Molecular Modeling and Molecular Dynamics Simulations of Biomembranes and Membrane Proteins

2007

Kenichi Mori

Molecular Modeling and Molecular Dynamics Simulations of Biomembranes and Membrane Proteins

2007

Kenichi Mori

Department of Physical Chemistry
Graduate School of Pharmaceutical Sciences
Chiba University

Contents

Figure & Table Contents

Tables

Figures

Abbreviations

Aβ	amyloid β-protein
AD	Alzheimer's disease
CE, CER	ceramide
Chol	cholesterol
COM	center of mass
CPE	ceramide phosphatidylethanolamine
CPI	ceramide phosphatidylinositol
CPG	ceramide phophatidylglycerol
CPS	ceramide phosphatidylserine
C1P	ceramide 1-phosphate
DG	*sn*-1,2-diacylglycerol
F-Cys	farnesylated cystein
Far-Ras	farnesylated H-Ras protein
Gal	D-galactose
GalNAc	N-acetyl-D-galactosamine
Glc	D-glucose
GM1	GM1 ganglioside
GPI	glycosylphosphatidylinositol
GSL	glycosphingolipid
H-bond	hydrogen bond
L_α	liquid crystalline phase
$L_\alpha I$	interdigitated liquid crystalline phase
$L_\beta I$	interdigitated gel phase
L_β'	tilted-gel phase
L_c	crystalline phase
MD	molecular dynamics
Neu5Ac	5-N-acetylneuramic acid
PA	1,2-diacyl-*sn*-glycero-3-phosphate
PC	1,2-diacyl-*sn*-glycero-3-phosphocholine
PE	1,2-diacyl-*sn*-glycero-3-phosphoethanolamine
PG	1,2-diacyl-*sn*-glycero-3-phospho-(1'-*sn*-glycerol)
PI	1,2-diacyl-*sn*-glycero-3-phospho-(1'-*myo*-inositol)
PS	1,2-diacyl-*sn*-glycero-3-phosphoserine
PME	particle mesh Ewald
RDF	radial distribution function
RMSD	root mean square deviation
SA	sialic acid
SASA	solvent accessible surface area
SM	sphingomyelin
vdW	van der Waals

I

Introduction

I.1 Biomembranes

Eukaryotic cells have a profusion of organelles (Figure 1.1 (a) and (b)).[1] Each organelle is surrounded by one or more biomembranes, in which a unique complement of proteins are embedded. These proteins enable each organelle to carry out its characteristic function. Biomembranes give the proteins a suitable reaction field in which their maximum performance can be exerted. In the proteins, there are many receptors that can detect chemical signals sent by other cells or other organelles in molecular forms in response to environmental changes. These signals are translated into other forms by the receptors, which enable the signals to pass into interior space partitioned by biomembranes. These series of signal transduction provide homeostasis to our body. Therefore, it is natural that many drugs target the receptors.

Biomembranes are comprised of lipid molecules such as phospholipids. These lipids form sheetlike lipid bilayers with all hydrocarbon chains sandwiched between the hydrophilic head groups. These hydrocarbon chains provide biomembranes with a 2-4 nm-thick hydrophobic core. The hydrophobic core functions as an impermeable barrio to hydrophilic molecules. Drug molecules need generally pass through the barrio, once at least, in order to reach their targets. Moreover, the bilayer structure provides a cell with a capability of a resistance to a variety of changes of an exterior environment such as ionic strength and pH. Certainly, the bilayer works as a physical barrio to invasion of other cells and other lives (Figure 1.1 (a)). Since the most abundant lipids in cell membranes are phospholipids, the term “phospholipid bilayer” is generally used to designate the bilayer. Electron microscopy of thin cell sections stained with osmium tetroxide, which binds strongly to the phosphate groups of phospholipids, can reveal the bilayer structure (Figure 1.1 (b)).

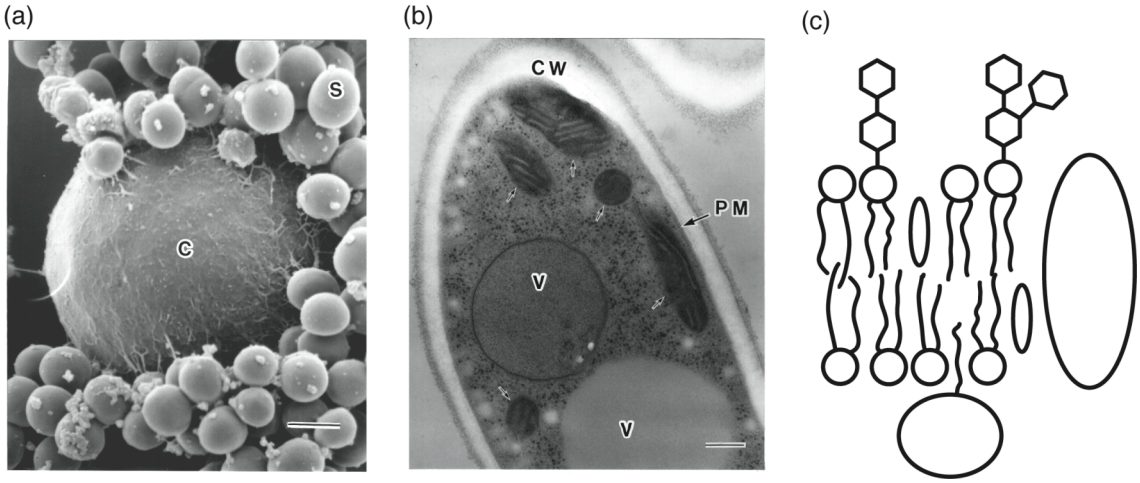


Figure 1.1 Cells and Bilayer structure of biomembranes.

(a) Scanning electron micrograph (SEM) of a *Cryptococcus neoformance* surrounded with the crowd of *Staphylococcus aureus* (C, *C. neoformance*; S, *S. aureus*). Scale is 1 μm. (b) Transmission electron micrograph (TEM) of a thin section of *Fusarium oxysporum* (CW, cell wall; PM, plasma membrane; V, vacuole). Scale is 0.2 μm. (c) Schematic representation of a biomembrane.

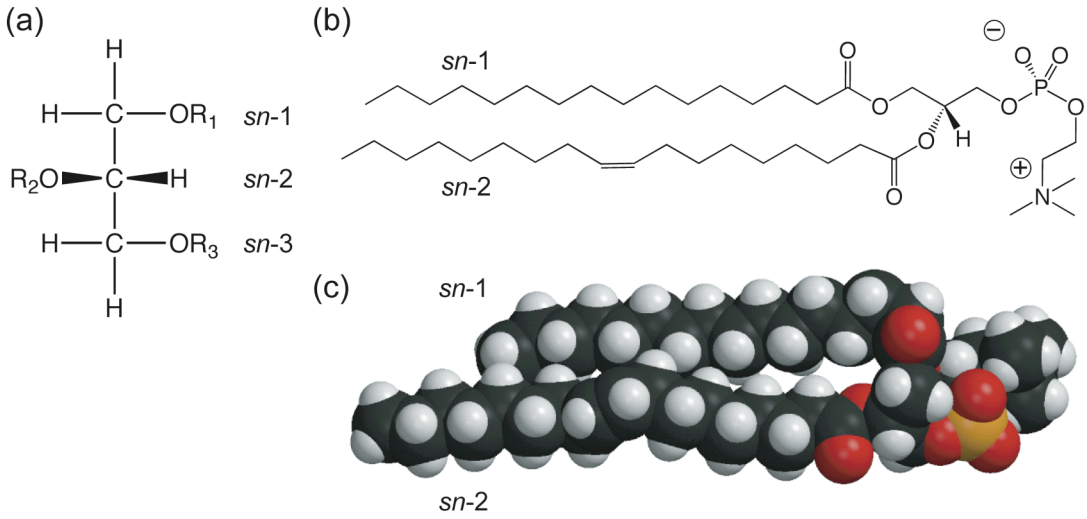


Figure 1.2 Nomenclature of a tri-substituted L-glycerol.

(a) Nomenclature of a tri-substituted L-glycerol. Glycerolipids have commonly two acyl chains in *sn*-1 and *sn*-2 positions, and a polar head group in *sn*-3. (b) The structure of 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphocholine (POPC), the most common glycerolipid. Molecular graphics were created by MolScript[2] and Raster3D.[3]

I.2 Lipids in Biomembranes

Each layer, called leaflet, consists of many types of lipids. Three types of lipids are common in general biomembranes: glycerolipids, sphingolipids, and steroids. All three classes are amphipathic (or ambiphilic) molecules with a polar (or hydrophilic) head group and a large hydrophobic tail group.[4]

Glycerolipids

Glycerolipids, the most abundant class of lipids in most membranes, are derivatives of trihydric alcohol, or L-glycerol. Each glycerolipid has a polar head group and two hydrocarbon chains attached to the two hydroxyl groups in glycerol (Figure 1.2 (a) and (b)). The two acyl chains generally differ in the number of carbons and their degree of unsaturation (0, 1, and more). In chemical terms, the resulting substituted glycerols are asymmetric and therefore can display optical activity. The stereochemistry of the glycerol can be described by the "stereospecific numbering" (*sn*) system (Figure 1.2 (a)). The *sn* system can clearly designate to which the two hydrocarbon chains are attached. For example, phosphatidylcholines, the most abundant glycerolipids in the plasma membrane, have a palmitoyl chain in the *sn*-1 and an oleoyl chain in the *sn*-2 position.

In glycerolipids, the most abundant are phosphoglycerides those are derivatives of 1,2-diacyl-*sn*-glycero-3-phosphate. Phosphoglycerides are classified according to their head groups (Figure 1.3).

Phosphatidylcholine (PC, 1,2-diacyl-*sn*-glycero-3-phosphocholine), one of the most widely distributed lipids in the plasma membrane of animals and plants, has a choline, a positively charged alcohol, esterified to the negatively charged phosphate. Hence PC is zwitterionic, but is neutral in its charge despite highly polarized. PC is often observed in exoplasmic face of biomembranes.[1]

Phosphatidylethanolamine (PE, 1,2-diacyl-*sn*-glycero-3-phosphoethanolamine) is the second most abundant lipids in animals and plants, and frequently the main lipid component of microbial membranes.[5] The head group consists of a ethanolamine that is similar to choline in its structure, but is much smaller than choline. PE exists mostly in the endoplasmic face of biomembranes.[1]

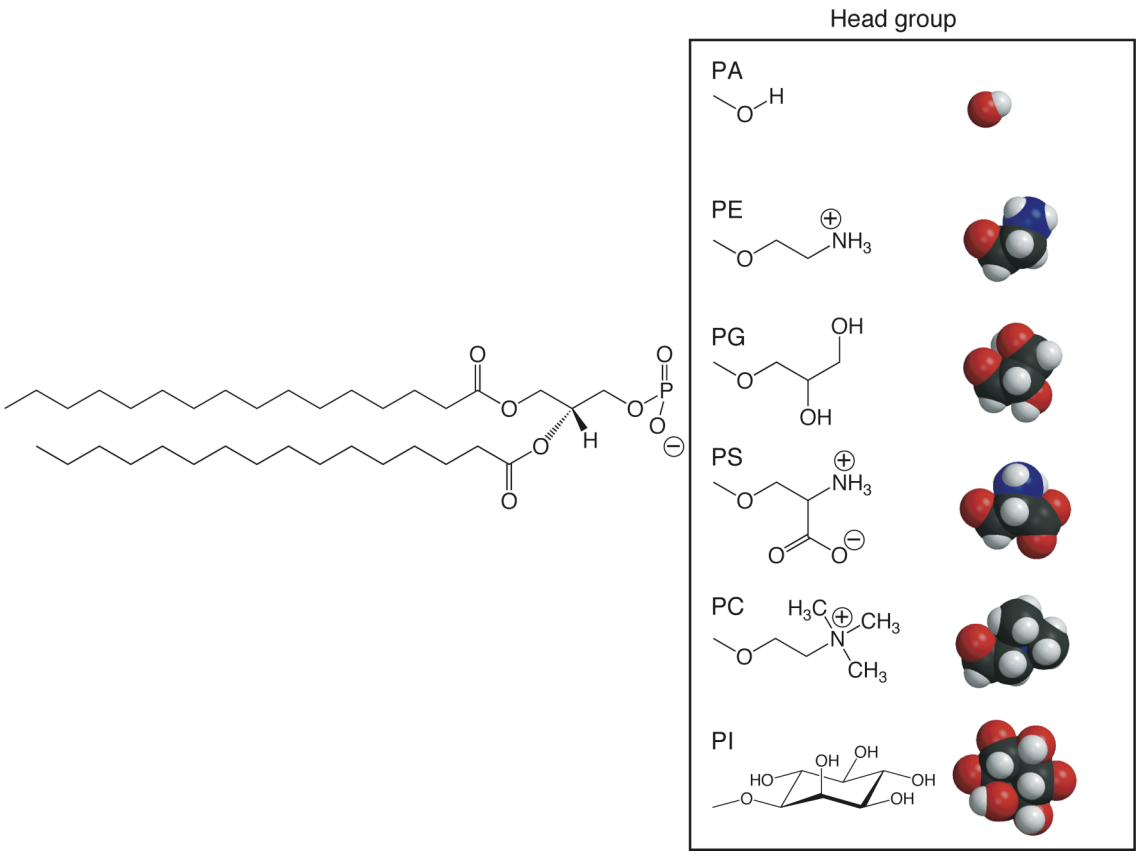


Figure 1.3 Major classes of phosphoglycerides.

The chemical structure and the molecular graphics of the head groups are shown. The volume of the head groups increases from top to down. Both of the acyl chains are represented as the palmitoyl moieties. Molecular graphics were created by MolScript[2] and Raster3D.[3]

In addition to these neutral lipids, there are many kinds of acidic (anionic) phosphoglycerides in biomembranes. Although the total amount is usually small compared with the neutral lipids, these anionic lipids have generally great significant functions for keeping cellular homeostasis.[6]

Phosphatidylserine (PS, 1,2-diacyl-*sn*- glycero-3-phosphoserine), having L-serine in its head group, is the most abundant anionic lipids. PS has an important role in the regulation of apoptosis,[7] moreover, it can affect the non-specific electrostatic interactions among membranes, ions, and membrane binding protiens.[7-9] PS is also located entirely in the endoplasmic leaflet.[1]

Phosphatidylglycerol (PG, 1,2-diacyl-*sn*-glycero-3-phospho-(1'-*sn*-glycerol)), having glycerol esterified to phosphate, is also anionic. PG is ubiquitous lipid in bacterial

membranes. For example, *Escherichia coli* has up to 20 % phosphatidylglycerol in its membranes (PE makes up much of the rest),[10] while the amount is 1-2% in most animal tissues. Moreover, PG is the second most abundant phospholipid in lung surfactant of human, whose concentration reaches up to 20 %.[11] But, its precise function has not been unveiled.

Phosphatidylinositol (PI, 1,2-diacyl-*sn*-glycero-3-phospho-(1'-*myo*-inositol)) is one of the most biologically important lipids in animals. PI has a sugar derivative inositol in its head group. The stereochemical form of inositol is *myo*-D-inositol with an axial hydroxyl group in position 2 with the others equatorial. Inositol can be often enzymatically phosphorylated (commonly in position 3, 4, and 5).[12] These phosphorylated derivatives, phosphatidylinositol phosphates (PIPs), are maintained at a steady state level in the endoplasmic leaflet, and have critical roles in signal transduction in the cell. Hydrolysis of PIPs by lipases such as the phospholipase C leads to generation of *sn*-1,2-diacylglycerols (DG) and inositol phosphates, both of which act as second messengers in the cell. PI is also important as a building block for glycosylphosphatidylinositol (GPI)-anchored proteins.[1,13]

Phosphatidic acid (PA, 1,2-diacyl-*sn*-glycero-3-phosphate) is generally a minor component of cells, but it is a key intermediate in the biosynthesis of all other phospholipids.[14] It is known to have signaling functions in animal cells, by specific binding to particular proteins.[9]

In addition to diacyl forms, phosphoglycerides with ether and vinyl ether bonds in *sn*-1 position are widely distributed in animal tissue. In these ether phosphoglycerides, the vinyl ether class, or plasmalogen forms predominate. These molecules constitute about 20 percent of the total phosphoglyceride content in humans. Plasmalogens are especially high in human brain and cardiac tissue, and immunological cells.[15,16] Nonetheless, it has been veiled what functions they have in these tissues through the additional chemical stability of the ether linkage and subtle differences in their three dimensional structures. Beside animals, human immunodeficiency virus (HIV) contains much amount of plasmalogen types of PE.[17]

Sphingolipids

Sphingolipids are the second major class of lipids in biomembranes. The sphingolipids comprise a complex range of lipids in which fatty acids are linked via amide bonds to a long-chain base or sphingoid base (Figure 1.4).[18] The base is long-chain aliphatic amines, containing two or three hydroxyl groups, and often a *trans*-double bond in position 4. To be more precise, they are 2-amino-1,3-dihydroxy-alkanes or alkenes with

(2S,3R)-*erythro* stereochemistry. The most abundant form in animal tissues is sphingosine, (2S,3R,4E)-2-amino-4-octadecen-1,3-diol or 4-sphingenine, with a C18 aliphatic chain, hydroxyl groups in positions 1 and 3, and an amine group in position 2. The double bond in position 4 has the *trans* configuration. The fatty acid constituents bonded to the amine group of the sphingoid base tend to be long, commonly ranging from 16 to 24 carbon atoms. These fatty acid chains are less unsaturated; moreover, hydroxylation in α position of the fatty acids sometimes occurs.

Sphingolipid is also classified according to the nature of its head group.

Ceramide (CER), having a hydroxyl moiety in its head group, is one of the simplest sphingolipids. Ceramides play roles in both structure and signaling, and relate to the regulation of apoptosis, cell differentiation, transformation, and proliferation.[19] Furthermore, the stratum corneum, the outer-most layer of the skin, contains relatively high levels of ceramides, which inhibit the vaporization of water from the skin.[20] Therefore, ceramides are also studied vigorously in the cosmetic and the dermatological fields.

In sphingomyelin (SM), one of the most abundant sphingolipids, phosphocholine is attached to the terminal hydroxyl group of sphingosine (Figure 1.5). Although SM has the same head group as PC, the sphingosine backbone, having both more hydrogen bond donors and acceptors than PC, enables SM to make extra intermolecular hydrogen bonds.[21,22] In addition, the extra hydrogen bonds allow SM to interact quite strongly with cholesterol, which is believed to drive a formation of lipid microdomain or "lipid raft".[23]

Other major class is glycosphingolipids (GSL) in which a single sugar residue or a branched oligosaccharide is attached to the terminal hydroxyl group of sphingosine (Figure 1.5). Cerebrosides, having a single galactose or glucose unit, are the principal GSL in the brain tissue. More complex GSL are gangliosides with a branched oligosaccharide containing one or more N-acetylneuraminic acids (sialic acids or SA or Neu5Ac). As a result, gangliosides are acidic so that they can have net negative charges on their head groups at pH 7.0. One of the common monosialo-gangliosides GM1 is illustrated in Figure 1.5. Other di- and oligoglycosphingolipids, such as Lactosylceramides, are found in animals, plants, and microorganisms, which play specific functions in the particular tissues.[4]

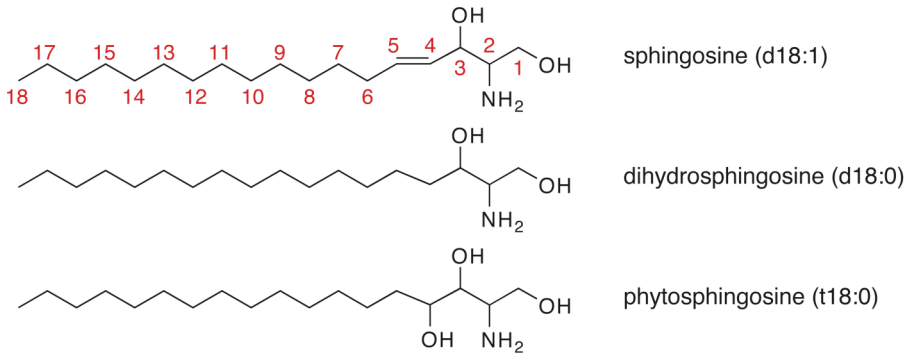


Figure 1.4 Nomenclature of sphingoid bases.

The chemical structures of the major three sphingoid bases follow their names. Nomenclature is similar to that of the fatty acid; the number of carbon atoms and the number of double bonds are denoted with the prefix 'd' or 't' to designate di- or tri-hydroxyl bases.

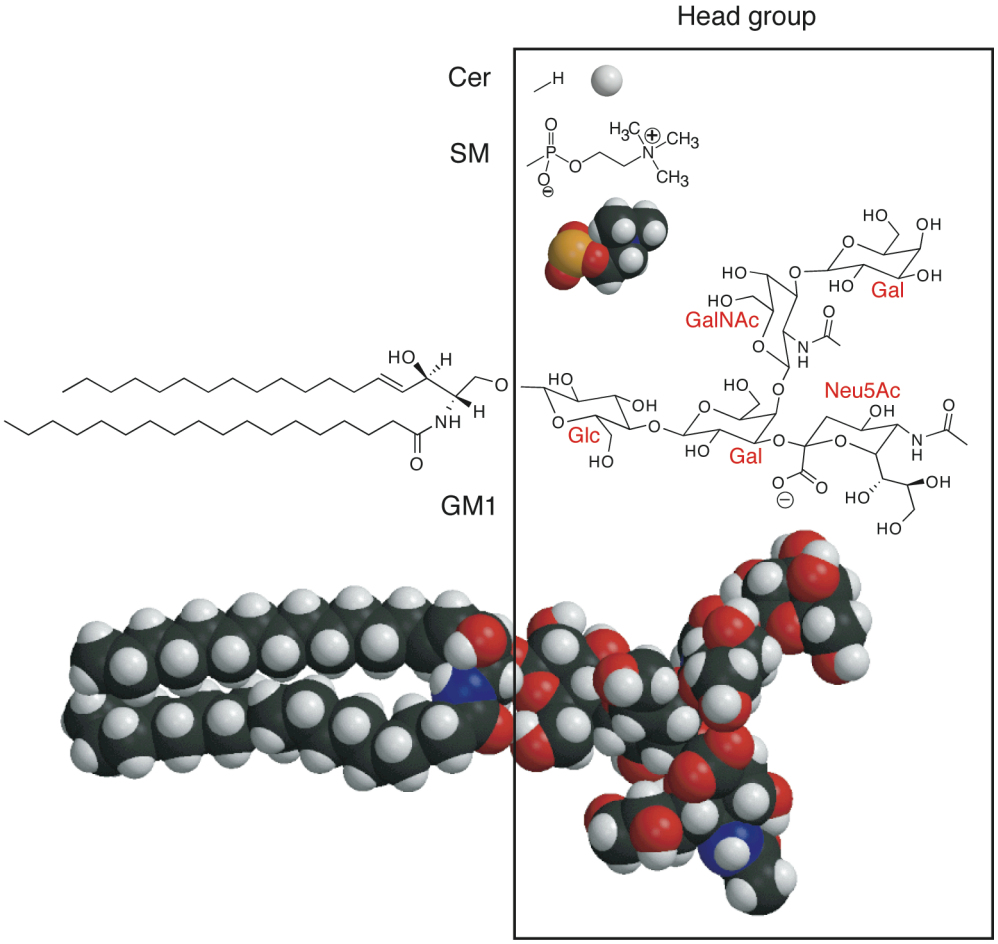


Figure 1.5 Major subclasses of sphingolipids.

