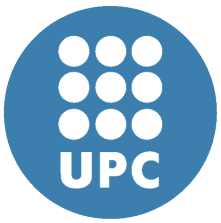
Exploring free-energy landscapes and microscopic interactions of selected small-molecules and proteins with cell membranes

Phd Program in Computational and Applied Physics



Huixia Lu

Supervisor: Jordi Martí Rabassa

Department of Physics
Universitat Politècnica de Catalunya

This dissertation is submitted for the degree of
Doctor of Physics

July 2020

Acknowledgements

Undertaking this Ph.D. has been a truly life-changing experience for me and it would not have been possible to do without the support that I received from many people I meet here in this beautiful city: Barcelona, especially my tutor Dr. Jordi Martí. I would like to first say a very big thank you to Jordi for the continuous support during my Ph.D. study and related research, for his patience, motivation, and immense knowledge. His guidance through many discussions we had helped me all the time of work and writing of this Thesis. Thank you for giving me the opportunity of getting started in the field of Biophysics that arises my huge amount of interest in the topic I did throughout the whole period time of my Ph.D. Without him this would not have been achievable. Jordi is the first teacher I met who shows me the true and pure tutoring spirit. Shifting from a pure chemist after finishing my master's study, it was Jordi who led me as a newborn baby into Biophysics, he taught me how to run simulations using different plugins, build biomembrane systems, and how to organize my research time more efficiently. It is Jordi who showed me what kind of researcher I shall be in the future.

I am grateful to all other staffs of the Department of Physics of Universitat Politècnica de Catalunya with special mention to the other researchers with whom I have shared hundreds of lunches, coffees and a few dinners: Grecia Guijarro, Juan Sanchez, Dr. Raúl Bombín, viktor Cikojević, Dr. Giulia De Rosi, and Dr. Gregory E. Astrakharchik. I will never forget your precious friendship and those happy moments we shared.

My sincere thanks also go to Ana, Cristina, etc. from the Secretaria office; Alfredo (who retired from UPC in the year 2019) and Agustí from the IT office. Cristina and Alfredo together cook the best paella in the world during every year's social lunch from the department. These people always give me extra energy during my Ph.D. study, they showed me unbelievable good intentions and kindness which make me always feel welcome during my stay far away from home. They showed me the amazing big hearts of the Spanish people, their good hearts, and positive spirit. Alfredo and Agustí offered me lots of help when using computing resources: INCA of the department of Physics.

I gratefully acknowledge the computing resources and assistance provided by Barcelona Supercomputing Center (BSC) and the Red Española de Supercomputación

(RES), with their help, I could conduct my work using supercomputers, such as CALENDULA (Activity Id: FI-2018-1-0007), CIBELES (Activity Id: FI-2018-2-0018), MARE NOSTRUM4 (Activity Id: FI-2018-3-0023, FI-2019-1-0004, and FI-2019-2-0004) and LA PALMA (Activity Id: FI-2019-3-0008). Many thanks to the financial support from China Scholarship Council (CSC) by a State Scholarship Fund (NOº 201604070059). Without their precious support it would not be possible to conduct this research.

I am indebted to my family in China who I love from the bottom of my heart. My father and mother are the typical Chinese parents who devote all they have to their children. With this unconditional love I had a happy childhood. Thanks to my little sister and little brother who I love a lot. They are always so helpful in numerous ways. Special thanks to my partner Pau Farré whose support seems to be endless and always taking care of me with his warm heart.

Thanks so much for this city, Barcelona. Thanks to the amazing weather here, beautiful beaches I went and all kind people I met here and never forget to mention, the mild and delicious Catalan food I ate. I am so grateful that I enjoyed so much, had so many happy moments with my colleagues, my dear tutor, and all my friends here.

Table of contents

List of figures **vii**

List of tables **xv**

1 Introduction **1**
1.1 General information on mammalian cell membranes 1
1.2 Small molecules studied in this Thesis 4
1.3 KRas-4B membrane protein 7
1.4 Objectives and outline 14

2 Methods **17**
2.1 CHARMM-GUI web-based platform 17
2.2 Force Fields 20
2.3 Molecular Dynamics Simulations 22
2.4 Well-tempered metadynamics simulations 26
2.5 Simulation packages used in this Thesis 29

3 Small molecules bound to di-palmytoil-phosphatidyl-choline bilayer membranes **31**
3.1 Preliminary study on the adsorption of tryptophan into DPPC bilayers 32
3.2 Effects of cholesterol on the binding of tryptophan to zwitterionic membranes 41
3.3 Binding free energies of small-molecules in phospholipid membranes . 56
3.4 Conclusions 60

4 Cellular absorption of small molecules: molecular dynamics and free energy landscapes of melatonin binding at phospholipid membranes **65**
4.1 Molecular dynamics of melatonin binding the interface of DMPC lipids 66
4.2 Free energy landscapes of melatonin binding at phospholipid membranes 81
4.3 Conclusions 98

5 Structural and energetic aspects on the anchoring of wild-type and oncogenic KRas-4B proteins at model cell membranes 101
5.1 Methods . . . 103
5.2 Results and discussion . . . 105
5.3 Well-tempered metadynamics simulations of the conformational transitions of KRas-4B . . . 119
5.4 Conclusions . . . 132

6 Conclusions and prospect 135
6.1 General conclusions . . . 135
6.2 Perspectives of future work . . . 140

Appendix A Related publications 143

References 145

List of figures

1.1 Simplified model of a bio-membrane. Water is attracted to the outsides (red) of the membrane but is prevented from going through the non-polar interior (yellow) layer. 2

1.2 General structures of five small molecules studied in this Thesis. . . . 5

1.3 A: Domain structures of KRas-4B along its sequence. As described in Ref. [115], in the sequence there are hydrophobic (black), polar/glycine (green), positively charged (blue), and negatively charged (red) residues. The same colors were used in the ribbon representation of its secondary structure, except for the hydrophobic residues (white). B: KRas-4B GDP-GTP cycling regulated by GAPs and GEFs. 9

1.4 PTMs steps of KRas-4B: prenylation, hydrolysis, carboxymethylation and decarboxymethylation. 11

3.1 Sketches of the backbone structures of DPPC, cholesterol, and TRP. Hydrogens bound to carbon and nitrogen are not shown. The highlighted sites of TRP ('H1', 'H2', 'N1', 'N2', 'C1', 'C2', 'O1', and 'O2') and of DPPC ('N', 'O2', and 'O8') will be referred in the text by the labels defined here. Due to the zwitterionic characteristics of TRP, its site 'H1' corresponds to any of the three hydrogens bound to 'N1', which share the positive charge. Sites 'O1' and 'O2' of TRP share the negative charge. Site 'C2' is close to the center of mass of the molecule. 33

3.2 Two snapshots of the aqueous DPPC membrane with an embedded TRP molecule. TRP inside DPPC (left) and inside water (right). Surrounding water is not shown for the sake of clarity. Atoms in TRP: Oxygen (red); hydrogen (white); carbon (cyan); nitrogen (blue). Sodium ions (yellow) and chloride ions (green). DPPC are depicted as lines. 34

3.3 Energy decomposition for the full system. 35

3.4 Temperature and pressure fluctuations for the full system. 36

3.5 Selected radial distribution functions for TRP (T) with water(W) and DPPC (charged sites 'N' and 'O2', see Fig. 3.1): OT-HW (bottom left), HT-OW (top left), OT-N^+ (bottom right) and HT-O^- (top right). . . 37

3.6 Potentials of mean force for water-TRP and DPPC-TRP sites. Labels as in Fig. 3.5. 39
3.7 Snapshot of a typical TRP-DPPC bond. Note that HB exits between sites of H1 of TRP and O8 of DPPC. 40
3.8 Snapshots of the aqueous DPPC and cholesterol membranes with an embedded TRP molecule (Top: 30% of cholesterol; bottom: 50% of cholesterol). Atoms in TRP: oxygen (red), hydrogen (white), carbon (cyan) and nitrogen (blue). DPPC and water molecules are shown as thin sticks. Cholesterol are depicted as thick sticks. Sodium ions (green) and chlorine ions (yellow). 43
3.9 Order parameter $|S_{CD}|$ for the acyl tails for three different percentages of cholesterol. 44
3.10 Area per lipid versus cholesterol contents of the system: 0%, 30% and 50%. Dashed lines indicate average values reported in Table 3.2. . . . 45
3.11 Radial distribution functions for TRP (T) with water(W) and cholesterol (sites 'O' and 'H' of the hydroxyl group): H1T-OW (top left), OT-HW (bottom left), H1T-OC (top right) and OT-HC (bottom right). 47
3.12 Radial distribution functions for TRP (T) with DPPC (charged sites 'H1', 'H2', 'O2' and 'O8', see Fig. 3.1): H1T-O2 (top left), H1T-O8 (bottom left), H2T-O2 (top right) and H2T-O8 (bottom right). 48
3.13 Distance distribution of selected sites in cholesterol-TRP bonding and DPPC-TRP bonding as a function of simulation time. Percentages of cholesterol are indicated on top of each graph. Dashed lines indicate the typical HB distance (1.8 Å). 'H1$_A$T' and 'H1$_B$T' are two of the three hydrogens bound to the site 'N1' of TRP, whereas 'O2' is one of two oxygens of TRP and 'OC' ('HC') is the hydroxyl's oxygen (hydrogen) of cholesterol. 50
3.14 Snapshots of typical DPPC-TRP-cholesterol bonds. Atoms in TRP colored as in Fig. 3.8. Binding sites have been highlighted in dashed orange lines. 51
3.15 Spectral densities of the full TRP molecule (left). Full arrows indicate the position of experimental peaks and the dashed arrow indicates a maximum of weak intensity. Decomposition of the spectrum for the 0% cholesterol case (right), indicates the relative contribution of several (most relevant) selected atoms. The inset reveals the contributions of high frequencies of the two relevant groups of hydrogen atoms ('H1' and 'H2'). 54
3.16 Sketches of the backbone structures of histidine, tyrosine, serotonin, and melatonin with highlighted sites which will be inferred in the following section. Hydrogens bound to carbon are not shown. 56

3.17 Potentials of mean force for the adsorption of small-molecules to water molecules. 58
3.18 Potentials of mean force for DPPC-small molecules. 61
3.19 Snapshots of typical small-molecule and DPPC bonds. Serotonin-DPPC (left, where the binding of a hydroxyl's hydrogen with 'O2' site is clearly seen) and tyrosine-DPPC (right, where the binding of hydrogen 'HC' to the 'O2' site and of the hydrogen from the hydroxyl group with 'O8' site are clearly seen). 62

4.1 Backbone structures. Sketches of molecular structures of MEL, DMPC, and cholesterol. Part of Hydrogen-Carbon bonds not shown. The highlighted sites of MEL ('C3', 'C4', 'H15', 'H16', 'N1', 'N2', 'O1', and 'O2') and of DMPC ('O1', 'O2', 'O6', and 'O8') will be referred in the text by the same labels. 67
4.2 Order parameter. $|S_{CD}|$ for the (sn1, sn2) acyl tails of DMPC at three different cholesterol concentrations. 68
4.3 Physical characteristics of the membranes. Area per lipid of systems with different cholesterol contents: 0% (black line), 30% (blue line), 50% (red line) as a function of simulation time. The green dashed lines indicate the average values for the last 150 ns at each concentration. . 69
4.4 Z-axis position of MEL. Penetration of MEL inside DMPC bilayer (green symbols indicate the position of the center of mass of MEL whereas red symbols stand for the position of phosphorus atoms atoms in each layer) at 0%, 30%, and 50% cholesterol concentrations. 71
4.5 Radial distribution functions. Selected radial distribution functions for hydrogens of MEL ('H15' and 'H16') with oxygens of water ('OW'), DMPC ('O2' (representing 'O1 & O2') and 'O6' (representing 'O6 & O8')) and cholesterol, belonging to hydroxyl group ('O Chol'). 72
4.6 Dihedral angles. Angles (a) θ, (b) Ψ and (c) ϕ defined for MEL. Top figures correspond to "folded" and bottom figures to "extended" configurations for each dihedral. The two snapshots (top, bottom) have been obtained from equilibrated configurations 75
4.7 Angular distributions. Distribution of selected dihedral angles in MEL as a function of simulation time, where labels (a), (b), and (c) have the same meaning as in Fig. 4.6. Percentages of cholesterol are: 0% (black circles), 30% (red squares), and 50% (green triangles). Dashed lines indicate average values and are a guide for the eye. 76

4.8 Spectral densities. $S_{MEL}(\omega)$ of MEL (left). Positions of main experimental peaks are indicated by full arrows (low-mid frequency range) and locations of a broad group of frequencies (high frequency range) have been indicated with initial and final vertical arrows linked by a horizontal double arrow. For the cholesterol-free system (right), atomic spectra $S_i(\omega)$ indicate the relative contribution of selected atoms to $S_{MEL}(\omega)$. Inset: high frequency region for most relevant atoms of MEL ('C3', 'N1', 'N2', 'H15', and 'H16'). 78
4.9 Density profiles of lipid backbone after 200 ns of MD equilibration (full lines: 0% cholesterol, dotted lines: 30% cholesterol, and dashed lines: 50% cholesterol) and water (circles: 0% cholesterol, squares: 30% cholesterol, and triangles: 50% cholesterol) 83
4.10 Area per lipid of systems as a function of simulation time in three cases: 0% cholesterol (black), 30% (red), and 50% (blue). Dashed green lines indicate the average values, which are shown in Table 4.6, of the last 150 ns at each concentration. 85
4.11 Fluctuations of the CV values as a function of time in different states: (a) 0% cholesterol (black), (b) 30% (red), and (c) 50% (blue). 85
4.12 Well-tempered Metadynamics hills height as a function of time in different states: (a) 0% cholesterol, (b) 30%, and (c) 50%. 86
4.13 Time cumulative free energy profiles at the cholesterol-free system. Bottom: CV1, top: CV2. 87
4.14 Time cumulative free energy profiles at the 30% cholesterol system. Bottom: CV1, top: CV2. 88
4.15 Time cumulative free energy profiles at the 50% cholesterol system. Bottom: CV1, top: CV2. 89
4.16 2D free energy landscapes $F(\Psi, z)$ (in kJ/mol) in the cholesterol-free case. Four stable state basins (A, B, C, and D) are indicated. TS indicates a (local) transition state between basins A and B. The minimum free energy path is shown in dark blue, indicating all the computed points of the path, as listed in Table 4.7. Snapshots 'A' and 'B' correspond to selected basins A and B on the free energy hypersurface, respectively. In between basins A and B we can locate a transition state corresponding to the conformational angular change of MEL between extended (basin A) and folded (basin B) configurations. DMPC (white), sodium ions (yellow), chlorine (dark orange), water (blue) and for MEL: carbon (cyan), nitrogen (dark blue), oxygen (red), hydrogen (white). 91

4.17 2D free energy landscapes $F(\Psi, z)$ (in kJ/mol) in the case of 30% cholesterol. Minimum free energy path pictured as in Fig. 4.16. Snapshot 'TS' corresponds to the global TS on the free energy surface. Colors as in Fig. 4.16, with cholesterol chains depicted in green. 92
4.18 2D free energy landscapes $F(\Psi, z)$ (in kJ/mol) in the case of 50% cholesterol. 'TS' corresponds to the global TS on the FEL. Snapshot 'D' corresponds to selected basin D. Colors as in Fig. 4.17. 92
4.19 Snapshots of representative bound states for MEL at the interface of the membrane. Colors of MEL as in Fig. 4.6. Colors for DMPC/cholesterol: carbon (cyan), nitrogen (dark blue), oxygen (red), hydrogen (white). Hydrogen bonds indicated as green solid lines. Left: case of 0% cholesterol concentration, with MEL bound to two DMPC molecules; Right: case of 30% cholesterol concentration, with MEL bound to one DMPC plus one cholesterol molecule. 95
4.20 Integrated free energy F(z) for 0%, 30% and 50% cholesterol. 97
4.21 Integrated free energy F(Ψ) for 0%, 30% and 50% cholesterol. 97

5.1 Three sequences of different KRas-4B structures. Mutated sites are in red color. Here C_f denotes the farnesylated Cys-185 and S_p represents the phosphorylated Ser-181 (see Fig. 5.3). Methyl group of KRas-4B-FMe is in blue color. 102
5.2 Graphical expressions of two CVs. the CD of KRas-4B (cyan), the HVR (green), GTP (red) and FAR (purple) are shown here. 104
5.3 Sketches of the backbone structures of cholesterol, GTP, DOPC, DOPS, and Cys-185 sites are shown here. Parts of Hydrogen-Carbon bonds aren't shown. The highlighted and renamed sites in red color will be referred to in the text. 105
5.4 Averaged S_{CD} for sn-1, sn-2 chains of DOPC in all systems. 106
5.5 Area per lipid of systems with different KRas-4B sequences as a function of simulation time. The blue dashed line indicates the average value for each system of the second half of the total 1000 ns production runs. . 107
5.6 Two mutated sites in structures of the oncogenic KRas-4B-Far/KRas-4B-FMe and the active amino acids of KRas-4B with their side chains framed. All highlighted and colored in red sites will be inferred in the following text. 109

5.7 Selected active sites of the CD (cyan) and the HVR (green) of the KRas-4B structure are depicted in red color. Thereinto, H_{Arg-CD} donates the cationic hydrogen atoms from the guanidinium ($RNHC(NH_2)_2^+$) of arginine of the CD that share the same atom type, H_{Lys-CD} and H_{HVR} represent the cationic ammonium (RNH_3^+) from lysine of the CD and of the HVR. O_{Glu-CD} and O_{Asp-CD} represent oxygen atoms of the anionic carboxylate ($RCOO^-$) of glutamate amino acid of the CD and Aspartate amino acid of the CD, respectively. O_{HVR} represents all oxygen atoms of the anionic carboxylate ($RCOO^-$) of the HVR. O_{Gln-CD} and O_{Gln-CD} donate active oxygen and hydrogen atoms of the side chain of glutamine amino acid of the CD. Moreover, O_{phos} and H_{phos} donate two active oxygen atoms and one hydrogen atom of the phosphate group from PHOS. All highlighted and colored in red sites will be inferred in the following text. 110

5.8 Selected RDFs for active atoms of GTP with selected sites of the CD and active oxygen atoms of head groups of lipids ('O2L'). 'H_{gtp}' and 'O_{gtp}' represent hydrogen and oxygen atoms from phosphorus group of GTP. 111

5.9 Selected RDFs for selected sites of the CD with head groups of lipids. In all panels of this figure, as indicated in panel "A", we use three colors to represent three systems in order to avoid the duplication and repetition, same as Fig. 5.10. Thereinto, H_{CD} from panel "A" refers to the averaged value of H_{Arg-CD} and H_{Lys-CD}. 112

5.10 Selected RDFs for selected sites of the HVR with atoms of lipids and the CD. Here O_{pt-CD} and H_{pt-HVR} represent oxygen and hydrogen atoms of the peptide bonds from the CD and the HVR. And O_{CD} stands for oxygen atoms of O_{Glu-CD} and O_{Asp-CD} from the CD. 113

5.11 Selected RDFs related to selected sites of FAR. 114

5.12 Distance distribution of selected sites of FAR and PHOS as a function of simulation time. Blue dashed lines indicate the typical HB distance (1.8 Å). O1P and O2P represent the two oxygen atoms belong to the phosphorus group of PHOS (see Fig. 5.6). And the corresponding snapshot is shown in Fig. 5.13. 114

5.13 Snapshot of long-lasting salt-bridges indicated as black lines of dashes between FAR and PHOS. The atoms forming the FAR and PHOS moiety are: carbon (cyan), oxygen (red), hydrogen (white), nitrogen (blue) and phosphorus (brown). 115

5.14 Localisation of the KRas-4B domains and GTP with respect to the center of the membrane along with the membrane normal as a function of simulation time. Geometric centers of the CD, the HVR, FAR, GTP, and phosphorus atoms of DOPC lipids from both leaflets are indicated as triangle down in orange, star in green, triangle up in violet and circle in turquoise, respectively. The four favored configurations described here are reported in Fig. 5.15. 116
5.15 Four preferential configurations of three KRas-4B-membrane systems. 117
5.16 Well-tempered Metadynamics hills height as a function of time in two cases. 120
5.17 Time evolution of two CVs as a function of time in different states for two systems. 121
5.18 Time cumulative free energy profiles. Lower panels correspond to the wild-type system and upper ones to the oncogenic case. Labels such as 'A, B, ...X' correspond to the relevant basins along with the CV within the free energy profile and will be described below in Fig. 5.19. Here we would like to point out that basins with the same labels in section 5.3.1 are different from section 5.3.2. 122
5.19 Convergence of well-tempered metadynamics based on the free energy barriers between two chosen basins. The selected basins ('A',...,'Y') have been defined in Fig. 5.18. 122
5.20 2D free energy landscapes F(gtp, far) (kJ/mol) for the oncogenic KRas-4B-Far system. Several stable configurations (A, B, C, D, E, and F) are indicated. All MFEPs between two selected basins are depicted in purple and the coordinates of the transition path have been reported in Table 5.5. 125
5.21 2D free energy landscapes F(gtp, far) (kJ/mol) of the wild-type KRas-4B system. Seven meta-stable configurations are indicated here. MFEPs between different two basins are depicted in purple and the relevant coordinates have been reported in Table 5.5. 128
5.22 1D integrated free energy profiles of GTP and FAR along with the membrane normal direction. Oncogenic case (top) and wild-type case (bottom). Basins marked with the same letters as in Fig. 5.18. Minima in four panels are set equal to zero. The minimum in each of the four figures has been set to zero. 130

6.1 Structures of FGTI-2734 and DBD shown here. In the two configurations (Config. 1 and 2), DBD is depicted in orange, whereas the remaining moieties are depicted with the same color prescriptions as in Fig. 5.2. 141

List of tables

1.1 Table of amino acids and their abbreviations. 8

3.1 Free energy barriers for the binding of TRP to water and DPPC. ΔF is the size (in k_BT) of the free energy barrier and $\left\langle r_{1(2)} \right\rangle$ are the position of the first (second) minima of PMF. 40
3.2 Mechanical properties of the membrane: area per lipid (A) and thickness (Δz). Estimated errors are in parenthesis. 45
3.3 Self-diffusion coefficients D (in cm^2/s) of water and of TRP as a function of the cholesterol's percentage of the membrane. 51
3.4 Area per lipid (A) of the membrane systems. Estimated errors are in parenthesis. 58
3.5 Free energy barriers ΔF (in kcal/mol) for the binding of small-molecules to water and to DPPC. 59

4.1 Area per lipid (A) and thickness (Δz) of the membrane for all three cholesterol concentrations studied in this work. Estimated errors in parenthesis. 69
4.2 Gibbs free energy differences PMF (in kcal/mol) for the binding of MEL (sites 'H15' and 'H16') to selected atomic sites. Since 'O1' and 'O2' sites in DMPC share the negative charge, their contributions have been averaged; the same situation for 'O6' and 'O8'. 74
4.3 Ratio of the two angular configurations for MEL in three systems with different cholesterol concentration. "Ext." stands for "extended configuration of MEL" and "Fol." stands for "folded configuration of MEL" in all cases. 76
4.4 D (in cm^2/s) of MEL and of water in three systems with different cholesterol percentages. Estimated errors in parenthesis. 77
4.5 Simulation parameters 83
4.6 Area per lipid (A) and thickness (Δz) of the membrane. Estimated errors in parenthesis. 84
4.7 Coordinates of segments forming each path. 90

5.1 Metadynamics simulation parameters. 104
5.2 Area per lipid A and thickness Δz of the anionic membrane for all the KRas-4B systems studied in this Chapter. Estimated errors in parenthesis. 106
5.3 Averaged values of z-locations of FAR, HVR, CD, GTP during the last 500 ns simulations, among them, for the wt. KRas-4B-Far system, values of location of FAR from configuration 1 and 2 have been provided. Estimated errors in parenthesis. 118
5.4 Minima on the FEL of the oncogenic KRas-4B-Far (coordinates given in nm, free energy in kJ/mol). Here each letter refers to one meta-stable state, with A standing for the global state when FAR is anchoring deeply into the membrane and GTP locates near the interface. Here locations of minima of FEL of Fig. 5.20 were obtained by using metadynminer package. 124
5.5 Coordinates of segments forming selected paths for the oncogenic and wild-type KRas-4B-Far systems. Along with each path 8 locations of fespot are displayed. 127
5.6 Minima on the FEL of the wild-type KRas-4B-Far. Here each letter refers to one meta-stable state, among which A stands for the global state which is set to zero. Here locations of minima of FEL of Fig. 5.21 were obtained by using metadynminer package. 129

Chapter 1

Introduction

Computer simulation is a powerful approach to study the properties of biological cell membrane systems. Bio-molecular simulations initially focused on proteins, have seen a long development over two decades[1]. Groundbreaking simulations of bilayer undulations[2] and bilayer self-assembly[3] showed that collective events can be studied in atomic detail. Plasma membranes (PM) are ubiquitous as limiting structures of cells, separating their contents from external environments. Many cellular functions, including the uptake and conversion of nutrients, synthesis of new molecules, signal transduction, and regulation of metabolic sequences, take place on the PM. In this Thesis, our main focus has been on membrane structure and dynamics, its functions, the bio-molecular organization, and the study of the characteristics of species such as small-molecules and proteins when bound to the cell membranes. Before introducing our findings and the particular issues that will be the main body of this Thesis, a brief historical overview of several systems that we have studied is very appropriate.

1.1 General information on mammalian cell membranes

Mammalian cell membranes are complex structures protecting cellular contents, which include a wide variety of organelles (mitochondria, nucleus, lysosomes, vacuoles, Golgi apparatus, centrioles, ribosomes, etc.) surrounded by the cytoplasm, see Fig. 1.1.

The principal components of human cellular membranes are lipids, cholesterol ($C_{27}H_{46}O$), and proteins, all of them surrounded by aqueous solutions including water and salts. Water can provide the driving force for the assembly and stability of many cell components which is fundamental in many biological processes. The water-membrane interface has a structural effect on the water bulk as far as $\simeq 2.5$ nm from the membrane according to Franzese et al.[4]. Water near biomembranes can be separated in *bound* and *unbound* hydration water, with the two classes of water diffusing one order of magnitude slower than bulk water[5]. Phospholipid membranes provide the framework to biological membranes, to which other molecules (such as proteins or carbohydrates)

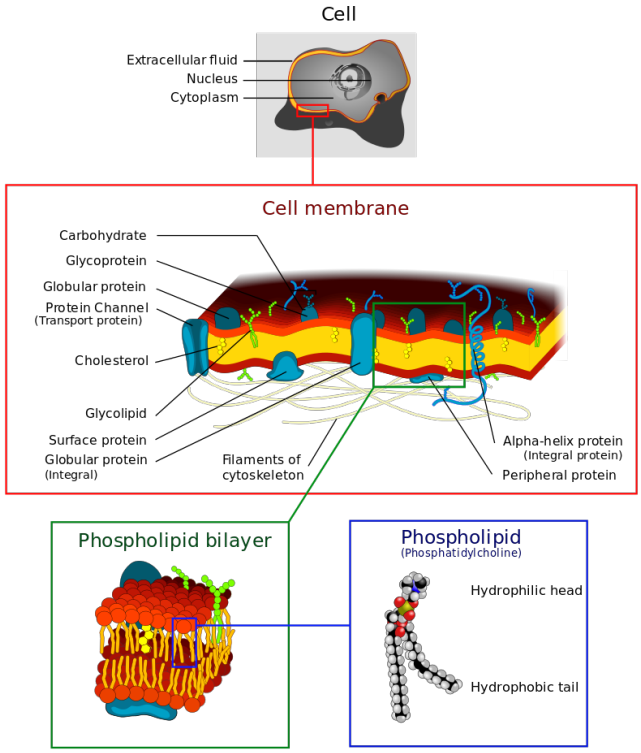


Fig. 1.1 Simplified model of a bio-membrane. Water is attracted to the outsides (red) of the membrane but is prevented from going through the non-polar interior (yellow) layer.

attach. They consist of two leaflets of amphiphilic lipids which self-assemble due to the hydrophobic effect[6]. Such lipids are molecules with a hydrophilic head and one or two lipophilic (and hydrophobic) tails. When placed inside aqueous ambient they can form three main different classes of structures: micelles, liposomes, and bilayers, with the latest hiding the tails from water[7]. The main function of phospholipids as building blocks of cell membranes was already discovered in 1925[8].

Phospholipids are widely distributed in animals and plants. Variation in the head-groups leads to different kinds, such as phosphatidylcholine (PC), phosphatidylserine (PS), phosphatidylinositol (PI), phosphatidylethanolamine (PE), phosphatidylglycerol (PG), cardiolipin (CL), and phosphatidic acid (PA), that share some general properties, with each group keeping their unique physical functions. For instance, PC is an important substrate of synthesis of the neurotransmitter acetylcholine and has the function of nourishing the brain and improving intelligence. As an example of a prototype membrane, dimyristoylphosphatidylcholine (DMPC, $C_{36}H_{72}NO_8P$) belongs to PCs family, basic components of lecithin, a substance forming egg yolk and soy. DMPC is a phospholipid species employed in many experimental works since its gel to liquid phase transition happens at a relatively low temperature (around 295.5 K)[9]. A large number of simulations have already been published on DMPC membranes[10–17], often

considering the influence of cholesterol in aqueous ionic environments[18–25] as well as numerous reports based on experimental techniques (neutron scattering, X-ray, infrared reflection absorption spectroscopy, NMR, etc.) have been published throughout the last decades (see for instance Refs. [26–32]). Di-palmitoyl-phosphatidyl-choline (DPPC, $C_{40}H_{80}NO_8P$) is a major constituent (about 40%) of pulmonary lungs[33]. When it comes to the molecular structure, DPPC is observed to have the same headgroup as DMPC but with two C16 hydrocarbon tails, and DMPC has two C14 hydrocarbon tails. In addition, human lungs are coated with a lattice-like structure formed by protein and lipid mixture called lung surfactant, preventing the lungs from collapsing and protecting us from bacterial and viral infections. Extensive simulations and experimental studies have already been performed on DPPC monolayer/bilayers, often including the influence of the cholesterol in water environments[34].

1,2-dioleoyl-sn-glycero-3-phosphocholine (DOPC, $C_{44}H_{84}NO_8P$) belongs to the unsaturated PC family which might be important in the formation of a lipid reservoir. Referring to DMPC, DOPC is a more typical constituent of real biological membranes with a transition temperature (T_m) of 253.15 K. T_m of lipids is defined as the temperature required to induce a change in the lipid physical state from the ordered gel phase, where the hydrocarbon chains are fully extended and closely packed, to the disordered liquid crystalline phase, where the hydrocarbon chains are randomly oriented and fluid. In PM, the lipid composition between the two leaflets is highly asymmetric and changes dynamically to trigger environmental responses. For instance, cells that undergo apoptosis increase the concentration of anionic PS in the outer leaflet of the PM to signal phagocytes to approach and digest them, whereas, in healthy cells, PS lipids are only present in the inner leaflet[35]. In between, 1,2-dioleoyl-sn-glycero-3-phospho-L-serine (DOPS, $C_{42}H_{77}NO_{10}P$) with a T_m of 262.15 K is the most common anionic lipid in the PM of mammalian cells which is preferentially targeted by the PM intracellular surface protein KRas-4B[36] for signal transduction.

Recently people used to conduct computational and experimental researches on component of anionic DOPS and zwitterionic DOPC bilayers and liposomes[37–40]. Determined by the method called nanoplasmonic sensing, the T_ms of DMPC and DPPC have been reported to be 296.65 K and 314.15 K, respectively, which are close to their theoretical T_m values, *i.e.*, 297.15 K for DMPC and 314.55 K for DPPC[9]. T_m can be affected by hydrocarbon length, charge, unsaturation, and headgroup species in the lipid's structure and has an important effect on the structure of the bio-membrane. T_m of DPPC is much higher than T_m of DMPC, the reason can be longer hydrocarbon tails. Increasing the hydrocarbon length makes van der Waals interactions become stronger which requires more energy to disrupt the ordered packing, thus the phase transition temperature increases. Likewise, introducing a *cis* double bond into the acyl group and charged headgroups which requires the much lower temperature to

induce an ordered packing arrangement[9], leading to lower T_ms for DOPC and DOPS lipids[41].

The addition of sterols to a phospholipid bilayer diminishes the differences between the fluid and gel phases. Cholesterol is a sterol, a type of lipid, playing a central role in maintaining the structure of the membrane and regulating their functions[42, 43]. Being an amphiphilic molecule, cholesterol inserts into the membrane through its hydroxyl group of cholesterol oriented toward the aqueous inside and aliphatic chain associated parallel toward the acyl chains in the center of the bilayers, and it increases the separation between choline head groups and eliminates the normal electrostatic and hydrogen-bonding interactions. The phospholipids are arranged in such a means that the hydrophilic head is exposed outside and the lipophilic tails are oriented inside[44]. It induces the membrane to adopt a liquid-ordered phase with the positional disorder and high lateral mobility[7]. Mammalian cells regulate the concentration of cholesterol in their PM (10 $\sim$ 40% and endoplasmic reticulum ($\sim$ 5%)[45, 46]. The fluidity of the membrane is mainly influenced by the amount of cholesterol, in such a way that membranes with high cholesterol contents are stiffer than those with low amounts but keeping the appropriate fluidity for allowing normal membrane functions. Model membranes with the constitution of cholesterol and different lipids have been extensively studied in the literature from the experimental and also computational points of view[47–53].

1.2 Small molecules studied in this Thesis

Besides the composition of the membrane, the role of proteins, drugs, and their interactions with cell membrane structures are undoubtedly a relevant field of research. Amino acids are organic compounds containing amine (-NH_2) and carboxyl (-COOH) functional groups, along with a side chain (R group) specific to each amino acid. It is well known that amino acids can either be *essential*, i.e. indispensable or *non-essential*. An essential amino acid cannot be synthesized *de novo* by the organism and should be necessarily supplied by the diet. The nine amino acids humans cannot synthesize are: phenylalanine, valine, threonine, tryptophan (TRP, $C_{11}H_{12}N_2O_2$), methionine, leucine, isoleucine, lysine, and histidine (HIS, $C_6H_9N_3O_2$), and structures of some of them are shown in Fig. 1.2.

We have devoted our interest in the amino acid TRP[54] which is commonly used to treat insomnia and sleep orders like sleep apnea[55, 56]. TRP can act as a building block in protein biosynthesis, while proteins are fundamentals required to sustain life. TRP is a precursor of the neurotransmitter serotonin (SER, $C_{10}H_{12}N_2O$)[57] which, in turn, is also the precursor of the hormone melatonin (MEL, $C_{13}H_{16}N_2O_2$)[58]. Since the human body cannot synthesize TRP, needed to prevent diseases and death, it has to be acquired from the diet. Besides, it helps in the regulation of human sleep.

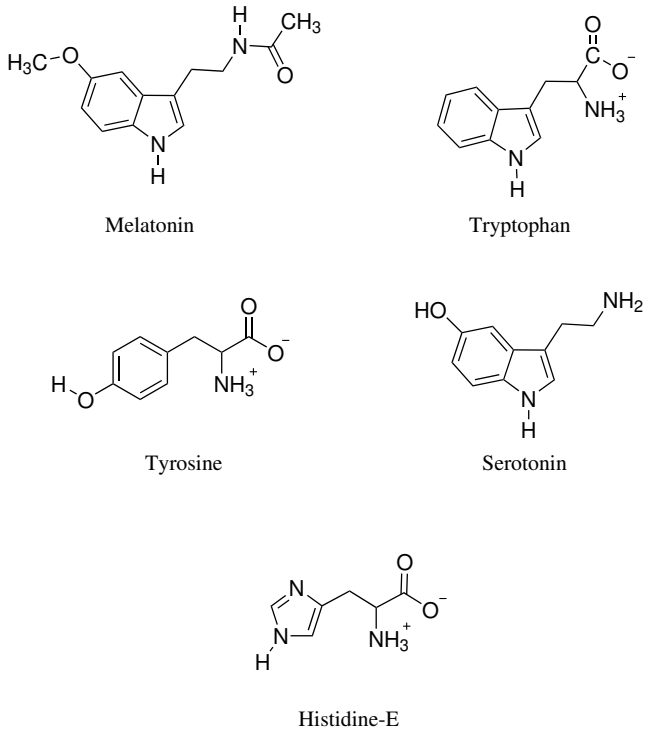


Fig. 1.2 General structures of five small molecules studied in this Thesis.

Recently, some studies indicated promising antibacterial properties of TRP when synthesized at polyionic membranes[59] and its major role in the preparation of the nanoparticles which are able to control membrane disruption[60]. TRP trimers and tetramers were reported to inhibit Dengue and Zika virus replication by interfering with viral attachment processes[61]. TRP also has a vital role in maintaining the structure and function of gramicidin channels at the surface of a cell membrane and its side chain has the greatest hydrophobic surface area among all the aromatic amino acids[62].

SER is a well-characterized neurotransmitter and hormone that is mainly produced by enterochromaffin cells in the gut and also by neurons of the brain stem[63]. It is primarily found in the gastrointestinal tract, blood platelets, and the central nervous system of animals, including humans. It is thought to be a contributor to the regulation of human mood and happiness. SER has been reported to play a role in promoting the development of pancreatic ductal adenocarcinoma in mice by reinforcing the so-called Warburg effect[64]. SER, in turn, can be converted to MEL *in vivo*, which may help humans to the regulation of biological rhythms, to induce sleep, to work as a strong antioxidant, and also to contribute to the protection of the organism from carcinogenesis and neurodegenerative disorders[65]. In addition to its physiological role, growing evidence suggests SER also regulates the connectivity of the brain by modulating developmental cellular migration and cytoarchitecture[66].

HIS is an alpha-amino-acid that is also used in the biosynthesis of proteins. Initially thought essential only for infants, longer-term studies have shown it is essential for adults also. Copper is an essential trace element required by all living organisms. Since the discovery in 1966 of copper(II)–l-histidine species in human blood[67], extensive research has proven HIS plays a very important role in copper transport and has been used in the treatment of infantile hypertrophic cardioencephalomyopathy[68].

Differently, tyrosine (TYR, $C_9H_{11}NO_3$) is a non-essential aromatic amino acid produced by hydroxylation of phenylalanine and acts as a biochemical precursor of dopamine and norepinephrine. It is responsible for the synthesis of melanin, thyroxin, adrenaline, and the causative agent for genetic, hormonal, and neurological disorders. It has been revealed that a high level of TYR in blood tissue can be used as a weathervane to foresee and analyzes the level of certain metabolic disorders, such as obesity, insulin resistance, and liver cancer[69].

MEL is a natural hormone secreted by the pineal gland well known to regulate biological rhythms[70], to induce sleep[71], and that can also contribute to protecting the organism from Alzheimer disease[65]. MEL is reported to induce/promote complex antioxidative and DNA repair systems which make it a very good candidate for curing several dermatoses associated with substantial oxidative damage. It helps for preventing skin cancer, skin photo, and radioprotection and also works as an inducer of repair mechanisms of human skin recovery from environmental damage[72, 73]. In recent years, the community of biologists, dermatologists, and physicists have published plenty of works on MEL and studied how it can affect the human body[74–77]. For instance, MEL was found to have a significant effect on reducing cholesterol absorption and causing greatly decreases in total cholesterol in membrane bilayers and concentrations of cholesterol in the liver[78]. MEL is not only important for humans but also for plants and animals. In particular, MEL works as a multifunctional signaling molecule that regulates broad aspects of responses to environmental changes[79]. It has been observed that MEL is able to cross most physiological barriers, such as the blood-brain barrier[80, 81]. So, MEL may help to control brain function[82], and it has also interesting immunotherapeutic potential in both viral and bacterial infections[83]. It is also able to contribute to the protection of the organism from carcinogenesis and neurodegenerative disorders[65, 84]. A number of works have been focused on the structure and interactions of MEL with zwitterionic membranes[85, 86], and several experiments and simulations suggest that small solutes such as TRP and MEL are bound to the phosphate and carbonyl regions of phospholipid species[58, 87–89]. Other studies based on the binding of MEL at ion channels in lymphocytes indicated an active role of MEL in reversible blocking of the channel and its possible activity as an immunological agent[90]. While bound to membranes in aqueous solution, strong hydrogen-bonds between the small molecule and the phospholipids belonging to the membrane have been located in a variety of possible configurations[91]. Permeation of

MEL across the cell membrane has been also a subject of study and debate. Recent studies indicated that cellular permeation rates in the pineal gland are of the order of 1.7 μm/s[92] and that they can occur by pure diffusion. Conversely, some studies found that active processes are required for the entrance of MEL inside cancer cells[93]. f In particular, glucose transporters (GLUT family) may play a central role in MEL uptake and even be able to inhibit tumor growth[94]. The safety of MEL in humans has been questioned. Andersen et al.[95] addressed that, in general, animal and human studies documented that short-term use of MEL is safe, even in extreme doses. But mild adverse effects, such as dizziness, headache, nausea, and sleepiness have been reported. Also due to a lack of human studies, pregnant and breast-feeding women should not take exogenous MEL. Plenty of experimental and computational work on mixtures of cholesterol and MEL at phosphatidylcholine membranes has been published to analyze the joint effects of the two species (see Refs. [96, 97]).

Since all these molecules are not chemically bound between them, not only their local (vibrational) motions can be analyzed but also their long-range correlated motions (diffusion) can be measured and studied with all details through computer simulations. There exist a big variety of experimental techniques useful to explore membrane organization and molecular interactions of small probes within the membrane, such as NMR, neutron diffraction, X-ray scattering or IR, Raman, and fluorescence spectroscopy[54, 98–100]. Among the latest techniques, fluorescence-lifetime microscopy[101] can be combined with spectral information to report basic information of aspects such as metabolic profiles, photophysics, or dipolar relaxations[102]. Our efforts in the study of phospholipid membranes and their interaction with the environment can shed light on the basic biological functions of the membrane. Also, the knowledge of the mechanisms responsible for the exchange of small peptides and drugs inside different cell membranes is of greatest scientific interest.

1.3 KRas-4B membrane protein

PM systems are particularly interesting when they are associated with proteins. Proteins are usually big molecules that are difficult to name by the list of amino acids in their sequences. Amino acids can be recognized using the three-letter code or one-letter code which are developed to describe protein sequences easily. The listing of amino acids and the corresponding three/single-letter code is shown in Table 1.1.

Recent studies have shown that the role of proteins and their interactions with components of PM is extremely important to understand the mechanisms of protein anchoring into the membrane that can lead to oncogenesis[103]. Ras proteins, a kind of the membrane surface protein, are important signaling hubs frequently dysregulated in cancer and in a group of developmental disorders called Rasopathies. GTPases are a large family of hydrolase enzymes that bind to the nucleotide guanosine triphosphate

Table 1.1 Table of amino acids and their abbreviations.

Full name	Abbreviations (three-letter code)	Abbreviations (one-letter code)
Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartate	Asp	D
Cysteine	Cys	C
Glutamate	Glu	E
Glutamine	Gln	Q
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V

(GTP) and hydrolyze it to guanosine diphosphate (GDP)[104]. Ras proteins are small molecular weight GTPases and function as GDP/GTP-regulated molecular switches controlling pathways involved in critical cellular functions like cell proliferation, signaling, cell growth, and anti-apoptosis pathways[105, 106]. GDP/GTP cycling is controlled by two main classes of regulatory proteins. Guanine-nucleotide-exchange factors (GEFs) promote the formation of the active, GTP-bound form, while GTPase-activating proteins (GAPs) inactivate Ras by enhancing the intrinsic GTPase activity to promote the formation of the inactive GDP-bound form[107–109]. Once activated, GTP-bound Ras can interact with upstream regulators and activate several downstream effectors[110]. The three Ras genes give rise to three base protein sequences: KRas, HRas, and NRas. Over 30% of cancers are driven by mutant Ras proteins, thereinto, one method called Catalog of Somatic Mutations in Cancer (COSMIC)[111] confirms that HRas (3%) is the least frequently mutated Ras isoforms in human cancers, where KRas (86%) is the predominantly mutated isoforms followed by NRas (11%)[112].

KRas can be found as two splice variants designated KRas-4A and KRas-4B. There is significant KRas-4A expression in some tissues and it contains a palmitoylation site at its Cys-180 that has been shown to assist membrane localization. KRas-4A is only a small fraction ($< 10\%$) of total KRas relative to KRas-4B isoforms and they differ only in their flexible C-terminal membrane-targeting regions[113], no functional distinctions among the splice variants have so far been established. KRas-4A and KRas-4B both

have polybasic sequences that facilitate membrane-association in acidic membrane regions[114], however, for KRas-4A it is covalently modified by a single palmitic acid. The catalytic domain (CD), which contains the catalytic lobe (residues 1-86) and the allosteric lobe (residues 87-166), is highly homologous and the structure is shared and identical for both KRas isoforms. KRas-4B is distinguished from KRas-4A isoform in the residue 181 that serves as a phosphorylation site within its hypervariable region (HVR, residues 167-185). Fig. 1.3 reported here is adopted from Ref. [115] for a sketch of the structure and sequence of KRas-4B. The HVR of KRas-4B contains multiple amino-acid lysines that act as an electrostatic farnesylated switch which guarantees KRas-4B's association with the negatively charged phospholipids in the inner PM leaflet. The inner PM leaflet is often enriched with anionic phospholipids, such as phosphatidic acids (PA), phosphatidylglycerol (PG), phosphatidylinositol phosphates (PIPs, also known as phosphoinositides), and phosphatidylserine (PS). It has been reported that the KRas-4B activation level in diseased cells is linked to PS contents[116]. Anionic lipids could influence the membrane potential which in turn regulates the orientation, location, and signaling ability of KRas-4B[117, 118].

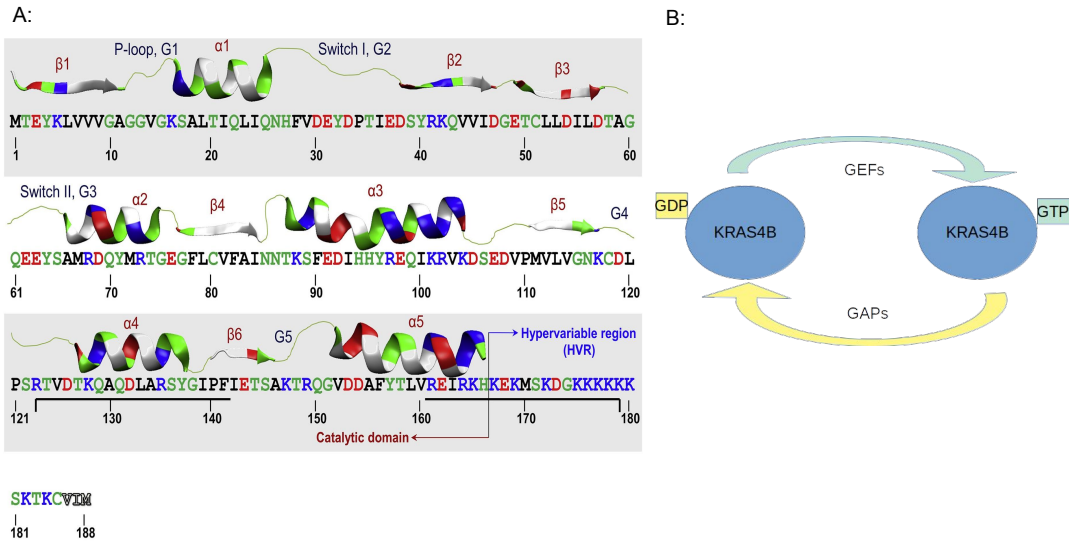


Fig. 1.3 A: Domain structures of KRas-4B along its sequence. As described in Ref. [115], in the sequence there are hydrophobic (black), polar/glycine (green), positively charged (blue), and negatively charged (red) residues. The same colors were used in the ribbon representation of its secondary structure, except for the hydrophobic residues (white). B: KRas-4B GDP-GTP cycling regulated by GAPs and GEFs.

Contributing to its focus on cancer studies, KRas-4B appears to be particularly prone to mutations and there exists an intimate link to residue substitutions[119–123]. In KRas-4B, G12 mutations (89%) are predominant in human cancers, followed by G13 (9%) and Q61 (1%) mutations. Moreover, the G12D mutation is arguably the most

prevalent mutation among the three frequent G12C (14%), G12D (36%), and G12V (23%) mutations and G13D (7%) and Q61H (0.6%) mutations are also observed[124]. In addition, KRas-4B mutations could also occur at residues Glu-63, Lys-117, Asp-119, and Ala-146 but with fewer frequency[125]. Structural studies have suggested that mutation at these positions often enable perturbation of GTPase activity of KRas-4B. Furthermore, the mutation within KRas-4B might exhibit significant differences between cancer types. For instance, G12C mutation hampers GAP binding and then controls the GTP hydrolysis to some extent[126]. In one study KRas-4B G12D mutation is revealed to promote colon cancer, supporting the ability of KRas-4B to initiate the formation of colon cancer[127]. Moreover, a relatively high express KRas-4B G13 mutation in combination with high expression of the metastasis-associated in colon cancer 1 (MACC1) indicates the poor survival for colorectal adenocarcinoma (CRD) patients[128]. Pancreatic ductal adenocarcinoma (PDAC) is dominated by KRas-4B G12D mutation followed by G12V, whereas Q61 and G13 mutations in this disease are rare to be found. In contrast, the G12C mutation is a hallmark of exposure to tobacco smoke and, accordingly, is the most common mutation in lung adenocarcinoma (LAC)[129]. KRas-4B G12C together with G12V mutations in LAC is able to preferentially activate the RalGDS pathway, and G12D mutation prefers the Raf/mitogen-activated protein kinase (MAPK) and PI3K pathways[130]. Moreover, KRas-4B has been observed to be able to be acetylated at Lys-104 which is proposed to impair GEF-mediated activation and consequently, to suppresses the oncogenic activity of mutationally activated KRas-4B. The observation indicates that acetylation at Lys-104 is a negative regulatory modification[131]. Even though considerable efforts on KRas-4B mutations have been invested in characterizing these distributions, their impact on the cell is still poorly understood.

All Ras proteins' signaling strongly depends on their correct localization in the cell membrane and it is essential for activating downstream signaling pathways. KRas-4B function, membrane-association and interaction with other proteins are regulated by post-translational modifications (PTMs)[132–134], including ubiquitination, acetylation, phosphorylation, carboxymethylation, and prenylation.

We represent the crystal structure of KRas-4B *in vivo* in Fig. 1.3, but it would still need further PTMs, that are shown in Fig. 1.4 to function properly in cells. Firstly the prenylation reaction, catalyzed by cytosolic farnesyltrasferase (FTase) or geranylgeranyltransferase (GGTase), proceeds through the addition of an isoprenyl group to the Cys-185 side chain. Then farnesylated KRas-4B is ready for further processing: hydrolysis, catalyzed by the endopeptidase enzyme called Ras-converting enzyme 1 (RCE1), during the process the VIM motif (HVR tail composed of three amino-acids: valine-isoleucine-methionine) of the C-terminal Cys-185 is lost in step 2, see Fig. 1.4. Later KRas-4B is transferred to the endoplasmic reticulum for carboxymethylation at the carboxyl terminus of Cys-185 catalyzed by isoprenylcysteine

carboxyl methyltransferase (ICMT), forming a reversible ester bond. The outcome of these modifications is the farnesylated and methylated KRas-4B (KRas-4B-FMe). The reversible ester bond can go through decarboxymethylation, catalyzed by Prenylated/polyisoprenylated methylated protein methyl esterases (PMPEases) giving rise to a farnesylated and demethylated KRas-4B (KRas-4B-Far) which is the product of step 2 and reactant of step 3. This reversible reaction can modulate the equilibrium of methylated/demethylated KRas-4B population in tumors and consequently can impact downstream signaling, protein-protein interactions, or protein-lipid interactions[135].

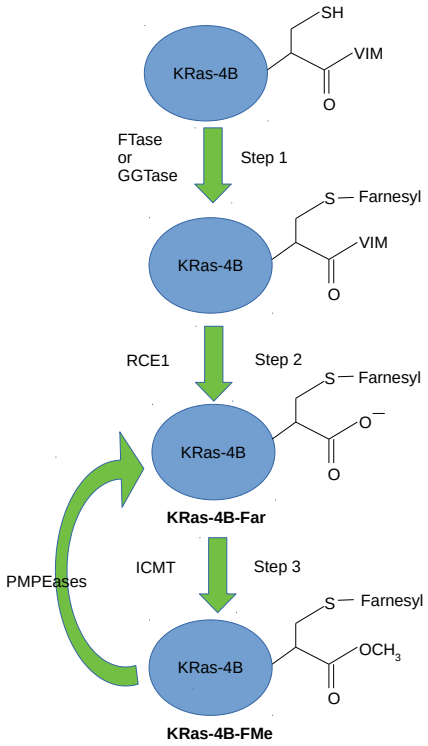


Fig. 1.4 PTMs steps of KRas-4B: prenylation, hydrolysis, carboxymethylation and decarboxymethylation.

