Structural adaptations of proteins to different biological membranes

Irina D. Pogozheva, Stephanie Tristram-Nagle [1], Henry I. Mosberg, Andrei L. Lomize *

College of Pharmacy, Department of Medicinal Chemistry, University of Michigan, Ann Arbor, MI 48109-1065, USA

ABSTRACT

To gain insight into adaptations of proteins to their membranes, intrinsic hydrophobic thicknesses, distributions of different chemical groups and profiles of hydrogen-bonding capacities (α and β) and the dipolarity/polarizability parameter (π^*) were calculated for lipid-facing surfaces of 460 integral α-helical, β-barrel and peripheral proteins from eight types of biomembranes. For comparison, polarity profiles were also calculated for ten artificial lipid bilayers that have been previously studied by neutron and X-ray scattering. Estimated hydrophobic thicknesses are 30–31 Å for proteins from endoplasmic reticulum, thylakoid, and various bacterial plasma membranes, but differ for proteins from outer bacterial, inner mitochondrial and eukaryotic plasma membranes (23.9, 28.6 and 33.5 Å, respectively). Protein and lipid polarity parameters abruptly change in the lipid carbonyl zone that matches the calculated hydrophobic boundaries. Maxima of positively charged protein groups correspond to the location of lipid phosphates at 20–22 Å distances from the membrane center. Locations of Tyr atoms coincide with hydrophobic boundaries, while distributions maxima of Trp rings are shifted by 3–4 Å toward the membrane center. Distributions of Trp atoms indicate the presence of two 5–8 Å-wide midpolar regions with intermediate π^* values within the hydrocarbon core, whose size and symmetry depend on the lipid composition of membrane leaflets. Midpolar regions are especially asymmetric in outer bacterial membranes and cell membranes of mesophilic but not hyperthermophilic archaebacteria, indicating the larger width of the central nonpolar region in the later case. In artificial lipid bilayers, midpolar regions are observed up to the level of acyl chain double bonds.

© 2013 Elsevier B.V. All rights reserved.

1. Introduction[2]

Biological membranes provide a functional platform for integral transmembrane (TM) proteins and more temporarily bound peripheral proteins and peptides. Integral membrane proteins constitute a large part of biological membranes ranging from 20% to 80% by mass. They play important roles in vital biological processes including protein synthesis, trafficking, ionic conductance, electron and molecular transport, signal transduction, cell adhesion, cell communication, immune response, respiration, and energy metabolism.

The unique feature of membrane proteins is that they evolve and function in the highly anisotropic lipid environment. Physical and chemical properties of the lipid bilayer are essential for protein structure, functional dynamics, spatial localization and interactions with other proteins and small molecules [1–4]. In particular, the stability of protein complexes is defined by the strength of hydrogen bonds, hydrophobic, electrostatic, and van der Waals forces [5,6], which depend on the local dielectric environment of protein atoms and, therefore, on spatial arrangement of proteins in membranes [7,8].

To ensure solubility of proteins in membranes, polarity of the lipidic phase should match the polarity of embedded proteins. To maintain the functionally required degree of structural flexibility of proteins, the membrane fluidity should be strictly regulated in different cells and in different environmental conditions by adjusting the lipid composition [9]. In addition, the presence of certain lipid species at distinct locations in membranes is essential for proper membrane protein folding, sorting, targeting, and functioning [10–12]. Therefore, maintenance and regulation of compositional diversity of lipids consume a considerable amount of ATP and require proteins encoded by up to 5% of the genome [13].

TM α-helices, β-barrels, and β-helices are the only known protein folds that fulfill the requirement to saturate the hydrogen bonding potential of the polypeptide main chain in the hydrophobic environment. TM α-helical proteins are highly abundant in all types of cellular and intracellular membranes and are encoded by ~25–30% of genes of all sequenced organisms [14]. In contrast, the TM β-barrels are mostly found in outer membranes of bacteria, mitochondria and chloroplasts, and are estimated to be encoded by less than 3% of bacterial genes [15,16]. TM β-barrels are also formed by a number of bacterial pore-forming toxins in host membranes [17]. Single- and double-stranded β-helices were reported for membrane peptides with alternating L- and D-amino acids, such as gramicidin A, B, and C [18].

Due to progress in protein engineering, crystallization, and X-ray diffraction techniques, the number of integral membrane proteins with known three-dimensional (3D) structures is constantly growing [19]. It has currently reached more than 1750 entries in the Protein Data Bank (PDB) [20], or approximately 2% of the PDB content. Most of these entries (82%) correspond to TM α-helical proteins, less than 17% are TM β-barrels, and only around 1.5% are TM β-helices.

What can we learn from available protein structures about their membrane environment? What common features of TM α-bundles and β-barrels allow their general adaptation to the anisotropic lipid environment? What structural features can provide fine-tuning and specific adaptation of proteins to different types of membranes? What topological rules and membrane-sorting signals can be deduced from analysis of protein structures destined to different cellular membranes? Is it possible to characterize physico-chemical properties of different biological membranes with a complex protein and lipid composition based on the structures of their proteins?

To answer these questions, we examined 460 representative structures of integral and peripheral membrane proteins from our OPM (Orientations of Proteins in Membranes) database [21]. The current analysis significantly differs from previously performed studies of statistical distributions of residues in membrane proteins [22–28] in the following aspects: (i) we analyzed separately proteins from eight types of biological membranes using a sufficiently large dataset for each membrane type; (ii) proteins were positioned in membranes by the sufficiently accurate PPM method which has been extensively verified against numerous experimental data; (iii) we analyzed distributions of atoms rather than residues and only on the lipid-facing protein surface; and (iv) we implemented commonly used polarity descriptors of organic solvents (α, β and π*) to define polarity of protein surface and of lipid bilayers.

Analysis of protein atoms rather than whole residues improves the precision and statistical reliability of data: the greater number of atoms allows building the histograms with a 2 Å-step. Previous verification of the PPM method demonstrated a sufficiently high accuracy of calculated intrinsic hydrophobic thicknesses of TM proteins and their tilt angles relative to the membrane plane (1 Å and 2°, respectively), judging from deviations of these parameters in different crystal forms of the same protein [29]. Characterization of biomembranes by polarity parameters α, β, and π* has an important advantage because these parameters have a clear physical meaning as descriptors of dielectric properties and hydrogen-bonding. Besides, these parameters represent integral properties of different lipid-facing atoms and, therefore, are less sensitive than distributions of individual residues to structural and topological biases.

Based on calculated polarity profiles of membrane proteins and model lipid bilayers, we highlight the multilayered organization of the hydrocarbon core with a central nonpolar and two peripheral midpolar regions. We also identified polarity parameters and other structural properties that may reflect general and specific adaptations of proteins to eight different types of biological membranes. These results can be used to quantify anisotropic properties of the lipid environment in these membranes and to improve protein modeling methods.

2. Methodology

2.1. Overall approach to analysis of polarity of membrane components

The analysis of membrane proteins and lipid bilayers presented here is based on general approach to describe molecular solubility that was implemented in the upgraded PPM (Positioning of Protein in Membranes) method [29,30]. PPM allows calculation of binding energies and spatial positions of molecules of different sizes ranging from small organic compounds to large multi-protein complexes in membranes. The method was successfully validated using data for 24 TM and 42 peripheral proteins and many peptides whose arrangements in membranes have been experimentally studied [29–31].

The PPM method combines an all-atom protein structure with an anisotropic solvent representation of the lipid bilayer and the universal solvation model [32]. The solvation model describes the transfer energy of an arbitrary chemical compound from water to an organic solvent or another fluid environment. It accounts for long-range electrostatic interactions and first-solvation-shell effects (van der Waals, hydrophobic and hydrogen bonding interactions).

We found that the polarity of the solvent can be adequately described by a few commonly used parameters: its dielectric constant (ε), the solvatochromic dipolarity/polarizability parameter (π^*) [33], and hydrogen-bonding donor (α) and acceptor (β) parameters of Abraham [34]. The α and β parameters have been previously used in SMx implicit solvation models developed for isotropic solvents [35]. We have extended this approach to anisotropic environments [30]. Accordingly, the lipid bilayer was represented as a fluid anisotropic solvent with polarity properties described by profiles of α, β, ε and/or π^* parameters.

Hence, in the present work, we examined solubility properties of membrane proteins by calculating profiles of polarity parameters, α, β, and π^*, for the lipid-facing surfaces of membrane proteins together with distributions of polar and nonpolar protein atoms, "hydrophobic dipoles" of Tyr and Trp residues, positively and negatively charged ionizable groups, crystallized lipids, detergents and water. The solvatochromic parameter π^* replaces the macroscopic dielectric constant because it better describes microscopic dielectric properties of the environment and can be more easily calculated than the dielectric constant. In addition, we calculated polarity profiles for ten model lipid bilayers and compared them with profiles of membrane proteins.

2.2. Calculations of polarity profiles for model lipid bilayers

Transbilayer profiles of parameters α (z), β (z), and π^*(z) and dielectric function F(ε)(z), describe the changes of polarity across the lipid bilayer [30]. These functions are used by the PPM method to define spatial positions of proteins in membranes. The profiles were previously calculated for the fluid dioleoyl-phosphatidylcholine bilayer (DOPC) using distributions of lipid quasi-molecular segments obtained from neutron and X-ray scattering data [36]. The concentration of water in the lipid acyl chain region of the DOPC bilayer was evaluated based on spin-labeling data [37].

Similar polarity profiles can be calculated for any other model lipid bilayer with known distributions of lipid components along the membrane normal. Here we compared ten lipid bilayers that have been previously studied in the fully hydrated fluid liquid-crystalline (L_α) phase that is biologically relevant (Table 2) [36,38–43]. Structural parameters for these bilayers were determined from X-ray scattering analysis, sometimes supplemented by a simultaneous fitting to neutron diffraction data [36]. The structure of each lipid bilayer is represented by Gaussian distributions of a number of lipid fragments with maxima indicating the most probable location of these fragments and width indicating range of their thermal motion along the bilayer normal. The distribution of water was obtained by subtracting concentrations of all

other lipid components and assuming that total probability is equal to 1 at each point across the bilayer.

Distributions of volume probability of lipid components and corresponding parameters $\alpha(z)$, $\beta(z)$ and $\pi^*(z)$ were calculated as previously described [30]. Most lipids were represented as a combination of total hydrocarbon ("CH2") component, carbonyl–glycerol groups ("CG"), and the remainder of lipid head group ("P") based on X-ray scattering data (Tables 2, S3 and S4). A more detailed structural representation of the lipid bilayer was made for DOPC and POPG bilayers. It includes the locations of double bonds ("CH") established by neutron scattering and an additional peak for lipid head group (for example, "PG1" and "PG2" in POPG). The presence of small amount of water in the hydrocarbon region observed in ESR studies [37] was not taken into account. Incorporation of this water, as in our previous work [30], leads to the increases of parameters α, β and π^* in the midpolar region of the lipid bilayer.

To understand the contribution of different factors to the polarity parameters, we compared bilayers formed by lipids with different acyl chain lengths, such as dilauroyl-phosphatidylcholine (diC12:0PC, DLPC), dimyristoyl-phosphatidylcholine (diC14:0PC, DMPC), dipalmitoyl-phosphatidylcholine (diC16:0PC, DPPC), DOPC (diC18:1 PC), dierucoyl-phosphocholine (diC22:1PC, *DEPC*); with fully saturated and monounsaturated acyl chains; with ester-linked lipids, such as DOPC, dipalmitoyl-phosphatidylcholine (di16:0PC, DPPC), palmitoyl-oleoyl-phosphatidylcholine (C16:0-18:1PC, POPC), palmitoyl-oleoyl-phosphatidylglycerol (C16:0-18:1PG, POPG), and ether-linked lipids, such as dihexadecyl-phosphocholine (diC16:0ePC, DHPC); with zwitterionic (PC) and anionic head groups (PG), and some with branched acyl chains, such as diphytanoyl-phosphocholine (di(16:0(3me,7me,11me,15me)PC). The multicomponent membrane included a mimic of the eukaryotic plasma membranes, LM$_3$ bilayer, which is composed of palmitoyl-oleoyl-phosphatidylcholine (C16:0-18:1PC, POPC), palmitoyl-oleoyl-phosphatidylethanolamine (C16:0-18:1PC, POPE), palmitoyl-oleoyl-phosphatidylserine (C16:0-18:1PS, POPS), phosphatidyl-inositol (PI), sphingomyelin (SM), and cholesterol at molar ratio 10:5:2:1:2:10.

2.3. Protein dataset

Membrane proteins used in this work were taken from the OPM database [21], which provides calculated spatial arrangements in a model lipid bilayer of integral and peripheral proteins from the PDB. OPM also includes classification of membrane proteins into different families, superfamilies and classes, their topology and intracellular localization, which greatly facilitates protein analysis. The OPM database currently contains 671 distinct representative structures of TM proteins and multi-protein complexes, 1088 distinct structures of peripheral proteins, and 291 peptide structures, which cover 6410 PDB entries (http://opm.phar.umich.edu, release as of 02/01/13). Though the database contains proteins from 23 types of biological membranes, statistically significant sets of these proteins can be found for only eight membrane types: plasma membranes (PM) of archaebacteria, eubacteria (Gram-negative and Gram-positive), and eukaryotic cells, membranes of endoplasmic reticulum (ER), thylakoid membranes, mitochondrial inner membranes (MIM), and outer membranes (OM) of Gram-negative bacteria.

The set of selected TM α-helical proteins includes 191 structures associated with seven membrane types (all types except the bacterial OM with β-barrel proteins) (Table 1). This set represents different functional classes of α-helical proteins, including receptors, channels, transporters and enzymes (Table S1). Families of closely homologous proteins (with sequence identity higher than 85%) were represented by only one structure. The chosen cutoff for sequence identity was relatively high, because lipid-facing residues are rather variable, even in closely homologous proteins. These selected proteins usually contain co-crystallized lipids, water, ligands, and co-factors.

Membrane α-helical proteins frequently oligomerize and create large multiprotein complexes involved in vital cellular functions. Protein oligomerization is biologically important, as it usually increases protein stability, creates additional active sites between subunits, allows cooperative interactions of subunits, increases enzymatic and transport efficiency, and provides an additional level of regulation [44]. Therefore, the majority of selected proteins represent functional complexes formed by homo- or heterooligomers (Fig. 1, Table S1).

The set of β-barrel proteins includes 68 single-chain β-barrels (β-I) from the OM of Gram-negative bacteria. The barrels are formed by 8 to 24 antiparallel β-strands that enclose a central pore. These proteins belong to several functional classes, such as channels, transporters, enzymes, adhesion molecules, and components of secretion systems. Several multi-chain β-barrels (β-II) from the TolC-like bacterial secretion system (1EK9, 1YC9, 1WP1, 3PIK) where analyzed separately, since they have a very different structure. These proteins fold into a trimeric 12-stranded TM β-barrel and a large α-helical water-soluble domain in the periplasmic space. The distribution of charged residues of TolC-like proteins is different from that in typical β-I barrels (see Section 3.2.3). TM β-barrel proteins also form oligomers [44]. While oligomerization of single-chain β-barrels in dimeric or trimeric complexes is quite frequent, but not obligatory, oligomerization of subunits of multi-chain β-barrels is mandatory for structural integrity and membrane insertion of these proteins.

