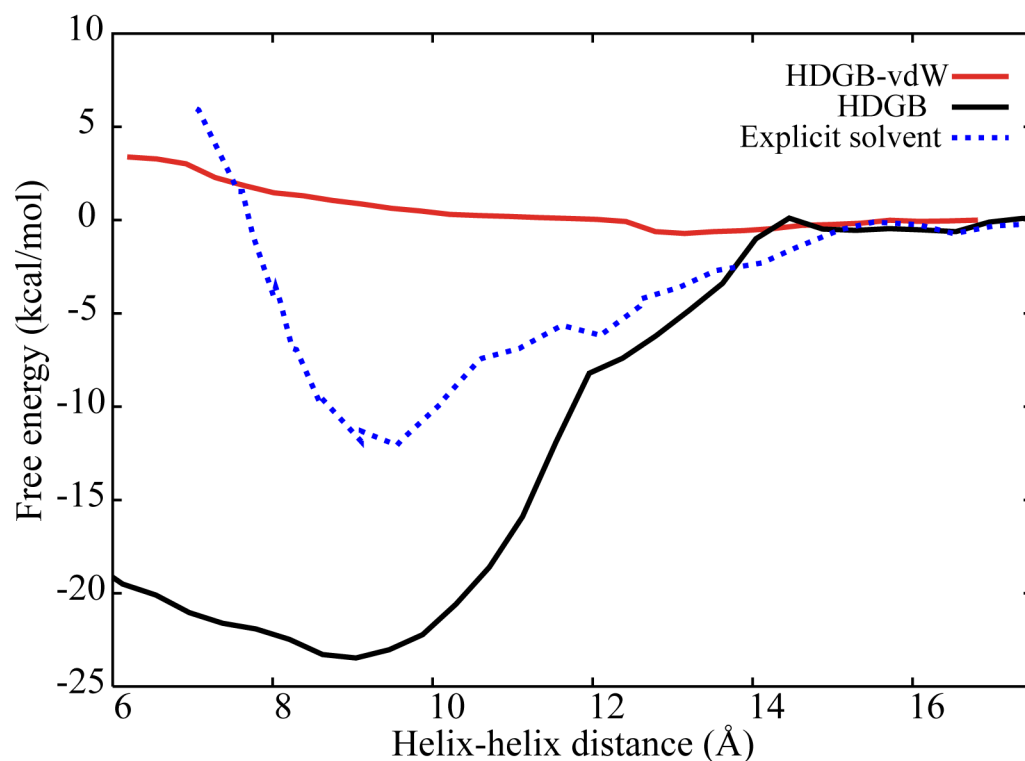
Figure A-12 Dissociation free energy profile for pVNVV peptide dimer calculated with the HDGB method with no dispersion term. Explicit solvent curve is presented for comparison as well [299].

shown in Figure A-12. The explicit solvent simulation estimates the dissociation free energy of the pVNVV dimer [299] to be almost 12 kcal/mol . The HDGB method overestimates this free energy by more than 10 kcal/mol. This overestimation could result in the aggregation or overestimation of the association free energy between oligomers. The HDGB-vdW method however, is not successful to capture the protein-protein interaction. It is possible that by adding the attractive dispersion term, the total interaction between the solute and solvent is not balanced. One suggestion for fixing this problem could be a positive shift of the non-polar profile. In the

current version of the non-polar profile, the cavity term shift to zero towards the center of the membrane. This means that close to the center of the membrane, the non-polar contribution of the solvation free energy merely comes from the attractive dispersion term. By shifting the non-polar profile, the non-polar contribution of the solvation free energy would be more balanced. It should be mention that shifting the non-polar profile will have no effect on the insertions profile for any peptide or amino acid side chain analog because these profiles are based on the relative free energy values. Further studies needs to be done to study the effect of shifting the non-polar profile on the protein-protein interactions inside the membrane.

A-4 Conclusions

This work presents the preliminary results for the modified HDGB method. The dispersion term has been modified for the DPPC bilayer system and is implemented in the HDGB method. The un-modified HDGB method cannot maintain the balance between the non-polar and polar terms in the solvation free energy because the non-polar term is only described with a model. The solvation free energies of the membrane proteins are therefore estimated to be more favorable in the water. By modeling the dispersion term in HDGB method, the favorability of the membrane proteins are corrected in the membrane relative to the aqueous solvent. The amino acid insertion profiles, which were used to parameterize the method, are also improved with the modified HDGB method. This modified method is proposed to better describe a membrane environment. More work is necessary to further optimize and study more applications of this technique.

REFERENCES

REFERENCES

1. Alder, B.J. and T.E. Wainwright, *Studies in molecular dynamics .1. General method.* Journal of Chemical Physics, 1959. **31**(2): p. 459-466.

2. McCammon, J.A., B.R. Gelin, and M. Karplus, *Dynamics of folded proteins.* Nature, 1977. **267**(5612): p. 585-590.

3. Holbrook, S.R., et al., *Crystal structure of an RNA double helix incorporating a track of non-Watson-Crick base-pairs.* Nature, 1991. **353**(6344): p. 579-581.

4. Mesecar, A.D., B.L. Stoddard, and D.E. Koshland, *Orbital steering in the catalytic power of enzymes: Small structural changes with large catalytic consequences.* Science, 1997. **277**(5323): p. 202-206.

5. Feig, M. and Z.F. Burton, *RNA polymerase II with open and closed trigger loops: Active site dynamics and nucleic acid translocation.* Biophysical Journal, 2010. **99**(8): p. 2577-2586.

6. Ramachandran, A., et al., *Coarse-grained molecular dynamics simulation of DNA trans location in chemically modified nanopores.* Journal of Physical Chemistry B, 2011. **115**(19): p. 6138-6148.

7. Pieniazek, S.N., M.M. Hingorani, and D.L. Beveridge, *Dynamical allosterism in the mechanism of action of DNA mismatch repair protein muts.* Biophysical Journal, 2011. **101**(7): p. 1730-9.

8. Perez, A., F.J. Luque, and M. Orozco, *Dynamics of B-DNA on the microsecond time scale.* Journal of the American Chemical Society, 2007. **129**(47): p. 14739-14745.

9. Liwo, A., et al., *Implementation of molecular dynamics and its extensions with the coarse grained UNRES force field on massively parallel systems: Toward millisecond scale simulations of protein structure, dynamics, and thermodynamics.* Journal of Chemical Theory and Computation, 2010. **6**(3): p. 890-909.

10.   Shaw, D.E., et al., *Atomic level characterization of the structural dynamics of proteins.* Science, 2010. **330**(6002): p. 341-346.

11.   Wuthrich, K., et al., *NMR studies of the hydration of biological macromolecules.* Faraday Discussions, 1996. **103**: p. 245-253.

12.   Mendelsohn, A.R. and R. Brent, *Protein biochemistry - Protein interaction methods - Toward an endgame.* Science, 1999. **284**(5422): p. 1948-1950.

13.   Wagner, G., S.G. Hyberts, and T.F. Havel, *NMR structure determination in solution. A critique and comparison with X-ray crystallography.* Annual Review of Biophysics and Biomolecular Structure, 1992. **21**: p. 167-198.

14.   Feig, M., A.D. MacKerell, and C.L. Brooks, *Force field influence on the observation of pi-helical protein structures in molecular dynamics simulations.* Journal of Physical Chemistry B, 2003. **107**(12): p. 2831-2836.

15.   Thompson, E.J., et al., *Evaluating molecular mechanical potentials for helical peptides and proteins.* Plos One, 2010. **5**(3).

16.   Shaw, D.E., et al., *Anton, a special-purpose machine for molecular dynamics simulation.* Communications of the Acm, 2008. **51**(7): p. 91-97.

17.   Sanbonmatsu, K.Y. and C.S. Tung, *High performance computing in biology: Multimillion atom simulations of nanoscale systems.* Journal of Structural Biology, 2007. **157**(3): p. 470-480.

18.   Kumar, S., et al., *Scalable molecular dynamics with NAMD on the IBM Blue Gene/L system.* Ibm Journal of Research and Development, 2008. **52**(1-2): p. 177-188.

19.   Freddolino, P.L., et al., *Molecular dynamics simulations of the complete satellite tobacco mosaic virus.* Structure, 2006. **14**(3): p. 437-449.

20.   Feig, M. and C.L. Brooks, *Recent advances in the development and application of implicit solvent models in biomolecule simulations.* Current Opinion in Structural Biology, 2004. **14**(2): p. 217-224.

21.   Roux, B. and T. Simonson, *Implicit solvent models.* Biophysical Chemistry, 1999. **78**(1-2): p. 1-20.

22.   Crawford, G.E. and J.C. Earnshaw, *Viscoelastic relaxation of bilayer lipid membranes. Frequency-dependent tension and membrane viscosity.* Biophysical Journal, 1987. **52**(1): p. 87-94.