One of the best known reversible PTMs in HVR is phosphorylation[134, 136, 137]. There are two sites (Ser-171 and Ser-181) within HVR that could be phosphorylated. Phosphorylation involves the addition of phosphate (PO_4^{3-}) group to the side chain of the amino acid serine, then the phosphorylated serine is obtained. Phosphorylation at Ser-181 operates a farnesyl-electrostatic switch that reduces but does not completely inhibit membrane-association and clustering of KRas-4B, leading to redistribution to the cytoplasm and endomembranes[133, 138, 139], whereas, phosphorylation at Ser-171 is less vital for its function[139]. Functionally, phosphorylation of KRas-4B can have either a negative[140, 141] or positive[139, 142] regulatory effect on the tumor cell growth, depending on the conditions[135]. For instance, from an MD simulation of the HVR peptide with the farnesyl group (FAR) at Cys-185 of KRas-4B in two types

of model membranes, it has been observed that phosphorylation at Ser-181 prohibits spontaneous FAR membrane insertion[143]. According to N.Agell, KRas-4B binding with calmodulin leads to different behaviors: short or prolonged signaling whether KRas-4B is at its phosphorylated state on residue Ser-181[139, 144]. In summary, phosphorylation of the HVR of KRas-4B can affect its function, membrane-association, and reacting with downstream effectors[135].

The translocation pathways of Ras proteins vary for different Ras isoforms, and signaling activity of KRas-4B is dependent on its enrichment level in the PM. Phosphodiesteraseδ (PDEδ) has been revealed to promote effective KRas-4B signaling by sequestering KRas-4B from the cytosol by binding the prenylated HVR and help to enhance its diffusion to the PM throughout the cell, where it is released to activate various signaling pathways required for the initiation and maintenance of cancer[145–148]. It is one of the common targets considered for oncological drug development. And investigation of the mechanistic inhibitory treatment of PDEδ has been carried out[147, 149], hoping to identify a panel of novel PDEδ inhibitors. According to S.Dharmaiah[145], the affinity between KRas-4B-FMe and PDEδ) is 78-fold compared with KRas-4B-Far by using ITC (isothermal titration calorimetry) measurements and also, a 5-amino-acid-long sequence motif in its HVR (K-S-K-T-K), which is shared by KRas-4B-Far and KRas-4B-FMe, may enable PDEδ to bind prenylated KRas-4B-FMe. However, it doesn't mean that KRas-4B-Far cannot be translocated to bind membrane bilayers. In a paper published in "Cell" in 2014[146] and another recent research in 2017[150], it was explained that KRas-4B could also be transferred to the PM through trapping and vesicular transport without the help of PDEδ. In the last step of carboxymethylation, there is a clear difference in the level of the reversible carboxymethylation of Cys-185 according to different kinds of tumor types related to KRas-4B. Interests have been focused largely on methylated KRas-4B-FMe, however, accumulating shreds of evidence indicate that demethylated KRas-4B-Far could play a big role in the signaling pathway that happens on the inner leaflet of the membrane bilayers. In 2018, 91% of the oncogenic KRas-4B and 51% of the wild-type KRas-4B can be found free from carboxymethylation in certain colorectal cell samples[110], which implies that in the tumor cells, demethylated KRas-4B-Far could be able to be translocated to bind the PM and functions well in its active state. In a mouse model studied by A.Wahlstrom et al., the myeloproliferative disease could be improved with higher demethylated KRas-4B-Far level in the cell caused by ICMT deficiency[151], however, in a mouse model of pancreatic cancer the disease could be accelerated. This indicates the behavior of demethylated KRas-4B-Far could differ in different diseased cells.

While there is a relatively high abundance of KRas-4B-Far (wild-type and mutant) lacking the methyl group of Cys-185 in tumors, the effects of demethylated KRas-4B-Far on downstream signaling have yet to be determined[110]. Only with the genomic sequence data can the oncologist not acknowledge the entire information in target

selection. As a well-studied membrane protein, there is still much to learn about KRas-4B's structure, dynamics, function, and ways of therapeutic targeting in both sides of clinical and theoretical aspects.

It has been reported that after these PTMs, a hydrophobic density is provided to KRas-4B-FMe after carboxymethylation so that its enhanced protein recruitment to the PM is revealed by most of the published works related to KRas-4B[132–135], which is debatable. With a preference for binding to anionic lipids of KRas-4B HVR, demethylated and farnesylated KRas-4B-Far shares the similar electrostatic potential with an anionic head and a hydrophobic tail in its FAR same with PM lipids. We propose that KRas-4B-Far might have more affinity for binding with anionic membrane bilayers which needs to be explored.

After extensive studies on experimental, clinical, and theoretical aspects, one relevant element is missing: the Free Energy Landscapes (FEL) as suggested by R.Nussinov et. al. in 2019[152]. She emphasized the need for considering every Ras isoform's completeness, especially in the changes in the protein conformation. And FEL describes the wide range of states, displays all the possible conformations that the molecule can adopt. FEL inspires and motivates approaches to molecular biology[153] and it helps us understand and propose testable hypotheses of how biomacromolecules have been exploited for life. Proteins continuously interconvert between states, and their fluctuations encode their functional behaviors under different conditions and environments[154]. Biomolecules interconvert their structures with carrying energies in nature. A proper physicochemical description of biological proteins is not as single structures but as conformational ensembles with dynamic distributions of states which change with varying conditions[153]. FEL could clarify how some mutations or abnormal PTMs can be oncogenic by crossing lower free energy barriers between two biologically relevant states, such as between destabilized inactive state ensemble and active state or between an inactive state and stabilized active state ensemble.

The dynamic FEL assimilates and embodies concepts from Physics and Chemistry which can help explain a range of biochemical mysteries, especially protein folding[155–157], binding to membranes and other proteins, conformational orientation, etc. Free energy change of membrane reorientation from relatively long simulations is believed that can be estimated from the 2D probability density distribution. For example, recent work revealed that the overall free energy surface of KRas-4B membrane-association is rugged, indicating a large diversity of conformations including at least two sub-states in each orientation state that differ in stability only by about 0.5-1.0 Kcal/mol and in between, two orientations of either the three C-terminal α-helices or N-terminal β-strands of KRas-4B catalytic domain facing the membrane are the most two common orientations with a barrier of $\sim$ 1 Kcal/mol for their interconversions[158].

Tighter binding of KRas-4B-GTP with lower binding free energy than KRas-4B-GDP has also been observed[159]. Binding free energies of KRas-4B-PDEδ were carried

out through the surface area continuum solvation (MMGBSA) approach[160, 161] provided in Amber16 suite[162]. Small molecules that bind and strengthen KRas-4B-PDEδ complex followed by colorectal cancer cell growth inhibited *in vitro* and *in vivo*[163] and their binding energies of the protein-ligand system have been explored using the same MMGBSA approach. Binding energies could provide useful information on selected conformations, orientations, and localizations of KRas-4B binding PMs. However, the relevant transition states are hardly traceable.

Extremely long simulation time scales that were caused by high free energy barriers serve as the roadblock during the investigations of the mechanism and kinetic aspects of the conformational transitions in Ras[164]. Overcoming these barriers could be accomplished by employing biasing potentials that drive the system towards the barrier region along with predefined collective variables. Through enhanced path sampling simulations, the conformational space and dynamic behavior of GTP-bound KRas-4B can be investigated. For instance, KRas-4B was shown visiting several states, which are the same for wild type and the oncogenic mutant Q61L and large differences of both systems, indicating the dramatic effect of Q61L mutation on KRas-4B dynamics[165].

Calculation of the FEL of conformational sub-states of KRas-4B could be astonishingly useful for medicine design for targeting KRas-4B related cancers, even though the binding free energies between selected conformations have been studied extensively by researchers in recent years[154, 157–159, 162]. The calculation of free energy barriers between its meta-stable states is poorly considered. Insight into the underlying biological mechanisms will play an important role in better defining 'actionability' for KRas-4B targets and bring new ideas to the "precision medicine"[152].

1.4 Objectives and outline

This Thesis presents the continuation and extension of a previous one, carried out by Dr. Jing Yang in 2015[166] in this same research group at UPC, who started to work in the modeling and simulation of phospholipid membranes in aqueous ionic solutions, for the first time in the SIMCON research group at UPC. The main object of the study of this Thesis is the research on the structure, dynamics, and FEL of small molecules and proteins binding PMs. The properties of these systems have been performed using different molecular dynamics simulations and well-tempered metadynamics simulations that allow us to calculate different microscopic properties of the system. The Thesis is structured as follows:

In Chapter 2 all techniques and computational packages employed in the present Thesis have been described. After introducing the method, we have done the research regarding the importance of the amino acids, neurotransmitters, hormones, and membrane surface proteins for the correct functions of the body.

In Chapter 3, firstly we have explored TRP's interactions with a typical zwitterionic phospholipid membrane formed by DPPC and cholesterol in an aqueous solution of sodium chloride at a temperature of 323.15 K under liquid-crystalline phase conditions. It is of central importance to elucidate their local structures and dynamics as well as the mechanisms responsible for the access of tryptophan to the interior of the cell. We have analyzed its mechanical properties, the local structures, and some relevant dynamical properties such as diffusion and vibrational spectra after performing molecular dynamics simulations. Microscopic properties such as the area per lipid, lipid thickness, radial distribution functions, hydrogen-bonding lengths, atomic spectral densities, and self-diffusion coefficients have been evaluated. Then four more biological systems of small molecules binding DPPC membrane bilayers have been studied by focusing our attention especially on the free energy barriers of the adsorption of small molecule species: histidine, tyrosine, serotonin, and melatonin. We have computed the radial distribution functions of all species for a variety of membrane and water-related sites and extracted potentials of mean force through the reversible work theorem.

In Chapter 4 we have focused our efforts on the study of the binding of melatonin at a simplified model cell membrane referring to the gradually increased interest in melatonin in recent years. Molecular dynamics simulations of melatonin inside mixtures of DMPC and cholesterol in NaCl solution at physiological concentration have been performed at 303.15 K to systematically explore melatonin-cholesterol, melatonin-lipid, and melatonin-water interactions. Properties such as the area per lipid and thickness of the membrane as well as selected radial distribution functions, binding free energies, angular distributions, atomic spectral densities, and translational diffusion of melatonin are reported.

To the best of our knowledge, the free energy landscape of melatonin on membrane surfaces is poorly studied. Free energy calculations are essential to unveil mechanisms at the atomic scale such as binding of small solutes and their translocation across cell membranes, eventually producing cellular absorption. We have used 1.4 μs well-tempered metadynamics simulations to perform Gibbs free energy calculations in order to precisely describe the characteristics of melatonin binding to specific sites in the membrane and reveal the role of cholesterol in free energy barrier crossing. A specific molecular torsional angle and the distance between melatonin and the center of the membrane along the normal to the membrane Z-axis have been considered as suitable reaction coordinates.

In Chapter 5 we shifted our interest to KRas-4B-membrane systems since KRas-4B is a challenging target for anti-cancer drugs, in part because that the dynamic behavior of flexible regions in the protein is difficult to characterize experimentally. To the best of our knowledge, this is the first time of the dedicated systematic work on the demethylated KRas-4B-Far. In order to reveal GTP's affinity to post-translationally modified KRas-4B-Far and its FAR group's binding free energy barriers anchoring

to/departing from membrane bilayers, especially for membranes with high cholesterol content, continuous MD simulations of three KRas-4B-membrane systems have been performed. Under the combining influence of mutation G12D, phosphorylation at site Ser-181, and demethylation of Cys-185, KRas-4B's behaviors in different systems have been analyzed and results have been introduced from a Structural and energetic perspective.

In this Thesis, we have discussed whether increasing its hydrophobic density in the structure of KRas-4B helps anchor the PM by studying three systems of KRas-4B-FMe and KRas-4B-Far from the atomic level. A better understanding of the localizations of different domains of wild-type/oncogenic KRas-4B proteins on the anionic membrane is important for its precision medicine and offering detailed free energy landscapes of FAR and GTP could hopefully provide useful insights for this field. In order to obtain atomistic insight into the dynamics of KRas-4B-Far, we have depicted several stable conformations that have not been captured by experiments through a well-tempered metadynamics method. For the first time through well-tempered metadynamics simulation using 2 collective variables (distances along the membrane normal between the mass centers of GTP and FAR group at Cys-185 site to the membrane mass center) we investigated their free energy landscapes as compared with the wild-type protein, the theoretical distinction among the mechanisms of G12D mutation and phosphorylation of oncogenic KRas-4B-Far and wild-type KRas-4B-Far is clarified.

Chapter 2

Methods

In this Chapter, we will describe all techniques and computational packages employed in the present Thesis. All input files were generated from a web-based platform called CHARMM-GUI, from which CHARMM36 and CHARMM36m force fields are adopted and applied to small molecule-membrane bilayer systems during molecular dynamics (MD) simulations. GROMACS and NAMD are used during standard MD simulations for different biological systems and GROMACS and PLUMED are used when performing well-tempered metadynamics simulations.

2.1 CHARMM-GUI web-based platform

Since its original development in 2006, CHARMM-GUI[167, 168], http://www.charmm-gui.org, has proven to be an ideal web-based graphical user interface to build complex molecular systems and prepare their input files for simulation packages such as CHARMM (Chemistry at Harvard Macromolecular Mechanics), NAMD (Nanoscale Molecular Dynamics), GROMACS (GROningen MAchine for Chemical Simulation), AMBER (Assisted Model Building with Energy Refinement), GENESIS, LAMMPS, Desmond, and OpenMM to facilitate and standardize the usage of wide-used and advanced simulation techniques. As a well-established and reproducible simulation protocol for state-of-the-art biomolecular simulations, CHARMM-GUI has been widely adopted for various purposes and now contains a number of different modules designed to set up a broad range of biomolecular simulation systems in Input Generator. In the year 2007, a Java applet integrated within CHARMM-GUI to induce a novel online visualization tool based on MarvinSpace (https://chemaxon.com/marvin-archive/4.1.6/marvin/doc/dev/space-map.html) molecular visualization software which allows one to interactively inspect the generated molecular system at each step. In this way, users could go back to the previous setup and regenerate the whole system again with generally accepted standards and well-qualified methods before quitting

the web browser. CHARMM-GUI has been used by many researchers, and it is a well-recognized tool in the biomolecular modeling and simulation communities.

The rest of the section will be organized as follows. We describe several currently available functional modules of CHARMM-GUI Input Generator that are designed to set up a broad range of simulations: PDB Reader, Membrane Builder, Solution Builder, Quick MD Simulator, Free Energy Calculator, ligand Reader & Modeler, et cetera[168–170]. We will focus on the first three modules which were used in this Thesis.

2.1.1 PDB Reader

Reading a PDB file into a simulation program is not straightforward and is generally considered the first hurdle in a simulation project of a biosystem. PDB Reader is the most basic, yet powerful web interface to convert a PDB file. It reads a PDB (Protein Data Bank[171]) file from RCSB (http://www.rcsb.org) PDB (with PDB entry ID) or uploads it from a local computer, then converts it to CHARMM readable format with the following options:

- partial selection of protein chains and model selection in the case of NMR structures,
- modification of engineered residues,
- terminal group selection,
- protonation selection,
- disulfide bond selection,
- phosphorylation selection,
- generation of a biologically functional unit
- generation of crystal packing.

For example, according to Jo et al.[168], PDB Reader automatically detects some of the engineered residues and converts them to corresponding natural residues. However, due to the complexity of PDB files with many kinds of heteroatoms and other engineered residues, users themselves need to check if all the residues to be read do exist in a CHARMM topology file (currently, top_all36_prot.rtf). If there are undetermined coordinates in selected chains, PDB Reader will simply build them using a predetermined internal coordinate table.

2.1.2 Membrane Builder

Simulating a realistic biomembrane system is very challenging. Since its first release in 2007, Membrane Builder[167, 168, 172–174] consists of 6 major functional sub-modules: Bilayer Builder, HMMM Builder, Nanodisc Builder, Monolayer Builder, Micelle Builder, and Hex Phase Builder. Here I will only introduce Bilayer Builder through which all biosystems are generated in this Thesis. Membrane Builder, especially Bilayer Builder, helps users to generate a series of CHARMM inputs necessary to build a protein-membrane complex for MD simulations. Through a series of steps of generating organic molecules plus bilayer membrane system, users need to provide validated PDB of properly oriented structures in the first building system step and also mutation, phosphorylation could also be carried out in step 1. After modifying the orientation with respect to membranes and its location along the normal direction for the molecule in step 2, water molecules and different kinds/amounts of sterols, lipids, fatty acids, detergents, et al., which allow users to build experimentally comparable and biologically realistic membrane systems. Based on the system size determined in step 3, positive and negative ions will be generated in step 4. All the pieces (protein, lipid bilayer, additional water, and ions) are assembled together in step 5. Finally and importantly input files with appropriate formats for various MD programs could be obtained in the last step.

Compared to other tools for protein-membrane model construction[175–179], Membrane Builder is much more flexible and robust in terms of producing reasonably packed systems which are proven to be of usage worldwide, providing simulation inputs, various user options, and availability of various lipids. An additional advantage of Membrane Builder is to have various functionalities available to read ligand structures and modify protein side chains with MTS reagents or unnatural amino acids during the PDB structure reading step. E.L.Wu[173] reported new features and major improvements: various kinds of lipids and robust and efficient algorithms could be adopted in building realistic biological membrane systems. Membrane Builder helps users to build sophisticated protein/membrane or membrane only systems through an easy and automated process.

2.1.3 Solution Builder

It is common simulating DNA/RNA, ions, proteins, small molecules in a realistic environment, especially in different aqueous solutions. Solution Builder[168, 174] of CHARMM-GUI helps users to generate a series of input files for MD simulation of the target molecule in aqueous solvent environments. Molecules can be solvated, for example, relax the protein structures in a certain aqueous ionic solution (NaCl, KCl, $CaCl_2$, and $MgCl_2$) or generating a simple water box only for other purposes. Solution

Builder also guides users to define its dimension and build the water boxes in different shapes: orthorhombic, hexagonal, octagonal, cylindrical, or spherical shapes.

The default ion concentration is 0.15 M KCl that is close to the physiological ion concentration (C) in human body. Numbers of ions are determined by the ion-accessible volume (V) and the total charge of the system (Q_{sys}) as follows:

$$\begin{aligned} N_+ &= CV - Q_{sys}/(2ez_+) \\ N_- &= z_+CV + Q_{sys}/2e \end{aligned} \tag{2.1}$$

where N_+ and N_- are numbers of positive and negative ions respectively, e is the charge of an electron, and z_+ is the valence of the positive ion. The number of ions roughly corresponds to the linearized Poisson-Boltzmann theory[180]. Solution Builder allows users to visualize the solvated system to validate if the system is built properly. Short steps of minimization are performed to remove bad contacts in step 3 and equilibrated through the later MD simulation.

2.2 Force Fields

Which force field to use in MD simulations using simulation packages? The most common choices are various versions of AMBER[181–183], CHARMM, and Optimized Potentials for Liquid Simulations (OPLS)[184–186]. All force fields rely on similar functional forms, however, each one has certain strengths and weaknesses. The CHARMM force field, adopted by this work, is the name of a widely used set of force fields for MD. Multiple versions of each force fields are available, they all use strikingly similar functional forms.

The potential energy function used in CHARMM force field is as follows:

$$\begin{aligned} U_{total} = &\sum_{bonds} K_b(b-b_0)^2 + \sum_{angles} K_\theta(\theta-\theta_0)^2 + \sum_{dihedrals}\sum_{j} K_{\varphi,j}[1+cos(n_j\varphi-\delta_j)] \\ &+ \sum_{vdW_{i,j}} \varepsilon_{ij}[(\frac{R_{min,ij}}{r_{ij}})^{12} - (\frac{R_{min,ij}}{r_{ij}})^{6}] + \sum_{Coulomb\,i,j} \frac{q_iq_j}{\varepsilon_{D^{r_{ij}}}} \end{aligned} \tag{2.2}$$

where the terms have their usual meanings. The bond and angle terms are harmonic, with force constants K_b and K_θ and equilibrium values b_0 and θ_0. The dihedral potential is a sum of sinusoids with force constant $K_{\varphi,j}$, multiplicity n, and offset φ, and j can range from 1 to 6. For the van der Waals (vdW) interactions, ε_{ij} is the potential energy minimum between two particles separated by r_{ij}, and $R_{min,ij}$ is the position of this minimum. Lastly, q_i and q_j are the partial atomic charges for the Columbic term;

the dielectric constant ε_D equals 1 in explicit solvent simulations. For CHARMM36, The bond and angle terms are left unchanged from previous modifications[187–189].

The following objective function used to obtain all new torsional terms for CHARMM36 force field:

$$\chi = \sum_{i}^{no.\,of\,QM\,points} \left[U_i^{QM} - U_i^{model}\right]^2 \tag{2.3}$$

where U_i is the energy for conformation i.

In CHARMM36 the new vdW parameters included the recently optimized oxygen vdW parameters for the ether oxygen in the ester linkage, but methyl vdW parameters stay the same. According to Ref [189], using CHARMM36 force field, deuterium order parameters for carbons in the glycerol and headgroup region of the lipid are well described, surface areas of a total 6 lipids are in very good agreement with experiment data, and intra-molecular conformations near the glycerol moiety could be soundly described during simulations. CHARMM36 is a significant update to the CHARMM27/27r lipid force field[190–192]. Hydration of lipids near the carbonylglycerol section of phospholipids is improved and it allows for the use of a tensionless ensemble (NPT) in MD simulations. It is shown that the CHARMM36 force field leads to better correlation with experimental data compared to the CHARMM22/CMAP force field (sometimes referred to as CHARMM27 force field[190]) and using CHARMM36 force field in protein simulations is suggested cause CHARMM36 force field could more accurately treat folded proteins and equilibrium of conformations observed in partially disordered peptides. The major improvements over CHARMM27 involve new side-chain dihedral parameters and refined backbone CMAP potentials. In CHARMM36, other minor improvements include revised parameters for the guanidinium moiety of arginine, aliphatic hydrogens, and tryptophan[189]. And CHARMM36 protein force field was found to generate a high population of left-handed α-helix (αL), inconsistent with NMR spectroscopy and small-angle X-ray scattering experimental measurements.

CHARMM36m force field comes from the all-atom additive CHARMM36. In CHARMM36m, the modified CHARMM36 protein parameter set (par_all36m_prot.prm), which corrects oversampling of the αL conformation and gives an improved sampling of disordered states of peptides, was available online in July 2017. Together with top_all36_prot.rtf, all simulations involving peptides and proteins, even for folded proteins, were improved. It also improves accuracy in generating polypeptide backbone conformational ensembles for intrinsically disordered peptides and proteins based on a refined backbone potential derived from reweighting calculation and a better description of specific salt bridge interactions. It was validated by Huang et al. in 2017[185]. In the present Thesis, we employed CHARMM36 and CHARMM36m force fields along with our research.

2.3 Molecular Dynamics Simulations

Molecular dynamics is a computer simulation method that mimics what atoms do in real life, assuming a given potential energy function. It provides detailed information on the fluctuations and conformational changes of atoms and molecules in materials[193–195], and it was firstly introduced in the late 1950s after Monte Carlo simulation[196]. The aim of MD simulation is to trace the classical trajectory of a N particles system or to describe its dynamics. The system has been prepared in a particular initial condition. We wish to predict the average behavior of a system along multiple trajectories[197]. There are two main families of MD methods: 'classical' mechanics approach to MD simulations (molecules are treated as classical objects) and 'quantum' or 'first-principles' MD simulations (the quantum nature of the chemical bond is taken into account), and the latter is outside of the scope of this Thesis.

2.3.1 Equations of motion

In classical MD it is assumed that each particle in the system behaves like a Newtonian particle and the quantum behavior is completely ignored. The potential energy $U(r_1, r_2, r_3, ..., r_n)$ of N interacting particles in the system depends on their positions $r_i = (x_i, y_i, z_i)$. The force acting on ith atom is determined as Ref. [198]:

$$F_i = - \bigtriangledown_{r_i} U(r_1, r_2, r_3, ..., r_n) = -(\frac{\partial U}{\partial x_i}, \frac{\partial U}{\partial y_i}, \frac{\partial U}{\partial z_i}). \tag{2.4}$$

Time evolution of a series of particles is obtained through Newton's Second law of motion:

$$F_i = m_i \frac{d^2 r_i(t)}{dt^2} \tag{2.5}$$

where $r_i(t) = (x_i(t), y_i(t), z_i(t))$ is the position of ith atom at time t and m_i is the mass of the atom i.

In MD, in order to integrate Newton's equations of motion, the most commonly used and time integration algorithm is the so-called Verlet algorithm, introduced by Verlet in 1967[199]. In order to derive it, Taylor expansion of the coordinate of a particle around time t is adopted:

$$r(t + \Delta t) = r(t) + v(t)\Delta t + \frac{F(t)}{2m}\Delta t^2 + \frac{\Delta t^3}{3!}\dddot{r} + O(\Delta t^4)$$

similarly,

$$r(t - \Delta t) = r(t) - v(t)\Delta t + \frac{F(t)}{2m}\Delta t^2 - \frac{\Delta t^3}{3!}\dddot{r} + O(\Delta t^4)$$

An error of order Δt^4 for the estimate of the new position exists in the two equations above. Summing these two equations, we could obtain:

$$r(t+\Delta t) + r(t-\Delta t) = 2r(t) + \frac{F(t)}{m}\Delta t^2 + O(\Delta t^4)$$

or

$$r(t+\Delta t) \approx 2r(t) - r(t-\Delta t) + \frac{F(t)}{m}\Delta t^2 \tag{2.6}$$

Where r(t) is the particle's position at time t, Δt is known as the time step in MD simulation. So using

$$r(t+\Delta t) - r(t-\Delta t) = 2v(t)\Delta t + O(\Delta t^3)$$

or

$$v(t) = \frac{r(t+\Delta t) - r(t-\Delta t)}{2\Delta t} + O(\Delta t^2),$$

$$v(t) \approx \frac{r(t+\Delta t) - r(t-\Delta t)}{2\Delta t} \tag{2.7}$$

We could derive the velocity from trajectories obtained from MD simulation which precision is of order Δt^2.

Once the first configuration of the system is given, the following configurations of the system can be computed, i.e. new positions and velocities also can be calculated. After repeating the procedure many times, a series of trajectories of atomic motions will be obtained. After analyzing meaningful trajectories of MD simulation, the movement of atoms: velocities, positions, and orientations as a function of simulation time could be obtained.

MD simulations are often used in comparison to and combination with many experimental structural biology techniques, including X-ray crystallography, neutron scattering, electron paramagnetic resonance (EPR), nuclear magnetic resonance (NMR), IR and Raman spectroscopy, cryoelectron microscopy (cryo-EM), and Förster resonance energy transfer (FRET). MD simulations are used both to interpret experimental results and to guide experimental work[200]. In important published papers about the study of proteins related to neuronal signaling[201–204]; providing the foundation for the design of improved optogenetics tools[205], and assisting in improving drugs targeting nervous systems[206, 207], MD simulations play an important role in the research.

Even though MD simulations themselves are becoming much more powerful and easier to use for non-experts over the past decades, limitations of the method still exist. They are related to the parameter sets used and to the underlying molecular mechanics force fields whose accuracy needs to be improved. In a wide variety of important biomolecular processes, such as conformational change, ligand binding,

and protein folding are difficult to be accessed within the limited time scales of the conventional MD simulation. Most high-impact works performed using MD simulations needed the support of supercomputer centers. Advances in computer power have enabled microsecond simulations, but simulation timescales remain to be a challenge which makes enabling longer-timescale simulations an active research area, among which algorithmic improvements, parallel computing, and hardware such as GPUs and specialized hardware, etc. are involved.

2.3.2 Periodic boundary conditions

In Biophysics and Structural Biology, MD simulation is widely applied to study proteins, nucleic acids, cell plasma membrane systems which can provide information on interpreting the results of certain biophysical experiments and modeling interactions with other molecules. In Physics, MD is used to examine the dynamics of atomic-level phenomena that can not be observed directly, such as thin-film growth, examine physical properties of nanotechnological devices that have not or can not yet be created, and so on. In all kinds of MD simulations, the simulation box size must be large enough to avoid boundary condition artifacts. Boundary conditions are often treated by choosing fixed values at the edges or by employing periodic boundary conditions (PBCs) in which one side of the simulation loops back to the opposite side, mimicking a bulk phase. Both may cause artifacts for the system[208]. PBCs are a set of boundary conditions chosen for approximating a large (infinite) system by using a small part called a unit cell. The size of the simulation box must also be large enough to prevent periodic artifacts from occurring due to the nonphysical topology of the simulation.

One example of periodic boundary conditions can be defined according to smooth real functions $\phi : \Re^n \to \Re$ by

$$\begin{aligned}
&\frac{\partial^m}{\partial_{x_1^m}}\phi(a_1, x_2, ..., x_n) = \frac{\partial^m}{\partial_{x_1^m}}\phi(b_1, x_2, ..., x_n),\\
&\frac{\partial^m}{\partial_{x_2^m}}\phi(a_1, x_2, ..., x_n) = \frac{\partial^m}{\partial_{x_2^m}}\phi(x_1, b_2, ..., x_n),\\
&...,\\
&\frac{\partial^m}{\partial_{x_n^m}}\phi(a_1, x_2, ..., x_n) = \frac{\partial^m}{\partial_{x_n^m}}\phi(x_1, x_2, ..., b_n)
\end{aligned} \tag{2.8}$$

For all m = 0, 1, 2, ... and for constants a_i and b_i. In MD simulation, PBCs are usually applied to simulate bulk gases, liquids, crystals or mixtures.

2.3.3 Ensembles

Generalized ensembles, Isothermal–isobaric (NPT) ensemble, Canonical ensemble (NVT), and Microcanonical (constant-NVE) ensemble are ensembles used in MD simulations, among which NPT and NVT ensembles are the most used ones.

NVT is the statistical ensemble that represents the possible states of a mechanical system with constant temperature (T), system's volume (V) and the number of particles in the system (N). The system can exchange energy with the heat bath[209], so that states of the system along the simulation time will differ in total energy. Under the NVT ensemble, the partition function of a bulk system with N particles and the volume V is given by[210, 211]:

$$\begin{aligned} \Lambda &= \sqrt{2\pi\hbar^2/mk_BT} \\ Q(N,V,T) &= \frac{V^N}{\Lambda^{3N}N!}\int_{D(V)} dq\, exp\,[-\frac{U(q)}{k_BT}] \end{aligned} \tag{2.9}$$

where Λ is the thermal de Broglie wavelength, U is the potential energy of the atomic coordinates $q = \{q_1, q_2, ..., q_N\}$, $D(V)$ denotes the spatial domain defined by the containing volume. The expression for the Helmholtz free energy of the system is thus[212]:

$$\begin{aligned} A(N,V,T) &= -k_BT\, lnQ(N,V,T) \\ &= -k_BT\, ln\frac{V^N}{\Lambda^{3N}N!} - k_BT\, ln\int_{D(V)} dq\, exp\,[-\frac{U(q)}{k_BT}] \end{aligned} \tag{2.10}$$

where $-k_BT\, ln\frac{V^N}{\Lambda^{3N}N!}$ is the free energy of an ideal gas.

The practical use of the NVT ensemble is usually justified in two ways: by assuming that the contact is mechanically weak, or by incorporating a suitable part of the heat bath connection into the system under analysis so that the connection's mechanical influence on the system is modeled within the system.

NPT is a statistical ensemble that maintains constant pressure P, temperature T and number of particles N in the system. It is often used for measuring the equation of state of model systems whose virial expansion for pressure cannot be evaluated, or systems near first-order phase transitions[197]. When NPT ensemble is used, the characteristic state function of this ensemble is the Gibbs free energy given by:

$$\begin{aligned} G(N,P,T) &= -k_BT\, ln\Delta(N,P,T) \\ &= -k_BT\, ln\int dV\, exp\,[-\frac{PV}{k_BT}]exp\,[-\frac{A(N,V,T)}{k_BT}] \end{aligned} \tag{2.11}$$

NPT ensemble is very useful in MD simulations, e.g. to model the behavior of water at ambient conditions. Computing directly the partition function for an arbitrary potential is impractical, and for this reason, practical routes for computing Helmholtz free energy or Gibbs free energy typically use a reference system for which the phase-space integral can be computed analytically, followed by one or more thermodynamic integration steps [212].

2.3.4 Limitations of MD simulations

It is crucial to be aware of the limitations of MD in order to make reasonable use of it. Dynamical events of important biologically processes which involve quantum effects such as changes in the presence of important noncovalent intermediates, chemical bonding, and tunneling of protons or electrons. Straightforward atomic force field simulations cannot be used to model such phenomena. Even the available force fields, such as AMBER, CHARMM, or GROMOS, have proved to be sufficiently accurate in terms of kinetic and thermodynamic properties. Still, there is vast room for improvement. The stability and accuracy of MD will be challenged during long-time MD simulations. So, improving the quality of the atomic force fields is highly desirable. Relevant time scales for biologically important processes that place limits on the desired increasing size of simulated biological systems.

2.4 Well-tempered metadynamics simulations

After its early beginning run by Mary Ann Mansigh in 1962, MD has become a standard tool used in many branches of Science. The problem of sampling complex free energy landscapes, of which meta-stable states are separated by large energy barriers, has recently attracted much attention. Also, these process of interesting events often take place on a large timescale. A number of methods could be adopted: transition path sampling[213], local elevation[214], steered MD[215], conformational flooding[216, 217], umbrella sampling[218–223], adaptive force bias[224], and metadynamics[225–230].

Metadynamics is a powerful technique which has proven its effectiveness in a variety of contexts, aiming at enhancing the sampling of the configurational space. In this method sampling is facilitated with an additional bias potential (or force) which acts on a selected number of degrees of freedom $\vec{s}(q)$: the so-called collective variables (CVs) along the trajectory. The potential is constructed as the sum of Gaussian functions[214]:

$$V_{mem}(\chi) = k_{mem} n_{\chi^0} exp^{-\frac{(\chi-\chi^0)^2}{2\omega^2}} \tag{2.12}$$

where χ is the current conformation, χ^0 is the conformation visited before, k_{mem} (>0) defining the magnitude and ω is the width of the memory penalty function, n_{χ^0}

is the number of times this conformation has been sampled before (including the dimension).

This potential could be expressed as follows[230–232]:

$$V(\overrightarrow{s},t)=\sum_{k\tau<t}W(k\tau)exp\,(-\sum_{i=1}^{d}\frac{(s_i-s_i(q(k\tau)))^2}{2\sigma_i^2}) \tag{2.13}$$

where τ is the Gaussian deposition stride, $W(k\tau)$ the height of the Gaussian, and σ_i the width of the Gaussian for the ith CV.

In the long time limit, the bias potential converges to minus the free energy as a function of the CVs:

$$V(\overrightarrow{s},t\to\infty)=-F(\overrightarrow{s})+C \tag{2.14}$$

After a transient, the bias potential compensates the underlying free energy landscape and provides an estimate of its dependence on the selected CVs. However, it is often difficult for users to decide when to terminate metadynamics simulation. In metadynamics, the free energy does not converge to a definite value but fluctuates around a certain result in a single run which leads to an average error which is proportional to the square root of the bias potential deposition rate[232]. And the time required to fill the free energy landscape increases when this rate is reduced.

In order to alleviate these problems and inspired by the self-healing umbrella sampling methods, well-tempered metadynamics was introduced by Michele Parrinello et al. (2008)[231]. In well-tempered metadynamics, the height of Gaussian is decreased with simulation time according to:

$$W(k\tau)=W_0exp\,(-\frac{V(\overrightarrow{s}(q(k\tau)),k\tau)}{k_B\Delta T}) \tag{2.15}$$

where k_B is the Boltzmann constant, W_0 is an initial Gaussian height, ΔT is an input parameter with the dimension of a temperature. By rescaling the Gaussian height during the simulation, the bias potential could converge in a smooth way, but it does not fully compensate the underlying free energy in the long limit:

$$V(\overrightarrow{s},t\to\infty)=-\frac{\Delta T}{T+\Delta T}F(\overrightarrow{s})+C \tag{2.16}$$

where T is the system temperature.

When simulation time approaches the long time limit, along with CVs, an ensemble is sampled at a temperature $T+\Delta T$. When $\Delta T=0$ corresponds standard MD, and $\Delta T\to\infty$ to standard metadynamics. By definition, bias factor γ:

$$\gamma=\frac{\Delta T}{T+\Delta T} \tag{2.17}$$

so, Eq. 2.16 is changed into:

$$V(\overrightarrow{s}, t \to \infty) = -\frac{F(\overrightarrow{s})}{\gamma} + C \tag{2.18}$$

So, γ shall be chosen carefully for the system to escape from any local minimum and into visiting new regions of the phase space in an effective way in the time scale if simulation.

2.4.1 Metadynminer

Metadynamics is a highly successful enhanced sampling technique for simulation of molecular processes and prediction of their free energy landscapes. It is as crucial as the simulation itself to do an in-depth analysis of data obtained by this method. R is a programming language and free software environment for statistical computing[233] and graphics[234–236] supported by the R Foundation for Statistical Computing. The R language is widely used among statisticians and data miners for developing statistical software and data analysis[237, 238]. R packages are collections of functions and data sets developed by the community. They increase the power of R by improving existing base R functionalities, or by adding new ones. Metadynminer is one R package for loading, reading, analyzing, and visualization of 1D and 2D hills files produced from metadynamics simulation. It uses a fast Bias Sum algorithm by P.Hosek and V.Spiwok[239] to calculate the free energy landscapes. Minima can be located and plotted on the free energy landscapes. Transition states can be analyzed by the Nudged Elastic Band (NEB) method by G.Henkelman and H.Jonsson[240]. Minimum energy paths for transitions such as atomic and/or spin rearrangements in thermalized systems are the transition paths of the largest statistical weight. Such paths are frequently calculated using the NEB method[241–245], where an initial path is iteratively shifted to the nearest minimum energy path. In order to effectively locate transition states, to frequently calculate the minimum energy paths for transitions, and to determine the atomistic reaction pathways, energy barriers, and binding potentials, the Nudged Elastic Band method has been largely adopted. In this method, an initial path is iteratively shifted to the nearest minimum energy path.

An estimation of the thermal stability of a state can be of important importance. The lifetime of a state at a given temperature can be estimated using harmonic transition state theory[246, 247], for some atomic, nuclear, or particle states, this lifetime can be very short. Halftime[248, 249], calculated through the metadynminer package, could be chosen to describe how quickly two (meta-)stable states undergo the transition state. With this method, free energy landscapes, minima, and transition paths can be computed.

2.5 Simulation packages used in this Thesis

2.5.1 NAMD

NAMD is one of the typical parallel MD program designed for high-performance simulation of large biomolecular systems. It is distributed free of charge with source code which scales to hundreds of cores for typical simulations and beyond 500,000 cores for the largest simulations and many important papers have been published using NAMD[250–260]. NAMD2 described in Ref. [261] is one of the programs aimed at utilizing large parallel machines in a scalable manner. Using intelligent periodic load balancing and maximally utilizing the available compute power are available. A combination of algorithms and numerical techniques guarantees that energy drifts are minimized to ensure the simulation's accuracy. In this Thesis NAMD2.7 was adopted.

2.5.2 GROMACS

GROMACS is a versatile package to perform MD for systems with hundreds to millions of particles. Now users could download the 5, 2016, 2018, 2019, and the latest 2020 release series of GROMACS on web pages: http://manual.gromacs.org/documentation. GROMACS was firstly introduced by Berendsen in 1995 [262] which was primarily designed for biochemical molecules like proteins, lipids, and nucleic acids that have a lot of complicated bonded interactions. Also, many groups use GROMACS to simulate non-biological systems because of its rapidness of non-bonded interactions.

In principle using the updated versions of GROMACS could give better performance and results. Here we have employed GROMACS/5.14, GROMACS/2018.3 in different systems. The GROMACS/2018 major release was out on January 11, 2018, in which several useful performances have been obtained: 1, many fewer CPU cores are needed for good performance; 2, a physical validation suite is added to verify expected properties; 3, a new dual-list dynamic-pruning algorithm for the short-ranged interactions has been adopted to permit longer-lived outer list with less work and more dependent on a certain parameter, et al.

2.5.3 PLUMED

PLUMED, firstly introduced in 2009, is a plugin that works with a large number of MD codes, like NAMD, GROMACS, AMBER, LAMMPS, ABIN, CP2K, Quantum ESPRESSO, and so on. And it can also be used together with analysis and visualization tools HTMD, and VMD. More importantly PLUMED is an open-source library implementing various free-energy methods with some of the most popular enhanced sampling techniques for MD simulations. In addition, PLUMED offers a large collection of CVs, for example, dihedral angles, distances, α-helical contents, positions, components of the simulation cell, coordination numbers, different energies,

and volumes of the simulation cell. The detailed theoretical background could be found in Ref[263]. PLUMED 2 which has been adopted in well-tempered metadynamics simulations in Chapter 4 shows greater flexibility and modularity and is far easier to add new methods and CVs than the original design. With a simpler interface with the MD engine and a single software library including tools and core facilities, it obviously provides much better service and brings new features for the community of users and developers.

2.5.4 VMD

Visual Molecular Dynamics (VMD)[176] is designed for modeling, visualization, and analysis of biological systems such as proteins, lipid bilayer systems, et al. Being released in 1995, VMD was developed for silicon graphics workstations and could also communicate with NAMD program in MD simulations. In general, VMD is used to read standard Protein Data Bank (PDB) files and display the contained structure and to animate and analyze the output files of MD simulations. It also includes tools for working with sequence data, volumetric data, and arbitrary graphics objects.

Chapter 3

Small molecules bound to di-palmytoil-phosphatidyl-choline bilayer membranes

In the present Chapter, microscopic structures of a fully DPPC lipid bilayer membrane with a small biological probe embedded inside, such as α-amino acids (TRP, HIS, and TYR), a neurotransmitter (SER) and a kind of hormone (MEL), have been analyzed through all-atom MD simulations based on the recently parameterized CHARMM36 force field. We have focused our efforts in two directions: on one side, the study of phospholipid membranes can help understand basic biological membrane functions and interactions with the environment. On the other side, the role of amino acids and drugs and their interactions with the membrane structure is undoubtedly a relevant field of research. As it is well known, cholesterol plays a key role in the structure of cell membranes. The effect of cholesterol on the binding of tryptophan to DPPC membranes is of great scientific interest. The characterization of the microscopical forces between the essential TRP and the basic components of cell membranes and their environments (phospholipids, cholesterol, ionic species, and water) is of central importance to elucidate their local structures and dynamics as well as the mechanisms responsible for the access of TRP to the interior of the cell. Finally, in the last section of this Chapter, free energy barriers associated to the binding of small-molecules at phospholipid zwitterionic membranes inside aqueous sodium chloride solution for a variety of species: TRP, HIS, TYR, SER, and MEL have been investigated, given their importance for the correct function of the body.

3.1 Preliminary study on the adsorption of tryptophan into DPPC bilayers

To shed light on the hot topic of the uptaken of drugs and proteins through the cell membranes *in vivo*, we have considered the introduction of a small biological probe into the lipid bilayer structure: the amino acid TRP[54, 58]. Since it has to be acquired from the diet, hence it is able to act as a building block in protein biosynthesis, while proteins are fundamentals required to sustain life. So, given the importance of TRP, we started our study exploring its interactions with DPPC and water in a salty solution of sodium chloride as well as its local structure, molecular bonding, and free energy profiles.

3.1.1 Computational information

Our system has been build by means of the CHARMM-GUI web-based tool. Here the bilayer membrane is composed of 204 lipids distributed in two leaflets of 102 flexible DPPC molecules surrounded by 4962 TIP3P[264] water (W) molecules plus 17 sodium and 17 chloride ions, corresponding to physiological concentration (0.15 M) and one TRP molecule. Sketches of the backbone structure of DPPC and of TRP are represented in Fig. 3.1. TRP and each DPPC molecule are described with atomic resolution (27 and 130 sites, respectively). MD simulations were performed with the NAMD2 simulation package[265] at a temperature of 310.15 K and the pressure of 1 atm. The simulation time step was set to 2 fs. The recently parameterized force field CHARMM36[189], which is able to reproduce the area per lipid in excellent agreement with experimental data, has been used. All molecular bonds have been left non-rigid, allowing fluctuations of bond distances and angles. Van der Waals interactions were cut off at 12 Å with a smooth switching function starting at 10 Å. Long ranged electrostatic forces were computed with the help of the particle mesh Ewald method[266], with a grid space of about 1 Å. Electrostatic interactions were updated every time step and periodic boundary conditions were applied in three dimensions of space. After thermal equilibration (for about 10 ns), two 25 ns production runs were generated, with a simulation box of size: 74.4 Å$\times$ 74.4 Å$\times$ 70.1 Å. The temperature was controlled by a Langevin thermostat[267] with a damping coefficient of 1 ps^{-1}, whereas the pressure was controlled by a Nosé-Hoover Langevin barostat[268] with a damping time of 50 fs. Finally, in order to eliminate any artificial drift of the center-of-mass of the system in the simulations, the coordinates of lipid atoms were corrected for the motion of the center-of-mass of the monolayer they belong[269]. In a previous simulation where DMPC lipid bilayers were modeled[17], a wide variety of properties of the simple lipid bilayer were satisfactorily reproduced (surface area per lipid, lateral pressure profile, order parameter of the lipid tails, etc.), so that we have not included these verifications here.

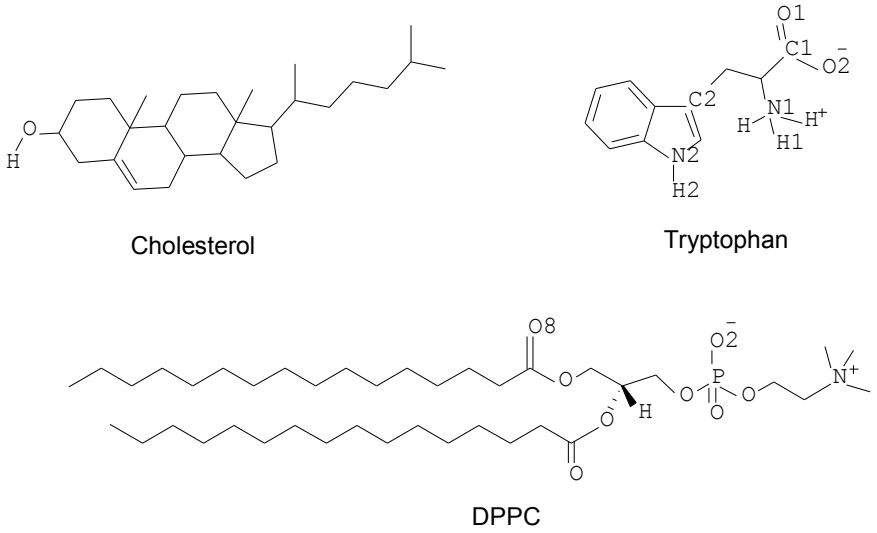


Fig. 3.1 Sketches of the backbone structures of DPPC, cholesterol, and TRP. Hydrogens bound to carbon and nitrogen are not shown. The highlighted sites of TRP ('H1', 'H2', 'N1', 'N2', 'C1', 'C2', 'O1', and 'O2') and of DPPC ('N', 'O2', and 'O8') will be referred in the text by the labels defined here. Due to the zwitterionic characteristics of TRP, its site 'H1' corresponds to any of the three hydrogens bound to 'N1', which share the positive charge. Sites 'O1' and 'O2' of TRP share the negative charge. Site 'C2' is close to the center of mass of the molecule.

A general view of the system is shown in Fig. 3.2. There, for the sake of clarity, water has been hidden so that only TRP, DPPC, and sodium and chloride ions have been shown. We have chosen two significant configurations, one for TRP adsorbed to sites located at the headgroup in DPPC (left side of Fig. 3.2) and another one for water solvating TRP (right side of Fig. 3.2). We should point out that TRP tends to stay close to the headgroup regions of the lipid chains during most time of the simulated trajectories: An estimation over the full trajectory indicated that around 70% of time TRP is attached to DPPC.

3.1.2 Results and discussion

3.1.2.1 Stability of the system

Before addressing the structural organization of the system, we checked the stability of the simulations and represented the contribution of all kinds of energies to the total energy of the system (Fig. 3.3) as well as the temperature and the pressure (Fig. 3.4), in order to ensure that our calculations were based on fully equilibrated MD runs.

On the one hand, the two main contributions to the total energy are the kinetic and the potential terms, as usual. In a deeper level of classification, potential energy has six contributions, namely those of: (a) molecular bonds; (b) molecular, dihedral and improper angular terms; (c) Van der Waals and (d) electrostatic (Coulomb) terms.

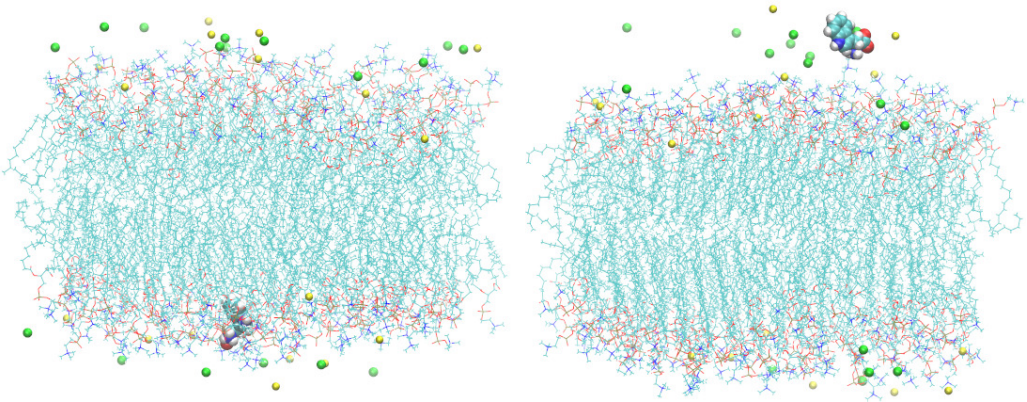


Fig. 3.2 Two snapshots of the aqueous DPPC membrane with an embedded TRP molecule. TRP inside DPPC (left) and inside water (right). Surrounding water is not shown for the sake of clarity. Atoms in TRP: Oxygen (red); hydrogen (white); carbon (cyan); nitrogen (blue). Sodium ions (yellow) and chloride ions (green). DPPC are depicted as lines.

Among all these terms, the electrostatic one is the largest by far, as expected. All contributions of the energy are shown in the last 10 ns of the simulation run, revealing stable profiles with fluctuations up to 3% of the average values. On the other hand, we found some fluctuations smaller than 2% for the temperature and of about 25 bar for the pressure. The mean pressure is zero, indicating that the bilayer is not affected by external or internal neat forces. From experimental and theoretical works, the expected surface tension of a lipid bilayer membrane should be zero[270–274], what it is consistent with the fact that the thermodynamic pressure of our system fluctuates around zero.

At the equilibration stage of the simulation, we placed TRP initially at the center of the membrane. After a few nanoseconds, we observed that TRP moved quickly to the water-DPPC interface and stayed there in close contact with the lipid headgroups. The most stable configurations of TRP at the interface and when solvated by water will be discussed below.