The chemical structure and the molecular graphics of the head groups of ceramide (Cer), sphingomyelin (SM), and GM1 ganglioside are shown. Overall structure of GM1 is drawn particularly. Both of the acyl chains are represented as 18:0. Molecular graphics were created by MolScript[2] and Raster3D.[3]

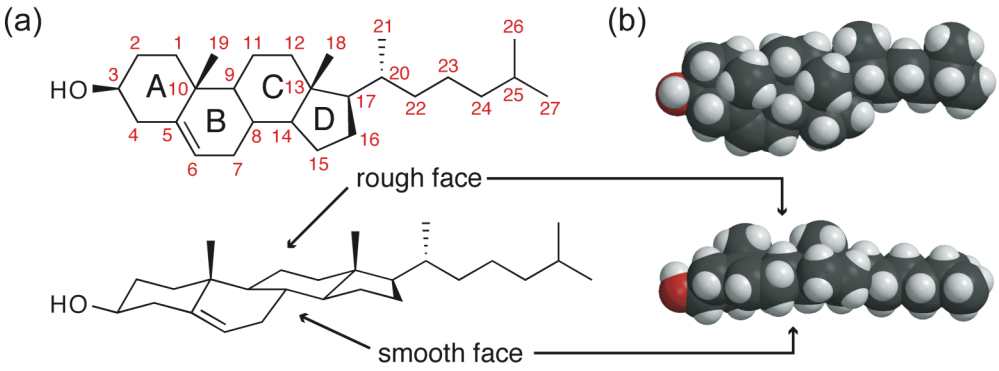


Figure 1.6 Structure of cholesterol.

(a) The chemical structures of cholesterol. The numbering system is also shown in red figures. (b) The molecular graphics of cholesterol (upper row, top views from the rough face; lower row, side views). Arrows indicate the rough and the smooth faces. Molecular graphics were created by MolScript[2] and Raster3D.[3]

Cholesterol

Third principal component of biomembranes is cholesterol and its derivatives. Cholesterol has a steroid backbone, which is a tetracyclic cyclopenta[a]phenanthrene (A, B, C, and D) with an iso-octyl substituent at carbon 17, methyl groups at carbon 10 and 13, and hydroxyl moiety at carbon 3 (Figure 1.6). As a result, the cholesterol molecule has a rough face and a smooth face in its structure. These faces make a difference in the preference for the interactions with lipids whether the hydrocarbon tails are saturated or not.[24] Cholesterol is abundant in the plasma membranes of mammalian cells, and has various physiological functions.[25,26] In plasma membranes, cholesterol is believed to form lipid raft,[23,27] and relate to cellular signaling described in the following sections.

I.3 Phase of Lipid bilayers

Lipid bilayers have various phases as shown in Figure 1.7. The nomenclature of phases of lipid bilayers is unique and does not relate to any terms used in the other fields, which sometimes causes confusion. L_c phase is a crystalline phase, in which the degrees of freedom of both the head groups and the lipid tails are restrained completely, in which all the lipids align with order in two dimensions. L_β phase means a gel phase, in which the degrees of freedom of the head groups are allowed but those of the lipid tails except for the rotation are restrained. L_β' phase is also called a tilted-gel phase, in which the lipid tails incline together with a specific tilt angle. $L_\beta I$ phase is called a interdigitated gel phase, in which the lipid tails interlocked with the lipids from the opposite leaflet, resulted in a remarkable decrease of membrane thickness and a conspicuous increase of area per lipid.

L_α phase means a liquid crystalline phase. This term does not relate to the term "liquid crystal" in other fields at all. In this phase, the head groups and the lipid tails are disordered with keeping a bilayer structure clearly. In $L_\alpha I$ phase, or interdigitated liquid crystalline phase, lipids tails insert themselves into the opposite leaflet partially or completely, which makes a bilayer structure unclear.

Cholesterol (Chol) makes these phases more complicated. A characteristic phase at addition of Chol is a liquid ordered phase, or l_o phase. In this phase, the lipid tails have ordered conformations like L_β phase, but their lateral diffusion is rapid like L_α phase. Hence, Chol makes the boundary between L_α and L_β phases ambiguous. These functions of Chol are due to its dual effects on lipids. When cholesterol is added in L_α phase, it compensates for the lipid-lipid interactions, which leads the membrane to have more packed structure. In contrast, added in L_β phase, cholesterol disturbs the packed structure and enlarges the distances among the adjacent lipids.

These phases of lipid bilayers can be changed with ion concentration, pH, pressure, water ratio, and interactions with other molecules such as alcohol and proteins. The phase transition relates deeply to the interactions among lipids.

There are other phases like P_β phase emerging between L_α and L_β phases and various forms other than bilayer structure like cubic phase and hexagonal II phase in three dimensions. In the cell, the phase of biomembranes is strictly controlled, because fluidity and flexibility of biomembranes are so critical for exerting adequate functions that lack of them breaks biomembranes and leads to cell death.

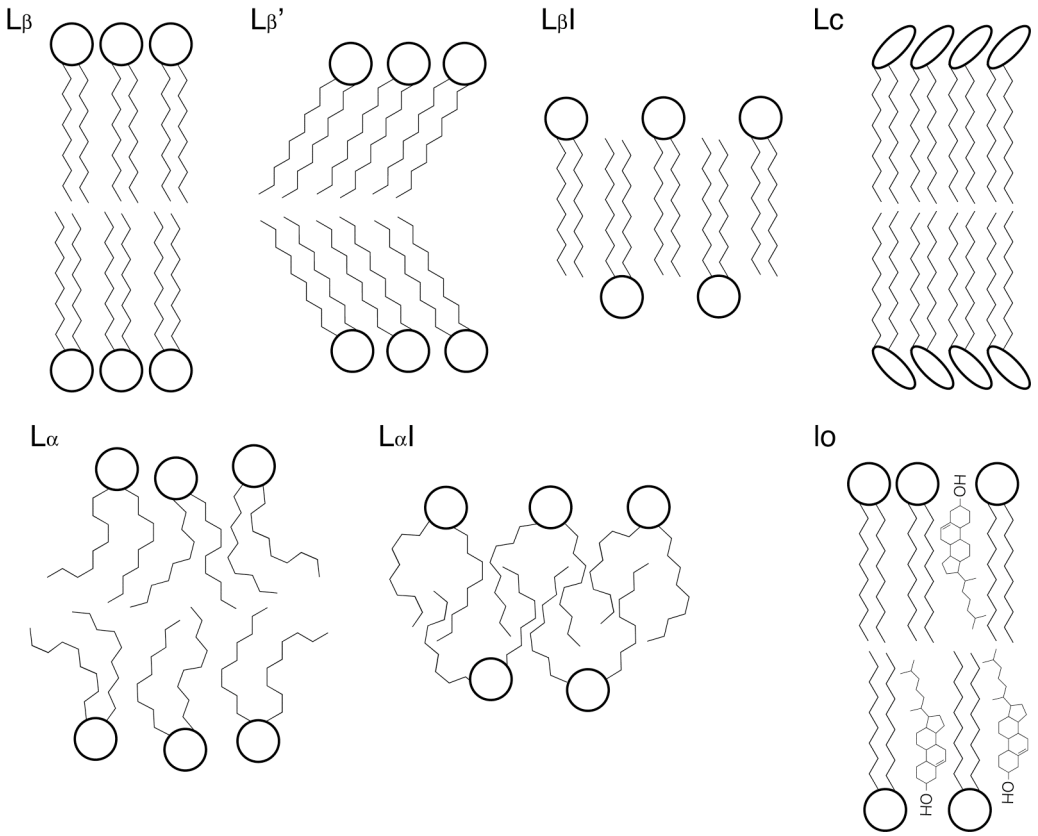


Figure 1.7 Phases of lipid membranes.

L_β: gel phase, or solid ordered phase, s_o, L_β': tilted gel phase, $L_\beta I$: interdigitated gel phase, L_c: crystalline phase, L_α: liquid crystalline phase or liquid disordered phase: l_d, $L_\alpha I$: interdigitated liquid crystalline phase, l_o: liquid ordered phase

I.4 Lipid Composition and Lipid microdomain

Figure 1.8 shows a schematic representation of a biomembrane. Usually, lipids in biomembrane distribute asymmetrically in its exoplasmic and endoplasmic leaflets. As shown in Figure 1.8, an exoplasmic face includes a profusion of glycosphingolipids and lipids with bulky head groups, while an endoplasmic face has many lipids with small head groups, negative charges, PIs, and so on. This asymmetric distribution is so essential for membrane proteins and cells to function appropriately that it is highly regulated in the cell.[28-30] For example, PS is usually rich in the endoplasmic face of a plasma membrane, but in the apoptosis, its distribution is rapidly changed and PS is exposed in the exoplasmic face. Accordingly, PS exposure enables other apoptosis-relating proteins to recognize the cells sending the apoptosis signal.[7]

In addition to the asymmetric distribution across the leaflets, lipids in biomembranes are segregated laterally within a leaflet. For example, SM and Chol interact so strongly

with each other that form a more ordered, less fluid membrane patch surrounded by other fluid lipids.[23] Moreover, lipids in the vicinity of membrane proteins are also different in their structure from other distant lipids. These adjacent lipids are called "annular lipids", and show specific interactions with membrane proteins.[31] The membrane proteins attract particular lipids around them, and create discrete membrane patches in a bilayer. Collectively, these laterally separated domains are called "lipid microdomain". In particular, the term "lipid raft" is used for the microdomain in which SM and Chol concentrate. Lipid microdomain does not have a fixed structure, but changes dynamically its size (an average size is believed from 1 nm to hundreds nm). Lipid microdomain gathers some kinds of receptors and proteins related to signal transduction.[27] Therefore, the lipid microdomain is believed to play important role on the signaling efficiency, and to relate with some diseases.[32] Until now, many combinations of lipids are known to form the microdomain,[5] nonetheless, how lipid microdomain arises is still unclear.

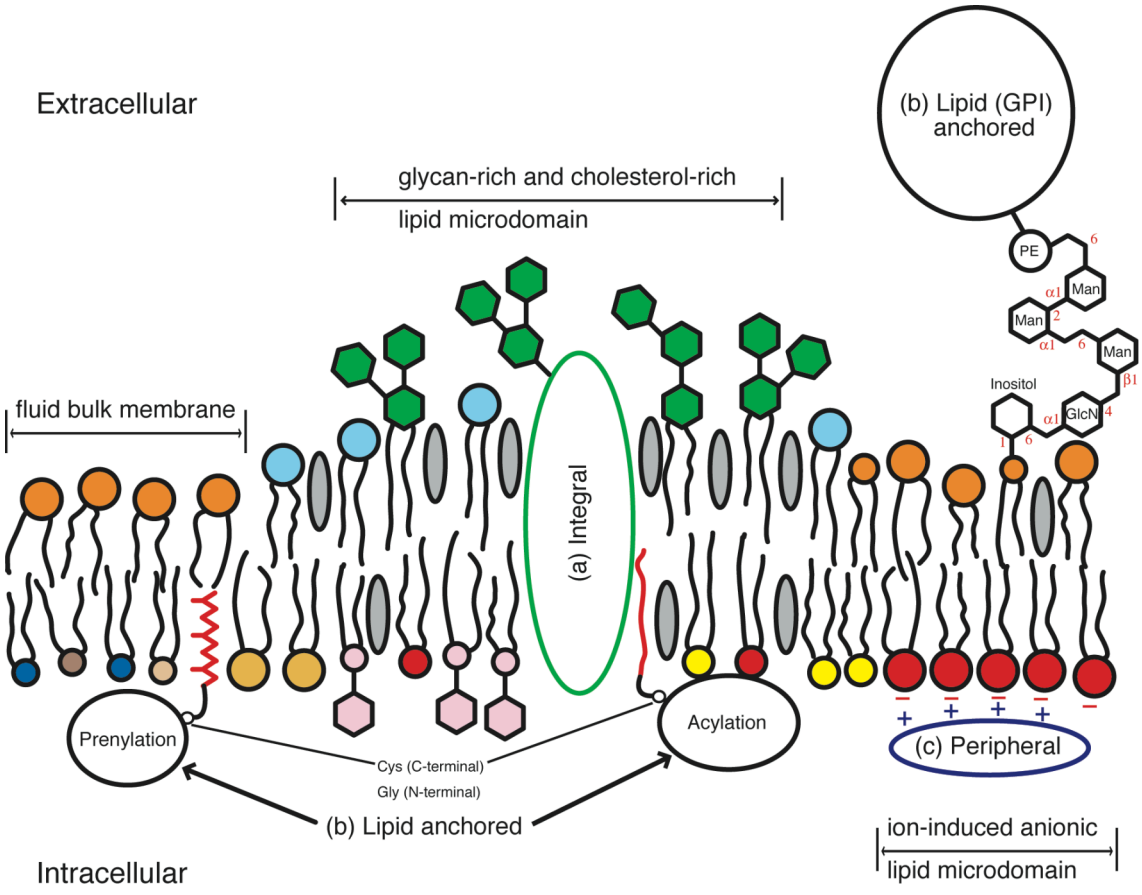


Figure 1.8 Schematic representation of biomembrane.

Biomembrane has many types of patches called lipid microdomains and has asymmetrical composition of lipids in the endoplasmic and the exoplasmic faces. This figure is drawn only for explanation about the heterogeneity of the biomembranes, so the details such as relative position of lipids are not considered.

I.5 Membrane Proteins

Membrane proteins can be classified into three classes according to the way of the interaction with membranes; integral, lipid-anchored, and peripheral (Figure 1.8).[1]

Integral membrane proteins (Figure 1.8 (a)), or transmembrane proteins, span a lipid membrane. Integral membrane proteins are further classified by the number of the span; a single-pass or a multipass transmembrane protein. A single-pass membrane protein is exemplified by cytokine receptors and receptor tyrosine kinases. A typical example of multipass transmembrane proteins is a seven-pass G-protein coupled receptor.

Lipid-anchored membrane proteins are bound to membranes by covalently attached lipid molecules (Figure 1.8 (b)). These proteins are classified into three groups according to the types of lipid anchors. A first group is anchored to the cytosolic face by fatty acyl groups attached to N-terminal glycine or C-terminal cystein. Myristoylation (14:0) and Palmitoylation (16:0) exemplify these acylation. A second group is also cytosolic ones anchored by prenylation of a cysteine residue at or near the C-terminus. In these proteins, a farnesyl (C15) or a geranylgeranyl (C20) moiety is bound through a thioester bond. The first and the second types of the lipid anchors are sometimes attached to one protein. For example, a small G-protein H-Ras is farnesylated at the 186 cystein and is further added two palmitoyl moieties to the 181 and 184 cysteines. Its addition and removal are enzymatically controlled, which regulates the localization of H-Ras in the cell.[33] A third group is bound to membranes by GPI (glycosylphosphatidylinositol) anchors. GPI anchor has a PI as a scaffold, and several sugar residues (glucosamine and mannose) attached to the inositol moiety of PI, and a phosphatidylethanolamine attached to the sugar. A protein is connected to a GPI anchor by an amide bond between the protein's C-terminus and the amino group of the ethanolamine. Additional sugar moieties can be attached to the core GPI backbone. The GPI-anchored proteins are localized at the exoplasmic face of the membranes.

Peripheral membrane proteins are also bound to the membrane surface, but do not insert themselves into the hydrophobic core of the membranes (Figure 1.8 (c)). The main force attaching the peripheral proteins to biomembranes is due to the electrostatic interactions. The peripheral proteins usually have a cluster of positive charges at their surface. These charges force the proteins move to the negatively charged lipids at the endoplasmic face. Localization of the peripheral proteins is often regulated by kinases. They add some negative charges to a peripheral protein by phosphorylation, which weaken the electrostatic interaction between the protein and the membrane.

I.6 Molecular Dynamics Simulations of Biomolecules and Biomembranes

Molecular dynamics (MD) simulations are powerful tools in order to investigate the time-series dynamics of biomolecules in the atom scales. MD simulation requires two things to start; parameters and an initial structure.

In MD simulations, some approximations are applied to; an atom is represented as a hard sphere with a single point charge; a covalent bond and an angle formed among three connected atoms are described as springs with the specific spring constants; moreover, interatomic forces are calculated based on the van der Waals (vdW) interaction energy and the electrostatic energy between any two non-bonded atoms. These point charges, spring constants, and vdW coefficients are called "force field parameters", and a system described by them is called "force field". AMBER[34] and CHARMM[35] exemplify the extensively used force field parameters (and the two terms also mean their names of the MD simulation program packages).

Another requirement to start MD simulations is getting an initial structure of a target biomolecule. In the case of a protein, its initial structure can be obtained from the Protein Data Bank (PDB) as a text file containing the three-dimensional coordinates of the atoms determined by X-ray crystallographic analysis or NMR. In contrast, there is little information about the coordinates of lipids and saccharides enough to carry out MD simulations of biomembranes, and there is no database about them. Hence, alternative way to get the initial structure must be taken; molecular modeling. However, the molecular modeling of biomembranes is a quite time-consuming task even for experts of MD simulations, because of the complexity of the biomembranes as mentioned above.

As a result, MD simulations in the biological field have been intensively focused on proteins so that most of the force field parameters were developed and optimized for proteins.[36,37] As for biomembranes, MD simulations of them have been more significant recently,[21,22,38-45] however, they are now still in a developmental stage.

I.7 This Thesis

This thesis describes the following three main subjects;

1. Molecular modeling of biomembranes and membrane-protein complexes. (*Chapter II*)

2. Molecular dynamics simulations of lipid bilayers. (*Chapter III, IV,* and *V*)

3. Molecular dynamics simulations of membrane-protein complexes. (*Chapter VI* and *VII*)

In the first topic, the computer program named GLYMM (GLYcan and Membrane Modeling) is introduced, which has been developed in our laboratory. GLYMM is a GUI (Graphical User Interface) based computer program that can build a three dimensional structure of complex biomembranes automatically. GLYMM is a GUI-based program, to say, what is needed is only clicking mouse buttons. Therefore, concentration of a specific lipid is graphically adjustable with only clicking the mouse buttons. GLYMM also enables molecular modeling of glycans, or sugars. Glycans require tremendous efforts for molecular modeling because of their branched complex structures. GLYMM is capable of reading a glycan sequence and predicting a possible structure in a computer. In addition, GLYMM provides automatic modeling of a membrane-protein complex. Up to now, embedding a protein in a biomembrane has been usually carried out manually. Details and theoretical background of GLYMM are described in *Chapter II*.

The second topic is MD simulations of three types of lipid bilayers;

Lipid bilayers with lipids distributed symmetrically in endoplasmic and exoplasmic leaflet. (*Chapter III*)

Lipid bilayers with PC, PE, and PS distributed asymmetrically in endoplasmic and exoplasmic leaflet with and without cholesterol. (*Chapter IV*)

A lipid microdomain model with ganglioside GM1, sphingomyelin and cholesterol. (*Chapter V*)

In *Chapter III*, The results of MD simulations and comparative analyses of 35 kinds of glycerolipid bilayers and 48 kinds of sphingolipid bilayers are shown. Such large-scale

computational comprehensive analyses of lipid bilayers have never been carried out in the world. In particular, we investigated water-membrane interfaces in detail. Water-membrane interface is the important region, where small molecules, such as antimicrobial drugs and small peptides, interact and adhere. Some quantities of the interfaces are described in *Chapter III*.

In *Chapter IV*, MD simulations of two kinds of asymmetric lipid bilayers are described.[40] The two models were based on the human erythrocyte membrane, lipid composition of which had been already known.[28,29] It is a unique case in that such biological models based on the actual lipid composition are used in MD simulations.

Chapter V describes ganglioside GM1-containing lipid bilayers. GM1 forms glyco-cluster in the presence of SM and Chol. Recently, it was revealed that GM1 cluster was recognized by amyloid β (Aβ) peptides, one of the main factors for Alzheimer's disease, and its aggregation occurred on the GM1 cluster. How do the Aβ peptides recognize the GM1? And what structural characters does the cluster have? The results from the MD simulations of GM1-containig bilayers, described in *Chapter V*, provide a key to understanding these questions.

The last two chapters, *Chapter VI* and *VII*, focus on the dynamics of small G-proteins. Small G-proteins are involved in many cellular functions such as signal transduction, membrane trafficking, and cytoskeleton.[46] Four of the five superfamilies (Ras, Rho, Rab, Arf, Ran) of small G-proteins are lipid-anchored membrane proteins. Their dynamics and interaction modes on membrane have been one of the most attractive targets in the molecular biology, and many studies revealed their interaction networks, localizations, and trafickings.[46] However, the physical aspects of their dynamics are little understood. At first, MD simulations of five small G-proteins in solution were carried out (*Chapter VI*).[47] Then, MD simulations of H-Ras on water-membrane interface were carried out (*Chapter VII*). The used lipid bilayers were based on the plasma membrane of a human red blood cell, the same one described in *Chapter IV*. The results revealed their dynamics and interactions with lipids clearly in the atom scale.

II

Molecular Modeling of Lipid Bilayers and Membrane proteins

Abstract

Molecular Dynamics (MD) simulations must require coordinates, usually Cartesian coordinates, of a target molecule to be calculated. However, coordinates of lipid bilayers and membrane proteins embedded in lipid bilayers are so scarce that a model structure needs to be manually constructed every calculation. The modeling process usually becomes a time-consuming step in MD simulation. This barrier prevents us from carrying out MD simulations of various types of lipid bilayers. Here, we developed a computer program called GLYMM that can construct a 3D structure of a glycan and a lipid bilayer automatically in computer. Concentration of a specific lipid is also graphically adjustable. In addition, insertion of a protein into a lipid bilayer was also programmed. GLYMM provides rapid modeling of lipid bilayers and membrane-protein complexes.

II.1 Introduction

As described in *Chapter I*, biomembranes usually have asymmetric structures built up with a mixture of various lipids. Owing to the complexity of biomembranes, membrane proteins exert their intrinsic functions effectively. Therefore, MD simulations of membrane proteins should consider the complexity of biomembranes, in particular, in the case of a protein existing in a lipid microdomain.

MD simulations of membrane proteins have been increasing more and more, however, the membranes used in the simulations often consisted of PC or PE only. MD simulations must require coordinates, usually Cartesian coordinates, of a target molecule to be calculated. However, the coordinates of lipid bilayers are so scarce that a model structure needs to be manually constructed before every calculation. This situation tells a reason why only PC and PE are used so frequently; the reason is just it is much easier to get coordinates of PC and PE membranes than other lipid membranes. When I directly asked a researcher performing a simulation of Ras on POPE membrane at an annual meeting why you chose the POPE membrane for the media surrounding the protein, then he answered, "Just because it was in our lab". Hence, even for an expert of MD simulation, one of the biggest problems is the difficulty to get the coordinates of a membrane.

In order to make it easier, we developed a computer program called GLYMM (GLYcan and Membrane Modeling) that can construct a 3D structure of a glycan and a lipid membrane automatically. Concentration of a specific lipid is also graphically adjustable. In addition, inserting of a protein into a membrane was also programmed. These computer programs provide rapid modeling of lipid bilayers and membrane proteins.