23.   Earnshaw, J.C. and G.E. Crawford, *Viscoelastic relaxation of bilayer lipid-membranes .2. Temperature-dependence of relaxation time.* Biophysical Journal, 1989. **55**(5): p. 1017-1021.

24.   Harvey, S.C., *Treatment of electrostatic effects in macromolecular modeling.* Proteins-Structure Function and Genetics, 1989. **5**(1): p. 78-92.

25.   Chen, S.-H., et al., *The violation of the Stokes-Einstein relation in supercooled water.* Proceedings of the National Academy of Sciences of the United States of America, 2006. **103**(35): p. 12974-12978.

26.   Feig, M., *Implicit membrane models for membrane protein simulation*, in *Methods in molecular biology:molecular modeling of proteins*, K. A., Editor 2008, Humana Press.

27.   de Pablo, J.J., *Coarse-grained simulations of macromolecules: From DNA to nanocomposites*, in *Annual Review of Physical Chemistry, Vol 62*, S.R.C.P.S.G.J.T.J.M.A. Leone, Editor 2011. p. 555-574.

28.   Kamerlin, S.C.L., et al., *Coarse-grained (multiscale) simulations in studies of biophysical and chemical systems*, in *Annual Review of Physical Chemistry, Vol 62*, S.R.C.P.S.G.J.T.J.M.A. Leone, Editor 2011. p. 41-64.

29.   Spiriti, J., H. Kamberaj, and A. Van Der Vaart, *Development and application of enhanced sampling techniques to simulate the long-time scale dynamics of biomolecular systems.* International Journal of Quantum Chemistry, 2012. **112**(1): p. 33-43.

30.   Sugita, Y. and Y. Okamoto, *Replica-exchange molecular dynamics method for protein folding.* Chemical Physics Letters, 1999. **314**(1-2): p. 141-151.

31.   Schlick T., *Molecular modeling and simulation*2002: Springer.

32.   MacKerell, A.D., et al., *All-atom empirical potential for molecular modeling and dynamics studies of proteins.* Journal of Physical Chemistry B, 1998. **102**(18): p. 3586-3616.

33.   Mackerell, A.D., J. Wiorkiewiczkuczera, and M. Karplus, *An all-atom empirical energy function for the simulation of nucleic-acids.* Journal of the American Chemical Society, 1995. **117**(48): p. 11946-11975.

34.   Cornell, W.D., et al., *A second generation force field for the simulation of proteins, nucleic acids, and organic molecules (vol 117, pg 5179, 1995).* Journal of the American Chemical Society, 1996. **118**(9): p. 2309-2309.

35.   Jorgensen, W.L., D.S. Maxwell, and J. TiradoRives, *Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids.* Journal of the American Chemical Society, 1996. **118**(45): p. 11225-11236.

36.   Schuler, L.D., X. Daura, and W.F. Van Gunsteren, *An improved GROMOS96 force field for aliphatic hydrocarbons in the condensed phase.* Journal of Computational Chemistry, 2001. **22**(11): p. 1205-1218.

37.   Scott, W.R.P., et al., *The GROMOS biomolecular simulation program package.* Journal of Physical Chemistry A, 1999. **103**(19): p. 3596-3607.

38.   MacKerell, A.D., M. Feig, and C.L. Brooks, *Improved treatment of the protein backbone in empirical force fields.* Journal of the American Chemical Society, 2004. **126**(3): p. 698-699.

39.   Affleck, R., C.A. Haynes, and D.S. Clark, *Solvent dielectric effects on protein dynamics.* Proceedings of the National Academy of Sciences of the United States of America, 1992. **89**(11): p. 5167-5170.

40.   Jaaskelainen, S., et al., *Identifying key electrostatic interactions in Rhizomucor miehei lipase: the influence of solvent dielectric.* Theoretical Chemistry Accounts, 1999. **101**(1-3): p. 175-179.

41.   Fraternali, F. and W.F. vanGunsteren, *An efficient mean solvation force model for use in molecular dynamics simulations of proteins in aqueous solution.* Journal of Molecular Biology, 1996. **256**(5): p. 939-948.

42.   Feig, M., et al., *Performance comparison of generalized born and Poisson methods in the calculation of electrostatic solvation energies for protein structures.* Journal of Computational Chemistry, 2004. **25**(2): p. 265-284.

43.   Krol, M., *Comparison of various implicit solvent models in molecular dynamics simulations of immunoglobulin G light chain dimer.* Journal of Computational Chemistry, 2003. **24**(5): p. 531-546.

44.   Wang, J., et al., *Poisson-Boltzmann solvents in molecular dynamics simulations.* Communications in Computational Physics, 2008. **3**(5): p. 1010-1031.

45.   Warshel, A. and M. Levitt, *Theoretical studies of enzymic reactions. Dielectric, electrostatic and steric stabilization of carbonium ion in reaction of lysozyme.* Journal of Molecular Biology, 1976. **103**(2): p. 227-249.

46.   Pickersgill, R.W., *A rapid method of calculating charge charge interaction energies in proteins.* Protein Engineering, 1988. **2**(3): p. 247-248.

47.   Gelin, B.R. and M. Karplus, *Side-chain torsional potentials: Effect of dipeptide, protein, and solvent environment.* Biochemistry, 1979. **18**(7): p. 1256-1268.

48.   Hassan, S.A., F. Guarnieri, and E.L. Mehler, *A general treatment of solvent effects based on screened Coulomb potentials.* Journal of Physical Chemistry B, 2000. **104**(27): p. 6478-6489.

49.   Warshel, A., *Conversion of light energy to electrostatic energy in the proton pump of halobacterium-halobium.* Photochemistry and Photobiology, 1979. **30**(2): p. 285-290.

50.   Lazaridis, T. and M. Karplus, *Effective energy function for proteins in solution.* Proteins-Structure Function and Genetics, 1999. **35**(2): p. 133-152.

51.   Mallik, B., A. Masunov, and T. Lazaridis, *Distance and exposure dependent effective dielectric function.* Journal of Computational Chemistry, 2002. **23**(11): p. 1090-1099.

52.   Hassan, S.A. and E.L. Mehler, *A critical analysis of continuum electrostatics: the screened Coulomb potential-implicit solvent model and the study of the alanine dipeptide and discrimination of misfolded structures of proteins.* Proteins-Structure Function and Genetics, 2002. **47**(1): p. 45-61.

53. Wesson, L. and D. Eisenberg, *Atomic solvation parameters applied to molecular dynamics of proteins in solution.* Protein Science, 1992. **1**(2): p. 227-235.

54. Eisenberg, D. and A.D. McLachlan, *Solvation energy in protein folding and binding.* Nature, 1986. **319**(6050): p. 199-203.

55. Ooi, T., et al., *Accessible surface areas as a measure of the thermodynamic parameters of hydration of peptides.* Proceedings of the National Academy of Sciences of the United States of America, 1987. **84**(10): p. 3086-3090.

56. Wolfenden, R., et al., *Affinities of amino acid side chains for solvent water.* Biochemistry, 1981. **20**(4): p. 849-855.

57. Schiffer, C.A., et al., *Inclusion of solvation free energy with molecular mechanics energy. Alanyl dipeptide as a test case.* Protein Science, 1992. **1**(3): p. 396-400.

58. Hasel, W., T.H. Hendrickson, and W.C. Still, *A rapid approximation to the solvent accessible surface areas of atoms.* Tetrahedron Computer Methodology, 1988. **1**(2): p. 103-116.

59. Ferrara, P., J. Apostolakis, and A. Caflisch, *Evaluation of a fast implicit solvent model for molecular dynamics simulations.* Proteins-Structure Function and Genetics, 2002. **46**(1): p. 24-33.

60. Guvench, O. and C.L. Brooks, *Efficient approximate all-atom solvent accessible surface area method parameterized for folded and denatured protein conformations.* Journal of Computational Chemistry, 2004. **25**(8): p. 1005-1014.

61. Gibson, K.D. and H.A. Scheraga, *Minimization of polypeptide energy .I. Preliminary structures of bovine pancreatic ribonuclease s-peptide.* Proceedings of the National Academy of Sciences of the United States of America, 1967. **58**(2): p. 420-&.

62. Hopfinge.Aj, *Polymer-solvent interactions for homopolypeptides in aqueous solution.* Macromolecules, 1971. **4**(6): p. 731-&.

63. Kang, Y.K., G. Nemethy, and H.A. Scheraga, *Free energies of hydration of solute molecules .1. Improvement of the hydration shell model by exact computations of overlapping volumes.* Journal of Physical Chemistry, 1987. **91**(15): p. 4105-4109.

