

Properties of the Membrane Binding Component of Catechol-O-methyltransferase Revealed by Atomistic Molecular Dynamics Simulations

Adam Orłowski,^{†,‡} Jean-François St-Pierre,[§] Aniket Magarkar,^{||} Alex Bunker,^{||,⊥} Marta Pasenkiewicz-Gierula,[†] Ilpo Vattulainen,^{‡,#,▽} and Tomasz Róż^{*,‡}

[†]Department of Computational Biophysics and Bioinformatics, Faculty of Biotechnology, Biochemistry and Biophysics, Jagiellonian University, Gronostajowa 7, Poland

[‡]Department of Physics, Tampere University of Technology, P.O. Box 692, FI-33101 Tampere, Finland

[§]Departement de Physique and Regroupement Quebecois sur les Materiaux de Pointe, Universite de Montreal, C.P. 6128, Succursale Centre-ville, Montreal (Quebec), Canada

^{||}Centre for Drug Research, Faculty of Pharmacy, University of Helsinki, P.O. Box 56, FI-00014, University of Helsinki, Finland

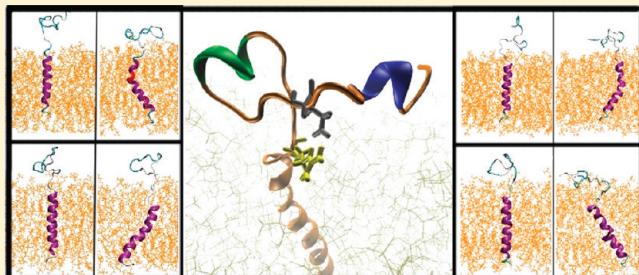
[⊥]Department of Chemistry, Aalto University School of Science, P.O. Box 6100, FI-02015, AALTO, Espoo, Finland

[#]Department of Applied Physics, Aalto University School of Science, Finland

[▽]MEMPHYS—Center for Biomembrane Physics, University of Southern Denmark, Odense, Denmark

Supporting Information

ABSTRACT: We used atomistic simulations to study the membrane-bound form of catechol-O-methyltransferase (MB-COMT). In particular we investigated the 26-residue transmembrane α -helical segment of MB-COMT together with the 24-residue fragment that links the transmembrane component to the main protein unit that was not included in our model. In numerous independent simulations we observed the formation of a salt bridge between ARG27 and GLU40. The salt bridge closed the flexible loop that formed in the linker and kept it in the vicinity of the membrane-water interface. All simulations supported this conclusion that the linker has a clear affinity for the interface and preferentially arranges its residues to reside next to the membrane, without a tendency to relocate into the water phase. Furthermore, an extensive analysis of databases for sequences of membrane proteins that have a single transmembrane helical segment brought about an interesting view that the flexible loop observed in our work can be a common structural element in these types of proteins. In the same spirit we close the article by discussing the role of salt bridges in the formation of three-dimensional structures of membrane proteins that exhibit a single transmembrane helix.



1. INTRODUCTION

About 30% of the genome codes for *membrane proteins*, which function in a number of ways, including diverse tasks such as signaling, recognition, and transport. Membrane proteins come in two varieties: *peripheral proteins*, which are loosely associated with the membrane, and *integral proteins*, which are embedded in a cell membrane. The domain anchoring the integral protein to a membrane is typically a highly symmetric structure composed of one or more transmembrane helices, which are discussed in this work.

The three-dimensional (3D) structure of proteins is usually determined by techniques that are based on X-ray or electron scattering. As the main principles of these techniques are well established, the challenge associated with protein structure determination is closely related to the crystallization of proteins, a crucial step prior to conducting scattering measurements. Despite many related difficulties, the progress in the determination of

water-soluble protein structures has nonetheless been substantial. The situation is much more complicated with membrane proteins, the main reason being that protein structure should in this case be determined under conditions where the protein is solvated by and crystallized in a lipid membrane. This is a major challenge, and while some techniques to this end have been developed,^{1,2} the number of solved 3D structures of membrane proteins has remained modest. The cases where the complete protein structure is known are particularly limited in number, as quite a few reported protein structures refer to incomplete systems where the structures of only a subset of the domains that make up the protein have been determined.

Received: July 27, 2011

Revised: October 9, 2011

Published: October 11, 2011

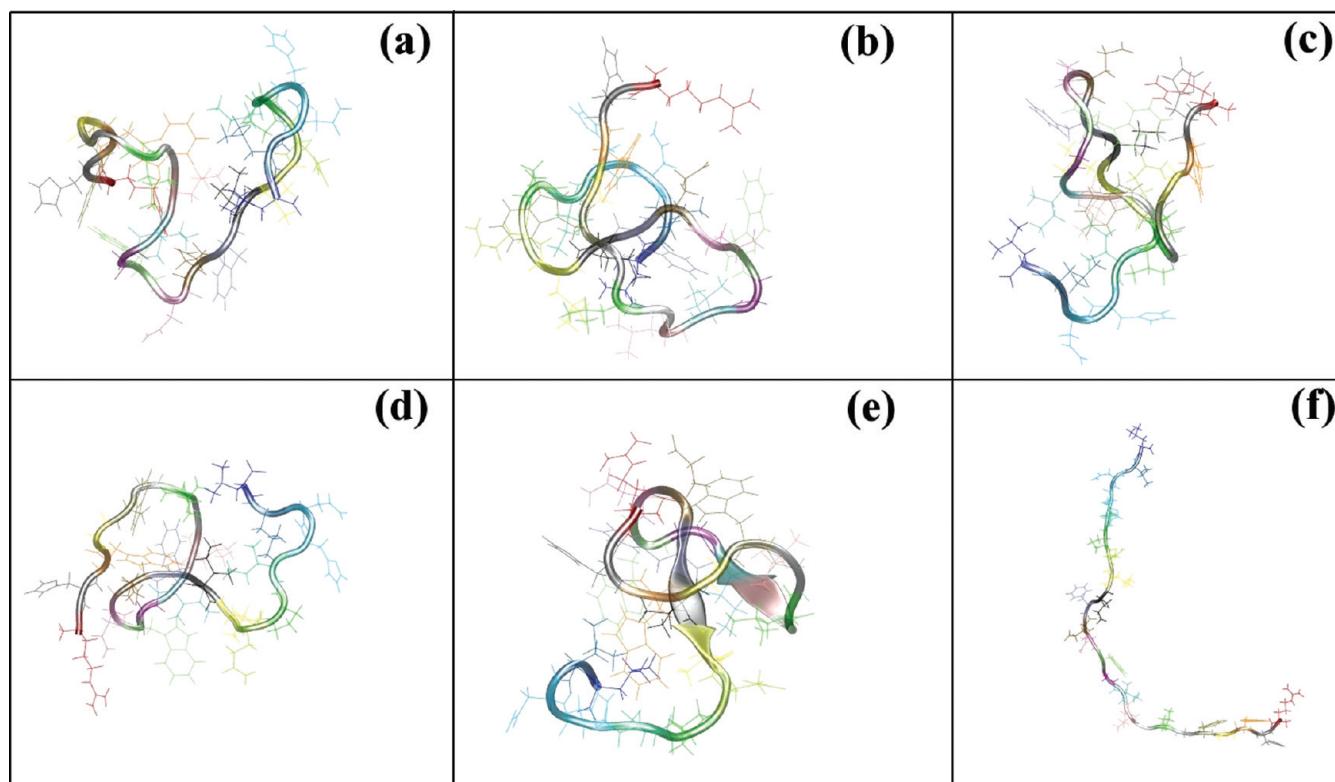


Figure 1. Initial structures of the linker fragments. Structures a–e originate from replica exchange simulation, while structure f is an extended conformation. Figures were generated with the VMD program.⁴⁹

These concerns are exemplified by catechol-O-methyltransferase (COMT), a protein of significant pharmaceutical interest: it has an important role in neurochemistry, where one of its main functions is the inactivation, through O-methylation, of catecholamine neurotransmitters such as dopamine, epinephrine, nor-epinephrine, and catechol steroids.³ It is also known to methylate a wide variety of xenobiotics that contain the catechol group. Among these xenobiotics is the leading drug for treatment of Parkinson's disease,³ L-Dopa, and it is this activity that drives the development of inhibitors for COMT (for a recent review, see the work of Männistö et al.⁴).