3.1.2.2 Structure of TRP around water and lipids

The local structure of the system can be analyzed by means of normalized atomic radial distribution functions (RDF) $g_{12}(r)$. For a species '2' close to a tagged species '1', they are given by

$$g_{12}(r) = \frac{V \langle n_2(r) \rangle}{4 N_2 \pi r^2 \Delta r}, \tag{3.1}$$

where $n_2(r)$ is the number of atoms of species '2' surrounding a given atom of species '1' inside a spherical shell of width $\Delta r = 0.1$ Å. V stands for the total volume

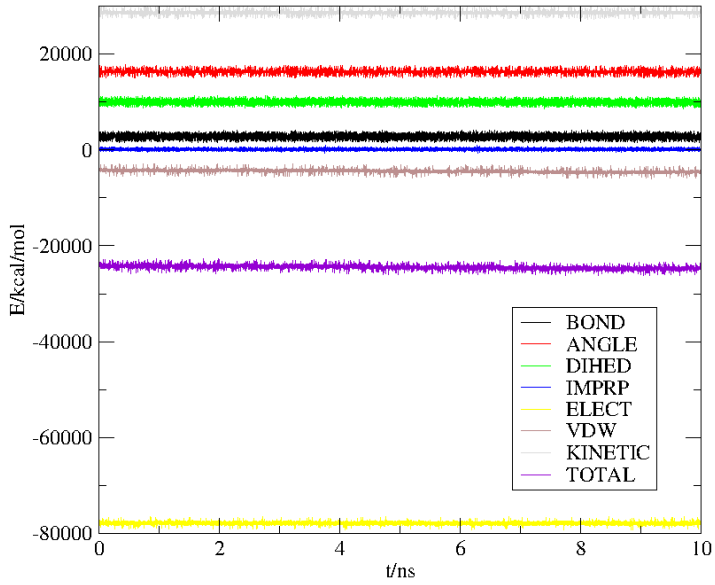


Fig. 3.3 Energy decomposition for the full system.

and N_2 is the total number of particles of species '2'. The four $g(r)$s considered were defined for the pairs:

1. The hydrogen of TRP ('HT', corresponding to the three hydrogens labeled 'H1' which represents the three hydrogen atoms attached to 'N1' atom, see Fig. 3.1) *versus* the oxygen of a water molecule ('OW');

2. The oxygens of TRP ('OT', corresponding to the sites labeled 'O1' and 'O2' of TRP, see Fig. 3.1) *versus* the hydrogen of a water molecule ('HW');

3. The hydrogen of TRP 'HT' *versus* the negatively charged oxygen of DPPC ('O^-'), located at the lipid headgroups (label 'O2' in Fig. 3.1);

4. The oxygens of TRP 'OT' *versus* the charged nitrogen of DPPC ('N^+'), located at the lipid's heads (see Fig. 3.1).

The results are shown in Fig. 3.5. The four radial distribution functions show some fluctuations in their profiles, the statistical noise observed in the association of TRP with N^+ of DPPC being more marked than for TRP-water and for TRP-O^- of DPPC. We can observe a first coordination shell in all cases, being the binding of TRP to water the one with the highest peaks of the corresponding $g(r)$ (those depicted in the left column of Fig. 3.5). In such a case, a second coordination shell can be also observed. Throughout our long simulation runs we observed periods of time of about 5 ns where TRP was fully solvated by water, indicating that the hydration of TRP is one stable state of the system, i.e. TRP can be fully hydrated for significant

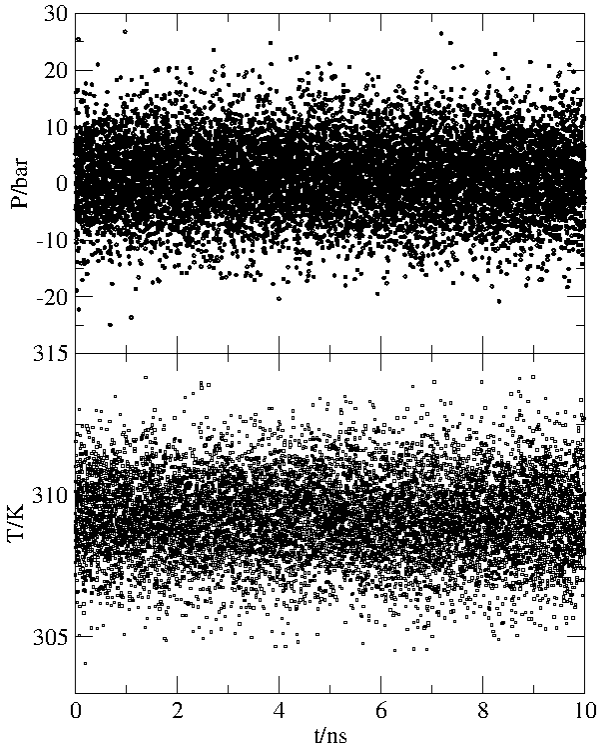


Fig. 3.4 Temperature and pressure fluctuations for the full system.

periods of time, essentially through hydrogen-bonding between the hydrogens of TRP (labeled as 'H1', 'H2', and 'H3' in Fig. 3.1 and the water's oxygen. If we focus on the TRP-DPPC binding, we can observe that the corresponding radial distribution functions $g_{HT-O^-}(r)$ and $g_{OT-N^+}(r)$ (depicted in the right column of Fig. 3.5) show marked first coordination shells and very smooth second shells. This fact indicates that the solvation of TRP by DPPC is carried out by a few lipid chains. From our simulation runs we observed that the average time of TRP inside the DPPC bilayer is of the order of 10-15 ns. The structure of water around DPPC is very similar to that of water around DMPC as it was described in the previous study[17, 275].

3.1.2.3 Hydrogen bonding of TRP with water and lipids

We can use a geometrical definition of a hydrogen bond, in the fashion as it is usually assumed in most computer simulations of water and associated liquids (see for instance a detailed study in Ref. [276]). It consists in considering that a hydrogen-bond is formed between two molecules when the next two geometrical conditions are fulfilled:

1. the distance R_{AH} between the "acceptor" molecule 'A' and the hydrogen "donor" atom 'H' is smaller than R^c_{AH} ,

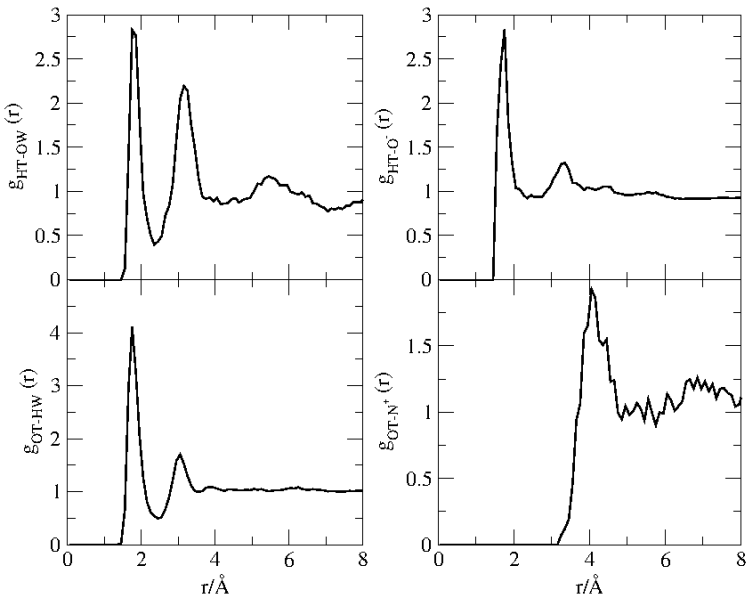


Fig. 3.5 Selected radial distribution functions for TRP (T) with water(W) and DPPC (charged sites 'N' and 'O2', see Fig. 3.1): OT-HW (bottom left), HT-OW (top left), OT-N^+ (bottom right) and HT-O^- (top right).

2. the H-B$\cdots$A angle φ is lower than φ^c. Here 'B' is the atom to which 'H' is chemically bound.

In our study, the threshold distances were taken from standard values in the literature. So, $R^c_{\mathrm{AH}} = 3.0$ Å for all types of hydrogen-bonds (HB) and the angular cutoff was chosen to be $\varphi^c = 30°$.

With this HB definition, we computed the average number of HBs between several bonded pairs. Our results indicate that we had (on average) a number of 3.2 water-water HBs per water molecule; around 1.9 HB for DPPC-water; about 2 HB between TRP and DPPC and ≈ 5 HBs for water-TRP. Furthermore, we can extract additional information from the radial distribution functions presented in Fig. 3.5. There we can observe a sharp maximum located around 1.85 Å in three cases: HT-OW, OT-HW, and HT-O^-. Since such distance is the signature of a typical oxygen-hydrogen HB, we can safely assume that HBs have been found between TRP and water as well as between TRP and the oxygen labeled 'O2' (see Fig. 3.1). Given the simulation data collected, we did find clear HB-signature peaks for TRP with oxygens 'O6' and 'O8' of DPPC as well, but to a lesser extent than for oxygen 'O2', which is the preferred site for TRP-DPPC binding. This would indicate that TRP is also able to stay bound to the inner part of lipid heads during significant periods of time.

3.1.2.4 Potentials of mean force for TRP solvation

Once the local structure around TRP has been obtained, a common way to analyze the microscopic forces acting on it is by means of the so-called potential of mean force between species '1' and '2', namely $PMF_{12}(r)$, that can be readily obtained from the pair (atom-atom) radial distribution function $g_{12}(r)$ given in Eq. 3.1:

$$PMF_{12}(r) = W(r_{12}) = -k_B T \ \ln g_{12}(r), \tag{3.2}$$

where k_B is Boltzmann constant, T is the temperature of the system, and $W(r_{12})$ is the reversible work. This remarkable theorem can be proved by considering the average force between two particles '1' and '2' inside the remaining system (particles '3', '4',...'N' being the solvent). The average (over all configurations) force between particles '1' and '2' fixed at their corresponding positions is given by:

$$\begin{aligned}
-\left\langle \frac{dU(r^N)}{d\vec{r}_1} \right\rangle_{\vec{r}_1,\vec{r}_2} &= \frac{-\int d\vec{r}_3 \cdots d\vec{r}_N (\frac{dU}{d\vec{r}_1})\ e^{-\beta U}}{\int d\vec{r}_3 \cdots d\vec{r}_N\ e^{-\beta U}} \\
&= +k_B T \frac{\frac{d}{d\vec{r}_1} \int d\vec{r}_3 \cdots d\vec{r}_N\ e^{-\beta U}}{\int d\vec{r}_3 \cdots d\vec{r}_N\ e^{-\beta U}} \\
&= k_B T \frac{d}{d\vec{r}_1} \ln \int d\vec{r}_3 \cdots d\vec{r}_N\ e^{-\beta U} \\
&= k_B T \frac{d}{d\vec{r}_1} \ln \left[(N(N-1) \left(\frac{\int d\vec{r}_3 \cdots d\vec{r}_N\ e^{-\beta U}}{\int d\vec{r}_1 \cdots d\vec{r}_N e^{-\beta U}} \right) \right] \\
&= k_B T \frac{d}{d\vec{r}_1} \ln\ g(\vec{r}_1, \vec{r}_2),
\end{aligned} \tag{3.3}$$

where $\beta = 1/(k_B T)$. Also, the definition of the radial distribution function from the statistical point of view has been assumed (see Ref. [277]).

Integration of the expression for the average force, introduced by Eq. 3.3, gives the reversible work,

$$W(r_{12}) = \int_{r_{12}}^{\infty} dr_1 (-\left\langle \frac{dV(r^N)}{dr_1} \right\rangle) = k_B T \int_{r_{12}}^{\infty} dr_1 \frac{d}{dr_1}\ \ln g(r_1), \tag{3.4}$$

Therefore, the reversible work $W(r_{12})$ associated with the process by which particles 1 and 2 are moved from infinite separation to a relative separation r_{12} is:

$$W(r_{12}) = k_B T\ \ln g(\infty) - k_B T\ \ln g(r_{12}) = -k_B T\ \ln g(r_{12}), \tag{3.5}$$

since $g(\infty) = 1$. In this Thesis, PMFs computed as the reversible work can be directly considered as the Gibbs free energy differences, considering r_{12} as the reaction coordinate[277].

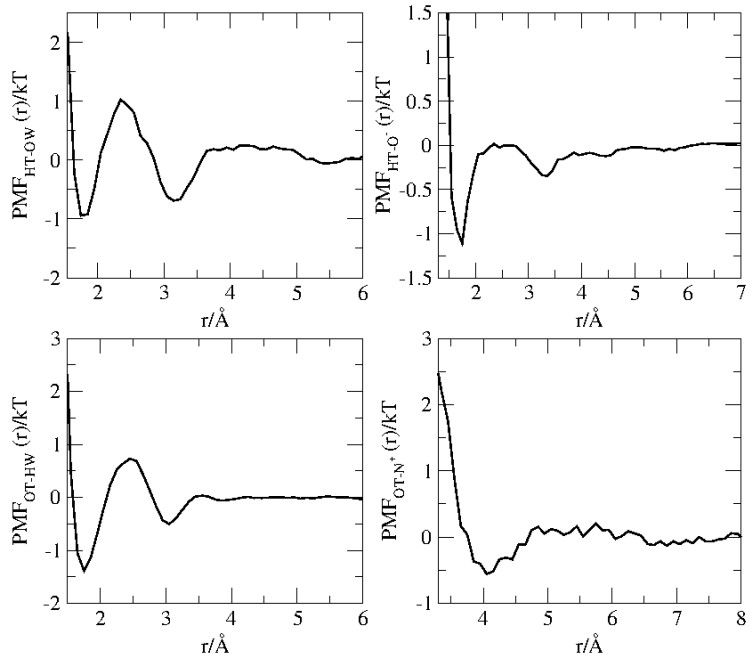


Fig. 3.6 Potentials of mean force for water-TRP and DPPC-TRP sites. Labels as in Fig. 3.5.

The use of a variety of methods to compute the PMF has been extensively discussed in the literature, as it was reported for instance in Ref. [278], where up to twelve methods based on one-dimensional coordinates were applied to the benchmark case of a methane pair in aqueous solution. The authors concluded that the best choice is a constraint-bias simulation combined with force averaging for Cartesian or internal degrees of freedom. The results from unbiased simulations, as those reported in the present work, were considered good at the qualitative level, with the PMF reasonably well reproduced.

The use of one-dimensional reaction coordinates is simply an approximation to the real ones[279], which may be in general multidimensional, presumably involving a limited number of water molecules. A method that does not assume any preconceived reaction coordinate is the so-called transition path sampling[280], with a high computational cost. So, since the determination of the true reaction coordinate for the adsorption of TRP is out of the scope of this Chapter, we will consider radial distances between two species as order parameters useful to account for mean forces between them, as a reasonably good approach.

The results of our calculation are displayed in Fig. 3.6. With the same order as in Fig. 3.5, we show PMF for TRP-water at the plots in the left column and those for TRP-DPPC in plots at the right column of Fig. 3.6. A free energy barrier is seen in

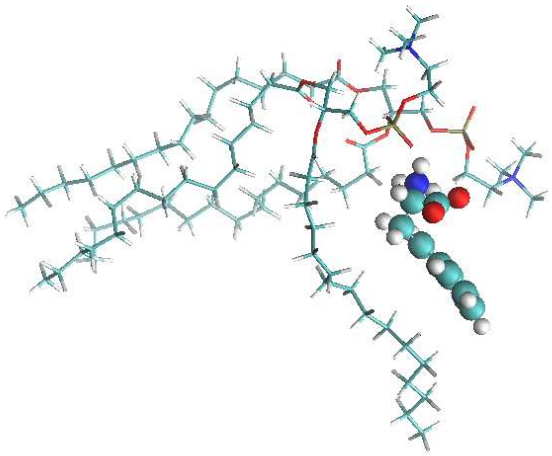


Fig. 3.7 Snapshot of a typical TRP-DPPC bond. Note that HB exits between sites of H1 of TRP and O8 of DPPC.

all cases, defined by a neat first minimum and a second minimum less clearly defined, especially for TRP-DPPC interactions. In order to quantify the height of all barriers, we included a numerical estimation of them in Table 3.1.

Table 3.1 Free energy barriers for the binding of TRP to water and DPPC. ΔF is the size (in k_BT) of the free energy barrier and $\left\langle r_{1(2)} \right\rangle$ are the position of the first (second) minima of PMF.

Pair	ΔF (k_BT)	$\langle r_1 \rangle$ (Å)	$\langle r_2 \rangle$ (Å)
HT/OW	2.0	1.8	3.2
OT/HW	2.0	1.7	3.0
HT/O^--DPPC	1.0	1.8	3.3
OT/N^+-DPPC	0.7	4.0	$\approx$ 7.0

Here we should note that in our system k_BT $\sim$ 0.616 kcal/mol so that the physical values for the free energy barriers are of between 0.4 and 1.2 kcal/mol, i.e. the same order of magnitude of the free energies of adsorption of metal ions in DMPC membranes[275]. From the data reported in Table 3.1 we observe that the highest barrier corresponds to the pairing of TRP and water and that the most stable (average) distance for water closest to TRP is of about 1.7-1.8 Å , i.e. of the order of the typical HB distance, as pointed out above in section 3.1.2.3. Conversely, the binding of TRP to DPPC happens with lower free energy barriers and at different distances between TRP and the binding site in DPPC. These data suggest that the configurational cost

of the association of TRP to a lipid molecule is lower than that of the association to water, what would explain why TRP tends to stay close to DPPC during periods of time about three-fold longer than the mean time remaining inside the water region of the system. To illustrate such association, we included a typical snapshot of this pair in Fig. 3.7.

3.2 Effects of cholesterol on the binding of tryptophan to zwitterionic membranes

Cholesterol plays a central role in maintaining the structure of the membrane and regulating their functions. A better and detailed study of zwitterionic DPPC bilayer membranes can help to understand their basic biological functions and interactions with the environment. We have considered the introduction of cholesterol into the lipid bilayer structure of TRP. TRP is able to act as a building block in protein biosynthesis, while proteins are fundamentals required to sustain life. Given the importance of TRP, we have explored its interactions with a typical zwitterionic phospholipid membrane formed by DPPC and cholesterol in an aqueous solution of sodium chloride, analyzing its mechanical properties, the local structure and some relevant dynamical properties such as diffusion and vibrational spectra.

3.2.1 Preparing systems for simulations

Similar to section 3.1, all systems were generated through the CHARMM-GUI web-based tool. The membrane consisted of a TRP molecule, 204 DPPC lipids, 4962 TIP3P water (W) molecules, with 17 sodium and 17 chlorine ions, corresponding to physiological concentration (0.15 M). Eventually, some of the lipids were replaced by cholesterol molecules in such a way that two cholesterol percentages were considered: 30 and 50%. Previous simulations with percentages of 10, 20 and 40% showed no further relevant physical changes compared to the reported states. A benchmark state of a cholesterol-free system (labeled as 0% throughout the paper) was also considered at the temperature of 323.15 K, in order to ensure that the simulation was performed at the liquid state, since the crossover temperature for pure DPPC to become liquid has been determined to be around 314 K[281, 282].

Sketches of the backbone structure of cholesterol, TRP, and DPPC are represented in Fig. 3.1. MD simulations in section 3.1 were performed at the fixed pressure of 1 atm and a fixed temperature of 310.15 K. The reason for considering 323.15 K for all systems in this section is to have the system always in liquid-crystalline phases as close as possible to the human body temperature: liquid-disordered for the cholesterol-free case and liquid-ordered (with headgroups showing liquid-like behavior, but with tail groups showing some gel-like state, with an ordering structure larger than that of

headgroups) for the setups containing 30 and 50% cholesterol, when the crossover temperature from gel to the liquid phase is below 310 K (see Ref. [281]). In all cases, the temperature was controlled by a Langevin thermostat[267] with a damping coefficient of 1 ps^{-1}, whereas the pressure was controlled by a Nosé-Hoover Langevin barostat[268] with a damping time of 50 fs.

Considering the NPT ensembles for equilibration periods for all simulations ran for more than 50 ns. After equilibration, we recorded statistically meaningful trajectories of more than 70 ns in all cases through several production runs. The simulation boxes had different sizes because of cholesterol concentrations. For instance, in the cholesterol-free system the average size was 80 Å× 80 Å× 90 Å. As cholesterol concentration increased, the size of the system was systematically reduced. The simulation time step was set to 2 fs. The CHARMM36 force field was adopted for all simulations. All bonds involving hydrogens were set to fixed lengths, allowing fluctuations of bond distances and angles for the remaining atoms. Nevertheless, during the calculation of spectral densities (see section 3.2.2.4) all bonds including those involving hydrogens were left flexible. Van der Waals interactions were cut off at 12 Å with a smooth switching function starting at 10 Å. Long ranged electrostatic forces were computed using the particle mesh Ewald method[266], with a grid space of about 1 Å. Such electrostatic interactions were updated every time step. Finally, periodic boundary conditions were applied in all three dimensions of space.

3.2.2 Results and discussion

3.2.2.1 General properties of all systems

A general view of half of the system at several cholesterol concentrations is shown in Fig. 3.8. There, for the sake of clarity, water and DPPC have been depicted with thin sticks, whereas cholesterol, TRP, and the sodium and chlorine ions have been shown more explicitly. We should point out that TRP tends to stay close to the headgroup regions of the lipid and cholesterol chains during most time of the simulated trajectories[283]: An estimation over the full trajectory length (cholesterol concentration of 30%) indicated that around 70% of time TRP is attached to DPPC with the remaining time being mostly solvated by water. Further, in the membrane with 30% of cholesterol, we observe that TRP at the interface is located slightly deeper than in the case with 50% of cholesterol so that a primary effect of cholesterol is a tendency to exclude TRP out of the water-membrane interface.

With the aim of ensuring that we are simulating the liquid phase of the model system as well as to efficiently characterize the order of the hydrated lipid bilayer, we computed the deuterium order parameter S_{CD} which can be obtained from ^{2}H NMR experiments. This quantity was first reported by Stockton and Smith[284] and later

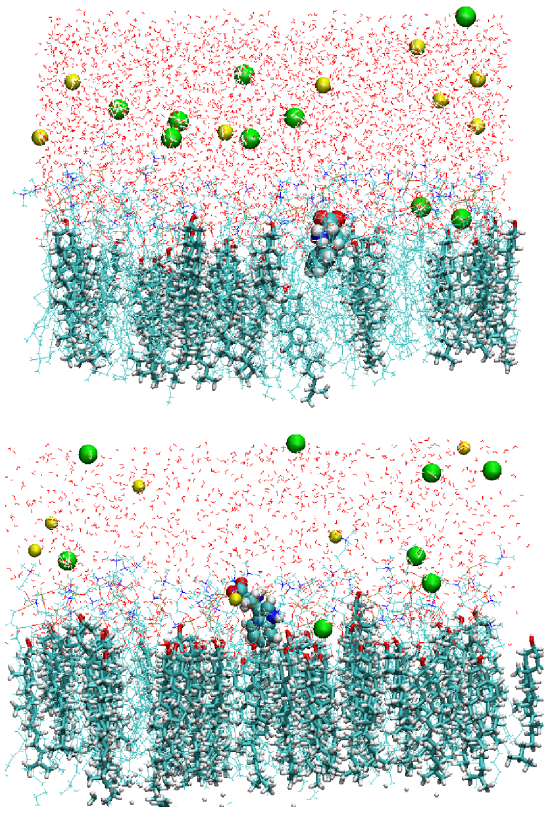


Fig. 3.8 Snapshots of the aqueous DPPC and cholesterol membranes with an embedded TRP molecule (Top: 30% of cholesterol; bottom: 50% of cholesterol). Atoms in TRP: oxygen (red), hydrogen (white), carbon (cyan) and nitrogen (blue). DPPC and water molecules are shown as thin sticks. Cholesterol are depicted as thick sticks. Sodium ions (green) and chlorine ions (yellow).

on adopted by Hofsäss et al.[285] by means of an order parameter, defined for each CH_2 group as follows:

$$S_{CD} = \frac{1}{2}(3 < \cos^2 \theta_{CD} > -1), \tag{3.6}$$

where θ_{CD} is the angle between the membrane normal and a CH-bond (a CD-bond in the experiments). Brackets in Eq. 3.6 indicate an ensemble average for all lipids. The results are shown in Fig. 3.9 for the two chains (sn1, sn2) of a DPPC molecule at three cholesterol concentrations. The results are in good overall agreement with both experimental[98, 286] and simulation works[285, 287] and confirm that the system, in the simulation conditions assumed in the present work, represents well the liquid-crystalline phase. We should note the tendency to higher ordering as cholesterol concentration increases, what is represented by profiles of S_{CD} having larger maxima (around 0.2 for the cholesterol-free system *versus* ∼0.4 for the cholesterol-rich setups), a similar tendency as observed by Hofsäss et al.[285] and Chiu et al.[20].

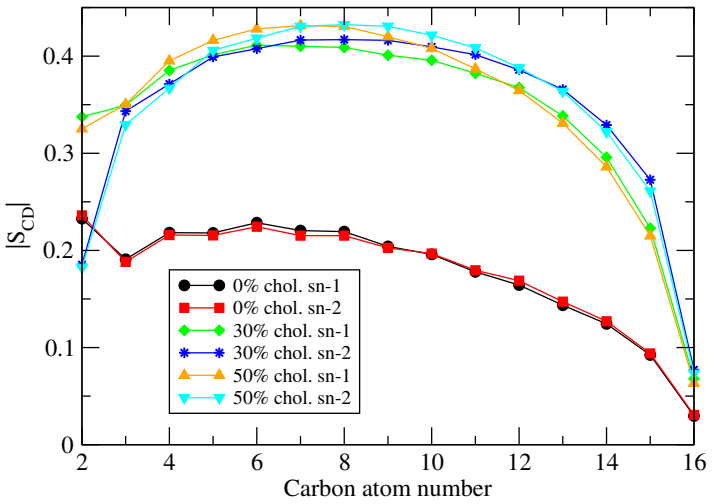


Fig. 3.9 Order parameter $|S_{CD}|$ for the acyl tails for three different percentages of cholesterol.

One primary output from any simulation of biomembranes is the area per lipid. For a series of equilibrated MD runs at the NPT ensemble we have monitored the surface area per lipid considering the total surface along the XY plane (plane parallel to the bilayer surface) divided by the number of lipids plus cholesterol in one lamellar layer[288], as defined in Eq. 3.7.

$$A = \frac{L_x \times L_y}{N_l + N_{chol}} \tag{3.7}$$

where A is the area per lipid of the simulation box, L_x and L_y are the length of the simulation box along x axis and y axis, and N_l and N_c are the numbers of lipids and cholesterol in each leaflet, respectively.

The area per lipid (for the final 70 ns of all trajectories) as a function of simulation time is shown in Fig. 3.10 and the average values are reported in Table 3.2. It should be pointed out that these values for the area per lipid arise naturally from the relaxation of the system at a given temperature, pressure, and number of particles and are not an imposition to fit the experimental value. The main trend observed from our data is that the areas decrease as cholesterol concentration increases, as expected. In addition, fluctuations of the area per lipid in the 0% cholesterol case are larger than those of the cholesterol-rich ensembles. We obtained an average value of 61.4 Å^2 for the 0% cholesterol system and smaller values around to 40 Å^2 when cholesterol is present at the membrane. These results are in overall good agreement with other computational data in a wide variety of thermodynamical conditions[20, 98, 285, 289, 290] where the values for pure DPPC ranged between 50 and 63 Å^2 and the trend of decreasing areas for increasing cholesterol percentages is clearly reported. From the experimental side, an influential review from Nagle et al.[6] reported values of the area per lipid of pure DPPC membranes obtained from a wide variety of methods (NMR, X-ray, and neutron

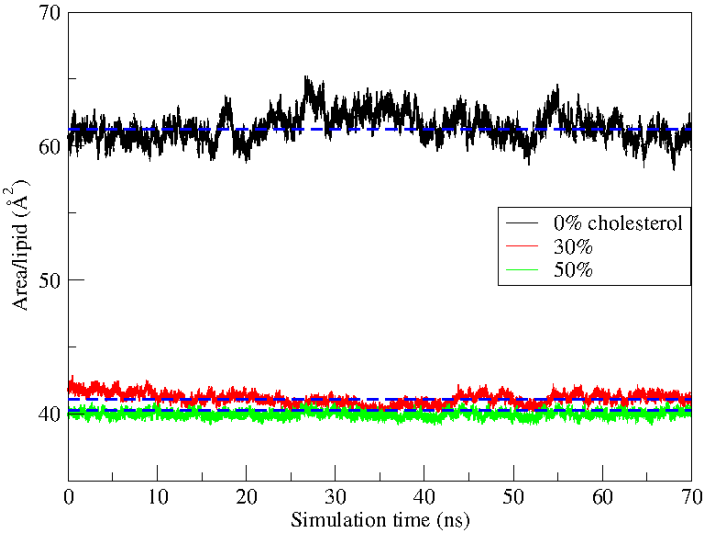


Fig. 3.10 Area per lipid versus cholesterol contents of the system: 0%, 30% and 50%. Dashed lines indicate average values reported in Table 3.2.

scattering) between 48 and 52 Å^2 in the gel phase (293 K) and between 57 and 71 Å^2 at the liquid phase (323 K). However, some of them were reported to be measured under wrong assumptions due to artificial undulations of the membrane sets. The best estimations were of 48 and 64 Å^2 for the gel phase and liquid phase respectively, in the cholesterol-free case. In one quite recent work, Kučerka et al.[282] found a value of 63.1 Å^2 for DPPC at 323.15 K by means of X-ray and neutron scattering techniques. In the present work, a sudden drop in the area per lipid is dramatically seen when cholesterol is introduced in the membrane, as expected. Edholm and Nagle[289] reported a summary of values from independent MD simulations where at the cholesterol concentration of 30% area per lipid were around 43 Å^2 and at 50% were between 37 and 39 Å^2, at 323 K in all cases. As a matter of fact, the agreement with our results is qualitatively good. Further, the results reported here are consistent with the observation that, in DMPC membranes analyzed at 303 K, cholesterol is able to favor a phase transition from a liquid-disordered phase to a liquid-ordered phase[291, 292], in particular when cholesterol contents are higher than 30%.

Table 3.2 Mechanical properties of the membrane: area per lipid (A) and thickness (Δz). Estimated errors are in parenthesis.

% of cholesterol	A (Å^2)	Δz (Å)
0	61.4(0.8)	39.7(0.5)
30	40.8(0.2)	48.9(0.4)
50	40.1(0.2)	47.8(0.3)

The variations of the thickness of the membrane may give some additional clues of the effect of cholesterol on its mechanical properties, such as its rigidity and its capability to allow the passing of species inside and outside the cell. We have obtained the thickness by computing the distance between the phosphate groups of the DPPC molecules located at the two sides of the membrane. The results are also reported in Table 3.2.

These values of the bilayer thickness are in good overall agreement with those reported by Kučerka et al.[282]. These authors found a Δz of 39 Å at 323 K for a cholesterol-free sample and a tendency to decrease with increasing temperatures. In the present work, we have obtained a value of 39.7 Å at the cholesterol-free system and values around 48 Å when cholesterol is present in the membrane. This fact is consistent with the findings of the area per lipid reported above where, at higher cholesterol contents, the values have been found smaller: at large cholesterol contents the bilayer structures are more compressed so that we have more lipid species per $Å^2$ what would eventually produce the extension of the lipid tails (which are normally showing a large extent of folding for the cholesterol-free system) and the corresponding larger bilayer thickness. This is also consistent with the results shown in Fig. 3.9 where cholesterol-rich systems have tails significantly more ordered (gel-like) than those corresponding to the cholesterol-free case.

3.2.2.2 Structure of the solvated TRP

We have evaluated the local structure of the TRP molecule solvated by water, DPPC, and cholesterol. We will restrict ourselves to report the most relevant *RDF* given in Eq. 3.1 which is also indicated as $g(r)$ in the rest of the text, namely those defined for the pairs:

1. The hydrogens 'H1' of TRP (see Fig. 3.1) *versus* the oxygen of a water molecule ('OW');
2. The oxygens of TRP ('OT', corresponding to the sites labeled 'O1' and 'O2', see Fig. 3.1) *versus* the hydrogen of a water molecule ('HW');
3. The hydrogens 'H1' of TRP *versus* the oxygen of cholesterol ('OC');
4. The oxygens of TRP 'OT' *versus* the hydrogen of cholesterol ('HC');
5. The hydrogens of TRP 'H1' and 'H2' (taken separately) *versus* the negatively charged oxygens of DPPC ('O2' and 'O8', taken separately).

The results of the $g(r)$ defined above are presented in Fig. 3.11 and Fig. 3.12 for the three relevant TRP concentrations (0, 30, 50%).

All eight radial distribution functions show some fluctuations in their profiles, the statistical noise observed in the association of TRP with cholesterol being more marked

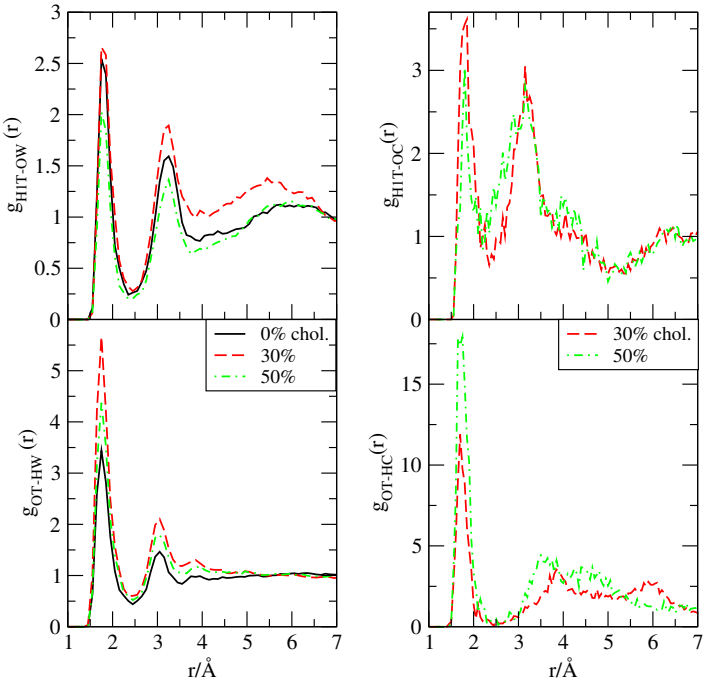


Fig. 3.11 Radial distribution functions for TRP (T) with water(W) and cholesterol (sites 'O' and 'H' of the hydroxyl group): H1T-OW (top left), OT-HW (bottom left), H1T-OC (top right) and OT-HC (bottom right).

than for TRP-Water and for TRP-DPPC. The fluctuations are essentially due to the fact that TRP is a mobile species and can be mostly solvated by DPPC or water. However, cholesterol solvation only happens during short periods of time (see below), which generates larger fluctuations of the corresponding $g(r)$s. As a general feature, the presence of cholesterol does not affect the positions of the maxima in the $g(r)$ between TRP and water and between TRP and DPPC, although significant reduction of the heights of the maxima has been observed, especially for g_{H1T-O2}(r). We can observe a clear first coordination shell in all cases, being the binding of TRP to DPPC the one with the highest peaks of the $g(r)$s (see Fig. 3.12). In such a case, a marked second coordination shell can be also observed. The peaks located around 1.7-1.8 Å in all $g(r)$ are the signature of hydrogen bonds (HB) established between TRP and water, TRP and DPPC, and also TRP with cholesterol. Using fluorescence spectroscopy, Liu et al.[100] obtained values for the hydrogen-bond lengths of TRP and water between 1.6 and 2.1 Å , i.e. in the same range of the values reported in the present work.

In the present work, we have not considered the usual density profiles defined by the distance of an atomic site to the center of the lipid bilayer, and instead, we computed radial distribution functions. The reason is twofold: (1) we are mainly interested in the solvation structure of TRP instead of their distribution across the bilayer and (2) RDF have the additional advantage to provide a direct route to the estimation of free energy barriers[277] but, since reaction coordinates for probe binding

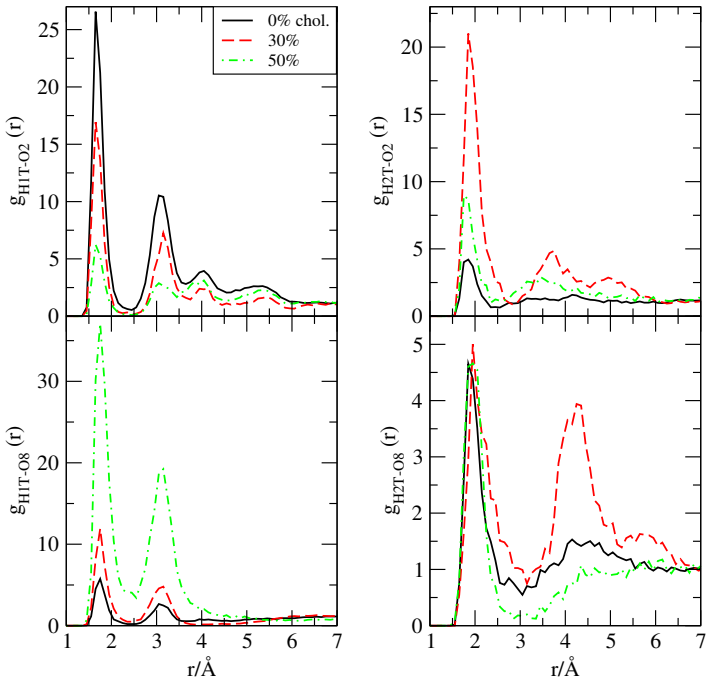


Fig. 3.12 Radial distribution functions for TRP (T) with DPPC (charged sites 'H1', 'H2', 'O2' and 'O8', see Fig. 3.1): H1T-O2 (top left), H1T-O8 (bottom left), H2T-O2 (top right) and H2T-O8 (bottom right).

to membranes will be generally unknown and not one-dimensional but multidimensional, methods not assuming preconceived reaction coordinates such as transition path sampling[213, 280, 293] or, those allowing to consider several complementary *collective variables*, such as metadynamics[231, 275] would be in order to obtain much more accurate free energy landscapes for TRP adsorption. Nevertheless, their drawback is the huge amount of computational time required in both cases and for this reason, they are not considered in the present Chapter.

During the long simulation runs used to collect statistically meaningful properties we observed periods of time of about 5 ns where TRP was fully solvated by water, suggesting that hydration of TRP is one stable state of the system. Full TRP hydration is essentially operated through hydrogen-bonding between the triplet of 'H1' hydrogens of TRP bound to 'N1' (see Fig. 3.1) and the water's oxygen or, equivalently, between oxygens of TRP and hydrogens in water, as it is shown in the left column of Fig. 3.11. There we can observe a sharp maximum located around 1.85 Å for both g_{H1T-OW} and g_{OT-HW}. When cholesterol is introduced in the systems, the binding of TRP to water remains essentially unchanged, although the first maxima of g_{OT-HW} shows larger values. Firstly, this could be due to the exclusion of TRP off the interface operated by the presence of cholesterol and, secondly, because of the binding of TRP to cholesterol as it is shown at the right column of Fig. 3.11. Interestingly, cholesterol can form hydrogen-bonds with TRP's hydrogens as indicated by the first maximum of g_{H1T-OC}

and also with TRP's oxygens (shown in g_{OT-HC}). The large statistical noise in those functions might be an indication of the short life of such HB.

Observing that TRP can establish some hydrogen-bonding with cholesterol, all data we got suggest that TRP is also able to stay bound to the inner part of lipid heads during significant periods of time, in agreement with previous results indicating that in a cholesterol-free DOPC bilayer membrane TRP tends to localize strongly in the interfacial region[89]. The preferential location of TRP at the lipid-water interface was observed by de Jesus and Allen[58] for membrane proteins. In the present work, we observed that the average lifetime of TRP at the interface of the DPPC bilayer is of the order of 10 ns. Regarding TRP-DPPC binding (see Fig. 3.12) we have observed hydrogen-bonding connections between sites 'H1' and 'H2' of TRP and DPPC sites 'O2' and 'O8' (labels according Fig. 3.1), as indicated by radial distribution functions g_{H1T-O2}, g_{H1T-O8}, g_{H2T-O2} and g_{H2T-O8}. Functions related to 'H1' hydrogens in TRP (left column of Fig. 3.12) show a clear maximum located around 1.7 Å when bound to 'O2' of DPPC and around 1.75 Å when bound to 'O8' of DPPC. The presence of cholesterol diminishes the 'H1-O2' binding but it enhances the 'H1-O8' one. This suggests that the former corresponds to the most stable bond when no cholesterol in present but that becomes weaker when the amount of cholesterol increases. Conversely, the HB between 'H1' and 'O8' becomes strongest at the membrane with 50% content of cholesterol. Then we can observe the influence of the cholesterol in the TRP-DPPC binding as one major effect. Concerning 'H2'-DPPC binding (right column of Fig. 3.12), we observed its existence in all analyzed setups, although the maxima are found at slightly larger distances (around 1.9-2.0 Å). In the same fashion as for 'H1'-DPPC binding, the presence of cholesterol showed a significant influence on the characteristics of 'H2'-DPPC bonds as well. Finally, the structure of water around DPPC has been found to be very similar to that of water around DMPC as was previously described in Ref. [17], Section III.

In order to estimate the average time intervals for the TRP-lipid association, we display the time evolution of selected atom-atom distances in Fig. 3.13. There we can see that typical hydrogen-bonding distances are reached in all cases. At the cholesterol-free membrane, one of the three hydrogens of TRP bound to its 'N1' site is almost permanently bound to the oxygen 'O2' of DPPC with some intervals of simultaneous binding ("shared" hydrogen-bonds). At the two cholesterol concentrations considered here the binding is less common, with a tendency to longer hydrogen-bonding periods at the 30% concentration. In all cases, the typical H-bond lifetimes are of the order 1 ns, with eventual intervals of breaking and re-formation of hydrogen-bonds.

When analyzing the hydrogen-bonding of TRP with cholesterol, we have observed that some periods of hydrogen-bonding are established between oxygen 'O2' of TRP and the hydroxyl's hydrogen of cholesterol ('HC') and that, simultaneously, one of

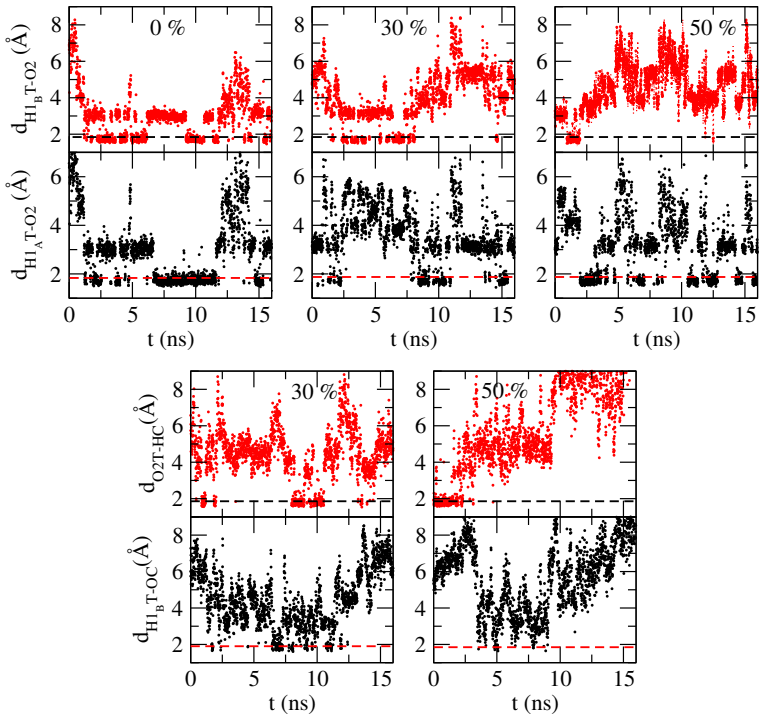


Fig. 3.13 Distance distribution of selected sites in cholesterol-TRP bonding and DPPC-TRP bonding as a function of simulation time. Percentages of cholesterol are indicated on top of each graph. Dashed lines indicate the typical HB distance (1.8 Å). 'H1$_A$T' and 'H1$_B$T' are two of the three hydrogens bound to the site 'N1' of TRP, whereas 'O2' is one of two oxygens of TRP and 'OC' ('HC') is the hydroxyl's oxygen (hydrogen) of cholesterol.

the TRP's hydrogens (bound to 'N1') forms some HBs with hydroxyl's oxygen of cholesterol. Again we see longer binding for cholesterol's 30% concentration. The time scale of hydrogen-bonding seems shorter than 1 ns, although more precise calculations of hydrogen-bond lifetimes using correlation functions (see Ref. [294] for instance), are out of the scope of this Chapter and have not been carried out.

In order to have a more detailed idea of TRP binding to the membrane, we are reporting two characteristic snapshots of TRP linked to DPPC and cholesterol simultaneously (see Fig. 3.14). There we can observe that either TRP can be bound to two DPPC in different sites ('O2' and 'O8') and one cholesterol (right) either it can be bound to one DPPC (site 'O2') and to two cholesterol molecules, the first through the hydroxyl's hydrogen and the second through hydroxyl's oxygen (left). These images help to enlighten the relatively complex multiple hydrogen-bonding connections between TRP, DPPC, and cholesterol described above and the key role of TRP as a bridging molecule between several lipids or cholesterol chains.

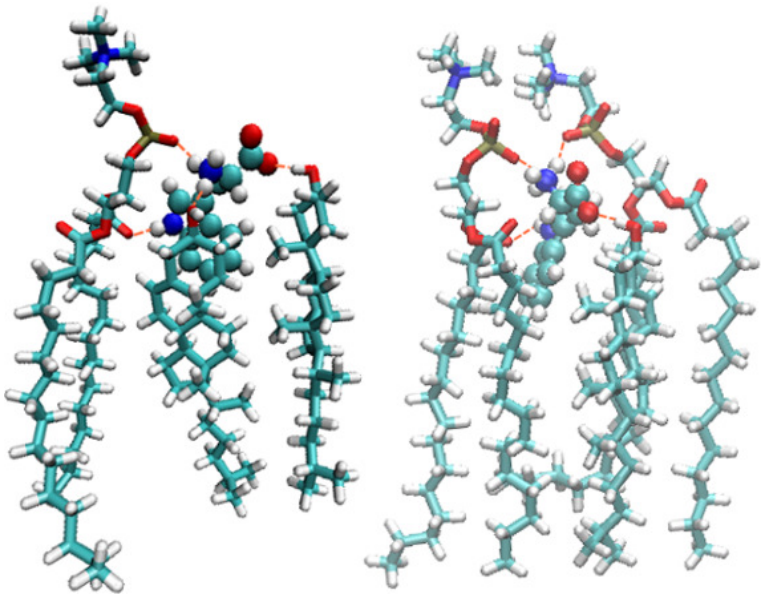


Fig. 3.14 Snapshots of typical DPPC-TRP-cholesterol bonds. Atoms in TRP colored as in Fig. 3.8. Binding sites have been highlighted in dashed orange lines.

3.2.2.3 Diffusion of TRP and water

Microscopic dynamics of TRP and water has been extensively considered. The first property evaluated has been the mean square displacement (MSD) of water (monitoring oxygen's trajectories) and of the carbon 'C2' in TRP (see Fig. 3.1). Diffusion coefficients of DPPC require simulations at the scale of hundreds of nanoseconds and have not been computed here, although a detailed study on DMPC was reported earlier[17]. From the long-time slopes of both MSD, we obtained the corresponding self-diffusion coefficients D through Einstein's formula:

$$D = \lim_{t\to\infty} \frac{< |\vec{r}_i(t) - \vec{r}_i(0)|^2 >}{2\,d\,\Delta t}, \tag{3.8}$$

where $\vec{r}_i$ is the instantaneous position of particle i. In this general procedure, the spatial dimension d is considered according to the type of diffusion. So, water undertakes isotropical diffusion ($d = 3$) meanwhile TRP showed lateral-like diffusion ($d = 2$). The results have been summarized in Table 3.3.

Table 3.3 Self-diffusion coefficients D (in cm^2/s) of water and of TRP as a function of the cholesterol's percentage of the membrane.

D	0%	30%	50%
Water	$5.03(0.01)\times10^{-5}$	$4.20(0.01)\times10^{-5}$	$4.45(0.07)\times10^{-5}$
TRP	$3.48(0.80)\times10^{-7}$	$2.91(0.35)\times10^{-7}$	$1.40(0.02)\times10^{-6}$

In the case of water, the results include all molecules in the system regardless of their location in the bulk or interfacial regions. In a previous work[17], we obtained D

separately for bound water, water at interfaces and water external to the membrane and found the mobility of water molecules being significantly faster as their distance to the interface increased (the water outside the membrane was about 14 times faster than water bound to lipid chains). In this work, we simply included the water diffusion for the sake of comparison with those previous results in order to test the influence of TRP and of cholesterol. At 0% cholesterol our results are of the same order of magnitude as the ones obtained for the simple aqueous DMPC membrane[17] at 303 K (2.66×10^{-5} cm^2/s) what indicates that, as expected, the presence of TRP does not affect significantly water's dynamics. Nevertheless, we can observe that when cholesterol is included in the system, water tends to diffuse about 33% faster. This can be attributed to the fact that at non-zero cholesterol concentrations, membranes are more compact (areas per lipid are smaller) and water's penetration in the interface is harder to occur, producing diffusion mainly along the surface of the bilayer.

TRP's diffusion coefficients were computed from simulation runs where TRP was permanently at the interface. We stress this point because in our simulation runs we observed that TRP is able to make excursions to the water bulk regions, especially when cholesterol concentration increases. For instance, at the highest concentration of 50%, TRP will stay eventually half of the time at the water bulk and the remaining time at the lipid-water interface. Along the periods when TRP was fully solvated by water, we estimated that its diffusion coefficient was 3.9×10^{-6} cm^2/s at 50% concentration of cholesterol. The latter value is significantly higher than those corresponding to TRP at the lipid-water interfaces (see Table 3.3).

Focusing on the diffusion of TRP when residing at the interface, the values reported in Table 3.3 are about two orders of magnitude smaller than those of water (for cholesterol concentration up to 30%). At the highest concentration (50%), the diffusion of TRP becomes significantly faster. The values of D for TRP (0% cholesterol's concentration) are of the same order of magnitude of those of DMPC molecules[17] (0.6×10^{-7} cm^2/s in absence of cholesterol), although the former is about six times smaller. This indicates that TRP's mobility is significantly higher than that of DPPC and this suggests that the mechanisms of diffusion of TRP would be qualitatively different of those of lipids (it was suggested that they are due to the motion of nanodomains or local groups of around 10 units[17] and also directly related to the slow motion of gel-like tails in the present case). We should point out that the effect of temperature is remarkable since, at complementary simulations at 310 K, TRP diffusion was of about 2×10^{-7} cm^2/s, i.e. significantly slower, due to the essentially gel-like state of the membrane at such conditions.

Regarding the reliability of the present calculations, we should indicate that the values reported in Table 3.3 have been obtained in the framework of the CHARMM36 force field. Whereas the number of lipids in our systems may be large enough[295] to avoid strong side effects from periodic boundary conditions, there are some known

deficiencies of the CHARMM36 (and the MARTINI force field among others) regarding its reliability when computing lipid diffusion[296] to be considered. Further, Yeh and Hummer suggested that in simulations using periodic boundary conditions, hydrodynamic interactions should be taken into account in order to obtain more accurate values of D, in such a way that a correction of the order $N^{\frac{1}{3}}$ might be in order[297]. In our case, we did not consider lipid diffusion but TRP's. Since TRP dynamics is clearly faster than DPPC's, we expect that our computed D are correct in order of magnitude, although they might presumably be affected to some extent by the deficiencies of the force field mentioned above. Only the existence of precise experimental values could shed light on this issue, but up to our best knowledge, no measured diffusion coefficients of TRP in zwitterionic membranes are currently available.

3.2.2.4 Spectral densities of TRP

We have computed and analyzed the spectral densities of the TRP molecule. Experimental infrared spectra are usually obtained through the absorption coefficient $\alpha(\omega)$ or the imaginary part of the frequency-dependent dielectric constant[298]. These properties are directly related to the absorption lineshape $I(\omega)$, which can be obtained from MD simulations[299–301] in certain cases. In most cases the physically relevant property to be computed is the so-called atomic spectral density $S_i(\omega)$, defined as:

$$S_i(\omega) = \int_0^\infty \mathrm{d}t \ < \vec{v}_i(t)\vec{v}_i(0) > \cos(\omega t) \tag{3.9}$$

where $\vec{v}_i(t)$ is the velocity of atom i. In our case, we have obtained the spectral density of the whole TRP molecule as the sum of the contributions of all atoms pertaining to the molecule. In Eq. 3.9, the brackets $\langle \cdots \rangle$ denote equilibrium ensemble average. Generally speaking, classical MD simulations are not able to fully reproduce experimental absorption coefficients, these being quantum properties. However, they can be used to locate the position of the spectral bands since, in the harmonic (oscillator) approximation, classical and quantum fundamental frequencies are the same. The model employed in the present work considers harmonic bond vibrations and for such a reason we will be able to locate the positions of the main experimental spectral bands. We are reporting the full spectral density $S_{TRP}(\omega)$ on the left side of Fig. 3.15.