The set of peripheral membrane proteins includes 196 structures of proteins from five membrane types: PM of eukaryotic cells (104 structures) and archaebacteria (6 structures), IM of Gram-negative bacteria (26 structures) and mitochondria (15 structures), and

Table 1
Sets of protein structures used for analysis of polarity profiles.[a]

Membrane type	TM protein # (subunit #)*		Peripheral protein #		$D_{aver} \pm$ r.m.s.d. (Å)	D_{min} (Å)	D_{max} (Å)
	OPM	Analysis	OPM	Analysis			
TM α-helical proteins							
Plasma membrane (PM) of eukaryotic cells	134	50 (110)	333	104	33.5 ± 3.1	27.2	40.8
Endoplasmic reticulum (ER) of eukaryotic cells	28	10 (18)	74	21	30.2 ± 1.9	27	33.8
Plasma membrane (PM) of Gram-positive bacteria	36	12 (43)	37	12	31.6 ± 3.4	27.4	38.2
Plasma membrane (PM) of archaeabacteria	35	20 (49)	7	7	30.6 ± 2.2	28.2	36.4
Inner membrane (IM) of Gram-negative bacteria	163	82 (301)	75	26	30.2 ± 1.9	26.8	37.2
Mitochondrial inner membranes (MIM)	19	9 (71)	25	15	28.6 ± 1.4	26.8	30.8
Thylakoid membrane	17	8 (107)	11	5	30.7 ± 2.1	27.4	33.6
TM β-barrel proteins							
Outer membrane (OM) of Gram-negative bacteria	104	68 (94)	16	6	23.9 ± 1.7	20.4	28.4

[a] Proteins from different membrane types were selected from the OPM database, which provides coordinates of the three-dimensional structures together with intrinsic hydrophobic thicknesses (D_{aver}, D_{min}, D_{max}) calculated by PPM. Proteins used for the structural analysis represent only a part of corresponding proteins from OPM, because structures of close homologues were excluded. TM α-helical proteins include both single-spanning and multi-spanning proteins. Bacterial OM proteins include TM β-barrels of β-I type. Many structures consist of multiple individual polypeptide chains (subunits). The total numbers of subunits in each set are indicated in parentheses.
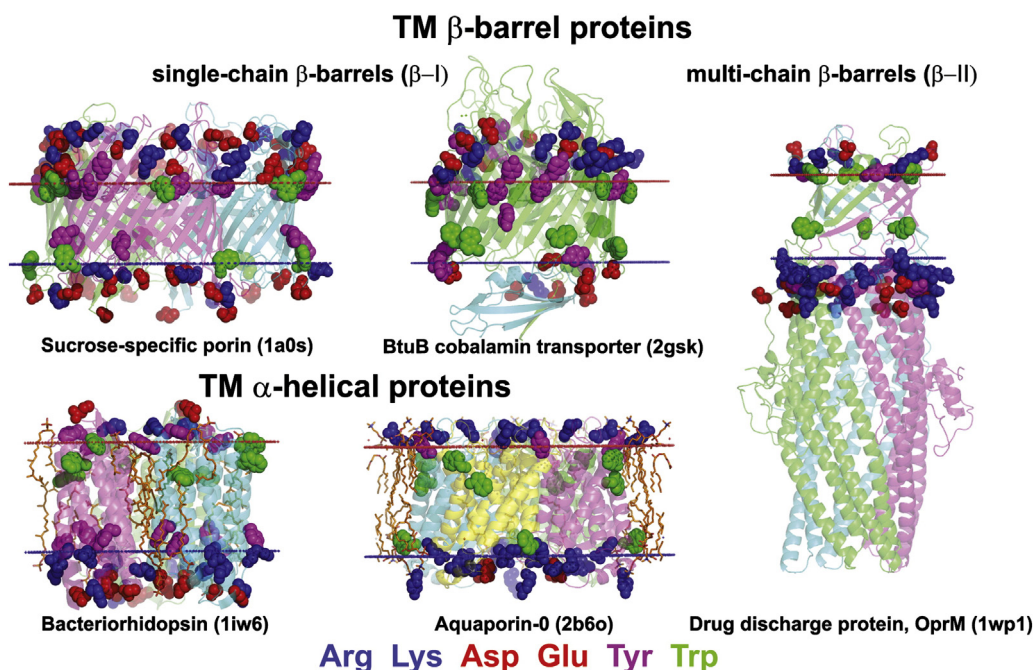
Fig. 1. Different types of integral membrane proteins positioned in membranes: TM α-helical proteins, single-chain β-barrels (β-I type) and multi-chain β-barrels (β-II type). Lipid-facing residues are shown by colored spheres: Lys and Arg (blue), Asp and Glu (red), Tyr (purple), Trp (green). Co-crystallized lipids are shown by sticks colored orange (C and P atoms), red (O-atoms) and blue (N-atoms). Hydrophobic boundaries calculated by PPM are shown by horizontal lines: blue (for PM cytoplasmic side or OM periplasmic side) and red (for PM extracellular side or IM periplasmic side). Cartoon representations of proteins are colored by chain. All selected proteins represent oligomers: homotrimer of sucrose-specific porin, heterodimer of BtuB cobalamin transporter, homotrimer of drug discharge proteins OprM, homotrimer of bacteriorhodopsin, homotetramer of aquaporin-0. Trimerization of OrpM is required to form TM β-barrel. Distributions of Tyr, Trp, Lys, Arg, Asp, and Glu on the surface of membrane proteins are clearly nonuniform with charged residues accumulated in the lipid headgroup regions, and Trp and Tyr residues located near hydrophobic boundaries inside and outside of the hydrocarbon region, respectively.

thylakoid membranes (5 structures). Among these proteins are monotopic proteins that are deeply inserted into the hydrocarbon core of the lipid bilayer, as well as peripheral proteins weakly bound to the membrane surface, some of which require the presence of specific lipids to ensure the efficient binding to the particular membrane type. The majority of selected peripheral proteins function as enzymes, carriers of nonpolar substances, electron carriers, or membrane-targeting domains.

2.4. Calculation of distributions of atoms and polarity parameters in protein structures

Distributions along the membrane normal were calculated for different atoms and atomic groups of lipid-facing residues in 3D structures of α-helical and β-barrel TM proteins from different biological membranes (Table S1). Distributions were analyzed separately for each category of protein atoms, such as "polar atoms" (N- and O-atoms of side chains and main chains of all residues), "nonpolar atoms" (C- and S-atoms from side chains of Val, Leu, Ile, Met, Cys, Phe, Tyr, and Trp residues), "aromatic atoms" (C-atoms from aromatic rings of Tyr. Phe and Trp), and charged atomic groups (amine group of Lys, guanidinium group of Arg, carboxyl group of Asp and Glu).

To obtain "intrinsic" hydrophobic thicknesses of proteins required for current analysis, spatial positions of proteins in membranes provided by OPM were recalculated by PPM 2.0 [29] while omitting penalty for hydrophobic mismatch. Separate distributions for single-spanning (bitopic) proteins were calculated by combining structures of individual bitopic proteins positioned in membranes with single-spanning TM subunits from large protein complexes using their orientations in complexes.

The atomic distributions describe changes in surface fraction (concentration) of lipid-facing protein atoms. All solvent-inaccessible protein groups and groups within internal cavities were excluded, as previously described [29]. The surface concentration of atom i was

determined by averaging lipid accessible surface area (ASA) of the corresponding atom in a protein set:

$$c_i(z) = ASA_i(z)/ASA_{total}(z) \qquad (1)$$

where $ASA_i(z)$ is the ASA of atoms in the slice [z-δ; z + δ] (δ = 1 Å), and $ASA_{total}(z)$ is the total ASA of all atoms in the slice for the protein set.

To analyze distributions of charged groups and net charge, the residue fraction (or number of charges) was used instead of surface fraction:

$$c_i(z) = N_i(z)/N_{total}(z) \qquad (2)$$

where $N_i(z)$ is the number of the corresponding solvent-accessible charged group in the slice, and $N_{total}(z)$ is total number of all charged residues in the slice.

Distributions of co-crystallized water were normalized by the number of lipid-facing protein residues in the slice. Distributions of co-crystallized lipids and detergents were not normalized and, therefore, are not based on surface concentrations but on number of atoms. Molecules of water, lipids and detergents within water-filled TM channels were excluded. Only polar (non-carbon) atoms of lipids and detergent were used for analysis of the distributions. Three distributions were generated for lipid atoms separated into the following categories: (a) glycerol/carbonyl groups; (b) P and O atoms of lipid phosphates or structurally equivalent groups, and (c) head group atoms, such as choline or ethanolamine.

Average values of parameters α, β and π* for the lipid-facing protein surface per Å² were calculated in a similar fashion. For example,

$$\alpha(z) = \frac{\sum\limits_{j\in[z;z+\delta]} \alpha_j ASA_j}{ASA_{total}(z)} \qquad (3)$$

where α_j represents the value of H-bond donor parameter α for protein group j that belongs to slice $[z-\delta; z + \delta]$. The values of α, β and π^* for different chemical groups (Table S2) were based on tabulated values [33,34,45–47].

2.5. Approximation of distributions for proteins by sigmoidal curves

Distributions of polarity parameters (α, β and π^*), different protein atoms and groups, co-crystallized water, lipids, and detergents were approximated by analytical curves, whose parameters were defined by fitting. We tested several types of functions for describing the distributions (Lorentz, Gauss and sigmoidal curves) and found that sigmoidal functions provide the best fitting. The sigmoidal curves are commonly used to describe transitions between two media, such as water and a nonpolar solvent [48]. Hence, a successful approximation by the sigmoidal curves may indicate that different membrane regions (lipid head group region, nonpolar hydrocarbon region, and midpolar region) can be treated as separate phases with different properties.

Each sigmoidal curve was defined by four adjustable parameters: background values in the corresponding media (a and b), the middle point separating two media (z_0), and steepness of the transition (exponential decay parameter λ):

$$c(z) = b + \frac{a}{1 + \exp(z-z_0)/\lambda} \qquad (4)$$

To describe complex distributions with two asymmetric peaks (e.g. for Tyr and Trp atoms, charged groups, lipid atoms, and water molecules), four sigmoidal curves were simultaneously fitted to match these distributions. The fitting was accomplished by grid scan to minimize root-mean-square deviations (r.m.s.d.) between the

observed (Eqs. (1)–(3)) and calculated (Eq. (4)) values of parameters $c(z)$, $\alpha(z)$, $\beta(z)$ or $\pi^*(z)$. The r.m.s.d. were usually calculated in the interval of ± 40 Å around the membrane center. Parameters a and b of four curves were restricted to provide a continuous curve and identical background values in water on the both sides of the lipid bilayer.

3. Results

3.1. Polarity profiles of artificial lipid bilayers

Polarity parameters (α, β, π^*) calculated for ten artificial bilayers abruptly change at the hydrocarbon boundary, which corresponds to the midpoint of distributions of aliphatic groups (Figs. 2, S1, S2). The inflection points on the profiles define the acyl chain boundary, the D_C distance from the bilayer center. The location of lipid carbonyl–glycerol groups ("CG") matches D_C only in certain bilayers, most notably, in LM$_3$ and POPG bilayers, which include negatively charged lipids (Table 2). In DOPC, DPPC, and DPhyPC bilayers, the maxima of "CG" groups are also rather close to the midpoints of aliphatic distributions (within 1 Å), while in other lipid bilayers (DHPC, DLPC, DMPC, DEPC), the "CG" maxima are shifted by 2–3 Å outside the hydrophobic boundaries. Unlike the carbonyl groups, the peaks of lipid phosphate distributions in all bilayers are located at approximately the same 5 Å distance outside the hydrocarbon boundaries.

The relative thickness of the hydrocarbon region primarily depends on the length of the lipid acyl chains: it increases in the order DLPC < DMPC < DPPC < DEPC (Fig. S1). A significant increase of the thickness was also observed in the multicomponent system, LM$_3$, which has a larger thickness ($2D_C = 32.6$ Å) than the single-component DOPC bilayer ($2D_C = 28.8$ Å), even though both bilayers mostly have acyl chains of similar length and saturation level (C18:1) (Fig. S2, Table 2). The larger thickness in LM$_3$ is due to its
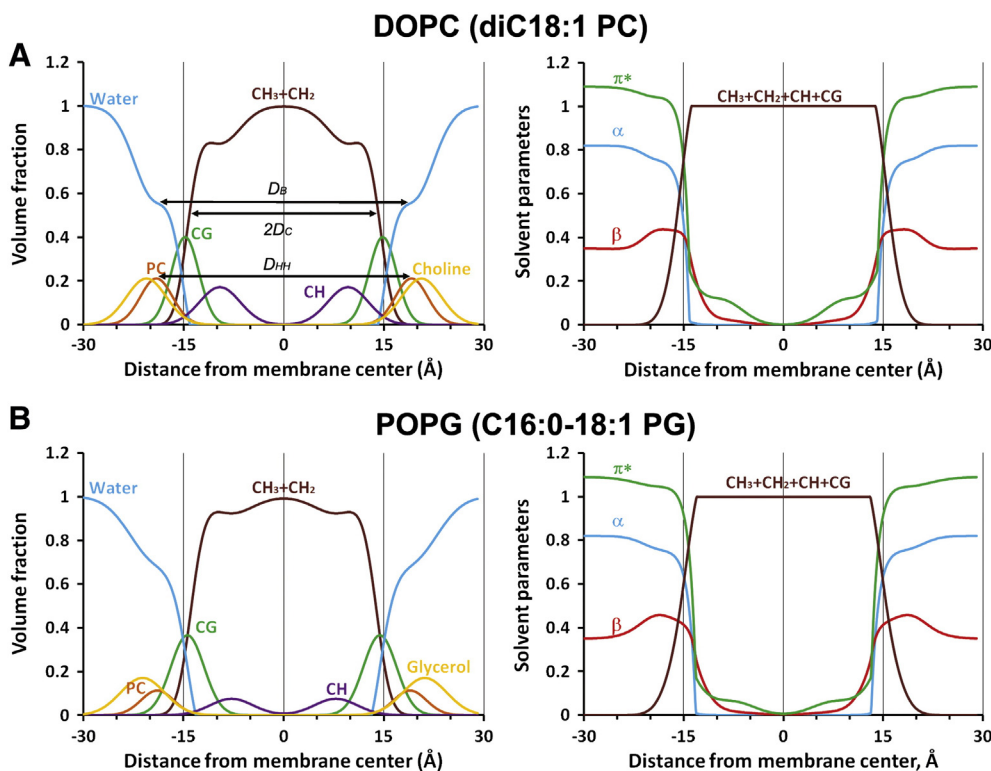

Fig. 2. Structures of DOPC (A) and POPG (B) bilayers. Distributions of lipid segments determined by the simultaneous analysis of X-ray and neutron scattering data (left panels). Volume probability distributions of various lipid components were determined assuming that the total probability is equal to 1 at each point across the bilayer. Changes of polarity parameters, hydrogen bonding donor (α), acceptor (β) capacities and dipolarity/polarizability parameter (π^*) along the membrane normal (right panels). Structural parameters are indicated by arrows: hydrophobic thickness ($2D_C$), distance between lipid head groups (D_{HH}), total bilayer thickness (hydrocarbon chains plus head groups) (D_B). Parameters used for calculation of these profiles and references are provided in Tables 2, S3 and S4.