COMT is known to exist in two forms: the water-soluble (S) form known as S-COMT, and the membrane-bound (MB) form known as MB-COMT. S-COMT is essentially a catalytic domain since it is identical in sequence and structure to the catalytic domain of MB-COMT. The difference between the two forms lies in the additional 50-residue fragment that is covalently attached to the N-terminus of the catalytic domain of MB-COMT. A sequence analysis of this additional fragment clearly indicates that its 26 N-terminal residues form a single transmembrane helix, while the remaining 24 residues constitute a linker segment that connects the helix to the catalytic domain. Interestingly, while the water-soluble and membrane-bound forms of COMT have an identical catalytic domain, enzymatic kinetics of the two forms are different.⁵ It can be hypothesized that the cause of this difference lies in the interaction of the protein with a membrane, possibly mediated by the linker segment.

The 3D structure of S-COMT has been determined,⁶ but the structure of the entire MB-COMT remains unclear. There is data in favor of a helical structure for the transmembrane part (see, e.g., ref 7), but finding the structure of the linker segment has

turned out to be more complicated. While bioinformatics tools have been used to reconstruct the structure of the whole protein,⁷ the relevance of this approach can be debated: almost all of the templates used in protein structure prediction through sequence identity are structures of water-soluble proteins. This is not the case for the COMT linker segment, which is instead located at a water–membrane interface.

In this study, we use atomistic molecular dynamics (MD) simulations to predict the behavior of the linker segment of MB-COMT at a water–membrane interface. This study is an extension of our previous work, which focused mainly on methodological issues of rat S-COMT,⁸ where the Apo enzyme was simulated for 70 ns starting from a structure crystallized with an inhibitor.

2. METHODS

We performed MD simulations of a lipid bilayer with the 50-residue fragment that differentiates MB-COMT from S-COMT. As described above, the fragment consists of a 26-residue transmembrane helix and a 24-residue linker segment that connects the helix to the catalytic domain (S-COMT).

The lipid bilayer was composed of 124 dilinoleylphosphatidylcholine (2-18:2c9 PC) (DLPC) molecules symmetrically distributed in the two leaflets. Its initial structure was that of our previous study^{9,10} from which four lipid molecules were removed to create a void for the transmembrane helix. The initial structure of the 26-amino acid transmembrane helix (MPEAPP-LLLAAVLLGLVLLVVLLLL) was chosen to be α -helical. As for the 24-amino acid linker segment (RHWGWGLCLIGWNE-FILQPIHNLL), its 3D structure was not known. Therefore, five

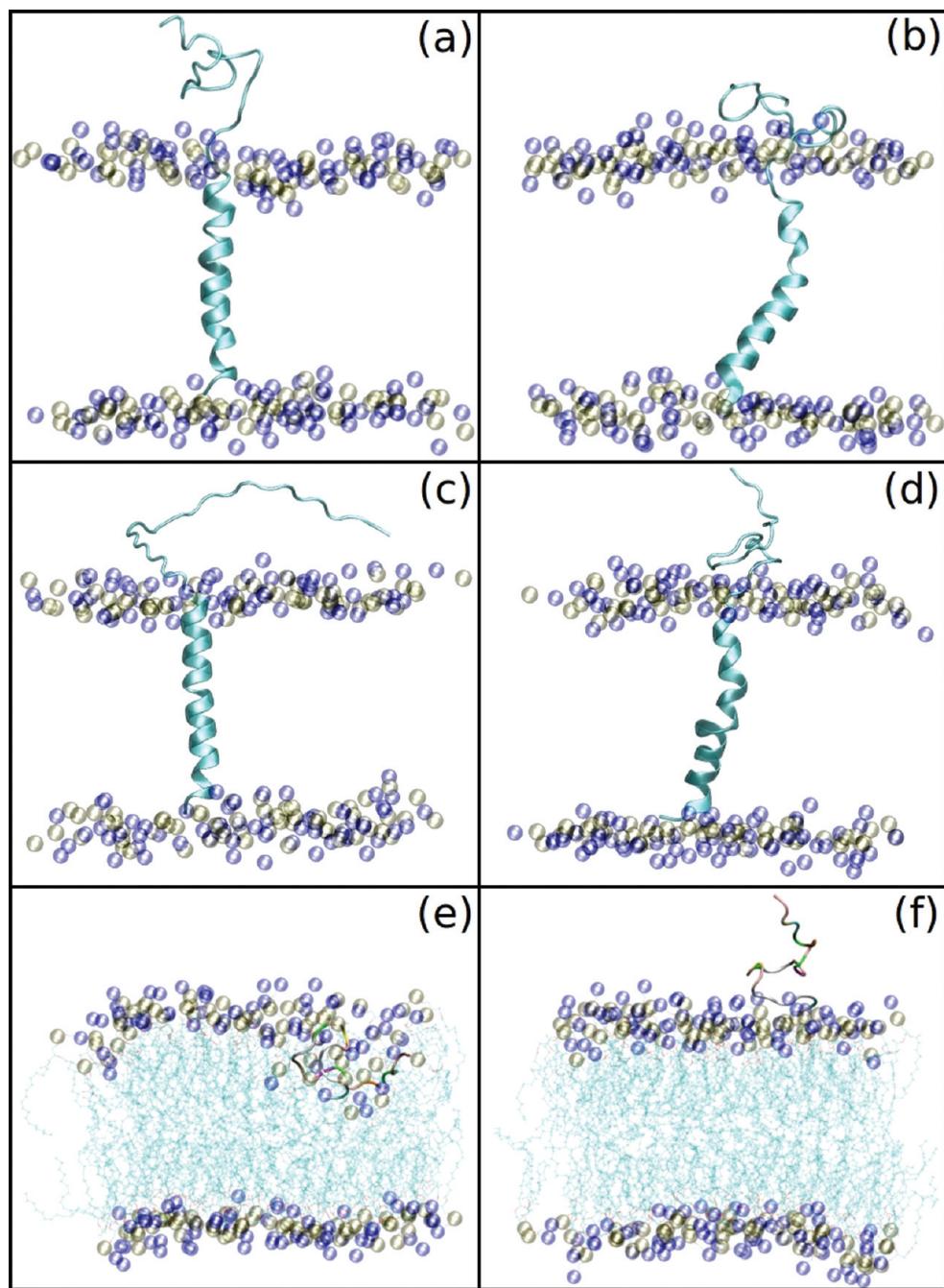


Figure 2. Initial (a,c,e) and final (b,d,f) structures of lipid bilayers with a transmembrane helix and a linker. While the peptide is shown in full, only the phosphorus atoms are shown for the lipids, and water is not shown for clarity. Panels a and b represent the simulation initiated using the initial peptide structure configuration shown in Figure 1a; panels c and d correspond to the structure depicted in Figure 1f. Panel e shows the initial structure of the system where the peptide was pulled into the hydrocarbon region of the bilayer (based on Figure 1e), and panel f shows this structure after a simulation of 40 ns. The figure was generated using VMD.⁴⁹