Our results reveal two main spectral regions, one for frequencies below 1800 cm^{-1} and another one for $2700 < \omega < 3400$ cm^{-1}. The agreement with available data from infrared and Raman spectroscopy of TRP in aqueous solution reported by Leyton et al.[99] (and references therein) is very good overall. These authors reported strong maxima located at 128, 163, 425 (weak), 744, 1014, (1351), 1432, 1566, (1647), 2938, 3056 and 3402 cm^{-1}. The numbers into parenthesis indicate a group of maxima centered at the reference value indicated here. Most of these maxima are included in the $S(\omega)$ reported in Fig. 3.15 and, at 0% cholesterol content, show excellent agreement

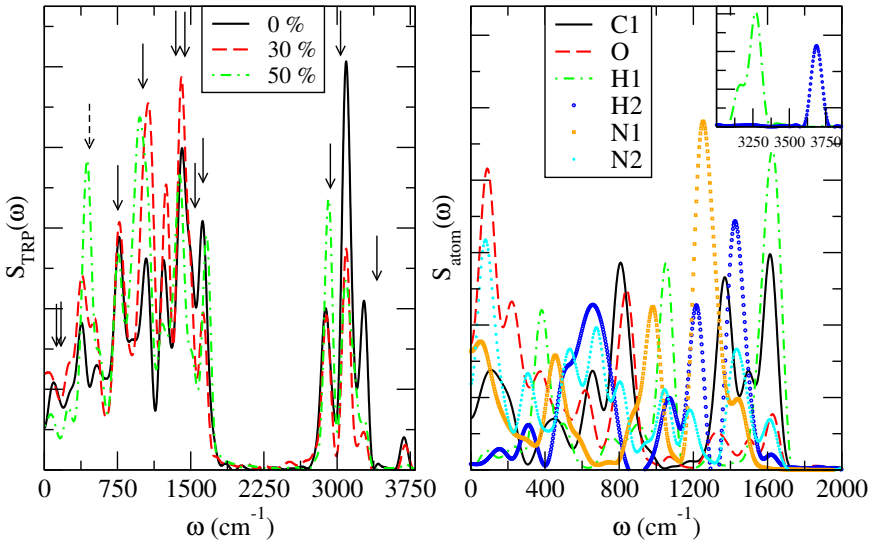


Fig. 3.15 Spectral densities of the full TRP molecule (left). Full arrows indicate the position of experimental peaks and the dashed arrow indicates a maximum of weak intensity. Decomposition of the spectrum for the 0% cholesterol case (right), indicates the relative contribution of several (most relevant) selected atoms. The inset reveals the contributions of high frequencies of the two relevant groups of hydrogen atoms ('H1' and 'H2').

with only three clear disagreements: (1) the Raman peak located at 163 cm^{-1} is not clearly seen in $S_{TRP}(\omega)$; (2) same fact happens with band located at 1566 cm^{-1} (unseen) and (3) the Raman peak at the highest frequency at 3402 cm^{-1} is strongly red-shifted in the computed spectral density (reported at 3276 cm^{-1}).

Having established the reliability of the potential model and method employed in the present work, we have obtained the partial spectra of each atom in TRP. The results are reported on the right side of Fig. 3.15. There we gain an advantage of one of the strongest points of MD simulations: we can discern the contribution of each atom to the full spectrum. We should remember that frequencies below 1000 cm^{-1} are usually related to rotational and vibrational motions of the molecule, those between 1000 and 1500 cm^{-1} are normally attributed to molecular bending motions and that frequencies of the order of 2000-3000 wavenumbers are due to molecular stretching motions (for liquid water, see, for instance[302]). The list of assignments of this work is as follows:

1. The peak experimentally found at 128 cm^{-1} corresponds, within our framework, to restricted translations of oxygens, since it corresponds to the high maximum (centered at 123 cm^{-1}) appearing at the oxygen's spectral density. Further, it can be also related to carbon vibrations (maxima at 95 and 210 cm^{-1}). In summary, we can attribute this band to 'C1-O' vibrations (see Fig. 3.1).

2. The peak located at 163 cm^{-1} is not seen in our spectral decomposition.

3. The weak band centered around 425 cm^{-1} is matched by the maxima at 385 cm^{-1} of our spectrum. This indicates that this vibration should be attributed to a low-frequency vibration of hydrogens 'H1' (see Fig. 3.1). Since this is a relatively low frequency it will probably correspond to rotational motion.

4. The maximum at 744 cm^{-1} is located around 810 cm^{-1} in Fig. 3.15 and it appears in both the contributions of atoms carbon 'C1' and oxygens 'O' (TRP has two of them), suggesting a "scissoring" motion of the corresponding triplet.

5. The peak at 1014 cm^{-1} is found at 1045 in the present work and it is directly related to the spectra of sites 'N1' and 'H1', suggesting, in this case, a bond stretch of those particles.

6. The group of experimental bands centered around 1351 cm^{-1} is matched in our spectra by one peak centered at 1225 cm^{-1}. Connected to particles 'N1' and one carbon, we expect that this band should correspond to the stretch of 'N1' along with its carbon bond.

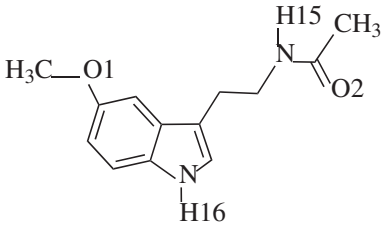
7. The maximum at 1432 cm^{-1} can be directly related to the maxima around 1410 cm^{-1} observed in the spectra of 'N2' and 'H2', suggesting that this vibration corresponds to a bond stretch between them.

8. The next band, centered at 1566 cm^{-1} is not well matched by any peak of the computed spectrum.

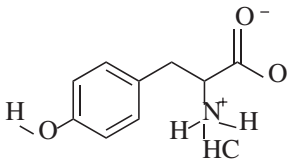
9. The remaining band centered at 1647 cm^{-1} is a group of maxima in the experimental spectrum and it can be also attributed to the contribution of 'H1' (1622 cm^{-1}). Since the label 'H1' accounts for the three hydrogens bound to 'N1', it should be attributed to a single stretching mode.

10. Finally, the three bands of highest frequency (2938, 3056, and 3402 cm^{-1}) must be fully attributed to stretching vibrations of hydrogens 'H1' and 'H2' and are located around 3170 (shoulder), 3270, and 3690 cm^{-1} in our spectral densities. The large deviations not only indicate a drawback in the characterizations of the vibrational interactions by the force field employed here, but they also indicate they derive from pure hydrogen vibrations.

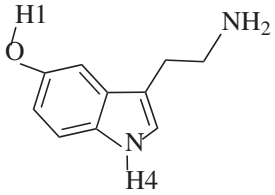
The influence of the cholesterol in the full spectrum of TRP reported in the left side of Fig. 3.15 is remarkable in some of the spectral bands, producing observable frequency shifts and it essentially indicates again that cholesterol interacts with TRP in a quite strong way and that these interactions are able to modify the vibrational motions of their atomic components producing frequency shifts up to 66 cm^{-1} (case of the band centered around 1045 cm^{-1} in the computed spectrum, red-shifted at the system with 50% of cholesterol).



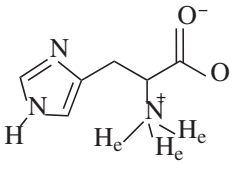
Melatonin



Tyrosine



Serotonin



Histidine-E

Fig. 3.16 Sketches of the backbone structures of histidine, tyrosine, serotonin, and melatonin with highlighted sites which will be inferred in the following section. Hydrogens bound to carbon are not shown.

3.3 Binding free energies of small-molecules in phospholipid membranes

Considering the importance of the DPPC lipid, a major constituent of pulmonary lungs, we kept using DPPC bilayer as the target membrane to study small molecules, for example, SER, MEL, HIS, TYR, and TRP. Focusing our attention on the free energy barriers of adsorption of small-molecule species, we have explored their interactions with the prototypical cell membrane formed by DPPC and water in sodium chloride solution using all-atom MD simulations, analyzing its local structure through free energy profiles based on the reversible work theorem. Sketches of the backbone structure of small molecules and DPPC are represented in Fig. 3.1 and Fig. 3.16.

In Fig. 3.1, the highlighted sites of TRP are 'H1', 'H2', 'O1', and 'O2'; 'O2' and 'O8' for DPPC. In Fig. 3.16, 'H1', 'H4', and 'O' atoms for SER; 'O1', 'O2', 'H15', and 'H16' for MEL; 'HC' stands now for 3 hydrogen atoms sharing one positive charge attached to a nitrogen atom of TYR; and same for 'H-TYR' representing 3 'H_e' atoms for HIS. Sites 'O1' and 'O2' of TRP share the negative charge, same for HIS and TYR. All highlighted sites will be referred to in this Chapter by the labels defined here.

In order to analyze the microscopic forces relevant for the binding process by obtaining the Gibbs free energy, here we use the potential of mean force (PMF), described in section 3.1.2.4.

3.3.1 Preparation of simulations

A model of a zwitterionic lipid bilayer membrane in aqueous sodium chloride solution has been build by means of the CHARMM-GUI tool. The membrane was composed of 204 lipids, distributed in two leaflets of 102 flexible DPPC molecules, surrounded by TIP3P water (W) molecules (enough to ensure full hydration in all cases), with 17 sodium and 17 chlorine ions, corresponding to physiological concentration, plus one small-molecule. In order to compare several probes of the different chemical structures and able to perform a variety of biological functions, we considered five species: three amino acids (TRP, HIS, and TYR), a neurotransmitter (SER), and a hormone (MEL). We should note that these three species have a common indole ring in their structures.

MD simulations were generated through the web-based CHARMM-GUI tool and performed with the NAMD2 simulation package at a fixed temperature of 323.15 K and at the average pressure of 1 atm. At this temperature, the DPPC membrane is fully at the liquid crystal state (see for instance Refs. [281, 282]). The temperature was controlled by a Langevin thermostat with a damping coefficient of 1 ps^{-1}. The CHARMM36 force field was applied for all simulations. After applying the NPT ensemble for equilibration periods which ran for more than 40 ns, we recorded statistically meaningful trajectories of more than 80 ns. A typical size of the system was of 80 Å$\times$ 80 Å$\times$ 81 Å, regardless of the probe considered, since the biggest part of the membrane was made of the same components, i.e. DPPC, water and ions in exactly the same concentrations. The simulation time step was set to 2 fs in all cases.

All bonds involving hydrogens were fixed to constant length, allowing fluctuations of bond distances and all sorts of angles for the remaining atoms. Van der Waals interactions were cut off at 12 Å with a smooth switching function starting at 10 Å, as usual. Long ranged electrostatic forces were taken into account by means of the particle mesh Ewald method, with a grid space of about 1 Å. Electrostatic interactions were updated every time step. Finally, PBCs were applied in the three directions of space.

3.3.2 Result and analysis

As a very first primary output, we have computed again the area per lipid for each system. We have monitored the surface area per lipid considering the total surface along the XY plane (plane parallel to the bilayer surface) divided by the number of lipids in one lamellar layer[288]. The final average areas per lipid are reported in Table 3.4. We needed to consider time scales of more than 40 ns in order to obtain a converged area per lipid in all cases. The main trend observed is that all values are very close to 61 $Å^2$, in overall good agreement with other computational results[290] where the values for pure DPPC range between 50 and 63 $Å^2$.

Table 3.4 Area per lipid (A) of the membrane systems. Estimated errors are in parenthesis.

Small-molecule	A (Å^2)
TRP	61.4(0.8)
HIS	60.8(1.5)
TYR	60.6(1.7)
SER	61.3(1.5)
MEL	61.1(0.9)

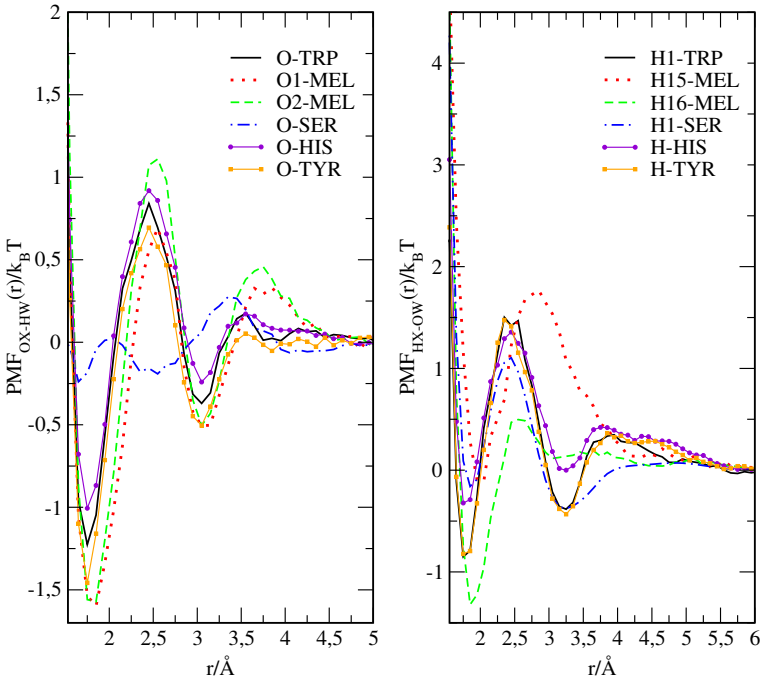


Fig. 3.17 Potentials of mean force for the adsorption of small-molecules to water molecules.

After fully equilibrated simulations were produced, we obtained a series of pair radial distribution functions $g_{12}(r)$ (not reported here) and applied the procedure explained earlier (see section 3.2.2.2) in order to obtain the PMFs. The results of PMF for water and DPPC *versus* small-molecules are displayed in Figs. 3.17 (water) and 3.18 (DPPC) in units of k_BT. In order to quantify the height of all barriers, we included the corresponding numerical estimation in Table 3.4 assuming that, for the present simulations, 1 $k_BT = 0.64185$ kcal/mol. The values reported in Table 3.4 are between 0.2 and 3.1 kcal/mol, i.e. of the same order of magnitude of the free energies of adsorption of metal ions in DMPC membranes[275].

We show PMF for water's oxygens at the plots in the left column and those for water's hydrogens in plots at the right side of Fig. 3.17. For oxygens of water, a free energy barrier is seen in all cases, defined by a neat first minimum and a second minimum clearly defined, although the barrier of water-SER is much smaller than

those corresponding to the rest of pairings. This finding is in good agreement with the results reported by Wood et al.[303], indicating that SER is normally anchored to the POPC membrane whereas TRP and other zwitterions have full access to the water region. In the case of hydrogens of water, the second minimum is not well defined for MEL's hydrogens 'H15' and 'H16' (see Fig. 3.16). The binding of small-molecules to water reveals, as a general fact, free energy barriers of between 0.2-2.8 k_BT with stable binding distances very close to the typical hydrogen-bond (HB) distances in water, given by the position of the first minimum of the oxygen-hydrogen radial distribution function (1.85 Å)[276]. However, the typical energy of water-water HBs estimated from *ab-initio* calculations is of about 5 kcal/mol[304], a value significantly larger than those for the HB observed in this work.

A closer look indicates that the largest barriers correspond to HB formed by oxygens of a small-molecule (acting as acceptors) and hydrogens of water, acting as donors. "Reverse" hydrogen-bondings composed by hydrogens of a small-molecule (donors) and oxygens of water (acceptors) are also possible but they are weaker than the former, with significantly smaller free energy barriers, up to 2.3 k_BT in size. In summary, the aqueous solvation of small probes revealed similar characteristics regardless of the type of the molecule (amino acid, neurotransmitter, and hormone), with the strongest pairing for MEL-water and weakest for SER-water.

Table 3.5 Free energy barriers ΔF (in kcal/mol) for the binding of small-molecules to water and to DPPC.

Probe (Active site)	O-water	H-water	O2-DPPC	O8-DPPC
TRP (O)	-	1.33	-	-
TRP (H1)	1.48	-	2.70	1.80
TRP (H2)	-	-	1.92	1.00
HIS (O)	-	1.24	-	-
HIS (H)	1.07	-	-	-
TYR (O)	-	1.39	-	-
TYR (H)	1.50	-	3.11	3.00
TYR (HC)	-	-	2.79	1.54
SER (O)	-	0.18	-	-
SER (H1)	0.84	-	1.90	1.56
SER (H4)	-	-	1.78	1.88
MEL (O1)	-	1.45	-	-
MEL (O2)	-	1.77	-	-
MEL (H15)	1.17	-	-	1.16
MEL (H16)	1.18	-	1.92	0.47

Regarding the interactions of small-molecules versus lipid atomic sites and from data reported in Table 3.5 and in Fig. 3.18, we observe the highest barrier (3.11 kcal/mol) corresponding to the pairing of TYR (through its hydroxyl's hydrogen) with the phosphate oxygen 'O2' of DPPC. In decreasing order, 'HC' of TYR and 'H1' of

TRP show also strong interactions with 'O2'. From a general point of view, all small molecules but HIS are able to establish HB with 'O2' and also with the site 'O8' of DPPC, much deeper in the membrane (see Fig. 3.1). It should be pointed out that the barrier of TYR to 'O8' is remarkable, of about 3 kcal/mol, and further indicates the stability of TYR at the membrane, compared to the other two amino acids: MEL and SER. As a general fact, the position of maxima of the first barrier is centered around 2.45 Å for small-molecule-'O2' binding, whereas barriers of ligands 'H' of TYR and 'H4' of SER associated to the 'O8' sites were centered around a slightly larger distance of 2.75 Å.

The most stable distance for 'O2' in DPPC bound to TRP is of about 1.75 Å, i.e. the position of the first minimum of the PMF between TRP and DPPC. As it has been stated before, such distance is of the order of the typical HB distance in water. Interestingly, the stable position for 'O8' sites of DPPC is centered in a wider distribution of values between 1.7 and 2 Å. For the sake of comparison, the PMF of TRP in a di-oleoyl-phosphatidyl-choline bilayer membrane shows a barrier of the order of 4 kcal/mol[89], whereas the barrier for the movement of TRP (attached to a poly-leucine α-helix) inside a DPPC membrane was reported to be of 3 kcal/mol[58]. Finally, neurotransmitters such as glycine, acetylcholine, or glutamate were reported to show small barriers of about 0.5-1.2 kcal/mol when located close to the lipid glycerol backbone[305].

In order to have a more detailed idea of the particular binding of some small-molecules to the membrane, we are reporting two characteristic snapshots of SER and TYR linked to two DPPC molecules (see Fig. 3.19). There we can observe that the most active sites are hydrogens belonging to hydroxyl groups, bound to DPPC at different sites ('O2' and 'O8' simultaneously for TYR, right and 'O2' for SER, left). These images are only significant configurations selected among a wide variety of possible choices (see Table 3.5) and may help the reader to enlighten the relatively complex multiple hydrogen-bonding connections between the small-molecules and DPPC described above.

3.4 Conclusions

A series of molecular dynamics simulations of zwitterionic lipid bilayer membranes in the aqueous ionic solution of NaCl with an embedded small molecule have been performed by MD using the CHARMM36 force field. Among them, cholesterol with different concentrations (30% and 50%) have been considered. In all simulations, systems have been well stabilized and then meaningful production trajectories at the time-scale of hundreds of nanoseconds are collected to analyze statistically meaningful properties.

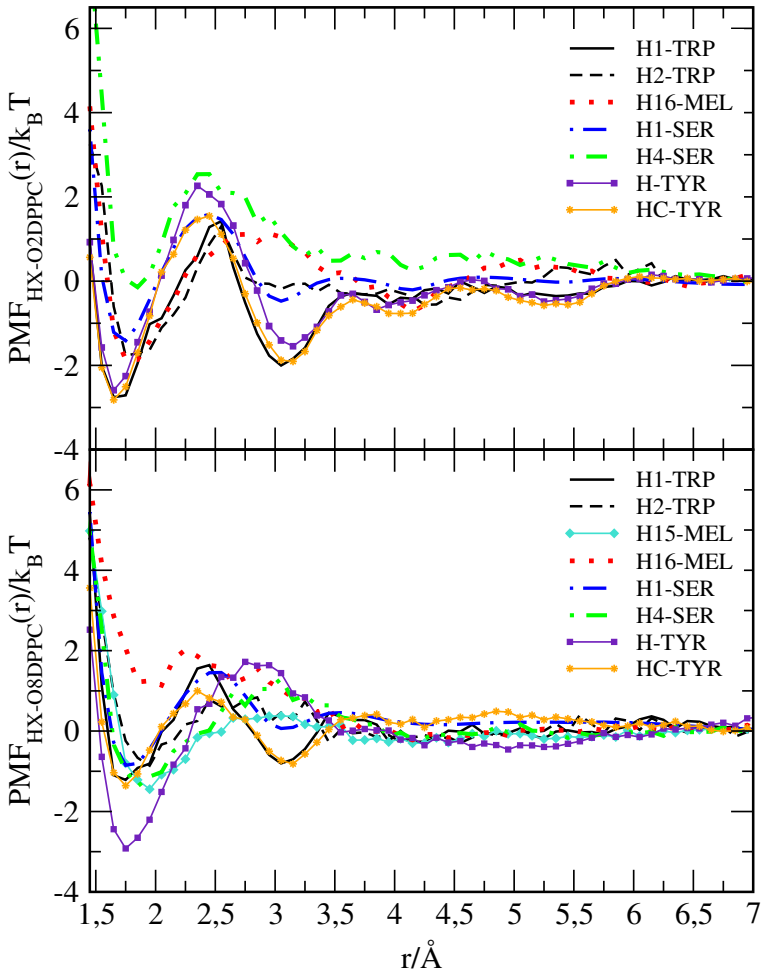


Fig. 3.18 Potentials of mean force for DPPC-small molecules.

In our preliminary study on tryptophan's adsorption to DPPC bilayer membrane at 310.15 K, we have carefully analyzed the local structure of tryptophan associated to water and to DPPC and evaluated the free energy of adsorption considering the usual one-dimensional reaction coordinates based on atomic distances for selected sites. Our data revealed the existence of a strong first coordination shell and a milder second coordination shell for the tryptophan-water association, which translated to deep minima in the corresponding PMFs, with energy barriers of less than 1 kcal/mol. Conversely, the binding of tryptophan to DPPC involves a single coordination shell for the two sites of possible association and energy barriers much lower than for the former TRP-water, i.e. of the order of 1 kcal/mol. Tryptophan shows a clear tendency to stay close to DPPC, during periods of time about 3 times longer than for TRP-water association.

Continuing our preliminary study, the temperature has been increased from 310.15 K to 323.15 K for three systems at three cholesterol concentrations (0, 30, and 50%) in order to systematically examine tryptophan-DPPC, tryptophan-cholesterol and

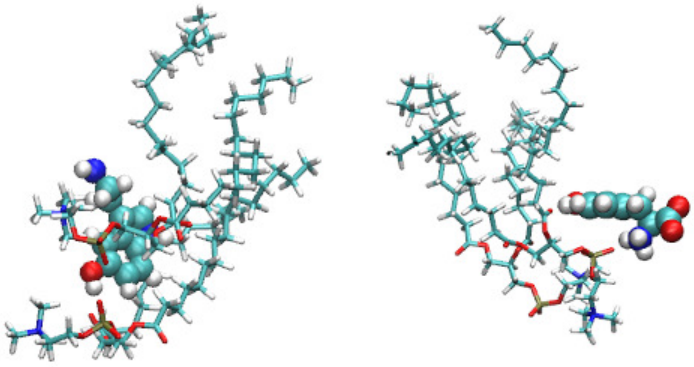


Fig. 3.19 Snapshots of typical small-molecule and DPPC bonds. Serotonin-DPPC (left, where the binding of a hydroxyl's hydrogen with 'O2' site is clearly seen) and tyrosine-DPPC (right, where the binding of hydrogen 'HC' to the 'O2' site and of the hydrogen from the hydroxyl group with 'O8' site are clearly seen).

tryptophan-water interactions at liquid-crystalline phase conditions. The effect of cholesterol on the tryptophan biding to the zwitterionic DPPC membrane has been explored. After the systematic analysis of meaningful data, we noted relevant changes in local structure and dynamics of tryptophan only for cholesterol concentrations above 30%. In agreement with MacCallum et al.[89], we observed that tryptophan tends to localize strongly in the interfacial region so that crossing the membrane by diffusion with no help of external fields or mediating particles is unlikely to happen in the time scale of our simulations. The binding of tryptophan to DPPC involved coordination shells for the different sites of possible association (charged oxygens 'O2' and 'O8' of DPPC *versus* the two tagged hydrogens in TRP, namely 'H1' and 'H2'). Also, the distribution functions of tryptophan-cholesterol revealed stable hydrogen-bonding configurations. These data indicate that tryptophan is able to establish stable interactions with all classes of solvating particles (water, DPPC, and cholesterol) including a sort of *bridging* between DPPC and cholesterol species. The typical hydrogen-bond distances between tryptophan and other species have been found between 1.7-2.0 Å, in good agreement with experimental data obtained from fluorescence measurements[100]. Tryptophan's self-diffusion coefficients have been found to be of the order of $1\text{-}10\times10^{-7}$ cm^2/s and they are strongly dependent on the presence of cholesterol in the system. The diffusion of tryptophan has revealed to be significantly faster than that of DPPC and slower than that of water. The computed spectral densities, in excellent agreement with experimental infrared and Raman data from Leyton et al.[99], allowed us to reveal the degree of participation of each atomic site of tryptophan to the complete spectrum of the molecule.

In the last part of this Chapter, considering the importance of studying DPPC lipids, we kept using DPPC bilayer to study the binding free energies of small molecules: serotonin, histidine, tyrosine, melatonin, and tryptophan at the condition of 323.15

K and 1 atm. We found that the area per lipid was practically not influenced by the presence of one particular probe and they were in all cases around 61 $Å^2$, in agreement with other computational and experimental data. The Gibbs free energy of adsorption has been evaluated through the potentials of mean force. We have considered the usual one-dimensional reaction coordinates based on atomic distances for selected sites. We chose six types of particles: (1) the hydrogens labeled 'H1' and (2) the double-bonded oxygens 'O1' and 'O2' of the small molecule; (3) the three water sites; (4) the charged oxygens labeled 'O2' and 'O8' of DPPC. Our data revealed the existence of a strong first coordination shell and a milder second coordination shell for small molecule-water structure, which correspond to two minima in the corresponding PMFs, with energy barriers for the TRP-water association of the order of 1-2 kcal/mol. Conversely, the binding to DPPC involves a single coordination shell for the two sites of possible association (oxygens 'O2' and 'O8' of DPPC *versus* hydrogens in the small-molecules) and energy barriers between 0.5-3 kcal/mol. Throughout our simulation runs, we did not observe any event of permeation of a small-molecule across the DPPC membrane. Concerning the essentiality of the two amino acids reported in the present work (histidine, tryptophan), we observed that tryptophan is able to enter the interfacial membrane, whereas histidine is not. Interestingly, tyrosine, a non-essential amino acid shows the highest free energy barriers, indicating that it is the most stable molecule for DPPC binding. Serotonin has revealed to be a molecule anchored at the membrane and with a low propensity to be solvated by water, whereas its derivative melatonin is able to equally interact with water and DPPC, showing similarly strong free energy barriers. Due to its medical purposes, we will investigate the dynamics of melatonin binding to the cell membranes in the next Chapter.

Chapter 4

Cellular absorption of small molecules: molecular dynamics and free energy landscapes of melatonin binding at phospholipid membranes

In this chapter we have focused our efforts on the study of the binding of MEL[306, 307] at a simplified model cell membrane. This is a process that aims to improve our understanding of the basic mechanisms of molecular binding and crossing of biological membranes by small solutes and the interactions with their surroundings. In the present work, we have considered a model membrane composed of cholesterol and DMPC lipids, extensively studied in the literature from the experimental and also computational points of view[47–49]. We have devoted our work to the analysis of the structure and dynamics of MEL, following the previous Chapter, where TRP[91, 283] and other similar solutes were simulated[308]. In order to investigate these relevant effects of MEL on the human body and more specifically on plasma membranes, we have focused our attention on the characterization of the structure and transport processes of MEL at the atomic scale with the aid of MD simulations at the time-scale of hundreds of nanoseconds.

Although the binding constant and Gibbs free energy of a probe adsorption can be estimated from experiments[58, 89], detailed studies of the relative stability of different bound states are still scarce in the literature. In addition, in a system with multidimensional reaction coordinates, usually several stable states (bound configurations) are separated by high free energy barriers corresponding to transition states of the system[309, 310], making it difficult for MD simulations to sample them

adequately[225]. Free energy calculations using enhanced sampling techniques provide a method to address the problem and provide a great deal of understanding of the passive diffusion phenomena of small solutes over a barrier, which requires a detailed view of the underlying free energy landscape (FEL). However, despite of the significant progress of free energy calculations achieved in recent years[311–313], the free energy landscape of MEL and most of the small solutes on membrane surfaces is still mostly unknown. The problem of computing free energy landscapes in the multidimensional quantum or classical systems has been extensively discussed in the literature. We have employed well-tempered metadynamics in the study of zwitterionic phospholipid membranes that can help understand basic biological membrane functions and their interaction with specific small molecules, such as MEL. Furthermore, we will also discuss the role of cholesterol in the cellular absorption of small molecules.

4.1 Molecular dynamics of melatonin binding the interface of DMPC lipids

4.1.1 Methods

A realistic model of DMPC-cholesterol membrane bilayers in a sodium chloride solution has been generated using CHARMM-GUI web-based tool. The cholesterol-free system was composed of 204 DMPC lipids molecules, 10250 TIP3P (water) molecules (allowing flexible bonds through harmonic springs), with 21 sodium and 21 chlorine ions of physiological concentration, along with one single MEL molecule obtained from pdb 4QOI through PDB Reader plugin of CHARMM-GUI website-based tool. From the previous work[91], only two significant cholesterol percentages of 30% and 50% should be considered such that the harvested data should be compared with the results obtained for the case with 0% cholesterol. MEL locates at the interface of membrane bilayer in all three cases at the beginning of the simulation. For the convenience, sketches of the backbone structures of MEL, DMPC, and cholesterol are represented in Fig. 4.1. The NAMD2 software package and the recently reparameterized CHARMM36m force field were used in all MD simulations at a fixed temperature of 303.15 K and at the fixed pressure of 1 atm (NPT), in order to make sure all simulations were performed at the liquid crystal phase[43], since the transition temperature of DMPC lipid from gel to liquid crystal phase of DMPC has been determined to be close to 297 K[9]. As usual in such kind of the simulations[283] the temperature was regulated with a Langevin thermostat with a damping coefficient of 1 ps^{-1}, whereas the pressure was controlled by a Nosé-Hoover Langevin piston with Langevin dynamics at an oscillation period of 50 fs.

All MD simulations used for analysis were operated in NPT conditions. After a run of 150 ps for relaxation and 100 ns of equilibration periods, several production

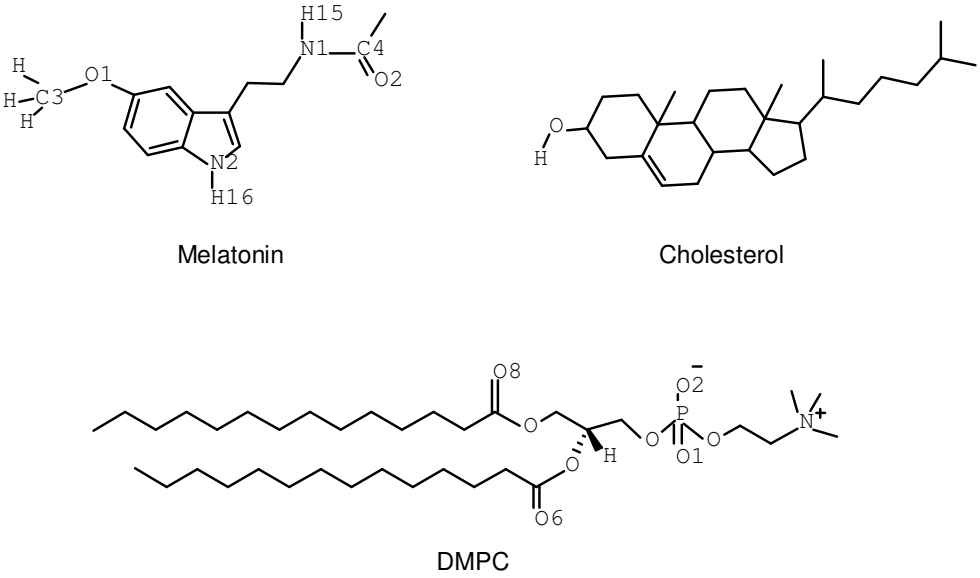


Fig. 4.1 Backbone structures. Sketches of molecular structures of MEL, DMPC, and cholesterol. Part of Hydrogen-Carbon bonds not shown. The highlighted sites of MEL ('C3', 'C4', 'H15', 'H16', 'N1', 'N2', 'O1', and 'O2') and of DMPC ('O1', 'O2', 'O6', and 'O8') will be referred in the text by the same labels.

runs were generated and statistically meaningful trajectories of more than 100 ns were recorded in all cases. The simulation boxes had different sizes because of different cholesterol concentrations. For instance, the size of the cholesterol-free system was of 79 Å× 79 Å× 85 Å. A time step of 2 fs was used and periodic boundary conditions were applied. All bonds involving hydrogens were set to fixed lengths, allowing fluctuations of bond distances and angles for the remaining atoms. During the calculation of spectral densities all bonds (including those involving hydrogens) were left flexible. The cutoff for the Van der Waals interactions was of 12 Å and a switching function was employed starting at 10 Å. Coulomb forces were computed using the particle mesh Ewald method, with a grid space of 1 Å. Every time step electrostatic interactions were updated. The usual periodic boundary conditions in all directions of space were taken.

4.1.2 Results and Discussion

4.1.2.1 Physical characteristics of the membranes

In order to explore some phase-diagram states of the model systems being simulated, as well as to efficiently characterize the ordering inside the hydrated lipid bilayer, a procedure already employed in the previous works[17, 91, 284, 285] was used. A deuterium order parameter S_{CD} was defined for each CH_2 group of the DMPC lipid tails as Eq. 3.6 and average results are shown in Fig. 4.2 for both tail chains of all DMPC lipids at the three cholesterol concentrations considered in this Chapter. S_{CD} can be

also obtained from ^{2}H NMR experiments[314]. The results for the cholesterol-free case were previously tested[17] and are in good agreement with both simulation[287, 315] and experimental works[98, 286], confirming the liquid crystal phase was represented well in the three systems adopted in this Chapter. We should note that as cholesterol concentration increases in the system, the tendency to higher ordering increases too, which is represented by profiles of S_{CD} having larger maxima (around ~ 0.4 for the cholesterol-rich setups *versus* ~ 0.2 for the cholesterol-free system), a tendency which was already observed by Petrache et al.[98].

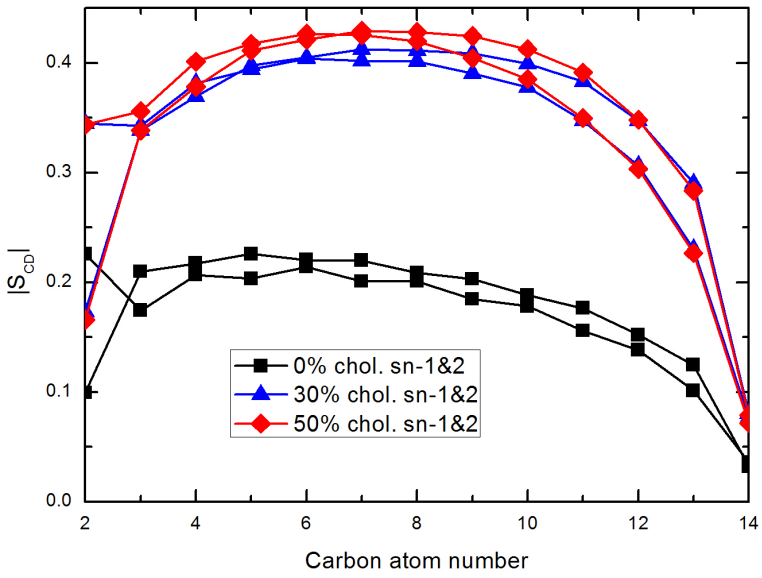


Fig. 4.2 Order parameter. $|S_{CD}|$ for the (sn1, sn2) acyl tails of DMPC at three different cholesterol concentrations.

The area per lipid is definitely a relevant output from the most molecular simulations of plasma membranes. We have calculated the area per lipid, as defined in Eq. 3.7, considering the membrane surface along the XY plane divided by the number of lipids and cholesterol[288]. For continuous MD production runs, area per lipid as a function of simulation time is reported in Fig. 4.3 whereas their average values together with the average thickness of membranes are reported in Table 4.1. Area per lipid decreases as cholesterol concentration increases. We obtained a value of around 62 Å^2 for a cholesterol-free system and smaller values down to 40 Å^2 for the system with a concentration of cholesterol of 50%. These results are in excellent agreement with other computational works[98, 282] where the value for pure DMPC is of about 60 Å^2 at 303 K. According to the review of Nagle et al.[6], values of area per lipid of pure DMPC membranes (303 K) can be obtained from multiple methods (neutron scattering, X-ray and NMR) and were reported to be between 59 and 62 Å^2 at the liquid phase. In our

case, the change in the area per lipid has been observed to be more marked when the concentration of cholesterol was above 20% (not reported). This is consistent with the observed fact that DMPC membranes in this work experienced the phase transition point from a liquid-disordered phase (cholesterol-free system) to a liquid-ordered phase (systems of cholesterol 30% and 50%)[291, 292].

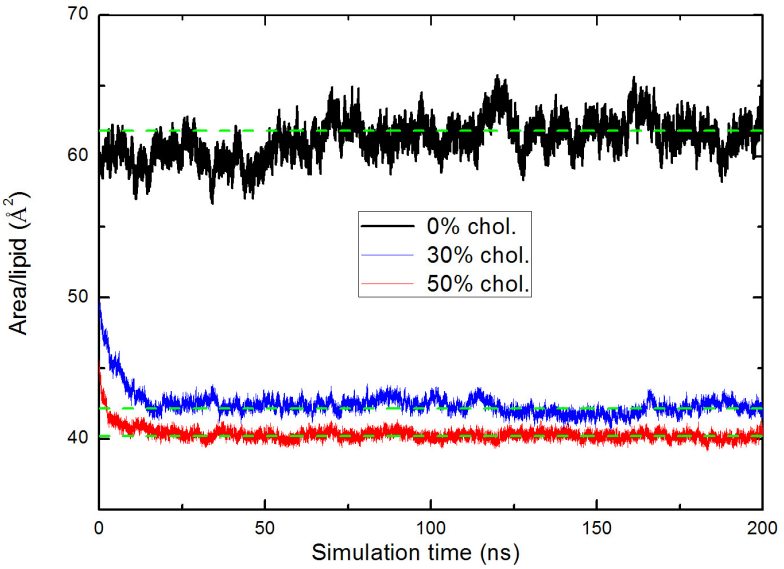


Fig. 4.3 Physical characteristics of the membranes. Area per lipid of systems with different cholesterol contents: 0% (black line), 30% (blue line), 50% (red line) as a function of simulation time. The green dashed lines indicate the average values for the last 150 ns at each concentration.

The thickness of the membrane may provide additional clues about the influence of cholesterol on its mechanical properties, such as rigidity and capability of allowing the movement of species in and out of the cell. We have obtained the thickness of the membrane Δz by computing the distance between phosphorus atoms (P) of the DMPC head groups from both layers.

Table 4.1 Area per lipid (A) and thickness (Δz) of the membrane for all three cholesterol concentrations studied in this work. Estimated errors in parenthesis.

Percentage of cholesterol	A (Å^2)	Δz (Å)
0%	61.8(1.2)	34.9(0.6)
30%	42.1(0.5)	44.3(0.3)
50%	40.2(0.3)	44.7(0.3)

The results of the thickness of the membrane are in good agreement with those reported by Kučerka et al.[282] by means of X-ray and neutron scattering. These authors reported a value of 36.7 Å at 303 K for the DMPC membrane at a cholesterol-free system. In this Chapter, we observe a tendency to larger bilayer thickness as cholesterol concentration increases. As was pointed out in the case of the binding of TRP at DPPC-cholesterol membranes[91], at higher cholesterol percentages the values of A are smaller: the larger the cholesterol contents the more compressed are the bilayer structures. This eventually can increase the rigidity of the membrane, extending the lipid tails, and producing larger bilayer thickness. In summary, the increase of the rigidity of the membrane is a fact already observed by several authors from both experimental and computational sides, such as Drolle et al.[96] or Choi et al.[97] for cholesterol-MEL mixtures in phosphatidylcholine membranes. In their studies these authors found out the effect of MEL reducing the thickness of the membrane and enhancing its fluidity, a compensating effect of the condensation introduced by cholesterol. In the present work we only considered a single MEL molecule that did not allow us to explore the joint effects of MEL and cholesterol on the thickness of the membrane.

The penetration of MEL in the membrane along its normal direction is also a relevant feature. We report in Fig. 4.4 the Z-axis position of MEL from the center of the bilayer (i.e. $z = 0$) using the last meaningful 80 ns of each production trajectory in all three cases, namely, concentrations of cholesterol set at 0%, 30% and 50%. Red symbols represent the average positions of phosphorus (P) atoms of the head groups of DMPC along the direction normal to the membrane (Z-axis); green symbols stand for the Z-axis distance between the center of mass of MEL and the center of DMPC bilayers, as a function of simulation time. It stands clear that the thickness of the DMPC bilayer membrane is increased when cholesterol is present in the membrane, in good agreement with the results of thickness reported in Table 4.1. At 0% and 30% cholesterol systems, MEL stays in the internal region of DMPC bilayers most of the time, with a few occasional visits to the interface of the membrane, i.e. surroundings of the P atoms; however, when more cholesterol is added into the system (i.e. at 50%), MEL can diffuse into water bulk and become fully solvated by water or the head groups of DMPC without changing the size of the system too much, indicating the competition between water and lipids to solvate MEL. These results are in good agreement with those of Drolle et al. that reported the preferential location of MEL in a DPPC bilayer at distances around $z = 1.2$ nm (see Fig. 8 of Ref.[96]), i.e. at the crossover region between lipid head groups and the fatty acid chains. In the case of 50%, Fig. 4.4 shows that MEL has the ability of either being adsorbed by the head groups of DMPC either to stay for long periods of time in the water bulk, so we can expect that it is not difficult for a small molecule like MEL to cross the free energy

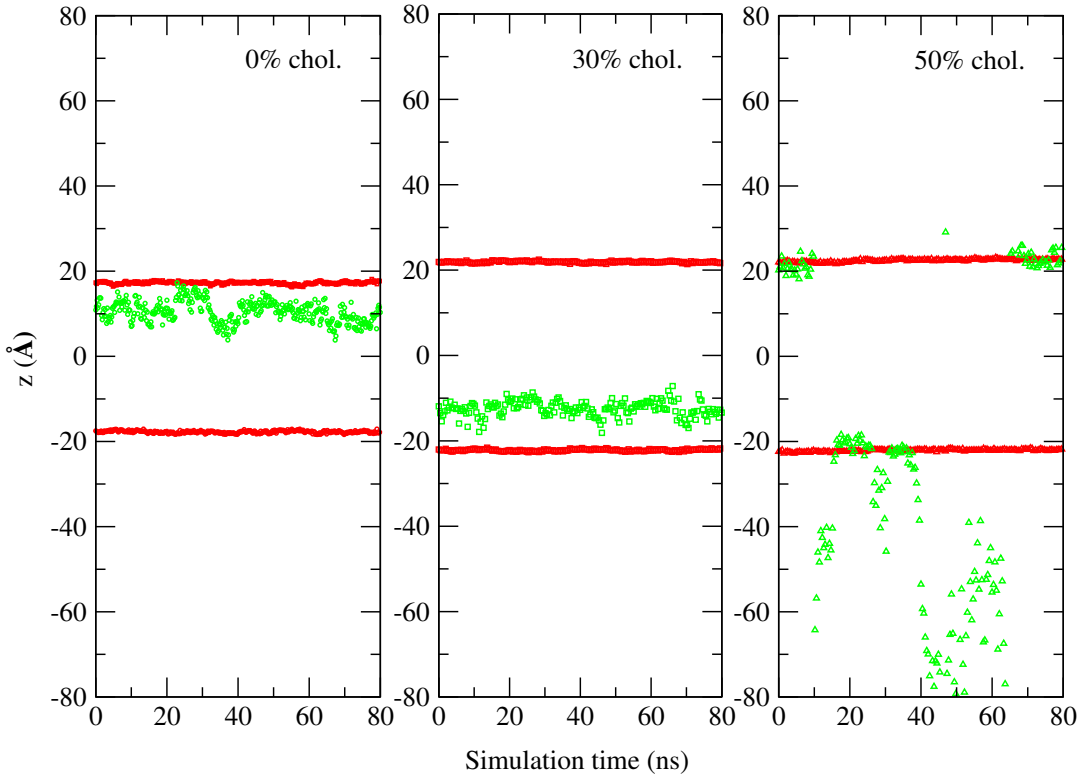


Fig. 4.4 Z-axis position of MEL. Penetration of MEL inside DMPC bilayer (green symbols indicate the position of the center of mass of MEL whereas red symbols stand for the position of phosphorus atoms atoms in each layer) at 0%, 30%, and 50% cholesterol concentrations.

barrier between the two states (solvated by water or solvated by the head groups of DMPC). This latter aspect will be addressed with more details in section 4.2.

4.1.2.2 Radial distribution functions of melatonin around DMPC, water, and cholesterol

A direct route to the characterization of the local structure of each atomic species of the system is usually obtained by means of normalized radial distribution functions (RDF) $g_{12}(r)$ for two different species '1' and '2', see Eq. 3.1. Among the wide variety of possible RDF that could be computed, we have considered only six relevant RDF based on the first coordination shells of 'H15' and of 'H16' of MEL. The remaining RDFs indicate low maxima at distances significantly longer than the typical HB values or show too noisy profiles which indicate that the corresponding local structures are

not stable enough. The selected $g(r)$s are reported in Fig. 4.5 for the three cholesterol percentages (0, 30, 50%).

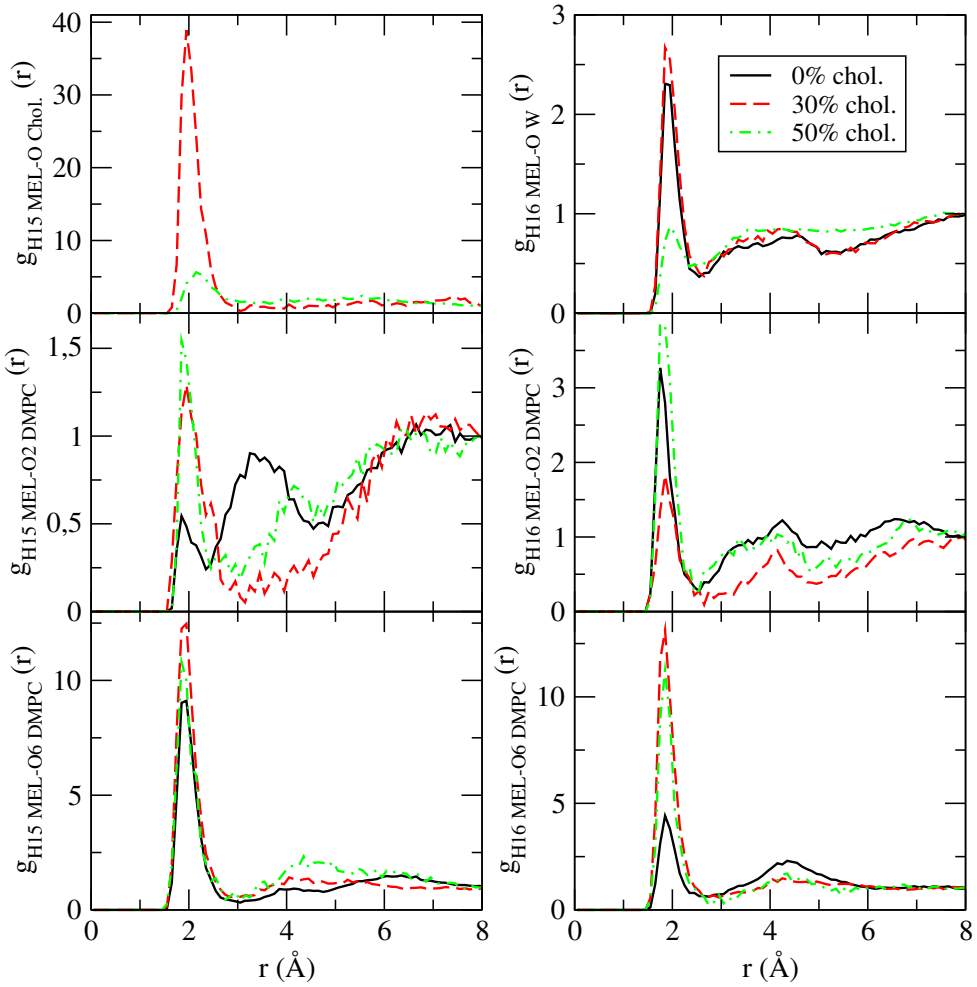


Fig. 4.5 Radial distribution functions. Selected radial distribution functions for hydrogens of MEL ('H15' and 'H16') with oxygens of water ('OW'), DMPC ('O2' (representing 'O1 & O2') and 'O6' (representing 'O6 & O8')) and cholesterol, belonging to hydroxyl group ('O Chol').

All six $g(r)$s show some fluctuations in their profiles, especially at the coordinates of $r = 3$ Å and beyond, i.e. those corresponding to second coordination shells. We could observe a neat first coordination shell in every case, located around 1.8-2.0 Å that should be essentially attributed to HB between MEL and the remaining species, cause such distance is the signature of typical oxygen-hydrogen HB in water[316]. Interestingly, the largest peak in all RDF is, by far, the one appearing at the MEL-cholesterol association, centered at 1.9 Å when the concentration of cholesterol is of 30%. Further, when we raised the concentration to 50% such band decreased dramatically and its position was shifted to about 2.2 Å . The interaction of MEL with cholesterol in DPPC bilayers, already reported by Choi et al.[97] produced a

fluidizing effect on the membrane for a MEL concentration high enough, opposite to the condensing effect of cholesterol.

In all the remaining cases, HB lengths are around 1.9 Å. The height of each maximum (related to the intensity of the HB) depended strongly on the concentration of cholesterol, as follows: (1) in the case of MEL-water association, strong HB were observed between 'H16' and water at 0% and 30% although at 50% they were much weaker; (2) both 'H15' and 'H16' were able to form HB with the DMPC sites 'O1' (or 'O2', both sites sharing the negative charge); (3) finally, MEL can establish HB between both 'H15' and 'H16' hydrogens with the DMPC sites 'O6' (or 'O8') in all three percentages of cholesterol. These findings of HB association between MEL and DMPC are in good agreement with those from Severcan et al.[85] who, by means of Fourier transform infrared spectroscopy, observed the existence of hydrogen bonding between the hydrogen in the N-H group of the furanose ring of MEL (labeled 'H16' in the present work) and the carbonyl (C=O) and phosphate (PO_4^-) groups in DPPC membranes. From our findings we have observed both HB of 'H15' and 'H16' of MEL with the phosphate group of DMPC ('O1') and also with the more internal C=O groups ('O6' and 'O8'). Thus, the novelty here is the hydrogen bond association of 'H15' with the two well-known acceptor groups in phosphatidyl-cholines indicated above, together with the already reported association of 'H16'. This fact allows MEL to be adsorbed deeper than TRP[91] at the membrane lipid bilayer with two selected donors ('H15' and 'H16') as well as through 'H15-O Chol.' bridges, which provides a variety of structures as it will be described in full details below. Results on MEL located close to the lipid head groups in studies of MEL inside DOPC and DPPC membranes were found by Drolle et al.[96] by means of small-angle neutron diffraction and MD simulations.

4.1.2.3 Estimation of Gibbs free energy differences

From a general perspective, the calculation of the Helmholtz or Gibbs free energy differences for binding processes or for configurational changes is a difficult task and it requires a considerable amount of computer time and a precise knowledge of the hypersurface of the potential energy of the system[317]. This can be explored by means of methods such as metadynamics[226, 275], hybrid quantum mechanics/molecular mechanics methods[318] or transition path sampling[213, 280, 293, 319]. However, a usual way to obtain free energy estimations is through the so-called potential of mean force (PMF), which is an estimation of the Gibbs free energy difference between two particles ('1' and '2') and together with pair radial distribution function (RDF) has been explained earlier in sections 3.1.2.4 and 3.2.2.2, respectively. The only restriction of this method is the fact that a single radial distance is always assumed as the pre-conceived reaction coordinate. Using this procedure on the RDF reported above, we have estimated the height of the barriers appearing in the Gibbs free energy differences

for MEL when bound to selected atomic sites from water, DMPC and cholesterol, assuming that at 303.15 K the equivalence is $k_BT = 0.602$ kcal/mol.