Table 2
Lipid structural data from X-ray and neutron scattering studies: hydrophobic thickness ($2D_C$), distance between lipid head groups (D_{HH}), total bilayer thickness (hydrocarbon chains plus head groups) (D_B), and lipid lateral area (A) for various bilayers studied at temperature corresponding to liquid-crystalline (L_α) phase.[a]

Lipid	T (°C)	"CG"		"P"		"CH$_3$ + CH$_2$ + CH"		D_B (Å)	A, (Å²)	Reference
		Z_m (Å)	S_m (Å)	$Z_m = \frac{1}{2}D_{HH}$ (Å)	S_m (Å)	$Z_{HDC} = Dc$ (Å)	S_{HDC} (Å)			
DOPC (diC18:1 PC)*	30	14.8	2.050	18.4	2.41	14.40	2.48	38.7	67.4	[36][b]
POPG (C16:0-18:1 PG)*	30	14.4	2.480	18.7	2.45	13.85	2.80	36.5	66.0	[42][c]
DLPC (diC12:0 PC)	30	10.9	1.730	15.4	2.00	10.50	2.80	31.3	63.2	[38]
DMPC (diC14:0 PC)	30	13.1	1.900	17.7	2.10	12.70	2.80	36.3	60.6	[38]
DPPC (diC16:0 PC)	50	14.7	2.100	19.0	2.55	14.20	2.80	39.0	63.1	[36][b]
DEPC (diC22:1 PC)	30	18.6	2.535	22.2	2.60	17.2	3.66	44.0	69.3	[43]
POPC (C16:0-18:1 PC)	30	14.9	2.555	18.5	2.75	13.60	2.90	36.8	68.3	[43]
DHPC (diC16:0e PC)	48	15.1	2.655	19.1	2.60	13.9	2.80	37.6	65.1	[39]
DPhyPC (diC16:0 (3me,7me,11me,15me)PC)	30	13.8	2.300	18.2	2.15	13.60	2.80	35.4	80.5	[41]
LM$_3$ (POPC:POPE:POPS: PI:SM:Chol = 10:5:2:1:2:10)	30	17.3	2.720	22.0	2.385	16.30	2.90	33.0	73.3	[40]

[a] Z_m and S_m define the locations and widths of corresponding Gaussians, respectively. Z_{HDC} and S_{HDC} are parameters of Gaussian error function. D_B equal to $2V_L/A$ defines the hydrophobic bilayer thickness, where V_L is volume per lipid, A is lateral area per lipid.
[b] "P" is P(O)$_4$CH$_2$–CH$_2$–N segment. Additional parameters were used for CH and CholCH3 groups.
[c] Experimental data based on 3G SDP model. "P" is PG1 (PO$_4$) segment. Additional parameters for CH and PG2 groups are shown in Table S4.

33 mol% cholesterol content, which is known to significantly thicken lipid bilayers [49,50].

The polarity parameters studied behave differently along the membrane normal: α rapidly drops from a relatively steady level in the head group region to zero value at the hydrocarbon boundary, β has a maximum in the head group region and changes more gradually, while π* also changes gradually at the boundary and may have a wide shoulder within the hydrocarbon region (Fig. 2, right panels). The maxima of β originate from the relatively high hydrogen bonding acceptor capacity of lipid head groups, primarily phosphates. The more gradual changes of β and π* are likely caused by the presence of ester/ether linkages between head groups and acyl chains, as well as by the remaining amount of water. The values of β and π* vary in different bilayers in accordance with changes in relative positions and amplitudes of lipid phosphate and carbonyl groups (Figs. 2, S1, S2).

An additional region within the hydrocarbon core with an increased value of π* originates from the peaks for lipid double bonds that are included in more advanced, neutron diffraction-based models describing DOPC and POPG bilayers. This region is located between the lipid carbonyls and double bonds. Its polarity can be described by π* values intermediate between those for wet octanol and dibutylether (π* = 0.4 and π* = 0.18, respectively [32]). We name this region with intermediate polarity as a midpolar region.

Thus, any lipid bilayer can be described as an anisotropic medium with five regions characterized by distinct dielectric and hydrogen-bonding parameters: two head group regions and two midpolar regions that enclose a central nonpolar region. We also found that lipid double bonds are responsible for the complex behavior of dipolarity/polarizability parameters π* within the hydrocarbon core.

3.2. Distributions of lipid-facing protein groups and polarity profiles of membrane proteins

Our studies of model lipid bilayers demonstrated that polarity of membrane systems with known lipid composition can be quantified using transbilayer profiles of a few polarity parameters: α, β, and π*. However, natural biological membranes are significantly more complex systems than model bilayers, because they have highly diverse and yet unidentified lipid and protein composition. To describe polarity profiles and structural asymmetry of different biomembranes, we analyzed the polarity parameters (α, β, and π*) of TM protein surfaces (Figs. 3JKL and 4JKL) rather than of surrounding lipids. These parameters can be easily calculated from distributions of lipid-facing protein atoms for a set of proteins originated from the corresponding membranes. The obtained polarity profiles of proteins are expected to match microscopic dielectric properties of their native biomembranes.

Thus, to characterize polarity, surface charge, and thicknesses of eight natural membranes, we performed analysis of relatively large sets of TM proteins with known 3D structures, defined topology, and accurately pre-calculated orientations with respect to the lipid bilayer plane that were associated with each of these membrane types. In particular, we studied distributions of lipid-facing protein atoms sorted by categories, such as polar, nonpolar, aromatic, and charged atoms or atomic groups, as described in Methodology (2.4). Analysis was performed separately for each membrane studied and for each structural protein type (α-helical, β-I barrel or β-II barrel). The major observations made during the analysis of distribution of lipid-facing protein atoms or atomic groups and polarity profiles of membrane proteins are described below.

3.2.1. Common features of membrane proteins

Analysis of distributions of protein atoms and polarity parameters of TM α-helical proteins from different membranes (Figs. 3–7) revealed several common features which apparently reflect the general adaptation of all TM proteins to the hydrophobic environment.

The first common feature of TM proteins is the presence of an extended hydrophobic surface of approximately 30 Å-thickness rich in aliphatic and aromatic amino acid residues. The surface fractions of polar protein atoms, nonpolar atoms, and crystallized water abruptly change at both sides of this hydrophobic zone (Figs. 3ABC, 4ABC). The midpoints of analytical curves used for approximation of these distributions coincide reasonably well with each other and with the intrinsic hydrophobic thickness of TM proteins calculated by PPM (Fig. 5). Importantly, the distances between all these midpoints change in a synchronized fashion, following changes in the average hydrophobic thickness for proteins from seven membrane types.

We assume that borders of the intrinsic hydrophobic surface of TM proteins calculated by the PPM method, as well as midpoints of curves for polar and nonpolar atoms, generally match hydrocarbon boundaries of lipid bilayers accommodating these proteins. This assumption is supported by the observed distribution of ether/ester oxygen atoms ("CG" groups) of lipids co-crystallized with TM α-helical proteins (Fig. 6) that shows maxima at approximately ±15 Å-distances from the membrane center. The distance between two "CG" maxima indicates the thickness of the hydrocarbon core of the lipid bilayer (2Dc in Fig. 2A), which is consistent with the average hydrophobic thickness of all TM α-helical proteins combined (31.2 ± 2.8 Å).

The second feature of TM proteins is the highly asymmetric distribution of polar residues along the membrane normal (Figs. 3AGH, S3). Polar residues are practically absent in the middle of the membrane, but are abundant in the area of lipid head groups. One exception is the distribution of surface-exposed Ser residues which are rather
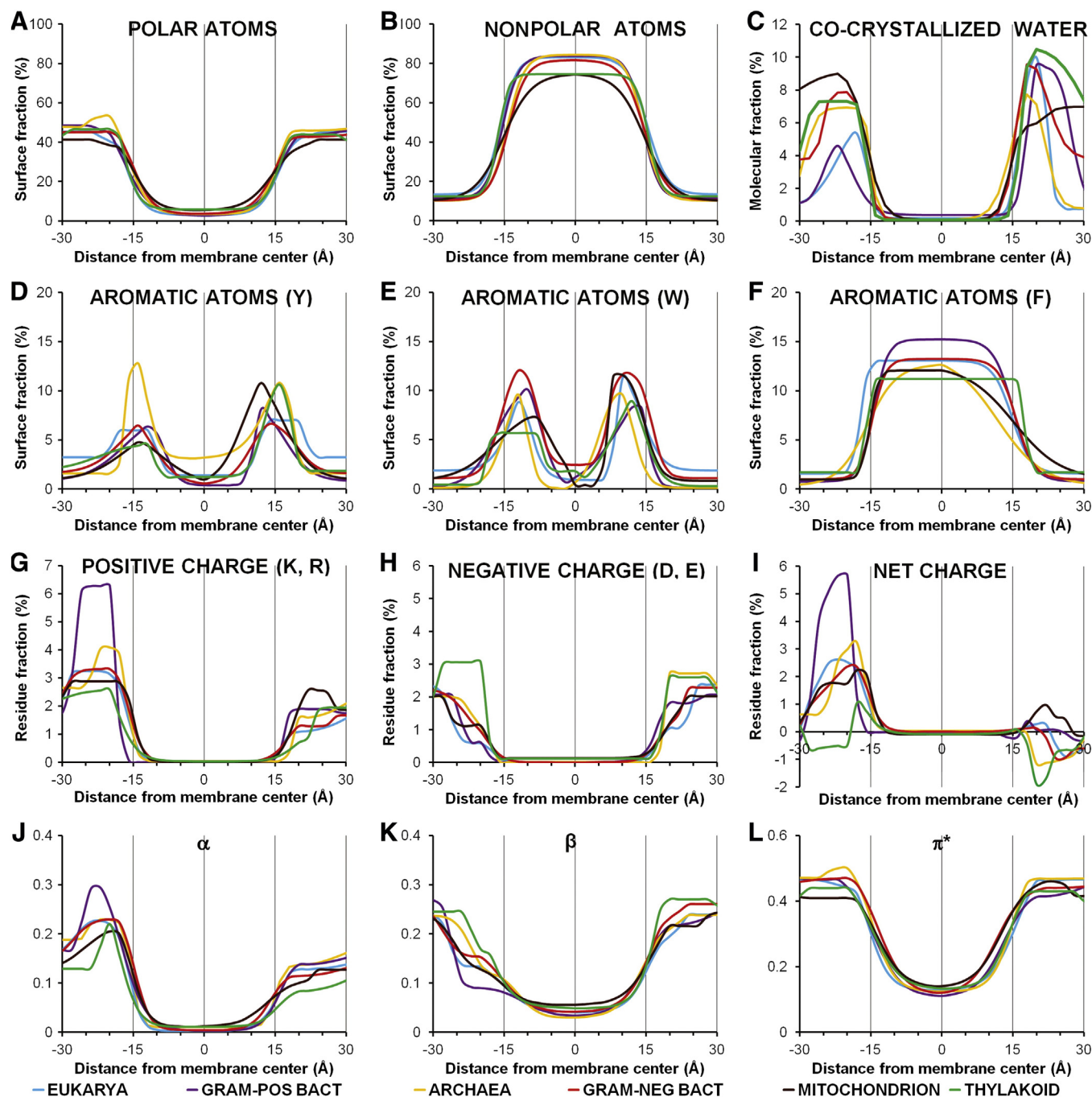
Fig. 3. (A–I) Distributions of lipid-facing protein atoms in structures of TM α-helical protein from six membrane types: eukaryotic PM (50 structures, blue line), PM of Gram-positive bacteria (12 proteins, purple line), PM of archaeabacteria (20 structures, yellow line), IM of Gram-negative bacteria (82 structures, red line), IM of mitochondria (9 structures, black line), thylakoid membrane (8 structures, green line). Polar atoms are N- and O-atoms; nonpolar atoms are C- and S-atoms from side chains of Val, Leu, Ile, Met, Cys, Phe, Tyr, and Trp residues. Aromatic atoms are C-atoms from benzene rings of Tyr and Phe and from indole ring of Trp. Charged groups are: amine group of Lys, guanidinium group of Arg, carboxyl group of Asp and Glu. (J–L) Transbilayer profiles of polarity parameters: hydrogen bonding donor (α) and acceptor (β) capacities and solvatochromic dipolarity/polarizability parameter (π*). Similarities in distributions and polarity profiles are observed inside the hydrophobic boundaries (± 15 Å from the membrane center), while most differences are seen outside these boundaries.

frequent at the hydrophobic membrane interior. Atypical distribution was also observed for His residues from thylakoid membrane proteins. In TM protein complexes from thylakoid membranes, such as photosystems I and II, light-harvesting and b6f complexes, His residues are present not only at the lipid head group and extramembrane regions, but also at the hydrocarbon region, where numerous histidines provide axial coordination of the heme cofactors at 5–10 Å-distances from both sides of the membrane center (not shown). Therefore, thylakoid membrane proteins were excluded from the protein set in Fig. S3.

The prominent and previously reported aspect of transbilayer distribution of ionizable residues of TM proteins is the higher occurrence of basic groups of Lys and Arg residues at the inner membrane side (i.e. cytoplasmic side, mitochondrial matrix side, thylakoid stroma side) than at the outer side (Figs. 3G, 7A–D). Thus, the distribution of positive charges follows the "positive inside" rule, an important factor that defines the TM topology of α-helical proteins [51–54].