64.   Nicholls, A. and B. Honig, *A rapid finite difference algorithm utilizing successive over-relaxation to solve the Poisson-Boltzmann equation.* Journal of Computational Chemistry, 1991. **12**(4): p. 435-445.

65.   Fogolari, F., A. Brigo, and H. Molinari, *Protocol for MM/PBSA molecular dynamics simulations of proteins.* Biophysical Journal, 2003. **85**(1): p. 159-166.

66.   Kollman, P.A., et al., *Calculating structures and free energies of complex molecules: Combining molecular mechanics and continuum models.* Accounts of Chemical Research, 2000. **33**(12): p. 889-897.

67.   Swanson, J.M.J., R.H. Henchman, and J.A. McCammon, *Revisiting free energy calculations: A theoretical connection to MM/PBSA and direct calculation of the association free energy.* Biophysical Journal, 2004. **86**(1): p. 67-74.

68.   Georgescu, R.E., E.G. Alexov, and M.R. Gunner, *Combining conformational flexibility and continuum electrostatics for calculating pK(a)s in proteins.* Biophysical Journal, 2002. **83**(4): p. 1731-1748.

69.   Nielsen, J.E. and J.A. McCammon, *On the evaluation and optimization of protein X-ray structures for pKa calculations.* Protein Science, 2003. **12**(2): p. 313-326.

70.   Lu, B.Z., et al., *Protein molecular dynamics with electrostatic force entirely determined by a single Poisson-Boltzmann calculation.* Proteins-Structure Function and Genetics, 2002. **48**(3): p. 497-504.

71.   Luo, R., L. David, and M.K. Gilson, *Accelerated Poisson-Boltzmann calculations for static and dynamic systems.* Journal of Computational Chemistry, 2002. **23**(13): p. 1244-1253.

72.   Still, W.C., et al., *Semianalytical treatment of solvation for molecular mechanics and dynamics.* Journal of the American Chemical Society, 1990. **112**(16): p. 6127-6129.

73.   Onufriev, A., D.A. Case, and D. Bashford, *Effective Born radii in the generalized Born approximation: The importance of being perfect.* Journal of Computational Chemistry, 2002. **23**(14): p. 1297-1304.

74.   Onufriev, A.V. and G. Sigalov, *A strategy for reducing gross errors in the generalized Born models of implicit solvation.* Journal of Chemical Physics. **134**(16).

75.   Bashford, D. and D.A. Case, *Generalized born models of macromolecular solvation effects.* Annual Review of Physical Chemistry, 2000. **51**: p. 129-152.

76.   Ghosh, A., C.S. Rapp, and R.A. Friesner, *Generalized born model based on a surface integral formulation.* Journal of Physical Chemistry B, 1998. **102**(52): p. 10983-10990.

77.   Grant, J.A., et al., *The Gaussian Generalized Born model: application to small molecules.* Physical Chemistry Chemical Physics, 2007. **9**(35): p. 4913-4922.

78.   Haberthur, U. and A. Caflisch, *FACTS: Fast analytical continuum treatment of solvation.* Journal of Computational Chemistry, 2008. **29**(5): p. 701-715.

79.   Lee, M.S., F.R. Salsbury, and C.L. Brooks, *Novel generalized Born methods.* Journal of Chemical Physics, 2002. **116**(24): p. 10606-10614.

80.   Lee, M.S., et al., *New analytic approximation to the standard molecular volume definition and its application to generalized born calculations.* Journal of Computational Chemistry, 2003. **24**(11): p. 1348-1356.

81.   Onufriev, A., D. Bashford, and D.A. Case, *Exploring protein native states and large-scale conformational changes with a modified generalized born model.* Proteins-Structure Function and Bioinformatics, 2004. **55**(2): p. 383-394.

82.   Tsui, V. and D.A. Case, *Molecular dynamics simulations of nucleic acids with a generalized born solvation model.* Journal of the American Chemical Society, 2000. **122**(11): p. 2489-2498.

83.   Feig, M., W. Im, and C.L. Brooks, *Implicit solvation based on generalized Born theory in different dielectric environments.* Journal of Chemical Physics, 2004. **120**(2): p. 903-911.

84.   Tanizaki, S. and M. Feig, *A generalized Born formalism for heterogeneous dielectric environments: Application to the implicit modeling of biological membranes.* Journal of Chemical Physics, 2005. **122**(12).

85.   Ulmschneider, M.B., et al., *A generalized Born implicit-membrane representation compared to experimental insertion free energies.* Biophysical Journal, 2007. **92**(7): p. 2338-2349.

86. Spassov, V.Z., L. Yan, and S. Szalma, *Introducing an implicit membrane in generalized Born/solvent accessibility continuum solvent models.* Journal of Physical Chemistry B, 2002. **106**(34): p. 8726-8738.

87. Gallicchio, E. and R.M. Levy, *AGBNP: An analytic implicit solvent model suitable for molecular dynamics simulations and high-resolution modeling.* Journal of Computational Chemistry, 2004. **25**(4): p. 479-499.

88. Gallicchio, E., K. Paris, and R.M. Levy, *The AGBNP2 implicit solvation model.* Journal of Chemical Theory and Computation, 2009. **5**(9): p. 2544-2564.

89. Dominy, B.N. and C.L. Brooks, *Development of a generalized born model parametrization for proteins and nucleic acids.* Journal of Physical Chemistry B, 1999. **103**(18): p. 3765-3773.

90. Scarsi, M., J. Apostolakis, and A. Caflisch, *Continuum electrostatic energies of macromolecules in aqueous solutions.* Journal of Physical Chemistry A, 1997. **101**(43): p. 8098-8106.

91. Archontis, G. and T. Simonson, *A residue-pairwise generalized Born scheme suitable for protein design calculations.* Journal of Physical Chemistry B, 2005. **109**(47): p. 22667-22673.

92. Calimet, N., M. Schaefer, and T. Simonson, *Protein molecular dynamics with the generalized Born/ACE solvent model.* Proteins-Structure Function and Genetics, 2001. **45**(2): p. 144-158.

93. Schaefer, M. and M. Karplus, *A comprehensive analytical treatment of continuum electrostatics.* Journal of Physical Chemistry, 1996. **100**(5): p. 1578-1599.

94. Im, W.P., M.S. Lee, and C.L. Brooks, *Generalized born model with a simple smoothing function.* Journal of Computational Chemistry, 2003. **24**(14): p. 1691-1702.

95. Hawkins, G.D., C.J. Cramer, and D.G. Truhlar, *Parametrized models of aqueous free energies of solvation based on pairwise descreening of solute atomic charges from a dielectric medium.* Journal of Physical Chemistry, 1996. **100**(51): p. 19824-19839.

96. Hawkins, G.D., C.J. Cramer, and D.G. Truhlar, *Pairwise solute descreening of solute charges from a dielectric medium.* Chemical Physics Letters, 1995. **246**(1-2): p. 122-129.

97.   Simeonov, A., et al., *Blue-fluorescent antibodies.* Science, 2000. **290**(5490): p. 307-313.

98.   Dzubiella, J., J.M.J. Swanson, and J.A. McCammon, *Coupling hydrophobicity, dispersion, and electrostatics in continuum solvent models.* Physical Review Letters, 2006. **96**(8).

99.   Cheng, L.T., et al., *Interfaces and hydrophobic interactions in receptor-ligand systems: A level-set variational implicit solvent approach.* Journal of Chemical Physics, 2009. **131**(14).

100.   Dzubiella, J., J.M.J. Swanson, and J.A. McCammon, *Coupling nonpolar and polar solvation free energies in implicit solvent models.* Journal of Chemical Physics, 2006. **124**(8).

101.   Lee, M.S., F.R. Salsbury, and M.A. Olson, *An efficient hybrid explicit/implicit solvent method for biomolecular simulations.* Journal of Computational Chemistry, 2004. **25**(16): p. 1967-1978.

102.   Beglov, D. and B. Roux, *Dominant salvation effects from the primary shell of hydration. Approximation for molecular dynamics simulations.* Biopolymers, 1995. **35**(2): p. 171-178.

103.   Fennell, C.J., C.W. Kehoe, and K.A. Dill, *Modeling aqueous solvation with semi-explicit assembly.* Proceedings of the National Academy of Sciences of the United States of America. **108**(8): p. 3234-3239.

104.   Onsager, L., *Electric moments of molecules in liquids.* Journal of the American Chemical Society, 1936. **58**: p. 1486-1493.

105.   Wagoner, J.A. and V.S. Pande, *A smoothly decoupled particle interface: new methods for coupling explicit and implicit solvent.* The Journal of chemical physics, 2011. **134**(21): p. 214103.