Table 4.2 Gibbs free energy differences PMF (in kcal/mol) for the binding of MEL (sites 'H15' and 'H16') to selected atomic sites. Since 'O1' and 'O2' sites in DMPC share the negative charge, their contributions have been averaged; the same situation for 'O6' and 'O8'.

Atomic site	H15			H16		
	0%	30%	50%	0%	30%	50%
O-water	-	-	-	1.08	1.20	0.42
O2-DMPC	0.48	1.69	1.14	1.45	1.56	1.62
O6-DMPC	1.99	1.87	1.93	1.14	1.87	2.23
O-chol.	-	2.53	0.90	-	-	-

From the results in Table 4.2 we can observe that all Gibbs free energy differences are of the order of 1 kcal/mol, in overall good agreement with preliminary calculations (case of a cholesterol-free DPPC membrane at 323 K)[308]. Density functional theory calculations of binding energy barriers of aqueous solvation of MEL in water clusters have been found to be of a few kcal/mol[320]. In most cases, a single coordination shell has been observed. As expected from the RDF reported above (see Fig. 4.5), the highest free-energy barrier corresponds to the association of MEL with cholesterol, at the 30% concentration, where the hydrogen bonding distance is very close to 2 Å and tends to increase when cholesterol concentration is of 50%. Conversely, all remaining HB distances have been found at the typical value of 1.85 Å . The MEL-water HB has been observed, in a significant amount, for 'H16' and not for 'H15' and they become very weak at the highest cholesterol concentration.

Focusing on DMPC-MEL HB, we have encountered a strong influence of cholesterol, with the overall highest barriers at 50% concentration. The typical bindings sites were O1-O2 (located at phosphate belonging to the head groups of DMPC) or O6-O8 (at carbonyl species on the tail groups of DMPC). Our results indicate that the presence of cholesterol increases the energetic cost of the binding of MEL to the membrane, with a preferential association of MEL 'H15' to the phosphate group of DMPC and of 'H16' to carbonyl groups, all depicted in Fig. 4.1. This is in agreement with the preferential location of MEL deep inside the 30% cholesterol membrane and also with the fact that MEL can move outside the membrane more easily as cholesterol concentration increases (see Fig. 4.4 as well as data from Refs. [96, 97]).

4.1.2.4 Angular orientations of melatonin at the membrane

In order to analyze the angular distributions of MEL at the interface of the membrane, we have defined three dihedrals, according to Fig. 4.6.

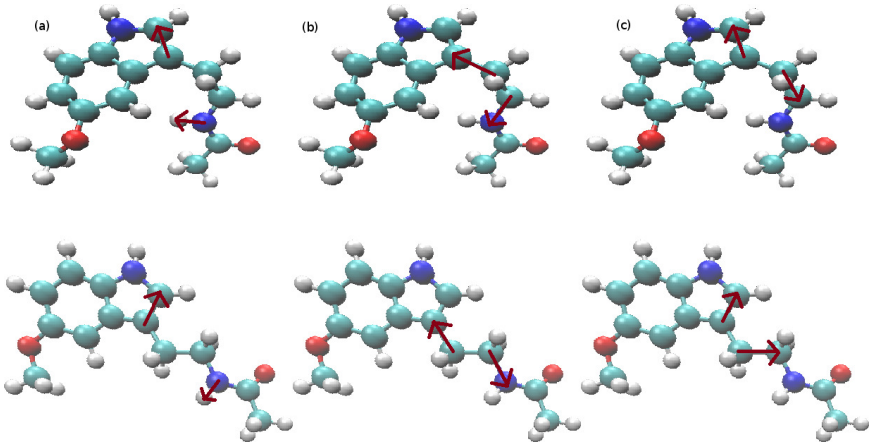


Fig. 4.6 Dihedral angles. Angles (a) θ, (b) Ψ and (c) ϕ defined for MEL. Top figures correspond to "folded" and bottom figures to "extended" configurations for each dihedral. The two snapshots (top, bottom) have been obtained from equilibrated configurations
.

We report in Fig. 4.7 angular distributions of MEL using the three selected dihedral angles defined in Fig. 4.6. In all cases we can observe that, on average, the angular distributions of MEL are centered around two preferential orientations, called "folded" and "extended" configurations, that are found at all cholesterol concentrations. The nitrogen atom involved in dihedrals θ and Ψ is the one labeled 'N1' in Fig. 4.1, namely the nitrogen chemically bound to the hydrogen labeled 'H15' in Fig. 4.1. For the dihedral angle θ, the average angles corresponding to stable configurations are of 1.08 rad (62°, folded configuration) and 2.22 rad (127°, extended configuration); for angle Ψ, the two stable configurations correspond to 1.42 rad (81°, folded) and 2.96 rad (170°, extended), whereas for ϕ the values are of 2.07 rad (119°, folded) and 1.17 rad (67°, extended).

From the distributions reported in Fig. 4.7, we can note that the dihedral angle θ can reach continuously all sorts of values between 0.44 and 2.70 rad, the angle ϕ can fluctuate between 0.98 and 2.60 rad, whereas the angle Ψ is much better defined and it reaches either value around 2.96 $\pm$ 0.18 rad or 1.42 $\pm$ 0.41 rad, regardless of the concentration of cholesterol of the system. After analyzing 100 ns of equilibrated trajectories (production runs), we calculated the ratios of the three dihedral angles for equilibrated systems with different cholesterol concentrations and reported the averages in Table 4.3.

From the results of dihedral angles θ and Ψ (see Table 4.3), we get that the extended configuration of MEL is more favored when adding cholesterol into the system. However, from the results of ϕ, it seems that at 50% cholesterol percentage, MEL stays much less at its extended configuration than 0% and 30%. Considering the behavior of MEL in all three cases (see Fig. 4.4), we found that MEL spent almost half of the simulation time in the water bulk at the 50% cholesterol system and it

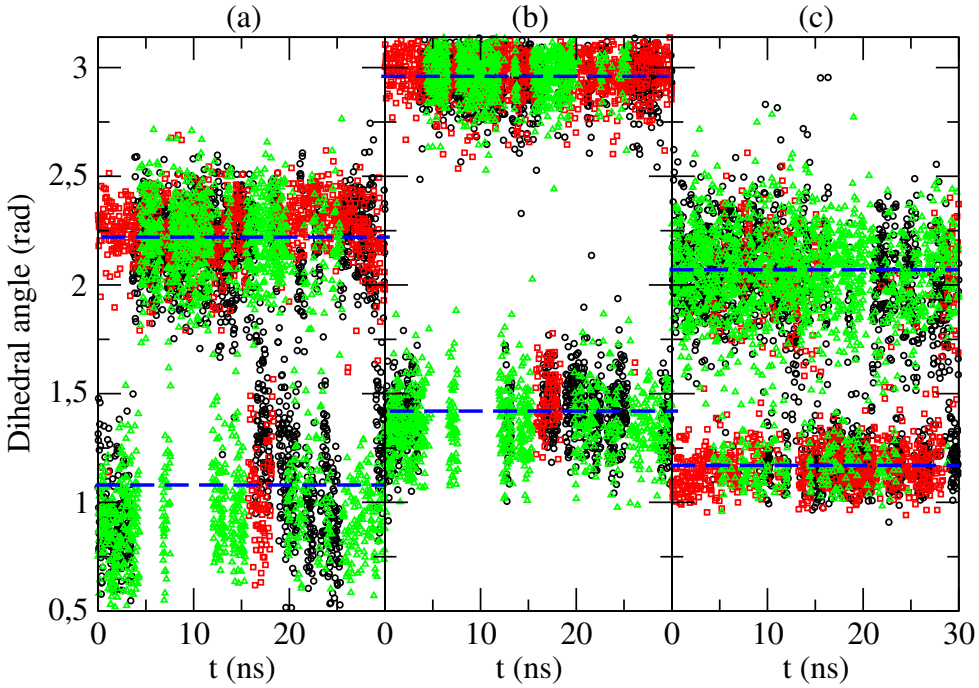


Fig. 4.7 Angular distributions. Distribution of selected dihedral angles in MEL as a function of simulation time, where labels (a), (b), and (c) have the same meaning as in Fig. 4.6. Percentages of cholesterol are: 0% (black circles), 30% (red squares), and 50% (green triangles). Dashed lines indicate average values and are a guide for the eye.

Table 4.3 Ratio of the two angular configurations for MEL in three systems with different cholesterol concentration. "Ext." stands for "extended configuration of MEL" and "Fol." stands for "folded configuration of MEL" in all cases.

	θ		Ψ		ϕ	
	Ext.	Fol.	Ext.	Fol.	Ext.	Fol.
0%	0.53	0.47	0.46	0.54	0.40	0.60
30%	0.53	0.47	0.52	0.48	0.46	0.54
50%	0.62	0.38	0.56	0.44	0.19	0.81

stayed at the interface for 0% and 30% cholesterol systems. So, we can suggest that the dihedral angle ϕ is significantly more sensible than the other two dihedral angles as being solvated by water. We can also suggest that introducing cholesterol into the system could help MEL change from its folded to its extended configuration more easily through hydrogen-bonding between MEL-DMPC and MEL-cholesterol. Also, according to this, Ψ is an excellent candidate for being used as a collective variable in metadynamics calculations[226, 231] of free energy landscapes for MEL binding in biomembranes, the results of such type of calculations will be reported in section 4.2. Our results confirm the suitability of Ψ as a collective variable.

4.1.2.5 Diffusion of melatonin and water

Dynamics of MEL and water have been extensively explored through translational diffusion and vibrational spectroscopy, as we will report below. Conversely, the dynamics of lipids and cholesterol are much slower and it has not been considered here. In particular, the mean square displacements (MSD) of water and center of mass of MEL, defined by Eq. 3.8, have been evaluated. From the long time slopes of MSD (not reported here), self-diffusion coefficients (D) have been calculated for three cases. The results are reported and summarized in Table 4.4.

Table 4.4 D (in cm^2/s) of MEL and of water in three systems with different cholesterol percentages. Estimated errors in parenthesis.

D	0%	30%	50%
Water	$4.0(0.1)\times10^{-5}$	$4.3(0.1)\times10^{-5}$	$4.4(0.1)\times10^{-5}$
MEL	$1.1(0.4)\times10^{-7}$	$3.9(0.6)\times10^{-7}$	$4.1(0.9)\times10^{-7}$

For the self-diffusion of water, all water molecules in the system are included in the results regardless of their location in the bulk or interfacial regions. At 0% cholesterol our results are in overall agreement with the ones obtained in our research group for the simple aqueous DMPC membrane[17] (2.7×10^{-5}) what indicates that when cholesterol concentration rises, water tends to diffuse slightly faster. Nevertheless, we can observe that the presence of cholesterol and the single MEL molecule does not affect the dynamics of water to a large extent. This could be due to the fact that with higher cholesterol concentration in the system plasma membranes are more packed and water cannot easily penetrate the interface of the membrane, having its main diffusion along the instantaneous surface of the bilayer.

Diffusion coefficients for MEL are approximately two orders of magnitude smaller than those of water and show a tendency to increase when cholesterol is mixed with DMPC, regardless of its concentration. In Table 4.4, at 30% cholesterol the value of D for MEL is six-fold larger than the value of D of DMPC molecules in pure DMPC bilayer membrane systems (0.6×10^{-7} cm^2/s), although the former is about 6.5-fold larger. Thus, the diffusion of MEL is significantly faster than that of DMPC. This fact would suggest that its mechanisms of diffusion may be similar to those of an individual particle (such as in Fickian diffusion) and qualitatively different of those of lipids, whose diffusion was observed to occur in a sort of collective way, associated in local groups of a few units (around 5-10 units)[17].

4.1.2.6 Spectral densities of melatonin

IR spectroscopy exploits the fact that molecules absorb specific frequencies that are characteristic of their structure and atom-atom interactions. IR spectrum is usually obtained on a spectrometer using the attenuated total reflection sampling technique

with a neat sample in the laboratory. Such experimental properties could also be obtained from certain MD simulations[300, 301]. We carefully calculated the atomic spectral density $S_i(\omega)$ defined as Eq. 3.9 in section 3.2.2.4. In the present section, we have obtained the spectral density of MEL and also of each atom site (only part of them are shown). Since the force field employed in the present work accounts for harmonic bond vibrations, we have been able to locate most of the positions of main experimental spectral bands (see Ref. [91] for more details). At this point we should remark that one powerful characteristic of computing molecular spectra in simulations is the possibility of locating the particular atomic sites contributing to each spectral band and thus be able to make a precise interpretation of experimental spectra. We will report this feature below. The full spectral density $S_{MEL}(\omega)$ and its decomposition into atomic contributions for the 0% cholesterol case is reported in Fig. 4.8.

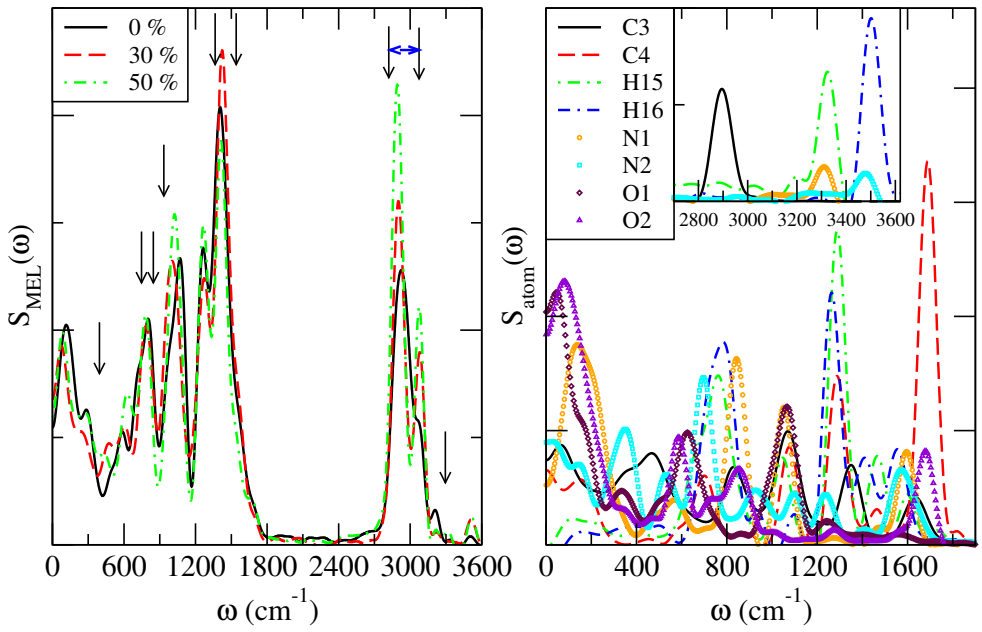


Fig. 4.8 Spectral densities. $S_{MEL}(\omega)$ of MEL (left). Positions of main experimental peaks are indicated by full arrows (low-mid frequency range) and locations of a broad group of frequencies (high frequency range) have been indicated with initial and final vertical arrows linked by a horizontal double arrow. For the cholesterol-free system (right), atomic spectra $S_i(\omega)$ indicate the relative contribution of selected atoms to $S_{MEL}(\omega)$. Inset: high frequency region for most relevant atoms of MEL ('C3', 'N1', 'N2', 'H15', and 'H16').

From left side of Fig. 4.8, two main spectral regions of the full IR spectrum of MEL have been located: (1) frequencies below 1800 cm^{-1} and (2) frequencies of $2700 < \omega < 3600$ cm^{-1}. As a general feature, the agreement of the calculated full spectrum with available data (at 0% cholesterol concentration) from infrared and Raman spectroscopy reported by Singh et al.[321], Fleming et al.[322], and Pieta et

al.[323] have revealed a very good overall agreement (see the left side of Fig. 4.8), although some discrepancies have been observed in the location of some peaks. As a general fact, the maxima reported by the three experimental groups match each other very well. Let us describe the comparison of experimental data with results from the present work as follows.

In the region up to 1800 wavenumbers, the maxima with strong signatures were located at 401, 756, 834, 926, 1353, and 1550 cm^{-1}, according to the recent measurements of Pieta et al.[323], whereas Fleming et al.[322] reported strong maxima at 404, 508, 836, 928, 1358, 1449, and 1553 cm^{-1}. Most of these maxima are included in the $S(\omega)$ showed in the present work (Fig. 4.8) at 0% cholesterol content. We observed maxima at: 120, 294, 592, 714 (shoulder), 813, 1077, 1265, and 1408 cm^{-1}. The main disagreements are related to frequency shifts in most bands:

1. The peaks observed by us at 120 and 592 cm^{-1} are not seen in experimental data (the Raman frequency range started at around 250-300 wavenumbers in all cases);

2. The Raman band located at $\sim$ 400 cm^{-1} is found at $\sim$ 300 cm^{-1} in our spectrum;

3. The Raman bands at 756 and 834 wavenumbers are red-shifted down to 714 and 813 cm^{-1} in Fig. 4.8;

4. The Raman band centered at 926 wavenumbers has been found around 1077 cm^{-1} in Fig. 4.8;

5. The Raman bands at 1353 and 1550 cm^{-1} are also red-shifted to 1265 and 1408 cm^{-1} in the computed spectra.

In the high frequency region ($2700 < \omega < 3600$ cm^{-1}) the strong maxima reported from infrared spectroscopy measurements[321] were located at 3280 and 3302 cm^{-1} (assigned by the authors to the N-H stretching region), whereas a group of thirteen bands were located at the C-H stretching region, between 2826 and 3079 wavenumbers. The computed spectra of Fig. 4.8 reveal four bands, one of them very strong at 2930 cm^{-1}, another as a shoulder around 3070 cm^{-1} and two weak bands at 3210 and 3500 cm^{-1}.

According to the overall agreement between our results and those from Raman and infrared data, we can assume the reasonably good reliability of the potential MEL model and method adopted in this work. Accordingly, we have computed the partial spectra of each atom in MEL, in order to identify and assign the microscopical participants and type of each vibrational mode. The results are depicted in the plot on the right side of Fig. 4.8. There we computed the contribution of each individual atom to the full spectrum at 0% cholesterol concentration. The list of assignments and their physical meaning is as follows:

1. The peak located at 120 cm^{-1} in the computed spectra is directly related to oxygens 'O1' and O2' as well as nitrogen 'N1' since it appears in the corresponding spectral densities. From the previous knowledge[17, 91, 324] it should be attributed to restricted translations of the full MEL molecule.

2. The weak band located around 294 cm^{-1} along with the maxima at 345 cm^{-1} are approximately assigned to the spectrum of 'N2'. Since this is a relatively low frequency it will probably correspond to a rotational motion of the indole group of the MEL molecule.

3. The maximum at 592 cm^{-1} directly matches a maximum of 'O2' and this suggests a librational motion of the 'C4-O2' bond.

4. The maxima at 714 cm^{-1} (shoulder) and 813 cm^{-1} (attributed to the experimental bands at 756 and 834 wavenumbers) are clearly connected to 'N-H' vibrational bending motions, since they are detected as maxima in the spectra of 'N2-H16' and 'N1-H15' pairs, respectively.

5. The band centered at 1077 cm^{-1} is matched in our detailed spectra by maxima of atomic pairs 'N1-C4' and 'O1-C3'. Being this a bending-like mode, it should account for scissoring vibrations of the mentioned pairs.

6. The maxima at 1265 and 1408 cm^{-1} are observed in the partial spectra of 'C3' and 'C4', suggesting a stretching vibration of both carbons, the lowest frequency band associated with 'C4' and the highest to 'C3'.

7. In the high-frequency range, the strongest band observed by us and located at 2930 cm^{-1} is well matched by the peak of 'C3' (see inset of S_{atom}) and, according to the previous works[17, 91] it indicates a stretching vibration of hydrogens bound to 'C3'.

8. The shoulder at 3070 cm^{-1} can only be (hardly) seen at the spectral densities of 'H15' and 'N1'.

9. Finally, the two weak maxima located around 3210 and 3500 cm^{-1} should be respectively attributed to the pairs 'N1-H15' and 'N2-H16' as stretching vibrational modes of the corresponding hydrogens.

When cholesterol is included in the system, we can observe at the left side of Fig. 4.8 that some observable frequency shifts are produced, suggesting that cholesterol is able to interact with MEL in a remarkable way, affecting the vibrational motions of their atomic components. This is in good agreement with the fact reported above (see section 4.1.2.2) of the HB of MEL to cholesterol, especially at the 30% concentration.

4.2 Free energy landscapes of melatonin binding at phospholipid membranes

The calculation of Helmholtz or Gibbs free energy differences for a realistic condensed matter system is a difficult task, mainly due to the fact that the partition function of a multidimensional system is normally unknown. In the present case, our main aim is to obtain free energy differences in the process of binding/crossing of a small molecule through/at the model cell membrane. This will involve configurational changes and, consequently, it will require a considerable amount of computational time in order to explore and obtain the precise knowledge of the hypersurface of the potential energy of the system[317]. As reported in Chapter 2, a wide variety of methods have been designed for the calculation of FEL. Among them, a method called "local elevation" (Huber et al.[214]) or "conformational flooding" (Grubmüller[216]) was the initial idea that lead to Laio and Parrinello[225, 226] to introduce the concept of metadynamics as a method to explore multidimensional free energy surfaces of complex systems as a function of a finite number of the so-called *collective variables* (CV) which are *a priori* unknown. CVs are arbitrarily chosen and act as effective reaction coordinates to drive the calculations when moving the probe between (meta-)stable states surrounded by free energy barriers at the multidimensional configuration space. The right selection of the CVs is crucial for the sake of the efficiency of the calculation. However, given some drawbacks of the method, a new version was released. Well-tempered metadynamics is a variant of metadynamics able to enhance the sampling of multiple CV dimensions.

The specific interactions of MEL with cell membranes are revealed and interpreted from a free energy perspective, providing a quantitative characterization of the barriers between stable states as well as of the most relevant binding states of MEL to water and lipids. Our work employs a general methodology to define meaningful reaction coordinates and to explore the free energy landscapes for small molecules or drugs at complex biological interfaces which can be extended to study other interactions of interest between such species and charged head groups in colloidal chemistry and biology[325]. Interaction between several classes of lipids and cholesterol can play a significant role as well[292]. In such a case, a bound state can be characterized by variables such as the coordination number of the ions with lipid/cholesterol binding sites and its simultaneous coordination number with water molecules[275].

We have observed from the full MD simulations of MEL in DMPC membranes reported in section 4.1.2.4 that two orientations of MEL can be associated with three different dihedral angles involving particular molecular directions. Surprisingly, one of such dihedral angles, namely a given angle Ψ has shown an angular distribution clear enough to serve as a suitable CV. This particular angle reaches the same two stable values (1.42 ± 0.41 rad and 2.96 ± 0.18 rad) at all three cholesterol concentrations, corresponding to two preferential molecular configurations adopted by MEL, that we

called "folded" and "extended", respectively. The remaining dihedral angles evaluated showed large fluctuations and were clearly inadequate as CVs.

4.2.1 Computational details

The three usual sets of lipid bilayer systems with different amounts of cholesterol (0%, 30%, 50%) have been considered. The CHARMM-GUI tool was employed to generate the three sets of lipid bilayer systems. Each system consisted of 204 lipid chains: DMPC lipid molecules and cholesterol, 5000 TIP3P water molecules, and sodium chloride ions with zero total charge at human body concentration (0.15 M). Three series of 200 ns long MD simulations adopting the CHARMM36 force field were considered for the equilibration of the systems in order to minimize and equilibrate the three sets of MEL-membrane systems, before launching the well-tempered metadynamics simulations. All MD runs were performed using the package GROMACS/2018.3. A time step of 2 fs was used in all cases. The particle mesh Ewald method with a Coulomb radius of 1.2 nm was employed to compute long-ranged electrostatic interactions. The cutoff for Lennard-Jones interactions was set to 1.2 nm, with a switching distance of 1 nm. The pressure was controlled by a Parrinello-Rahman piston with a damping coefficient of 5 ps^{-1} and temperature was controlled by a Nosé-Hoover thermostat with a damping coefficient of 2 ps^{-1} [326, 327]. Standard MD simulations were performed at the NPT ensemble at pressure 1 atm and temperature 303.15 K, well above the transition temperature for DMPC, in order to ensure that we were simulating the liquid crystalline state. Periodic boundary conditions in three directions of space have been taken. The initial size of the simulation box, where Z-axis is the direction normal to the instantaneous plane of the membrane, was as follows: (a) for the cholesterol-free system it was of 7.9 nm $\times$ 7.9 nm $\times$ 8.5 nm; (b) for the cholesterol-rich setups it was of 7.9 nm $\times$ 7.9 nm $\times$ 11.6 nm.

After equilibrating well the system in all cases, we have run 1.4 μs metadynamics simulations to perform Gibbs free energy calculations of the binding states of MEL at phospholipid membrane bilayers. Three serials of metadynamics simulations were performed using the PLUMED2 within the joint GROMACS/2018.3-plumed tool, the NPT ensemble of the temperature of 303.15 K and pressure of 1 atm were adopted in all cases. Periodic boundary conditions in the three directions of space have been considered. The size of each simulated system has been found to be the same as those reported for MD simulations. The range of values for the positions of lipids, cholesterol, and water molecules have been found to be exactly the same as those reported in Fig. 4.9. We have defined two main CVs to describe the binding states of MEL: CV1 is the dihedral angle Ψ, as represented in Fig. 4.6 and CV2 is the distance $z_{\mathrm{MEL-DMPC}}$ between the center of mass of the MEL molecule and the center of the DMPC bilayer membrane, defining Z-axis as the direction normal to the instantaneous plane of the membrane, i.e. the plane formed by axes X and Y. From here on, we will refer to

CV2 simply as z. We should point out that in section 4.1 we have observed that MEL is normally bound to cholesterol and/or to the carboxyl group of DMPC and in a less usual manner to the phosphate groups of DMPC, but forming always a dihedral angle of $\pm\pi$ rad, i.e. at the extended configuration. Parameters of well-tempered metadynamics simulations are listed in Table 4.5.

Table 4.5 Simulation parameters

Parameter	0%	30%	50%
Gaussian width of CV1 [rad]	0.35	0.35	0.35
Gaussian width of CV2 [nm]	0.30	0.30	0.25
Starting (Gaussian) hill [kJ/mol]	1.0	1.0	1.0
Deposition stride [ps]	1	1	1
Bias factor	10	10	20
Simulation time [ns]	1100	1400	1400

4.2.2 Results and discussion

4.2.2.1 Physical properties of equilibrated systems

The values for the average z-length of the full systems can be also estimated from the density profiles, as reported in Fig. 4.9.

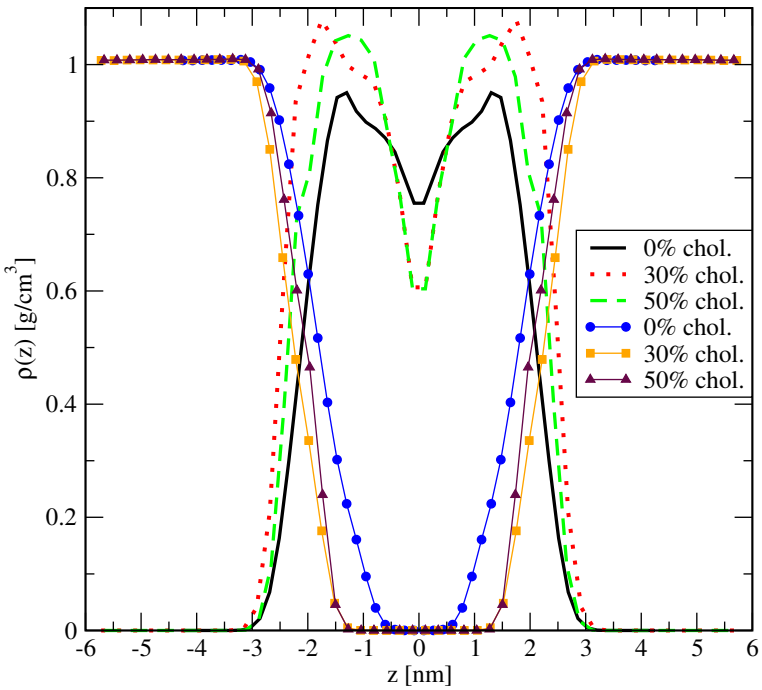


Fig. 4.9 Density profiles of lipid backbone after 200 ns of MD equilibration (full lines: 0% cholesterol, dotted lines: 30% cholesterol, and dashed lines: 50% cholesterol) and water (circles: 0% cholesterol, squares: 30% cholesterol, and triangles: 50% cholesterol)

We can highlight several general results:

1. Water is able to access deep positions inside the membrane: up to $z = \pm 1.3$ nm at the 30% and 50% cholesterol concentrations and up to $z = 0.6$ nm at 0% cholesterol, i.e. very close to the center of the membrane.

2. According to the size of the systems and density profiles, the Z-axis lengths were 8.5 nm at the cholesterol-free setup and 11.6 nm for the two setups including cholesterol. We can observe that DMPC/cholesterol molecules can reach z values up to ± 3.2 nm in all cases. This indicates that two "bulk" water boxes of about 7.9 nm $\times$ 7.9 nm $\times$ 1.1 nm i.e. 68.65 nm^3 surround the membrane for the 0% cholesterol case and two water boxes of about 7.9 nm $\times$ 7.9 nm $\times$ 2.6 nm i.e. 162.27 nm^3 enclose the membrane for the 30% and 50% cholesterol setups.

3. As the additional information, cholesterol-rich setups have membrane sizes thicker than those at the cholesterol-free system, a feature already observed in section 4.1. From these observations, we obtained average values of the area per lipid and thickness of the membrane, as reported in Table 4.6. Such values are in good agreement with the density profiles reported in Fig. 4.9. Further, the thickness of the membranes is in good agreement with the findings of Kučerka et al.[282] obtained by X-ray and neutron scattering, who reported a value of 3.67 nm at 303 K for the DMPC membrane at a cholesterol-free system.

Here we report the area per lipid and thickness of the membrane in all cases. Hence, with the information reported here, we will be able to characterize and approximately locate the states from the free energy landscapes corresponding to the water-membrane interfaces or MEL fully solvated by water/sodium-chloride molecules. Results obtained from well-tempered metadynamics simulations are reported in the following text[328].

Table 4.6 Area per lipid (A) and thickness (Δz) of the membrane. Estimated errors in parenthesis.

Percentage of cholesterol	A (nm^2)	Δz (nm)
0%	0.587(0.011)	3.65(0.05)
30%	0.413(0.004)	4.49(0.02)
50%	0.397(0.003)	4.47(0.02)

As we can see, three systems are well equilibrated and results of area per lipid generated using GROMACS/2018.3 are similar to results using the NAMD2 package (shown in Table 4.1).

4.2.2.2 Study of the free energy convergence

We report in Fig. 4.11 the fluctuations of the two CVs (Ψ, $z_{\mathrm{MEL-DMPC}}$) along the final 1 μs time span of our well-tempered metadynamics simulations. We can clearly

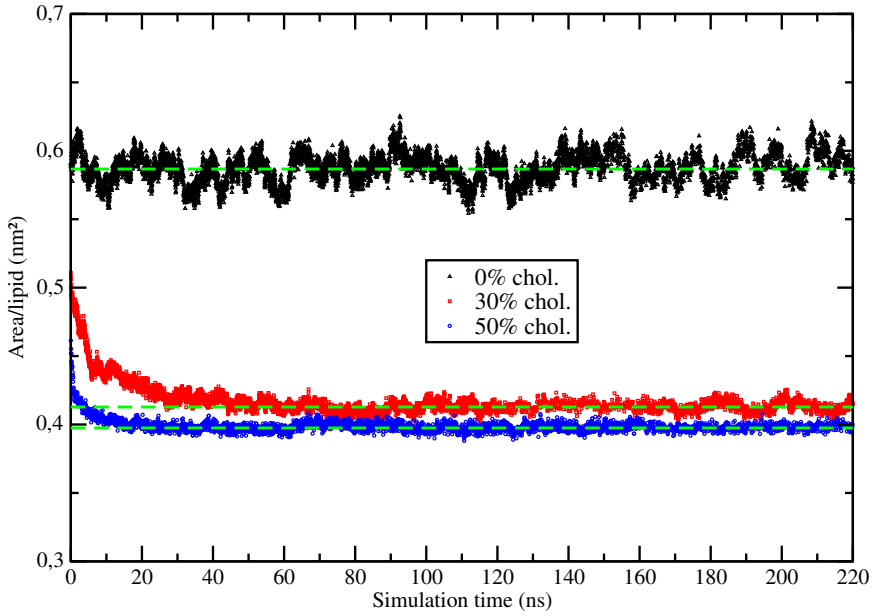


Fig. 4.10 Area per lipid of systems as a function of simulation time in three cases: 0% cholesterol (black), 30% (red), and 50% (blue). Dashed green lines indicate the average values, which are shown in Table 4.6, of the last 150 ns at each concentration.

identify a few events where MEL permeated the membrane from one leaflet to the other one at all three cholesterol concentrations. Two selected CVs potentially reached all possible values within given ranges.

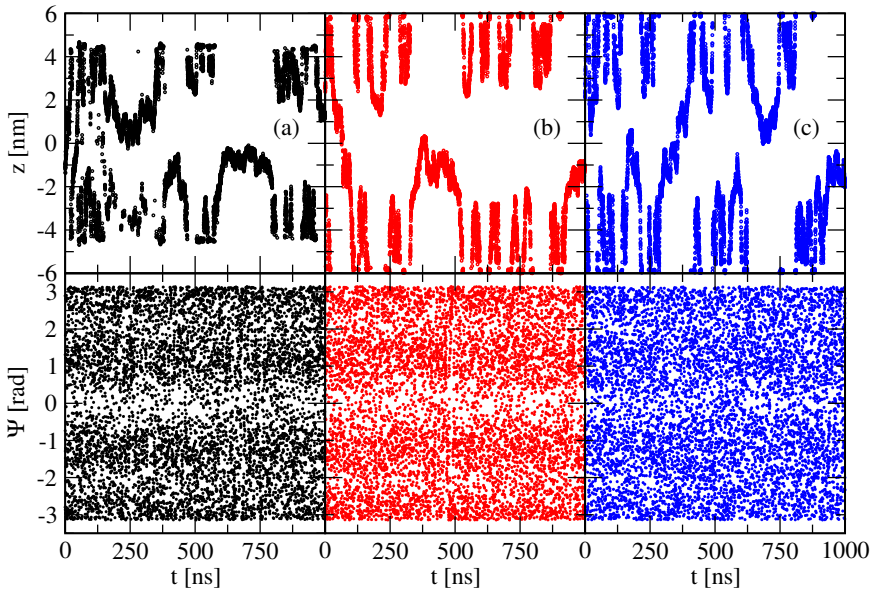


Fig. 4.11 Fluctuations of the CV values as a function of time in different states: (a) 0% cholesterol (black), (b) 30% (red), and (c) 50% (blue).

On one hand, Ψ (bottom row of Fig. 4.11) shows regular fluctuation profiles, covering all possible angular orientations in the full range $\Psi \in [-\pi, \pi]$. We note that

at 0% and 30% cholesterol concentrations, dihedral angles around 0 are slightly less common than at 50% cholesterol. On the other hand, the patterns of $z_{\mathrm{MEL-DMPC}}$ evolution in time (top row of Fig. 4.11) indicate that, although the whole distance span ($z_{\mathrm{MEL-DMPC}} \in [-4.3, 4.3]$ nm at 0%, $z_{\mathrm{MEL-DMPC}} \in [-5.8, 5.8]$ nm at 30%, and 50% cholesterol concentrations) can be reached, only some distances between the center of mass of MEL and the center of the DMPC bilayer will be accessed in a regular fashion. So, MEL is clearly biased to stay close to 1 nm away from the center of the bilayer or around 3-4 nm, which suggest that at least two preferential Z-axis locations of MEL should arise from free energy calculations. In addition, we can observe that only a few crossings of the membrane center by MEL are observed, which indicates that high free-energy barriers will prevent such crossings in a regular way.

It is usual to monitor the size of the hills of the Gaussian kernels deposited along with the simulation. We can visualize the decrease of the Gaussian height during the simulation and eventually displays low-height spikes. In the present case, the height of the biased potential decreased accordingly along the simulation runs, as indicated in Fig. 4.12. In all cases, a quasi-flat profile is already seen around 600-800 ns.

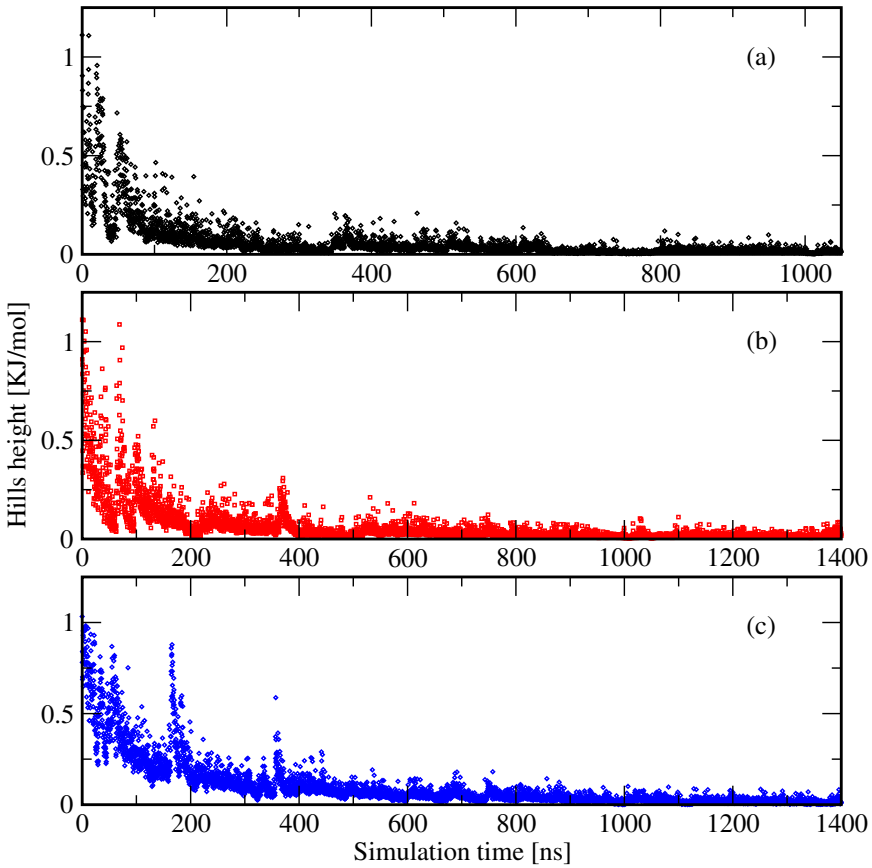


Fig. 4.12 Well-tempered Metadynamics hills height as a function of time in different states: (a) 0% cholesterol, (b) 30%, and (c) 50%.

The fact that the Gaussian height is decreasing to zero should not be used as a measure of convergence of metadynamics simulation. Therefore, in order to assess the convergence of a metadynamics simulation, we have calculated the estimate of the free energy as a function of simulation time. At convergence, the reconstructed profiles should be similar. Here we report the time cumulative average of 1D free energy profiles as defined in the main text, i.e. averaging the two leaflets and projecting onto (integrating out) the alternative CV in a range larger than 1 microsecond in all cases. From the results of Figs. 4.13, 4.14, and 4.15, we can see that after long cumulative time lengths, the differences between a profile and the one immediately before are very small (up to 2.5 kJ/mol) which lead us to fully converged free energies for the two CV considered in the present study.

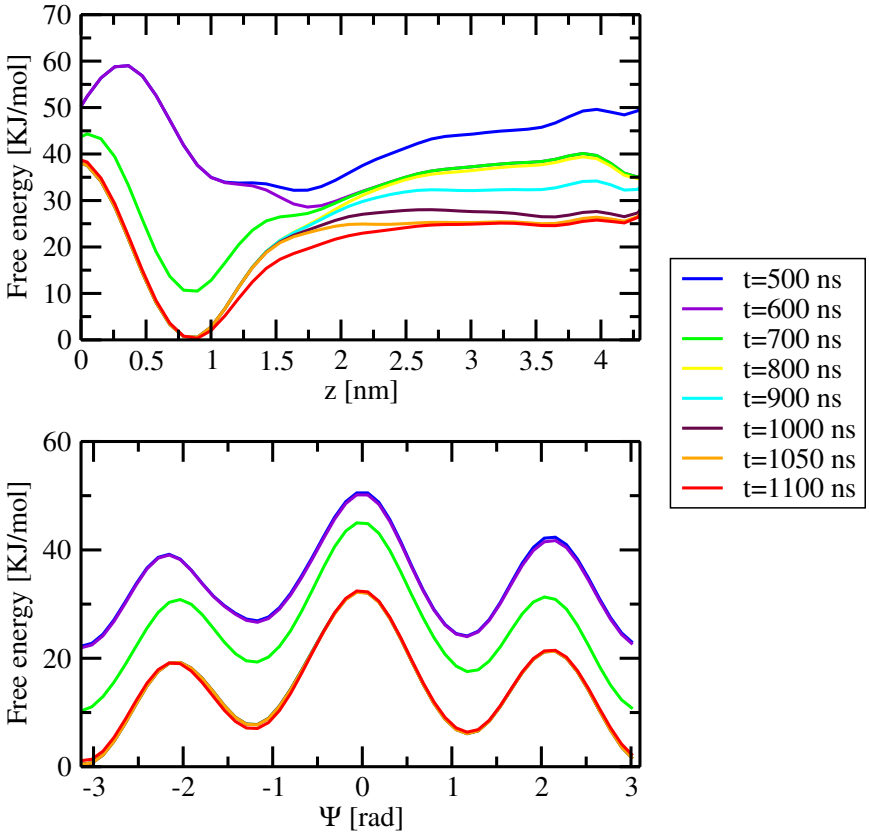


Fig. 4.13 Time cumulative free energy profiles at the cholesterol-free system. Bottom: CV1, top: CV2.

In order to obtain and evaluate the paths connecting two free energy minima located on the 2D free energy surfaces reported here, we have considered the practical guidelines reported by B.Ensing et al. in Ref. [329]. There, the authors describe precisely how to trace the minimum free energy pathway between two stable states in a multidimensional free energy landscape computed using metadynamics. The process includes three steps: (a) optimizing two local minima; (b) locating a coarse

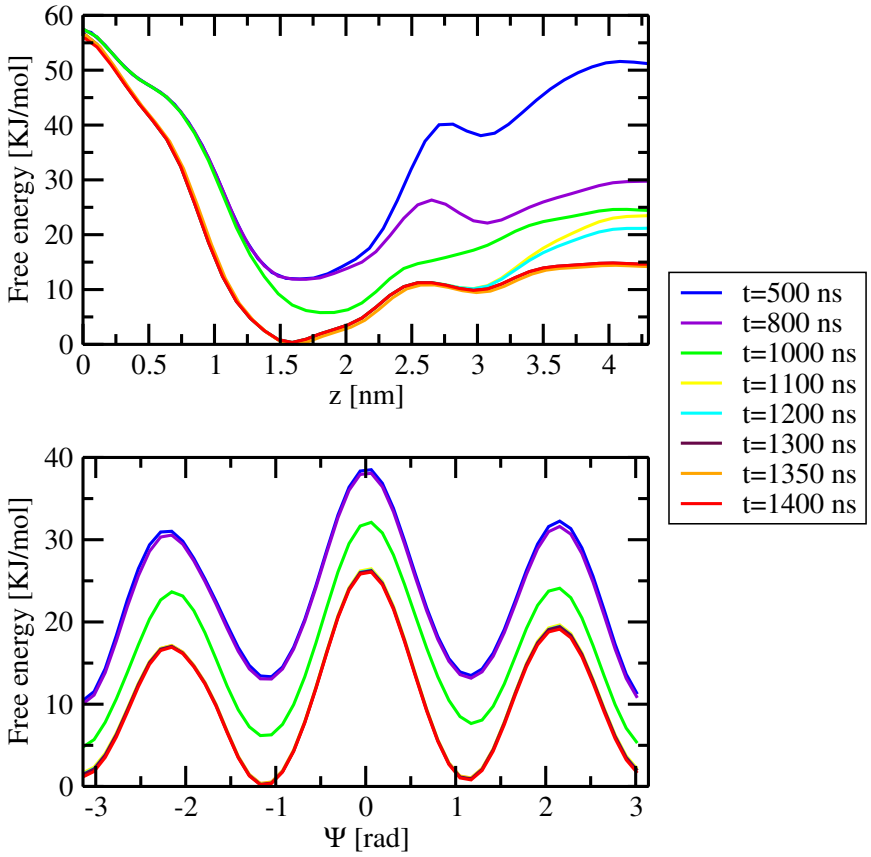


Fig. 4.14 Time cumulative free energy profiles at the 30% cholesterol system. Bottom: CV1, top: CV2.

path, and (c) refining the path. In all calculations, we considered between 8 and 10 points per path. The results for a series of paths between consecutive minima for the three cholesterol concentrations considered in the present work have been depicted in Figs. 4.16-4.18 of the manuscript and their coordinates are numerically reported in Table 4.7. The first coordinate of each point corresponds to CV1 (angle Ψ) and the final coordinate to CV2 (distance $z_{\mathrm{MEL-DMPC}}$).

4.2.2.3 2D Free energy landscapes

Three sets of two-dimensional (2D) well-tempered metadynamics simulations based on the specific CVs defined above were performed to calculate free-energy surfaces of MEL at cholesterol contents of 0%, 30%, and 50% at neutral zwitterionic DMPC membranes. The resulting 2D free energy surfaces of MEL bound to DMPC membranes are shown in Figs. 4.16, 4.17 and 4.18. Each state can be indexed by the two CVs. A pattern including several regions with clear minima is present in the FEL in all cases. The main features are the global minima of the FEL located between $z \in [0.7, 2.5]$ nm and around two distinctive values for the dihedral angle, namely those around $|\Psi| \sim [1.2,\ \pi]$ rad. Such orientations are in excellent agreement with the average values of Ψ obtained

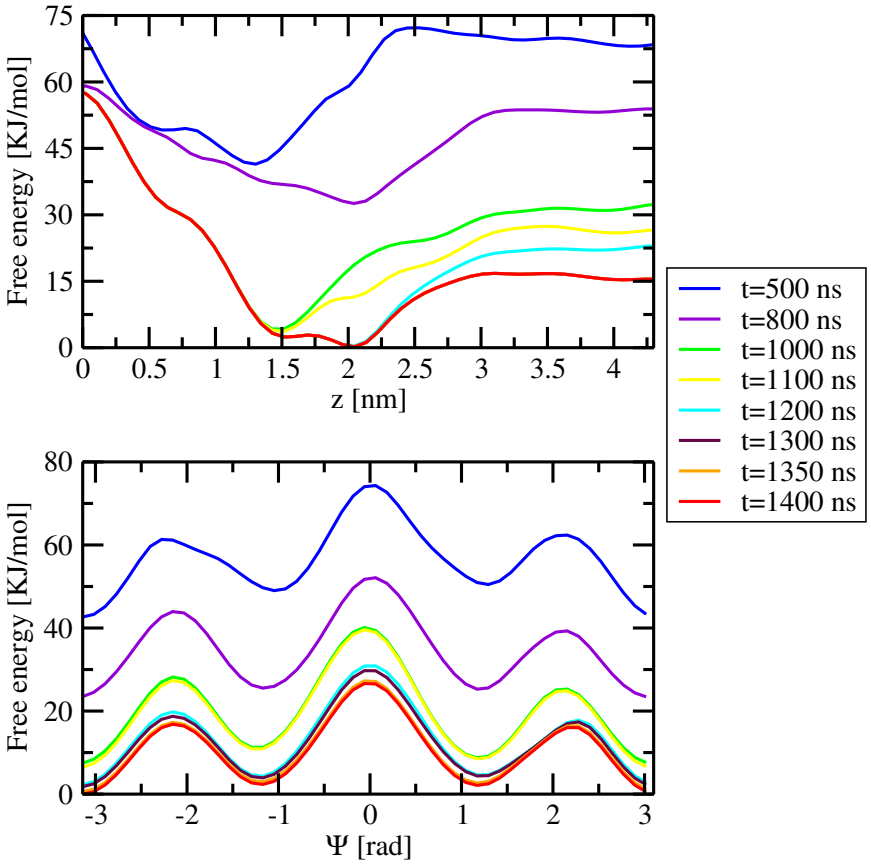


Fig. 4.15 Time cumulative free energy profiles at the 50% cholesterol system. Bottom: CV1, top: CV2.

from ordinary MD simulations[330] (see section 4.1.2.4 therein) and they correspond to the two folded and extended geometries of MEL previously observed. In addition, we observe from Figs. 4.17 and 4.18 that a membrane containing cholesterol shows the highest free energy barrier from the interface to $z \sim 0$ is of the order of 60 kJ/mol that prevents MEL crossing membrane bilayers in a spontaneous easy way.

we should point out that the zero of each 2D plot can be arbitrarily fixed since the physically meaningful quantities are not the absolute values of $F(\Psi, z)$ but the free energy differences. In the present work, the zeroes of each 2D plot have been set at the configuration of MEL fully solvated for water, since the extracellular bulk is a common feature of all sets, regardless of their cholesterol contents. In particular, we set the zero free energy for values of CV2 (z positions) at $\sim$ 4 nm (cholesterol-free case), $\sim$ 5 nm (for the two cholesterol-rich cases) i.e. with MEL fully solvated by water, and the values of CV1 corresponding to the lowest value of F, that was for $\Psi = 0$ in all cases. This means that, since we did not set the zero at the absolute minimum value of F for each 2D surface, we are reporting positive and negative values of F.

The overall range of absolute free energies reported in Figs. 4.16-4.18 is rather wide, similar to the range reported in the work of Jämbeck and Lyubartsev[312] for

Table 4.7 Coordinates of segments forming each path.

Minimum free energy paths						
Stable states	0% cholesterol		30% cholesterol		50% cholesterol	
	Ψ (rad)	z (nm)	Ψ (rad)	z (nm)	Ψ (rad)	z (nm)
A	-3.14	0.79	-3.14	1.64	-3.10	2.17
	-2.96	0.66	-2.89	1.80	-2.91	2.29
	-2.64	0.77	-2.70	1.78	-2.75	2.24
	-2.26	0.99	-2.37	1.64	-2.48	2.22
	-1.97	0.97	-2.04	1.62	-2.25	2.08
	-1.58	0.88	-1.62	1.64	-1.98	2.05
	-1.32	0.73	-1.45	1.63	-1.45	1.98
B	-1.17	0.88	-1.19	1.57	-1.18	2.11
	-0.87	0.84	-0.81	1.66	-0.83	2.07
	-0.59	0.89	-0.64	1.65	-0.49	2.12
	-0.29	0.95	-0.43	1.63	-0.15	2.19
	-0.09	0.97	-0.14	1.58	0.16	2.25
	0.16	1.08	0.07	1.59	0.42	2.39
	0.44	1.02	0.33	1.60	0.70	2.47
	0.74	0.95	0.66	1.39	0.96	2.41
	0.92	0.93	0.86	1.36	1.08	2.26
C	1.17	0.84	1.12	1.51	1.20	2.02
	1.29	0.90	1.32	1.57	1.37	2.24
	1.58	0.98	1.49	1.70	1.50	2.33
	1.80	0.97	1.76	1.56	1.80	2.45
	2.30	0.99	2.02	1.50	2.05	2.35
	2.59	1.00	2.31	1.41	2.48	2.37
	2.87	1.06	2.52	1.62	2.73	2.26
D	3.03	0.80	3.03	1.69	3.05	2.19

small-molecules (ibuprofen, aspirin, and diclofenac) around lipid bilayers, who found free energy ranges of the order of 70 kJ/mol and barriers of around 20-40 kJ/mol. On the other hand, we estimated from reversible work calculations (based on considering radial distances as reaction coordinates) barriers for MEL forming or breaking hydrogen bonds to DMPC lipids to be in between 2 and 10 kJ/mol, when MEL is located inside the interfacial region, but free energy barriers related to orientational changes of MEL and to global displacements of MEL to the center of the membrane or to the extracellular bulk were not considered. Given this wide free energy range, we will focus especially on the characterization of the free energy barriers between the particular states of MEL at interfaces and MEL fully solvated by water, where the raw differences are in the range of 10-25 kJ/mol.