initial conformations were generated (Figure 1) for the linker using replica exchange molecular dynamics (REMD) in conjunction with the optimized potential for efficient structure prediction (OPEP).¹¹ REMD simulations were executed at 16 temperatures ranging from 222.5 to 525 K. The simulation was executed for 100 ns at a 1.5 fs time step, and temperature exchange was attempted every 15 ps. In each of these conformations, the linker segment was covalently attached to the transmembrane helix. A similar procedure was performed for an extended conformation

(Figure 1) that was the sixth case considered in our work, and all of these six conformations were finally inserted into the DLPC bilayer. Examples of system structures are shown in Figure 2.

To further study the linker segment in particular, we constructed six additional systems where the linker in each of the above conformations was placed at the water–bilayer interface (without the transmembrane helix).

Altogether, 12 systems were constructed, each hydrated with ~8000 water molecules. The systems also had 20 K⁺ and 20 Cl⁻

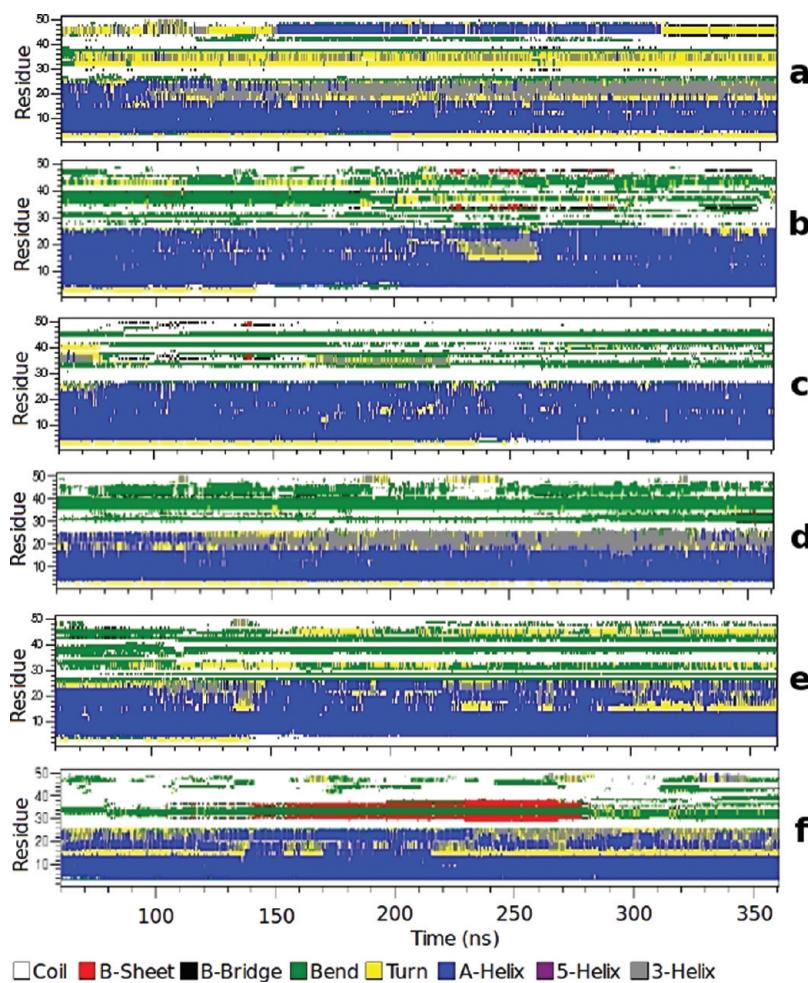


Figure 3. Secondary structure elements of a transmembrane helix (residues 1–26) and a linker (residues 27–50) shown as a function of time determined with DSSP.²⁹ Panels a–f correspond to simulations started using as the initial configuration the structures shown in Figure 1 labeled a–f, respectively.

ions that were added to the water phase to mimic the KCl concentration of 140 mM in the intracellular environment.

For the lipids, peptides, and ions, we used the all-atom OPLS force field.^{12–15} Details of the parametrization are discussed in our previous articles.^{9,10,16} Partial charges on the PC head groups were taken from Takaoka et al.¹⁷ as they were derived in compliance with the OPLS methodology. Definitions of the charge groups are given in our previous work.¹⁶ For water, we employed the TIP3 model.¹⁸

Prior to MD simulations, the structures of the systems were optimized to remove van der Waals contacts between atoms. Energy minimizations and MD simulations were performed using the GROMACS 4 software package.^{19,20} The simulations of the six bilayer systems with the transmembrane helix and the linker were carried out for 360 ns. The last 300 ns fragment of each trajectory was used for analysis. The remaining six bilayer systems with the linker segment placed at the water–bilayer interface were simulated for 100 ns each, and of these simulations the last 60 ns were used for analysis. Periodic boundary conditions were employed in all three directions. The LINCS algorithm was used to preserve covalent bond lengths.²¹ The time step was set to 2 fs. The temperature (300 K) and pressure (1 bar) were controlled using the Nosé–Hoover^{22,23} and Parrinello–Rahman methods,²⁴ respectively. The temperatures of the solute and the

solvent were controlled independently. For pressure, a semi-isotropic control was used. The Lennard-Jones interactions were cut off at 1.0 nm, and the electrostatic interactions were calculated using the particle mesh Ewald method with a real space cutoff of 1.0 nm.^{25,26} This protocol was successfully used in our previous studies.^{9,27,28}

3. RESULTS

3.1. Characteristics of the Peptide Structure. To characterize the conservation of the 50-residue peptide structure in the bilayer systems, we monitored the temporary secondary structure of both fragments (the helix and the linker) using the DSSP program.²⁹ Results are shown in Figures 3 and 4. As can be seen from Figure 3, the structure of the transmembrane helix is generally preserved; however, there are short periods when the helicity is temporarily either lost or altered from an α -helix to a more extended 3–10 helix. The spatial structure of the linker was not stable in any of the simulated systems. The dominant structural motives of the linker segment covalently attached to the transmembrane helix were bends, coils, and turns; however, the formation of a helical fragment was also observed in one of the simulations. In the systems where the linker was loosely associated with the bilayer, the dominant structural motives were