106.   Im, W., M. Feig, and C.L. Brooks, *An implicit membrane generalized born theory for the study of structure, stability, and interactions of membrane proteins.* Biophysical Journal, 2003. **85**(5): p. 2900-2918.

107.   Stern, H.A. and S.E. Feller, *Calculation of the dielectric permittivity profile for a nonuniform system: Application to a lipid bilayer simulation.* Journal of Chemical Physics, 2003. **118**(7): p. 3401-3412.

108.   Lazaridis, T., *Effective energy function for proteins in lipid membranes.* Proteins-Structure Function and Genetics, 2003. **52**(2): p. 176-192.

109.   Sharp, K.A., et al., *Extracting hydrophobic free energies from experimentaldata. Relationship to protein folding and theoretical models.* Biochemistry, 1991. **30**(40): p. 9686-9697.

110.   Lum, K., D. Chandler, and J.D. Weeks, *Hydrophobicity at small and large length scales.* Journal of Physical Chemistry B, 1999. **103**(22): p. 4570-4577.

111.   Wallqvist, A. and D.G. Covell, *Free energy cost of bending n-dodecane in aqueous solution. Influence of the hydrophobic effect and solvent exposed area.* Journal of Physical Chemistry, 1995. **99**(35): p. 13118-13125.

112.   Ashbaugh, H.S., E.W. Kaler, and M.E. Paulaitis, *Hydration and conformational equilibria of simple hydrophobic and amphiphilic solutes.* Biophysical Journal, 1998. **75**(2): p. 755-768.

113.   Ashbaugh, H.S., E.W. Kaler, and M.E. Paulaitis, *A "universal" surface area correlation for molecular hydrophobic phenomena.* Journal of the American Chemical Society, 1999. **121**(39): p. 9243-9244.

114.   Gallicchio, E., M.M. Kubo, and R.M. Levy, *Enthalpy-entropy and cavity decomposition of alkane hydration free energies: Numerical results and implications for theories of hydrophobic solvation.* Journal of Physical Chemistry B, 2000. **104**(26): p. 6271-6285.

115.   Levy, R.M., et al., *On the nonpolar hydration free energy of proteins: Surface area and continuum solvent models for the solute-solvent interaction energy.* Journal of the American Chemical Society, 2003. **125**(31): p. 9523-9530.

116.   Fennell, C.J., C. Kehoe, and K.A. Dill, *Oil/water transfer is partly driven by molecular shape, not just size.* Journal of the American Chemical Society. **132**(1): p. 234-240.

117.   Kelly, C.P., C.J. Cramer, and D.G. Truhlar, *SM6: A density functional theory continuum solvation model for calculating aqueous solvation free energies of neutrals, ions, and*

solute-water clusters. Journal of Chemical Theory and Computation, 2005. **1**(6): p. 1133-1152.

118. Purisima, E.O. and T. Sulea, *Restoring charge asymmetry in continuum electrostatics calculations of hydration free energies.* Journal of Physical Chemistry B, 2009. **113**(24): p. 8206-8209.

119. Dorairaj, S. and T.W. Allen, *On the thermodynamic stability of a charged arginine side chain in a transmembrane helix.* Proceedings of the National Academy of Sciences of the United States of America, 2007. **104**(12): p. 4943-4948.

120. Lwin, T.Z. and R. Luo, *Force field influences in beta-hairpin folding simulations.* Protein Science, 2006. **15**(11): p. 2642-2655.

121. Mihajlovic, M. and T. Lazaridis, *Membrane-bound structure and energetics of alpha-synuclein.* Proteins-Structure Function and Bioinformatics, 2008. **70**(3): p. 761-778.

122. Ulmschneider, M.B. and J.P. Ulmschneider, *Membrane adsorption, folding, insertion and translocation of synthetic trans-membrane peptides.* Mol Membr Biol, 2008. **25**(3): p. 245-57.

123. Efremov, R.G., et al., *A solvent model for simulations of peptides in bilayers. I. Membrane-promoting alpha-helix formation.* Biophys J, 1999. **76**(5): p. 2448-59.

124. Bechor, D. and N. Ben-Tal, *Implicit solvent model studies of the interactions of the influenza hemagglutinin fusion peptide with lipid bilayers.* Biophys J, 2001. **80**(2): p. 643-55.

125. Kessel, A., D.P. Tieleman, and N. Ben-Tal, *Implicit solvent model estimates of the stability of model structures of the alamethicin channel.* Eur Biophys J, 2004. **33**(1): p. 16-28.

126. Tsien, R.W., *Cyclic AMP and contractile activity in heart.* Advances in cyclic nucleotide research, 1977. **8**: p. 363-420.

127. Tada, M. and A.M. Katz, *Phosphorylation of the sarcoplasmic-reticulum and sarcolemma.* Annual Review of Physiology, 1982. **44**: p. 401-423.

128. Jones, L.R., *Sarcolemmal enzymes mediating beta-adrenergic effects on the heart.* Current Topics in Membranes and Transport, 1985. **25**: p. 11-41.

129. Simmerman, H.K. and L.R. Jones, *Phospholamban: protein structure, mechanism of action, and role in cardiac function.* Physiol Rev, 1998. **78**(4): p. 921-47.

130. Tada, M., M.A. Kirchberger, and A.M. Katz, *Phosphorylation of a 22,000-dalton component of cardiac sarcoplasmic-reticulum by adenosine 3'-5'-monophosphate-dependent protein kinase.* Journal of Biological Chemistry, 1975. **250**(7): p. 2640-2647.

131. Tada, M., et al., *Calcium-transport by cardiac sarcoplasmic-reticulum and phosphorylation of phospholamban.* Molecular and Cellular Biochemistry, 1982. **46**(2): p. 73-95.

132. Bilezikjian, L.M., et al., *Studies on phosphorylation of canine cardiac sarcoplasmic reticulum by calmodulin-dependent protein kinase.* Circ Res, 1981. **49**(6): p. 1356-62.

133. Imagawa, T., T. Watanabe, and T. Nakamura, *Subunit structure and multiple phosphorylation sites of phospholamban.* Journal of Biochemistry, 1986. **99**(1): p. 41-53.

134. Colyer, J. and J.H. Wang, *Dependence of cardiac sarcoplasmic-reticulum calcium-pump activity on the phosphorylation status of phospholamban.* Journal of Biological Chemistry, 1991. **266**(26): p. 17486-17493.

135. Chu, G., et al., *A single site (Ser16) phosphorylation in phospholamban is sufficient in mediating its maximal cardiac responses to beta -agonists.* J Biol Chem, 2000. **275**(49): p. 38938-43.

136. Vittone, L., et al., *Camp and calcium-dependent mechanisms of phospholamban phosphorylation in intact hearts.* American Journal of Physiology, 1990. **258**(2): p. H318-H325.

137. Lindemann, J.P. and A.M. Watanabe, *Phosphorylation of phospholamban in intact myocardium. Role of Ca2+-calmodulin-dependent mechanisms.* J Biol Chem, 1985. **260**(7): p. 4516-25.

138. Lindemann, J.P., et al., *beta-Adrenergic stimulation of phospholamban phosphorylation and Ca2+-ATPase activity in guinea pig ventricles.* J Biol Chem, 1983. **258**(1): p. 464-71.

139.   Mattiazzi, A., et al., *Role of phospholamban phosphorylation on Thr(17) in cardiac physiological and pathological conditions.* Cardiovascular research, 2005. **68**(3): p. 366-375.

140.   Ahmed, Z., et al., *A solid-state NMR study of the phospholamban transmembrane domain: local structure and interactions with Ca2+-ATPase.* Biochimica Et Biophysica Acta-Biomembranes, 2000. **1468**(1-2): p. 187-198.

141.   Toyoshima, C. and T. Mizutani, *Crystal structure of the calcium pump with a bound ATP analogue.* Nature, 2004. **430**(6999): p. 529-535.

142.   Olesen, C., et al., *The structural basis of calcium transport by the calcium pump.* Nature, 2007. **450**(7172): p. 1036-U5.

143.   Zamoon, J., et al., *NMR solution structure and topological orientation of monomeric phospholamban in dodecylphosphocholine micelles.* Biophys J, 2003. **85**(4): p. 2589-98.

144.   Traaseth, N.J., et al., *Structural dynamics and topology of phospholamban in oriented lipid bilayers using multidimensional solid-state NMR.* Biochemistry, 2006. **45**(46): p. 13827-34.

145.   Mascioni, A., et al., *Solid-state NMR and rigid body molecular dynamics to determine domain orientations of monomeric phospholamban.* Journal of the American Chemical Society, 2002. **124**(32): p. 9392-9393.