The 2D free energy landscapes reveal that the most favorable stable states of MEL binding to the membrane (basins A, B, C, and D) correspond to z-distances around 0.8 nm at the cholesterol-free system, whereas such distance tends to increase significantly around to 1.6 nm for the 30% cholesterol concentration and up to 2.2 nm

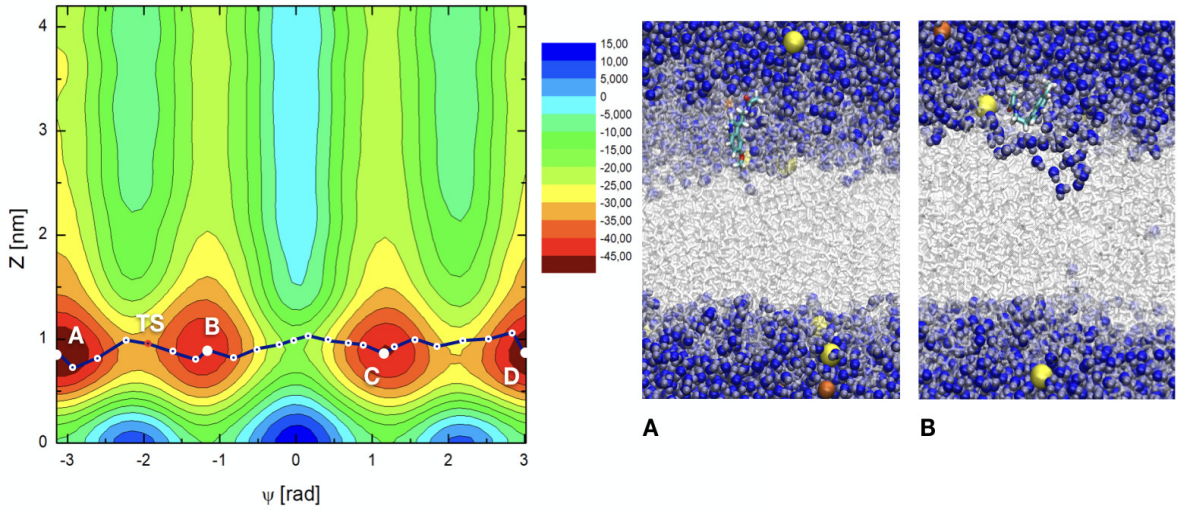


Fig. 4.16 2D free energy landscapes $F(\Psi, z)$ (in kJ/mol) in the cholesterol-free case. Four stable state basins (A, B, C, and D) are indicated. TS indicates a (local) transition state between basins A and B. The minimum free energy path is shown in dark blue, indicating all the computed points of the path, as listed in Table 4.7. Snapshots 'A' and 'B' correspond to selected basins A and B on the free energy hypersurface, respectively. In between basins A and B we can locate a transition state corresponding to the conformational angular change of MEL between extended (basin A) and folded (basin B) configurations. DMPC (white), sodium ions (yellow), chlorine (dark orange), water (blue) and for MEL: carbon (cyan), nitrogen (dark blue), oxygen (red), hydrogen (white).

when cholesterol reaches 50% of the total amount of lipids in the system. Considering the information revealed by CV1, i.e. the torsional angle Ψ, we can distinguish two sets of minima: (1) For $|\Psi| \sim 1.2$ rad (basins B and C) and (2) for $|\Psi| = \pi$ rad (basins A, and D). These minima are related to the two preferential configurations of MEL close to a DMPC-cholesterol bilayer (folded, extended) indicated above. According to CV2, i.e. the distance between the center of mass of the MEL molecule and the center of the DMPC bilayer membrane, MEL is preferentially located at the interface of the DMPC-cholesterol bilayer (regions with $0.7 < z < 3.0$ nm).

Locations of MEL outside the interface i.e. ($z < 0.7$ and $z > 3.0$ nm) show large free energies. These regions will be considered as (1) "water-solvated" ($z > 3.0$ nm, assuming that it in such regions MEL is fully solvated by the electrolyte solution surrounding the membrane) or (2) "internal" regions of the membrane ($z < 0.7$ nm, assuming that in such cases MEL is fully embedded into the body of the membrane). In Chapter 3 of this Thesis, we observed that TRP, a molecule quite similar to MEL (and one of its precursors) spent about 30% of unbiased MD simulation runs at the "water-solvated" state and the remaining time adsorbed at the interface. This state can be considered as a normal, accessible configuration of MEL. As it can be observed in Figs. 4.16, 4.17, and 4.18 when MEL is fully solvated by water ($4.2 > z > 3$ nm for

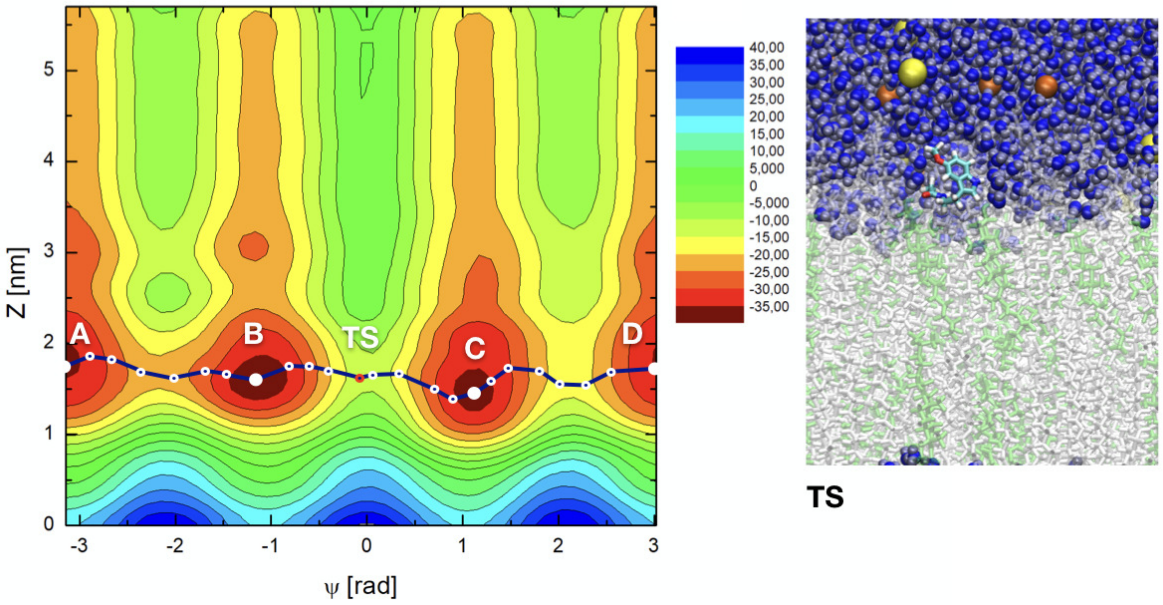


Fig. 4.17 2D free energy landscapes $F(\Psi, z)$ (in kJ/mol) in the case of 30% cholesterol. Minimum free energy path pictured as in Fig. 4.16. Snapshot 'TS' corresponds to the global TS on the free energy surface. Colors as in Fig. 4.16, with cholesterol chains depicted in green.

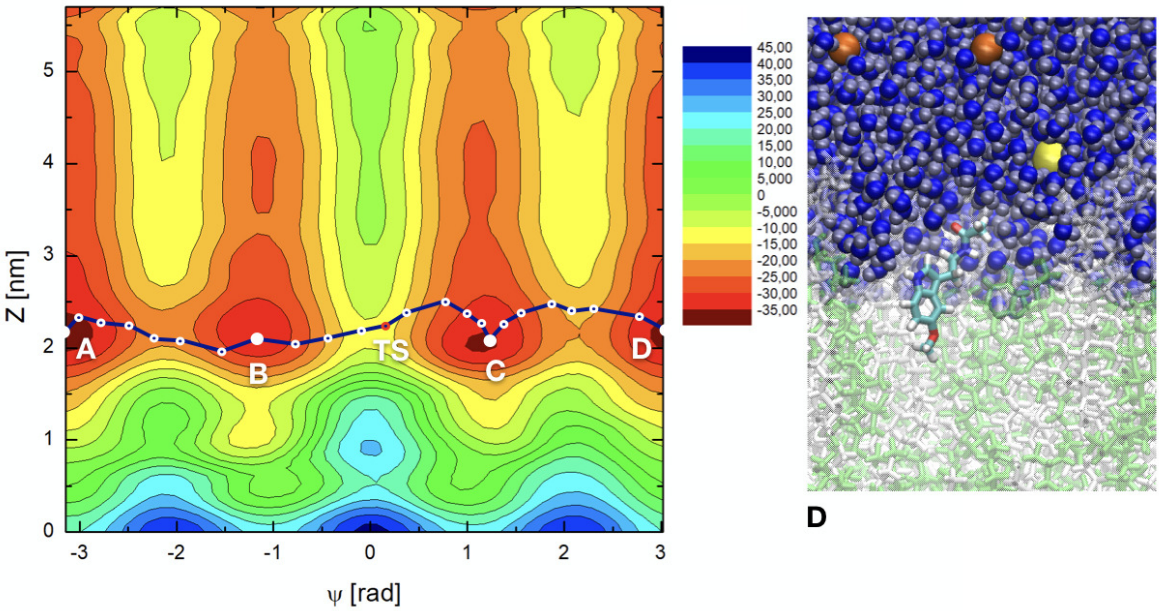


Fig. 4.18 2D free energy landscapes $F(\Psi, z)$ (in kJ/mol) in the case of 50% cholesterol. 'TS' corresponds to the global TS on the FEL. Snapshot 'D' corresponds to selected basin D. Colors as in Fig. 4.17.

the cholesterol-free case, $5.7 > z > 3$ nm at 30%, and 50% cholesterol) it tends to stay in a quite high free-energy conformation, far away from the corresponding minima,

located at the stable state basins. In this situation, MEL can change its orientational order with a moderately low free energy cost. Something similar happens when MEL reaches very internal regions of the membrane ($z < 0.7$ nm).

Since the present 2D representation includes a wide variety of low free energy states related to conformational and structural changes, some specific numerical calculations are in order. In particular, it is very useful to compute the so-called *minimum free energy path* (MFEP), that can be determined by iteratively refine a pathway connecting stable states that converges to the minimum free energy trajectory between them (see References[329, 331–333]).

From MFEP we can extract information associated with the most probable trajectories in the CV space followed by the system when evolving between stable states and when eventually crossing high energy barriers that can be associated with local or global transition states (TS) located at the free energy hypersurface. Here we assume the canonical interpretation give by Transition State Theory of the (global) TS as a saddle point in between two stable states located at the minimum free energy path and having the largest energy[334, 335]. In the following list, we will report the size of free energy barriers at two given states. These barriers can be compared to experimental data (see for instance Ref. [333]). The main features and numerical estimations can be summarized as follows:

1. Measuring the value of the absolute free energy directly from the 2D plots, we can identify the approximated location of TS revealed by the MFEP:

   (a) In Fig. 4.16, the local maximum labeled TS is at the pair of coordinates ($\Psi = -1.97$ rad, $z = 0.97$ nm);

   (b) The global TS in Fig. 4.17 is located at ($\Psi = -0.14$ rad, $z = 1.60$ nm) and

   (c) The global TS in Fig. 4.18 corresponds to ($\Psi = 0.16$ rad, $z = 2.42$ nm).

2. We can estimate the location and values of the free energy barriers between the significant states of "MEL at interface" and "water-solvated" (taken in all cases at $z = 4$ nm) respectively for each cholesterol concentration. Let us note coordinates as (CV1 in rad, CV2 in nm):

   (a) At 0% cholesterol, MEL at interface corresponds to the coordinate (-π, 0.79) and when water-solvated corresponds to (-π, 4.0), with a barrier is of 25.3 kJ/mol;

   (b) At 30% cholesterol, the two coordinates are (-1.19, 1.64) and (-1.19, 4.0) with a barrier of 14.1 kJ/mol;

   (c) At 50% cholesterol, the two coordinates are (-π, 2.27) and (-π, 4.0) with a barrier of 9.1 kJ/mol.

3. Other relevant free energy barriers are located between the free energy minima (let us take basin 'A' as the reference) and the center of the membrane ($z = 0$) in all cases, with values of the order of 40 kJ/mol (0% cholesterol) and of 50 kJ/mol (30% cholesterol and 50% cholesterol). These barriers are difficult to be crossed spontaneously by MEL, but they are interesting in order to estimate the amount of free energy such that MEL can permeate the membrane. As expected, when cholesterol is present in the membrane, the barriers are considerably higher than for the cholesterol-free case.

4. Concerning angular-related barriers, there is a relevant one between the two angular stable states of MEL (basins A and B), i.e. folded ($|\Psi| = 1.17$ rad) and extended ($|\Psi| = \pi$) conformations of 18.8 kJ/mol for the cholesterol-free case, 19.7 kJ/mol at 30%, and 17.6 kJ/mol at 50% cholesterol concentrations. This corresponds to MEL at the interface. This conformational barrier corresponds to surmount the local TS indicated in Fig. 4.16. The corresponding barrier when MEL is solvated by water is of 18.3 kJ/mol.

5. Other TS can be obtained for the conformational change between angles $\Psi = -1.17$ rad and $\Psi = 1.17$ rad (i.e. between basins B and C in Figs. 4.17 and 4.18). The corresponding barriers are of 22.0 kJ/mol (30% cholesterol) and 28.3 kJ/mol (50% cholesterol). These can be considered the absolute TS of the system since they show the maximum free energy along the reaction coordinate, indicated by the MFEP. The equivalent global TS of Fig. 4.16 (0% cholesterol), has a barrier of 26.5 kJ/mol. In all cases, MEL crossing of a TS can be carried out through an intermediate configuration of MEL with $\Psi = 0$, as represented in the snapshot reported in Fig. 4.17. When MEL is solvated by water, the corresponding barrier is of about 21 kJ/mol for each case.

The visible shifts to larger distances z of basins A, B, C, and D shown in Figs. 4.16-4.18 clearly reveal that the binding competitiveness of lipid head groups has been diminished by cholesterol contents and that the affinity of MEL to DMPC membranes becomes less favorable as cholesterol concentration increases, which is in agreement with previous experiments and simulations on MEL and cholesterol at variable concentrations close to DOPC and DPPC membranes[96, 97], where it was observed that cholesterol helps a membrane to increase its thickness, reduce its area per lipid, whereas MEL operates in the reverse way and tends to decrease the thickness of the membrane and increase its area per lipid.

In order to show characteristic molecular configurations of MEL when located at the interface of the membrane, two characteristic snapshots at 0% and 30% cholesterol concentrations are reported in Fig. 4.19. We have included in the figure only MEL and its closest solvating lipids, namely two DMPC at 0% and one DMPC plus one

cholesterol at 30%. Both configurations correspond to the dihedral angle $\Psi = \pi$ and to z distances of 0.8 (0%) and 1.6 nm (30%). With the aid of these snapshots we can see that MEL is able to be simultaneously bound to both DMPC and cholesterol chains through HBs. The binding sites were determined to be oxygens 'O6' of DMPC and the oxygen from the hydroxyl group of cholesterol and hydrogens 'H15' and 'H16' of MEL (see Reference[330] for further details).

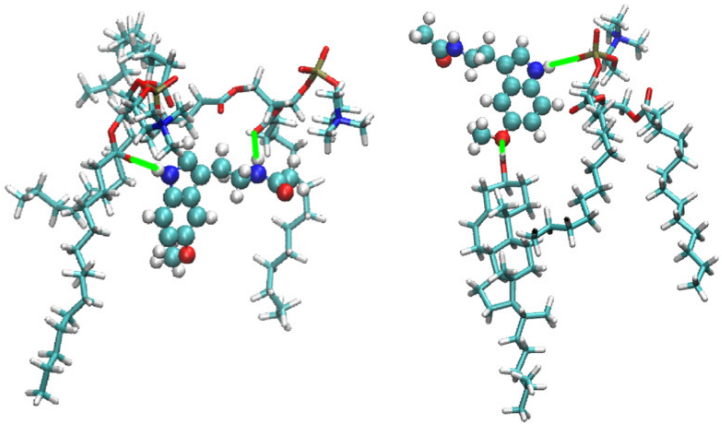


Fig. 4.19 Snapshots of representative bound states for MEL at the interface of the membrane. Colors of MEL as in Fig. 4.6. Colors for DMPC/cholesterol: carbon (cyan), nitrogen (dark blue), oxygen (red), hydrogen (white). Hydrogen bonds indicated as green solid lines. Left: case of 0% cholesterol concentration, with MEL bound to two DMPC molecules; Right: case of 30% cholesterol concentration, with MEL bound to one DMPC plus one cholesterol molecule.

According to point (1) of the list above, the energetic cost for MEL to move from the interface to the water-solvated region diminishes when cholesterol concentration increases. This means that cholesterol favors MEL to move outside the membrane and eventually to stay in the outer regions of the interface. To understand these changes, we should remember the well-known condensing effect of cholesterol on lipid bilayers, which produces higher membrane rigidity and ordering[336]. Interestingly, such transition in the binding behavior of MEL at cholesterol $\approx$ 30% might be correlated with the phase transition point of DMPC-cholesterol membranes, in which membranes change from a liquid-disordered phase to a liquid-ordered phase[291, 337]. Concerning the angular configurations, our results indicate that changing from extended to folded forms does not require to surmount a high energy barrier, but one of about 19 kJ/mol when MEL is located at the interface of the membrane. Gibbs free energy barriers of about 15 kJ/mol have been also obtained in simulations of the translocation of ethanol across a lipid bilayer membrane[338].

4.2.2.4 1D Free energy profiles

As an alternative to the 2D FEL and MFEP reported above, we considered the calculation of 1D free energy profiles as a function of only one of the two CV. This will allow us to directly compute free energy differences for meaningful MEL conformations, especially for MEL at the interface and at the center of the membrane. In Figs. 4.20 and 4.21 we represent the dependence of the integrated binding free energy $F(s1)$ on one CV for membranes with three cholesterol concentrations after integrating out the second CV according to Refs. [312, 339]:

$$F(s_1) = -k_{\mathrm{B}}T \ln \left[\frac{\int e^{-\beta\, F(s_1,s_2)}\, ds_2}{\int e^{-\beta\, F(s_1,s_2)}\, ds_1 ds_2} \right], \tag{4.1}$$

where s_1 and s_2 are the CVs, $\beta = 1/(k_{\mathrm{B}}T)$, k_{B} is the Boltzmann constant and T is the temperature. This means that all possible paths for the CV labeled as s_2 have been averaged. $F(s_1)$ reveals additional information about the most stable states only as a function of one CV. This also allows us to compare the relative stability of MEL at setups with different cholesterol contents separately as a function of z or Ψ. These barriers can be also directly compared to experimental findings[312].

In Fig. 4.20, we set the zero of each free energy at the extracellular bulk ($z = 4$ nm at 0% cholesterol and at $z = 5$ nm at 30% and 50%), as we did in the 2D free energy surfaces. Since in Fig. 4.21 the spatial dependence has been integrated out, we have fixed the zero to the minimum free energy in each case.

In agreement with the preliminary findings reported in the previous section, the observed behavior is that the position of stables states of MEL depends on the contents of cholesterol in the membrane: minima of $F(z)$ are located at $z = 0.8$ nm for 0% cholesterol, at $z = 1.6$ nm for 30% cholesterol, and at $z = 2.2$ nm for 50% cholesterol.

From plots in Figs. 4.20 and 4.21 we can obtain meaningful free energy barriers, i.e. free energy differences related to particular processes. We can highlight the following:

1. MEL needs about 25.5 kJ/mol to move from inside the interface of the membrane to the water-solvent regions at 0% cholesterol concentration, whereas it requires 14.3 kJ/mol at 30% cholesterol and 16.7 kJ/mol at 50% cholesterol (Fig. 4.20). The net effect of cholesterol is to reduce the energetic cost of MEL moving across the membrane from the interface to the aqueous solvent.

2. The barrier that needs to be surmounted for MEL to move from the interface to the center of the membrane is significantly high, between 40 (0%) and 50 kJ/mol (30% and 50%) (see Fig. 4.20). Even though these large numbers can be a drawback of the method employed here, it is very clear that in the process of crossing the membrane from one leaflet to the other, cholesterol enhances the free energy cost. We expect that increasing temperature[338] or pressure[92] it would be possible to observe the crossing of such barriers by MEL.

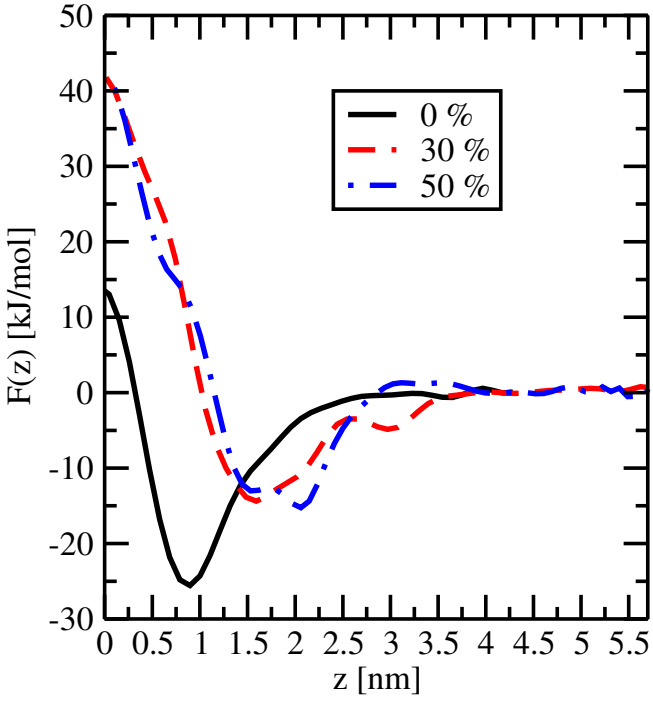


Fig. 4.20 Integrated free energy F(z) for 0%, 30% and 50% cholesterol.

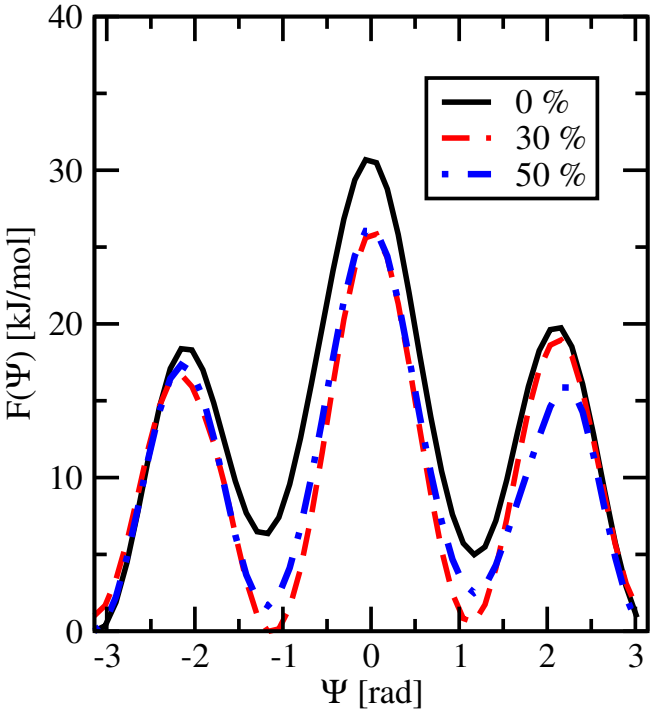


Fig. 4.21 Integrated free energy F(Ψ) for 0%, 30% and 50% cholesterol.

3. Further, in order to reach a folded configuration ($|\Psi| = 1.17$ rad) from the extended configuration ($|\Psi| = \pi$ rad) (Fig. 4.21), MEL has to surmount barriers of 18.4 (0%), 15.7 (30%), and 17.3 kJ/mol (50%). This means that for MEL it is

quite easy to adjust its molecular geometry in order to exchange its configuration between two orientational stable states. We should add that in standard molecular dynamics[330], it is easy to observe that MEL can change its configuration varying the dihedral angle Ψ and that it can move from interfaces of the DMPC-cholesterol membrane bilayer to the water-solvated region and vice-versa.

4. Finally, the access to the conformational structure with $\Psi = 0$ demands to surmount a barrier of about 25 kJ/mol from a folded configuration. We should point out that this is essentially the cost of crossing a global TS because in such a case the coordinate z changes very little (see point (5) of the list reported in section 4.2.2.3).

In processes such as interactions of MEL with free radicals[340], it has been estimated from DFT methods that Gibbs free energy changes from vacuum to aqueous solution were in between 33 and 670 kJ/mol. For the mechanism of MEL nitrosation (reaction of MEL with nitric oxide)[341] the Hartree-Fock estimated free energies were of 50-59 kJ/mol *in vacuo*. Further, the Gibbs free energy of MEL binding to the protein calmodulin has been estimated to be around 36 kJ/mol, 3-fold larger than the experimental value[342]. These findings indicate the difficulty of obtaining precise values of free energy barriers. Thus, these numerical estimations report values for different binding processes of a similar order of magnitude to the results presented in this work. From the experimental side, Florio et al.[343] used a variety of techniques such as a combination of two-color resonant two-photon ionization, laser-induced fluorescence excitation, resonant ion-dip infrared spectroscopy, fluorescence-dip infrared spectroscopy, and UV-UV hole-burning spectroscopy, to explore the conformational preferences of an isolated MEL molecule under molecular beams. In such a system, these authors found MEL three *trans* and two *cis* conformers showing a free energy gap of about 12.5 kJ/mol, quite close to the values obtained in the present work for the conformational angular changes (folded to extended and vice-versa). Finally, Florio and Zwier[320], analyzed the solvation of MEL by water clusters using infrared and ultraviolet spectroscopy and found barriers of about 63 kJ/mol for the *cis/trans* isomerization process. All these findings indicate the difficulty of obtaining precise values of free energy barriers in general and for processes where MEL is involved, in particular. Thus, some of these numerical estimations report values for different binding and conformational processes of the order of magnitude similar to the results presented in this Chapter.

4.3 Conclusions

In this Chapter, results of all-atom molecular dynamics simulations of a bilayer membrane made up with the zwitterionic phospholipid DMPC at three cholesterol

concentrations (0%, 30%, and 50%) including a single MEL molecule, all in aqueous NaCl ionic solution at 303 K and at the fixed pressure of 1 atm have been reported. In the first part, our main interest was focused on the local structure and angular distributions of MEL, especially when associated with DMPC and cholesterol molecules. After this, the Gibbs free energy differences of MEL binding onto different membrane bilayer interfaces have been evaluated through the potentials of mean force. As a general fact, MEL is able to establish HBs with water, cholesterol, and DMPC lipids. Typical HB through different sites between MEL-DMPC and MEL-cholesterol have been observed.

Two relevant MEL structures have been observed from angular distributions: "folded" and "extended" configurations. We defined three different dihedral angles to account for the two preferential angular configurations. Dihedral Ψ defined in Fig. 4.6 has revealed to be a meaningful order parameter (i.e. may act as a reliable reaction coordinate) to describe the dynamics of MEL, with preferential angles of ~ 1.42 rad and ~ 2.96 rad. We suggest that introducing cholesterol into the system could help MEL change from its folded configuration to the extended configuration more easily, using HBs between MEL-DMPC and MEL-cholesterol. The self-diffusion coefficient of MEL was found to be of the order of 10^{-7} cm^2/s and the presence of cholesterol in the system has little influence on it. The spectral densities of MEL computed in this work are in overall good agreement with experimental Raman and infrared data and have revealed the degree of participation of each atomic site of MEL to complete its whole molecular spectrum giving some clues to understand the microscopic origin of molecular vibrations and also giving evidence of the good reliability of the model we adopted in the present work.

Nevertheless, in the time span of 200 ns of auxiliary MD simulations, we did not record any spontaneous crossing of MEL through the DMPC membrane. Given the open debate on the mechanisms of MEL permeation through cells (see References[93, 94]), our results appear to be in qualitative disagreement with the reported experimental permeability of MEL across the pineal gland plasma membrane[92], of the order of 1.7 μm/s. We believe that the disagreement is essentially due to the pressure difference between both sides of the plasma membrane applied in the experiments by Yu et al.[92] and that we did not consider in the present work where, as indicated in SI, we performed our simulations at the NPT ensemble considering a constant pressure of 1 atm at the two sides of the membrane and for the whole system. In addition, the experimental setups consider real plasma membranes containing a large variety of species such as large trans-membrane proteins that we have not considered here, and that may favor the crossing of the membrane by MEL. Finally, in the recent work of Wang et al.[338] small solutes such as glycerol, caffeine, iso-propanol, or ethosuximide were simulated nearby a model cell membrane. These authors found that, in order to observe trans-membrane crossings of such small solutes in the time length of a

simulation at the atomic level of description, they needed to run trajectories of 10 μs at low temperatures (310 to 330 K) or, alternatively, raise the temperatures to more than 400 K (for simulation times of 1 μs).

To the best of our knowledge, we conducted the first quantitative characterization to date of the binding states of MEL at model DMPC-cholesterol phospholipid cell membranes through the calculation of free energy landscapes. With the help of the well-tempered metadynamics simulations we have calculated 2D free energy landscapes and located the binding stable states and several (local and global) transition states of the system. Zero of all free energies were set at the extracellular bulk, i.e. for MEL fully solvated by water and the ionic solution. Two CVs have been considered: dihedral angle Ψ and the distance $z_{\mathrm{MEL-DMPC}}$ between the center of mass of MEL and of the center of the lipid bilayer (given by $z = 0$). Our results indicate that MEL can be bound to the internal side of the membrane, at distances $z \sim 1 - 2$ nm, and in several stable state configurations where the dihedral $|\Psi| = 1.17$ rad and $|\Psi| = \pi$ rad. From the calculation of the minimum free energy paths, we have observed that the most probable trajectory of MEL is along with the interface of the membrane, changing its conformation between extended and folded configurations by means of surmounting free energy barriers of about 15-20 kJ/mol. In addition, the methodology employed in the present work allowed us to locate several global and local TS of the system and to estimate the values of their free energy barriers.

When cholesterol is added to the system, it helps pull MEL escape from the interface of the membrane to being solvated by water/sodium chloride molecules more easily by decreasing its corresponding free energy barriers. The effect is more marked as the concentration of cholesterol rises. The energetic cost for MEL to leave the interface of the membrane and to be fully solvated by water/sodium chloride molecules (z-distances around 4 nm), has been estimated to be of between 10 and 25 kJ/mol. The less common situation was found to be with MEL accessing regions around the center of the membrane, process requiring to cross free energy barriers above 40 kJ/mol. We believe that the findings presented in this Chapter could be of practical use in designing new reaction coordinates (such as for instance coordination numbers of MEL) for multidimensional, more accurate free energy calculations able to explore a wide variety of relevant small solute species of biochemical interest such as amino acids, neurotransmitters, drugs or hormones. The same methodology can be applied to systems of larger size such as big peptides or proteins: for instance, a detailed study of the absorption and binding of the oncogenic protein KRas-4B in DOPC/DOPS/cholesterol membranes is currently under investigation in our laboratory. Detailed results will be reported in the next Chapter.

Chapter 5

Structural and energetic aspects on the anchoring of wild-type and oncogenic KRas-4B proteins at model cell membranes

As I described in section 1.3 of this Thesis, Ras is a family of related proteins which is expressed in all animal cell lineages and organs. They play an essential role in signal transduction pathways which ultimately turn on genes involved in cell growth, differentiation, and survival so that overactive Ras signaling can lead to cancer[344]. As a surface protein anchored in the inner leaflet of the cell membrane, Ras proteins are normally in the inactive state. They are activated following an incoming signal from their upstream regulators. Among Ras proteins, KRas-4B has been found being expressed at higher levels in lung, colorectal, and pancreatic cancer cells[112, 120, 345]. These insights suggest the importance of fully understanding the regulation of the oncogenic KRas-4B activity when binding to the membrane.

Accumulating evidence indicates that demethylated and farnesylated KRas-4B-Far could play an important role in the signaling pathway that happens on the inner leaflet of the membrane bilayers. KRas-4B-Far has been reported to be able to be transferred to bind the inner PM leaflet[146, 346, 347]. While there is a relatively high abundance of KRas-4B-Far (wild-type and oncogenic) in tumors, the effects on downstream signaling have yet to be determined[110]. While most efforts have been focused on the characterization of the methylated KRas-4B-FMe binding to the PM[117, 142, 146, 158, 348, 349], the study on the KRas-4B-Far is an emerging area of research. According to Barcelo at el.[142], phosphorylation at Ser-181 of oncogenic KRas is required for tumor growth. Three sequences of the KRas-4B proteins are listed in Fig. 5.1, in which in the present Thesis we applied mutation G12D at Gly-12 and phosphorylation at Ser-181 (PHOS) for oncogenic ones.

Fig. 5.1 Three sequences of different KRas-4B structures. Mutated sites are in red color. Here C_f denotes the farnesylated Cys-185 and S_p represents the phosphorylated Ser-181 (see Fig. 5.3). Methyl group of KRas-4B-FMe is in blue color.

Our work in this Chapter shall provide essential information in gaining a precise understanding of the effects of oncogenic mutations on the effects of localization of oncogenic mutations in KRas-4B proteins. Understanding the structural specifics of KRas-4B in its GTP-bound form will help to design the oncogenic KRas-4B inhibitors. Finding a method of stabilizing the interactions between KRas-4B and the membrane with a good knowledge of the structural mechanisms at the atomic level might be a promising target for anticancer drug discovery. Also, the GTP affinity of KRas-4B is reported to be extremely high, with a dissociation constant (K_d) at around 10^{-11} mol/L[350], yet the corresponding binding free energy has been poorly explored. On its own, genomic sequence data may not provide the entire information to the oncologist in the targets selection. The idea of calculation of free energy landscapes is compelling and this enables us to map many possible conformations in which the protein could populate at the interface of the cell membrane. The treatment recommendations might derive from the FEL calculation. Insight into the underlying biological mechanisms of protein-membrane binding could foster molecular targeting in a significant way[152].

Ras association with membranes is not a one-way street and they undergo a cycle of delivery to the PM followed by returning to endomembranes for recycling[351]. In this Chapter we have applied well-tempered metadynamics to reveal GTP's affinity to KRas-4B and lipids and FAR's binding free energy barriers of anchoring to/departure from the membrane bilayers from a free energy perspective, especially for membranes with the high cholesterol content. The calculations reported here offer very detailed

FELs of FAR and GTP of different KRas-4B sequences, hoping to provide novel and useful insights for this field.

5.1 Methods

We conducted MD simulations of model cell membranes constituted by DOPC (56%), DOPS (14%), and cholesterol (30%) for three systems including a KRas-4B protein (see below). Each system contains a total of 304 lipid molecules and one protein in aqueous solution, yielding a system size of 222,000 atoms. All MD inputs were generated using the CHARMM-GUI membrane builder and CHARMM36m force field. The crystal structures of KRas-4B with a partially disordered hypervariable region (pdb 5TB5) and GTP (pdb 5VQ2) were used to generate full-length KRas-4B proteins. Both pdb files were downloaded from RCSB PDB Protein Data Bank[352]. Three sets of the full-length KRas-4B proteins (oncogenic KRas-4B-FMe and KRas-4B-Far, wild-type KRas-4B-Far) and a GTP molecule were solvated in a water box, with runs of 20 ns that generated fully equilibrated setups. KRas-4B was set anchoring to the membrane at the beginning of simulation in each system. Systems were solvated and KCl added to a final concentration of 0.15 M which serves as the intracellular liquid, and final systems contained about 60,000 H_2O. The TIP3P model was used for water molecules. All systems were energy minimized for 5000 steps followed by three 125 ps simulations and then four other 500 ps equilibrium runs while gradually reducing the harmonic constraints on the systems. Production runs were performed with an NPT ensemble for 1 μs. The pressure and temperature were set as 1 atm and 310.15 K, respectively, well above the corresponding T_m for DOPC and DOPS lipids, in order to ensure that we were simulating the liquid crystalline state using the GROMACS/2018.3 software package. Time steps of 2 fs were used in all production simulations and the particle mesh Ewald method with a Coulomb radius of 1.2 nm was employed to compute the long-ranged electrostatic interactions. The cutoff for Lennard-Jones interactions was set to 1.2 nm. The pressure was controlled by a Parrinello-Rahman piston with a damping coefficient of 5 ps^{-1} and the temperature was controlled by a Nosé-Hoover thermostat with a damping coefficient of 1 ps^{-1}. Periodic boundary conditions in three directions of space have been taken.

After equilibrating well the system in two cases of wild-type/oncogenic KRas-4B-Far from the production runs of 1 μs for each system, we switched to run another 1 μs well-tempered metadynamics simulations to perform Gibbs free energy calculations of the KRas-4B binding at anionic phospholipid membrane bilayers from the last configuration of MD simulations. Well-tempered metadynamics simulations were performed using the PLUMED2 within the joint GROMACS/2018.3-plumed tool. The NPT ensemble of the temperature of 310.15 K and the pressure of 1 atm was adopted in all cases. Periodic boundary conditions in the three directions of space have been

considered. Two CVs were considered to describe the conformational changes along with free energy paths. CV1 is the distance "gtp" between the center of mass of GTP and the center of mass of the membrane, and CV2 is the distance "far" between the farnesyl group of Cys-185 and the center of mass of the membrane along with the direction normal to the membrane center as depicted in Fig. 5.2. And "wt." and "onc." represent the "wild-type" and "oncogenic", respectively in all Figures and Tables reported throughout this Chapter.

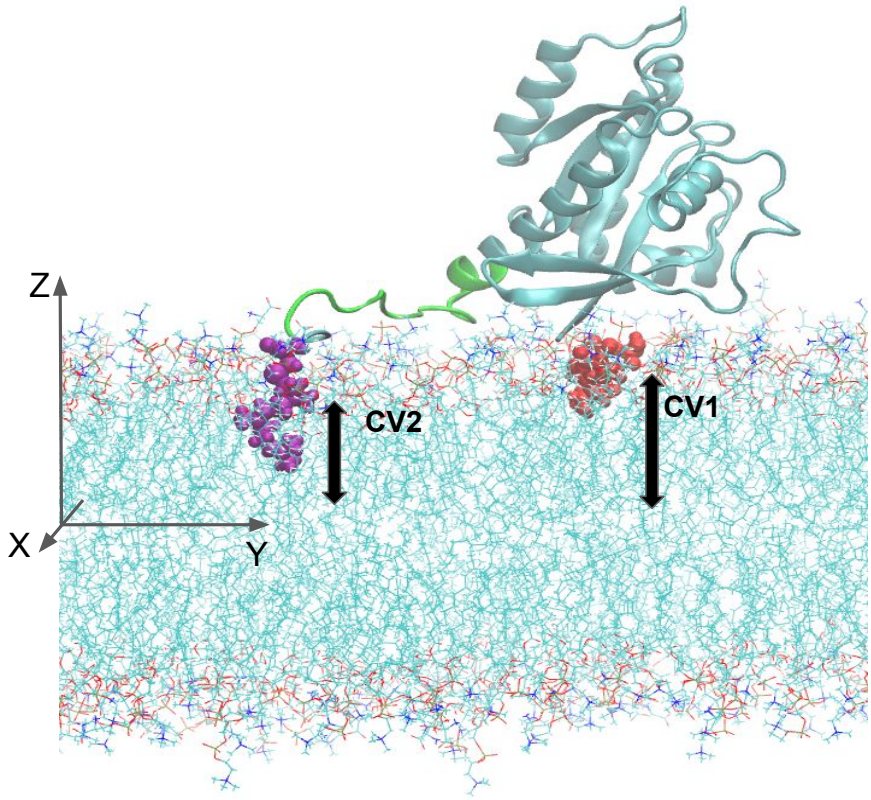


Fig. 5.2 Graphical expressions of two CVs. the CD of KRas-4B (cyan), the HVR (green), GTP (red) and FAR (purple) are shown here.

Parameters of well-tempered metadynamics simulations are are listed in Table 5.1.

Table 5.1 Metadynamics simulation parameters.

System	onc. KRas-4B-Far	wt. KRas-4B-Far
Gaussian width of CV1 [nm]	0.10	0.10
Gaussian width of CV2 [nm]	0.35	0.35
Starting (Gaussian) hill [kJ/mol]	2.0	1.2
Deposition stride [ps]	1	1
Bias factor	10	5
Simulation time [ns]	1000	1000

5.2 Results and discussion

5.2.1 Molecular dynamics of the KRas-4B binding at anionic membranes

5.2.1.1 Physical characteristics of the membrane in two systems

In order to explore the structural characteristics of the anchoring of the KRas-4B proteins at anionic membranes, several physical properties of the membrane have been calculated. Selected sketches of the structures of DOPC, DOPS, cholesterol, GTP, et al. are represented in Fig. 5.3.

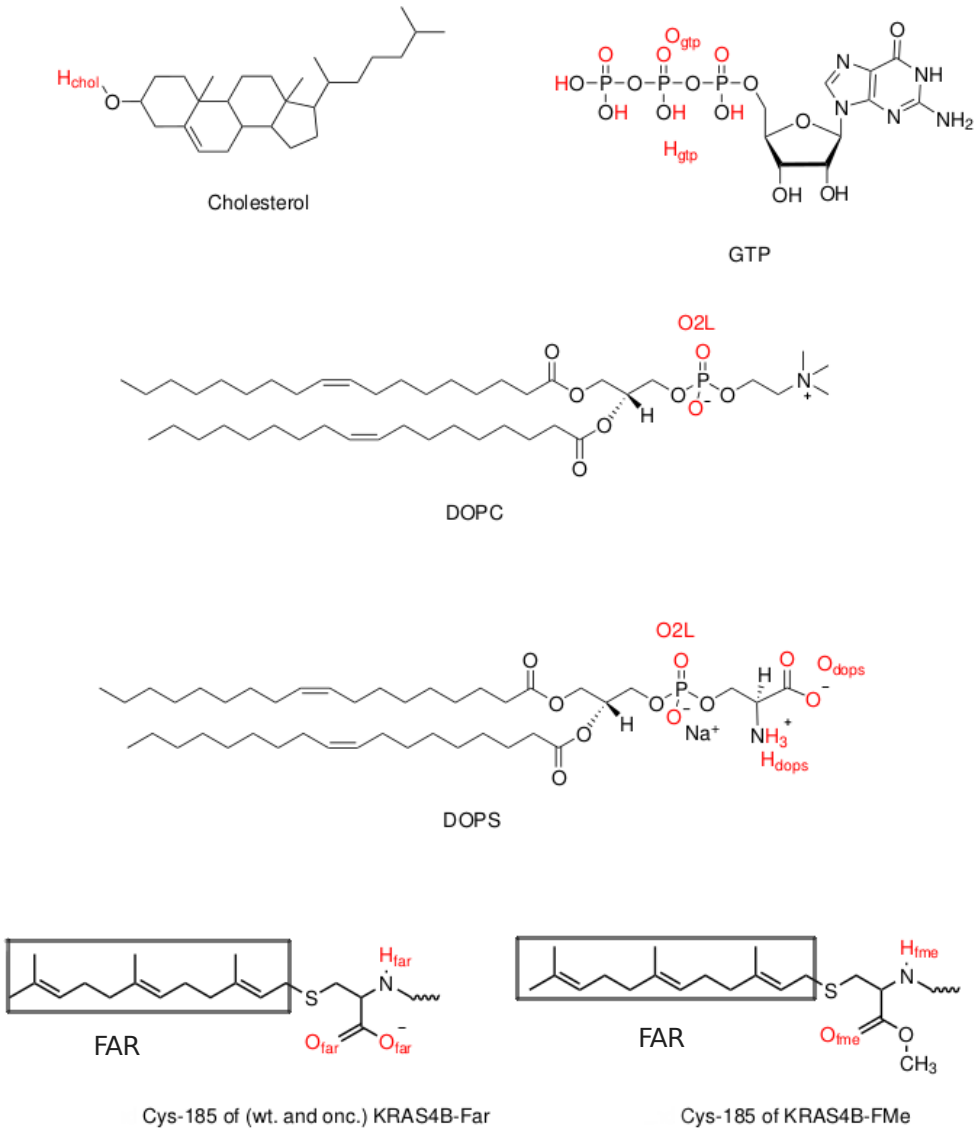


Fig. 5.3 Sketches of the backbone structures of cholesterol, GTP, DOPC, DOPS, and Cys-185 sites are shown here. Parts of Hydrogen-Carbon bonds aren't shown. The highlighted and renamed sites in red color will be referred to in the text.

A deuterium order parameter S_{CD} was defined for each CH_2 and CH groups of the DOPC lipid tails as Eq. 3.6 and averaged results are shown in Fig. 5.4 for both tail chains of all DOPC lipids in all KRas-4B systems studied in this Thesis. The order parameters of the first half of the chain of DOPC are directly related to the area per lipid, whereas the whole chain is directly related to its length, i.e. to the thickness of the membrane bilayers. By comparing with S_{CD} results of pure DOPC bilayers, the tendency has a great agreement with results provided in Refs.[353, 354] in all cases. Given the addition of 30% cholesterol in all systems studied in this Chapter, at 1 atm and 310.15 K the liquid phase was safely reached in all cases, since the transition temperature of DOPC T_m = 253.15 K and for DOPS T_m = 262.15 K[41].

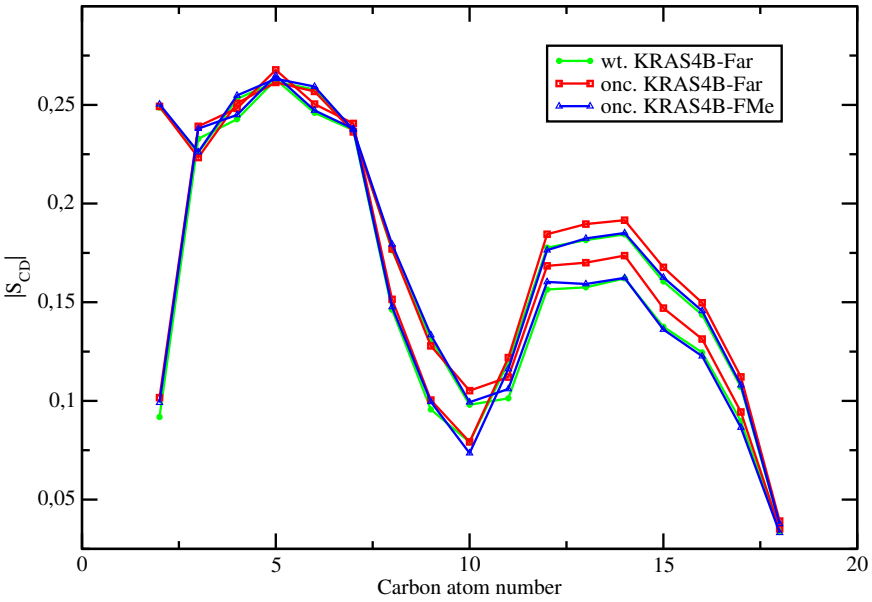


Fig. 5.4 Averaged S_{CD} for sn-1, sn-2 chains of DOPC in all systems.

Area per lipid is often used as the key parameter when assessing the validity of MD simulations. It has been proposed that a good test for such validation is the comparison of the area per lipid and thickness of the membrane with experimental data obtained from scattering density profiles[355]. For continuous MD production runs the area per lipid, defined in Eq. 3.7, as a function of simulation time, has been reported in Fig. 5.5 and their averaged values together with the related thickness and the averaged values are reported in Table 5.2.

Table 5.2 Area per lipid A and thickness Δz of the anionic membrane for all the KRas-4B systems studied in this Chapter. Estimated errors in parenthesis.

Systems	A (nm^2)	Δz (nm)
wt. KRas-4B-Far	0.523 (0.007)	4.35 (0.05)
onc. KRas-4B-Far	0.525 (0.006)	4.23 (0.04)
onc. KRas-4B-FMe	0.524 (0.006)	4.34 (0.05)

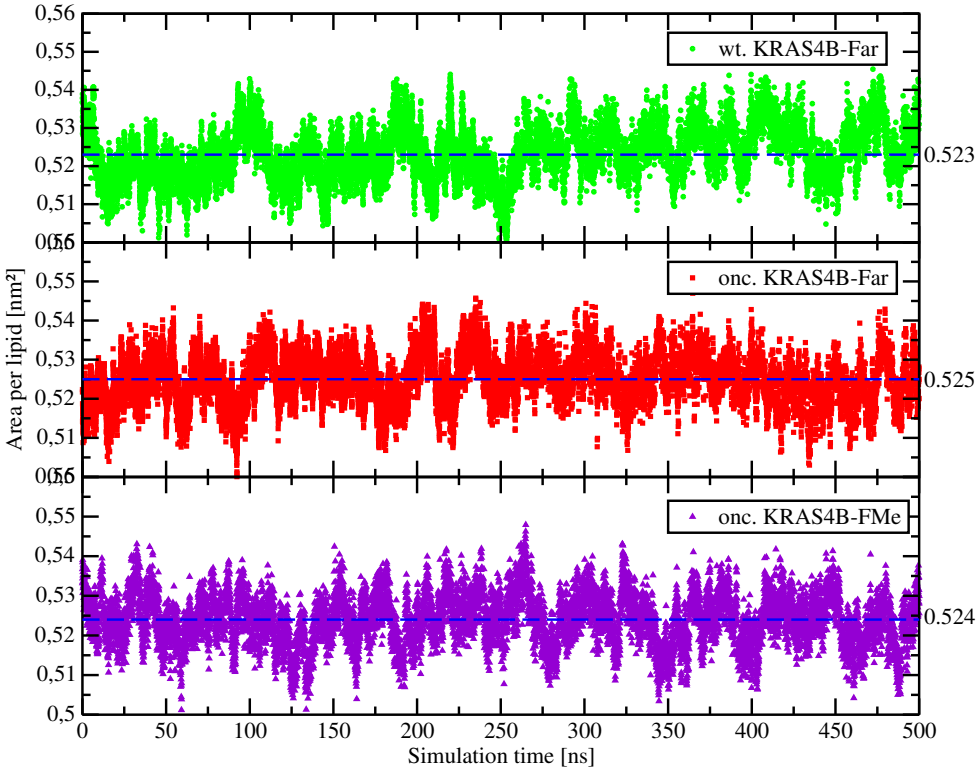


Fig. 5.5 Area per lipid of systems with different KRas-4B sequences as a function of simulation time. The blue dashed line indicates the average value for each system of the second half of the total 1000 ns production runs.

Obviously, different sequences of KRas-4B do not have much influence on the area per lipid of the membrane. X-ray results obtained for the DOPC lipid[356] suggested that temperature has a pronounced effect on the lipid area. As temperature increases, the fluorine atom oscillates more perpendicularly to its bonds. So increasing the temperature of the system leads to an increased area per lipid of certain phosphatidylcholine lipids as it was observed for temperatures below 420 K[357]. In Chapter 4 we have observed that adding cholesterol to the membrane could decrease its area per lipid and increase its thickness, which we call the condensing effect. Upon the addition of cholesterol, the area per lipid decreases more than one would expect from ideal mixing[22]. According to an earlier research[358], in the case of relative high cholesterol concentration, 10 $\sim$ 20% smaller area per lipid will be considered to be reasonable and close to equilibrium ones. From our work reported in section 4.1.2.1, compared to pure DMPC bilayer, the area per lipid of DMPC with cholesterol (30%) has been decreased by 32% from 0.62 to 0.42 nm^2, see Table 4.1. In the regime ($\leq$ 30%), the area per lipid has been reported to decrease sharply as cholesterol is added into the system[46]. As obtained from the analysis of X-ray scattering in the low angle and wide-angle regions at 303 K, the area per lipid of the membrane composed of DOPC:cholesterol (7:3) was reported by Nagle et al.[359] with a value of 0.544 nm^2 which is close to the value we

are reporting here for all systems. The DOPS bilayer was found to be thicker than the DOPC bilayer with a correspondingly smaller cross-sectional area which gives a smaller value of area per lipid and DOPS is suggested to have a condensing effect on DOPC bilayers as well[358]. We obtained a value of 0.52 nm^2 which is 26.7% less than the experimental value of 0.71 nm^2 for area per lipid of DOPC/DOPS (4:1) at 297 K[40] and is consistent with the experimental analysis by Nagle et al. The reason shall be the contribution of the joint effect of raising the system temperature and condensing effect by the addition of cholesterol and DOPS molecules. Adding cholesterol to the bilayer has much more influence on the area per lipid than raising the system temperature and the substitution of DOPS.