II.2 About GLYMM and Membrane Modeling

GLYMM is a GUI (graphical user interface) based program developed as a plug-in for VMD,[48] one of the most famous molecular viewer programs. GLYMM is written in Tcl/Tk, one of the script program languages. Because Tcl/Tk is independent of a machine platform, GLYMM can be run on any operating system, e.g. Windows, Mac, Unix, and Linux. GLYMM is composed of three main functions: a basic membrane modeling function (Figure 2.1), a modeling function for more complex bilayers "*Membrane Editor*" (Figure 2.1), and a glycan modeling function (Figure 2.2).

The basic modeling function has a single main window with some text boxes and buttons (Figure 2.1 (b)). Composition and asymmetric distribution of lipids can be chosen in this window. Pushing the "make membrane" button at the bottom works automatically the *Psfgen* and the *MEMBRANE* plug-ins in VMD, which results in making a topology (PSF) file and a coordinate (PDB) file of the membrane, both of which are required for the MD simulation program NAMD.[49] Lipids that can be used in this window are only PC, PE, PS, SM, CER and Chol. Moreover, only 16:0/18:1 (1-palmitoyl-2-oleoyl, or PO) is allowed for glycerolipids and 18:0 (stearoyl) for sphingolipids. In the case of modeling a bilayer composed of more than two kinds of lipids, each lipid is distributed to fulfill a given concentration of the lipid.

The second function is "*Membrane Editor*". *Membrane Editor* is composed of two sub-functions. The first sub-function has a capability of transforming POPC to another one, for example, POPC to DPPE, and POPC to lauryl-SM and so on. In addition, *Membrane Editor* has two canvases for showing lipid positions (Figure 2.1 (c)). Using the canvas, one can change POPC into any lipid on the list at the right side of the canvas with only drawing and clicking mouse buttons. Therefore, any types of lipid bilayers composed of tens of lipids can be modeled, such as lipid microdomain (Figure 2.4). The transformation of lipids is carried out using the *Psfgen* plug-in in VMD. The second sub-function of *Membrane Editor* is for embedding a protein in a membrane, which is described in detail in the following section.

The glycan modeling function initially reads a glycan sequence (KCF and CCSD) file, and then construct its initial structure in VMD according to the sequence, which simultaneously makes its PSF and PDB files (Figure 2.2). The modeled glycan structure with its PSF and PDB files can be used for MD simulation of itself. A KCF and a CCSD files can be obtained from "GLYCAN", the database of glycans on the GenomeNet server at Kyoto University[50] (Figure 2.2 (a)). A KCF file includes "node" coordinates in 2D and "edge" information, which are consistent with "residue" and "bonding mode", respectively. The terms, node and edge, originate in the graph theory. According to the graph theory, every 2D structure of a glycan can be depicted as graph. Graphviz,[51] one of the most famous software using the graph theory, is freely available tool which can convert a graph file written in *dot* language (DOT file) into an image file such as a PNG and a postscript files. GLYMM can change a KCF file into a DOT file. Therefore, 2D structures can be quite easily drawn with GLYMM (Figure 2.2 (b)), which takes only several seconds. Furthermore, GLYMM can construct three-dimensional initial structures of glycans from KCF and CCSD files (Figure 2.2 (c)). Using a CCSD file, GLYMM separates the sequence into each saccharide residue and bonding mode, and then, constructs the three-dimensional structure with the *Psfgen* plug-in of VMD. In the case with a KCF file, GLYMM extracts the information of

residues and bonds directly from the KCF file, and builds the structure with VMD. In both case, GLYMM rotates dihedral angles (φ, ψ, and ω) of glycans and converge in a probable structure. Moreover, the modeled glycan can be attached to a membrane and a protein by only clicking a mouse button (Figure 2.2 (d)). These tools remarkably reduce time for drawing, modeling, and preparing initial structures of a glycoprotein and a glycolipid in computer.

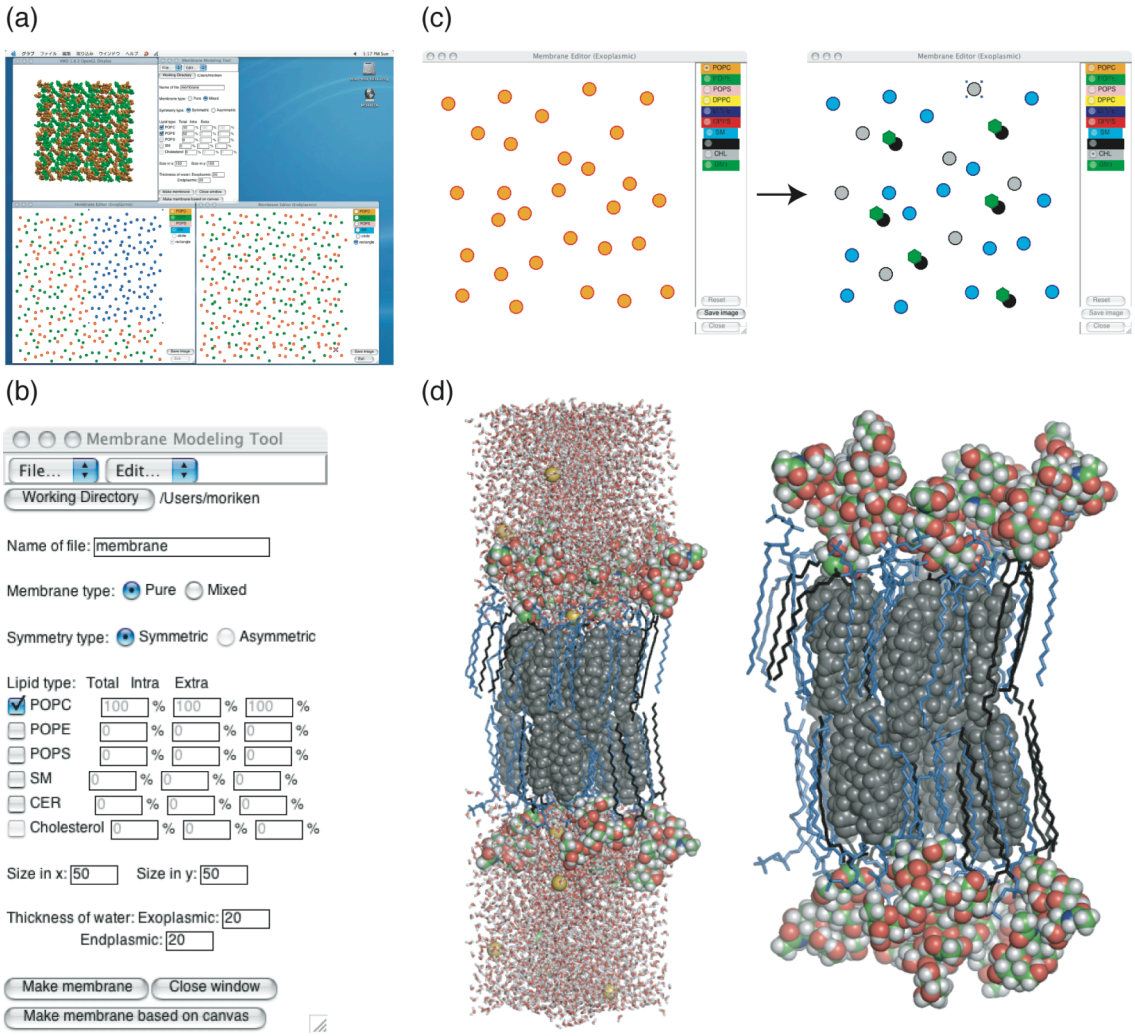


Figure 2.1 Modeling a lipid bilayer by GLYMM.

(a) Screenshot of GLYMM running on Mac OS X. (b) Main controlling panel of GLYMM (c) Panel window of GUI-based lipid converting tool "Membrane Editor" (left, an exoplasmic face of a bilayer showing the locations of POPC molecules represented as the orange circles; right, the locations of SM (cyan), GM1 (ceramide part, black; glycan part, green hexagonals), and cholesterol (gray) after the lipid conversion). (d) Molecular graphics of a modeled structure of the membrane. The lipid composition and the color scheme are the same as (c) [left, the image of all molecules in the system showing water and lipids as sticks, ions and cholesterols as spheres, glycan part of GM1 as vdW spheres colored in atom-based color (C, green; N, blue; O, red; H, white); right, the close-up of the membrane part]. The molecular graphics were drawn by PyMOL.

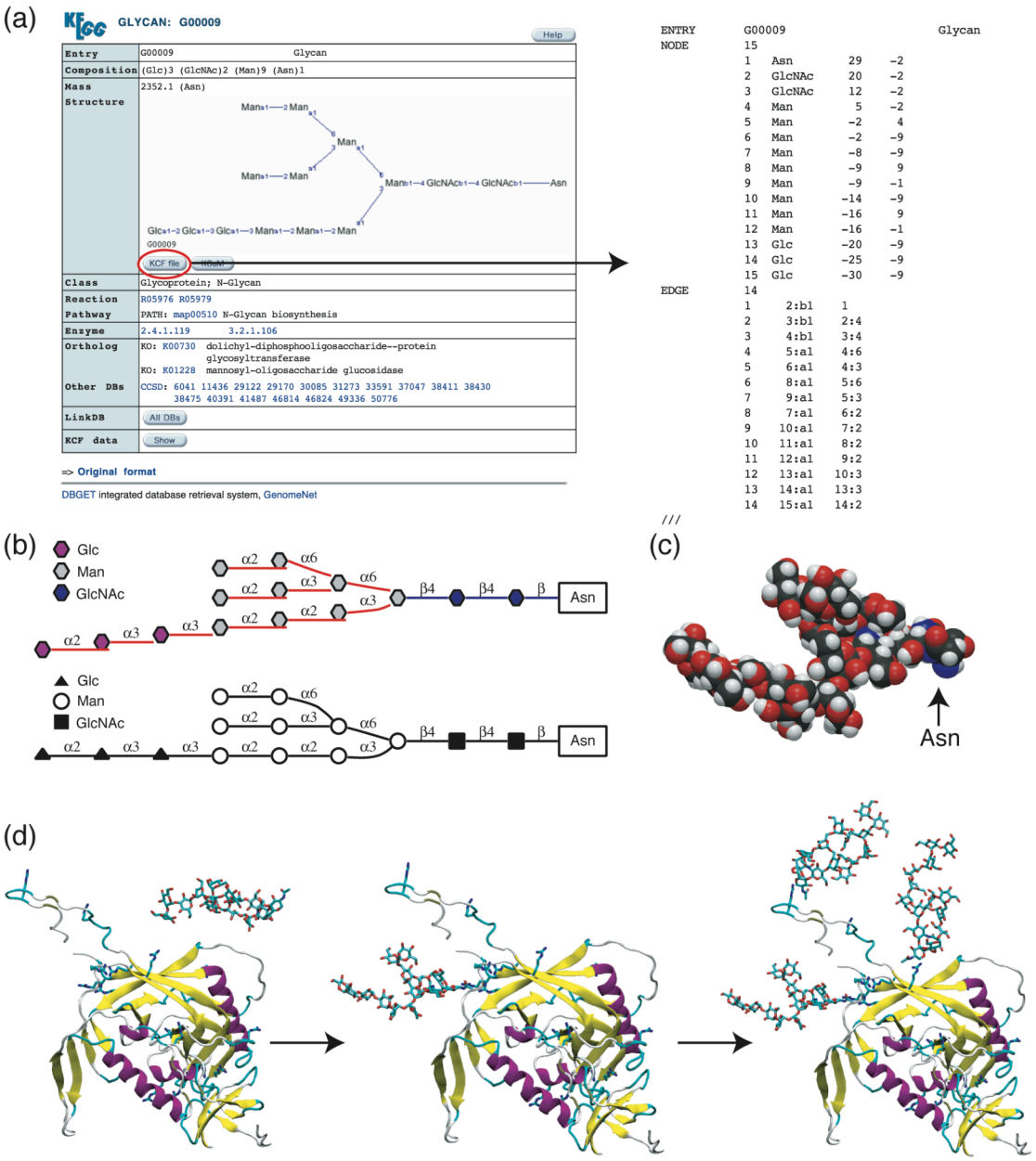


Figure 2.2 Modeling a glycan by GLYMM.

(a) Screenshots of KEGG GLYCAN database and a KCF (glycan sequence) file (ID: G00009). (b) Primary structures of G00009 created by GLYMM. The upper figure shows a 2D structure indicating relatively accurate bonding position. The coloring scheme is based on the residue type. The lower one shows the counterpart depicted with the Oxford style. (c) Ternary structure of G00009. A black arrow indicates the root, asparagine residue. The molecular graphics were created by MolScript[2] and Raster3D.[3] (d) The structure made by GLYMM can be attached to proteins by clicking a mouse button. The molecular graphics were created by VMD[48].

II.3 Embedding a Protein in a Membrane

Figure 2.3 shows how to embed a membrane protein in a lipid bilayer.

First of all, a membrane protein needs rotation so as to be embedded in a bilayer with an appropriate direction. The rotation matrix can be determined by matching the membrane normal vector with one of the principal axes of inertia of the membrane protein. The principal axes of inertia can be calculated by solving the eigenvectors of the moment of inertia tensor of the target membrane protein. For an object composed of N atoms whose positions are given in Cartesian coordinate, the inertia tensor is given as follows;

$$\mathbf{I} = \begin{pmatrix} I_{xx} & I_{xy} & I_{xz} \\ I_{yx} & I_{yy} & I_{yz} \\ I_{zx} & I_{zy} & I_{zz} \end{pmatrix} \quad \text{..... (2.1)}$$

$$I_{ij} = \sum_{a=1}^{N} m_a \left[(\mathbf{r}_a \cdot \mathbf{r}_a)\, \delta_{ij} - (\mathbf{r}_a \otimes \mathbf{r}_a)_{ij} \right], \quad \mathbf{r}_a = (x_a, y_a, z_a)^{\mathbf{T}} \quad \text{..... (2.2)}$$

where I represents an inertia tensor; a is an atom number; r_a is the coordinate of the atom; m_a is the mass of the atom; N is the total number of the atoms; δ_{ij} is Kronecker's delta; • means dot product and ⊗ means matrix product. Since this tensor is a symmetric, real matrix, it is possible to find a transformation matrix that converts I into a diagonal form. Then, the tensor has the form;

$$\mathbf{I} = \begin{pmatrix} I_1 & 0 & 0 \\ 0 & I_2 & 0 \\ 0 & 0 & I_3 \end{pmatrix} \quad \text{..... (2.3)}$$

where the coordinate axes are called the principal axes and the eigenvalues I_1, I_2 and I_3 are called the principal moments of inertia. Furthermore, unit vectors along these principal axes are usually denoted as e_1, e_2 and e_3. Each eigenvalue means potential of rotation along each principal axis (the smaller the eigenvalue becomes, the more easily the object can rotate along the principal axis). Because most of the integral membrane proteins are vertically long against the membrane, the principal axis to be fitted to the membrane normal vector is the axis with the smallest eigenvalue in order to embed the protein vertically against the membrane. These functions can be carried out using *orient* plug-in of VMD.[48]

After the rotation, it requires translation of the protein to be embedded. The translation is performed by fitting the center of mass (COM) of the transmembrane

(TM) region of the protein to the COM of the bilayer. In order to do this, the TM region must be determined at first. The surface of a TM region is generally covered with many hydrophobic amino residues. Therefore, the TM region can be searched only by determining the hydrophobic rich domain, which is calculated by moving a prove around the surface.

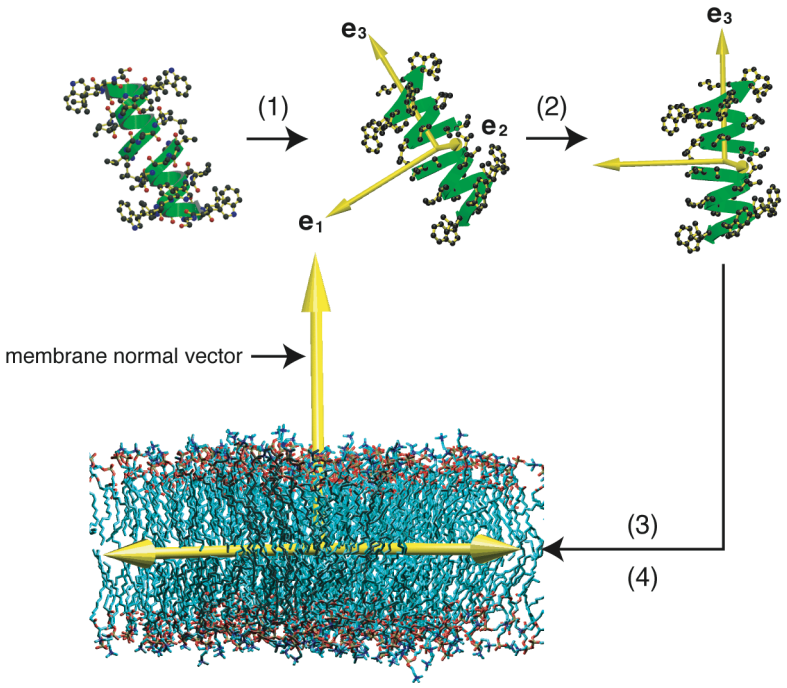


Figure 2.3 How to embed a membrane protein in a bilayer.

(1) Determination of the principal axes of inertia. (2) Rotation of the protein to make the axis parallel to the membrane normal. (3) Determination of the transmembrane (TM) region. (4) Translation of the protein to fit the center of mass of the TM region to that of the membrane. (5) Removal of overlapped lipids. Yellow arrows indicate the principal axes of inertia. The molecular graphics were created by VMD[48].

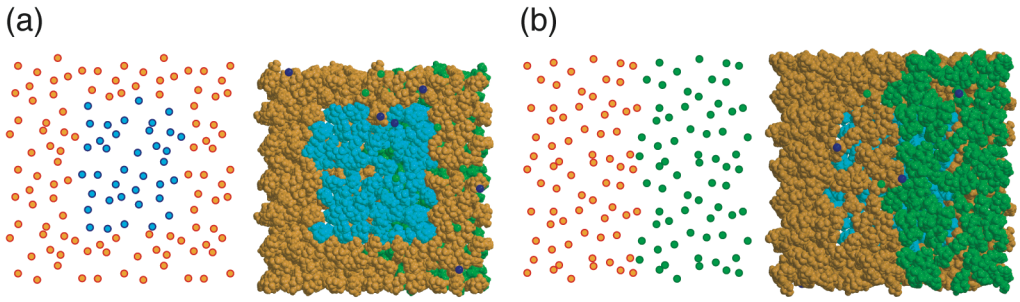


Figure 2.4 Modeled structure of a bilayer built by GLYMM.

(a) Exoplasmic face containing a microdomain of SM (cyan) surrounded with POPC (orange). (b) Endoplasmic face with a stripe pattern formed by POPC and POPE (green). The molecular graphics were created by VMD[48].

II.4 Results

The examples of membrane modeling by GLYMM are shown in Figure 2.4. GLYMM can automatically model a lipid microdomain of SM surrounded with more fluid POPC lipids (Figure 2.4 (a)). Moreover, a POPE/POPC striped pattern can be also modeled in the opposite face of the membrane by GLYMM as shown in Figure 2.4 (b). Hence, GLYMM can make any lipid bilayers with various lipid compositions with only clicking a mouse button. Figure 2.5 shows results of the 2D structures drawn by the glycan drawing tool. Even randomly selected in the GLYCAN database, every glycan can be depicted successfully. It takes only some seconds to draw them, which will be an useful tool for preparing an image for publishing and comparing structures of glycans. Moreover, when the edges are colored shown in the right column of Figure 2.5, it can be easily determined whether there is any cluster of similar edges or not. For example, it can be recognized at a glance that α anomers (red edges) are rich at the terminal of the G00010 glycan.

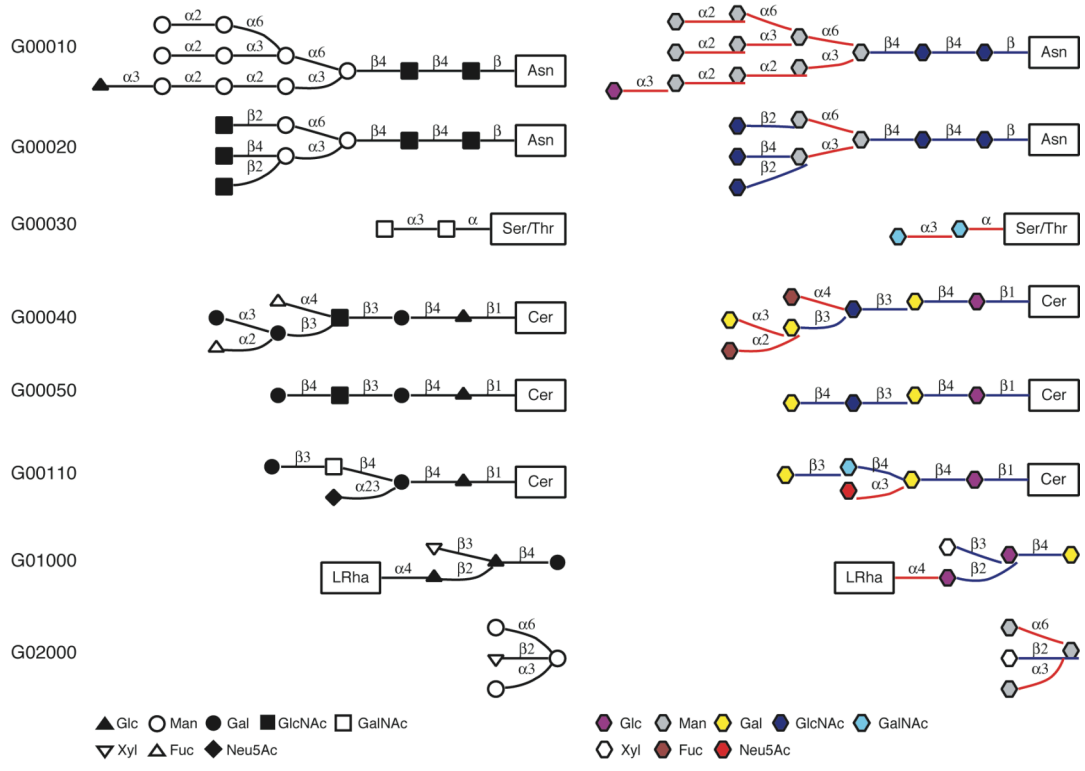


Figure 2.5 Two-dimensional drawing of glycans depicted by GLYMM.

The left column indicates each glycan ID in GLYCAN database. The middle and the right column show Oxford-style and original-style drawings.

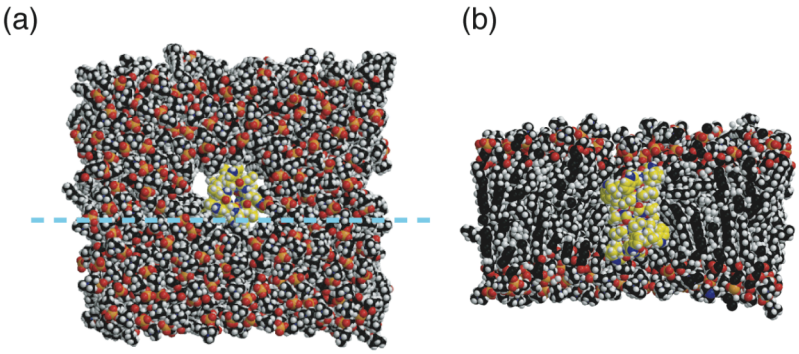


Figure 2.6 Modeled structure of a gramicidin A embedded in a DPPC bilayer.