Interestingly, all lipid-facing polar residues of TM α-helical proteins, except Arg and Lys, are more abundant at the outer membrane
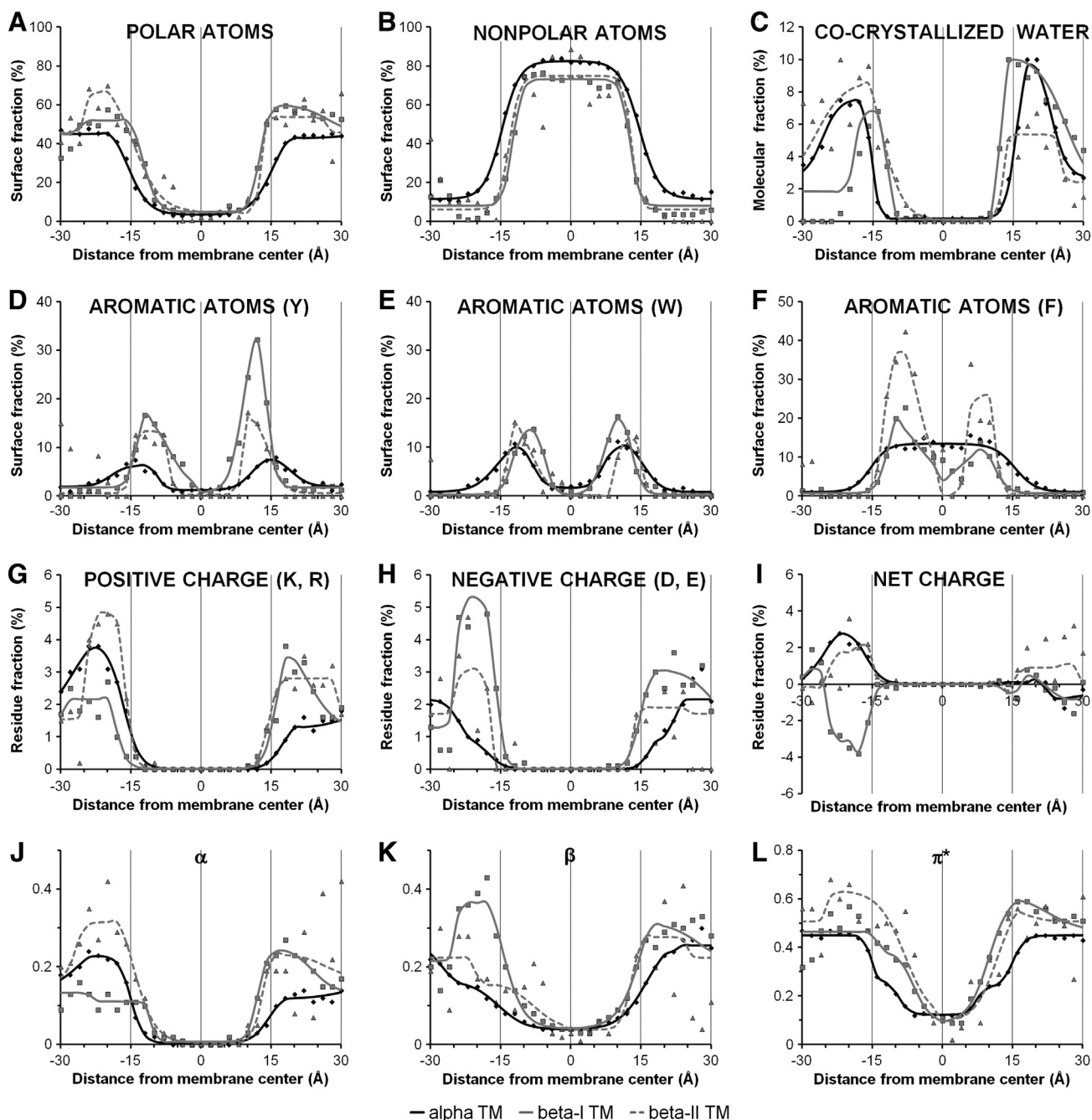
Fig. 4. (A–I) Comparison of distributions of lipid-facing protein atoms in structures of 191 TM α-helical protein from six membrane types (solid black line), 68 single-chain TM β-barrels (β-I type, solid gray line), and 5 multi-chain TM β-barrels (β-II type, dashed gray line). Distributions across the membrane were analyzed for polar atoms (N- and O-atoms of main and side chains), nonpolar atoms (C- and S-atoms from side chains of Val, Leu, Ile, Met, Cys, Phe, Tyr, Trp), aromatic atoms (C-atoms from benzene rings of Tyr, Phe and indole ring of Trp), and charged groups (amine group of Lys, guanidinium group of Arg, carboxyl group of Asp and Glu). (J–L) Comparison of transbilayer profiles of corresponding polarity parameters (α, β, π*). Differences in atom distributions and in polarity profiles for α-helical and β-barrel proteins are observed inside and outside the hydrophobic boundaries. These boundaries for β-barrels are shifted by 3–4 Å toward the membrane center as compared to α-helical proteins, indicating the smaller hydrophobic thickness of TM β-barrels. The distributions of charged residues and net charges are similar for α-helical and β-II proteins, but different for β-I proteins.

side (Fig. S3). This may compensate for the lower occurrence of Arg and Lys residues in this region. Indeed, transbilayer distributions of all polar atoms combined are almost symmetric for all membrane types studied (Fig. 3A).

In contrast, distributions of Arg and Lys charged groups of peripheral proteins are almost symmetric at both membrane sides (Fig. 7EF). Maxima of these distributions are located in both membrane leaflets at approximately 20–22 Å-distance from the bilayer center (Figs. 3G, 5, 7), which corresponds to locations of phosphodiester groups of bulk lipids (Fig. 2). Indeed, numerous ion pairs are observed between head groups

of co-crystallized lipids and Arg and Lys residues in the structures of membrane proteins (Fig. 6B). Hence, the formation of the corresponding protein-lipid hydrogen bonds and ionic bridges is an important factor not only for defining the protein topology, but also for binding and positioning of integral and peripheral proteins on both sides of membranes.

The third common feature of TM proteins is the presence of girdles of Tyr and Trp residues that enclose the hydrophobic zone. Consistent with previous studies, we can conclude that Tyr and Trp residues serve as membrane anchors that help to optimize spatial arrangements
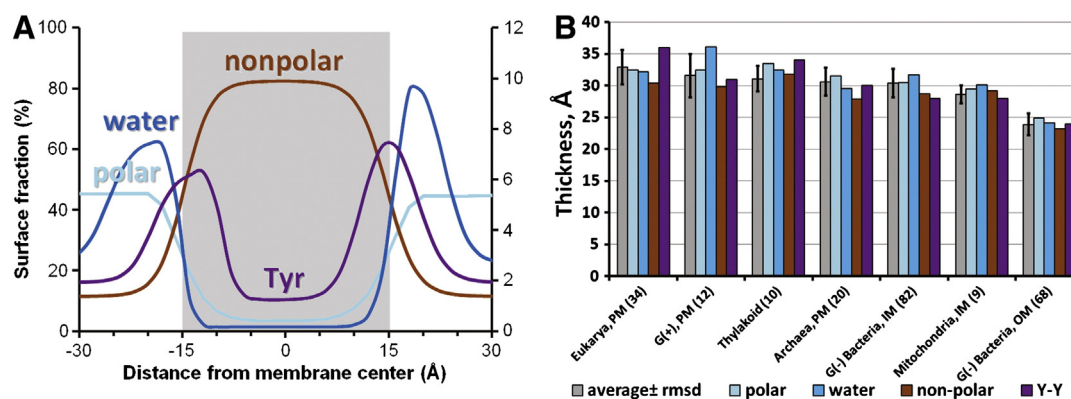
Fig. 5. Comparison of calculated hydrophobic thicknesses (gray) with midpoints of distributions of polar atoms (light blue), nonpolar atoms (brown), co-crystallized water (dark blue), and maxima of distribution of Tyr atoms (purple) in 191 selected TM α-helical proteins and protein complexes (A). Comparison of calculated hydrophobic thicknesses (gray) with distances between midpoints of distributions of polar atoms (light blue), nonpolar atoms (brown), and co-crystallized water (dark blue), and with distances between maxima of Tyr distributions (B). Analysis was performed for TM β-barrels from OM of Gram-negative bacteria and TM α-helical proteins from six membrane types: PM of eukaryotic cells, Gram-positive bacteria, archaeabacteria, thylakoid membranes, IM of mitochondria, and Gram-negative bacteria. Numbers of protein structures in each set are indicated in parenthesis. Calculated hydrophobic boundaries match the positions of midpoints of distribution curves of polar, nonpolar atoms and co-crystallized water, as well as maxima of Tyr distributions.

of proteins in membranes. The interfacial Tyr and Trp residues are present in both leaflets of nearly all TM and peripheral proteins, with very few exceptions. In particular, these residues are lacking in the inner, but not outer leaflet in the following TM proteins: (a) bacterial pilins; (b) viral proteins; (c) α-helical proteins from outer bacterial membranes; and (d) bacterial α-helical and β-barrel hemolysins (1WCD, 7AHL, 2B07, 3O44). All these exceptions are either multi-chain β-barrels or bitopic α-helical proteins. Perhaps the lack of anchoring residues at the intracellular side facilitates insertion of these proteins into membranes.

For most TM proteins distributions of aromatic rings of Tyr and Trp residues demonstrate two maxima (Figs. 3DE, 4DE, S4). Peaks of Tyr benzene rings match the positions of calculated membrane boundaries at 13 to 16 Å-distances from the membrane center. Maxima of Trp indole rings are located inside the hydrocarbon core at 10 to 12 Å-distances from the membrane center, which is deeper by approximately the indole ring size than the average location of lipid carbonyl groups and benzene rings of Tyr residues.

Although maxima of distributions of aromatic rings of Trp and Tyr residues do not overlap for either α-helical or β-barrel TM proteins from different membranes, peaks of their polar groups (NεH of Trp and OηH of Tyr) often coincide (Fig. S4). This may indicate that both

Trp NεH and Tyr OηH groups are preferentially oriented toward membrane boundaries (but from different sides), where they may form hydrogen bonds with lipid ester/ether lipid groups. The notable exceptions represent proteins from the outer leaflet of archaeabacterial PM and from both leaflets of ER and MIM, where peaks of Trp NεH groups are shifted by 4–6 Å closer to the membrane center than peaks of Tyr OηH groups.

Distributions of aromatic rings of Phe residues of most TM α-helical proteins differ from those of Tyr and Trp residues and resemble distributions of aliphatic residues: they have just one large flat maximum at the hydrocarbon core region (Fig. 3BF). However, for TM β-barrels (β-I and β-II types), distributions of Phe aromatic rings are quite similar to distributions of Trp indole rings: they have two maxima at approximately 7 Å-distance from each side of the membrane center (Fig. 4EF).

The last common feature of TM proteins is the behavior of polarity parameters (α, β, π*) that were calculated from the distribution of atoms on protein surfaces (Figs. 3JKL, 4JKL). The observed profiles of these polarity parameters appeared to be rather similar for proteins from different membrane types, especially within the hydrocarbon core. Therefore, all membrane proteins seem to be equally hydrophobic in the middle of their hydrophobic zone. This may explain the
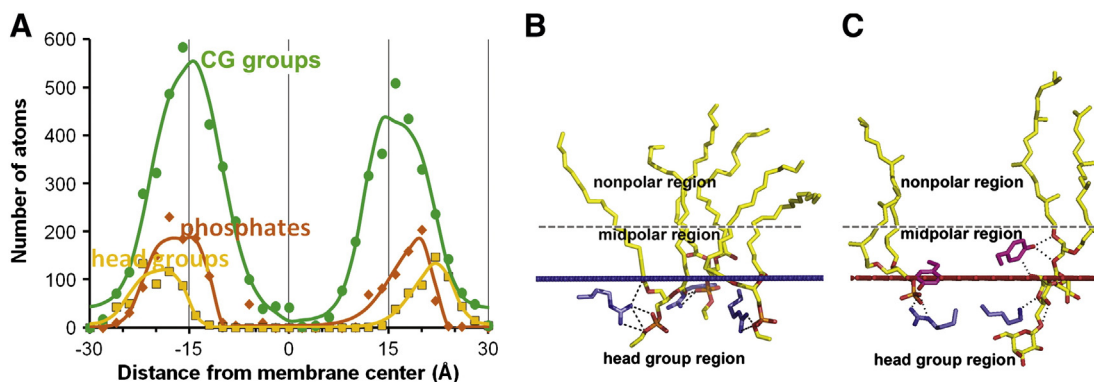

Fig. 6. Distribution of different groups of lipids co-crystallyzed with 164 TM α-helical proteins: ester and ether oxygens (green line and dots), phosphates (orange line and diamonds) and head group atoms (yellow line and squares) (A). Right pictures show position of 1,2-stearoyl-sn-glycero-3-phosphatidylethanolamine co-crystallized with bacterial cytochrome C oxidase (1M56) (B) and position of 2,3-Di-O-Phytanly-3-sn-Glycero-1-Phosphoryl-3′-sn-Glycerol-1′-Phosphate co-crystallized with bacteriorhodopsin (1IW6) (C) relative to the hydrophobic protein boundaries calculated by PPM (marked by blue and red lines). Lipid molecules are colored yellow (C atoms), red (O-atoms), and orange (P-atoms). Arg and Lys residues are colored blue (C-atoms), dark blue (N-atoms). Tyr residues are colored purple (C-atoms) and red (O-atoms). Hydrogen bonds between lipid head groups and Arg residues are indicated by black dashes.
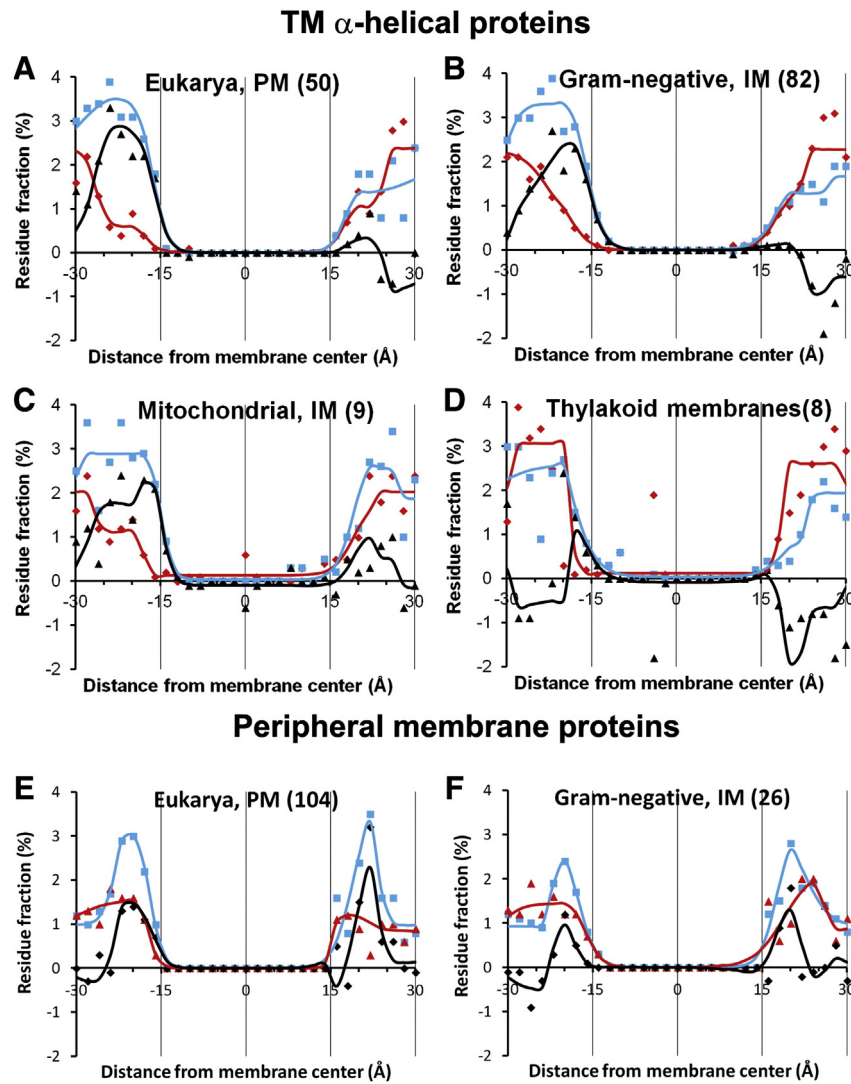
TM α-helical proteins

Fig. 7. Distributions of lipid-facing charged groups of Lys and Arg (blue lines and squares), Asp and Glu (red lines and triangles) and net charges (black lines and rhombs) in structures of TM α-helical proteins (A–D) and peripheral membrane proteins (E–F) from different membrane types: eukaryotic PM (A,E), IM of Gram-negative bacteria (B,F), IM of mitochondria (C), and thylakoid membranes (D). Numbers of protein structures in each set are indicated in parenthesis.

well-known tolerance of TM proteins to alteration of the lipid composition in native cells, in artificial bilayers, and during protein expression in different host organisms [55].