The thickness of the membrane may provide additional clues about its mechanical properties, such as rigidity and capability of allowing the anchoring/release of FAR from the membrane. Together with locations of FAR and GTP along with the membrane normal, a better understanding of how deep FAR and GTP insert into the anionic membrane could be achieved. We have obtained the thickness of the membrane Δz by computing the mean distance between phosphorus atoms of the DOPC head groups from both leaflets. From Table 5.2 we could see that Δz of the oncogenic KRas-4B-Far system has decreased slightly. Since we suppose that interactions between the mutated KRas-4B-Far and lipid head groups might play a role in the decreasing of thickness, a detailed explanation will be provided in the next section. Reported by E.Novakova[360], for DOPC/DOPS (4:1) bilayers at 303 K the experimental value of thickness is of $\Delta z = 3.94$ nm. The results make much sense when compared with the experimental values confirming cholesterol's condensation effect on DOPC/DOPS membrane bilayers. Additional data on the thickness of DOPC:cholesterol (7:3) bilayers reported by Nagle et al.[359] of 3.99 nm appears to be larger than the value we report here. We believe that this is due to the condensing effect produced by the presence of DOPS[358].

The results of the area per lipid and thickness of membrane bilayers we have investigated are in good agreement with experimental values. Hence, the validity of MD simulations reported here has been established, regarding the structural characteristics of the membrane.

5.2.1.2 Radial distribution functions of certain atomic species

As described in section 3.2.2.2, normalized radial distribution functions (RDF) $g_{12}(r)$ for two different species '1' and '2' have been investigated. We have considered several relevant RDFs between highlighted sites of molecules (see Fig. 5.3) and involved sites of the CD and the HVR together with the corresponding molecular structure of amino acids of KRas-4B (see Fig. 5.6 and Fig. 5.7).

The selected $g(r)$s are reported in Fig. 5.8-Fig. 5.11.

In three panels of Fig. 5.8, HBs between GTP and active sites of specific amino acids of the CD have been observed. The clear first coordination shell which located around

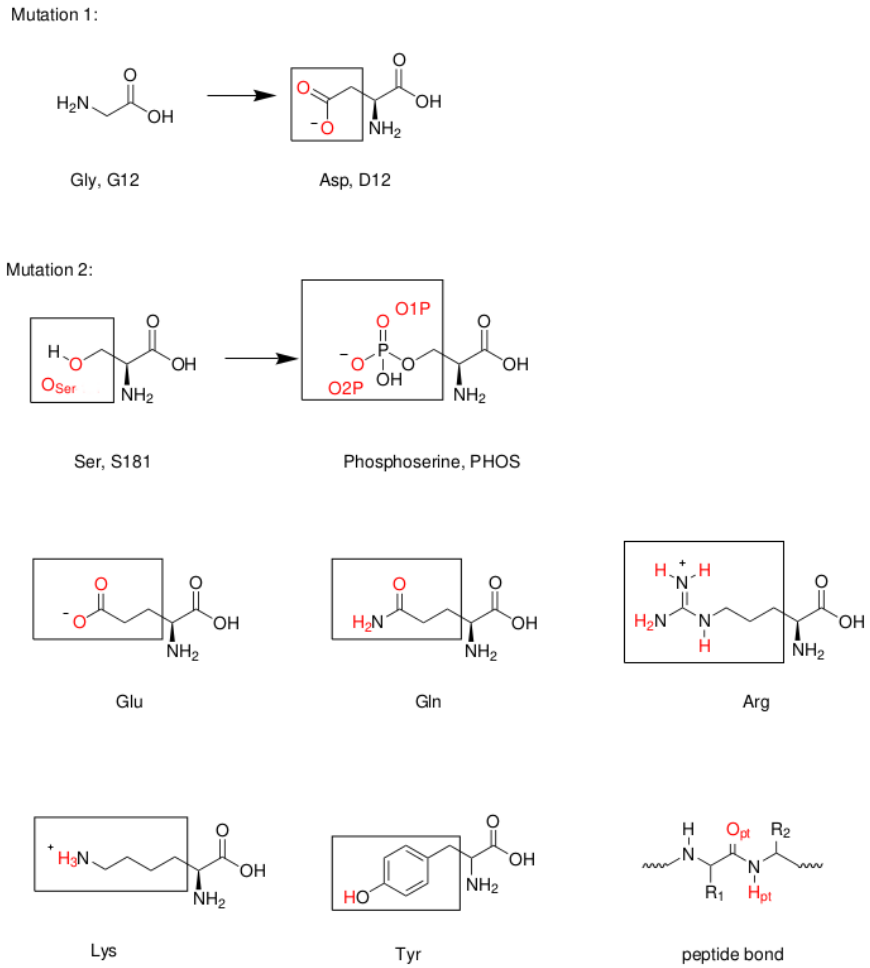


Fig. 5.6 Two mutated sites in structures of the oncogenic KRas-4B-Far/KRas-4B-FMe and the active amino acids of KRas-4B with their side chains framed. All highlighted and colored in red sites will be inferred in the following text.

1.75 Å shall be essentially attributed to HBs between selected atom sites indicated in the Fig. 5.3, Fig. 5.6 and Fig. 5.7. The typical signature of oxygen-hydrogen HB is 1.8 Å[100]. For the oncogenic KRas-4B systems, GTP tends to bind its the CD through HBs and no HBs between GTP and lipids have been observed. We could see that GTP prefers binding to the CD of the oncogenic KRas-4B proteins regardless of carboxymethylation at site Cys-185 before anchoring to the inner membrane leaflet. In particular, for the wild-type KRas-4B-Far, HBs between GTP and the CD are much stronger ($\sim$ 15-fold) than HB between GTP and lipids. By analyzing Fig. 5.8, we could propose that in oncogenic systems, GTP preferentially binds with the CD of KRas and locates away from lipid bilayers. However, in the wild-type system, GTP may prefer to locate at the interface but much closer to the CD. Especially strong interactions between GTP and the CD of the KRas-4B-FMe indicate less efficient nucleotide exchange for this mutant (see panel "C").

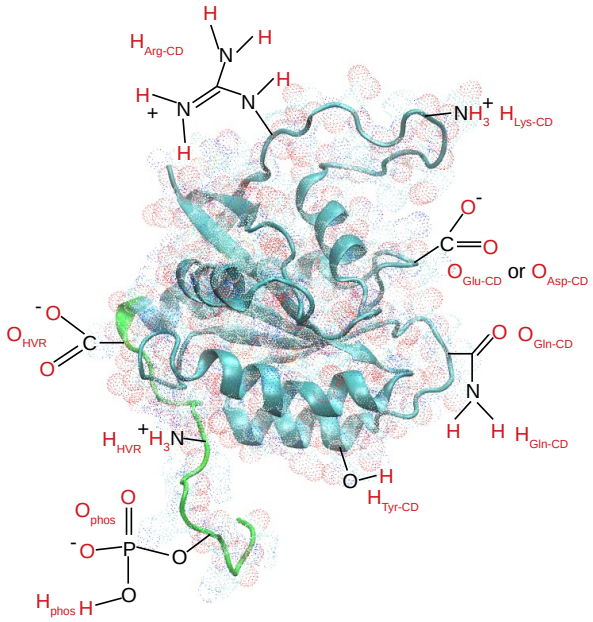


Fig. 5.7 Selected active sites of the CD (cyan) and the HVR (green) of the KRas-4B structure are depicted in red color. Thereinto, H_{Arg-CD} donates the cationic hydrogen atoms from the guanidinium ($RNHC(NH_2)_2^+$) of arginine of the CD that share the same atom type, H_{Lys-CD} and H_{HVR} represent the cationic ammonium ($RNH_3{}^+$) from lysine of the CD and of the HVR. O_{Glu-CD} and O_{Asp-CD} represent oxygen atoms of the anionic carboxylate ($RCOO^-$) of glutamate amino acid of the CD and Aspartate amino acid of the CD, respectively. O_{HVR} represents all oxygen atoms of the anionic carboxylate ($RCOO^-$) of the HVR. O_{Gln-CD} and O_{Gln-CD} donate active oxygen and hydrogen atoms of the side chain of glutamine amino acid of the CD. Moreover, O_{phos} and H_{phos} donate two active oxygen atoms and one hydrogen atom of the phosphate group from PHOS. All highlighted and colored in red sites will be inferred in the following text.

To further analyze the effect of the side chain of the CD, calculations were performed and reported in Fig. 5.9. We have revealed that the CD is able to associate head groups of lipids at the atomic level which proves that the CD of KRas-4B plays a role in binding to bilayers for all cases. In the $g(r)$s, when the two selected oxygen and hydrogen sites are ionized, the location of the first shell usually is smaller than 1.8 Å, here we call the contribution of the first shell is from the so-called $salt-bridge$[361, 362] and the coupling of salt-bridges was reported to be of general importance to the stability and function of proteins. A salt-bridge has two components: a hydrogen bond and an electrostatic interaction and of all the non-covalent interactions, salt-bridges are among the strongest. Here we report that the first shell of $g(r)$ of salt-bridges between O_{CD} and H_{dops} locates at 1.65 Å for three systems. Also in panel "C", we could obtain that HB between the cationic ammonium ($RNH_3{}^+$) from lysine of the CD of KRas-4B-FMe

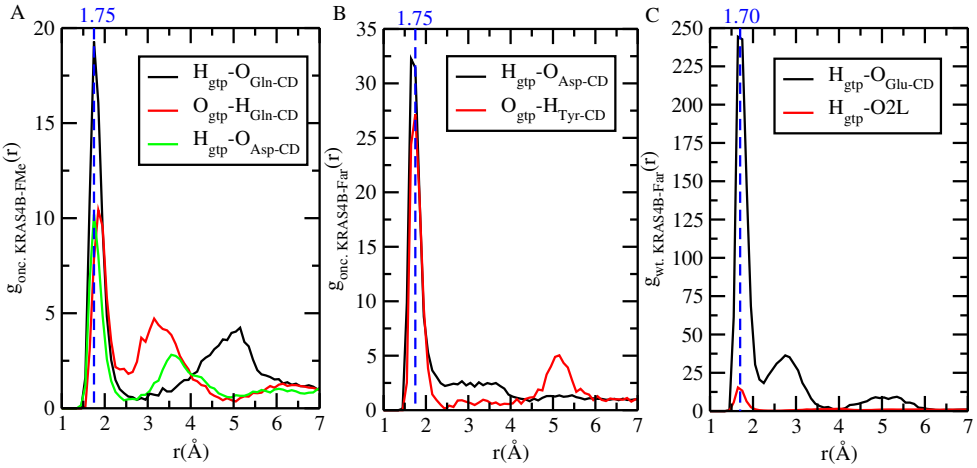


Fig. 5.8 Selected RDFs for active atoms of GTP with selected sites of the CD and active oxygen atoms of head groups of lipids ('O2L'). 'H_{gtp}' and 'O_{gtp}' represent hydrogen and oxygen atoms from phosphorus group of GTP.

and anionic oxygen atoms from DOPS is much stronger than demethylated KRas-4B isoforms.

Consecutively, RDFs related to the HVR have been displayed (Fig. 5.10). There we can see that the HVR is able to form stable HBs/salt-bridges with the CD, lipids, and PHOS. Obviously, carboxymethylation and mutations show their influence on the interactions of the HVR, DOPS, and PHOS as it can be observed in panels "D" and "F". With the first shell located at 1.65 Å, salt-bridge between O_{HVR} and H_{dops} is much weaker for the oncogenic KRas-4B-FMe comparing with the demethylated ones. After phosphorylation (as shown in Fig. 5.6), for the oncogenic KRas-4B proteins O_{phos} could form stable salt-bridges with hydrogen atoms of cationic ammonium from the HVR (H_{HVR}). For the wild-type KRas-4B-Far, only weak HB exits between H_{HVR} and oxygen atom from the side chain of site Ser-181. Different preferential interactions between the HVR and lipids and the CD could play a role in the different behaviors of three types of KRas-4B and its conformational shifting. As pointed out in Refs. [363, 364], we observed that the CD and the HVR also participate in PM binding but engaging differently during the process.

From Fig. 5.3 we can expect that in oncogenic system FAR will form strong and long-lasting salt-bridges with PHOS through negatively charged O_{far} of FAR interacting strongly with the positively charged H_{dops} of DOPS.

We have also investigated meaningful RDFs of active sites of FAR with PHOS, CD, and lipid molecules (see Fig. 5.11).

For the mutated KRas-4B-FMe, typical HB is observed between the carbonyl oxygen of the ester group of Cys-181 (O_{fme}) and positive hydrogen of the hydroxyl group of Tyr from the CD (H_{Tyr-CD}). FAR can interact with cholesterol through typical HB and DOPS through strong salt-bridge for the wild-type KRas-4B-Far which

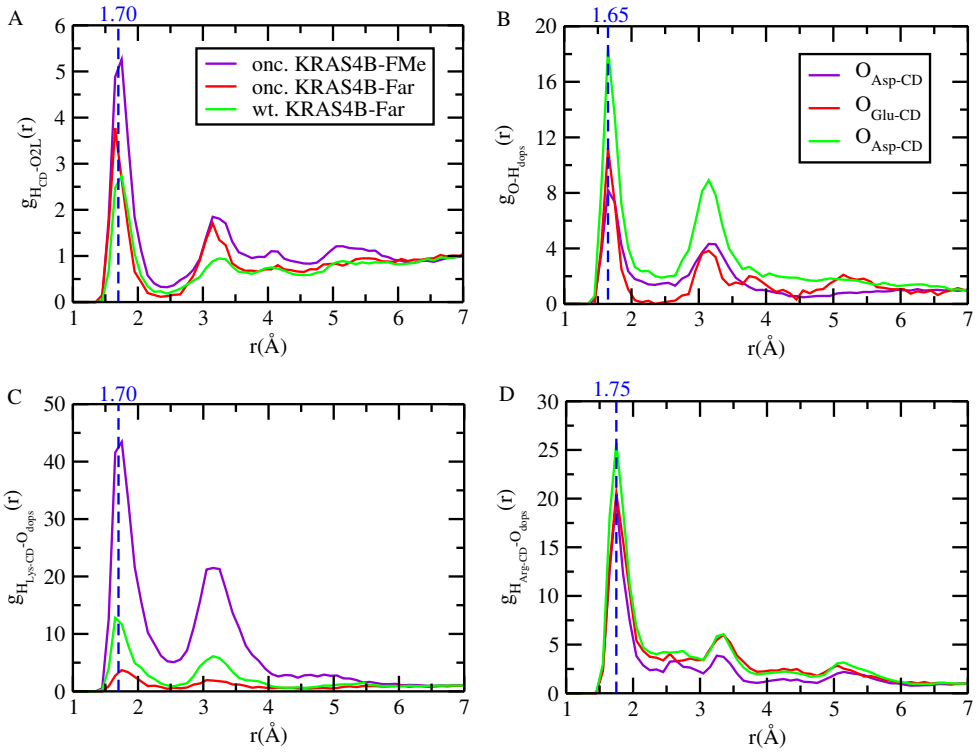


Fig. 5.9 Selected RDFs for selected sites of the CD with head groups of lipids. In all panels of this figure, as indicated in panel "A", we use three colors to represent three systems in order to avoid the duplication and repetition, same as Fig. 5.10. Thereinto, H_{CD} from panel "A" refers to the averaged value of H_{Arg-CD} and H_{Lys-CD}.

guarantees that the wild-type KRas-4B-Far can anchor into anionic membrane bilayers easily during its GTP-bound state. For the oncogenic KRas-4B-Far, long-lasting salt-bridge has been observed between positively charged O_{far} and negatively charged H_{dops} and typical HBs are found between oxygen and hydrogen atoms from FAR and PHOS, indicating that KRas-4B-Far possess the stronger ability to pull the negatively charged POPS lipids together, leading to the clustering of the POPS lipids around itself than other two proteins[349]. In order to monitor the interactions between FAR and PHOS for the oncogenic KRas-4B-Far system, we display the time evolution of selected atom-atom distances in Fig. 5.12 regarding the strong salt-bridge observed in Fig. 5.11, panel "D".

From Fig. 5.12 we can verify that long-lasting and strong salt-bridges do exist between active sites between FAR and PHOS for the oncogenic KRas-4B-Far protein. $O1P_{phos}$ and $O2P_{phos}$ could shift to form salt-bridge with H_{far}, but only one of them could interact with H_{far} hauling the typical HB distance, with one of them favored by H_{far} because of the conformational restrictions in the structure even though both of them share one negative charge. After analyzing production trajectories carefully, we have found that O_{far}, H_{phos}, H_{far}, $O1P_{phos}/O2P_{phos}$ and other relevant atoms could form a very stable 7-membered ring (see Fig. 5.13) which endows the oncogenic

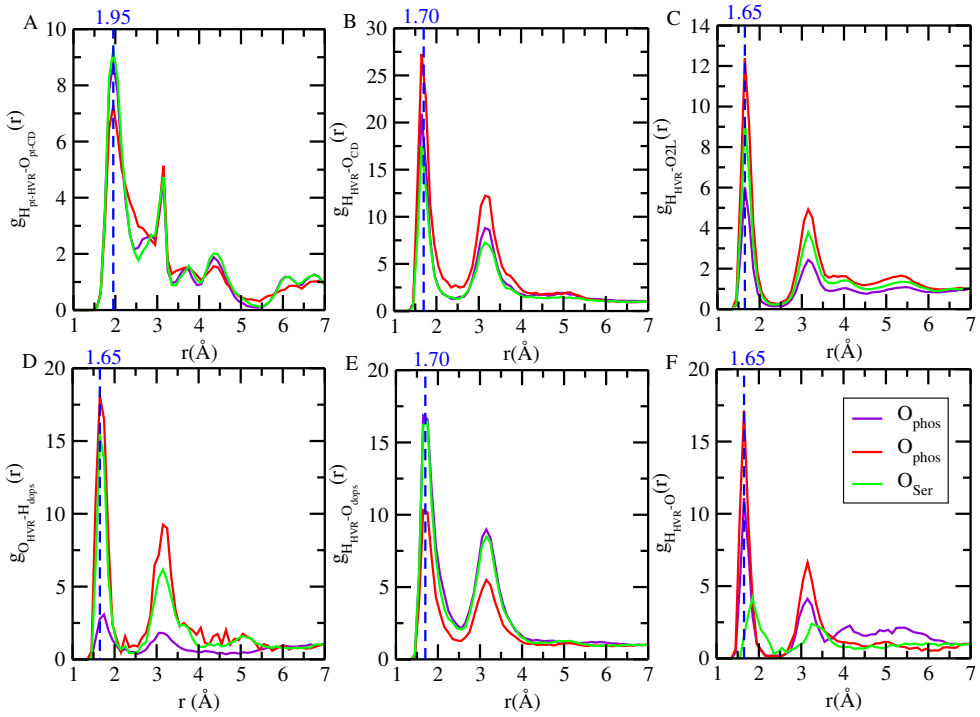


Fig. 5.10 Selected RDFs for selected sites of the HVR with atoms of lipids and the CD. Here O_{pt-CD} and H_{pt-HVR} represent oxygen and hydrogen atoms of the peptide bonds from the CD and the HVR. And O_{CD} stands for oxygen atoms of O_{Glu-CD} and O_{Asp-CD} from the CD.

KRas-4B-Far a specific structure in its tail which may play a role in its other activities, such as, anchoring membranes or reacting with other proteins *in vivo*. This stable 7-membered ring in the structure of the mutated KRas-4B-Far also involves a 4-amino-acid-long sequence in the HVR (S_p-K-T-K) and the farnesylated Cysteine at the site-185 (Cys-185), thereinto, S_p of the 4-amino-acid-long sequence stands for the phosphorylated Ser-181. Our observation agrees with the hypothesis conjectured by S.Dharmaiah[145] that the 5-amino-acid-long sequence motif in the HVR of wild-type KRAS4B-FMe (K-S-K-T-K) may enable PDEδ to bind the prenylated KRas-4B-FMe (see section 1.3). So is this special structure of the 7-membered ring responsible for the different interactions of wild-type KRas-Far and KRas-FMe with PDEδ reported in Ref. [145]? An investigation of the interactions within the 5-amino-acid-long sequence motif (K-S-K-T-K) of the wild-type KRAS4B-FMe should be conducted, which is beyond the scope of this Thesis.

5.2.1.3 Preferential localization of KRas-4B on membranes

Along with the simulation time, tracking the movement of different domains of KRas-4B and GTP along with the membrane normal could give us the direct information of how does the KRas-4B proteins and GTP molecule regulate. We report in Fig. 5.14

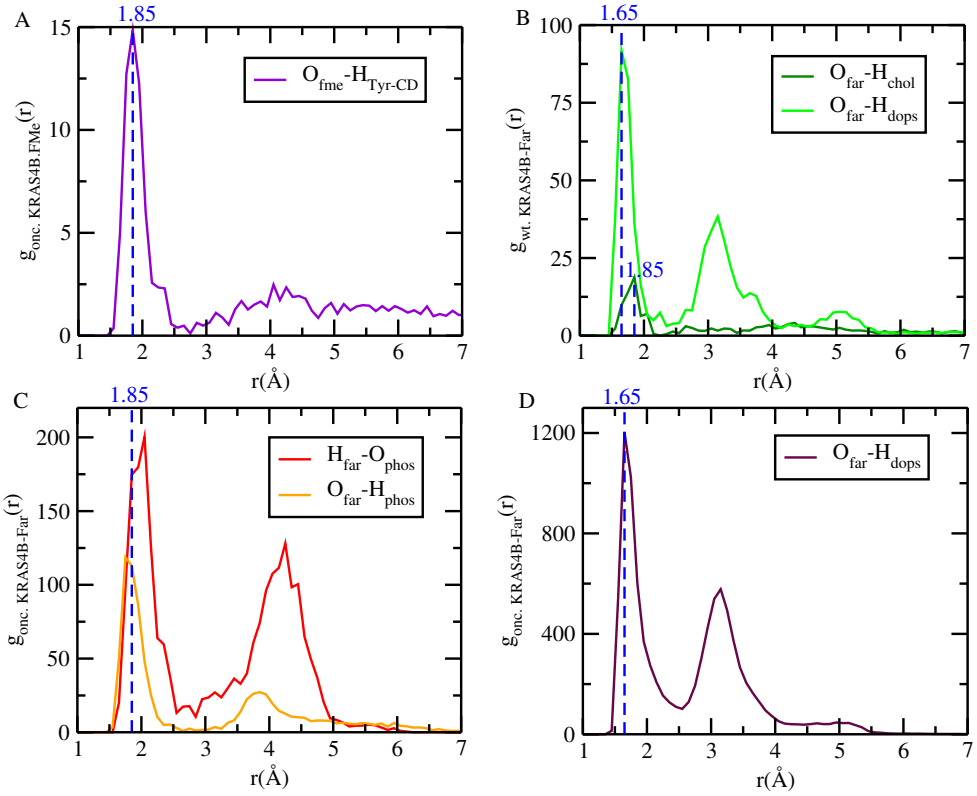


Fig. 5.11 Selected RDFs related to selected sites of FAR.

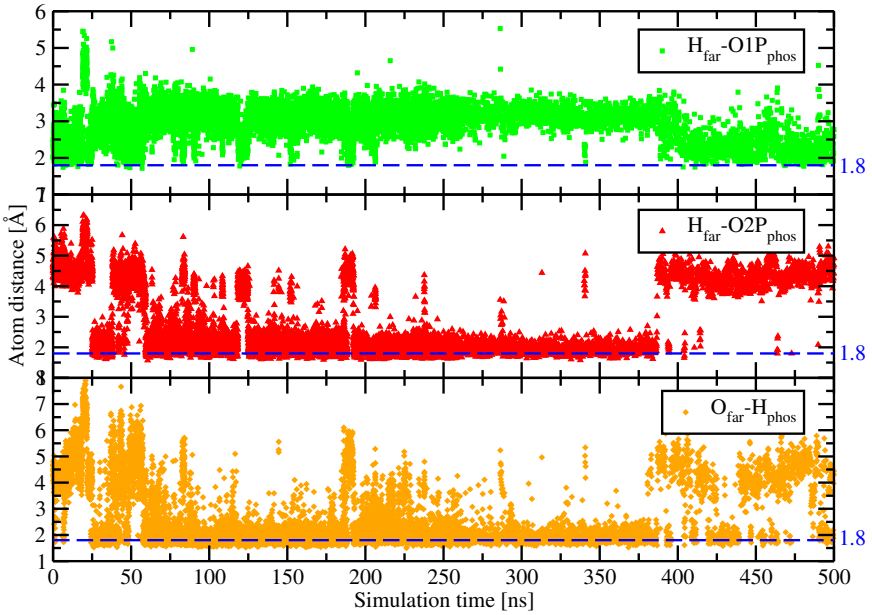


Fig. 5.12 Distance distribution of selected sites of FAR and PHOS as a function of simulation time. Blue dashed lines indicate the typical HB distance (1.8 Å). O1P and O2P represent the two oxygen atoms belong to the phosphorus group of PHOS (see Fig. 5.6). And the corresponding snapshot is shown in Fig. 5.13.

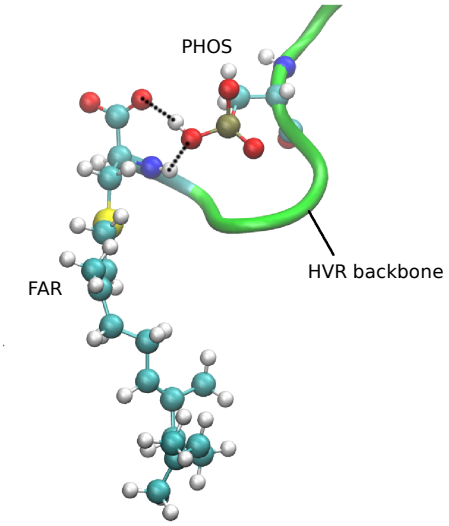


Fig. 5.13 Snapshot of long-lasting salt-bridges indicated as black lines of dashes between FAR and PHOS. The atoms forming the FAR and PHOS moiety are: carbon (cyan), oxygen (red), hydrogen (white), nitrogen (blue) and phosphorus (brown).

the Z-axis positions of the centers of FAR and GTP from the center of lipids (i.e. z = 0) using the second half of 1000 ns simulation for all cases.

After a thorough inspection of Fig. 5.14 and Fig. 5.15 a list of remarkable observations are as follows:

1. In all cases the inner leaflet with KRas-4B binding to the membrane tends to be slightly thicker and tails hold 0.2 nm more stretched-out than the other leaflet.

2. When the wild-type KRas-4B-Far is in its configuration 1 (see Fig. 5.15), the HVR and FAR tend to wander around 3.90 nm in parallel along with the XY plane at the interface of the membrane, whereas in configuration 2 the HVR locates roughly 0.5 nm closer to the membrane center while FAR anchoring to the membrane, located around z = 1.73 nm. GTP buries itself in the interface located around z = 2.39 nm whereas the CD is mainly placed above the membrane XY plane at z = 4.48 nm.

3. For the oncogenic KRas-4B-Far, FAR is revealed to be anchoring constantly into the anionic membrane as a function of the simulation time locating around 1.37 nm with the HVR locating around 3.56 nm away from the membrane center along with the membrane normal direction. GTP keeps binding with the CD and stays around 5.16 nm from the membrane center while the CD diffuses roughly 4.7 nm away.

4. For the oncogenic KRas-4B-FMe, the HVR interacts with lipids and the CD moiety, FAR flipped away from the surface of the membrane with an averaged

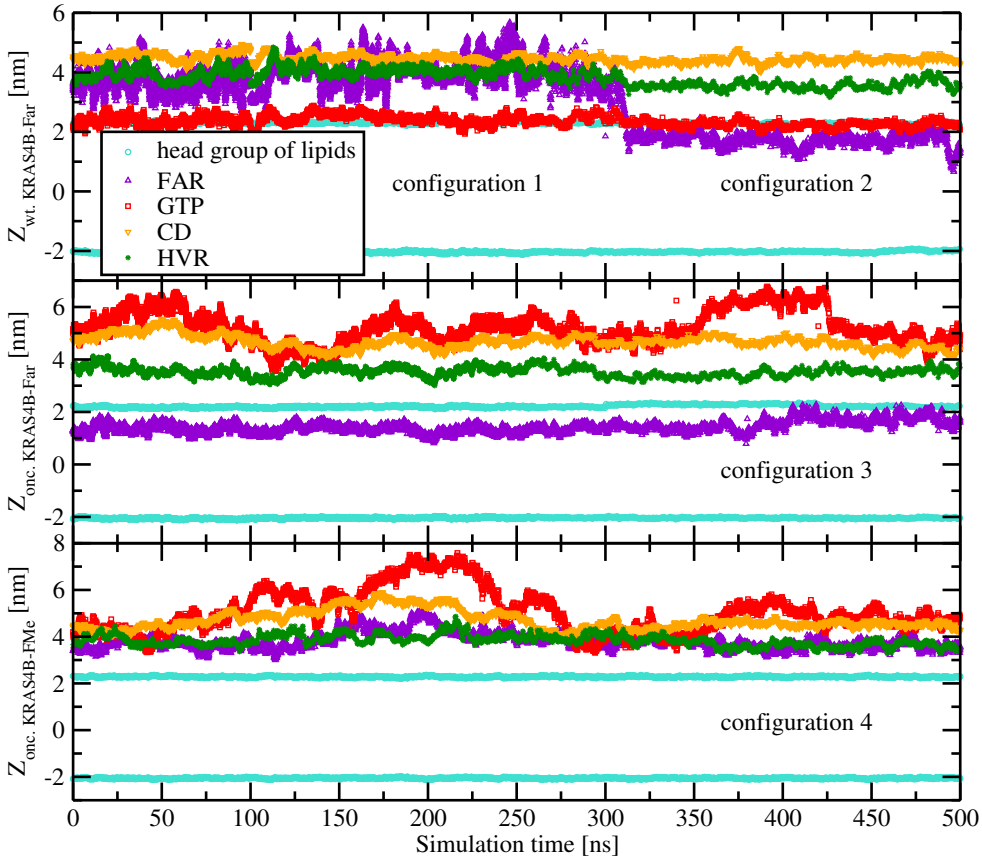


Fig. 5.14 Localisation of the KRas-4B domains and GTP with respect to the center of the membrane along with the membrane normal as a function of simulation time. Geometric centers of the CD, the HVR, FAR, GTP, and phosphorus atoms of DOPC lipids from both leaflets are indicated as triangle down in orange, star in green, triangle up in violet and circle in turquoise, respectively. The four favored configurations described here are reported in Fig. 5.15.

location of 3.82 nm. The FAR doesn't insert into the membrane bilayer. Instead, it is exposed and available for antibody recognition. No spontaneous FAR insertion was observed during the simulations, suggesting that the phosphorylation at Ser-181 prohibits FAR insertion into the lipid bilayer.

5. For the demethylated KRas-4B proteins FAR is able to anchor to the cholesterol-enriched (30%) membrane bilayers without much difficulty. The HVR is reported to be 0.86 nm closer to the interface than the CD with FAR binding the CD moiety (also shown in configuration 4), in which the HVR is sandwiched between the effector binding site of the CD and the membrane. Comparing to the oncogenic KRas-4B-Far, the PHOS group strengthens the auto-inhibition of KRas-4B-FMe by the HVR, the same as the wild-type GDP-bound KRas-4B suggested by *in vitro* and in silico observations[365].

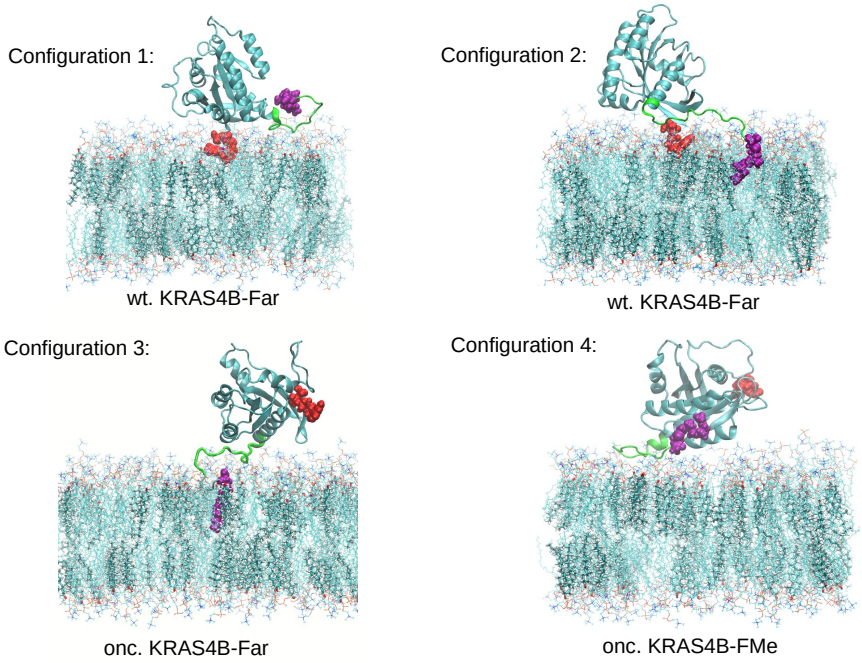


Fig. 5.15 Four preferential configurations of three KRas-4B-membrane systems.

6. According to Ref.[143], phosphorylation of Ser-181 prohibits spontaneous FAR membrane insertion for the methylated KRas-4B-FMe which is consistent with our results of the oncogenic KRas-4B-FMe system. However, the prohibition effect of phosphorylation has been diminished for the demethylated KRas-4B-Far proteins.

7. GTP of the wild-type KRas-4B system locates always at the interfacial region while GTP chooses to bind the CD region of the oncogenic KRas-4B proteins. From $g(r)$s reported in section 5.2.1.2, it is revealed that in the stable state we have observed, GTP serves as a bridge between the CD and lipids through pure HBs and salt-bridges between H_{gtp} and negatively charged oxygen in the CD's side chain and head groups of lipids.

8. Comparing to the wild-type KRas-4B-Far, the locations of the HVR from the oncogenic KRas-4B proteins studied in this Thesis are 0.22 nm higher along with the membrane normal, weakening the HVR interactions with the related the CD, increasing the fluctuations of the CD, and thus resulting in the population of exposing binding site for each oncogenic case.

We have obtained the localization of different moieties of KRas-4B by computing the distance between their geometric centers with the center of the membrane in our models and reported the results in Table 5.3.

Table 5.3 Averaged values of z-locations of FAR, HVR, CD, GTP during the last 500 ns simulations, among them, for the wt. KRas-4B-Far system, values of location of FAR from configuration 1 and 2 have been provided. Estimated errors in parenthesis.

Systems	FAR	GTP	HVR	the CD
wt. KRas-4B-Far	3.9(0.48)/1.73(0.23)	2.40(0.17)	3.93(0.25)	4.48(0.13)
onc. KRas-4B-Far	1.37(0.18)	5.16(0.60)	3.56(0.19)	4.70(0.28)
onc. KRas-4B-FMe	3.82(0.36)	5.07(0.94)	3.84(0.24)	4.70(0.41)

FAR of the wild-type KRas-4B-Far is able to regulate between configuration 1 and 2 easily and the reason might be the corresponding small free energy barriers between these two stable states. It is clear that for the methylated KRas-4B-FMe, the FAR cannot bind the membrane in the presence of PHOS, showing a blocking effect. However, for the demethylated KRas-4B-Far proteins (wild-type and mutated) PHOS could help FAR inserting more deeply and stably into the membrane, indicating a high free energy barrier for FAR departure from the membrane.

Based on an experimental research[348] of GppNHp (a purine nucleotide) bound the wild-type KRas-4B-FMe, the HVR of the KRas-4B-FMe appears to be $\sim$ 1.5 nm closer than the CD of KRas-4B-FMe to the plane of the L-α-PC:DOPS:cholesterol (3:1:1) monomer lipid when the KRas-4B-FMe anchors into the membrane in trimer units. In our wild-type demethylated model, the HVR has two preferred localization on the membrane: around 4 nm (configuration 1) and 3.5 nm (configuration 2) from the center of the membrane, while the CD locates stably near the interface of the anionic membrane bridged through GTP molecule. the HVR locates $\sim$ 0.5 nm closer than the CD when FAR solvated by water molecules and $\sim$ 1.0 nm closer than the CD when FAR anchoring to the membrane which is smaller than the value from Ref.[348], likely due to interactions between the HVRs and interfaces of membranes could be influenced in the trimer.

However, in the oncogenic KRas-4B-Far system FAR keeps anchoring deeply into the system while the HVR is buried near the interface. the HVR stays close to interface and the CD locates 0.22 nm further away from the membrane center than the CD of wild-type case. the CD acts much more actively which gains more accessibility to expose its binding sites for interacting with its upstream regulators and downstream effectors. The interaction between PHOS and DOPS lipid from the oncogenic KRas-4B-Far system, such as O_{far} and H_{dops} (see Fig. 5.11), is formidable, almost 12-fold stronger than the same interaction from the wild-type KRas-4B-Far system, which leads FAR of the oncogenic KRas-4B-Far anchoring the membrane during all the simulation time.

These observations indicate that the wild-type KRas-4B-Far can easily anchor into and departs from the anionic inside membrane and be released from the anionic membrane without crossing too high free energy barriers to reach the next stable state,

however for the oncogenic KRas-4B-Far during the whole simulation time, we haven't observed any tendency of FAR to leave the membrane. However, the oncogenic KRas-4B-Far might face much higher free energy barriers in order to release its C-terminus tail from the anionic membranes. This permanent anchoring is undoubtedly a key factor to keep the protein in its active stay. It helps GTP-bound KRas-4B always anchor to the membrane, and to stay in its active state.

In addition to the aforementioned GTP/GDP binding, the activation of KRas-4B signaling still requires proper membrane-localization, post-translation, and interaction with effector proteins[350]. When the wild-type KRas-4B-Far being in its GTP-bound state doesn't mean it is always in its active state. When FAR wonders near the interface, the membrane-interacting the HVR diffuses near the CD and helps bury the effector biding sites of the CD to the membrane surface. And our observation on the wild-type system has a good agreement with the results reported in Ref.[365] which indicates that demethylation does not affect the CD and the HVR's behavior for the wild-type KRas-4B. For the oncogenic KRas-4B-Far, it always stays in its active state along with the simulation time, its the CD moiety fluctuates more freely and liberates its the HVR binding with the membrane surface, exposing the effector binding sites (see Fig. 5.15). Our data points out that methylation makes a difference in the influence of phosphorylation at Ser-181 on the structural mechanisms of KRas-4B-FMe and KRas-4B-Far.

It is common knowledge that the all-atom MD simulation usually won't start in the global stable state. Along with the simulation, several local stable states will be visited and might include the global stable state. In order to explore conformational states that selected CVs have explored and to understand the structural specifics of KRas-4B-Far binding to the membranes in its GTP-bound form, we have used the enhanced well-tempered metadynamics to sample the conformational transitions of the active KRas-4B along with the membrane normal.

5.3 Well-tempered metadynamics simulations of the conformational transitions of KRas-4B

5.3.1 Study on the convergence of the simulations

In this Chapter, FEL of anionic membrane-anchored (wild-type and mutated) KRas-4B-Far proteins have been investigated for the first time, using well-tempered metadynamics in order to elucidate the joint effect of PHOS and G12D on the KRas-4B-membrane binding.

The convergence of well-tempered metadynamics simulations depends on the number of transition events between states and sampling of all physically interesting regions of the CV[366]. The challenge of well-tempered metadynamics is well known,

especially when applied to protein-bilayers systems with hundreds of thousands of atoms, in order to sample all relevant conformations and the full range of CVs.

It is usual to monitor the size of the hills of the Gaussian kernels deposited along with the simulation. As the simulation progresses and the bias added grows, the Gaussian height is progressively reduced. We can visualize the decrease of the Gaussian height during the simulation. In the long term, the Gaussian height becomes smaller and smaller while the system diffuses in the entire CVs space. The height of the biased potential decreased accordingly along with the simulation runs, as indicated in Fig. 5.16. In all cases, a quasi-flat profile is already seen after 600 ns, although some high spikes are appearing due to large fluctuations in the values of CVs when covering all configurational space.

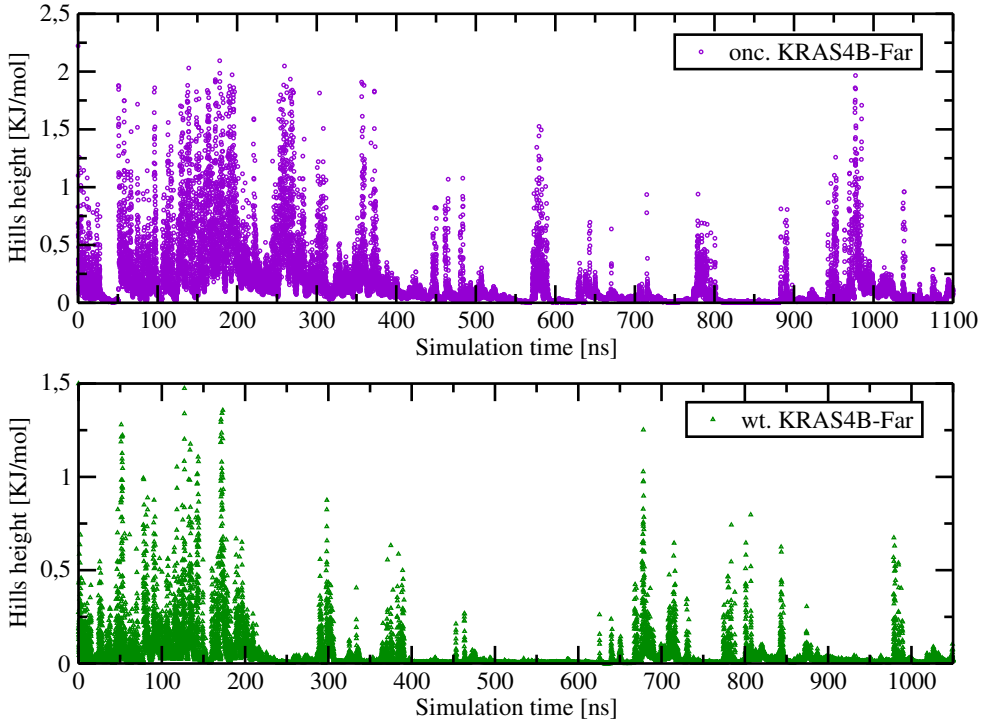


Fig. 5.16 Well-tempered Metadynamics hills height as a function of time in two cases.

The fact that the Gaussian height is decreasing to zero should not be used as a measure of convergence of metadynamics simulation. By inspecting Fig. 5.17 of the well-tempered metadynamics trajectories of the two CVs ("gtp" and "far") along with the simulation time span of our simulations, we can see that the two systems were both initialized in one of their meta-stable states. After ~ 0.1 ns, the systems were pushed by the metadynamics bias potential to visit another local minimum. As simulations continue, the bias potential fills the underlying free energy landscape, and CVs are able to diffuse in possible phase space along with the final 1050 ns simulation time span.

We can clearly identify that for both systems GTP has diffused efficiently in the full collective variable space but the center of the membrane. No permeation for

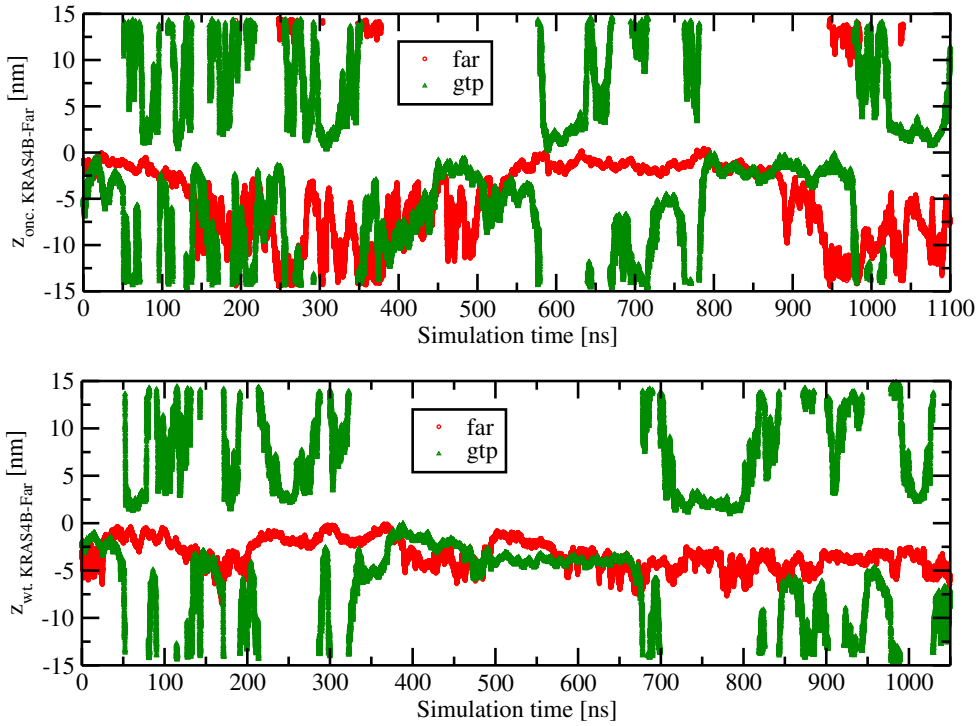


Fig. 5.17 Time evolution of two CVs as a function of time in different states for two systems.

GTP has been observed which indicates the corresponding free energy barrier shall be too high for GTP to cross the membrane bilayer from one leaflet to another and KRas-4B's existence does not have any impact on GTP's permeation. Given the different simulation parameters, for the oncogenic KRas-4B-Far system, FAR visited all the possible space in one leaflet. However, for wild-type KRas-4B-Far system, FAR only reached out in the CV space of around 8.5 nm away from the membrane center along with the membrane normal direction, which is enough for our aim of exploring FAR departure from the membrane and describing its FELs along with its CV space because of the bias employed in the well-tempered metadynamics runs.

Besides reporting phase space that CVs visited during the simulation, convergence can be evaluated in different ways including by monitoring the stability of free energy barriers between different states or by plotting integrated free energy profiles along with the simulation time. Following the previous work in section 4.2.2.2, we checked convergence by two ways: (1) reporting the time cumulative average of 1D free energy profiles along with the simulation time, and (2) calculating the free-energy difference between the two local minima in the one-dimensional free energy along with each CV as a function of simulation time.

At convergence, the reconstructed free energy profiles should be similar, apart from a constant offset. From the results of Fig. 5.18, we can see that the integrated 1D free energy profiles are very similar after a simulation time of 800 ns for the CV of

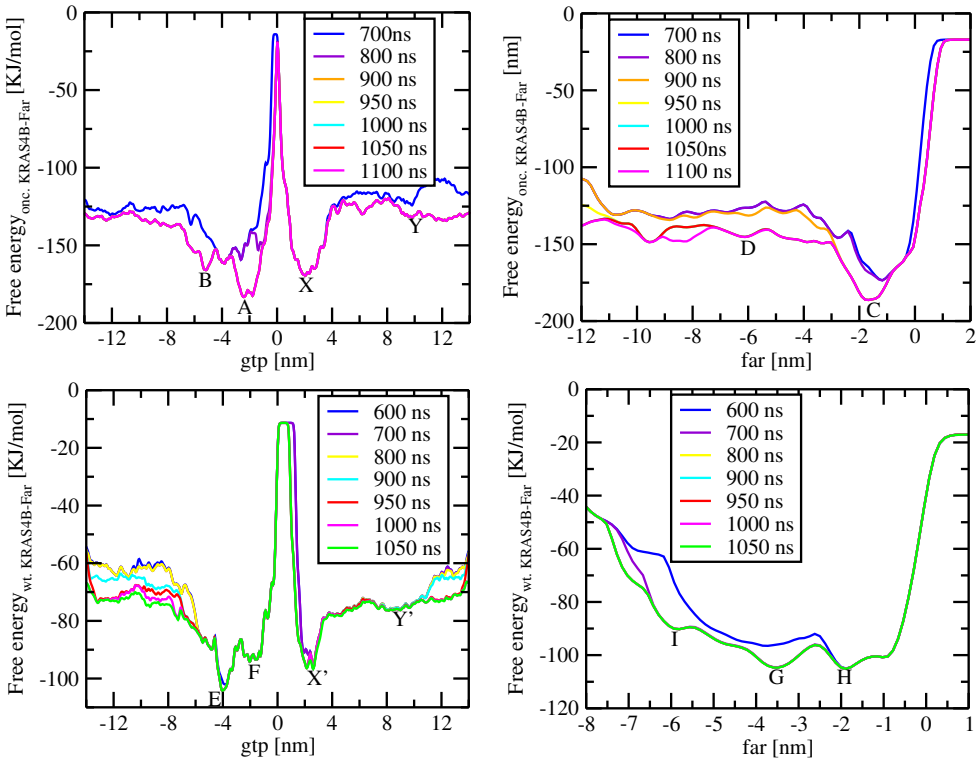


Fig. 5.18 Time cumulative free energy profiles. Lower panels correspond to the wild-type system and upper ones to the oncogenic case. Labels such as 'A, B, ...X' correspond to the relevant basins along with the CV within the free energy profile and will be described below in Fig. 5.19. Here we would like to point out that basins with the same labels in section 5.3.1 are different from section 5.3.2.

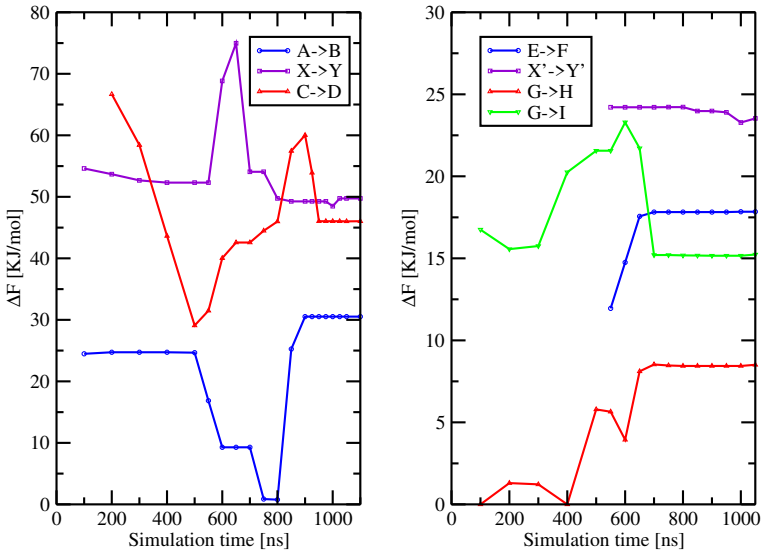


Fig. 5.19 Convergence of well-tempered metadynamics based on the free energy barriers between two chosen basins. The selected basins ('A',...,'Y') have been defined in Fig. 5.18.

"gtp" and 950 ns for the CV of "far" in the oncogenic system. For the wild-type system, convergence was realized much sooner: since 600 ns for "gtp" and 700 ns for "far".

To assess the convergence of the simulation more quantitatively, we have calculated the free energy barriers (ΔF) between two selected free energy basins representing two meta-stable configurations for KRas-4B-Far as a function of simulation time, whose results are reported in Fig. 5.19. As we can see, ΔF between the two chosen free energy basins tends to be similar after long cumulative time spans which lead us to fully converged free energies for the two CVs considered in this Chapter. The estimated free energy profiles and corresponding ΔFs do not change significantly in the final part of our simulations which also suggests that the simulations were fully converged.

5.3.2 2D free energy landscapes

We are interested in the free energy barriers of FAR and GTP binding to the anionic membrane: this information is stored in the free energy landscape. FELs are dynamical so that conformational states can change in response to intra-/extra-molecular events[154, 367], such as the localizations of the FAR and GTP. Two sets of 2D FELs of the oncogenic and wild-type KRas-4B-Far binding to anionic membrane composed of DOPC:DOPS:cholesterol (28:7:15) based on two selected CVs defined above are reported in Figs. 5.20 and 5.21, respectively. Each meta-stable state can be indexed by a pair of CVs. Several regions with clear minima are present in the FELs and clearly, no GTP permeation crossing membrane has been observed in both cases. Minima of the two FELs both locate when GTP is bound to the KRas-4B-Far protein in the inner leaflet. When GTP moves to the outer leaflet, minima correspond to GTP located at the interface of the anionic membrane. Since the range of CV space represented in Figs. 5.20 and 5.21 is rather wide, we will focus especially in the characterization of free energy barriers between the particular states of localizations of FAR (represented by "far" in FELs) and GTP (represented by "gtp" in FELs) along with the membrane normal binding to the membrane and fully solvated by water molecules. Let us note that in both well-tempered metadynamics simulations HVR has not been considered as a CV due to the high expense of computational resources, however, its effect of binding to the membrane interfaces could be shown through the free energy profiles of FAR which we will discuss below.