(a) Top view. (b) Side view of the vertical section at the dashed line in (a). C atoms of proteins and the membrane are colored in yellow and black; H in white; P in orange; N in blue; and O in red. The molecular graphics were created by MolScript[2] and Raster3D.[3]

Figure 2.6 shows a result of embedding a dimer of gramicidin A (gA) in a DPPC bilayer. It can be seen that the gA is located at the center of the membrane. gA has a hole that passes water molecules and ions cross a bilayer. In the model, the hole can be observed at the center of the gA, which indicates that the gA has a proper configuration against the membrane. Hence, it demonstrates that the modeling by GLYMM can produce an appropriate initial structure automatically. Figure 2.7 shows two modeled structures of 12mer of gA proteins and a bovine rhodopsin complexed with a POPC bilayer. It can be observed that automatic modeling with GLYMM can remove lipids along each protein's surface and that the center of each protein is located appropriately in the membrane. Moreover, both of them are embedded in bilayers vertically.

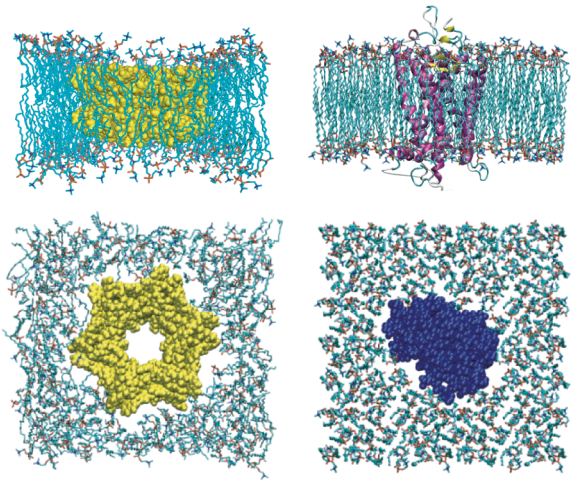


Figure 2.7 Modeled structures of more complexed membrane proteins embedded in bilayers.

The left column shows 12mer of gramicidin A proteins surrounded with a DPPC bilayer, and the right shows a GPCR, a bovine rhodopsin, embedded in a POPC bilayer (upper row, side views; lower row, top views). The molecular graphics were created by VMD[48].

II.5 Discussions

Modeling Glycans

Most glycans have complex structures in which they have branches, anomers, and many hydroxyl residues that can be often modified with some chemical functional groups. This complexity causes to make it difficult to model a glycan structure in a computer. That is why there is little software for automatic modeling of a glycan although a protein can be modeled easily with full automatic procedure. GLYMM, as shown in Figure 2.2, constructs an initial model of the glycan automatically by reading a KCF file that can be downloaded from GLYCAN database. Therefore, GLYMM avoid the difficulty from the glycan modeling. Moreover, the modeled glycan structure can be used for building a glycoprotein and a glycolipid. GLYMM implement these modeling procedures by GUI; by only clinking mouse buttons. These functions of GLYMM will be quite useful for modeling an initial structure. However, the initial model made by GLYMM is not optimized one. Therefore, energy optimization must be carried out after the modeling. This will be achieved by incorporating an energy optimization tool into GLYMM in the next version.

In the world, there are Web services relating to the glycan modeling. SweetDB[52] is the most famous Web service that can make a glycan structure in three dimension and a glycoprotein through the Internet. In contrast, KegDraw[50] is an application program working in a client computer developed at Kyoto University with the database KEGG GLYCAN in order to draw a two dimensional structure of a glycan. GLYMM now can work in the client computers. Therefore, in order for GLYMM to be widely used in the scientific fields, GLYMM will also need to work on the server connecting the Internet.

Modeling Lipid Microdomains

As described in *Chapter I*, lipid microdomains widely exist on the biomembranes, which plays many critical roles in the functions of membrane proteins. However, details about them remain unknown. Association and dissociation of lipids are the thermodynamical and kinetic issues. Therefore, the research of the lipid microdomain is quite suitable as a subject to be investigated by MD simulations. Nonetheless, MD simulations of lipid microdomain are accompanied by the difficulty of modeling, because the lipid microdomain is composed of, at least, two types of lipids in addition to the same types of lipids forming domains (Figure 1.7). This complex structure of a lipid microdomain conspicuously prevents researchers from modeling and MD simulations of lipid microdomains, which causes to delay the study in this field.

As shown in Figure 2.1, GLYMM can make a model of lipid microdomain by GUI with two canvases. Therefore, researchers can make the model quite easily as if they draw a painting on the canvases. This function of GLYMM provides modeling various types of lipid microdomains (Figure 2.4). Furthermore, GLYMM has the function for attaching glycans to lipids (Figure 2.1 and 2.2). Using these functions, we made models of ganglioside-containing membranes and calculated them (described in *Chapter V*). Because the two canvases are independent with each other, one can also model asymmetric types of lipid bilayers in which the endoplasmic and the exoplasmic leaflets have different compositions. With this function, we made models of asymmetric bilayers including cholesterols and ions (described in *Chapter IV*). These examples demonstrate that GLYMM is a quite useful tool for modeling a variety of lipid bilayers.

Modeling Membrane-Protein Complexes

A modeling method of a membrane-protein complex was also of great problem. In embedding a protein into a membrane, it is generally performed by hand; manually rotating a protein to adjust their direction, and manually translate it to the membrane. This must not be said "scientific".

GLYMM provides a scientific solution to the modeling problem. As shown in Figure 2.4, the rotation and the translation can be carried out automatically with two criteria; the principal axes of inertia for the rotation and the center of mass for the translation. In most cases, the principal axis with the smallest eigenvalue is a correct choice for the adjustment of the orientation of a membrane protein, because the integral type membrane proteins are usually long in the vertical direction against the membrane. Fitting the center of mass of a protein to that of a bilayer is also applicable to most cases for the translation as seen in Figure 2.6 and 2.7. However, exceptions are also happened. For example, in the case of Figure 2.7 (a), the membrane protein, 12mer of gA, is long horizontally against the membrane, so the principal axis to be used is different from the usual cases. In that case, since gA in itself (in the case that gA is dimer) is vertically long, the problem could be resolved easily. Hence, GLYMM has the case dependency yet. But, this problem will be avoided with consideration about the directions of the transmembrane α–helixes or the helical β–sheets, because these helixes and sheets must go up or go down through a lipid bilayer.

Before the protein-ligand docking programs emerged, such as Autodock,[53] DOCK,[54] and GOLD,[55] each ligand were manually docked into each binding site of a protein with "almost appropriate position and conformation". Similar kinds of things have remained in embedding the membrane proteins into the lipid bilayer. GLYMM will be a useful

tool for modeling a membrane-protein complex, and in the future, will build a position like the current docking programs of small ligands, such as Gold and so on.

II.6 Conclusion

GLYMM is a program for modeling biomembranes and membrane-protein complexes automatically. Using GLYMM, we can make a model as following;

(1) A three dimensional structure of a lipid microdomain with specific ion and cholesterol concentration

(2) A two dimensional glycan structure from KCF format file

(3) A three dimensional structure of a protein embedded in a membrane

(4) A three dimensional glycan structure from KCF or CCSD format file

(5) A glycoprotein and a glycolipid membrane from (1), (2) and (4) structures

These functions accelerate molecular modeling and molecular dynamics simulations of biomembranes with lipid microdomains and the membrane-protein embedded in these membranes.

III

MD Simulations of Symmetric Lipid Bilayers

Abstract

MD simulations were performed for glycerolipid and sphingolipid bilayers with symmetric distribution of lipids in leaflets. 35 kinds of glycerolipid bilayers and 48 kinds of sphingolipid bilayers were calculated. The results of comparative analyses are shown. The MD simulations reproduced various phases. Some of new indexes were introduced in order to distinguish the phase. The new indexes were more sensitive to the difference of the phases than some conventional indexes such as membrane thickness, area per lipid, and order parameters. Moreover, we found that the distribution of water molecules at the membrane-water interface depends on the head groups of lipids, while the differences of the phases and the acyl chains did not affect the distribution so much. In addition, we revealed that the water molecules at the water-membrane interface had a preference in their orientation against the membrane. These data will be useful for modeling biomembranes more accurately.

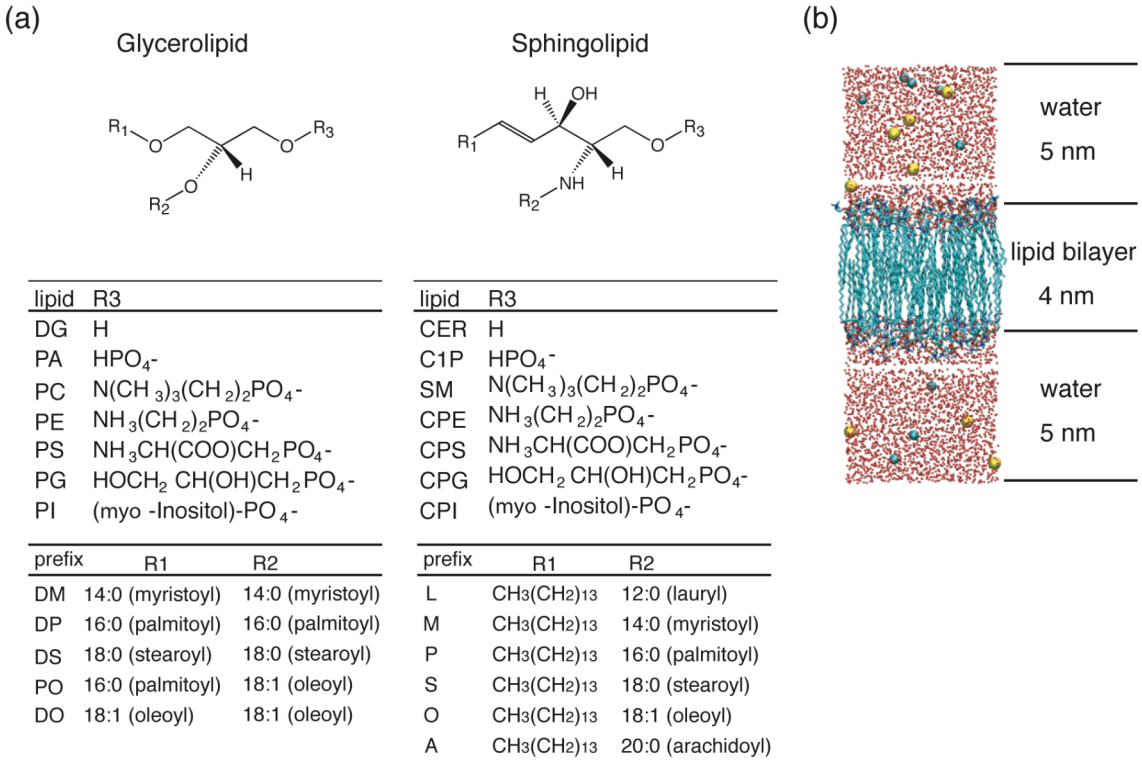
III.1 Introduction

Most of the drug targets are membrane proteins, or receptors. Recent improvement of the structural analysis provided us with many three dimensional structures of membrane proteins. These structures always show us unexpected lipid-protein interactions at their interfaces, where some lipids lay complementarily on grooves of membrane proteins' surfaces. These findings demonstrates importance of surrounding lipids for structural stability of the membrane proteins (see for review the paper written by Lee).[31] Furthermore, some experiments indicated that appropriate thickness of bilayer is essential for the membrane proteins to exert their intrinsic functions sufficiently.[31,56-58]

Numerous MD simulations of membrane proteins have been carried out.[59-65] In these simulations, membrane proteins are embedded in "explicit" or "implicit" lipid bilayer. "Explicit" means that the lipid bilayer is composed of atoms, in contrast to "implicit" meaning that the bilayer is represented as a continuum solvation model with a low dielectric constant, such as $\varepsilon = 1$ or 2,[66] or with smoothly changing dielectric constants by a switching function.[67,68]

These simulations, however, did not consider the variety of lipid composition of biomembranes. A difference of a lipid composition gives rise to a different interaction and a different dielectric field around the membrane. Therefore, lipid composition should be carefully considered in an accurate MD simulation of membrane proteins. As described in *Chapter II*, the main reason why only simple lipid bilayers are used in MD simulations is that it is difficult to get the structures of membranes with a different lipid compositions.

In order to solve these problems, we carried out MD simulations of various lipid bilayers with symmetric distribution of lipids in leaflets and we are now building the database of these coordinates. Here, we describe the results of MD simulations of 35 kinds of glycerolipid and 48 kinds of sphingolipid bilayers (Figure 3.1) under the same condition. Large-scale comparative analyses showed some common features between glycerolipid and sphingolipid bilayers. In addition, to distinguish phases of bilayers automatically, we introduced Correlation-Interdigitation map. This map was able to classify the bilayers into four phases; L_α, $L_\alpha I$, L_β, and $L_\beta I$. This classification was useful for further comparative analyses of the bilayers.



lipid	R3
DG	H
PA	HPO_4-
PC	$N(CH_3)_3(CH_2)_2PO_4$-
PE	$NH_3(CH_2)_2PO_4$-
PS	$NH_3CH(COO)CH_2PO_4$-
PG	$HOCH_2$ $CH(OH)CH_2PO_4$-
PI	(myo -Inositol)-PO_4-

lipid	R3
CER	H
C1P	HPO_4-
SM	$N(CH_3)_3(CH_2)_2PO_4$-
CPE	$NH_3(CH_2)_2PO_4$-
CPS	$NH_3CH(COO)CH_2PO_4$-
CPG	$HOCH_2$ $CH(OH)CH_2PO_4$-
CPI	(myo -Inositol)-PO_4-

prefix	R1	R2
DM	14:0 (myristoyl)	14:0 (myristoyl)
DP	16:0 (palmitoyl)	16:0 (palmitoyl)
DS	18:0 (stearoyl)	18:0 (stearoyl)
PO	16:0 (palmitoyl)	18:1 (oleoyl)
DO	18:1 (oleoyl)	18:1 (oleoyl)

prefix	R1	R2
L	$CH_3(CH_2)_{13}$	12:0 (lauryl)
M	$CH_3(CH_2)_{13}$	14:0 (myristoyl)
P	$CH_3(CH_2)_{13}$	16:0 (palmitoyl)
S	$CH_3(CH_2)_{13}$	18:0 (stearoyl)
O	$CH_3(CH_2)_{13}$	18:1 (oleoyl)
A	$CH_3(CH_2)_{13}$	20:0 (arachidoyl)

Figure 3.1 Lipids and a bilayer used in MD simulations.

(a) Glycerolipids and sphingolipids calculated in this study (DG, diacylglycerol; PA, phosphatidic acid; PC, phosphatidylcholine; PE, phosphatidylethanolamine; PS, phosphatidylserine; PG, phophatidylglycerol; PI, phosphatidylinositol; CER, ceramide; C1P, ceramide 1-phosphate; SM, sphingomyelin; CPE, CPS, CPG, and CPI are the sphingolipid analogues of PE, PS, PG, and PI, respectively). (b) The initial structure used in this study.

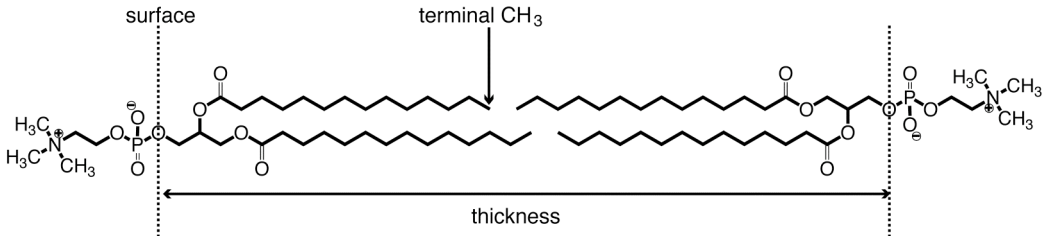


Figure 3.2 Definition of membrane surface and thickness

III.2 Computational Details

Details of MD Simulations

Computational models of lipid membranes were built by GLYMM described in *Chapter II*. MD simulations were carried out by NAMD 2.5 and 2.6.[69] The calculation protocol is as follows; (1) 10,000 steps of minimization (2) 6 ps of heating (3) 1.0 ns of equilibration at 310 K and 1 atm with the head groups of lipids restrained (4) 0.5 ns of equilibration at 310 K and 1 atm without restraint (5) 2 ns of production run at 310 K and 1 atm. The concentration of NaCl was set to 0.15 M. In the case of lipids with negative charges, extra ions were added in order to make net charge zero for PME calculations. Other detailed parameters for MD simulations were the same as described in the *Computational Details* section of *Chapter V*.

Statistics

Most of the analyses and the graphics were produced with the following functions in the statistical program R.[70] Balloon plots were drawn by the *balloonplot* function in the *gregmisc* library. All of the plots classified according to the lipid head groups were drawn using the *xyplot* function in the *lattice* library. The non-linear modeling function, *Nls*, was used for the prediction of parameters to fit a distribution function of water molecules to a sigmoid curve. The binned two dimensional kernel density estimation method[71] was used for the analyses of orientation of water molecules, which was implemented in the *bkde2D* function in the *KernSmooth* library and was drawn by the *contourplot* function in the *lattice* library.

Membrane Thickness and Area per Lipid

Membrane surface was defined as time-averaged z coordinates of oxygen atoms in *sn*-1 position for glycerolipids and oxygen atoms in position 1 of sphingoid base for sphingolipids (Figure 3.2). Membrane thickness was calculated as the distance from a membrane surface to the opposite surface. Area per lipid was calculated by dividing time-averaged x and y length of a periodic box by the number of lipid molecules in one leaflet.

Methods for Phase Determination

To determine the phase of the membrane automatically, we used correlation between all pairs of acyl chains and intensity of interdigitation, and then made a Correlation-Interdigitation map for each lipid bilayer (Figure 3.3). Correlation of acyl chains was defined as the following equation:

$$C_{ij} = \frac{|\vec{l_i}|}{l_{i0}} \frac{|\vec{l_j}|}{l_{j0}} \cos\theta_{ij} = \vec{u_i} \cdot \vec{u_j} \qquad where \;\; \vec{u_i} = \frac{\vec{l_i}}{l_{i0}} \tag{3.1}$$

where *i, j* denote *i*th and *j*th acyl chains; l_0 means the maximum length when the acyl chain is in all *trans* conformations; each vector starts from the carbonyl carbon and ends at the terminal methyl group of the acyl chain. Therefore, the correlation represents how the two acyl chains stretch toward the same direction. Because it forms dot product of two vectors with the length less than or equal to 1, its range is from -1 to 1.

Intensity of interdigitation was defined as follows:

$$I_{ij} = 2(d_{ij}^{z} - d_{ij}^{mid}) \quad \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \tag{3.2}$$

$$d_{ij}^{z} = \frac{\vec{u_i} + \vec{u_j}}{2} \cdot \vec{e_z} \;, \qquad d_{ij}^{mid} = \frac{z_i^{start} + z_j^{start}}{2} - z_{mid} \quad \cdots\cdots\cdots \tag{3.3}$$

where the vector $\boldsymbol{u}$ is the same one in the equation (3.1); the vector $\boldsymbol{e_z}$ is the unit vector toward z axis; z_i denotes z coordinate of an *i*th atom; z_{mid} means a distance between a carbonyl atom of an acyl chain and a center plane of a bilayer. The value of interdigitation is adjusted so as to be 1 when the center of the acyl chain is just located on the center plane. When a lipid membrane forms a typical bilayer structure, each leaflet of the bilayer is separated from the opposite leaflet, so some space exists at the center of the membrane. In this case, the bilayer takes minus value of interdigitation.

Using these indexes, all combinations of acyl chains are mapped on a Correlation-Interdigitation map as a point (Figure 3.3 (c)). Counting the number of points in the separated region by two threshold lines, the phase of each lipid bilayer can be determined automatically by computer. For example, DSPE bilayer had many points concentrating on the high correlation and the low interdigitation region; therefore, DSPE bilayer can be classified into L_β phase. In DSPI bilayer, the points were concentrated on the high correlation and the high interdigitation, which indicates that its phase is $L_\beta I$ (Figure 3.3 (c)).

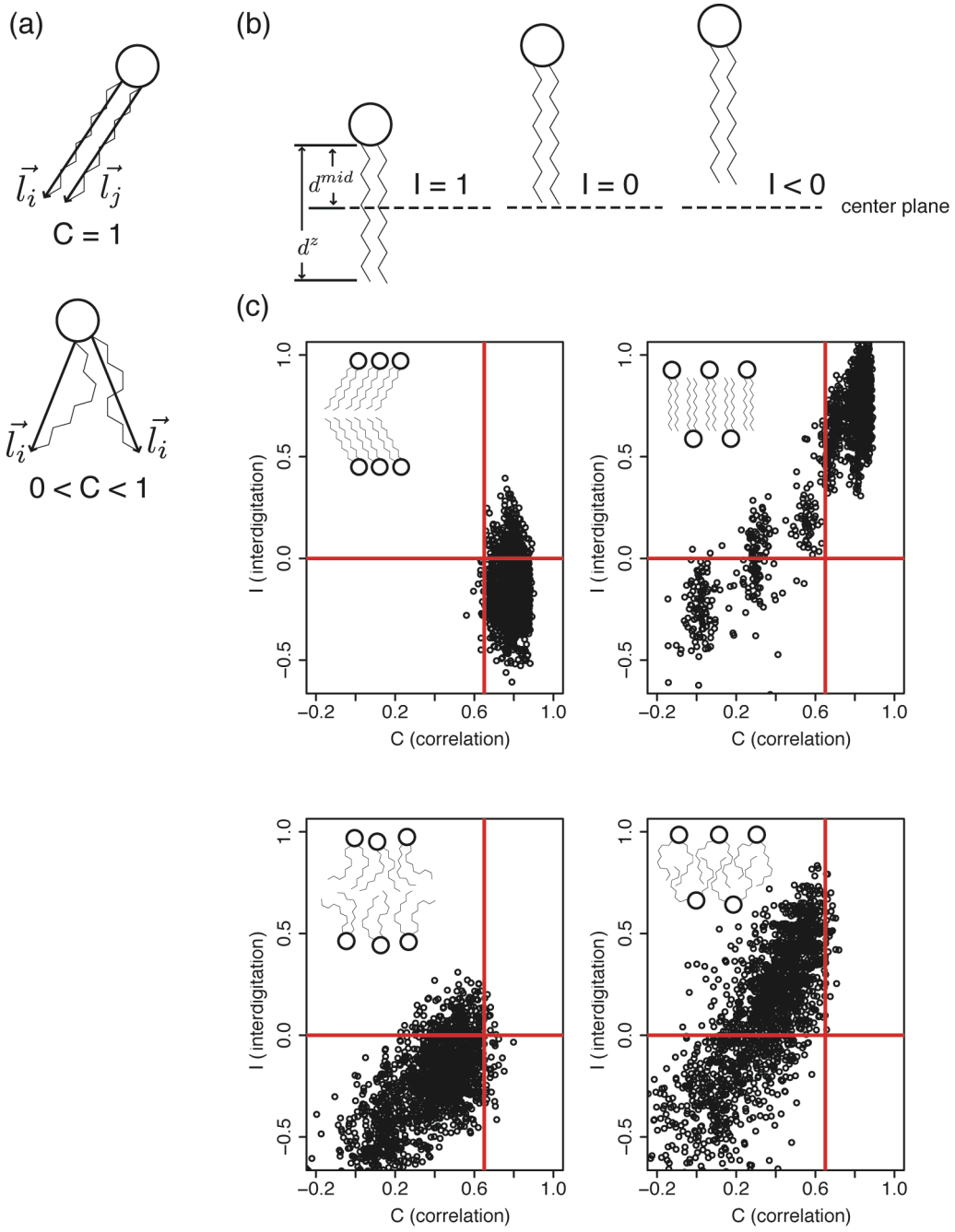


Figure 3.3 Correlation-Interdigitation map.