3.2.2. Differences in hydrophobic thickness

To compare properties of different membrane types, we analyzed the average intrinsic hydrophobic thicknesses (D_{aver}) of TM proteins from different membrane types and separately for α-helical or β-barrel proteins (Table 1, Fig. 8). We found that D_{aver} of TM α-helical proteins are rather similar and close to 30–31 Å for proteins from PMs of archaebacteria, Gram-positive and Gram-negative bacteria, the eukaryotic ER, and thylakoid membranes. The D_{aver} is slightly higher (33.5 ± 3.1 Å) for α-helical proteins from the eukaryotic PM and lower (28.6 ± 1.4 Å) for proteins from the mitochondrial IM (MIM). Largely increased hydrophobic thicknesses (up to 40 Å) were observed for integrins (2K1A, 2L8S, 2K9J) from eukaryotic PMs and V-type and F-type ATP-synthases from the PM of Gram-positive bacteria (2X2V), the IM of Gram-negative bacteria (2BL2, 1YCE), and the MIM (2XOK). The α-helical protein with the smallest hydrophobic thickness (22.6 Å) forms the 28-α-helical pore of T4S secretory system (3JQO) [56]. It is located in the OM of *E. coli*, which is largely enriched by β-barrel proteins.

Average hydrophobic thicknesses of typical TM β-barrels from bacterial and mitochondrial OMs are significantly smaller (23.9 ± 1.7 Å and 23.0 ± 0.4 Å, respectively) than the D_{aver} of most TM α-helical proteins. The OM proteins with the smallest hydrophobic thicknesses are: 24-stranded β-barrel of the usher protein FimD (3OHN, 21.2 Å) and 12-stranded β-barrel of the autotransporter-2 (2GR7, 20.4 Å).

Interestingly, an unusually large value of the hydrophobic thickness (40.7 Å) is observed for a multi-chain β-barrel of MspA porin from *Mycobacterium smegmatis* (1UUN) [57]. This value seems to be reasonable, as MspA resides in the OM of mycobacteria, which is rich in long-chain (from C_{30} to C_{90}) mycolic acids and was previously characterized by a large thickness, low fluidity, and low membrane permeability [58–62]. Our estimations of the hydrophobic thickness of the MspA porin are also supported by chemical labeling studies showing that the 40 Å-long β-barrel "steam domain" of MspA is protected from modification by a water-soluble label because it is surrounded by tightly bound lipids not removable by protein extraction [63].

We also analyzed variability of protein thicknesses in different membranes based on the corresponding minimal and maximal values and root-mean-square deviations (r.m.s.d.) (Table 1, Fig. 8). The largest variations are seen for eukaryotic PM proteins, whose hydrophobic thicknesses range from 27.2 to 40.8 Å and r.m.s.d. is 3.1 Å. Smaller
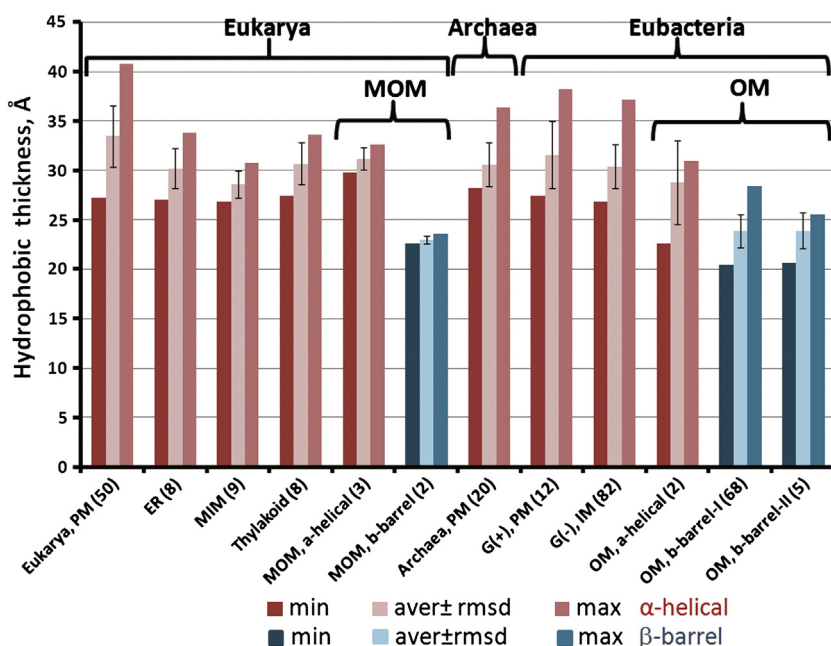
Fig. 8. Intrinsic hydrophobic thickness of membrane proteins from nine membrane types: PM of eukaryotic cell, Gram-positive (G(+)) bacteria, and archaeabacteria, endoplasmic reticulum (ER) membranes, thylakoid membranes, mitochondrial IM and OM (MIM, MOM) and Gram-negative bacteria (G(−)). Numbers of protein structures in each set are indicated in parenthesis.

dispersions in thicknesses are observed for α-helical proteins from MIM (r.m.s.d. is 1.4 Å) and β-barrels from bacterial OM (r.m.s.d. is 1.7 Å).

Separate analysis of single-spanning (bitopic) and multi-spanning (polytopic) proteins shows that there is no difference between these proteins in distributions of various atom types and polarity profiles in most membranes (Figs. S5, S6). However, polytopic and bitopic proteins from eukaryotic PM and MIM have slightly different hydrophobic thicknesses, judging from distances between midpoints in distributions of their aliphatic and polar atoms (Fig. S6, AG and BH). In particular, hydrophobic thickness seems to be larger for bitopic than polytopic proteins in PM, but smaller in MIM.

3.2.3. Differences in distributions of positive charges and net charge

The most significant differences between proteins from different membranes, and especially, between TM α-helical and β-barrel proteins, were observed in distributions of ionizable groups and hydrogen-bonding capacities (α and β) of protein polar atoms facing the head group region (Figs. 3G–K, 4G–K, 7A–D). As mentioned before, the distributions of basic Lys and Arg residues in TM α-helical proteins are highly asymmetric and follow the "positive inside" rule. However, this trend is observed only for TM α-helical proteins and multi-chain β-II barrels of

TolC-like proteins (Figs. 3G, 4G), but not for typical OM β-barrels (β-I type) (Fig. 4G) and peripheral proteins (Fig. 7E–F). On the contrary, the distribution of basic residues of OM β-barrels demonstrates a higher peak at the outer leaflet of the OM ("positive outside" rule). The distributions of basic residues in peripheral proteins are more symmetric.

Interestingly, TM proteins from thylakoid membranes, which are rich in non-phosphorous glycolipids [64], demonstrate a unique pattern of charge distribution (Fig. 7D). Similar to other α-helical TM proteins, they follow the "positive inside" rule having an excess of basic residues (i.e. Arg, Lys) in the inner (stroma) membrane side. However, positive charged groups of Arg and Lys residues are almost counterbalanced by negatively charged groups of Asp and Glu residues. Hence, TM proteins in thylakoid membranes do not have distinct maxima for the net positive charge at both membrane sides.

3.2.4. Midpolar regions in different biological membranes

The comparison of transbilayer profiles of polarity parameters (α, β, π*) calculated for artificial lipid bilayers and lipid-facing atoms of membrane proteins (Fig. 9) demonstrates the existence of two interfacial regions inside the hydrocarbon core where all polarity parameters sharply change, so-called "midpolar regions". In particular, for either
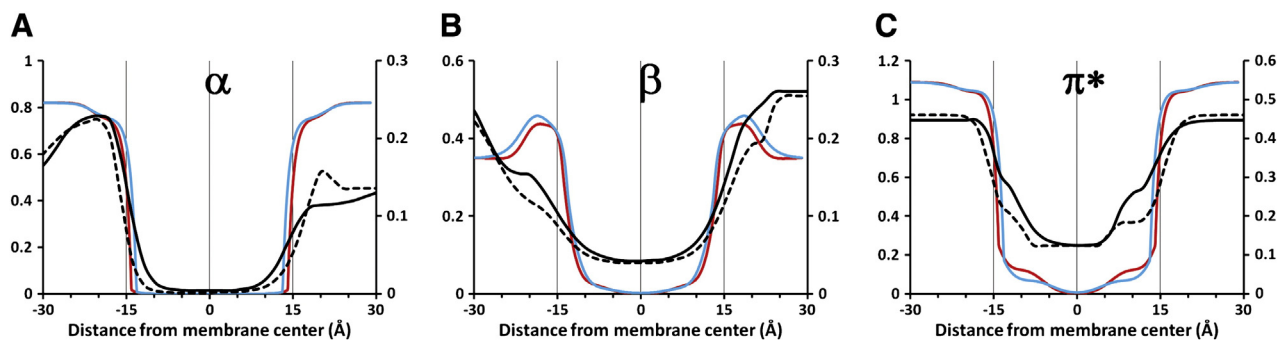

Fig. 9. Comparison of transbilayer profiles of polarity parameters (α, β, and π*) calculated for artificial lipid bilayer, DOPC (red lines), POPG (blue lines), and for lipid-facing atoms of TM α-helical proteins from IM of Gram-negative bacteria (82 structures, black solid lines) and from PM of eukaryotic cells (34 structures, black dashed lines). In each panel vertical axis corresponds to a polarity parameter calculated for artificial lipid bilayers (left axis) and TM proteins (right axis).
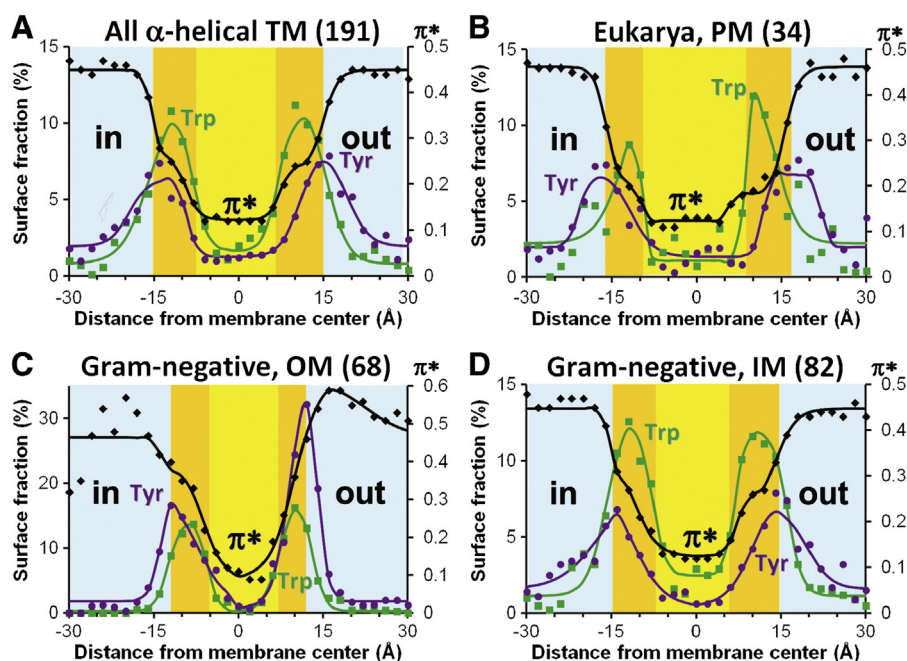
Fig. 10. Localization of midpolar regions based on distributions of lipid-facing atoms of TM α-helical proteins. (A) Analysis of distributions of polar atoms, N and O (blue line), Tyr atoms (purple line), Trp atoms (green line), and polarity parameter π* calculated for 191 structures of TM α-helical proteins. Midpolar regions are colored orange, head group regions are colored light blue, central nonpolar regions are colored yellow. (B) Comparison of polarity profile of parameter π* calculated for 34 structures of TM α-helical proteins from eukaryotic PM (blue line and squares) and 82 structures of TM α-helical proteins from IM of Gram-negative bacteria.

model bilayers or TM proteins the hydrogen-bond acceptor parameter β substantially drops at 10 to 15 Å-distance from the membrane center, while the value of dipolarity parameter π* significantly changes at 7 to 15 Å-distance from the membrane center. In artificial bilayers the midpolar region originates from the presence of double bonds in lipid acyl chains (Fig. 2). In proteins, this region is characterized by the presence of aromatic groups; especially Trp indole rings (Fig. 10).

We also observed that all polarity parameters of membrane proteins change more gradually than those calculated for artificial lipid bialyers (Fig. 9). This result probably reflects the significantly more heterogeneous lipid and protein composition of biological membranes. Moreover, the profiles of α and β parameters in proteins are highly asymmetric, whereas in model bilayers they are symmetric. The asymmetry of α and β is likely attributed to the preferred accumulation of basic protein residues with high hydrogen-bonding donor capacity (α) at the inner leaflet and of anionic residues with high hydrogen-bonding acceptor capacity (β) at the outer leaflet.

To better understand the nature of the midpolar region in biomembranes and its correlation with divergence in lipid composition of these membranes, we compared distributions of Trp atoms and profiles of the polarity parameter π* for protein sets from different membrane types (Fig. 10). Considering the midpolar region as a preferential location for Trp indole rings, we assume that it extends from the inner midpoint of Trp distribution to the boundary of the hydrophobic core.

Our study shows that midpolar regions are frequently asymmetric and their sizes differ for proteins from different membranes. The highest asymmetry is observed for bacterial OM (Fig. 10C). This correlates with asymmetric lipid composition of OM where phospholipids are present in the inner leaflet, while lipid A of lipopolysaccharide (LPS) forms the outer leaflet [65]. The central hydrophobicity barrier for OM is significantly more narrow, though the value of π* in the middle is almost the same in all membranes.