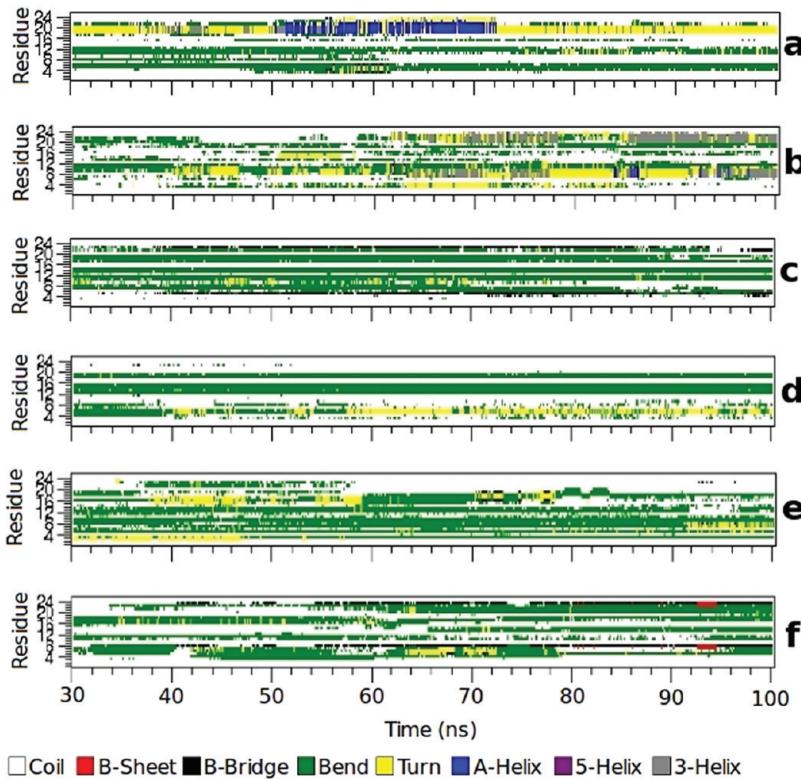


Figure 4. Secondary structure elements of a linker loosely associated with a membrane (without the transmembrane part) shown as a function of time determined with DSSP.²⁹ Panels a–f correspond to simulations started using as the initial configuration the structures shown in Figure 1 labeled a–f, respectively.

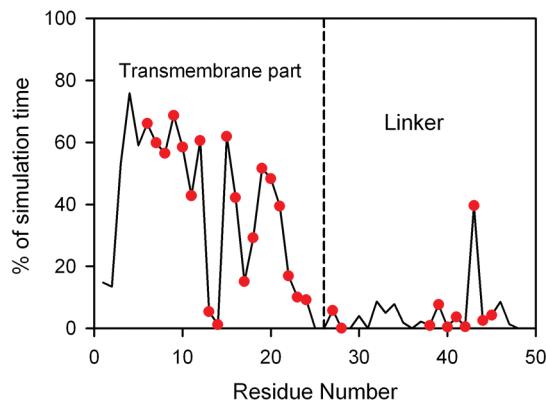


Figure 5. Percentage of simulation time for which a given residue adopts helical conformation (data averaged over all 6 runs and time). In red are marked points whose conformation is predicted to be helical in the studies of Bui et al.⁷

transient bends and turns (Figure 4). The percentage of the simulation time for which a given residue participates in a particular local conformation is given in Tables S1 and S2 presented in the Supporting Information (SI).

The structures of the helix and the linker segment have been previously analyzed using bioinformatics tools by Bai et al.⁷ They predicted, for each peptide residue, the probability of its participation in the helical structure and found that residues 38–45 of the 50-residue fragment (residues 12–19 of the 24-residue linker) can adopt a helical structure. We thus, for each residue, recorded cases when the residue is an element of a helical

structure during the simulation (Figure 5). Red points in Figure 5 indicate the residues for which a helical structure was predicted by ref 7. Our analyses indeed indicate that residues 38–45 have a tendency to form a helix, but residues 30–35 also display the same tendency, albeit less apparent. Examples of the helical conformations in the linker segment are shown in Figure 6.

3.2. Peptide Location. To assess the location of the linker segment relative to the bilayer surface, we first calculated the average position of the phosphorus atoms in the membrane leaflet from which the linker protrudes, and then we computed the time averaged position of each Ca atom of the linker relative to the average phosphorus atom position (Figure 7). Figure 7a shows large variations in the linker behavior but also some common characteristics. The ensemble-averaged (over six linker segments) distance profile is as follows: for residues 27–35, there is a gradual increase of the distance from 0.15 to 1.3 nm followed by a decrease to 0.9 nm for the three next residues, 38–40. For the remaining 10 residues, the distance increases to 1.6 nm. Meanwhile, the ensemble-averaged distance profile for the systems where the linker is loosely associated with the bilayer gradually increases from 0.9 to 1.7 nm with an increasing residue number (Figure 7b). No dip is observed in this profile. It is interesting to note that in each system the linker segment remains associated with the water–membrane interface for the whole simulation run, even though initially it was placed randomly in the water phase near the bilayer surface.

Figure 2e shows the initial conformation of the additional starting structure, which was obtained by pulling the peptide down (we used the structure shown in Figure 1e) into the interface. This initial structure was generated to find out whether the linker could

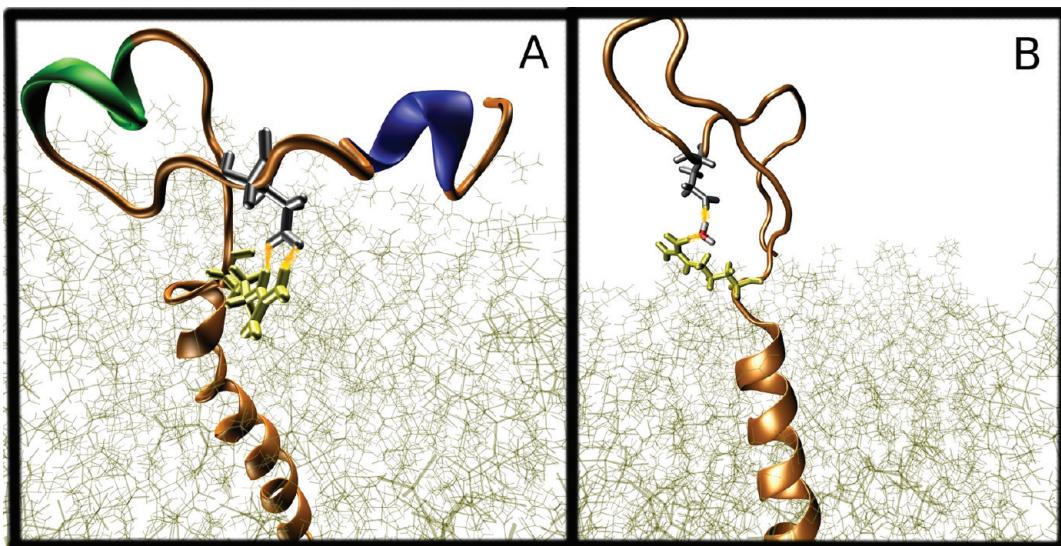


Figure 6. (A) Structure of the peptide at the end of simulation (structure a from Figure 1) with a salt bridge between GLU40 (silver) and ARG27 (yellow). Residues 33–36 (green) and 44–47 (blue) have helical structure. Water is not shown for clarity. (B) Illustration of a water bridge within the linker region; see text for discussion.