146.   Toyoshima, C., et al., *Modeling of the inhibitory interaction of phospholamban with the Ca2+ ATPase.* Proc Natl Acad Sci U S A, 2003. **100**(2): p. 467-72.

147.   Jones, L.R. and L.J. Field, *Residues 2-25 of phospholamban are insufficient to inhibit Ca2+ transport atpase of cardiac sarcoplasmic-reticulum.* Journal of Biological Chemistry, 1993. **268**(16): p. 11486-11488.

148.   Karim, C.B., et al., *Synthetic null-cysteine phospholamban analogue and the corresponding transmembrane domain inhibit the Ca-ATPase.* Biochemistry, 2000. **39**(35): p. 10892-10897.

149.   Traaseth, N.J., et al., *Structural and dynamic basis of phospholamban and sarcolipin inhibition of Ca2+-ATPaset.* Biochemistry, 2008. **47**(1): p. 3-13.

150.   Arkin, I.T., et al., *Structural perspectives of phospholamban, a helical transmembrane pentamer.* Annual Review of Biophysics and Biomolecular Structure, 1997. **26**: p. 157-179.

151.   Stokes, D.L. and T. Wagenknecht, *Calcium transport across the sarcoplasmic reticulum - Structure and function of Ca2+-ATPase and the ryanodine receptor.* European Journal of Biochemistry, 2000. **267**(17): p. 5274-5279.

152.   Kimura, Y., et al., *Phospholamban inhibitory function is activated by depolymerization.* Journal of Biological Chemistry, 1997. **272**(24): p. 15061-15064.

153.   Reddy, L.G., L.R. Jones, and D.D. Thomas, *Depolymerization of phospholamban in the presence of calcium pump: A fluorescence energy transfer study.* Biochemistry, 1999. **38**(13): p. 3954-3962.

154.   Traaseth, N.J., et al., *Structure and topology of monomeric phospholamban in lipid membranes determined by a hybrid solution and solid-state NMR approach.* Proc Natl Acad Sci U S A, 2009. **106**(25): p. 10165-70.

155.   Shi, L., et al., *A refinement protocol to determine structure, topology, and depth of insertion of membrane proteins using hybrid solution and solid-state NMR restraints.* Journal of Biomolecular Nmr, 2009. **44**(4): p. 195-205.

156.   Sayadi, M., S. Tanizaki, and M. Feig, *Effect of membrane thickness on conformational sampling of phospholamban from computer simulations.* Biophysical Journal, 2010. **98**(5): p. 805-14.

157.   Autry, J.M. and L.R. Jones, *Functional Co-expression of the canine cardiac Ca2+ pump and phospholamban in Spodoptera frugiperda (Sf21) cells reveals new insights on ATPase regulation.* J Biol Chem, 1997. **272**(25): p. 15872-80.

158.   Li, J., D.J. Bigelow, and T.C. Squier, *Phosphorylation by cAMP-dependent protein kinase modulates the structural coupling between the transmembrane and cytosolic domains of phospholamban.* Biochemistry, 2003. **42**(36): p. 10674-82.

159.   Schmitt, J.P., et al., *Dilated cardiomyopathy and heart failure caused by a mutation in phospholamban.* Science, 2003. **299**(5611): p. 1410-1413.

160. Simmerman, H.K.B., et al., *A leucine zipper stabilizes the pentameric membrane domain of phospholamban and forms a coiled-coil pore structure.* Journal of Biological Chemistry, 1996. **271**(10): p. 5941-5946.

161. Thomas, D.D., et al., *Direct spectroscopic detection of molecular dynamics and interactions of the calcium pump and phospholamban.* Cardiac Sarcoplasmic Reticulum Function and Regulation of Contractility, 1998. **853**: p. 186-194.

162. Li, M., et al., *A fluorescence energy transfer method for analyzing protein oligomeric structure: Application to phospholamban.* Biophysical Journal, 1999. **76**(5): p. 2587-2599.

163. Lamberth, S., et al., *NMR solution structure of phospholamban.* Helvetica Chimica Acta, 2000. **83**(9): p. 2141-2152.

164. Metcalfe, E.E., et al., *H-1/N-15 heteronuclear NMR spectroscopy shows four dynamic domains for phospholamban reconstituted in dodecylphosphocholine micelles.* Biophysical Journal, 2004. **87**(2): p. 1205-1214.

165. Zamoon, J., et al., *Mapping the interaction surface of a membrane protein: unveiling the conformational switch of phospholamban in calcium pump regulation.* Proc Natl Acad Sci U S A, 2005. **102**(13): p. 4747-52.

166. Karim, C.B., et al., *Phospholamban structural dynamics in lipid bilayers probed by a spin label rigidly coupled to the peptide backbone.* Proc Natl Acad Sci U S A, 2004. **101**(40): p. 14437-42.

167. Li, J.H., et al., *Phospholamban binds in a compact and ordered conformation to the Ca-ATPase.* Biochemistry, 2004. **43**(2): p. 455-463.

168. Sugita, Y., et al., *Structural changes in the cytoplasmic domain of phospholamban by phosphorylation at Ser16: a molecular dynamics study.* Biochemistry, 2006. **45**(39): p. 11752-61.

169. Paterlini, M.G. and D.D. Thomas, *The alpha-helical propensity of the cytoplasmic domain of phospholamban: A molecular dynamics simulation of the effect of phosphorylation and mutation.* Biophysical Journal, 2005. **88**(5): p. 3243-3251.

170.   Pantano, S. and E. Carafoli, *The role of phosphorylation on the structure and dynamics of phospholamban: A model from molecular simulations.* Proteins-Structure Function and Bioinformatics, 2007. **66**(4): p. 930-940.

171.   Kim, T., J. Lee, and W. Im, *Molecular dynamics studies on structure and dynamics of phospholamban monomer and pentamer in membranes.* Proteins, 2009. **76**(1): p. 86-98.

172.   Houndonougbo, Y., K. Kuczera, and G.S. Jas, *Structure and dynamics of phospholamban in solution and in membrane bilayer: computer simulations.* Biochemistry, 2005. **44**(6): p. 1780-92.

173.   Chocholousova, J. and M. Feig, *Balancing an accurate representation of the molecular surface in generalized born formalisms with integrator stability in molecular dynamics simulations.* Journal of Computational Chemistry, 2006. **27**(6): p. 719-729.

174.   Brooks, C.L., M. Berkowitz, and S.A. Adelman, *Generalized Langevin theory for many-body problems in chemical dynamics, gas surface collisions, vibrational energy relaxation in solids, and recombination reactions in liquids.* Journal of Chemical Physics, 1980. **73**(9): p. 4353-4364.

175.   Feig, M., *Kinetics from implicit solvent simulations of biomolecules as a function of viscosity.* Journal of Chemical Theory and Computation, 2007. **3**(5): p. 1734-1748.

176.   Ryckaert, J.P., G. Ciccotti, and H.J.C. Berendsen, *Numerical integration of cartesian equations of motion of a system with constraints molecular dynamics of n-alkanes.* Journal of Computational Physics, 1977. **23**(3): p. 327-341.

177.   Brooks, B.R., et al., *CHARMM - a program for macromolecular energy, minimization, and dynamics calculations.* Journal of Computational Chemistry, 1983. **4**(2): p. 187-217.

178.   Feig, M., J. Karanicolas, and C.L. Brooks, *MMTSB Tool Set: enhanced sampling and multiscale modeling methods for applications in structural biology.* Journal of Molecular Graphics & Modelling, 2004. **22**(5): p. 377-395.

179.   Tanizaki, S. and M. Feig, *A generalized Born formalism for heterogeneous dielectric environments: Application to the implicit modeling of biological membranes.* Journal of Chemical Physics, 2005. **122**(12): p. -.

180.   Marrink, S.J. and H.J.C. Berendsen, *Permeation process of small molecules across lipid membranes studied by molecular dynamics simulations.* Journal of Physical Chemistry, 1996. **100**(41): p. 16729-16738.

181.   Kumar, S., et al., *The weighted histogram analysis method for free energy calculations on biomolecules .1. The method.* Journal of Computational Chemistry, 1992. **13**(8): p. 1011-1021.

182.   DeLano, W.L., *The PyMOL molecular graphics system* 2002, DeLano Scientific Palo Alto, CA.

183.   MacCallum, J.L., W.F.D. Bennett, and D.P. Tieleman, *Partitioning of amino acid side chains into lipid bilayers: Results from computer simulations and comparison to experiment.* Journal of General Physiology, 2007. **129**(5): p. 371-377.