The smaller but noticeable asymmetry is also seen for midpolar regions of eukaryotic PM proteins, which mainly originate from mammalian cells (Fig. 10B). These membranes have uneven distributions of different lipid species at both membrane leaflets: PS and PE are accumulated in the inner leaflet at the cytoplasmic side, PC and SM in the outer leaflet [13], while cholesterol may be distributed more uniformly with slight preference for the inner leaflet [66].

In contrast, proteins from IM of Gram-negative bacteria demonstrate a more symmetric profile of polarity parameter π* (Fig. 10D). This correlates with smaller lipid diversity of bacterial IM as compared to eukaryotic PM [9]. Finally, the asymmetry in the fine structure of the hydrocarbon region almost disappears when 191 α-helical proteins from all membrane types are considered simultaneously (Fig. 10A) because the opposite trends in different membranes tend to cancel each other.

To further understand the influence of the lipid composition on the polarity of the lipid bilayer, we compared polarities of proteins from membranes of mesophilic and hyperthermophilic archaebacteria (Fig. 11). The latter thrive under harsh environmental conditions (temperature maximum 121° C, high pressure >120 MPs). We found that α-helical proteins from PM of mesophilic and hyperthermophilic archaebacteria have very different distributions of Trp residues and polarity parameter π*. In particular, in hyperthermophilic archaebacteria, the central nonpolar region, which likely serves as the permeability barrier for ions and polar molecules, is more symmetric, wider and has a higher hydrophobicity (π*$_{aver}$ = 0.10). In contrast, midpolar regions of mesophilic archaebacteria are highly asymmetric and less hydrophobic in the central nonpolar region (π*$_{aver}$ = 0.13). These polarity profiles correlate with properties of the corresponding membranes. Indeed, in thermophilic Archaea, membranes are composed of symmetric C_{40} cyclic tetraether lipids that are more densely packed, more rigid and stable, consistent with the larger thickness and higher hydrophobicity of the main permeability barrier [67,68]. On the other hand, lipids of mesophilic Archaea are based on C_{20}–C_{25} di-phytanyl-sn-glycerol and are rather variable: have numerous head groups, acyl chain may have double bonds and hexane rings that disturb lipid packing [68–70]. Thus, these membranes are more loosely packed and known to have a higher permeability for water than bipolar tetraether lipids.
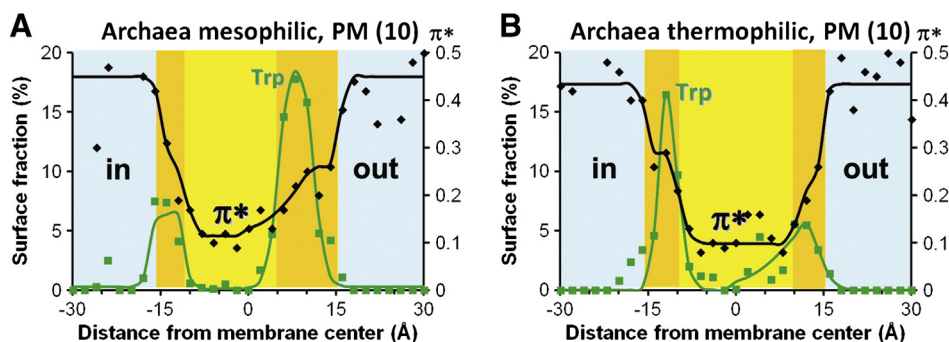
Fig. 11. Comparison of the fine structure of the hydrocarbon core region in mesophilic and thermophilic archaeabacteria. Localization of the midpolar regions is based on the distributions of lipid-facing Trp atoms and polarity parameter π* calculated for structures of TM α-helical proteins from mesophilic (A) and thermophilic (B) archaeabacteria. Numbers of protein structures in each set are indicated in parenthesis. Midpolar regions are colored orange, head group regions are colored light blue, central nonpolar regions are colored yellow.

4. Discussion

4.1. General adaptation of proteins to the lipid bilayer

Co-evolution of membrane proteins and lipids in diverse types of biological membranes is expected to develop a number of structural features of membrane proteins that provide their general adaptation to the anisotropic membrane environment, as well as their specific adaptation to a particular membrane type. Many of these features have been previously discussed [10,11,71–74]. Here, for the first time, we provide systematic analysis and comparison of hydrophobic thicknesses, surface charge distribution, and transbilayer polarity profiles of membrane proteins from different structural classes (α-helical and β-barrel) and eight types of biological membranes (IM and OM of Gram-negative bacteria, PMs or archaeabacteria, Gram-positive, and eukaryotes, ER membrane, MIM, and tylakoid membranes). These membranes have a complex lipid and protein composition that is not fully established and, therefore, cannot be easily simulated in vitro or in silico.

Several common features of all TM proteins are essential for their solubility, structural stability, proper topology, and orientation in the lipid bialyer. They include the presence of extensive hydrophobic surfaces, the asymmetry of distributions of polar and charged residues outside the hydrophobic zone, and the interfacial locations of Trp and Tyr residues.

Any integral membrane protein has a large solvent-facing hydrophobic zone that defines its intrinsic hydrophobic thickness (Fig. 1). This thickness and location of hydrophobic boundaries can be calculated by minimizing transfer energy of a protein from water to the lipid bilayer, as performed by our PPM 2.0 method. The current study demonstrates that the calculated boundary planes correspond to midpoints of sharp sigmoidal curves describing distributions of protein polar and nonpolar atoms and co-crystallized water (Fig. 4). The hydrophobic boundaries also correspond to maxima of distributions of Tyr rings and glycerol groups of co-crystallized lipids (Figs. 5, 6).

The hydrophobic surfaces of membrane proteins are enclosed by girdles of residues Tyr and Trp and regions rich in ionizable residues and co-crystallized water (Fig. 1). This apparently reflects favorable protein–lipid interactions at membrane interfaces [10,11,71–74]. The polar group of both Trp and Tyr residues often point to the hydrophobic boundaries (Fig. S4) where Tyr O$^\eta$H groups may form hydrogen bonds with glycerol groups of co-crystallized lipids (Fig. 6C). However, maxima of distributions of aromatic rings of Tyr and Trp residues do not coincide (Figs. 4S, 10). Indole rings of Trp accumulate at 10–12 Å-distance from the membrane center, which is 3–5 Å closer to the membrane center than maxima of Tyr ring distributions.

In both integral and peripheral membrane proteins, regions rich in ionizable residues correspond to the lipid head group area, since the peaks of positively charged groups of Arg and Lys residues coincide with locations of lipid phosphate groups at approximately 20–22 Å from the membrane center (Figs. 5, 7). Indeed, multiple hydrogen bonds and ionic bridges between basic protein residues and lipid phosphates can be found in crystal structures of membrane proteins (Fig. 6B). Interestingly, transbilayer distribution of polar and ionizable residues is highly asymmetric (Fig. S4), which translates in the asymmetry of hydrogen-bonding acceptor and donor capacity of protein residues from both membrane sides (Figs. 3JK, 4JK).

Importantly, the described common features are generally valid for both α-helical and β-barrel membrane proteins (Figs. 3, 4). The major distinction of TM β-barrels is in the two-peak distribution of Phe residues (see above) and in the enrichment of their hydrophobic surface by aromatic (Tyr, Trp, Phe) residues, whose relative occurrence in β-barrels is almost doubled as compared to α-helical proteins (Fig. 7F). This may be related to high β-sheet propensity and low α-helix propensity of aromatic residues. Indeed, substitution of Ala by Tyr, Phe, and Trp stabilizes the β-sheet structure by 0.96, 0.86 and 0.54 kcal/mol, respectively [75], while destabilizing α-helix by 0.3–0.6 kcal/mol [76].

4.2. Adaptations of proteins to specific membranes

In addition to the mentioned general features, there are certain differences in protein structures that may reflect their adaptation to specific membrane types. In the course of our study we found that proteins from distinct biological membranes differ in the average values of their intrinsic hydrophobic thicknesses, in distributions of ionizable and aromatic groups, and in the structure of midpolar regions within the hydrocarbon region.

4.2.1. Intrinsic hydrophobic thicknesses of membrane proteins

There are significant variations in hydrophobic thicknesses of proteins from different membranes (Table 1, Fig. 8). The observed differences in average thicknesses of eukaryotic proteins (PM > ER > MIM proteins) correlate with X-ray scattering data indicating that the bilayer thickness of the apical PM is approximately 5 Å larger than that of ER membranes, and can be modulated primarily by membrane proteins [77]. These differences may be important to facilitate sorting of proteins between the ER and PM or between ER and MIM. Indeed, studies of polarized epithelial cells provided evidence of the lipid-raft-based sorting in trans-Golgi network as an important mechanism for delivery to the cell surface of TM proteins with longer TM α-helices without involvement of coat or adaptor proteins [78,79]. It was also shown that unassisted post-translational targeting of tail-anchored proteins to mitochondrial OM (MOM) requires shorter and less hydrophobic helices with positive flanking changes, while ER-targeting is less restrictive [80,81]. We also found that the difference in hydrophobic thicknesses

between bitopic proteins from MIM and eukaryotic PM membranes is greater than between polytopic proteins from the same membranes (Fig. S6). This may also be important for sorting of single-spanning proteins between ER, PM and mitochondrial membranes.

We observed that TM α-helical proteins typically have larger hydrophobic thickness than TM β-barrels (Table 1, Fig. 8). This result is consistent with experimental observations [82–86], including NMR studies of detergent-embedded residues of OmpX [82], analysis of the thickness of the detergent belt in crystals of OmpLA [83], and neutron scattering studies using contrast variation of LPS bilayers [84,85].

We noticed a significant variability in hydrophobic thicknesses within the set of proteins from eukaryotic PMs. This may be related to a heterogeneous lipid composition of cell membranes in different tissues. Another possible reason might be the lateral heterogeneity of the PM due to the presence of lipid domains rich in cholesterol and sphingomyelin with locally increased width, so-called lipid rafts [87,88]. Indeed, the calculated hydrophobic thicknesses of several receptors, such as integrins (2K1A, 2K9J, 2L8S) or receptor tyrosine kinases (2L6W) that may by associated with lipid rafts, are usually bigger, ranging from approximately 34 to 40 Å.

On the other hand, the significantly smaller variations in hydrophobic thicknesses within the sets of proteins from bacterial OM and MIM may be attributed to a more stable lipid composition and properties of these membranes. Indeed, in contrast to the variable lipid composition of eukaryotic PMs, the level of mitochondrial phospholipids is rather similar in different tissues [89]. Particularly important is the proper amount of cardiolipin, an anionic tetra-acyl phospholipid that stabilizes the structure of protein complexes involved in cell bioenergetics: its alteration causes disease states, such as Barth syndrome [90]. The asymmetric bacterial OM also has a rather constant structure, as its outer leaflet is primarily composed of lipid A, a part of LPS, while its inner leaflet contains only a few types of glycerophospholipids (primarily PE, PG, and CL). The LPS-rich leaflet has a highly ordered gel-like arrangement of saturated acyl chains [91], which is additionally stabilized by hydrogen bonds and ionic interactions in the lipid head group area [65].

In contrast to the relatively small variations in thicknesses of OM β-barrels, the observed difference between minimal and maximal thicknesses of α-helical proteins from bacterial OMs is substantial (8.4 Å). In particular, one α-helical 14-meric protein, the VirB7/VirB9/VirB10 core complex of T4S secretion system (3JQO) [56], has thickness of 22.6 Å, which is close to D_{aver} ~24 Å of highly abundant OM β-barrels. Another α-helical OM protein, translocon of capsular polysaccharides Wza (2J58) [92], has a relatively large hydrophobic thickness of 31 Å. The unusually large hydrophobic thicknesses of Wza may indicate potential conformational transitions of Wza between states with small and large hydrophobic thicknesses that could change the diameter of the central pore of its octameric α-helical barrel.

A similar example represents proteins from MOM (Fig. 8), where three α-helical proteins with known 3D structures have average hydrophobic thickness that is approximately ~8 Å larger that the thickness of the most abundant MOM β-barrel protein, the voltage-dependent anion channel VDAC-1 (2JK4, 3EMN) [93,94].

Divergence was also found between hydrophobic thicknesses of mitochondrial F_1F_0-ATP synthase (38.2 Å) and numerous proteins forming large respiratory complexes (D_{aver} ~29 Å) which likely define the overall thickness of the MIM near this value. Thus, the hydrophobic mismatch exists between hydrophobic thickness of F_1F_0-ATP synthase and the surrounding lipid bilayer. This may enhance the known tendency of the F_1F_0-ATP synthase to form dimers, tetramers, and even regular arrays of dimers in membranes [95]. Indeed, it has been observed by EM microscopy [95] that ribbon-like complexes of mitochondrial F_1F_0-ATP synthase are stacked in parallel along the crystae. It was suggested that these quaternary structures may participate in membrane morphogenesis by changing membrane curvature and promoting formation of tubular structures [96].

Judging from the comparison of average values of intrinsic hydrophobic thicknesses of TM proteins from different membranes, we can conclude that this parameter may serve as an appropriate characteristic of a particular membrane type, though it also depends on the protein structural type (α-helical or β-barrel). The high content of proteins with defined hydrophobic thickness, such as the presence of thicker α-helical proteins or thinner β-barrels, may be regarded as a major factor that defines or modulates the average thickness of a particular biological membrane, as was previously suggested [77]. However, our observations also indicate that natural membranes may accommodate proteins with rather variable thicknesses. Hence, a significant hydrophobic mismatch can be tolerated by adjusting the local thickness of the fluid lipid bilayer to the geometry of residing proteins. Such mismatches may be structurally and functionally important for example by enhancing oligomerization or structural transitions in membrane proteins.

4.2.2. Charge distributions on membrane protein surfaces

There are significant differences in distributions of ionizable groups between membrane proteins of different structural types, e.g. TM α-helical and β-barrel proteins of β-I or β-II types (Fig. 4G–I), TM and peripheral proteins (Fig. 7), as well as between TM α-helical proteins from different membranes (Figs. 3G–I, 7A–D).

The well-recognized "positive-inside" rule [51–54] was observed only for TM α-helical proteins from seven membranes studied, but not for OM β-barrels (β-I type) or for peripheral proteins (Figs. 4G, 7EF). The distributions of positively charged residues in peripheral proteins are more symmetric. This may indicate that peripheral proteins interact similarly with lipid phosphates at both membrane sides.

Unlike TM α-helical proteins, OM β-I barrel proteins have a higher peak of positively charged residues at the outer membrane side, where they may interact with negatively charged LPS, as it is observed in crystal structure of FhuA receptor (1QFG). This "positive-outside" trend for bacterial OM β-barrels has been previously reported [97]. Further, β-I barrels display a net negative charge at the periplasmic side ("negative-inside" rule) due to the presence of acidic residues in periplasmic turns. This acidic residues may form ionic interactions with cationic periplasmic Skp chaperone that facilitates proper protein insertion and folding into the OM [98].