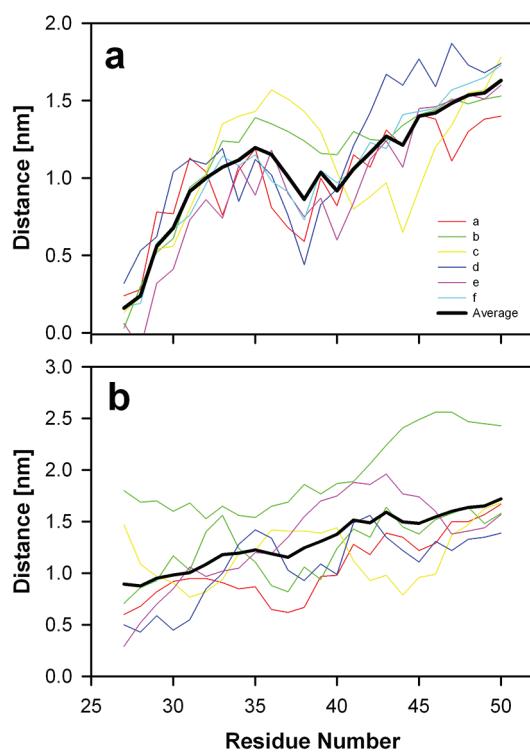


Figure 7. Distance between the membrane–water interface and C α atoms of each linker residue: (a) the linker connected to a transmembrane part and (b) the linker loosely associated with a membrane. Average position of a phosphorus atom at the interface is set to zero. Residue 25 is the first one after the transmembrane helix. Lines a–f correspond to simulations started using as the initial configuration the structures shown in Figure 1 labeled a–f, respectively.

possibly relocate closer to the membrane core in a stable fashion. However, as shown in Figure 2f, the peptide left the bilayer core region relatively quickly and remained associated with the membrane interface in a way similar to that observed in other simulations.

3.3. Peptide–Lipid Interactions. To characterize interactions between peptide and lipids, we evaluated the number of the peptide–lipid hydrogen bonds and charge pairs. In these calculations, we used geometrical criteria established in our previous studies: hydrogen bonds exist if the distance between donor (D) and acceptor (A) is less than 0.325 nm and the angle between a chemical bond D–H and a hydrogen bond D–A is less than 35°.³⁰ Charge pairs are electrostatic interactions between the positively charged methyl groups in the choline moiety of PC and the negatively charged oxygen atoms of neighboring groups of the same or other molecule. These interactions are defined by a distance cutoff of 0.4 nm. The numbers of these peptide–lipid interactions per peptide are given in Table 1. In general, the data indicate that there are more peptide–lipid hydrogen bonds and charge pairs when the linker is covalently attached to the transmembrane helix than otherwise.

Since the carboxyl group of the peptide is not present when a whole protein is attached, we performed additional analysis to estimate how its presence can affect the observed results. As the carboxyl group is negatively charged, it can interact with the positively charged choline group via the charge pairs. Thus first we calculated the frequency of the occurrence of charge pairs (measured as a percentage of simulation time during which any given pair exists). For the case of the covalently bonded peptide, the charge pairs exist over 7.3% of the simulation time, while for the case of the loosely associated peptide, it was 2.5% of the simulation time. Next we calculated the average distance between any atom of lipids and the carboxyl group. The distance is 1.1 and 1.3 nm for the case of covalently bonded and loosely associated peptide, respectively. In Supporting Information Table S3, the values obtained in each MD run are given. Obtained results suggest that the interactions of the C-terminal carboxyl group with the membrane are rather weak and do not affect the observed results. This is consistent with data presented in Figure 7, which indicate that the C-terminus is located further away than other residues from the membrane and does not indicate a specific interaction with lipids.

3.4. Peptide Intramolecular Interactions. Although we did not observe any specific and stable secondary structure elements

Table 1. Average Numbers of Hydrogen Bonds and Charge Pairs between Peptide and Bilayer Lipids Per Peptide Molecule

structure number	hydrogen bonds		charge pairs	
	helix 1–25	linker 26–50	whole 1–50	linker 26–50
Linker Covalently Attached to Helix				
1	1.6	1.5	3.1	6.37
2	1.3	4.0	5.3	5.00
3	1.2	2.5	3.7	6.20
4	1.7	2.0	3.7	6.43
5	1.8	4.9	6.7	7.00
6	2.6	2.1	4.7	3.75
average	1.7	2.8	4.5	5.79
Loosely Associated Linker				
1		2.5		2.62
2		0.15		0.73
3		1.5		3.02
4		1.9		3.05
5		2.5		4.78
6		0.15		3.48
average		1.45		2.95

in the linker segment, we found a tendency to form a specific salt bridge between ARG27 located at the end of the transmembrane helix and GLU40 of the linker segment (Figure 6). This salt bridge was predicted in four selected initial structures regarding the peptides with a transmembrane part. Figure 8a shows the time development of the distance between a nitrogen atom in the ARG27 side chain (N^{γ}) and an ϵ oxygen atom in the GLU40 side chain (O^{ϵ}) (the shortest distance between possible N^{γ} – O^{ϵ} pairs is ~ 0.3 nm). As can be seen in Figure 8a, the above-mentioned salt bridge is observed in all simulations, including the two where it was not present in the initial structures of the peptide. The salt bridge displays a dynamic behavior: although it is present in all simulation systems, it frequently breaks for long periods of time.

The salt bridge is also formed between ARG27 and GLU40 of the linker segment loosely associated with the bilayer. The time profiles of distance between the N^{γ} and O^{ϵ} of these two amino acids are shown in Figure 8b. The salt bridge is, however, created less frequently than in the case of the linker bonded to the transmembrane helix (Figure 8a).

A salt bridge is commonly defined as an interaction between two charged residues when they are hydrogen bonded, thus the distance between nitrogen and oxygen is below 0.3 nm.³¹ In our simulations, we observed two very stable distances for the N^{γ} – O^{ϵ} pair, one of about 0.27 nm, which corresponds to the presence of a hydrogen bond, and the second of about 0.5 nm (see, e.g., the gray curve on Figure 8b, first 20 ns). In these situations the classically understood salt bridge is broken; however, the residues still interact via a water bridge-type interaction (a water molecule is simultaneously hydrogen bonded with two different residues³⁰). As can be seen in Figure 8b, these water bridges are stable. A snapshot of such a water-mediated salt bridge is shown in Figure 6b.