184.   Radzicka, A.P., L.; R. Wolfenden, *Influences of solvent water on protein folding free energies of solvation of cis and trans peptides are nearly identical.* Biochemistry, 1988. **27**(12): p. 4538-4541.

185.   Trump, B.F., Duttera, S. D.; William, L. B.; and Arstila, A. U, *Membrane structure: Lipid protein interactions in microsomal membranes.* Proceedings of the National Academy of Sciences of the United States of America, 1970. **66**(2): p. 433-440.

186.   Mascioni, A. and G. Veglia, *Theoretical analysis of residual dipolar coupling patterns in regular secondary structures of proteins.* Journal of the American Chemical Society, 2003. **125**(41): p. 12520-12526.

187.   Senes, A., et al., *E-z, a depth-dependent potential for assessing the energies of insertion of amino acid side-chains into membranes: Derivation and applications to determining the orientation of transmembrane and interfacial helices.* Journal of Molecular Biology, 2007. **366**(2): p. 436-448.

188.   Andronesi, O.C., et al., *Determination of membrane protein structure and dynamics by magic-angle-spinning solid-state NMR spectroscopy.* Journal of the American Chemical Society, 2005. **127**(37): p. 12965-12974.

189.   Holtzwarth, J.F.E., V.; and Genz, A, *Iodine laser temperature jump from picosecond to seconds: Relaxation processes of phospholipid bilayers, in spectroscopy and dynamics of bioligical systems*, in *Spectroscopy and Dynamics of Biological Systems*, P.M.D. Bayley, R., Editor 1984, Academic Press London. p. 351-377.

190. Tieleman, D.P. and D. Van der Spoel, *Molecular dynamics simulations of dodecylphosphocholine micelles at three different aggregate sizes: Micellar structure and chain relaxation.* Journal of Physical Chemistry, 1996. **104**(27): p. 6380-6388.

191. Tieleman, D.P. and H.J. Berendsen, *Molecular dynamics simulations of a fully hydrated dipalmitoyl phosphatidylcholine bilayer with different macroscopic boundary conditions and parameters.* Journal of Chemical Physics, 1996. **105**(11): p. 4871-4880.

192. Bick, R.J., et al., *Membrane asymmetry in isolated canine cardiac sarcoplasmic reticulum: comparison with skeletal muscle sarcoplasmic reticulum.* Journal of Membrane Biology, 1998. **164**(2): p. 169-175.

193. James, P., et al., *Nature and site of phospholamban regulation of the Ca2+ pump of sarcoplasmic reticulum.* Nature, 1989. **342**(6245): p. 90-92.

194. Chen, Z., et al., *Spatial and dynamic interactions between phospholamban and the canine cardiac Ca2+ pump revealed with use of heterobifunctional cross-linking agents.* J Biol Chem, 2003. **278**(48): p. 48348-56.

195. Cheah, A.M., *Effect of long chain unsaturated fatty acids on the calcium transport of sarcoplasmic reticulum.* Biochimica et Biophysica Acta, 1981. **648**(2): p. 113-119.

196. Lee, A.G.J.M.E.R.J.F., *Are essential fatty acids essential for membrane function?* Progress in Lipid Research, 1986. **25**(1-4): p. 41-46.

197. Radzicka, A., L. Pedersen, and R. Wolfenden, *Influences of Solvent Water on Protein Folding - Free-Energies of Solvation of Cis and Trans Peptides Are Nearly Identical.* Biochemistry, 1988. **27**(12): p. 4538-4541.

198. Sayadi, M., Feig, M.,, *Role of conformational sampling of Ser16 and Thr17-phosphorylated phospholamban in interactions with SERCA.* BBA Biomembranes, 2012. **In Press**.

199. Wegener, A.D., et al., *Phospholamban phosphorylation in intact ventricles. Phosphorylation of serine 16 and threonine 17 in response to beta-adrenergic stimulation.* J Biol Chem, 1989. **264**(19): p. 11468-74.

200.   Sham, J.S., L.R. Jones, and M. Morad, *Phospholamban mediates the beta-adrenergic-enhanced Ca2+ uptake in mammalian ventricular myocytes.* Am J Physiol, 1991. **261**(4 Pt 2): p. H1344-9.

201.   Kuschel, M., et al., *Ser16 prevails over Thr17 phospholamban phosphorylation in the beta-adrenergic regulation of cardiac relaxation.* Am J Physiol, 1999. **276**(5 Pt 2): p. H1625-33.

202.   Mundina-Weilenmann, C., et al., *Immunodetection of phosphorylation sites gives new insights into the mechanisms underlying phospholamban phosphorylation in the intact heart.* J Biol Chem, 1996. **271**(52): p. 33561-7.

203.   Hagemann, D., et al., *Frequency-encoding Thr17 phospholamban phosphorylation is independent of Ser16 phosphorylation in cardiac myocytes.* J Biol Chem, 2000. **275**(29): p. 22532-6.

204.   Bassani, R.A., A. Mattiazzi, and D.M. Bers, *CaMKII is responsible for activity-dependent acceleration of relaxation in rat ventricular myocytes.* Am J Physiol, 1995. **268**(2 Pt 2): p. H703-12.

205.   Hagemann, D. and R.P. Xiao, *Dual site phospholamban phosphorylation and its physiological relevance in the heart.* Trends Cardiovasc Med, 2002. **12**(2): p. 51-6.

206.   Kemi, O.J., et al., *Aerobic interval training enhances cardiomyocyte contractility and Ca2+ cycling by phosphorylation of CaMKII and Thr-17 of phospholamban.* J Mol Cell Cardiol, 2007. **43**(3): p. 354-61.

207.   Kranias, E.G., *Regulation of Ca2+ transport by cyclic 3',5'-AMP-dependent and calcium-calmodulin-dependent phosphorylation of cardiac sarcoplasmic reticulum.* Biochim Biophys Acta, 1985. **844**(2): p. 193-9.

208.   Pollesello, P., A. Annila, and M. Ovaska, *Structure of the 1-36 amino-terminal fragment of human phospholamban by nuclear magnetic resonance and modeling of the phospholamban pentamer.* Biophys J, 1999. **76**(4): p. 1784-95.

209.   Metcalfe, E.E., N.J. Traaseth, and G. Veglia, *Serine 16 phosphorylation induces an order-to-disorder transition in monomeric phospholamban.* Biochemistry, 2005. **44**(11): p. 4386-96.

210. Traaseth, N.J., D.D. Thomas, and G. Veglia, *Effects of Ser16 phosphorylation on the allosteric transitions of phospholamban/Ca(2+)-ATPase complex.* J Mol Biol, 2006. **358**(4): p. 1041-50.

211. Mortishire-Smith, R.J., et al., *Structural studies on phospholamban and implications for regulation of the Ca2+-ATPase.* Cardiac Sarcoplasmic Reticulum Function and Regulation of Contractility, 1998. **853**: p. 63-78.

212. Tatulian, S.A., et al., *Secondary structure and orientation of phospholamban reconstituted in supported bilayers from polarized attenuated total reflection FTIR spectroscopy.* Biochemistry, 1995. **34**(13): p. 4448-56.

213. Pollesello, P. and A. Annila, *Structure of the 1-36 N-terminal fragment of human phospholamban phosphorylated at Ser-16 and Thr-17.* Biophys J, 2002. **83**(1): p. 484-90.

214. Li, M., et al., *Phosphorylation-induced structural change in phospholamban and its mutants, detected by intrinsic fluorescence.* Biochemistry, 1998. **37**(21): p. 7869-77.

215. Asahi, M., et al., *Physical interactions between phospholamban and sarco(endo)plasmic reticulum Ca2+-ATPases are dissociated by elevated Ca2+, but not by phospholamban phosphorylation, vanadate, or thapsigargin, and are enhanced by ATP.* J Biol Chem, 2000. **275**(20): p. 15034-8.

216. Negash, S., et al., *Phospholamban remains associated with the Ca2+- and Mg2+-dependent ATPase following phosphorylation by cAMP-dependent protein kinase.* Biochem J, 2000. **351**(Pt 1): p. 195-205.

217. Karim, C.B., et al., *Phosphorylation-dependent conformational switch in spin-labeled phospholamban bound to SERCA.* J Mol Biol, 2006. **358**(4): p. 1032-40.

218. Chen, Z., B.L. Akin, and L.R. Jones, *Mechanism of reversal of phospholamban inhibition of the cardiac Ca2+-ATPase by protein kinase A and by anti-phospholamban monoclonal antibody 2D12.* J Biol Chem, 2007. **282**(29): p. 20968-76.

219. Thomas, D.D., et al., *Direct spectroscopic detection of molecular dynamics and interactions of the calcium pump and phospholamban*, in *Cardiac Sarcoplasmic Reticulum Function and Regulation of Contractility*, R.G.K.E.G. Johnson, Editor 1998. p. 186-194.