(a) Concept of correlation, C. (b) Concept of intensity of interdigitation, I. (c) Typical examples of Correlation-Interdigitation maps for $L_{\beta'}$ (upper left, DSPE's data are plotted), $L_{\beta}I$ (upper right, DSPI's data), L_{α} (lower left, POPC's data), and $L_{\alpha}I$ (lower right, DODG's data) phase. Red lines show threshold for classification.

Quantification of Distribution of Water Molecules at Membrane Interface

Distribution of water molecules at the membrane-water interface is shown in Figure 3.8 and 3.15. At every interface, the number of water molecules gradually increases, and the shape of every distribution function results in the sigmoid curve. A sigmoid curve can be written as follows;

$$y = f(x, a, b) = \frac{1}{1 + e^{(-ax+b)}} \quad \text{..... (3.4)}$$

where a is a parameter for the degree of increase of y, and b is one for deviation. These a and b values are more easily understand in the following expression;

$$y = g(x, a, b) = \frac{a}{4}(x - d_{50}) + 0.5, \quad d_{50} = \frac{b}{a} \quad \text{..... (3.5)}$$

This expression means the tangent line at the point where x equals d_{50} and y equals 0.5.

The non-linear modeling function, *Nls*, was used for the prediction of these parameters, a and b.

Quantification of Orientation of Water Molecules at Membrane Interface

An orientational angle θ was defined as an angle formed between a membrane normal vector (**n**) and a cross product (**a×b**) of two OH bonds of a water molecule (**a** and **b**) (Figure 3.4). Then, all of cos θ with respect to the distance from the membrane surface were mapped, and the two dimensional probability density was estimated with the binned two dimensional kernel density estimation method. The results were drawn as in Figure 3.9 and 3.16.

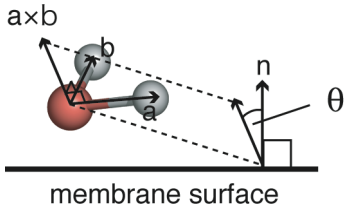


Figure 3.4 Definition of orientational angle of a water molecule.

III.3 Results

III. 3.1 Glycerolipids

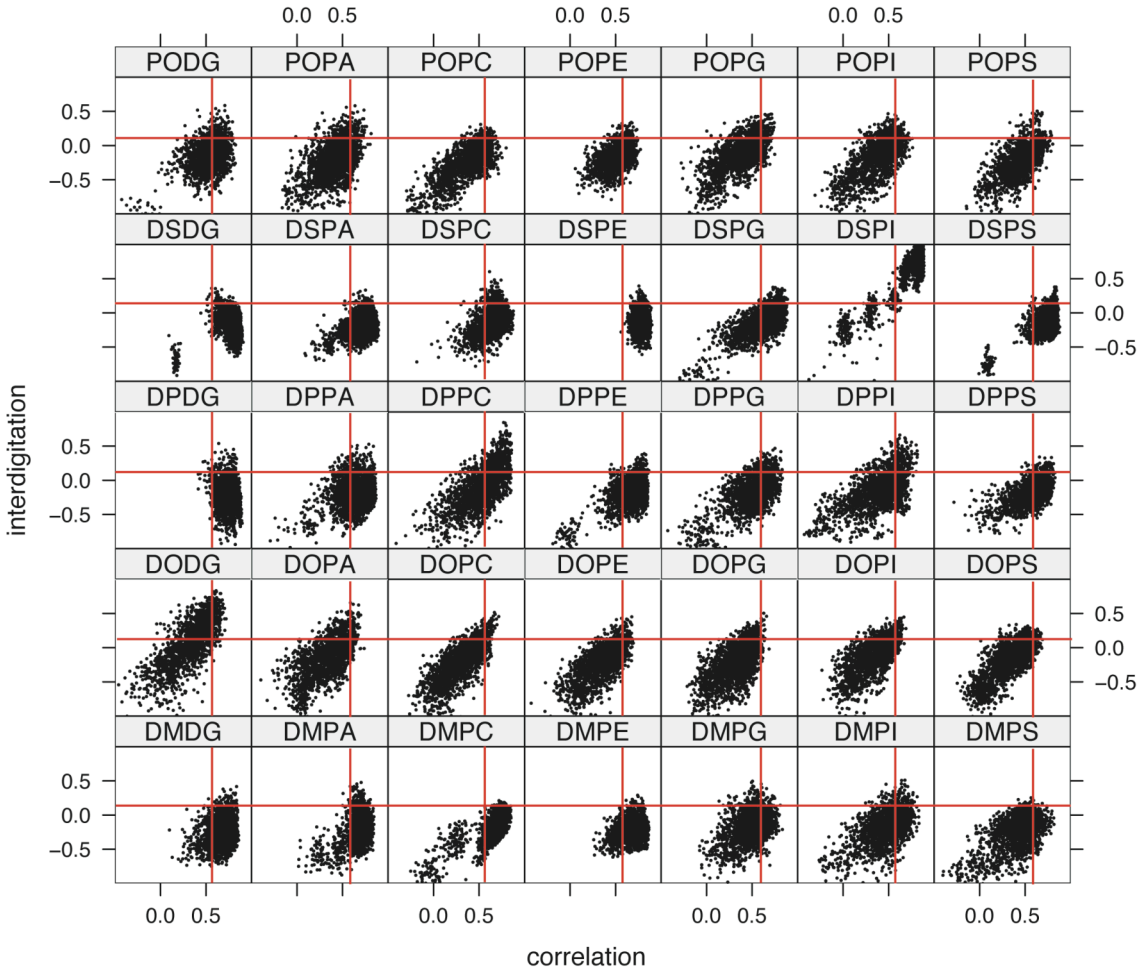


Figure 3.5 Correlation-Interdigitation map of glycerolipid bilayers.

Vertical and horizontal red lines indicate the thresholds for automatic determination of the phase of the bilayers.

Figure 3.5 shows the Correlation-Interdigitation map (described in the previous section) of glycerolipid bilayers and Figure 3.6 shows typical examples of L_β, $L_\beta I$, L_α, and $L_\alpha I$ bilayers. Comparing the map to the structures, it can be found that the map reflected the structures quite well. Therefore, it demonstrates that the mapping method is appropriate for determination of the phase.

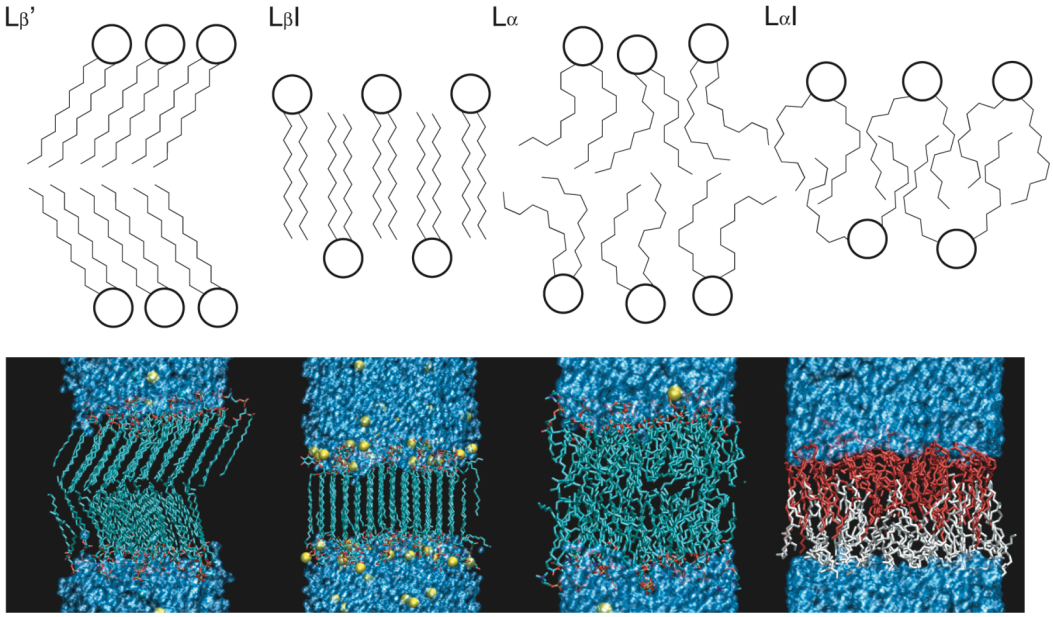


Figure 3.6 Snapshots of typical examples of L_β, $L_\beta I$, L_α, and $L_\alpha I$ phase.

Upper row shows schematic drawings of phases of lipid bilayers. Lower row shows snapshots of bilayer structures resulted from MD simulations (DSPE, DSPI, POPC, and DODG bilayers from left to right). Lipids are represented as sticks. The blue regions above and below the membranes are water layers. Yellow and green spheres are Na^+ and Cl^- ions. In $L_\alpha I$ phase, lipids in upper and lower leaflet are colored red and white for clarity.

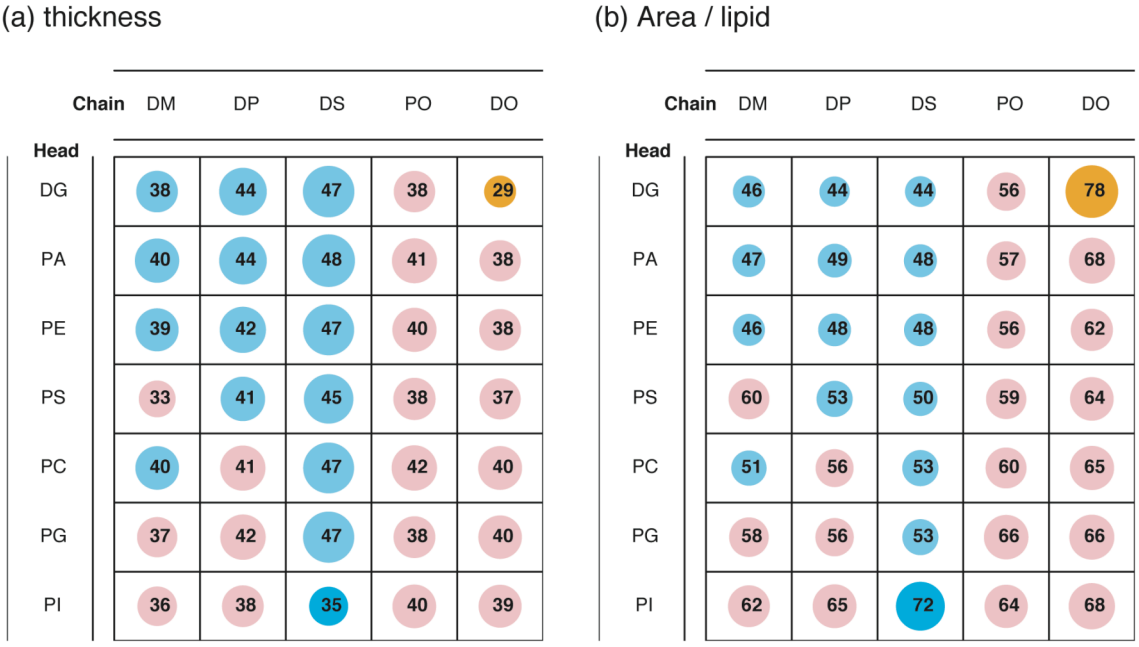


Figure 3.7 Balloon plots of thickness and area per lipid of glycerolipid bilayers.

The diameter of a circle is proportional to each value. Pink, orange, sky-blue, and blue colors indicate L_α, $L_\alpha I$, L_β, and $L_\beta I$ phase judged by the Correlation-Interdigitation maps, respectively. (a) Values of the average thicknesses of glycerolipids. Unit is Å. (b) Values of the average area per lipid of glycerolipids. Unit is $Å^2$.

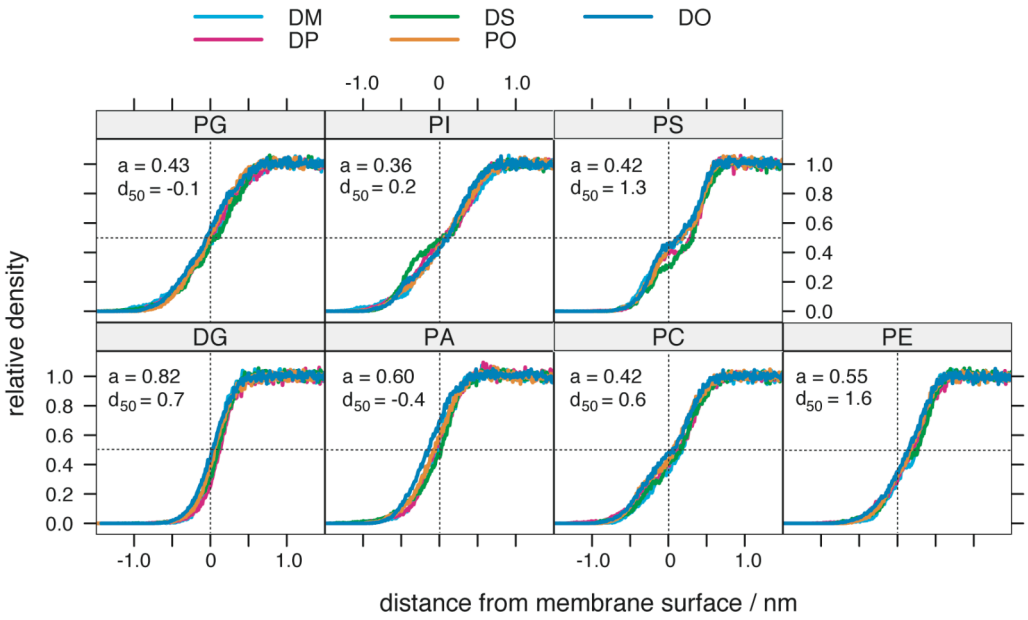


Figure 3.8 Distribution functions of water molecules at the glycerolipid bilayer interface.

Bulk density of water molecules is set to 1.0. "a" and "d_{50}", indicate the coefficients in equation (3.4).

Figure 3.7 shows values of the average thickness and area per lipid of glycerolipids, and simultaneously the figure shows the phase of the bilayer by color. At a glance, the lipid membranes in L_β phase have larger thickness and smaller area per lipid than in L_α phase. Regardless of the head groups, all of the lipid bilayers with unsaturated acyl chains (PO and DO series) were in L_α phase. PG and PI had also some preference for forming the L_α phase. Both the head groups of PG and PI are polyols (Figure 1.3), which suggests that polyols may induce a glycerolipid bilayer to change their phases from L_β to L_α. In contrast, lipids with a small head group such as DG and PA, and lipids with neutral charge such as PC and PE, preferred to form L_β phase. DSPI and DODG bilayer, where interdigitation was occurred, showed the quite small values of thickness and the largest values of area per lipid.

Figure 3.8 shows the distribution functions of water molecules at the membrane interfaces and the values of a and d_{50} averaged over the five curves (DM, DP, DS, PO, and DO) in each lipid. The two parameters, a and d_{50}, enable us to deal with each curve quantitatively, although it is difficult to recognize the difference only from their appearances. Then, it can be found that DG has the largest a value and PI has the smallest one, where the difference is more than double. Interestingly, the values are only dependent on the kind of the head group. It indicates that the water distribution at the membrane-water interface can be practically described by the difference of the head group.

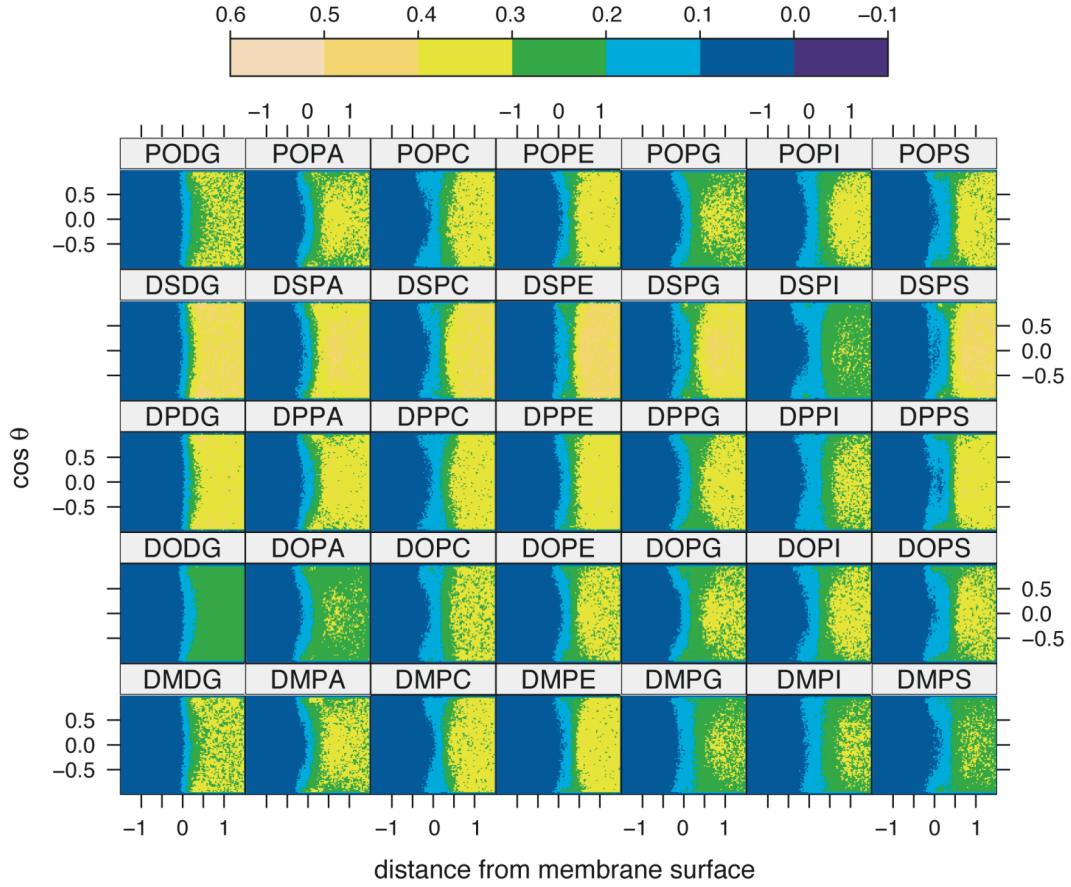


Figure 3.9 Density functions of orientational angels of water at the glycerolipid interface.

Abscissa is the distance from the surface of the membrane in nm.

Figure 3.9 shows two-dimensional density functions of orientational angles of water molecules at water-membrane interfaces. The figure shows that the water molecules near the membrane surface have a preference in their orientation that the absolute value of cosθ is nearly 1. This result indicates that the plane of the water molecule tend to parallel the membrane surface at the interface. Furthermore, it appears that the preference becomes stronger as the head group becomes larger.

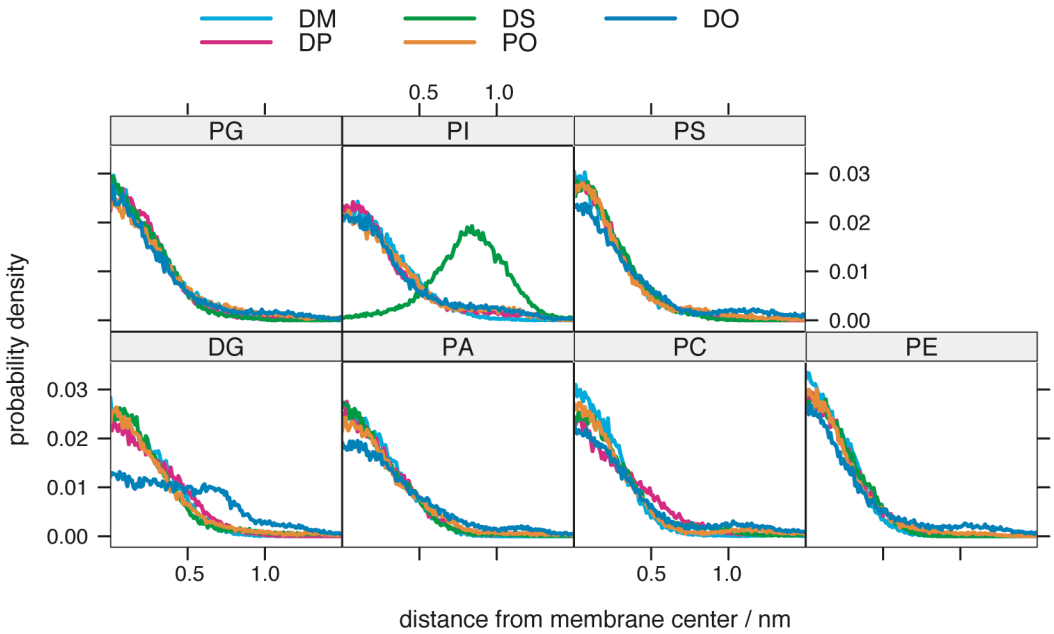


Figure 3.10 Probability density of terminal CH_3 of glycerolipids.

Abscissa is the distance from the geometrical center of the z-coordinates of the membrane in nm.

Figure 3.10 shows distributions of the terminal CH_3 groups of acyl chains. The terminal methyl groups are concentrated at the center of the bilayer when the bilayer is separated clearly into two leaflets. Therefore, a typical distribution should be the gaussian form with its mean at the membrane center. As seen in Figure 3.10, most lipids have the gaussian distribution except for DODG and DSPI. DODG and DSPI membranes formed interdigitated structures as shown in Figure 3.6. The DSPI membrane was in the fully interdigitated gel ($L_{\beta}I$) phase, while the DODG membrane was in the partially interdigitated liquid-crystalline ($L_{\alpha}I$) phase. Thus, analysis of the distribution of terminal CH_3 group can be used for distinguish a clearly separating bilayer from a interdigitated one.