Since this salt bridge seems to be an important structural element, we analyzed its environment more carefully, calculating the solvent-accessible surface area (SASA) as a percentage of the

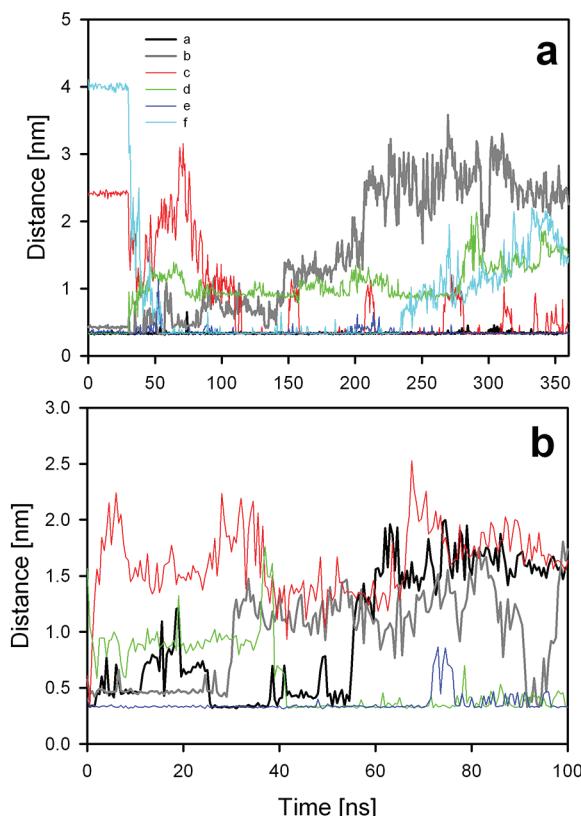


Figure 8. Time development of the distance between nitrogen N^{γ} in ARG27 and oxygen O^{ϵ} in GLU40 in (a) a linker bonded with transmembrane helix and (b) the loosely associated linker. Structure names (a–f) refer to the initial structures shown in Figure 1. Lines a–f correspond to simulations started using as the initial configuration the structures shown in Figure 1 labeled a–f, respectively.

total surface areas of these two groups. We found that the exposure of ARG27 to the water is $(36 \pm 2)\%$ and $(63 \pm 4)\%$ in the bonded and free linker, respectively, and for GLU40 the values are $(40 \pm 4)\%$ and $(49 \pm 5)\%$ in the bonded and free linker, respectively.

3.5. Peptide–Ion Interactions. Ions are known to affect the properties of macromolecules and the interactions between them. Thus, we also analyzed the interactions of K^+ and Cl^- ions with the peptide in addition to their interactions with the bilayer lipids. The analysis of the distribution of both ion types along the bilayer normal revealed that both ions are almost evenly distributed in the water phase and thus do not interact strongly with the lipids headgroups. This behavior is in sharp contrast to that of Na^+ and Cl^- of NaCl, as Na^+ binds strongly to the PC headgroup, while Cl^- remains loosely associated with the membrane.²⁸ This difference between Na^+ and K^+ interactions with the bilayer was previously observed in MD simulation studies where different force field parameters were used.^{32,33}

We next analyzed interactions of K^+ ions with the peptide oxygen atoms. We considered that an ion and an oxygen create a “bonded pair” when the distance between them is less than 0.325 nm, which corresponds to the position of the first minimum of the radial distribution function of ions and such oxygen atoms.³⁴ Our analysis revealed that, on average, there are only 0.20 K^+ ions bound to the entire peptide molecule (transmembrane helix and linker segment) and 0.10 with the loosely attached linker segment.

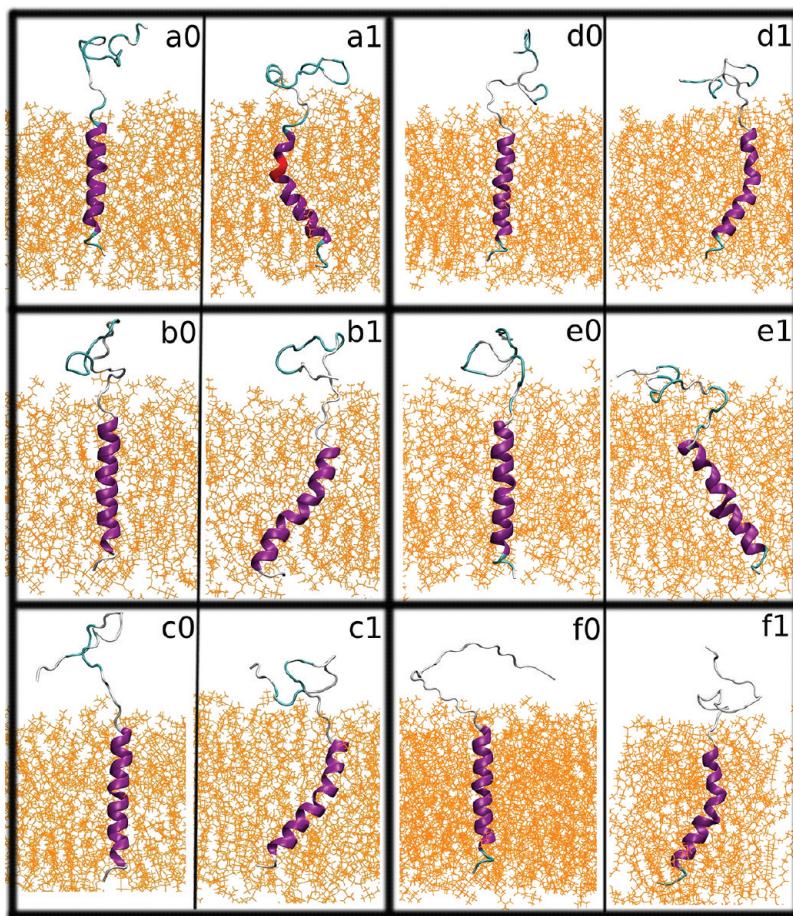


Figure 9. Snapshots of the initial ($a_0, b_0, c_0, d_0, e_0, f_0$) and final ($a_1, b_1, c_1, d_1, e_1, f_1$) simulation structures of the tilted transmembrane helix and linker. The kink at ALA17 is shown in red. Panels a–f correspond to simulations started using as the initial configuration the structures shown in Figure 1 labeled a–f, respectively.

Obviously, ions bind to residues located outside the membrane hydrophobic core, and somewhat preferentially to acidic residues: GLU3, GLU40, and LEU50 (C-terminal). However, the number of K^+ ions bound to these residues is small: a maximum of 0.03 K^+ per residue. In summary, we did not find strong interactions of K^+ and Cl^- with either the peptide or the bilayer lipids.

3.6. Hydrophobic Mismatch. The difference between the thickness of the hydrophobic core of the bilayer and the transmembrane peptide length predispose the peptide considered here to undergo positive hydrophobic mismatch as described in several prior studies.^{35–37} In all of the present simulations, we observed tilting of the transmembrane helix of $18\text{--}30^\circ$ relative to the bilayer normal. The second, parallel effect of the mismatch was the presence of a kink in the helix around the residue GLY15 in four cases, and in the other two, around ALA17 (Figure 9). It is possible that the presence of the linker segment covalently attached to the transmembrane helix and its interactions with the lipids can affect hydrophobic mismatch adjustments. In systems where the linker is located closer to the bilayer surface (Figure 7a in runs a–d), the kink is observed closer to the end of the transmembrane helix at ALA17 and thus closer to the linker.

3.7. Sequence Analysis. The above-discussed results for COMT indicate the formation of a loop at the water–membrane interface. To investigate whether such a loop appears frequently in membrane proteins in general, we analyzed databases for

membrane proteins that have a single transmembrane helix. Of particular interest in this analysis were those sequences that were next to the transmembrane helix. For the sequence analysis, we used the data set created by Sharpe et al.,³⁸ which includes proteins of humans taken from the RefSeq database.³⁹ Information about transmembrane domains was extracted from the literature and the TopDb database.⁴⁰ For those proteins whose transmembrane domain is not known, it was predicted using the TMHMM program.⁴¹ The data set obtained in this fashion consists of 223 protein sequences. To estimate the probability of loop formation in a given peptide sequence, we focused on a 20 amino acid long sequence that followed the transmembrane helix and computed the probability for finding oppositely charged amino acids in the sequence. In the analysis, the linker was divided into five regions, which included amino acids 1–4, 5–7, 7–10, 10–15, and 15–20, counting away from the membrane. The analysis shows that positively charged residues are preferentially placed in the first region. In agreement with Sharpe et al.³⁸ we observed them in 71% of the analyzed sequences. Table 2 shows the probabilities for finding a residue of the opposite charge in fragments 7–10, 7–15, and 7–20 under a condition that in region 1–4 there is a charged residue. As the occurrence of such sequences is high, we conclude that a loop at the interface region can be a common structural element of transmembrane proteins in general.