All these results clearly indicate that the asymmetric distributions of positively charged residues in TM α-helical and β-barrel proteins reflect the topological bias important for membrane protein biogenesis rather than asymmetric lipid composition in these membranes. However, some influence of lipid composition on protein charge distribution across the membrane cannot be ruled out. The role of ionic interactions between lipid phosphates and Lys/Arg residues may be more or less pronounced, depending on the level of phospholipids in membranes.

In particular, thylakoid proteins demonstrate the unusual pattern of transbilayer charge distributions lacking peak of net positive charge at the inner membrane side (Fig. 7D). This may be explained by the unique lipid composition of thylakoid membranes, where phospholipids are mainly substituted by non-phosphorous glycolipids, monogalactosyl diacylglycerol (MGDG) and digalactosyl diacylglycerol, and to a lesser extent (<15%) by sulfolipid, sulfoquinovosyl diacylglycerol [64]. Therefore, the amount of lipid phosphate groups and its role in lipid–protein interactions is greatly reduced in these membranes.

4.2.3. Multilayered organization of hydrophobic region

The analysis of polarity profiles of artificial lipid bilayers and membrane proteins indicates that the lipid hydrocarbon core is not a uniform environment, but includes two peripheral zones of 5–8 Å width each, so-called midpolar regions, which are characterized by intermediate values of dipolarity/polarizability parameter π* (Fig. 9).

The existence of midpolar regions within the lipid hydrocarbon core is consistent with previous studies. The term "midpolar region"

has been proposed to emphasize an intermediate value of the dielectric constant, which ensures the stronger ionic interactions and hydrogen-bonds in this region [99]. It was also shown that the charge density profile changes gradually at the membrane interface that includes both head group and midpolar regions [6].

It has been previously observed that the lipid acyl chain region consists of ordered and disordered domains separated by lipid double bonds [100]. The midpolar regions likely correspond to the more ordered "soft polymer" domains. These domains have a larger number of small structural defects that allow an easier penetration of water [101,102]. Indeed, studies of spin-labeled lipids indicate the presence of the appreciable amount of water in peripheral regions of the lipid hydrocarbon core close to the glycerol backbone [37,48,103,104]. The locations and properties of more polar peripheral regions in the hydrocarbon core of the lipid bilayer were shown to depend on the presence of lipid double bonds, cholesterol and carotenoids [48,104–106]. The amphipathic indole ring was shown to accumulate in the glycerol region, penetrating up to the level of lipid C1–C3 atoms [107].

In artificial bilayers, midpolar regions arise due to the presence of double bonds in lipid acyl chains (Fig. 2) and some amount of residual water molecules [30]. In natural membranes, midpolar regions can be identified based on the increased concentration of indole rings of Trp residues of membrane proteins that penetrate deeper into the hydrocarbon region than aromatic rings of Tyr residues (Fig. 10). Trp and Tyr residues behave differently because the $O^\eta H$ group of Tyr has higher H-bonding donor and acceptor capacities than $N^\varepsilon H$ of indole ring and, therefore, is more involved in hydrogen-bonding with water in the head group region, while the Trp indole ring has a two times larger dipole moment and, therefore, is more sensitive to electrostatic interactions in the hydrocarbon region.

Trp indole ring can serve as a "hydrophobic dipole" reporter: the localization of the indole ring inside the hydrocarbon core indicates a relatively small electrostatic penalty for the dipole due to an intermediate value of the dielectric parameter (π^* or ε) and preferential solvation of polar groups by small amount of water present in this region [30]. Hence the borders of midpolar regions located on both sides of the central nonpolar region can be defined using either inner midpoints of distributions of Trp dipoles or inflection points of the π^* profiles (Figs. 10, 11).

We found that midpolar regions may have different sizes and be asymmetric, depending on lipid composition of membrane leaflets. High asymmetry of midpolar regions in OM β-barrels likely reflects the apparent asymmetry of the OM (Fig. 10C). Larger thickness of the central nonpolar region of TM proteins from hyperthermophilic archaeabacteria correlates with decreased penetrability of corresponding membranes formed by the monolayer of C_{40} cyclic tetraether lipids (Fig. 11B) [68], although packing density and viscosity of lipid bilayers in archaebacteria also must play a role.

Thus, using commonly used polarity parameters, we were able to quantify transbilayer polarity of membrane proteins, establish the multilayer organization of membranes and evaluated size and asymmetry of midpolar regions in different biological membrane. Assuming an approximate matching of polarity of integral membrane proteins and surrounding lipidic phase, the profiles of parameter π^* calculated for the surfaces of membrane proteins can be used to derive dielectric properties of the corresponding native membranes.

5. Conclusions

This work represents the first comparative study of hydrophobic thicknesses and polarity profiles of proteins from eight types of biological membranes. We observed that average hydrophobic thicknesses of TM α-helical proteins from different membranes are 4 to 9 Å larger than that of β-barrel proteins. Calculated hydrophobic boundaries correspond to sharp polarity transitions on the protein surface and match the locations of Tyr aromatic rings and ester/ether

groups of co-crystallized lipids, but not maxima of Trp indole rings, which are shifted by 3–4 Å inside the hydrophobic region. The positively charged protein residues are preferentially located in the area of phosphate groups of co-crystallized lipids, which indicates the importance of ionic protein–lipid interactions.

Polarity of surfaces of membrane proteins was characterized by transbilayer profiles of hydrogen bonding donor and acceptor capacities (α, β) and dipolarity/polarizability (π^*), parameters commonly used to describe solubility of molecules in organic solvents. Behavior of polarity parameters within hydrophobic membrane boundaries indicates the multilayered organization of the hydrocarbon core that consists of a central nonpolar region and two 5–8 Å-wide peripheral midpolar regions with intermediate values of parameter π^*. In different biological membranes, size and location of midpolar regions were derived from distributions of indole rings of Trp that served as a convenient marker of these regions. In artificial lipid bilayers, boundaries of midpolar regions were defined by locations of double bonds of lipid acyl chains. The observed asymmetry of midpolar regions in proteins from different membranes correlates with known lipid asymmetry of corresponding membranes.

Hydrophobic thicknesses and profiles of polarity parameter π^* obtained in the current study for proteins from different membranes can be used for development of more advanced methods for computational studies of membrane proteins and amphiphilic molecules. Computational studies may include modeling of membrane protein structure, spatial localization of proteins and organic molecules in the lipid bilayer, and predicting permeability of peptides and small drug-like molecules, specifically for these eight types of biological membranes.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.bbamem.2013.06.023.

References

[1] H.E. Findlay, P.J. Booth, The biological significance of lipid–protein interactions, J. Phys. Condens. Matter 18 (2006) S1281–S1291.
[2] R.M. Epand, Lipid polymorphism and protein–lipid interactions, Biochim. Biophys. Acta Biomembr. 1376 (1998) 353–368.
[3] R. Phillips, T. Ursell, P. Wiggins, P. Sens, Emerging roles for lipids in shaping membrane–protein function, Nature 459 (2009) 379–385.
[4] J.L. MacCallum, D.P. Tieleman, Interactions between small molecules and lipid bilayers, Curr. Top. Membr. 60 (2008) 227–256.
[5] S. Fiedler, J. Broecker, S. Keller, Protein folding in membranes, Cell. Mol. Life Sci. 67 (2010) 1779–1798.
[6] S.H. White, A.S. Ladokhin, S. Jayasinghe, K. Hristova, How membranes shape protein structure, J. Biol. Chem. 276 (2001) 32395–32398.
[7] A.L. Lomize, M.Y. Reibarkh, I.D. Pogozheva, Interatomic potentials and solvation parameters from protein engineering data for buried residues, Protein Sci. 11 (2002) 1984–2000.
[8] A.L. Lomize, I.D. Pogozheva, H.I. Mosberg, Quantification of helix–helix binding affinities in micelles and lipid bilayers, Protein Sci. 13 (2004) 2600–2612.
[9] T.J. Denich, L.A. Beaudette, H. Lee, J.T. Trevors, Effect of selected environmental and physico-chemical factors on bacterial cytoplasmic membranes, J. Microbiol. Methods 52 (2003) 149–182.
[10] A.G. Lee, Lipid–protein interactions in biological membranes: a structural perspective, Biochim. Biophys. Acta Biomembr. 1612 (2003) 1–40.
[11] H. Palsdottir, C. Hunte, Lipids in membrane protein structures, Biochim. Biophys. Acta Biomembr. 1666 (2004) 2–18.
[12] W. Dowhan, M. Bogdanov, Lipid-dependent membrane protein topogenesis, Annu. Rev. Biochem. 78 (2009) 515–540.

[13] G. van Meer, D.R. Voelker, G.W. Feigenson, Membrane lipids: where they are and how they behave, Nat. Rev. Mol. Cell Biol. 9 (2008) 112–124.

[14] M.S. Almén, K.J.V. Nordström, R. Fredriksson, H.B. Schiöth, Mapping the human membrane proteome: a majority of the human membrane proteins can be classified according to function and evolutionary origin, BMC Biol. 7 (2009).

[15] Y.F. Zhai, M.H. Saier, The beta-barrel finder (BBF) program, allowing identification of outer membrane beta-barrel proteins encoded within prokaryotic genomes, Protein Sci. 11 (2002) 2196–2207.

[16] W.C. Wimley, Toward genomic identification of beta-barrel membrane proteins: composition and architecture of known structures, Protein Sci. 11 (2002) 301–312.

[17] I. Iacovache, M.T. Degiacomi, F.G. van der Goot, Pore-forming toxins, Compr. Biophys. 5 (2012) 164–188.

[18] B.A. Wallace, Gramicidin channels and pores, Annu. Rev. Biophys. Biophys. Chem. 19 (1990) 127–157.

[19] R.M. Bill, P.J.F. Henderson, S. Iwata, E.R.S. Kunji, H. Michel, R. Neutze, S. Newstead, B. Poolman, C.G. Tate, H. Vogel, Overcoming barriers to membrane protein structure determination, Nat. Biotechnol. 29 (2011) 335–340.

[20] H.M. Berman, T. Battistuz, T.N. Bhat, W.F. Bluhm, P.E. Bourne, K. Burkhardt, L. Iype, S. Jain, P. Fagan, J. Marvin, D. Padilla, V. Ravichandran, B. Schneider, N. Thanki, H. Weissig, J.D. Westbrook, C. Zardecki, The Protein Data Bank, Acta Crystallogr. Sect. D: Biol. Crystallogr. 58 (2002) 899–907.

[21] M.A. Lomize, I.D. Pogozheva, H. Joo, H.I. Mosberg, A.L. Lomize, OPM database and PPM web server: resources for positioning of proteins in membranes, Nucleic Acids Res. 40 (2012) D370–D376.

[22] M.B. Ulmschneider, M.S.P. Sansom, Amino acid distributions in integral membrane protein structures, Biochim. Biophys. Acta Biomembr. 1512 (2001) 1–14.

[23] A. Senes, D.C. Chadi, P.B. Law, R.F.S. Walters, V. Nanda, W.F. DeGrado, E-z, a depth-dependent potential for assessing the energies of insertion of amino acid side-chains into membranes: derivation and applications to determining the orientation of transmembrane and interfacial helices, J. Mol. Biol. 366 (2007) 436–448.

[24] M.B. Ulmschneider, M.S.P. Sansom, A. Di Nola, Properties of integral membrane protein structures: derivation of an implicit membrane potential, Proteins Struct. Funct. Bioinf. 59 (2005) 252–265.

[25] D. Hsieh, A. Davis, V. Nanda, A knowledge-based potential highlights unique features of membrane α-helical and β-barrel protein insertion and folding, Protein Sci. 21 (2012) 50–62.

[26] L. Adamian, V. Nanda, W.F. DeGrado, J. Liang, Empirical lipid propensities of amino acid residues in multispan alpha helical membrane proteins, Proteins Struct. Funct. Bioinf. 59 (2005) 496–509.

[27] T.A. Eyre, L. Partridge, J.M. Thornton, Computational analysis of alpha-helical membrane protein structure: implications for the prediction of 3D structural models, Protein Eng. Des. Sel. 17 (2004) 613–624.

[28] C.A. Schramm, B.T. Hannigan, J.E. Donald, C. Keasar, J.G. Saven, W.F. Degrado, I. Samish, Knowledge-based potential for positioning membrane-associated structures and assessing residue-specific energetic contributions, Structure 20 (2012) 924–935.

[29] A.L. Lomize, I.D. Pogozheva, M.A. Lomize, H.I. Mosberg, Positioning of proteins in membranes: a computational approach, Protein Sci. 15 (2006) 1318–1333.

[30] A.L. Lomize, I.D. Pogozheva, H.I. Mosberg, Anisotropic solvent model of the lipid bilayer. 2. Energetics of insertion of small molecules, peptides, and proteins in membranes, J. Chem. Inf. Model. 51 (2011) 930–946.

[31] A.L. Lomize, I.D. Pogozheva, M.A. Lomize, H.I. Mosberg, The role of hydrophobic interactions in positioning of peripheral proteins in membranes, BMC Struct. Biol. 7 (2007) 44.

[32] A.L. Lomize, I.D. Pogozheva, H.I. Mosberg, Anisotropic solvent model of the lipid bilayer. 1. Parameterization of long-range electrostatics and first solvation shell effects, J. Chem. Inf. Model. 51 (2011) 918–929.

[33] C. Laurence, P. Nicolet, M.T. Dalati, J.L.M. Abboud, R. Notario, The empirical-treatment of solvent solute interactions — 15 years of π*, J. Phys. Chem. 98 (1994) 5807–5816.

[34] M.H. Abraham, Scales of solute hydrogen-bonding — their construction and application to physicochemical and biochemical processes, Chem. Soc. Rev. 22 (1993) 73–83.

[35] C.J. Cramer, D.G. Truhlar, Implicit solvation models: equilibria, structure, spectra, and dynamics, Chem. Rev. 99 (1999) 2161–2200.

[36] N. Kučerka, J.F. Nagle, J.N. Sachs, S.E. Feller, J. Pencer, A. Jackson, J. Katsaras, Lipid bilayer structure determined by the simultaneous analysis of neutron and x-ray scattering data, Biophys. J. 95 (2008) 2356–2367.

[37] D. Marsh, Membrane water-penetration profiles from spin labels, Eur. Biophys. J. Biophys. 31 (2002) 559–562.

[38] N. Kučerka, Y.F. Liu, N.J. Chu, H.I. Petrache, S.T. Tristram-Nagle, J.F. Nagle, Structure of fully hydrated fluid phase DMPC and DLPC lipid bilayers using X-ray scattering from oriented multilamellar arrays and from unilamellar vesicles, Biophys. J. 88 (2005) 2626–2637.

[39] S.D. Guler, D.D. Ghosh, J.J. Pan, J.C. Mathai, M.L. Zeidel, J.F. Nagle, S. Tristram-Nagle, Effects of ether vs. ester linkage on lipid bilayer structure and water permeability, Chem. Phys. Lipids 160 (2009) 33–44.