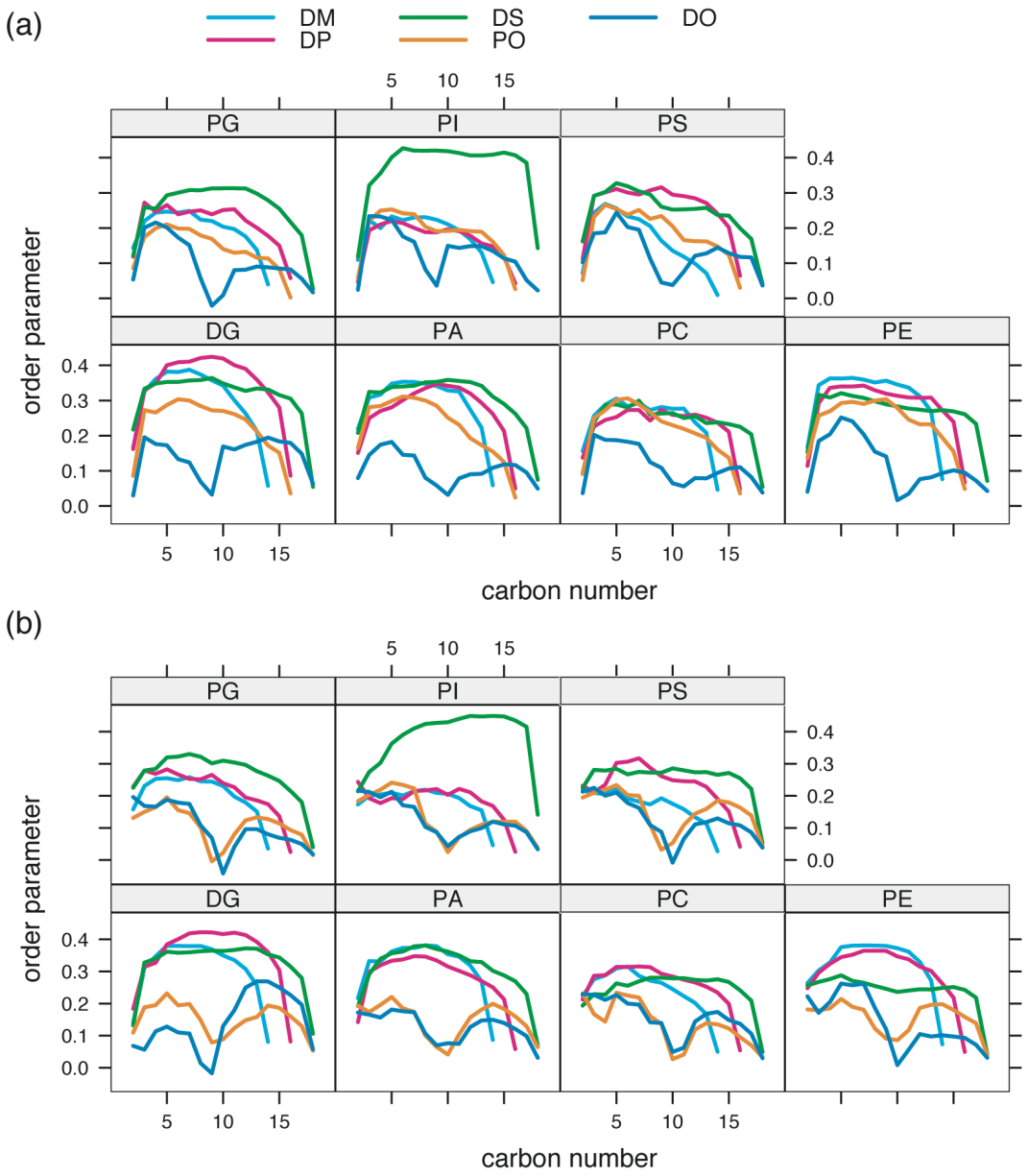


Figure 3.11 Order parameters of glycerolipids.

(a) *sn*-1 chains (b) *sn*-2 chains.

Figure 3.11 shows the order parameters of the acyl chains. Large values are observed in L_β phase while the values of L_α phase are relatively low. In both of them, the values decrease gradually at the end of the acyl chains, which are the typical results in a usual bilayer. However, the interdigitated bilayers (DSPI and DODG) have unusual large values from the middle methylene groups to terminal methyl groups. However, the boundary between L_α and L_β is not so clear.

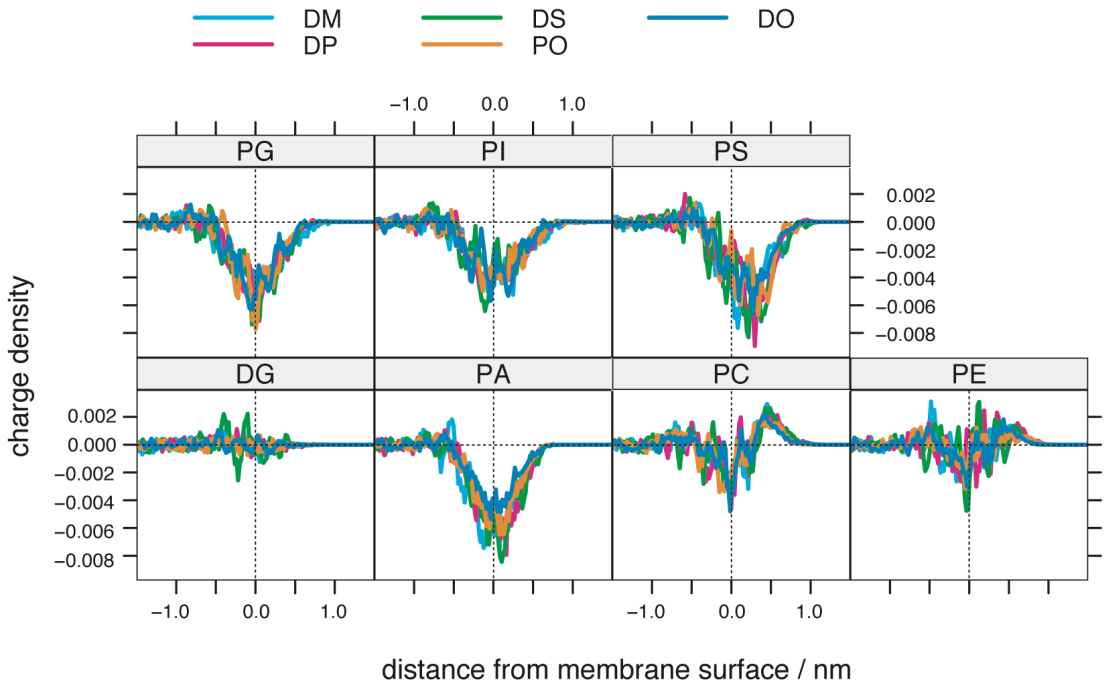


Figure 3.12 Charge densities of glycerolipid bilayers.

Charge densities of glycerolipids measured from the membrane surface. Unit is C/Å^3.

Charge density at the membrane-water interface is one of the key factors for the interactions of amphipathic peptides, such as antimicrobial peptides[72,73] and the amyloid β-peptide,[74,75] with biomembranes. Therefore, it is important to investigate them. The relationship between the amyloid β-peptide and charge densities of membranes is discussed in detail in *Chapter V*.

Figure 3.12 shows the charge densities of the glycerolipid bilayers without the contribution from ions. As was expected, the acidic lipids, PA, PG, PI, and PS, had negative peak at the membrane surface. The peak of PS moves toward the water layer a little, which would comes from its carboxyl groups those are located away from the surface. The absolute values at the minimum of PI are relatively small owing to the large surface area of the PI membranes (Figure 3.7). PC has a clear positive peak at the 5 nm away from the surface, which would cause the dipole moment along the membrane normal at the PC membrane-water interface.

III.3.2 Sphingolipids

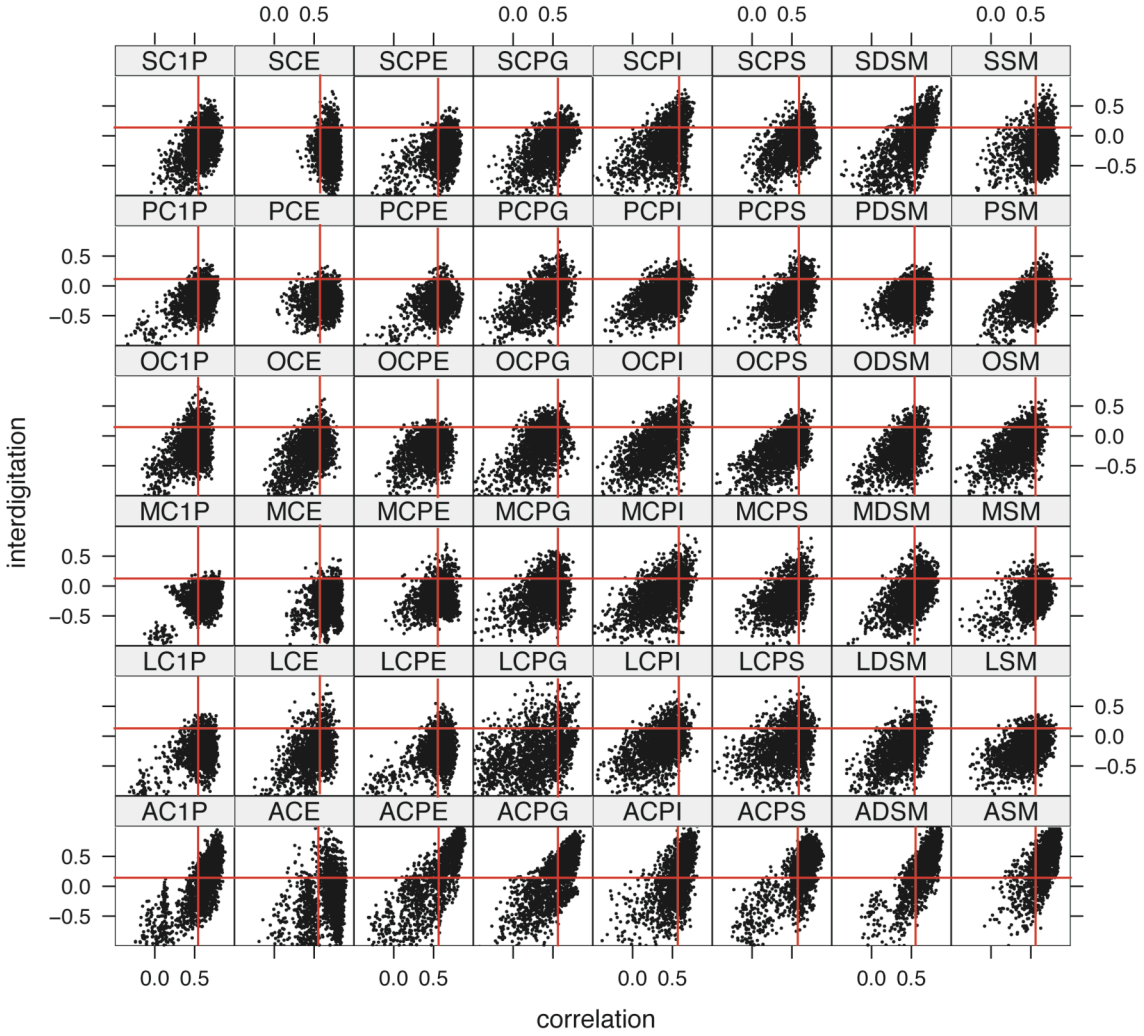


Figure 3.13 Correlation-Interdigitation map of glycerolipid bilayers.

Vertical and horizontal red lines indicate the thresholds for automatic determination of the phase of the bilayers. Prefix "L" means that the sphingolipid has a 12:0 acyl chain; "M", 14:0; "P", 16:0; "S", 18:0; "O", 18:1; "A", 20:0. Abbreviations for lipid names are shown at the page 6.

In contrast to the glycerolipids, most of the sphingolipids bilayers were assigned to L_α phase. However, many bilayers are on the border of L_α-L_β even in the case of bilayers assigned to L_β phase (Figure 3.13). It should be noted that most lipids with the arachidoyl (20:0) chain formed interdigitated bilayers. This would be caused by the large difference in the length of the acyl chain and the hydrophobic part of the sphingosine backbone.

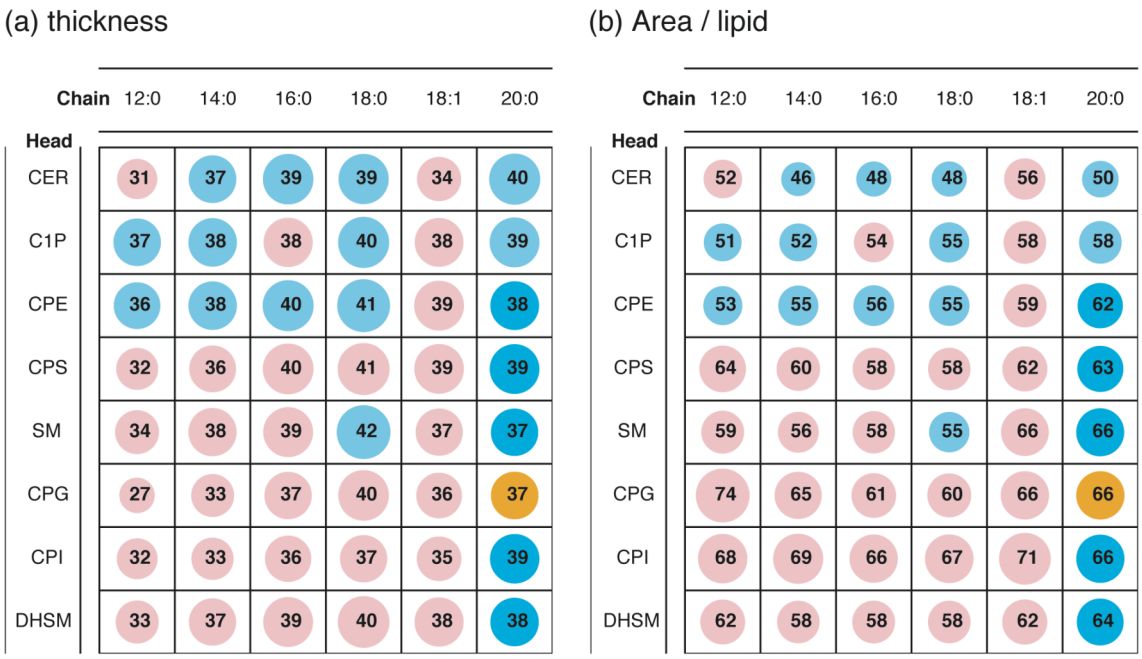


Figure 3.14 Thickness and area per lipid of sphingolipid bilayers.

The diameter of a circle is proportional to each value. Pink, orange, sky-blue, and blue colors indicate L_{α}, $L_{\alpha}I$, L_{β}, and $L_{\beta}I$ phase judged from the criteria defined above, respectively. (a) Values of the average thicknesses of sphingolipid bilayers. Unit is Å. (b) Values of the average area per lipid of sphingolipid bilayers. Unit is $Å^2$.

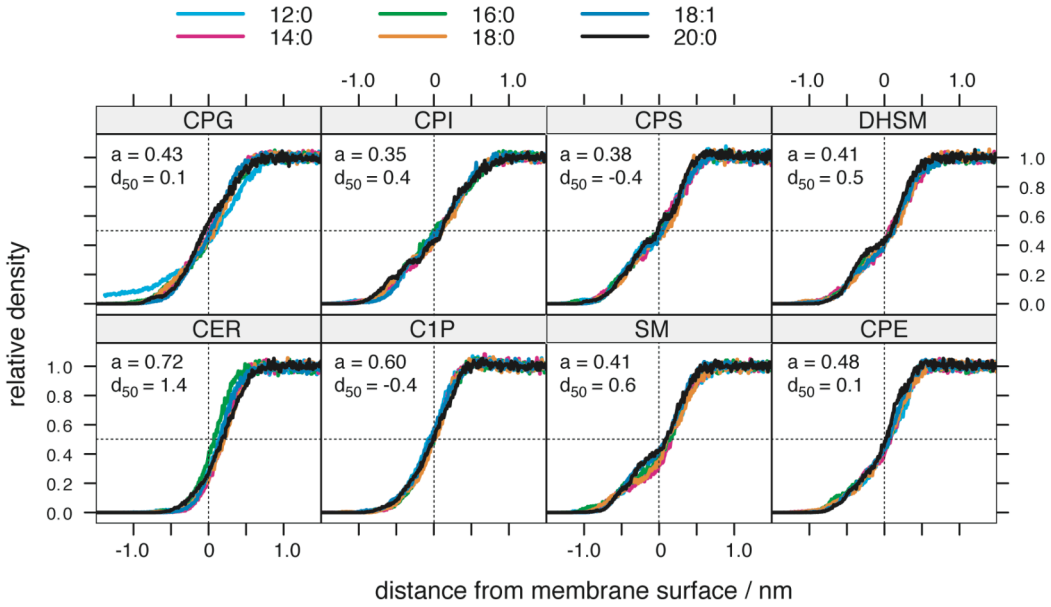


Figure 3.15 Distribution functions of water molecules at the sphingolipid bilayer interface.

Bulk density of water molecules is set to 1.0. "a" and "d_{50}", indicate the coefficients in equation (3.4).

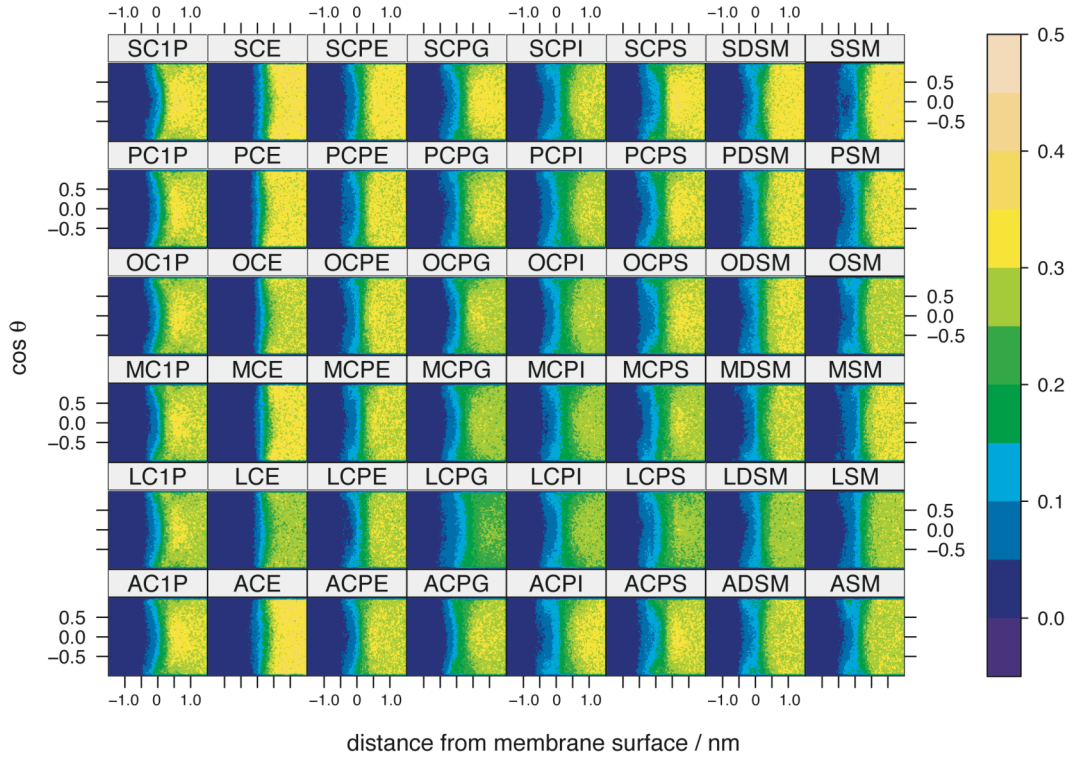


Figure 3.16 Density functions of orientational angle of water at the sphingolipid interface.

Abscissa is the distance from the membrane surface in nm.

Figure 3.14 shows the average thickness and area per lipid of sphingolipid bilayers. Because of most of the sphingolipid bilayers in the L_α phase, the values of thickness are around 35 ~ 40 Å. Interdigitation makes area per lipid larger than the other phases as similar to interdigitation of glycerolipids.

Figure 3.15 shows the distribution functions of water molecules at the membrane-water interfaces. Compared to the glycerolipids, all of the distribution functions also resulted in the sigmoid curves, independent of the length of acyl chains. This result also support the result described in the previous section that the water distribution at the interface is only dependent on the head group of lipid.

Figure 3.16 shows the two-dimensional density functions of orientational angles of the water molecules at the sphingolipid membrane-water interface. As similar to the glycerolipids, the water molecules at the interface had the same preference in their orientation. These results reveal that the water distribution and the orientation are strongly affected by the head groups of the lipids but not so affected by the acyl chains, backbones, and their phase.

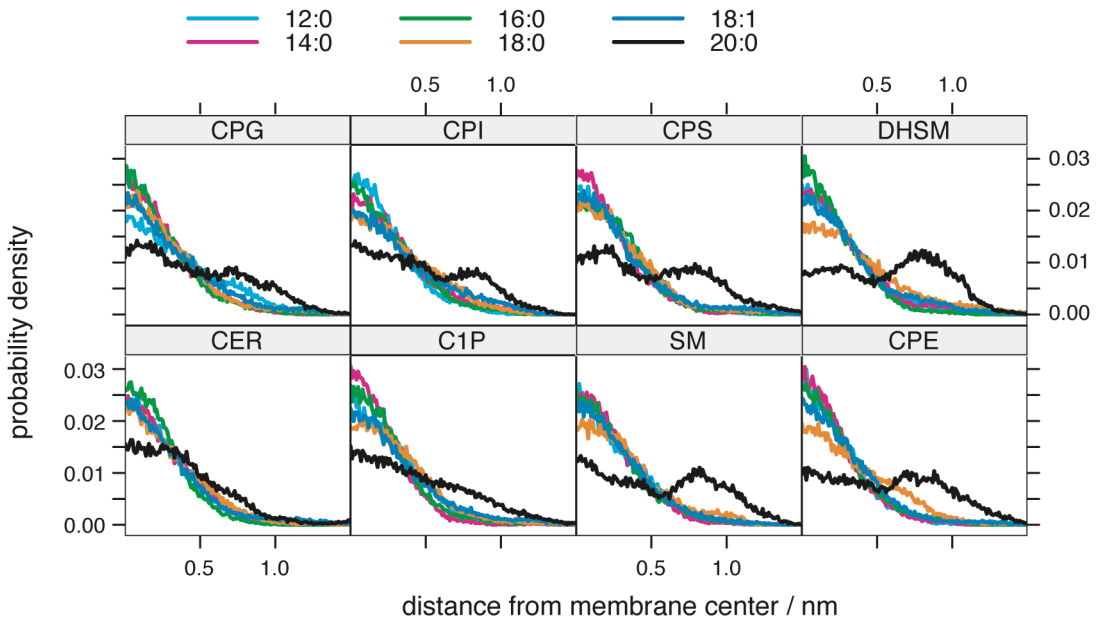


Figure 3.17 Probability density of terminal CH_3 of lipid tails of sphingolipids.

Abscissa is the distance from the membrane center in nm.

Figure 3.17 shows the probability density functions of the terminal CH_3 groups of the lipid tails of sphingolipids. As similar to the result of the glycerolipids, most of the distributions were the gaussian. However, almost all of the lipids with the arachidoyl (20:0) chain showed the characteristic pattern indicating the partially interdigitated phase. This can be explained from the heterogeneity of tail parts of lipids. Generally, the larger the difference between the lengths of acyl chain and the part of sphingosine in the membrane becomes, the more frequently the interdigitated phase occurs. Interestingly, although the heterogeneity also exists in the other acyl chains, the interdigitated phase was observed only in the membrane with 20:0 chains.