Table 2. Probability of the Presence of the Oppositely Charged Amino Acids in the Various Fragments of the Transmembrane Protein Sequence

amino acid fragments	occurrence of oppositely charged amino acid [%]
1–3 and 7–10	22.45
1–3 and 7–15	33.97
1–3 and 7–20	44.10

4. DISCUSSION

COMT can exist in two forms: cytosolic (S-COMT) and membrane bound (MB-COMT). In MB-COMT, the catalytic domain (which is equivalent to the S-COMT structure) is anchored to a membrane by a 26-residue transmembrane helix and connected to the helix through a 24-residue linker segment. In the present study, we performed six MD simulations of a membrane with the transmembrane helix and linker segment of MB-COMT, and six MD simulations of a membrane with the linker segment only. These simulations were designed to show the difference in the behavior of the linker either bonded to the transmembrane helix or loosely placed in the water phase near the bilayer surface.

The simulations showed that the linker has a clear affinity for the membrane–water interface and preferentially arranges its residues next to it without a tendency to relocate into the water phase. Such a trend was observed for the linker both bonded with the transmembrane helix and loosely placed at the bilayer surface.

The results of the linker secondary structure analysis carried out in our study are not entirely conclusive. While we observed transient helical structures in the linker, we did not find any *consistent* pattern of the secondary structure that could suggest some specific folding of the linker. This might be interpreted as an absence of stable secondary structure in the linker. However, there are two grounds on which an argument can be made that this conclusion is incorrect: (1) the time scale of the simulation, and (2) the lack of the main protein unit, that is, the catalytic domain, S-COMT. Our simulations, although long in terms of current possibilities (360 ns), may be too short to observe peptide folding,^{42,43} which additionally can be slower in the environment of the water–membrane interface.^{42,43} The presence of the catalytic domain can also influence conditions of folding through altering the local environment, e.g., the hydration level, as well as by excluding some conformations that are observed in the current simulations where the end of the peptide is directed toward the bilayer. At the same time, in a subset of the simulations, we observed the formation of a helical structure (Figure 5), which was predicted in the previous bioinformatics study by Bai et al.⁷ The fact that the membrane interface can stabilize helical structures was documented in another MD simulation study of different peptides⁴⁴ and is known, e.g., for antimicrobial peptides such as magainin.^{45,46} Considering the above, one can hypothesize that the linker segment adopts the structure shown in Figure 6a, with a short helical fragment in the loop that is closed by a salt bridge. To confirm this hypothesis, future studies that include the entire COMT protein are needed.

Although we could not provide a final description of the secondary structure of the linker segment, we were able to observe the presence of a salt bridge between ARG27 located at the end of the transmembrane helix and GLU40 closing a 13-residue long loop. This salt bridge was present in all simulations of the linker for some period of time. In some structures, it was

present in the initial structure but in some it was formed during the simulation. The behavior of the salt bridge was dynamic; it could be broken and reformed periodically. In some cases it could be replaced by a water-mediated interaction known as a water bridge.³⁰ The same type of a salt bridge was occasionally also present in the structure of the linker loosely associated with the bilayer; however, the probability of its formation was clearly lower. This salt bridge seems to be a key element of maintaining the close proximity of the peptide to the membrane. It is possible that this interaction is the cause of a characteristic dip around residue GLU40 in Figure 6, and is also the reason why the rest of the peptide remains in close proximity to the membrane.

In the study of phospholamban, a small 52-residue protein composed of a transmembrane helix and a short peptide outside the membrane, which in many respects resembles the MB-COMT fragment simulated in this study, the formation of a similar salt bridge was observed.⁴⁷ In that case the loop formed was shorter and included only six residues. That observation, together with the results of our simulations, can be connected with an extensive statistical analysis of salt bridges in proteins.⁴⁸ This statistical analysis of known protein structures has shown the preferential formation of salt bridges in an environment of ~30% exposure of the residue surface to the water. This level of partial exposure being found to be optimal has been hypothesized to result from two competing factors: (a) the insertion of charged residues into the hydrophobic core of the protein is energetically unfavorable due to the cost of charge group dehydration, and (b) full exposure to the water might decrease the probability of salt bridging due to the competition for hydrogen bonding with water.⁴⁸ The water–membrane interface in fact creates conditions that lead to lower hydration for the residues located at the end of the transmembrane helix, thus conditions for the formation of a salt bridge seem to be optimal. This lower hydration is documented by our analysis of the SASA that shows exposure of these two residues to the water at a level of 36–40%. At the same time, water exposure of the same residues in a freely placed linker is higher, and thus the frequency of the salt bridge formation is lower. In our present simulations, the main part of the protein was not present, thus we can expect that when it is present it can decrease the number of water molecules in the linker neighborhood to make the formation of the salt bridge more favorable.

An interesting question remains: Is the formation of similar loops stabilized by salt bridges a common phenomenon in the case of transmembrane proteins? As we have discussed above, the water–membrane interface creates an optimal condition for the salt bridge formation between charged residues, and such a salt bridge has been observed in two different MD simulation studies.⁴⁷ Additionally, it is known that positively charged residues, arginine or lysine, are preferentially located at the end of transmembrane peptides in membrane proteins.³⁸ Our additional analysis of the sequences of the single-helix transmembrane proteins showed that in 44% of cases when a charged residue is located at the end of a transmembrane helix, a residue of opposite charge is present at the position 7–20. This indicates that a similar loop might be a common structural element of the transmembrane proteins.

■ ASSOCIATED CONTENT

S Supporting Information. Percentage of simulation time over which a given type of secondary structure was observed in

the system with the linker attached to a transmembrane helix and loosely attached to the bilayer surface. Average shortest distance between the carboxyl group of C-terminus and lipids atoms, the frequency of the contacts between carboxyl group lipids atoms, and the frequency of charge pairs between carboxyl group and lipids choline groups. This material is available free of charge via the Internet at <http://pubs.acs.org>.

AUTHOR INFORMATION

Corresponding Author

*E-mail: tomasz.rog@tut.fi; tomasz.rog@gmail.com.

ACKNOWLEDGMENT

Computational resources were provided by the Finnish IT Centre for Science (CSC) and the SharcNet grid computing facility (www.sharcnet.ca). We wish to thank the Academy of Finland, the Natural Sciences and Engineering Research Council of Canada (NSERC), the Erasmus program, and the Emil Aaltonen (A.B.) and CIMO (A.B.) foundations for financial support.