[40] S. Tristram-Nagle, R. Chan, E. Kooijman, P. Uppamoochikkal, W. Qiang, D.P. Weliky, J.F. Nagle, HIV fusion peptide penetrates, disorders, and softens T-Cell membrane mimics, J. Mol. Biol. 402 (2010) 139–153.

[41] S. Tristram-Nagle, D.J. Kim, N. Akhunzada, N. Kučerka, J.C. Mathai, J. Katsaras, M. Zeidel, J.F. Nagle, Structure and water permeability of fully hydrated diphytanoylPC, Chem. Phys. Lipids 163 (2010) 630–637.

[42] N. Kučerka, B.W. Holland, C.G. Gray, B. Tomberli, J. Katsaras, Scattering density profile model of POPG bilayers as determined by molecular dynamics simulations and small-angle neutron and X-ray scattering experiments, J. Phys. Chem. B 116 (2012) 232–239.

[43] N. Kučerka, S. Tristram-Nagle, J.F. Nagle, Structure of fully hydrated fluid phase lipid bilayers with monounsaturated chains, J. Membr. Biol. 208 (2005) 193–202.

[44] G.Y. Meng, R. Fronzes, V. Chandran, H. Remaut, G. Waksman, Protein oligomerization in the bacterial outer membrane, Mol. Membr. Biol. 26 (2009) 136–145.

[45] J.P. Hickey, D.R. Passino-Reader, Linear solvation energy relationships — rules of thumb for estimation of variable values, Environ. Sci. Technol. 25 (1991) 1753–1760.

[46] Y. Marcus, The properties of organic liquids that are relevant to their use as solvating solvents, Chem. Soc. Rev. 22 (1993) 409–416.

[47] M.H. Abraham, Y.H. Zhao, Determination of solvation descriptors for ionic species: hydrogen bond acidity and basicity, J. Org. Chem. 69 (2004) 4677–4685.

[48] D. Marsh, Polarity and permeation profiles in lipid membranes, Proc. Natl. Acad. Sci. U. S. A. 98 (2001) 7777–7782.

[49] J.C. Mathai, S. Tristram-Nagle, J.F. Nagle, M.L. Zeidel, Structural determinants of water permeability through the lipid membrane, J. Gen. Physiol. 131 (2008) 69–76.

[50] J.J. Pan, S. Tristram-Nagle, J.F. Nagle, Effect of cholesterol on structural and mechanical properties of membranes depends on lipid chain saturation, Phys. Rev. E Stat. Nonlin. Soft Matter Phys. 80 (2009) 021931.

[51] G. von Heijne, The distribution of positively charged residues in bacterial inner membrane-proteins correlates with the trans-membrane topology, EMBO J. 5 (1986) 3021–3027.

[52] G. von Heijne, Control of topology and mode of assembly of a polytopic membrane-protein by positively charged residues, Nature 341 (1989) 456–458.

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

[54] S.H. White, G. von Heijne, The machinery of membrane protein assembly, Curr. Opin. Struct. Biol. 14 (2004) 397–404.

[55] C.R. Sanders, K.F. Mittendorf, Tolerance to changes in membrane lipid composition as a selected trait of membrane proteins, Biochemistry 50 (2011) 7858–7867.

[56] V. Chandran, R. Fronzes, S. Duquerroy, N. Cronin, J. Navaza, G. Waksman, Structure of the outer membrane complex of a type IV secretion system, Nature 462 (2009) 1011–1015.

[57] M. Faller, M. Niederweis, G.E. Schulz, The structure of a mycobacterial outer-membrane channel, Science 303 (2004) 1189–1192.

[58] C. Hoffmann, A. Leis, M. Niederweis, J.M. Plitzko, H. Engelhardt, Disclosure of the mycobacterial outer membrane: cryo-electron tomography and vitreous sections reveal the lipid bilayer structure, Proc. Natl. Acad. Sci. U. S. A. 105 (2008) 3963–3967.

[59] B. Zuber, M. Chami, C. Houssin, J. Dubochet, G. Griffiths, M. Daffe, Direct visualization of the outer membrane of mycobacteria and corynebacteria in their native state, J. Bacteriol. 190 (2008) 5672–5680.

[60] M. Niederweis, O. Danilchanka, J. Huff, C. Hoffmann, H. Engelhardt, Mycobacterial outer membranes: in search of proteins, Trends Microbiol. 18 (2009) 109–116.

[61] G.M. Cook, M. Berney, S. Gebhard, M. Heinemann, R.A. Cox, O. Danilchanka, M. Niederweis, Physiology of mycobacteria, Adv. Microb. Physiol. 55 (2009) 81–182.

[62] M. Niederweis, Mycobacterial porins — new channel proteins in unique outer membranes, Mol. Microbiol. 49 (2003) 1167–1177.

[63] M. Mahfoud, S. Sukumaran, P. Hulsmann, K. Grieger, M. Niederweis, Topology of the porin MspA in the outer membrane of *Mycobacterium smegmatis*, J. Biol. Chem. 281 (2006) 5908–5915.

[64] C. Benning, Mechanisms of lipid transport involved in organelle biogenesis in plant cells, Annu. Rev. Cell Dev. Biol. 25 (2009) 71–91.

[65] H. Nikaido, Outer membranes, Gram-negative bacteria, in: M. Schaechter (Ed.), Encyclopedia of Microbiology, Academic Press, Oxford, UK, 2009, pp. 439–452.

[66] F.R. Maxfield, G. van Meer, Cholesterol, the central lipid of mammalian cells, Curr. Opin. Cell Biol. 22 (2010) 422–429.

[67] Y. Koga, Thermal adaptation of the archaeal and bacterial lipid membranes, Archaea 2012 (2012) 789652, http://dx.doi.org/10.1155/2012/789652.

[68] A. Gliozzi, A. Relini, P.L.G. Chong, Structure and permeability properties of biomimetic membranes of bolaform archaeal tetraether lipids, J. Membr. Sci. 206 (2002) 131–147.

[69] M. De Rosa, A. Gambacorta, A. Gliozzi, Structure, biosynthesis, and physicochemical properties of archaebacterial lipids, Microbiol. Rev. 50 (1986) 70–80.

[70] R. Bartucci, A. Gambacorta, A. Gliozzi, D. Marsh, L. Sportelli, Bipolar tetraether lipids: chain flexibility and membrane polarity gradients from spin-label electron spin resonance, Biochemistry 44 (2005) 15017–15023.

[71] J.A. Killian, G. von Heijne, How proteins adapt to a membrane–water interface, Trends Biochem. Sci. 25 (2000) 429–434.

[72] W.M. Yau, W.C. Wimley, K. Gawrisch, S.H. White, The preference of tryptophan for membrane interfaces, Biochemistry 37 (1998) 14713–14718.

[73] H.D. Hong, S. Park, R.H.F. Jimenez, D. Rinehart, L.K. Tamm, Role of aromatic side chains in the folding and thermodynamic stability of integral membrane proteins, J. Am. Chem. Soc. 129 (2007) 8320–8327.

[74] W. Liu, M. Caffrey, Interactions of tryptophan, tryptophan peptides, and tryptophan alkyl esters at curved membrane interfaces, Biochemistry 45 (2006) 11713–11726.

[75] D.L. Minor, P.S. Kim, Measurement of the beta-sheet-forming propensities of amino-acids, Nature 367 (1994) 660–663.

[76] C.N. Pace, J.M. Scholtz, A helix propensity scale based on experimental studies of peptides and proteins, Biophys. J. 75 (1998) 422–427.

[77] K. Mitra, T. Ubarretxena-Belandia, T. Taguchi, G. Warren, D.M. Engelman, Modulation of the bilayer thickness of exocytic pathway membranes by membrane proteins rather than cholesterol, Proc. Natl. Acad. Sci. U. S. A. 101 (2004) 4083–4088.

[78] M.A. Surma, C. Klose, K. Simons, Lipid-dependent protein sorting at the trans-Golgi network, Biochim. Biophys. Acta Mol. Cell Biol. Lipids 1821 (2012) 1059–1067.

[79] X.W. Cao, M.A. Surma, K. Simons, Polarized sorting and trafficking in epithelial cells, Cell Res. 22 (2012) 793–805.

[80] N. Borgese, S. Colombo, E. Pedrazzini, The tale of tail-anchored proteins: coming from the cytosol and looking for a membrane, J. Cell Biol. 161 (2003) 1013–1019.

[81] N. Borgese, S. Brambillasca, S. Colombo, How tails guide tail-anchored proteins to their destinations, Curr. Opin. Cell Biol. 19 (2007) 368–375.

[82] C. Fernandez, C. Hilty, G. Wider, K. Wüthrich, Lipid–protein interactions in DHPC micelles containing the integral membrane protein OmpX investigated by NMR spectroscopy, Proc. Natl. Acad. Sci. U. S. A. 99 (2002) 13533–13537.

[83] H.J. Snijder, P.A. Timmins, K.H. Kalk, B.W. Dijkstra, Detergent organisation in crystals of monomeric outer membrane phospholipase A, J. Struct. Biol. 141 (2003) 122–131.

[84] T. Abraham, S.R. Schooling, M.P. Nieh, N. Kučerka, T.J. Beveridge, J. Katsaras, Neutron diffraction study of *Pseudomonas aeruginosa* lipopolysaccharide bilayers, J. Phys. Chem. B 111 (2007) 2477–2483.

[85] N. Kučerka, E. Papp-Szabo, M.P. Nieh, T.A. Harroun, S.R. Schooling, J. Pencer, E.A. Nicholson, T.J. Beveridge, J. Katsaras, Effect of cations on the structure of bilayers formed by lipopolysaccharides isolated from *Pseudomonas aeruginosa* PAO1, J. Phys. Chem. B 112 (2008) 8057–8062.

[86] L.K. Tamm, H. Hong, B.Y. Liang, Folding and assembly of beta-barrel membrane proteins, Biochim. Biophys. Acta Biomembr. 1666 (2004) 250–263.

[87] E. London, Insights into lipid raft structure and formation from experiments in model membranes, Curr. Opin. Struct. Biol. 12 (2002) 480–486.

[88] D. Lingwood, K. Simons, Lipid rafts as a membrane-organizing principle, Science 327 (2010) 46–50.

[89] A.J. Chicco, G.C. Sparagna, Role of cardiolipin alterations in mitochondrial dysfunction and disease, Am. J. Physiol. Cell Physiol. 292 (2007) C33–C44.

[90] M. Schlame, J.A. Towbin, P.M. Heerdt, R. Jehle, S. DiMauro, T.J.J. Blanck, Deficiency of tetralinoleoyl-cardiolipin in Barth syndrome, Ann. Neurol. 51 (2002) 634–637.

[91] H. Labischinski, G. Barnickel, H. Bradaczek, D. Naumann, E.T. Rietschel, P. Giesbrecht, High state of order of isolated bacterial lipopolysaccharide and its possible contribution to the permeation barrier property of the outer-membrane, J. Bacteriol. 162 (1985) 9–20.

[92] C.J. Dong, K. Beis, J. Nesper, A.L. Brunkan-LaMontagne, B.R. Clarke, C. Whitfield, J.H. Naismith, Wza the translocon for *E.coli* capsular polysaccharides defines a new class of membrane protein, Nature 444 (2006) 226–229.

[93] M. Bayrhuber, T. Meins, M. Habeck, S. Becker, K. Giller, S. Villinger, C. Vonrhein, C. Griesinger, M. Zweckstetter, K. Zeth, Structure of the human voltage-dependent anion channel, Proc. Natl. Acad. Sci. U. S. A. 105 (2008) 15370–15375.

[94] R. Ujwal, D. Cascio, J.P. Colletier, S. Faham, J. Zhang, L. Toro, P.P. Ping, J. Abramson, The crystal structure of mouse VDAC1 at 2.3 angstrom resolution reveals mechanistic insights into metabolite gating, Proc. Natl. Acad. Sci. U. S. A. 105 (2008) 17742–17747.

[95] D. Thomas, P. Bron, T. Weimann, A. Dautant, M.F. Giraud, P. Paumard, B. Salin, A. Cavalier, J. Velours, D. Brethes, Supramolecular organization of the yeast F1Fo-ATP synthase, Biol. Cell 100 (2008) 591–601.

[96] R.J. Devenish, M. Prescott, A.J.W. Rodgers, The structure and function of mitochondrial F(1)F(0)-ATP synthases, Int. Rev. Cell Mol. Biol. 267 (2008) 1–58.

[97] R. Jackups, J. Liang, Interstrand pairing patterns in beta-barrel membrane proteins: the positive-outside rule, aromatic rescue, and strand registration prediction, J. Mol. Biol. 354 (2005) 979–993.

[98] J. Qu, S. Behrens-Kneip, O. Holst, J.H. Kleinschmidt, Binding regions of outer membrane protein A in complexes with the periplasmic chaperone Skp. A site-directed fluorescence study, Biochemistry 48 (2009) 4926–4936.

[99] R.G. Hanshaw, R.V. Stahelin, B.D. Smith, Noncovalent keystone interactions controlling biomembrane structure, Chemistry 14 (2008) 1690–1697.

[100] T.-X. Xiang, Y.-H. Xu, B.D. Anderson, The barrier domain for solute permeation varies with lipid bilayer phase structure, J. Membr. Biol. 165 (1998) 77–90.

[101] S.J. Marrink, H.J.C. Berendsen, Permeation process of small molecules across lipid membranes studied by molecular dynamics simulations, J. Phys. Chem. 100 (1996) 16729–16738.

[102] T.H. Haines, Do sterols reduce proton and sodium leaks through lipid bilayers? Prog. Lipid Res. 40 (2001) 299–324.

[103] D.A. Erilov, R. Bartucci, R. Guzzi, A.A. Shubin, A.G. Maryasov, D. Marsh, S.A. Dzuba, L. Sportelli, Water concentration profiles in membranes measured by ESEEM of spin-labeled lipids, J. Phys. Chem. B 109 (2005) 12003–12013.

[104] O.H. Griffith, P.J. Dehlinge, S.P. Van, Shape of hydrophobic barrier of phospholipid bilayers (evidence for water penetration in biological-membranes), J. Membr. Biol. 15 (1974) 159–192.

[105] E.A. Disalvo, The role of water in the surface properties of lipid bilayers and its influence on permeability, in: T.S. Sörensen (Ed.), Surface Chemistry and Electrochemistry of Membranes, CRC Press, New York, NY, 1999, pp. 837–870.

[106] W.K. Subczynski, A. Wisniewska, J.J. Yin, J.S. Hyde, A. Kusumi, Hydrophobic barriers of lipid bilayer-membranes formed by reduction of water penetration by alkyl chain unsaturation and cholesterol, Biochemistry 33 (1994) 7670–7681.

[107] H.C. Gaede, W.M. Yau, K. Gawrisch, Electrostatic contributions to indole–lipid interactions, J. Phys. Chem. B 109 (2005) 13014–13023.