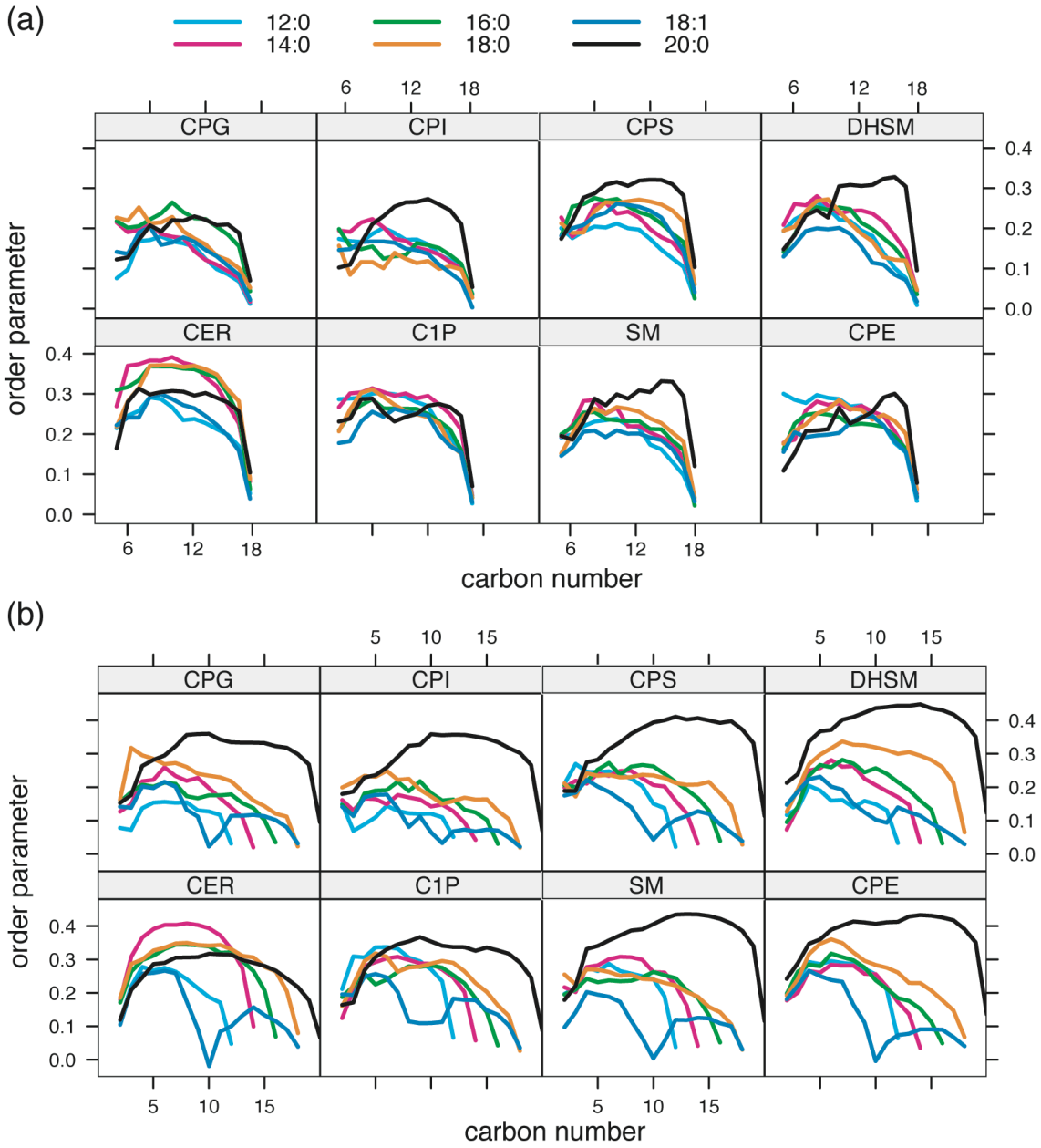


Figure 3.18 Order parameters of sphingolipids

(a) sphingosine base (b) acyl chain

Figure 3.18 shows the order parameters of the lipid tails of sphingolipids. Reflected the interdigitation, the sphingolipids with arachidoyl chains had large order parameters. CER in particular 14:0-ceramides, also showed larger values than other lipids, which is also observed in the DG as shown in Figure 3.11.

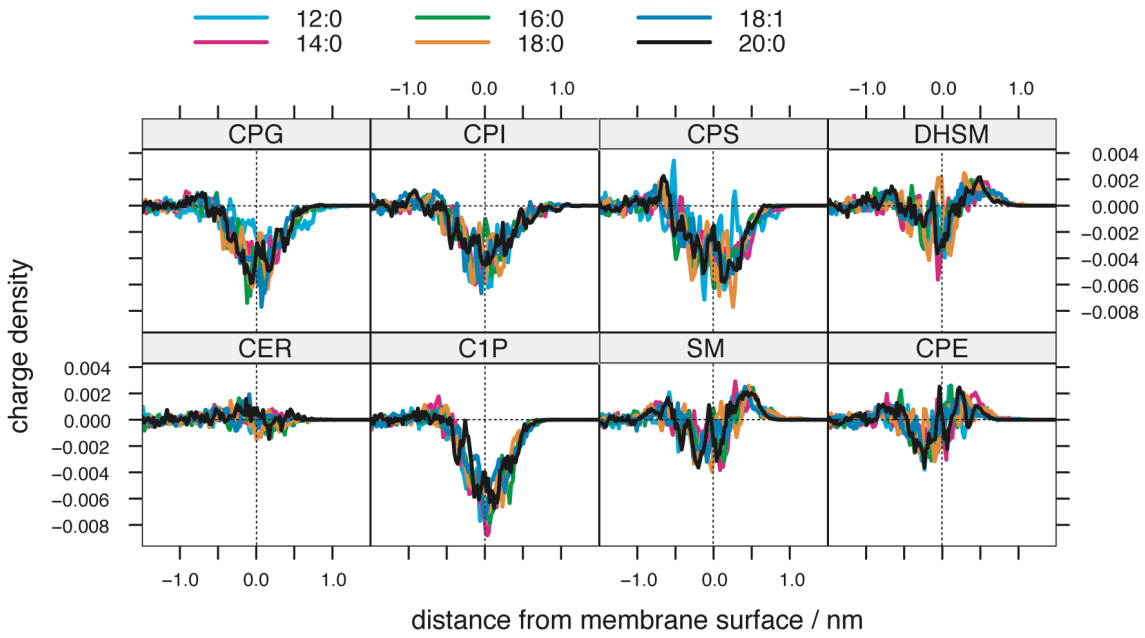


Figure 3.19 Charge densities of sphingolipids.

Abscissa is the distance from the membrane surface in nm. Unit of charge density is C/Å^3.

Charge densities of sphingolipid bilayers are shown in Figure 3.19. The profiles are quite similar to those of the corresponding glycerolipid bilayers. It indicates that the charge densities are also dependent on the head groups.

III.4 Discussion

Quantitative Analyses of Lipid bilayers

MD simulations of 83 kinds of lipid bilayers in this study enable comprehensive comparative analyses of glycerolipid and sphingolipid bilayers. These analyses reveal some common features in glycerolipid and sphingolipid bilayers.

As seen from Figure 3.7 and 3.14, there seems correlation between thickness and area per lipid. Therefore, linear regression analysis was performed (Figure 3.20). Because the thickness of the membrane is dependent on the number of carbon atoms composing the hydrophobic core of the membrane, the values of thickness need to be divided by the number of the carbon atoms for the sake of comparison. As expected, the correlation coefficients are high, -0.883 and -0.847, and the regression formulas are similar (Figure 3.20). As thickness increases, area per lipid decreases, which is common in both glycerolipid and sphingolipid bilayers, and furthermore, the decreasing rate is almost consistent. The difference between glycerolipid and sphingolipid bilayers becomes larger at the higher values of thickness per carbon atom. This difference suggests that the sphingolipid has larger areas in the gel phase. Interestingly, all of the ceramide bilayers have smaller area per lipid than the others, and clearly, they deviate from the estimated regression line (shown in the red circle in Figure 3.20 right). This deviation may be caused by the ceramide structure that has two hydroxyl moieties at its head group. The hydroxyl groups formed quite strong intra- and inter-molecular H-bonds in our simulations. These H-bonds may pull up their lipid tails from the hydrophobic core of the membrane, which lead us to overestimation of the number of carbon atoms in the hydrophobic core.

Although the thickness of the membrane is dependent on both the length of the acyl chains and its phase state, area per lipid would depend strongly on its phase as shown in Figure 3.7 and 3.14. So, the correlation between the gel percentage and the area per lipid was investigated (Figure 3.21). Then, the high negative correlations are observed in glycerolipid and sphingolipid bilayers. Moreover, the regression lines are quite similar, which reveals the area per lipid can be described only by the gel percentage of the membrane, and not dependent on the kinds of lipids. This information will be very useful for modeling of a lipid bilayer with a specific gel percentage. It may be interesting to add such a function to GLYMM.

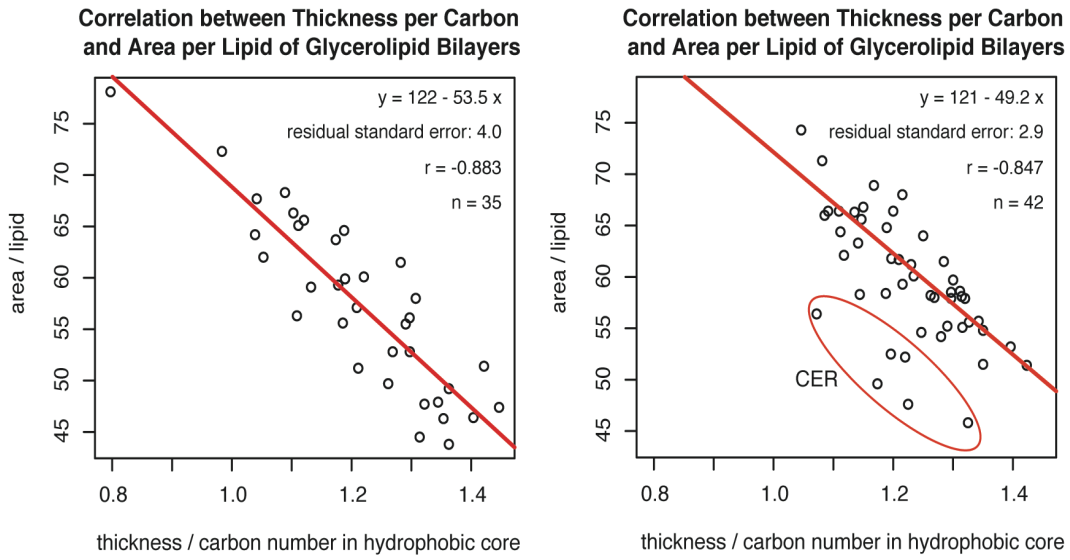


Figure 3.20 Correlation between thickness per carbon atom and area per lipid.

Regression line is colored by red. Regression formula, residual standard error, correlation coefficient, and the number of samples are shown in the upper right corner in each figure. Red circle indicate the data of ceramides that was not used for linear regression.

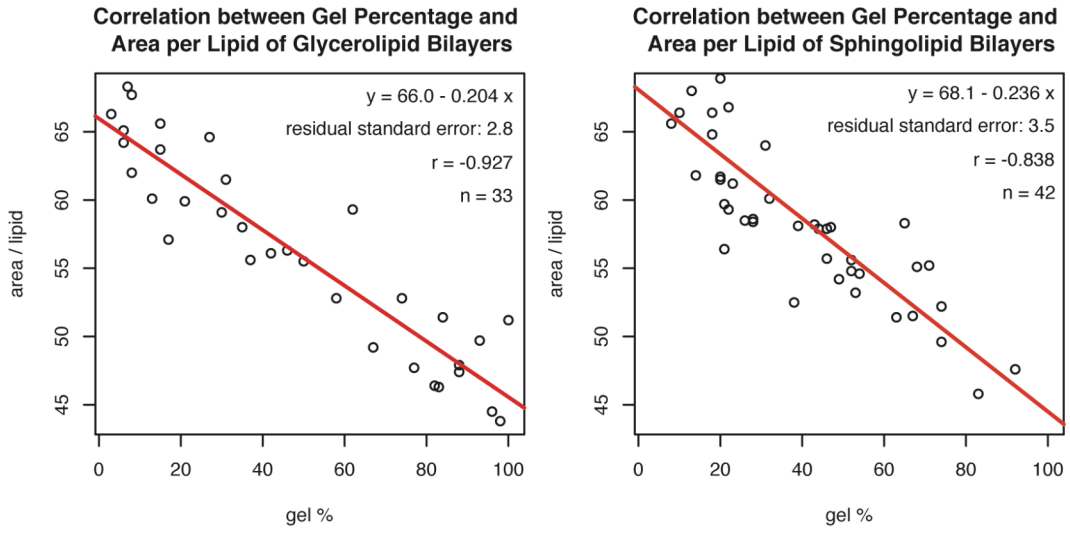


Figure 3.21 Correlation between gel percentage and area per lipid.

Regression line is colored by red. Regression formula, residual standard error, correlation coefficient, and the number of samples are shown in the upper right corner in each figure. These data does not include the data of the interdigitated membranes.

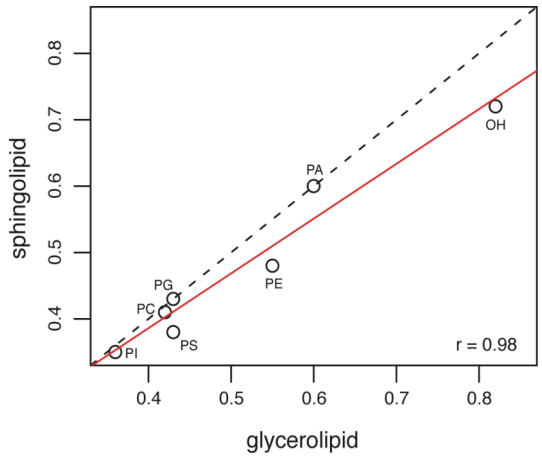


Figure 3.22 Correlation between *"a"* values of glycerolipids and sphingolipids.

Dashed line represents the line with the gradient = 1 and the intercept = 0. Solid line is the regression line. Correlation coefficient, *r*, is shown in the lower right corner. Head groups are also designated. OH means DG and CER.

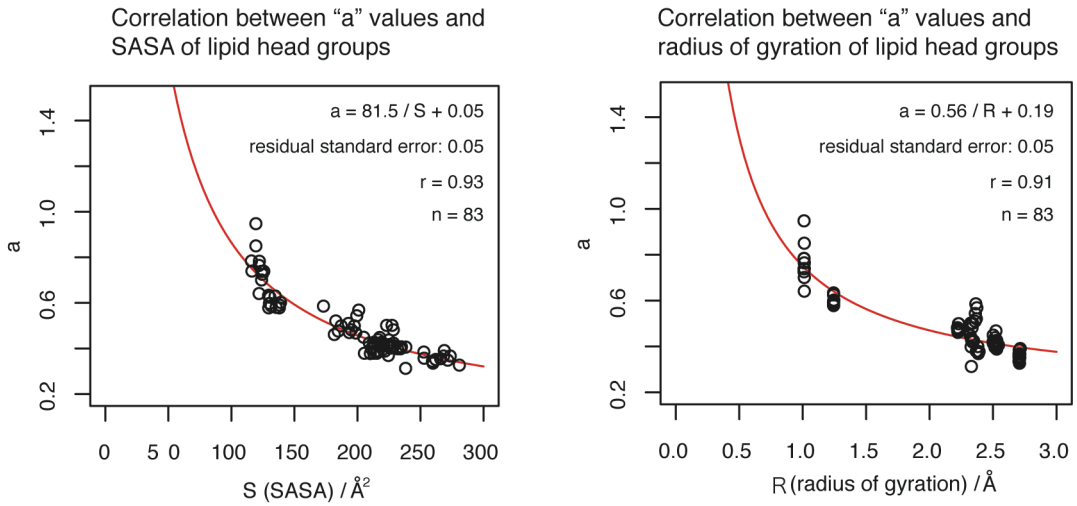


Figure 3.23 Correlation between *"a"* values and sizes of lipid head groups.

Regression curve is colored by red. Regression formula, residual standard error, correlation coefficient, and the number of samples are shown in the upper right corner in each figure.

The distribution of water molecules along the membrane normal at the interface is another common feature (Figure 3.8 and 3.15). We investigated the correlation between the "*a*" values of the glycerolipid and the sphingolipid bilayers (Figure 3.22). The result shows that the correlation is quite high: the correlation coefficient r = 0.98. It demonstrates the strong dependency of the water distribution on the head groups and its independency of the difference in the part of acyl chains and backbones (glycerol or sphingosine). Furthermore, we investigated the correlation between the "*a*" values and the sizes of lipid head groups. Figure 3.23 shows the results when solvent accessible surface area (SASA) per lipid and radius of gyration of a lipid head group are used as the amount of its size. The results shows that the correlation is high in both cases, which also indicates that the water distribution at the membrane-water interface is dependent on the lipid head groups, in particular, on their sizes. It should be noted that all of the distribution of water at the interface resulted in the sigmoidal curve that is a mathematical function of the distance from the membrane surface. This is quite important because it can be applicable to the modeling of a biomembrane. In general, water molecules are generated uniformly to solvate a membrane with a constant density at 300 K. In fact, GLYMM now generates extra water molecules on the membrane with using the bulk density of the water. However, using the sigmoidal curve to generate the water molecules, it would be capable of generating more accurate modeling of water molecules at the membrane interface. SOLVATE,[79] one of the most accurate computer program for solvating a protein with water molecules written by Grubmüller nearly a decade ago, uses a gaussian-form density function to place water molecules around a protein. Because a sigmoidal distribution function is given as an integral or a functional of a gaussian density function, the idea that uses a sigmoidal curve for solvation of a membrane surface may relate strongly with that used in SOLVATE.

We are now building membrane structure database in order to get the various types of bilayers from all over the world. This database will be contributed to the advance of MD simulations of membrane proteins and biomembrane.

III.5 Conclusion

We carried out molecular dynamics simulations of many types of symmetric lipid bilayers composed of glycerolipids or sphingolipids. The characteristics of their water-membrane interfaces, such as distribution and orientation of water molecules and charge densities, were strongly dependent on their head groups. The phase state of the membrane can be distinguished by the correlation between alkyl chains of lipids and the intensity of interdigitation.

IV

MD Simulations of Asymmetric Lipid Bilayers

Abstract

MD simulations of two modeled asymmetric lipid bilayers (model 1 and model 2) were carried out. In model 1, the exoplasmic leaflet comprises POPC and cholesterol, and the opposite face consists of POPE, POPS, and cholesterol. Model 2 has the same concentrations of lipids as model 1 but without cholesterol. The results showed clearly that the cholesterol made the acyl chains of the lipids more ordered and thickened the bilayer. It was also revealed that the asymmetric distribution of lipids affected the distribution of ions strongly, so that more sodium ions were found on the endoplasmic side along the POPS molecules. These results demonstrate that the asymmetric and the mixed characters have significant effects on the lipid-water interface, which would also affect the membrane proteins.

IV.1 Introduction

Simulation studies on membrane proteins have been increasing for the past several years, as described in the previous chapters. Although the membrane proteins in these studies were embedded in explicit lipid bilayers, the bilayers consisted of only one kind of lipid (for example, DMPC, DOPC, POPC, POPE and so on). In cells, however, biomembranes are not composed of a single phospholipid but rather a mixture of many kinds of glycerolipids, sphingolipids, and cholesterol. For example, the human erythrocyte membrane was revealed to include PC, PE, PS, and PI, in addition to SM and cholesterol.[28]

Furthermore, these lipids are not randomly distributed in both sides of the bilayers. Generally, the exoplasmic leaflet of the plasma membrane consists of PC, SM, and GSLs, while PE and PS are concentrated on the endoplasmic side, followed by PI and its derivatives, PIPs. PS, PI, and PIPs are acidic phospholipids with electrostatically negative charges. These negatively charged lipids are the main driving force to attract positively charged molecules such as ions, peptides, and a peripheral type of membrane proteins to the membrane interface. The other types of membrane proteins also exploit the negative charges of lipids for binding to the membrane surface. Rhodopsin, one of the G-protein coupled receptors, has positive charges on the intracellular α-helix running parallel to the membrane surface, and utilizes the electrostatic attractive energy for anchoring the helix strongly. K-Ras, a member of the small G-protein Ras subfamily, has positively charged Lys cluster at the N-terminus region, and makes use of it for localization on a negatively charged lipid microdomain in the endoplasmic side of the membrane surface. In addition, asymmetric distribution of lipids causes the differences of lateral pressure (surface tension) and curvature between the two leaflets. Some proteins are activated through the recognition of these differences, and others keep the appropriate structures owing to them. Thus, it is quite important for membrane proteins to be embedded in membranes with the asymmetric distribution of lipids, at least, with the negatively charged lipids in the endoplasmic leaflet.

Cholesterol (Chol) is also an important molecule to represent a biomembrane as described in *Chapter I*. In contrast to other lipids described above, Chol is distributed uniformly to both leaflets and has unique structure (Figure 1.6). Moreover, Chol has many effects on biomembranes.[42,76,77] When cholesterol is added in L_α phase, it compensates for the lipid-lipid interactions, which leads the membrane to have more packed structure. In contract, added in L_β phase, cholesterol disturbs the packed structure and enlarges the distances among the adjacent lipids. Thus, Chol induce a characteristic phase, l_o phase. Furthermore, it is known that Chol thickens bilayer and

produces lipid microdomain through strong interactions with SM and other sphingolipids. Lipid microdomain plays important role on the membrane proteins, which is also described in *Chapter I*. Moreover, Chol can directly interact with lipid anchored membrane proteins such as small G-proteins, and relate to their localization. Hence, it should be considered when MD simulations of membrane proteins are carried out, in particular, when the study focuses on their localization.

Proteins can exert their functions only under an adequate physiological ion concentration. Hence, it is also necessary for the MD simulations of the biomembranes and membrane proteins to be calculated under the physiological condition.

Thus, there are many things to be considered for MD simulations of biomembranes and membrane proteins. However, no one had reported MD simulation with the asymmetric types of lipid bilayers under the adequate physiological concentration of ions and cholesterols. In this study, the MD simulations of the asymmetric phospholipid bilayers containing POPC, POPE, POPS, and cholesterols (Figure 4.1) were performed under the physiological concentration of NaCl. MD simulations revealed that these properties are effective for stabilizing the membrane structure.

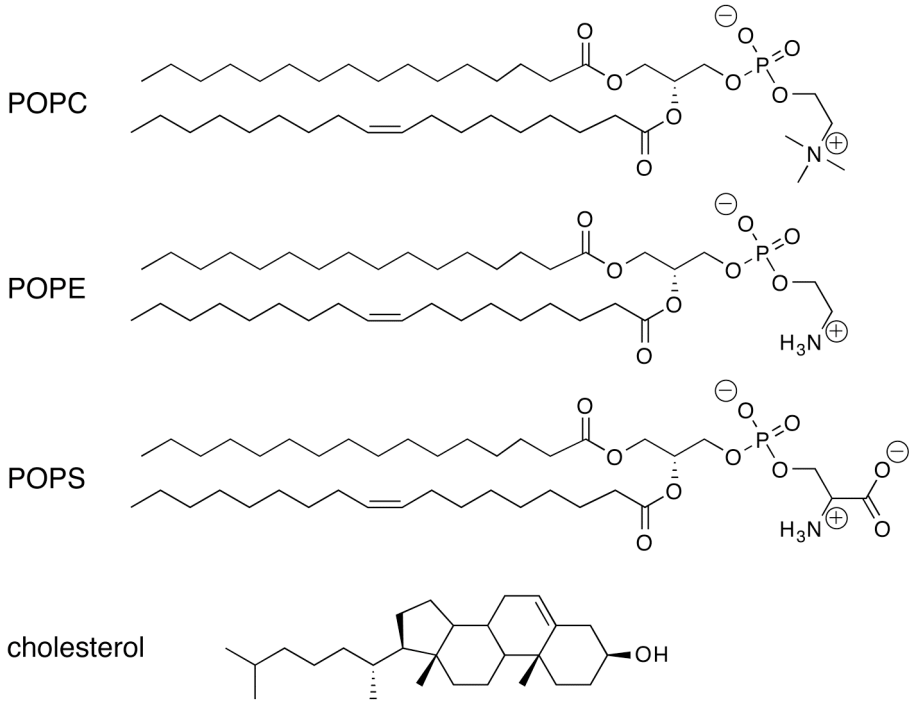


Figure 4.1 Lipids and cholesterols used in our simulations.