REFERENCES

- (1) Vinothkumar, K. R.; Henderson, R. *Q. Rev. Biophys.* **2010**, *43*, 65–158.
- (2) McLuskey, K.; Roszak, A. W.; Zhu, Y.; Isaacs, N. W. *Eur. Biophys. J.* **2009**, *39*, 723–755.
- (3) Guldberg, H. C.; Marsden, C. A. *Pharmacol. Rev.* **1975**, *27*, 135–206.
- (4) Männistö, P. T.; Kaakkola, S. *Pharmacol. Rev.* **1999**, *51*, 593–628.
- (5) Reenilä, I.; Männistö, P. T. *Med. Hypotheses* **2001**, *57*, 628–632.
- (6) Vidgren, J.; Svensson, L. A.; Liljas, A. *Nature* **1994**, *368*, 354–358.
- (7) Bai, H.-W.; Shim, J.-Y.; Yu, J.; Zhu, B. T. *Chem. Res. Toxicol.* **2007**, *20*, 1409–1425.
- (8) Bunker, A.; Männistö, P.; St Pierre, J.-F.; Rög, T.; Pomorski, P.; Stimson, L.; Karttunen, M. *SAR QSAR Environ. Res.* **2008**, *19*, 179–189.
- (9) Rog, T.; Martinez-Seara, H.; Munck, N.; Orešić, M.; Karttunen, M.; Vattulainen, I. *J. Phys. Chem. B* **2009**, *113*, 3413–3422.
- (10) Pöyry, S.; Rog, T.; Karttunen, M.; Vattulainen, I. *J. Phys. Chem. B* **2009**, *113*, 15513–15521.
- (11) Derreumaux, P. *J. Chem. Phys.* **1999**, *111*, 2301–2310.
- (12) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1996**, *118*, 11225–11236.
- (13) Rizzo, R. C.; Jorgensen, W. L. *J. Am. Chem. Soc.* **1999**, *121*, 4827–4836.
- (14) Price, M. L. P.; Ostrovsky, D.; Jorgensen, W. L. *J. Comput. Chem.* **2001**, *22*, 1340–1352.
- (15) Kaminski, G. A.; Friesner, R. A.; Tirado-Rives, J.; Jorgensen, W. L. *J. Phys. Chem. B* **2001**, *105*, 6474–6487.
- (16) Stepniewski, M.; Bunker, A.; Pasenkiewicz-Gierula, M.; Karttunen, M.; Rog, T. *J. Phys. Chem. B* **2010**, *114*, 11784–11792.
- (17) Takaoka, Y.; Pasenkiewicz-Gierula, M.; Miyagawa, H.; Kitamura, K.; Tamura, Y.; Kusumi, A. *Biophys. J.* **2000**, *79*, 3118–3138.
- (18) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926.
- (19) Lindahl, E.; Hess, B.; van der Spoel, D. *J. Mol. Model.* **2001**, *7*, 306–317.
- (20) van der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C. *J. Comput. Chem.* **2005**, *26*, 1701–1718.
- (21) Hess, B.; Bekker, H.; Berendsen, H.; Fraaije, J. *J. Comput. Chem.* **1997**, *18*, 1463–1472.
- (22) Nosé, S. *J. Chem. Phys.* **1984**, *81*, 511.
- (23) Hoover, W. G. *Phys. Rev. A* **1985**, *31*, 1695.
- (24) Parrinello, M. *J. Appl. Phys.* **1981**, *52*, 7182.
- (25) Essmann, U.; Perera, L.; Berkowitz, M.; Darden, T.; Lee, H.; Pedersen, L. *J. Chem. Phys.* **1995**, *103*, 8577–8593.
- (26) Patra, M.; Karttunen, M.; Hyvönen, M. T.; Falck, E.; Vattulainen, I. *J. Phys. Chem. B* **2004**, *108*, 4485–4494.
- (27) Kaszuba, K.; Rög, T.; St. Pierre, J.-F.; Männistö, P. T.; Karttunen, M.; Bunker, A. *SAR QSAR Environ. Res.* **2009**, *20*, 595–609.
- (28) Hall, A.; Rög, T.; Karttunen, M.; Vattulainen, I. *J. Chem. Phys. B* **2010**, *114*, 7797–7807.
- (29) Kabsch, W.; Sander, C. *Biopolymers* **1983**, *22*, 2577–2637.
- (30) Rög, T.; Murzyn, K.; Milchoud, J.; Karttunen, M.; Pasenkiewicz-Gierula, M. *J. Phys. Chem. B* **2009**, *113*, 2378–2387.
- (31) Kumar, S.; Nussinov, R. *J. Mol. Biol.* **1999**, *293*, 1241–1255.
- (32) Vácha, R.; Berkowitz, M. L.; Jungwirth, P. *Biophys. J.* **2009**, *96*, 4493–4501.
- (33) Gurtovenko, A. A.; Vattulainen, I. *J. Chem. Phys.* **2009**, *130*, 215107.
- (34) Zhao, W.; Rog, T.; Gurtovenko, A.; Vattulainen, I.; Karttunen, M. *Biophys. J.* **2007**, *92*, 1114–1124.
- (35) Killian, J. A. *FEBS Lett.* **2003**, *555*, 134–138.
- (36) Gil, T.; Ipsen, J. H.; Mouritsen, O. G.; Sabra, M. C.; Sperotto, M. M.; Zuckermann, M. *J. Biochim. Biophys. Acta* **1998**, *1376*, 245–266.
- (37) Rög, T.; Murzyn, K.; Karttunen, M.; Pasenkiewicz-Gierula, M. *J. Pept. Sci.* **2008**, *14*, 374–382.
- (38) Sharpe, H. J.; Stevens, T. J.; Munro, S. *Cell* **2010**, *142*, 158–169.
- (39) Pruitt, K. D.; Tatusova, T.; Klimke, W.; Maglott, D. R. *Nucleic Acids Res.* **2009**, *37*, D32–36.
- (40) Tusnady, G. E.; Kalmar, L.; Simon, I. *Nucleic Acids Res.* **2007**, *36*, D234–D239.
- (41) Krogh, A.; Larsson, B.; von Heijne, G.; Sonnhammer, E. L. *J. Mol. Biol.* **2001**, *305*, 567–580.
- (42) Lindahl, E.; Sansom, M. S. *Curr. Opin. Struct. Biol.* **2008**, *18*, 425–431.
- (43) Seibert, M. M.; Patriksson, A.; Hess, B.; van der Spoel, D. *J. Mol. Biol.* **2005**, *354*, 173–183.
- (44) Polverini, E.; Coll, E. P.; Tielemans, D. P.; Harauz, G. *Biochim. Biophys. Acta* **2011**, *1808*, 674–683.
- (45) Khandelia, H.; Ipsen, J. H.; Mouritsen, O. G. *Biochim. Biophys. Acta* **2008**, *1778*, 1528–1536.
- (46) Murzyn, K.; Rög, T.; Pasenkiewicz-Gierula, M. In *Computational Science - ICCS 2004*; Bubak, M., Albada, G. D., Sloot, P. M. A., Dongarra, J., Eds.; Springer: Berlin/Heidelberg, 2004; Vol. 3037, pp 325–331.
- (47) Kim, T.; Lee, J.; Im, W. *Proteins* **2009**, *76*, 86–98.
- (48) Sarakatsannis, J. N.; Duan, Y. *Proteins* **2005**, *60*, 732–739.
- (49) Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graphics* **1996**, *14*, 33–38.