220.   Chen, L.T., et al., *Phospholamban modulates the functional coupling between nucleotide domains in Ca-ATPase oligomeric complexes in cardiac sarcoplasmic reticulum.* Biochemistry, 2009. **48**(11): p. 2411-21.

221.   Panahi, A. and M. Feig, *Conformational sampling of influenza fusion peptide in membrane bilayers as a function of termini and protonation states.* J Phys Chem B. **114**(3): p. 1407-16.

222.   Tanizaki, S. and M. Feig, *Molecular dynamics simulations of large integral membrane proteins with an implicit membrane model.* J Phys Chem B, 2006. **110**(1): p. 548-56.

223.   Feig, M., W. Im, and C.L. Brooks, 3rd, *Implicit solvation based on generalized Born theory in different dielectric environments.* J Chem Phys, 2004. **120**(2): p. 903-11.

224.   Toyoshima, C. and H. Nomura, *Structural changes in the calcium pump accompanying the dissociation of calcium.* Nature, 2002. **418**(6898): p. 605-611.

225.   Jones, L.R., R.L. Cornea, and Z. Chen, *Close proximity between residue 30 of phospholamban and cysteine 318 of the cardiac Ca2+ pump revealed by intermolecular thiol cross-linking.* J Biol Chem, 2002. **277**(31): p. 28319-29.

226.   Chen, Z.H., D.L. Stokes, and L.R. Jones, *Role of leucine 31 of phospholamban in structural and functional interactions with the Ca2+ pump of cardiac sarcoplasmic reticulum.* Journal of Biological Chemistry, 2005. **280**(11): p. 10530-10539.

227.   Chen, Z.H., et al., *Cross-linking of C-terminal residues of phospholamban to the Ca2+ pump of cardiac sarcoplasmic reticulum to probe spatial and functional interactions within the transmembrane domain.* Journal of Biological Chemistry, 2006. **281**(20): p. 14163-14172.

228.   Sayadi, M., S. Tanizaki, and M. Feig, *Effect of membrane thickness on conformational sampling of phospholamban from computer simulations.* Biophys J. **98**(5): p. 805-14.

229.   Mattiazzi, A., et al., *Role of phospholamban phosphorylation on Thr17 in cardiac physiological and pathological conditions.* Cardiovasc Res, 2005. **68**(3): p. 366-75.

230.   Lindemann, J.P., *Alpha-adrenergic stimulation of sarcolemmal protein phosphorylation and slow responses in intact myocardium.* J Biol Chem, 1986. **261**(11): p. 4860-7.

231. Luo, W., et al., *Transgenic approaches to define the functional role of dual site phospholamban phosphorylation.* J Biol Chem, 1998. **273**(8): p. 4734-9.

232. Fukunishi, H., O. Watanabe, and S. Takada, *On the Hamiltonian replica exchange method for efficient sampling of biomolecular systems: Application to protein structure prediction.* Journal of Chemical Physics, 2002. **116**(20): p. 9058-9067.

233. Liu, P., et al., *Replica exchange with solute tempering: A method for sampling biological systems in explicit water.* Proceedings of the National Academy of Sciences of the United States of America, 2005. **102**(39): p. 13749-13754.

234. Su, L. and R.I. Cukier, *Hamiltonian and distance replica exchange method studies of Met-enkephalin.* Journal of Physical Chemistry B, 2007. **111**(42): p. 12310-12321.

235. Buck, M., *Trifluoroethanol and colleagues: cosolvents come of age. Recent studies with peptides and proteins.* Quarterly Reviews of Biophysics, 1998. **31**(3): p. 297-355.

236. Hong, D.P., et al., *Clustering of fluorine-substituted alcohols as a factor responsible for their marked effects on proteins and peptides.* Journal of the American Chemical Society, 1999. **121**(37): p. 8427-8433.

237. Shibata, A., et al., *Biphasic effects of alcohols on the phase transition of poly(l-lysine) between alpha helix and beta sheet conformations.* Biochemistry, 1992. **31**(25): p. 5728-5733.

238. Bhattacharjya, S., et al., *Fluoroalcohols as structure modifiers in peptides and proteins: hexafluoroacetone hydrate stabilizes a helical conformation of melittin at low pH.* Journal of Peptide Research, 1999. **54**(2): p. 100-111.

239. Kovacs, H., et al., *The effect of environment on the stability of an integral membrane helix . Molecular-dynamics simulations of surfactant protein C in chloroform, methanol and water.* Journal of Molecular Biology, 1995. **247**(4): p. 808-822.

240. Liu, H.L. and C.M. Hsu, *The effects of various alcohols on the stability of melittin: A molecular dynamics study.* Journal of the Chinese Chemical Society, 2003. **50**(6): p. 1235-1240.

241. Nguyen, H.D., A.J. Marchut, and C.K. Hall, *Solvent effects on the conformational transition of a model polyalanine peptide.* Protein Science, 2004. **13**(11): p. 2909-2924.

242. Searle, M.S., et al., *Native like beta hairpin structure in an isolated fragment from ferredoxin: NMR and CD studies of solvent effects on the N-terminal 20 residues.* Protein Engineering, 1996. **9**(7): p. 559-565.

243. Tanizaki, S., et al., *Conformational sampling of peptides in cellular environments.* Biophysical Journal, 2008. **94**(3): p. 747-759.

244. Akhadov, I.A.I.U., *Dielectric properties of binary solutions*1981: Pergamon Press.

245. Yoda, T., Y. Sugita, and Y. Okamoto, *Hydrophobic core formation and dehydration in protein folding studied by generalized ensemble simulations.* Biophysical Journal, 2010. **99**(5): p. 1637-1644.

246. Chakrabartty, A., T. Kortemme, and R.L. Baldwin, *Helix propensities of the amino acids measured in alanine based peptides without helix stabilizing side chain interactions.* Protein Science, 1994. **3**(5): p. 843-852.

247. Scholtz, J.M., et al., *The energetics of ion pair and hydrogen-bonding interactions in a helical peptide.* Biochemistry, 1993. **32**(37): p. 9668-9676.

248. Scholtz, J.M., et al., *A neutral, water-soluble, alpha-helical peptide. The effect of ionic strength on the helix coil equilibrium.* Journal of the American Chemical Society, 1991. **113**(13): p. 5102-5104.

249. Shalongo, W., L. Dugad, and E. Stellwagen, *Distribution of helicity within the model peptide acetyl(aaqaa)(3)amide.* Journal of the American Chemical Society, 1994. **116**(18): p. 8288-8293.

250. Shirley, W.A. and C.L. Brooks, *Curious structure in "canonical" alanine based peptides.* Proteins-Structure Function and Genetics, 1997. **28**(1): p. 59-71.

251. Hughes, J., et al., *Identification of 2 Related Pentapeptides from Brain with Potent Opiate Agonist Activity.* Nature, 1975. **258**(5536): p. 577-579.

252. Sutto, L., M. D'Abramo, and F.L. Gervasio, *Comparing the Efficiency of Biased and Unbiased Molecular Dynamics in Reconstructing the Free Energy Landscape of Met-Enkephalin.* Journal of Chemical Theory and Computation. **6**(12): p. 3640-3646.

253. Klepeis, J.L. and C.A. Floudas, *Free energy calculations for peptides via deterministic global optimization.* Journal of Chemical Physics, 1999. **110**(15): p. 7491-7512.

254. Carlacci, L., *Conformational analysis of [Met(5)]-enkephalin: Solvation and ionization considerations.* Journal of Computer-Aided Molecular Design, 1998. **12**(2): p. 195-213.

255. Shen, M.Y. and K.F. Freed, *Long time dynamics of met-enkephalin: Comparison of explicit and implicit solvent models.* Biophysical Journal, 2002. **82**(4): p. 1791-1808.

256. Zaman, M.H., et al., *Computer simulation of met-enkephalin using explicit atom and united atom potentials: Similarities, differences, and suggestions for improvement.* Journal of Physical Chemistry B, 2003. **107**(7): p. 1685-1691.

257. Montcalm, T., et al., *Simulated annealing of Met-enkephalin. Low energy states and their relevance to membrane bound, solution and solid state conformations.* Theochem-Journal of Molecular Structure, 1994. **114**: p. 37-51.

258. Eisenmenger, F. and U.H.E. Hansmann, *Variation of the energy landscape of a small peptide under a change from the ECEPP/2 force field to ECEPP/3.* Journal of Physical Chemistry B, 1997. **101**(16): p. 3304-3310.

259. Hansmann, U.H.E., M. Masuya, and Y. Okamoto, *Characteristic temperatures of folding of a small peptide.* Proceedings of the National Academy of Sciences of the United States of America, 1997. **94**(20): p. 10652-10656.