IV.2 Computational Details

Modeling of Asymmetric Phospholipid Bilayers

The asymmetric phospholipid bilayers were modeled with GLYMM and VMD.[48] The concentration of POPS was set to match the result of the previous study[28] on the lipid bilayer constitution of a human red blood cell, in which the rate of POPE versus POPS was about 2:1. The system was solvated with TIP3P water molecules.[78] Finally, the membrane had the form in a box by 100 Å × 100 Å × 40Å, and the thickness of water layer was 5Å on each side. Number of lipids, water molecules, and ions are shown in Table 4.1.

All of the counter ions were Na^+ or Cl^- to simplify the effects of the ions on the membrane, although there were other types of ions in a cell. The concentration of NaCl was 460 mM and that of the cholesterols was 22%. The ions were generated by the program called *SOLVATE*[79] and the *LEaP* module of AMBER.[34] There came to be 37 Na+ and 37 Cl- ions in each membrane. Because the system has to be electrically neutral for MD simulation with the PME method under periodic conditions, extra Na^+ ions were added.

The structure of cholesterol molecule was modeled with the program *GaussView*. The electrostatic potential charge of Chol was calculated by Gaussian 98 at the B3LYP/6-311+g(d,p) level and the atomic charge was fitted by the *RESP* module of AMBER. Lipids were replaced at random with Chol with GLYMM.

Table 4.1 Number of lipids, cholesterol, water molecules, and ions in our bilayer models.

	POPC	POPE	POPS	Chol	Water	Na^+	Cl^-
Model 2	137	91	46	0	4432	83	37
Model 1	106	70	35	62	4432	72	37

Details of MD Simulations

MD simulations were carried out with NAMD version 2.4.[69] All atom force fields in CHARMM27[80] were used. A periodic boundary condition was applied, and the pressure and temperature were kept constant by the Nose-Hoover Langevin piston method[81,82] during the simulations. Our simulation procedure was as follows. First, the system was minimized by the conjugate gradient method. Second, the system was gradually heated from 0K to 300K and then kept at 300K until the volume of the system became equilibrated at the NPT ensemble. The non-bonded interactions were calculated by the PME method. The spline interpolation order was 4, and the density of the charge grid spacing was approximately 1 Å. The Ewald error estimation was less than 1×10^{-3} during the simulations. The cutoff distance for the vdW and Coulomb potentials in real space was 12 Å. The integration time step was 1 fs. The total simulation time was 1 ns.

Analyses of MD Simulations

The data were collected every 1 ps for thickness, every 10 ps for distribution functions during 1 ns MD simulation.

A two-dimensional Voronoi tessellation[44] was used for the analysis of the distribution of lipids and the probability density of the Na^+ ions. The tessellation was calculated based upon the time-averaged centers of masses of the lipid molecules projected on the bilayer plane. The coordinates of the sodium ions were collected from the trajectories of MD simulations, and then the two-dimensional binned kernel density estimation with a Gaussian kernel function[71] was used to calculate the probability densities.

All images of the membranes were created by using MolScript[2] and Raster3D.[3]

IV.3 Results

Thicknesses and Whole Structures

The values of thickness of model 1 and model 2 are shown in Figure 4.2. They were approximately constant during our MD simulations. The thickness of model 2 was 39.0 Å on average during the MD simulation (Figure 4.2 and Table 4.2). In comparison, the thickness of model 1 was thicker than that of model 2, which was 40.0 Å on average (Figure 4.2 and Table 4.2).

The whole structures after MD simulation are shown in Figure 4.3. All lipid molecules in both models were randomly oriented. It is notable that many of ions are located at the water-membrane interface (Figure 4.3). Chol molecules are aligned along lipids and oriented with their hydrophilic group to water layers.

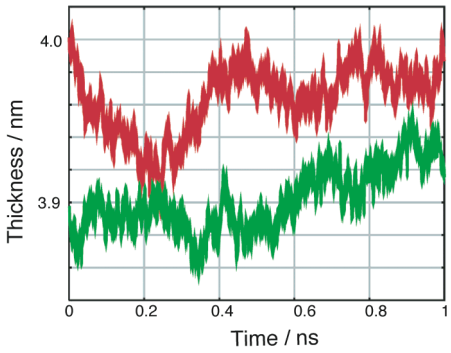


Figure 4.2 Thicknesses of the asymmetric membranes.

The red and green lines show the thickness of the bilayer (model 1, red; model 2 is green)

Table 4.2 Time-averaged thicknesses of the membranes

	Hydrophobic layer			Interface		
	exo	endo	total	exo	endo	total
Model 1	14.2	13.8	28.0	6.0	6.0	12.0
Model 2	14.3	12.4	26.7	6.2	6.0	12.2

Unit is Å.

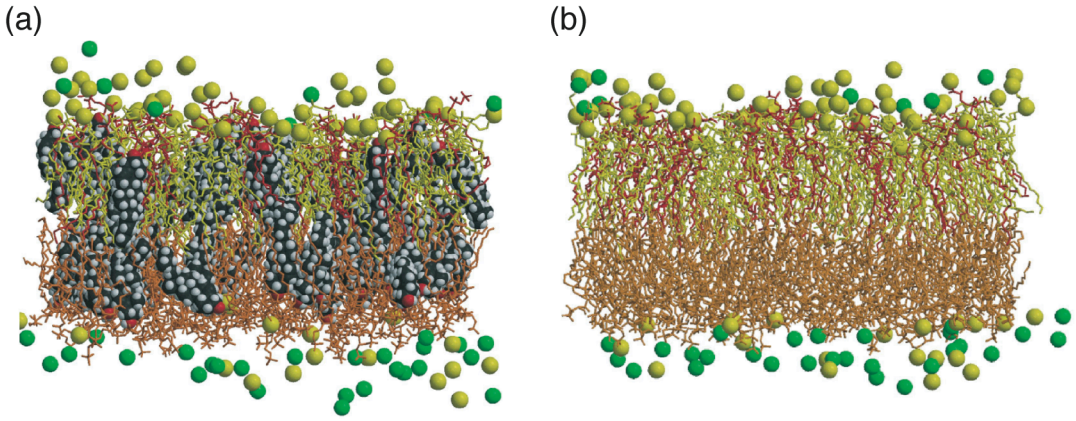


Figure 4.3 Structures of two asymmetric bilayers after MD simulations.

(a) Model 1. (b) Model 2. POPC, POPE, and POPS are shown in orange, yellow, and red lines, respectively. The counter ions and the cholesterols are represented as vdW spheres (C, black; O, red; H, white; Na^+, yellow; Cl^-, green).

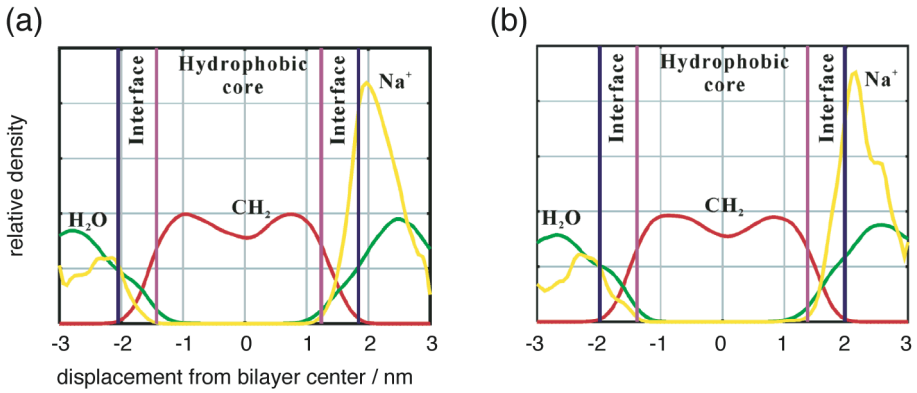


Figure 4.4 Distribution of CH_2 groups, water molecules, and Na^+ ions.

(a) The membrane with ions and without cholesterol. (b) The membrane with ions and cholesterols. The center of the membrane is set at 0 in the abscissa. The area of the positive values endoplasmic side, and that of the minus values is the exoplasmic side. The blue line indicates time-averaged z-coordinate of the phosphorus atoms in the head group of the lipids, and line indicates the time-averaged z-coordinate of the first carbon atoms of the sn-1 position.

Distribution Functions

In order to investigate the effect of the asymmetric distribution of lipids, distribution functions of water molecules and CH_2 groups of acyl chains were calculated. The results are plotted in Figure 4.4. Both membranes have a similar feature that the maximum peak of the distribution functions of water is a little higher in the endoplasmic side than the exoplasmic one. This result indicates that more water molecules are distributed near the surface in the endoplasmic side than in the exoplasmic one. Furthermore, it is surprising that the distribution functions of the CH_2 groups are a little asymmetric in both membranes, although both leaflets have completely the same acyl chains. The result indicated that the CH_2 groups of the endoplasmic side were distributed in a narrower

region than those of the exoplasmic side. The asymmetric distributions are shown in the difference in the thickness of the hydrophobic layers. The values of thickness of the hydrophobic layers of the endoplasmic and the exoplasmic leaflets were 12.4 Å and 14.3 Å, respectively, while those of the interface layers were almost the same (Table 4.2). In both membranes, the thickness of the hydrophobic layer of the endoplasmic side was narrower than that of the exoplasmic side.

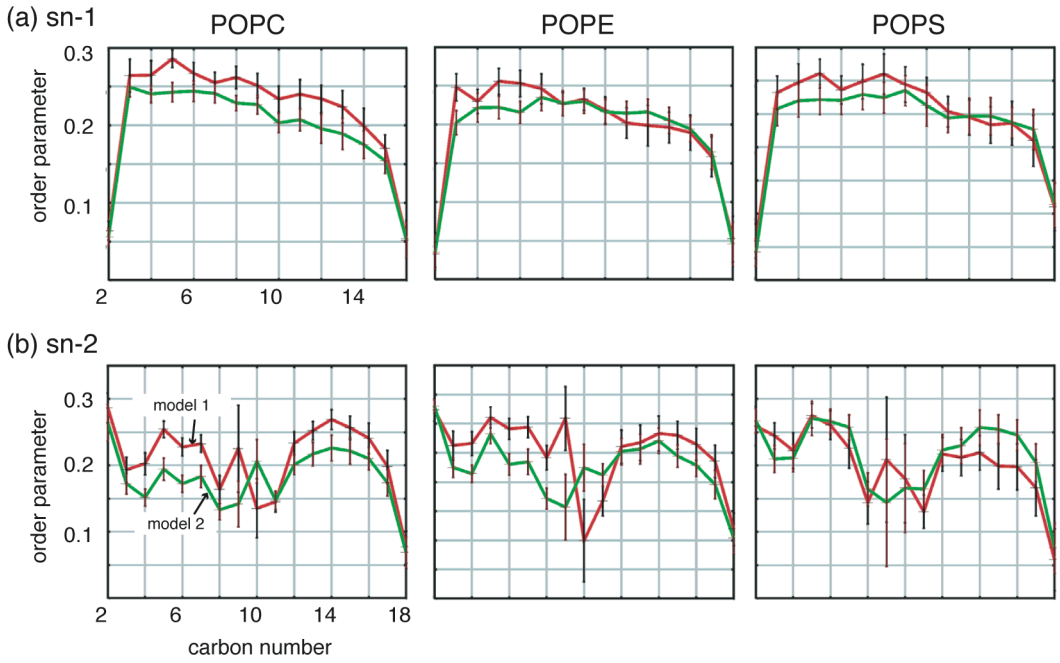


Figure 4.5 Order parameters of each carbon atom in sn-1 and sn-2 chains.

The black and brown lines indicate 95% confidence intervals.

Order Parameters

To investigate effects of Chol on the phospholipid bilayers, we calculated the order parameters, which corresponds to S_{CD} in NMR experiments (Figure 4.5). The plots of the order parameters of lipids have similar shapes in spite of different head groups. The shapes of the *sn*-1 chains are also similar, which decrease gradually from the start in position 2 to the terminal end in position 16. The order parameters of the *sn*-2 chains have two peaks around position 6 and 14, and have the lowest values in position 18 and the second lowest values around position 10, which is the carbon atom forming the unsaturated double bond. It is noted that almost all the order parameters with Chol are larger than without Chol. This result means that Chol makes the lipids in the membrane more ordered.

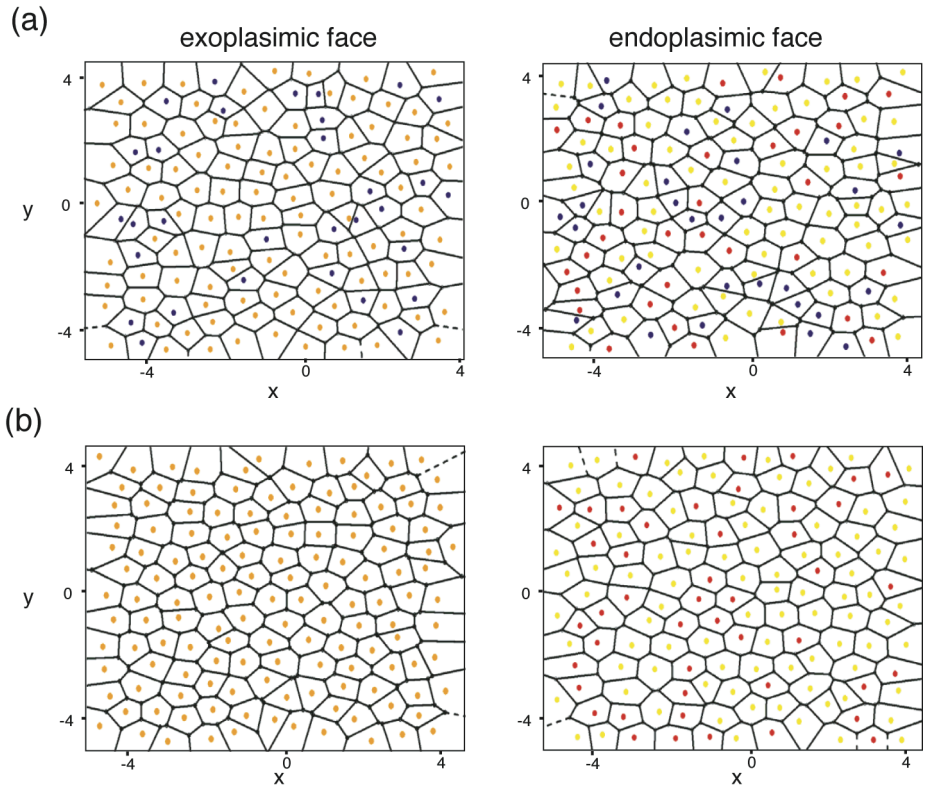


Figure 4.6 Voronoi diagrams of the membranes.

(a) Model 1. (b) Model 2. POPC, POPE, POPS, and Chol are colored orange, yellow, red, and dark blue, respectively. Center of mass of a lipid is represented as a colored point.

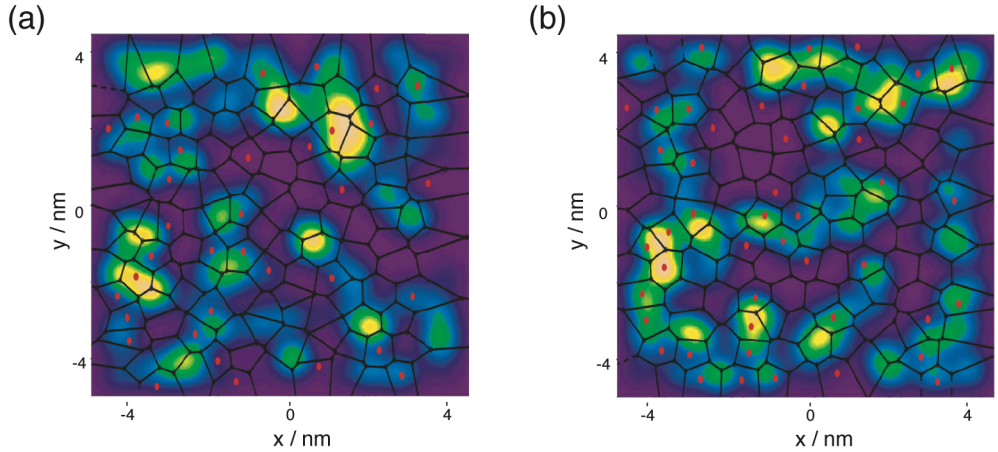


Figure 4.7 Probability densities of Na^+ ions.

(a) Model1. (b) Model 2. The centers of mass of POPS are colored in red and the Voronoi tessellation regions are overlapped.

Voronoi Analysis and Probability Densities of Na^+ Ions

In order to calculate area per lipid in the mixture of lipids and the effects of the Chol on ion distribution, the territories of the lipids were calculated and illustrated by using Voronoi tessellation (Figure 4.6). In the Voronoi diagrams, the lipids are located at random, which shows the character of L_α phase. It is also observed that Chol molecules are frequently very close to a lipid in the adjacent territory.

To investigate the influence of POPS on ion distribution, the probability densities of the Na^+ ions at the endoplasmic sides are calculated. Figure 4.7 shows the results overlapped on the Voronoi tessellation maps. In the figure, the Na^+ ions are mainly distributed along POPS, and are not observed so much around POPE. The result demonstrated that Na^+ ions were likely to localize around the POPS molecules. A close observation confirmed that there are some regions where the probability densities are remarkably large without POPS in model 1. These were accidentally caused by the presence of the Na^+ ions in bulk water, and such Na^+ ions were located far apart from the membrane. Furthermore, a comparison of the probability densities with the Voronoi diagrams shows that the probability densities do not depend upon the distribution of the cholesterols.

Table 4.3 shows area per lipid calculated based on the Voronoi tessellation area. It shows clearly that model 1 has smaller values than model 2. This result is quite interesting when compared to each area per lipid in the symmetric bilayer (Figure 3.7). POPC and POPS in model 2 have approximately same values in their symmetric forms, while POPE has a much larger value than the intrinsic one. However, in the model 1 with Chol, the values of POPC and POPS are smaller than those in their symmetric forms, while the value of POPE approaches its intrinsic one. In both cases, it reveals that Chol occupies the smallest area and decreases the area per lipid as a whole.

Table 4.3 Area per lipid based on Voronoi tessellation

	Model 1					Model 2		
	exo		endo			exo	endo	
	PC	Chol	PE	PS	Chol	PC	PE	PS
mean	56	52	55	56	50	61	62	61
sd	8	10	10	6	9	6	8	7

Unit is $Å^2$.

IV.4 Discussions

We carried out MD simulations of the asymmetric distribution of lipids comprising POPC, POPE, and POPS with Na^+, Cl^-, and Chol. The investigation of the data revealed the effects of Chol and the asymmetric composition of lipids on the membranes.

As shown in Figure 4.2 and Table 4.2, the membrane with Chol is thicker than that without Chol. This result corresponds to the results from the other MD simulation studies of membranes containing Chol.[42,76,77] While the previous studies used concentrations of Chol as high as 50 % in order to strengthen the effects on the membranes, the current study observes the effects on the membrane with only 22 % of cholesterols. The effect of Chol is also recognized in another analysis. Figure 4.5 shows the calculated order parameters of the lipid chains. The analysis demonstrates that the acyl chains in the membrane with Chol are more ordered than those without Chol. These results indicate the effects of Chol making acyl chains align. In fact, this effect of Chol has been reported by other studies.[42,76,77]

It should be noted that the order parameters of POPS are not so affected by the presence of Chol, although POPS molecules directly interact with Chol as shown in Figure 4.6. Figure 4.7 demonstrates that more Na^+ ions are found around POPS molecules. This localization of Na^+ ions results from the coordination of the ions to the negatively charged head groups of POPS. Therefore, a reason why Chol does not influence POPS molecules is possibly due to the coordination of many counter ions to POPS. As the number of Na^+ ions around POPS increases, the motion of the head groups of POPS becomes inert and more rigid. These rigid head groups might affect the conformation of the lipid chain. But there still remains a question why only Chol did not influence the *sn*-2 chains of POPS. Figure 4.5 shows that the *sn*-1 chain of the POPS molecule is affected by Chol the same as those of POPC and POPE. The *sn*-1 and *sn*-2 chains of the lipids used in our simulation have two differences. One is, certainly, the difference in the acyl chains (16:0 or 18:1), and the other is the position bound to the head group. The difference in the types of the acyl chain would affect the interaction with cholesterols because the saturated (16:0) chains would interact more strongly with them than the unsaturated (18:1) ones. This preference of Chol would make the *sn*-1 chains more sensitive to the ordering effect due to Chol than the *sn*-2 chains. The difference of the position can be a factor for the difference in the effect from ions. The *sn*-2 position is nearer the head group than *sn*-1 position. Therefore, this would cause the difference in the strength of the influence from the coordination of Na^+ ions around the head group. Accordingly, it is considered that the *sn*-2 chain would be more ordered without Chol.

In the current study, the asymmetric feature also influences the distribution of the water molecules. The analysis of the distribution functions revealed that the water molecules more frequently appeared near the endoplasmic side of the membrane than the exoplasmic one (Figure 4.4). This would attribute to the fact that the endoplasmic side has large minus charges (in our simulation, -46 and -40 charge units) on the surface of the membrane while the exoplasmic side is electrically neutral. More Na^+ ions are coordinated to the head groups of the POPS, which would result in bringing more water to the endoplasmic side due to the coordination of water molecules to the Na^+ ions.

As shown in this study, it is important for analyzing various characteristics of the membrane to use the asymmetric phospholipid bilayer. Previous MD simulation studies of the membrane proteins were carried out with a symmetric membrane comprising only one kind of lipid. But in the cell, the membrane is composed of many kinds of lipid molecules, and, in addition, its composition is asymmetric. So, the pattern of the interaction of the membrane proteins in each leaflet is very different. Furthermore, it is important for the membrane proteins to be placed correctly in the membrane, because the intracellular domains of membrane proteins interact with the head groups of the endoplasmic lipids. A model of an asymmetric bilayer can be applied to a MD simulation of a membrane protein, which is sensitively affected by the composition and the charge of the membrane surface. As MD simulations of the membrane proteins much increase in the research on their structural changes or the drugs design, it will become more and more meaningful to consider the membrane in itself such as its asymmetry and mixed lipid composition. In this study, we have analyzed one of the aspects of the asymmetric membrane. But more information would be required for elucidation of the characters of the asymmetric membrane. Additional studies in this field should be continued in future.

IV.5 Conclusion

The membrane with Chol is thicker than that with Chol. The order parameters of the lipids in the Chol-containing membrane are larger than those without Chol. These results show that Chol makes lipids align. Asymmetric distribution of lipids also affects water and ion distribution. The probability density of Na^+ ions is high along the distribution of the POPS molecules, and the water molecules are more concentrated in the endoplasmic side where POPS is rich.