260. Sanbonmatsu, K.Y. and A.E. Garcia, *Structure of Met-enkephalin in explicit aqueous solution using replica exchange molecular dynamics.* Proteins-Structure Function and Genetics, 2002. **46**(2): p. 225-234.

261. Karvounis, G., D. Nerukh, and R.C. Glen, *Water network dynamics at the critical moment of a peptide's beta-turn formation: A molecular dynamics study.* Journal of Chemical Physics, 2004. **121**(10): p. 4925-4935.

262. Nielsen, B.G., M.O. Jensen, and H.G. Bohr, *The probability distribution of side-chain conformations in [Leu] and [Met]enkephalin determines the potency and selectivity to mu and delta opiate receptors.* Biopolymers, 2003. **71**(5): p. 577-592.

263. Aburi, M. and P.E. Smith, *A conformational analysis of leucine enkephalin as a function of pH.* Biopolymers, 2002. **64**(4): p. 177-188.

264.   Harada, R. and A. Kitao, *Multi-scale free energy landscape calculation method by combination of coarse-grained and all-atom models.* Chemical Physics Letters. **503**(1-3): p. 145-152.

265.   Ramya, L. and N. Gautham, *Effects of hydration on the conformational energy landscape of the pentapeptide Met-enkephalin.* Journal of Chemical Theory and Computation, 2009. **5**(8): p. 2180-2190.

266.   Lee, M.S., et al., *New analytic approximation to the standard molecular volume definition and its application to generalized born calculations (vol 24, pg 1348, 2003).* Journal of Computational Chemistry, 2003. **24**(14): p. 1821-1821.

267.   Fink, A.L., *Chaperone-mediated protein folding.* Physiological reviews, 1999. **79**(2): p. 425-449.

268.   Oshea, E.K., et al., *X-ray structure of the GCN4 leucine zipper, a 2-stranded, parallel coiled coil.* Science, 1991. **254**(5031): p. 539-544.

269.   Finney, J.L., *Hydration processes in biological and macromolecular systems.* Faraday Discussions, 1996. **103**: p. 1-18.

270.   Sokolov, A.P., et al., *Role of hydration water in dynamics of biological macromolecules.* Chemical Physics, 2008. **345**(2-3): p. 212-218.

271.   Sitkoff, D., K.A. Sharp, and B. Honig, *Accurate calculation of hydration free energies using macroscopic solvent models.* Journal of Physical Chemistry, 1994. **98**(7): p. 1978-1988.

272.   Marten, B., et al., *New model for calculation of solvation free energies: Correction of self-consistent reaction field continuum dielectric theory for short-range hydrogen-bonding effects.* Journal of Physical Chemistry, 1996. **100**(28): p. 11775-11788.

273.   Simonson, T. and A.T. Brunger, *Solvation Free-Energies Estimated from Macroscopic Continuum Theory - an Accuracy Assessment.* Journal of Physical Chemistry, 1994. **98**(17): p. 4683-4694.

274.   Pellegrini, E. and M.J. Field, *A generalized-born solvation model for macromolecular hybrid-potential calculations.* Journal of Physical Chemistry A, 2002. **106**(7): p. 1316-1326.

275.   Lee, M.R., Y. Duan, and P.A. Kollman, *Use of MM-PB/SA in estimating the free energies of proteins: Application to native, intermediates, and unfolded villin headpiece.* Proteins-Structure Function and Genetics, 2000. **39**(4): p. 309-316.

276.   Hermann, R.B., *Theory of Hydrophobic Bonding .2. Correlation of Hydrocarbon Solubility in Water with Solvent Cavity Surface-Area.* Journal of Physical Chemistry, 1972. **76**(19): p. 2754-&.

277.   Chothia, C., *Hydrophobic bonding and accessible surface area in proteins.* Nature, 1974. **248**(5446): p. 338-339.

278.   Kang, Y.K., G. Nemethy, and H.A. Scheraga, *Free energies of hydration of solute molecules .2. Application of the hydration shell model to nonionic organic molecules.* Journal of Physical Chemistry, 1987. **91**(15): p. 4109-4117.

279.   Kang, Y.K., et al., *Free energies of hydration of solute molecules .4. revised treatment of the hydration shell model.* Journal of Physical Chemistry, 1988. **92**(16): p. 4739-4742.

280.   Kang, Y.K.G.N.a.H.A.S., *Free energies of hydration of solute molecules .3. Application of the hydration shell model to charged organic molecules. .* Journal of Physical Chemistry, 1987. **91**(15): p. 4118-4120.

281.   Huang, D.M., P.L. Geissler, and D. Chandler, *Scaling of hydrophobic solvation free energies.* Journal of Physical Chemistry B, 2001. **105**(28): p. 6704-6709.

282.   Rajamani, S., T.M. Truskett, and S. Garde, *Hydrophobic hydration from small to large lengthscales: Understanding and manipulating the crossover.* Proceedings of the National Academy of Sciences of the United States of America, 2005. **102**(27): p. 9475-9480.

283.   Huang, D.M. and D. Chandler, *Temperature and length scale dependence of hydrophobic effects and their possible implications for protein folding.* Proceedings of the National Academy of Sciences of the United States of America, 2000. **97**(15): p. 8324-8327.

284.   Zacharias, M., *Continuum solvent modeling of nonpolar solvation: Improvement by separating surface area dependent cavity and dispersion contributions.* Journal of Physical Chemistry A, 2003. **107**(16): p. 3000-3004.

285. Pitera, J.W. and W.F. van Gunsteren, *The importance of solute-solvent van der Waals interactions with interior atoms of biopolymers.* Journal of the American Chemical Society, 2001. **123**(13): p. 3163-3164.

286. Su, Y. and E. Gallicchio, *The non-polar solvent potential of mean force for the dimerization of alanine dipeptide: the role of solute-solvent van der Waals interactions.* Biophysical Chemistry, 2004. **109**(2): p. 251-260.

287. Gallicchio, E., L.Y. Zhang, and R.M. Levy, *The SGB/NP hydration free energy model based on the surface generalized born solvent reaction field and novel nonpolar hydration free energy estimators.* Journal of Computational Chemistry, 2002. **23**(5): p. 517-529.

288. Pitarch, J., et al., *Can hydrophobic interactions be correctly reproduced by the continuum models?* Journal of Physical Chemistry, 1996. **100**(23): p. 9955-9959.

289. Weeks, J.D., D. Chandler, and H.C. Andersen, *Role of repulsive forces in determining equilibrium structure of simple liquids.* Journal of Chemical Physics, 1971. **54**(12): p. 5237-+.

290. Wagoner, J.A. and N.A. Baker, *Assessing implicit models for nonpolar mean solvation forces: The importance of dispersion and volume terms.* Proceedings of the National Academy of Sciences of the United States of America, 2006. **103**(22): p. 8331-8336.

291. Floris, F. and J. Tomasi, *Evaluation of the dispersion contribution to the solvation energy. A simple computational model in the continuum approximation.* Journal of Computational Chemistry, 1989. **10**(5): p. 616-627.

292. Floris, F.M., J. Tomasi, and J.L.P. Ahuir, *Dispersion and repulsion contributions to the solvation energy. Refinements to a simple computational model in the continuum approximation.* Journal of Computational Chemistry, 1991. **12**(7): p. 784-791.

293. MacCallum, J.L., W.F.D. Bennett, and D.P. Tieleman, *Distribution of amino acids in a lipid bilayer from computer simulations.* Biophysical Journal, 2008. **94**(9): p. 3393-3404.

294. Jorgensen, W.L. and J. Tiradorives, *The OPLS potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin.* Journal of the American Chemical Society, 1988. **110**(6): p. 1657-1666.

295. Jo, S., et al., *CHARMM-GUI membrane builder for mixed bilayers and its application to yeast membranes.* Biophysical Journal, 2009. **97**(1): p. 50-58.

296. Inaba, K., et al., *Dynamic nature of disulphide bond formation catalysts revealed by crystal structures of DsbB.* Embo Journal, 2009. **28**(6): p. 779-791.

297. Gonnet, P., *P-SHAKE: A quadratically convergent SHAKE in O(n(2)).* Journal of Computational Physics, 2007. **220**(2): p. 740-750.

298. Lee, J. and W. Im, *Implementation and application of helix-helix distance and crossing angle restraint potentials.* Journal of Computational Chemistry, 2007. **28**(3): p. 669-680.

299. Lee, J. and W. Im, *Role of hydrogen bonding and helix-lipid interactions in transmembrane helix association.* Journal of the American Chemical Society, 2008. **130**(20): p. 6456-6462.