Since the present 2D FELs include a wide variety of (meta-)stable states related to conformational and structural changes and the free energy surface as a function of CVs is a natural outcome of well-tempered metadynamics method, methods that allow researchers to trace the minimum free energy path (MFEP) with high accuracy and that could be successfully employed to describe complex bio-molecular processes are highly needed in this field. MFEP can be determined by iteratively refine a pathway connecting stable states that converges to the minimum free energy trajectories between them. MFEP could also be obtained through the Path Collective Variables (PCVs)

description coupled with metadynamics[368]. Alternatively to biased-MD simulation methods, certain spontaneous binding events can also be achieved through extensive standard MD simulations, for example, running for μs to ms simulation time[369, 370]. From MFEP we can extract information associated with the most probable trajectories in the CV space followed by the system involving between (meta-)stable states and it also allows us the determination of local and global transition states, as it was described in Chapter 4.

In a different fashion to the method, we adopted in section 4.2.2.2, here we choose to trace the minimum free energy paths along with the FEL using the R-package metadynminer[239]. As we introduced the R-package metadynminer in section 2.4, it reads HILLS files from PLUMED, calculates free energy surface by fast Bias Sum algorithm, finds minima and analyses transition paths by Nudged Elastic Band method. Several MFEPs have been depicted in Figs. 5.20 and 5.21 where the global minimum is set equal to zero. The coordinates of minimum free energy paths are included in Table 5.5.

Our results suggest that there are many local and global stable states along with the 2D FEL with two CVs indexed for each system. For the oncogenic KRas-4B-Far protein, the most stable state and five more meta-stable states have been revealed and shown in Table 5.4.

Table 5.4 Minima on the FEL of the oncogenic KRas-4B-Far (coordinates given in nm, free energy in kJ/mol). Here each letter refers to one meta-stable state, with A standing for the global state when FAR is anchoring deeply into the membrane and GTP locates near the interface. Here locations of minima of FEL of Fig. 5.20 were obtained by using metadynminer package.

States	CV1(gtp)	CV2(far)	Free energy
A	-2.28	-1.37	0.00
B	2.28	-1.15	12.85
C	-5.08	-1.72	15.00
D	-1.81	-4.00	34.28
E	-1.93	-9.59	36.56
F	-7.77	-1.49	41.59

As we pointed out earlier, all-atom MD simulations usually won't start at the global minimum so that along with the simulation several local stable states will be visited, normally including the global minimum. Comparing with Fig. 5.18, we can observe that in the time span of 1000 ns of MD simulations, the oncogenic GTP-bound KRas-4B-Far kept staying at its meta-stable state C, not in its global state A, mainly due to relative high free energy barrier between the states C and A. One should be aware that locations reported in Fig. 5.18 and Table 5.3 are the geometric centers of FAR and GTP whose value could be slightly different from the results of the mass center of GTP and FAR from Table. 5.4. FAR of the oncogenic KRas-4B-Far tends to

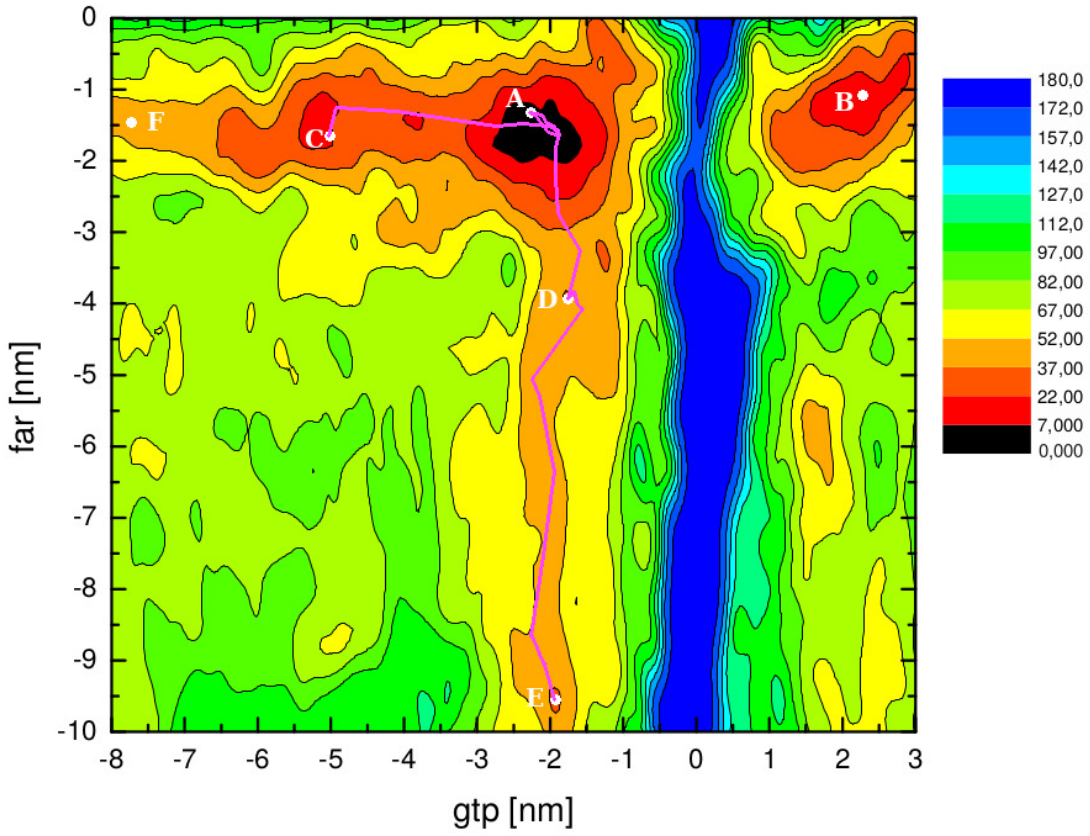


Fig. 5.20 2D free energy landscapes F(gtp, far) (kJ/mol) for the oncogenic KRas-4B-Far system. Several stable configurations (A, B, C, D, E, and F) are indicated. All MFEPs between two selected basins are depicted in purple and the coordinates of the transition path have been reported in Table 5.5.

anchor 0.2 nm deeper (state B) into the "internal" regions of the membrane bilayer when GTP diffuses in the interface of the other leaflet comparing with the most favorable state A when GTP and FAR bind to the inner leaflet of the membrane.

Taken all these remarks into consideration, we propose the stable or meta-stable configurations that the oncogenic KRas-4B-Far may adopt when binding to anionic membrane: in the state A, FAR anchors into the "internal" regions of the membrane when GTP binds the interface and binds the CD moiety, as well; in the state C, FAR anchors into the "internal" regions of the membrane when GTP only binds the CD moiety; in the state D, FAR is solvated by water molecules and GTP only binds the membrane interface; in the state E, FAR reaches out further away from the lipids; and in the state F, GTP is only solvated by water while FAR keeps anchoring to the membrane.

From 2D FEL of the mutated KRas-4B-Far system, we can estimate the values of the free energy barriers between selected (meta-)stable configurations reported in this section:

1. While FAR anchoring to the membrane, GTP dissociates from the most stable state A then binds to the CD moiety (state C), requiring crossing a free energy barrier of 27.04 kJ/mol, and from a meta-stable state C to the global state A needs crossing a barrier of 12.04 kJ/mol which is much lower.

2. The difference between the states D and E is most likely due to the HVR's location: HVR binding the head groups of lipids (state D) and HVR solvated in the aqueous region (state E). The estimated free energy barrier ΔF from the state D to the state E is of 17.20 kJ/mol required for the HVR departure from our model membrane bilayer.

3. GTP diffuses from binding to the CD of the mutated KRas-4B-Far (state C) to being solvated in the water region (state F) with a high barrier ΔF of 27.51 kJ/mol along with the MFEP (not reported in this section) which explains GTP's high affinity of the CD of the mutated KRas-4B-Far.

4. We estimated a free energy barrier of 42.44 kJ/mol from the state A to the state D. For the oncogenic KRas-4B-Far, the FAR's affinity of the anionic membrane is extremely high. Thereupon, the release of FAR from the anionic membrane becomes unlikely because of the high ΔF, accordingly, it is very easy for FAR anchoring back into the bilayer (configuration shifts from the state D to the state A) with a barrier of 8.16 kJ/mol.

5. When GTP moves in the outer membrane leaflet, only one local stable state B is identified in the related 2D FEL, and we can obtain from the 2D FEL that it is very difficult for FAR to depart from the anionic membrane, indicating GTP's huge impact on the mutated system.

After locating local minima from the 2D FEL, by changing parameters while calculating transition paths, we are able to refine the path and eventually obtain the MFEPs. Detailed coordinates of MFEPs are reported in Table 5.5.

Accordingly, the FEL of the wild-type KRas-4B-Far system has been calculated and analyzed from data reported in Fig. 5.21.

Measuring the value of the free energy directly from the 2D FEL of the wild-type KRas-4B-Far system, two most stable states and four more meta-stable states have been identified and shown in Table 5.6, important MEFPs are shown in Fig. 5.21.

We could also estimate the values of free energy barriers between the significant states of the wild-type KRas-4B-Far system, and the main features are summarized as follows:

Table 5.5 Coordinates of segments forming selected paths for the oncogenic and wild-type KRas-4B-Far systems. Along with each path 8 locations of fespot are displayed.

Minimum free energy paths							
onc. KRas-4B-Far				wt. KRas-4B-Far			
Stable states	path x	path y	fespot	Stable states	path x	path y	fespot
C	-5.08	-1.72	15.00	C	-3.35	-0.97	5.14
	-4.82	-1.24	19.50		-3.36	-1.07	7.00
	-4.03	-1.34	27.04		-3.80	-1.10	8.61
	-2.68	-1.67	11.55		-3.79	-1.18	8.02
	-2.29	-1.55	1.39		-3.77	-1.25	7.68
	-2.00	-1.59	1.79		-3.76	-1.41	6.69
	-1.89	-1.64	0.36		-3.76	-1.63	4.06
	-2.01	-1.61	1.78		-3.78	-1.77	2.48
	-2.19	-1.49	0.65		-3.75	-1.88	1.92
A	-2.28	-1.37	0.00	B	-3.94	-1.93	1.92
	-2.05	-1.44	2.48		-3.80	-1.93	2.02
	-2.07	-1.69	2.25		-3.77	-1.99	2.53
	-1.92	-1.61	0.36		-3.95	-2.27	9.43
	-1.98	-1.78	4.48		-3.64	-2.61	10.76
	-1.98	-1.97	8.38		-4.05	-3.00	8.29
	-1.99	-2.34	24.97		-3.94	-3.48	0.04
	-1.95	-2.86	35.40		-3.96	-3.50	0.04
	-1.63	-3.33	42.44		-3.93	-3.49	0.04
D	-1.81	-4.00	34.28	A	-4.06	-3.49	0.00
	-1.75	-3.93	34.28		-3.71	-3.75	6.27
	-1.62	-4.04	41.80		-3.92	-3.98	7.10
	-2.28	-5.03	47.22		-4.07	-4.26	10.73
	-2.11	-5.36	44.82		-3.92	-4.78	13.13
	-1.98	-6.43	40.92		-3.88	-5.04	15.57
	-2.06	-7.54	51.48		-3.73	-5.30	16.93
	-2.38	-8.51	45.08		-3.75	-5.51	20.10
	-2.05	-9.13	42.67		-3.76	-5.67	20.04
E	-1.93	-9.59	36.56	F	-3.82	-5.91	15.25

1. Differently to the case of the oncogenic KRas-4B-Far, two meta-stable states D and E are established when GTP and FAR move in different leaflets, with a barrier ΔF of 11.02 kJ/mol.

2. When GTP and FAR are bound to the inner leaflet, a barrier of 10.76 kJ/mol is likely to be crossed by FAR shifting from the global stable state A to another stable state B while GTP keeps bound to the CD, and the barrier for FAR's release from the membrane is about 8.84 kJ/mol. Both values are similar to barriers between states D and E, suggesting GTP doesn't have much influence on FAR's movement for the wild-type case.

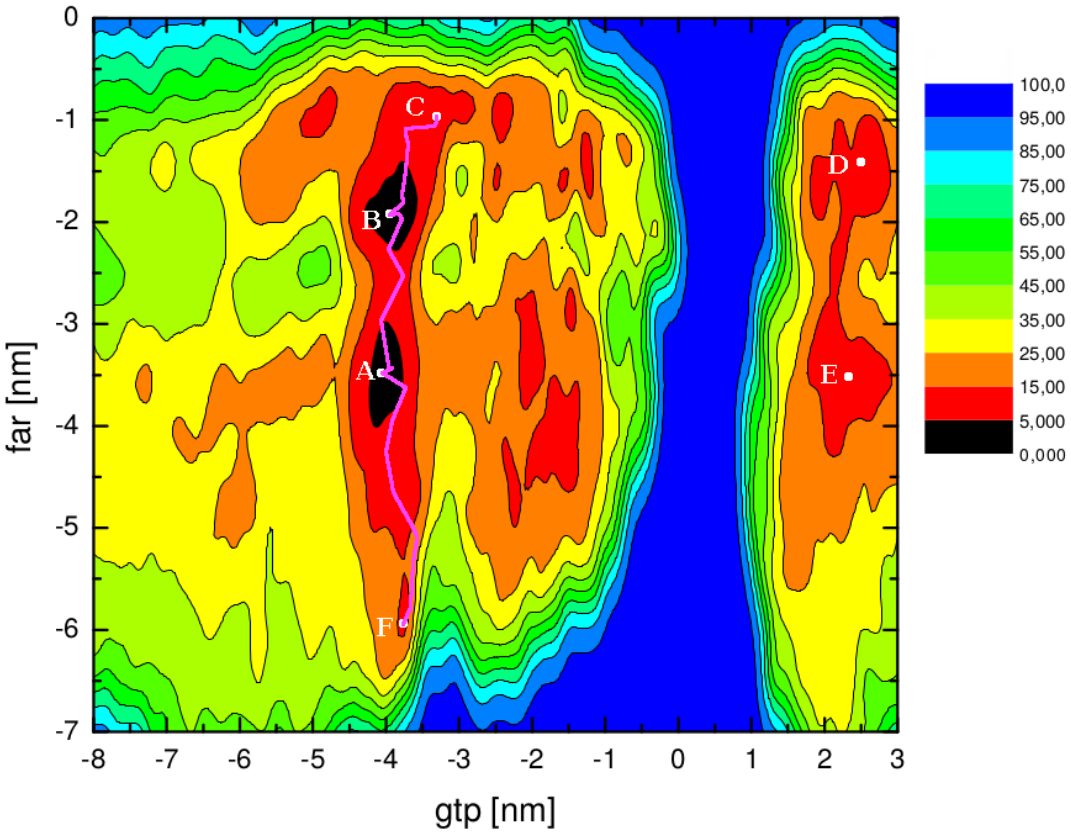


Fig. 5.21 2D free energy landscapes F(gtp, far) (kJ/mol) of the wild-type KRas-4B system. Seven meta-stable configurations are indicated here. MFEPs between different two basins are depicted in purple and the relevant coordinates have been reported in Table 5.5.

3. Shifting from the state B to the state C can be realized if a free energy barrier of 6.69 kJ/mol is surmounted.

4. When the system shifts its states from A to F, the crossing of a free energy barrier of 20.10 kJ/mol corresponds to the energy needed for the HVR to be released from the anionic membrane when FAR is solvated in the aqueous region.

5. There are several meta-stable states when GTP locates ∼2 nm away from the membrane center, indicating the existence of multiple configurations when GTP is around head groups of lipids, which may play a role in KRas-4B-Far's signal transduction and interactions with other proteins *in vivo*.

6. For the wild-type KRas-4B-Far system, the location of GTP does not play a role in FAR's movement along with the membrane normal directions. FAR has the

Table 5.6 Minima on the FEL of the wild-type KRas-4B-Far. Here each letter refers to one meta-stable state, among which A stands for the global state which is set to zero. Here locations of minima of FEL of Fig. 5.21 were obtained by using metadynminer package.

States	CV1(gtp)	CV2(far)	Free energy
A	-4.06	-3.49	0.00
B	-3.94	-1.93	1.92
C	-3.35	-0.97	5.14
D	2.53	-1.40	8.70
E	2.17	-3.52	8.75
F	-3.82	-5.91	15.25

ability to anchor to and be released from the anionic membrane whether GTP is bound to the interface of the outer leaflet; bound to the CD moiety by crossing low barriers (from observation of well-tempered metadynamics simulation), or when GTP is bound to the interface of the inner leaflet (Fig. 5.14 from all-atom MD simulation).

From the 2D FEL we can locate the two most stable states as the state A (FAR diffuses in the water region while GTP is bound to the CD moiety of the wild-type KRas-4B-Far) and the state B (FAR anchored to the membrane while GTP keeps being bound to the CD). To our surprise, two configurations "1" and "2" reported in Fig. 5.15 do not belong to any (meta-)stable state detected by the well-tempered metadynamics method. Two reasons for this lack of matching might be: (a) as pointed out earlier, MD simulations, especially for big complex biological systems, start more likely in one of its meta-state stable states and only during the extremely long simulation time spans other local stable states can be visited, including its global stable state (states A and B); (b) Well-tempered metadynamics simulation with the same two CVs should be extended for another 1 μs or even much longer which is very expensive to conduct. We expect that in such extended simulations MD and metadynamics simulations would describe similar results of stable states.

5.3.3 One-dimensional free energy profiles

It is also possible to calculate one-dimensional free energies from the two-dimensional well-tempered metadynamics simulation. The calculation of 1D free energy profiles could allow us to directly check how does the free energy change along with one CV which can be directly compared to experimental findings. 1D free energy profile F(s_1) for both systems (oncogenic, wild-type) after integrating out the second CV s_2, as it is defined in Eq. 4.1, have been represented in Fig. 5.22.

A newly adopted scientific method should be consistent with validation. Here we have the opportunity to validate the results of MFEP generated by metadynminer, in

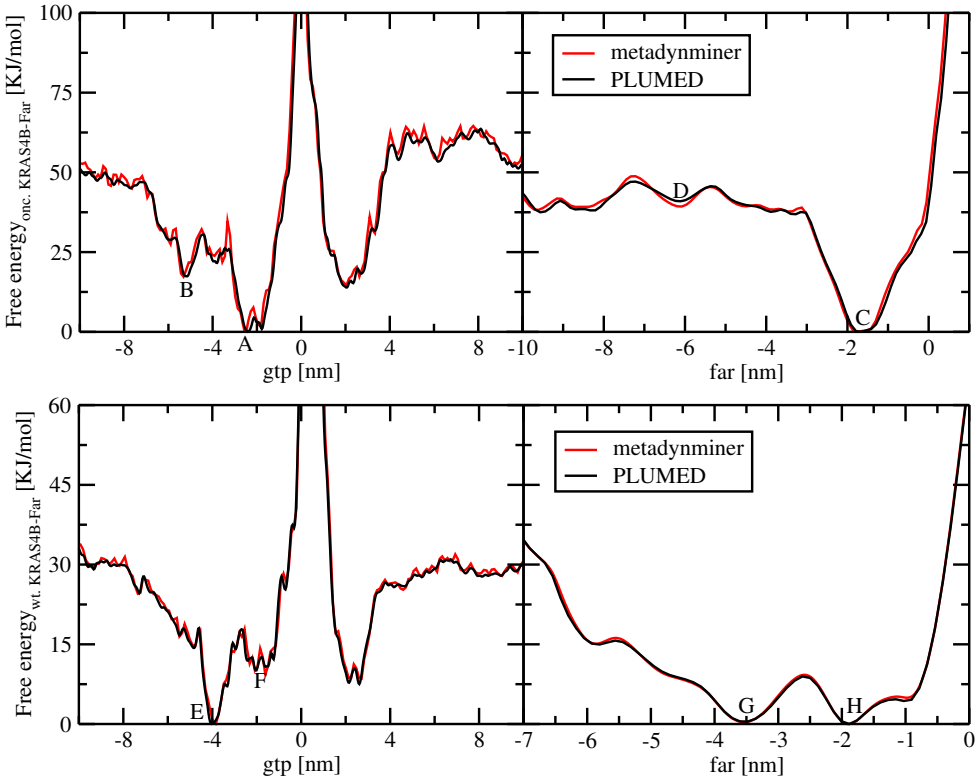


Fig. 5.22 1D integrated free energy profiles of GTP and FAR along with the membrane normal direction. Oncogenic case (top) and wild-type case (bottom). Basins marked with the same letters as in Fig. 5.18. Minima in four panels are set equal to zero. The minimum in each of the four figures has been set to zero.

comparison with 1D-free energy profiles generated using the well known PLUMED package. We also calculated the 1D free energy profile using the R-package metadynminer and results are presented in Fig. 5.22, as well. As we could see for both systems, free energy profiles related to GTP generated by metadynminer contain a degree of noise and interference, but the locations of basins, the corresponding barriers, and values of free energies are quite similar to energy profiles calculated with PLUMED. Hence, the equivalence of 1CV integrated free energy profiles and the related ones generated from MFEP profiles (Table 5.5) establish a good degree of internal consistency.

From plots in Fig. 5.22 we can obtain meaningful free energy barriers related to particular processes related to movements of GTP and FAR along with the Z-axis. Here we choose to calculate barriers precisely calculated using PLUMED which contains less noise. The main findings are highlighted as follows:

1. In the case of the for the oncogenic KRas-4B-Far system, when diffusing in a leaflet together with FAR, GTP needs 30.53 kJ/mol to move from the interface of the bilayer to outside, in order to bind with the CD moiety. However, only a free energy barrier of 7.74 kJ/mol is required for the wild-type case.

2. Similar barriers that need to be surmounted for GTP to dissociate from the CD to be solvated by water molecules are of 33.98 kJ/mol and 35.36 kJ/mol for oncogenic and wild-type system, respectively, indicating mutation at G12D doesn't have a significant influence on GTP's movement.

3. The displacement of GTP from the interface of the outer leaflet of the membrane to the aqueous region requires about 49.74 kJ/mol (onc.) and 23.53 kJ/mol (wt.), and numbers corresponding to the protein-populated leaflet are about 49.49 kJ/mol (onc.) and 30.56 kJ/mol (wt.).

4. In the same fashion, free energy needed for FAR release from the membrane to be solvated in the aqueous region is of 46.01 kJ/mol for the oncogenic case and of 15.16 kJ/mol for wild-type case, respectively.

5. In its GTP-bound state, the wild-type KRas-4B-Far could easily change its configurations between the state H (KRas-4B anchoring to the membrane by FAR aligning with lipid molecules of the membrane) and the state G (interacting with the lipids' interface by its the HVR while FAR diffusing in the water region). In order to decline interactions between FAR and lipids, FAR crosses a barrier of 8.44 kJ/mol to reach the state H from the state G, and a barrier of 8.98 kJ/mol to shift from the state H to G.

6. Here we are presenting two stable states for the wild-type KRas-4B. The wild-type KRas-4B-Far interacts with the membrane firmly through its the HVR and FAR moieties. However, once its FAR dissociates from the membrane it has the ability to anchor back and act as a switch, which ensures its normal function.

7. In the oncogenic case, comparing the free energy profiles, relative to the wild-type protein, PHOS at Ser-181 strongly favors FAR anchoring to the membrane while GTP is located at the interface. This is consistent with our MD simulation results. This provides additional support to our statement that increasing its hydrophobic density through the carboxymethylation at Cys-185 does not always favor its protein recruitment to the PM when PHOS gets involved in the KRas-4B structure.

In 2017, Y.Zhou et al.[366] published their results of metadynamics simulations of K-Ras in plasma membranes on the conformational changes using CV1 (the C-atom RMSD of residue 177-182 from a helical reference structure) and CV2 (one end-to-end distance involving the C-atom of residue 176 and 184) at the 1 μs timescale. Herein, PHOS at Ser-181 of mutated KRas-G12V was found to favor a specific orientation state.

To the best of our best knowledge, (well-tempered) metadynamics simulations to calculate the localization of different moieties of the KRas-4B-Far proteins had not been

investigated until now. The results reported in this Chapter have been the first-hand publication about FELs of the GTP-bound wild-type/oncogenic KRas-4B-Far binding to anionic membrane systems.

5.4 Conclusions

KRas-4B belongs to a family of small GTPase that regulates cell growth, differentiation, and survival which is frequently mutated in cancer cells such as lung, colon, and pancreatic cancers. However, the structural mechanisms at the atomic level of the KRas-4B-membrane association are not fully understood. Many important results have been reported and most efforts have been focused on the methylated KRas-4B-FMe protein after decades of research. However, there is a relatively high abundance of KRas-4B-Far (wild-type and oncogenic) in tumors. In this Chapter observations of three wild-type and mutated KRas-4B proteins obtained by using molecular dynamics simulations coupled with the up-to-date free energy landscapes for two cases have been presented, hoping to inspire and dynamize contributions in this research field.

Firstly, we conducted MD simulations of three systems of the GTP-bound KRas-4B (the oncogenic KRas-4B-FMe, KRas-4B-Far, and the wild-type KRas-4B-Far) binding cell membranes constituted by DOPC (56%), DOPS (14%) and cholesterol (30%) including KCl ionic solution at 310.15 K and at the fixed pressure of 1 atm. Three sets of continuous MD simulations all reached a scale of 1000 ns. Each oncogenic KRas-4B protein contains two mutations: G12D and phosphorylation at its Ser-181 site. Our main interest was focused on the local structures and localizations of different moieties of the KRas-4B proteins. As a general fact, the KRas-4B association with the PM requires the FAR's penetration into the membrane and the HVR's interactions with lipids in order to allow KRas-4B a proper localization in the membrane. After the systematic analysis of meaningful data, we noted relevant differences between different cases. We observed that for the wild-type KRas-4B-Far, demethylated Cys-185 allows FAR spontaneously to insert into and depart from the anionic membrane bilayer without much difficulty along with the simulation; the oncogenic KRas-4B-Far keeps anchoring to the membrane and staying in its active site in order to bind upstream regulators and downstream effectors; and the oncogenic KRas-4B-FMe stays in the state of auto-inhibition by the HVR in which the HVR is sandwiched between the effector binding site of the CD and the membrane, blocking the signal transduction pathways. In order to explain our observations, microscopic properties such as the area per lipid, lipid thickness, radial distribution functions, penetrations of FAR, GTP, etc. to the membrane, and deuterium order parameter have been evaluated. Interactions, such as HBs and long-lived salt-bridges between different active sites of GTP, the CD, the HVR, and lipids, have revealed that to play a central role in the stabilization of the KRas-4B membrane proteins. Among the two mutations in the two oncogenic

KRas-4B protein structures, the effect of G12D mutation on the localization of FAR and GTP of the KRas-4B systems is still unknown. However, PHOS has an important influence on the behavior of the oncogenic KRas-4B proteins. Furthermore, PTMs such as carboxymethylation always play a role in KRas-4B's localization on the PM. By investigating penetrations along with the membrane normal of different moieties of the KRas-4B structures, we have revealed that bigger distances of the CD from the membrane center along with the membrane normal direction of the oncogenic KRas-4B proteins, relative to the wild-type one, help to expose the binding site of the CD to bind upstream regulators and downstream effectors, finally populating their active state for the mutated proteins.

To the best of our knowledge, the free energy profiles of GTP binding to anionic membrane bilayers have been reported for the first time, giving first-hand information about FELs of the GTP-bound wild-type/oncogenic KRas-4B-Far binding to anionic membrane systems. With the help of well-tempered metadynamics simulations, we have calculated 2D free energy landscapes and the 1D free energy profiles along with one CV. We have located global and local stable states of the two systems that have been chosen to elucidate the joint effect of PHOS and G12D on the KRas-4B-membrane binding. Two CVs have been taken into consideration to describe the conformational changes along with the free energy paths. CV1 is the distance "gtp" between the center of mass of GTP and the center of mass of the membrane, and CV2 is the distance "far" between the FAR group of Cys-185 and the center of mass of the membrane along with the direction normal.

Our results indicate that for the GTP-bound oncogenic KRas-4B-Far, a free energy barrier of 42.44 kJ/mol has been estimated which explains that the release of the FAR moiety of the oncogenic KRas-4B-Far from the anionic membrane becomes unlikely because of the high ΔF. Correspondingly, a barrier of 10.76 kJ/mol is likely to be crossed for FAR of the wild-type KRas-4B-Far shifting from anchoring to the PM to being solvated by water while GTP keeps binding to the CD. Differently to one local stable state for the oncogenic KRas-4B-Far, two local stable states are established when GTP and FAR move in the different leaflets for the wild-type case. Likely, when FAR of the wild-type case shifts from $\sim$4 nm to 6 nm away from the membrane center, we have estimated a free energy barrier of 20.10 kJ/mol corresponding to the energies needed for the HVR dissociating from the anionic membrane.

From the calculation of 1D integrated free energy profiles of both systems, important observations have been presented. When diffusing in the inner leaflet with FAR, GTP needs 30.53 kJ/mol to move from the interface of the bilayer to bind with the CD moiety for the oncogenic KRas-4B-Far system, whereas 7.74 kJ/mol free energy is required for the wild-type case. Similar barriers that need to be surmounted for GTP to dissociate from the CD to be solvated by water molecules are of 33.98 kJ/mol and

35.36 kJ/mol for oncogenic and wild-type system, respectively, indicating mutation at G12D doesn't have a significant influence on GTP's movement.

Most importantly, the free energy needed for FAR release from the membrane to be solvated in the aqueous region is of 46.01 kJ/mol for the oncogenic case and of 15.16 kJ/mol for wild-type case, respectively. In its GTP-bound state, the wild-type KRas-4B-Far could easily change its configurations between the state H (KRas-4B anchoring to the membrane by FAR aligning with lipid molecules of the membrane) and the state G (interacting with the lipids' interface by its the HVR while FAR diffusing in the water region). In order to decline interactions between FAR and lipids, FAR crosses a barrier of 8.44 kJ/mol to shift from the state G to the state H. Comparison of free energy profiles, relative to the wild-type protein, PHOS at Ser-181 strongly favors FAR anchoring to the membrane while GTP locates in the interface for the oncogenic case, consistent with the MD simulation results. This provides additional support to our suggestion that adding its hydrophobic density through the carboxymethylation at Cys-185 does not always favor its protein recruitment to the PM when PHOS gets involved in the KRas-4B structure.

Chapter 6

Conclusions and prospect

6.1 General conclusions

In this Thesis we have performed all-atom molecular dynamics and well-tempered metadynamics simulations of different cell membrane model systems. In what follows, we will summarize the main conclusions of this Thesis.

In Chapter 3, a series of molecular dynamics simulations of different lipid bilayer membranes in an aqueous ionic solution of NaCl with an embedded molecule have been performed by MD using the CHARMM36 force field. Among them, cholesterol with different concentrations (30% and 50%) has been considered. In our preliminary study on tryptophan's adsorption to DPPC bilayer membrane at 310.15 K, our data revealed the existence of a strong first coordination shell and a milder second coordination shell for the tryptophan-water association, which translated to deep minima in the corresponding PMFs, with energy barriers of less than 1 kcal/mol. Conversely, the binding of tryptophan to DPPC involves a single coordination shell for the two sites of possible association and energy barriers much lower than for the former TRP-water, i.e. of the order of 1 kcal/mol. Tryptophan shows a clear tendency to stay close to DPPC, during the periods of time about 3 times longer than for the TRP-water association.

Continuing our preliminary study, the temperature was increased from 310.15 K to 323.15 K for three systems at three cholesterol concentrations (0, 30, and 50%) at liquid-crystalline phase conditions. We noted relevant changes in local structure and dynamics of tryptophan only for cholesterol concentrations above 30%. The binding of tryptophan to DPPC involved coordination shells for the different sites of possible association (charged oxygens 'O2' and 'O8' of DPPC *versus* the two tagged hydrogens in TRP, namely 'H1' and 'H2'). Also, the distribution functions of tryptophan-cholesterol revealed stable hydrogen-bonding configurations. These data indicate that tryptophan is able to establish stable interactions with all classes of solvating particles (water, DPPC, and cholesterol) including a sort of *bridge* between DPPC and cholesterol species. The typical hydrogen-bond distances between tryptophan and other species

have been found between 1.7-2.0 Å, in good agreement with experimental data obtained from fluorescence measurements[100]. Tryptophan's self-diffusion coefficients have been found to be of the order of 1-10$\times 10^{-7}$ cm^2/s and they are strongly dependent on the presence of cholesterol in the system. The computed spectral densities, in excellent agreement with experimental infrared and Raman data from Leyton et al.[99], allowed us to reveal the degree of participation of each atomic site of tryptophan to the complete spectrum of the molecule. Considering the importance of studying DPPC lipids, we kept using DPPC bilayer to study the binding free energies of small molecules: serotonin, histidine, tyrosine, melatonin, and tryptophan at the condition of 323.15 K and 1 atm. Our data revealed the existence of a strong first coordination shell and a milder second coordination shell for small molecule-water structure, which correspond to two minima in the corresponding PMFs, with energy barriers for the TRP-water association of the order of 1-2 kcal/mol. Conversely, the binding to DPPC involves a single coordination shell for the two sites of possible association (oxygens 'O2' and 'O8' of DPPC *versus* hydrogens in the small-molecules) and energy barriers between 0.5-3 kcal/mol. Throughout our simulation runs, we did not observe any event of permeation of a small-molecule across the DPPC membrane. Concerning the essentiality of the two amino acids reported in the present work (histidine, tryptophan), we observed that tryptophan is able to enter the interfacial membrane, whereas histidine is not. Interestingly, tyrosine, a non-essential amino acid shows the highest free energy barriers, indicating that it is the most stable molecule for DPPC binding. Serotonin has revealed to be a molecule anchored at the membrane and with a low propensity to be solvated by water, whereas its derivative melatonin is able to equally interact with water and DPPC, showing similarly strong free energy barriers.

The result of this work has been published in the following three works:

- Jordi Martí, Huixia Lu. Molecular Dynamics of Di-palmitoyl-phosphatidyl-choline Biomembranes in Ionic Solution: Adsorption of the Precursor Neurotransmitter Tryptophan. Procedia Computer Science, 2017, 180C, 1242-1250
- Huixia Lu, Jordi Martí. Effects of cholesterol on the binding of the precursor neurotransmitter tryptophan to zwitterionic membranes. The Journal of Chemical Physics, 2018, 149, 164906(1-9)
- Huixia Lu, Jordi Martí. Binding free energies of small-molecules in phospholipid membranes: amino acids, serotonin, and melatonin. Chemical Physics Letters, 2018, 712, 190-195

Due to its medical purposes, we have investigated the dynamics of melatonin binding to the cell membranes in Chapter 4. Results of all-atom molecular dynamics simulations of a bilayer membrane composed of DMPC at three cholesterol concentrations (0%, 30% and 50%) including a single MEL molecule in aqueous NaCl ionic solution at 303 K and the fixed pressure of 1 atm have been reported. Our main interest was

firstly focused on the local structure and angular distributions of MEL, especially when associated with DMPC and cholesterol molecules. Typical hydrogen-bonds through different sites between MEL-DMPC and MEL-cholesterol have been observed. Two relevant MEL structures have been observed from angular distributions: "folded" and "extended" configurations. We defined three different dihedral angles to account for the two preferential angular configurations. Dihedral Ψ defined in Fig. 4.6 has revealed to be a meaningful order parameter (i.e. may act as a reliable reaction coordinate) to describe the dynamics of MEL, with preferential angles of ~ 1.42 rad and ~ 2.96 rad. We suggest that introducing cholesterol into the system could help MEL change from its folded configuration to the extended configuration more easily, using hydrogen-bonds between MEL-DMPC and MEL-cholesterol. The self-diffusion coefficient of MEL was found to be of the order of 10^{-7} cm^2/s and the presence of cholesterol in the system has little influence on it. The spectral densities of MEL computed in this work are in overall good agreement with experimental Raman and infrared data and have revealed the degree of participation of each atomic site of MEL to complete its whole molecular spectrum giving some clues to understand the microscopic origin of molecular vibrations and also giving evidence of the good reliability of the model we adopted in the present work. Nevertheless, in the time span of 200 ns of auxiliary MD simulations, we did not record any spontaneous crossing of MEL through the DMPC membrane. Our findings agree that in order to observe trans-membrane crossings of such small solutes in the time length of a simulation at the atomic level of description, they needed to run trajectories of 10 μs at low temperatures (310 to 330 K) or to increase the temperatures to more than 400 K (for simulation times of 1 μs).

Continuously, we conducted the first quantitative characterization to date of the binding states of MEL at model DMPC-cholesterol phospholipid cell membranes through the calculation of free energy landscapes. With the help of well-tempered metadynamics simulations, we have calculated 2D free energy landscapes and located the binding stable states and several (local and global) transition states of the system. Two CVs have been considered: dihedral angle Ψ and the distance z between the center of mass of MEL and the center of the lipid bilayer (given by $z = 0$). Our results indicate that MEL can be bound to the internal side of the membrane, at distances $z \sim 1 - 2$ nm, and in several stable state configurations where the dihedral $|\Psi| = 1.17$ rad and $|\Psi| = \pi$ rad. Through the calculation of the minimum free energy paths, we have observed that the most probable trajectory of MEL is along with the interface of the membrane, changing its conformation between extended and folded configurations by means of surmounting free energy barriers of about 15-20 kJ/mol. In addition, the methodology employed in the present work allowed us to locate several global and local TS of the system and to estimate the values of their free energy barriers. When cholesterol is added to the system, it helps pull MEL escape from the interface of the membrane to being solvated by water/sodium chloride molecules more

easily by decreasing its corresponding free energy barriers. The effect is more marked as the concentration of cholesterol rises. The energetic cost for MEL to leave the interface of the membrane and to be fully solvated by water/sodium chloride molecules (z-distances around 4 nm), has been estimated to be of 10 $\sim$ 25 kJ/mol. The less common situation was found to be with MEL accessing regions around the center of the membrane, process requiring to cross free energy barriers above 40 kJ/mol. We believe that the findings presented in this chapter could be of practical use in designing new reaction coordinates (for instance coordination numbers of MEL) for multidimensional, more accurate free energy calculations able to explore a wide variety of relevant small solute species of biochemical interest such as amino acids, neurotransmitters, drugs or hormones.

This work has appeared in the following publications:

- Huixia Lu, Jordi Martí. Binding and dynamics of melatonin at the interface of phosphatidylcholine-cholesterol membranes. PloS one, 2019, 14(11), e0224624.
- Huixia Lu, Jordi Martí. Cellular absorption of small molecules: free energy landscapes of melatonin binding at phospholipid membranes. Scientific Reports, 2020, 10, 9235. https://doi.org/10.1038/s41598-020-65753-z

In this Thesis we have also studied the regulation of oncogenic KRas-4B activity when binding to the membrane. Most efforts have been focused on the methylated KRas-4B-FMe protein after decades of research. However, there is a relatively high abundance of KRas-4B-Far (wild-type and oncogenic) in tumors. In Chapter 5 observations of three wild-type and mutated KRas-4B proteins obtained by using molecular dynamics simulations coupled with the up-to-date free energy landscapes for two cases have been presented, hoping to inspire and dynamize contributions in this research field.

Firstly, we conducted MD simulations of three systems of GTP-bound KRas-4B (oncogenic KRas-4B-FMe, KRas-4B-Far, and wild-type KRas-4B-Far) binding cell membranes constituted by DOPC (56%), DOPS (14%) and cholesterol (30%) including KCl ionic solution at 310.15 K and the fixed pressure of 1 atm. Three sets of continuous MD simulations all reached a scale of 1000 ns. Two mutations of G12D and phosphorylation at its Ser-181 site have been considered for mutated KRas-4B proteins. Our main interest was focused on the local structures and localizations of different moieties of KRas-4B proteins. The association of KRas-4B with the PM requires the FAR's penetration into the membrane and the HVR's interactions with lipids to allow KRas-4B a proper localization in the membrane. We observed that for wild-type KRas-4B-Far, demethylated Cys-185 allows FAR spontaneously to insert into and departure from the anionic membrane bilayer without much difficulty along with the simulation; oncogenic KRas-4B-Far keeps anchoring to the membrane and staying in its active site to bind upstream regulators and downstream effectors; and oncogenic KRas-4B-FMe stays in the auto-inhibition by the HVR state in which the HVR is

sandwiched between the effector binding site of the CD and the membrane, blocking the signal transduction pathways. In order to explain our observations, microscopic properties such as the area per lipid, lipid thickness, radial distribution functions, penetrations of FAR, GTP, etc. to the membrane, and deuterium order parameter have been evaluated. Interactions, such as HBs and long-lived salt-bridges between different active sites of GTP, the CD, the HVR, and lipids, have revealed that to play a central role in the stabilization of KRas-4B membrane proteins. Among the two mutations in the two oncogenic KRas-4B protein structures, the effect of G12D mutation on the localization of FAR and GTP of KRas-4B systems is still unknown. However, PHOS has an important influence on the behavior of oncogenic KRas-4B proteins. Furthermore, PTMs such as carboxymethylation always play a role in KRas-4B's localization on the PM. By investigating penetrations along the membrane normal of different moieties of the KRas-4B structures, we have revealed that bigger distances of the CD from the membrane center along the membrane normal direction of oncogenic KRas-4B proteins, relative to the wild-type one, help to expose the binding site of the CD to bind upstream regulators and downstream effectors, finally populating their active state for the mutated proteins.

To the best of our knowledge, the free energy profiles of GTP binding to anionic membrane bilayers have been reported for the first time, giving first-hand information about FELs of GTP-bound wild-type/oncogenic KRas-4B-Far binding to anionic membrane systems. With the help of well-tempered metadynamics simulations, we have calculated 2D free energy landscapes and the 1D free energy profiles along with one CV. We have located global and meta-stable states of the two systems that have been chosen to elucidate the joint effect of PHOS and G12D on the KRas-4B-membrane binding. Two CVs have been taken into consideration to describe the conformational changes along free energy paths. CV1 is the distance "gtp" between the center of mass of GTP and the center of mass of the membrane, and CV2 is the distance "far" between the FAR group of Cys-185 and the center of mass of the membrane along the direction normal. Our results indicate that for GTP-bound oncogenic KRas-4B-Far, a free energy barrier of 42.44 kJ/mol has been estimated which explains that the release of the FAR moiety of oncogenic KRas-4B-Far from the anionic membrane becomes unlikely because of the high ΔF. Correspondingly, a barrier of 10.76 kJ/mol is likely to be crossed for FAR of wild-type KRas-4B-Far shifting from anchoring to the PM to being solvated by water while GTP keeps binding to the CD. Differently to one local stable state for oncogenic KRas-4B-Far, two local stable states are established when GTP and FAR move in the different leaflets for the wild-type case. Likely, when FAR of the wild-type case shifts from $\sim$4 nm to 6 nm away from the membrane center, we have estimated a free energy barrier of 20.10 kJ/mol corresponding to the energies needed for the HVR dissociating from the anionic membrane.

From the calculation of 1D integrated free energy profiles of both systems, important observations have been presented. When diffusing in the inner leaflet with FAR, GTP needs 30.53 kJ/mol to move from the interface of the bilayer to bind with the CD moiety for oncogenic KRas-4B-Far system, whereas 7.74 kJ/mol free energy is required for the wild-type case. Similar barriers that need to be surmounted for GTP to dissociate from the CD to be solvated by water molecules are of 33.98 kJ/mol and 35.36 kJ/mol for oncogenic and wild-type system, respectively, indicating mutation at G12D doesn't influence GTP's movement. Most importantly, the free energy needed for FAR release from the membrane to be solvated in the aqueous region is of 46.01 kJ/mol for the oncogenic case and 15.16 kJ/mol for wild-type case, respectively. In its GTP-bound state, wild-type KRas-4B-Far could easily change its configurations between the state H (KRas-4B anchoring to the membrane by FAR aligning with lipid molecules of the membrane) and the state G (interacting with the interface of the membrane by its the HVR while FAR diffusing in the aqueous region). In order to decline interactions between FAR and lipids, FAR crosses a barrier of 8.44 kJ/mol to shift from state G to state H. Comparison of free energy profiles, relative to the wild-type protein, PHOS at Ser-181 strongly favors FAR anchoring to the membrane while GTP locates in the interface for the oncogenic case, consistent with the MD simulation results. This provides additional support to our suggestion that adding its hydrophobic density through the carboxymethylation at Cys-185 does not always favor its protein recruitment to the PM when PHOS gets involved in the KRas-4B structure.

Results will be reported in the following works:

- Huixia Lu, Jordi Martí. Anchoring mechanisms of oncogenic KRas-4B proteins at cell membranes. In preparation.

- Huixia Lu, Jordi Martí. Influence of cholesterol in the anchoring of mutated KRas-4B proteins in model cell membranes. In preparation.

6.2 Perspectives of future work

In this Thesis we studied several membrane systems, especially KRas-4B-Far and KRas-4B-FMe binding to anionic bilayers, which allowed us to perform an in-depth study of the dynamics and free energy landscapes of these model cell membrane systems. We should keep in mind, by comparing the results of MD and well-tempered metadynamics simulations, that the favored configurations of the studied system observed by MD may not directly correspond to the global minimum of the 2D free energy landscape obtained from the well-tempered metadynamics simulation. Consequently, to increase the accuracy of the results of MD simulations, simulations from different starting configurations for the same protein-membrane system should be taken into consideration.

Concerning metadynamics simulations, when we compared the energetic properties between oncogenic and wild-type KRas-4B-Far systems, the 2D free energy landscapes of GDP-bound wild-type KRas-4B-FMe and GTP-bound oncogenic KRas-4B-FMe were not considered. But surely it is worth to be explored. Hence, a natural way to continue the research carried out in the present Thesis would be to perform 1 μs metadynamics simulations of methylated KRas-4B-FMe and compare their output with the corresponding results from MD simulations.

Despite the failures in developing anti-KRas therapies, targeting KRas remains one of the most promising directions in KRas-related cancer research. The design of small molecules that disrupt KRas-4B-regulators/effectors interactions may provide chemical probes and therapeutic agents for cancer treatment. A small drug named by FGTI-2734 reported in a recent research[371] inhibits the membrane association of mutated KRas protein, but not the wild-type ones, in pancreatic, lung, and colon human cancer cells. After further preclinical and clinical studies, a direct KRAS inhibitor-FGTI-2734 (structure shown in Fig. 6.1) may be on the horizon.

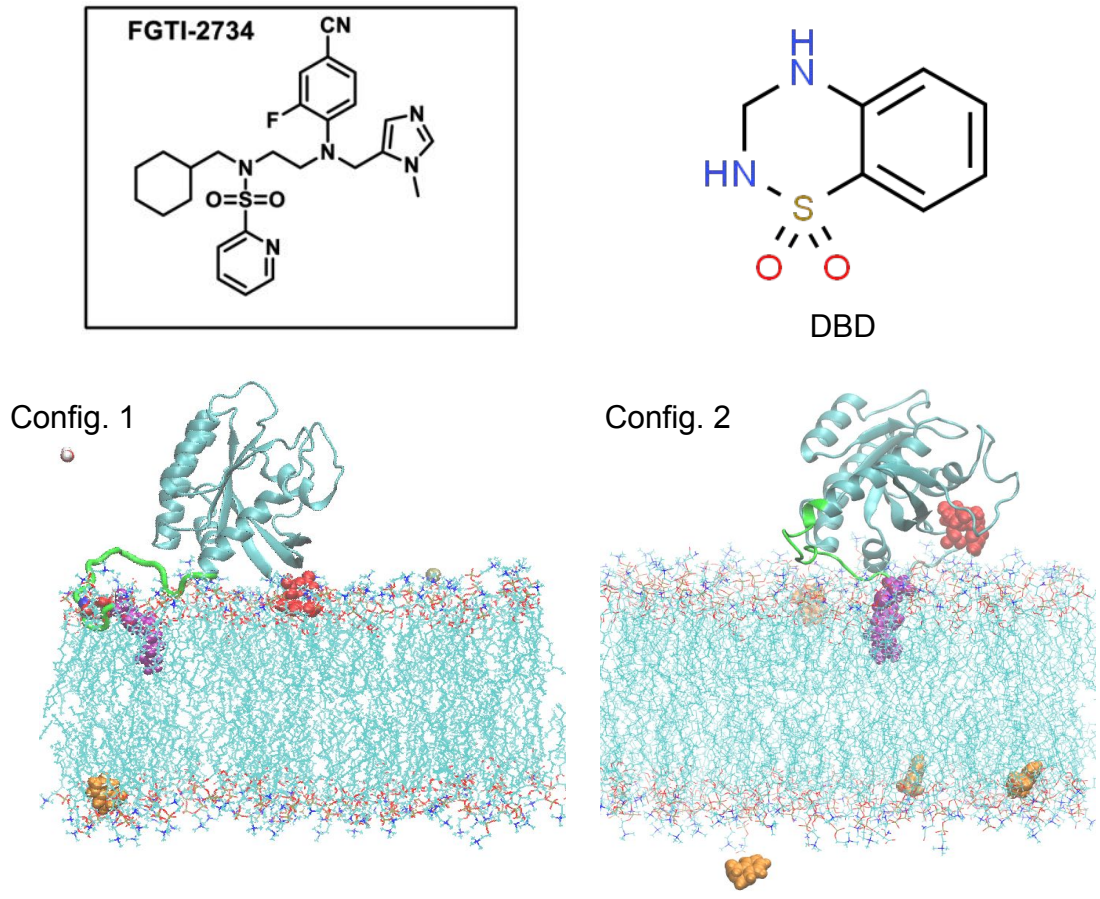


Fig. 6.1 Structures of FGTI-2734 and DBD shown here. In the two configurations (Config. 1 and 2), DBD is depicted in orange, whereas the remaining moieties are depicted with the same color prescriptions as in Fig. 5.2.

The substance 3,4-dihydro-1,2,4-benzothiadiazine-1,1-dioxide (DBD) has been usually applied in the treatment of Alzheimer disease[372, 373], in neurocognitive disorders[372] and in diabetes mellitus[374], among others. Considering DBD shares part of the molecular structure of FGTI-2734, we have started to analyze the effect of DBD molecules on oncogenic KRas-4B's binding to anionic membranes composed of DOPC/DOPS/cholesterol by conducting MD simulations. From Fig. 6.1 we can obtain that the HVR's autoinhibition might be induced with a large dose of DBD (Config. 2: four DBD molecules adopted in the system) comparing to the small dose case (Config. 1: only one DBD molecule adopted in the system). Longer time scales in MD simulations and a more detailed analysis should be conducted for more accurate results.

Appendix A

Related publications

The papers published in international peer-reviewed journals that have been the basis of this Thesis are:

- Martí J., Lu H. Molecular dynamics of di-palmitoyl-phosphatidyl-choline biomembranes in ionic solution: adsorption of the precursor neurotransmitter tryptophan. Procedia computer science. 2017 Jun 9;108(C):1242-50.
- Lu H., Martí J. Effects of cholesterol on the binding of the precursor neurotransmitter tryptophan to zwitterionic membranes. The Journal of chemical physics. 2018 Oct 28;149(16):164906.
- Lu H., Marti J. Binding free energies of small-molecules in phospholipid membranes: Aminoacids, serotonin and melatonin. Chemical Physics Letters. 2018 Nov 16;712:190-5.
- Lu H., Martí J. Binding and dynamics of melatonin at the interface of phosphatidylcholine-cholesterol membranes. PloS one. 2019;14(11).
- Lu H., Martí J. Cellular absorption of small molecules: free energy landscapes of melatonin binding at phospholipid membranes. Scientific Reports, 2020; 10, 9235. https://doi.org/10.1038/s41598-020-65753-z
- Lu H., Martí J. Anchoring mechanisms of oncogenic KRas-4B proteins at cell membranes. In preparation.
- Lu, H., Martí, J. Influence of cholesterol in the anchoring of mutated KRas-4B proteins in model cell membranes. In preparation.

References

[1] M. Karplus and J. A. McCammon, Nature structural biology **9**, 646 (2002).

[2] E. Lindahl and O. Edholm, Biophysical journal **79**, 426 (2000).

[3] S. J. Marrink, E. Lindahl, O. Edholm, and A. E. Mark, Journal of the American Chemical Society **123**, 8638 (2001).

[4] F. Martelli, H.-Y. Ko, C. C. Borallo, and G. Franzese, Frontiers of Physics **13**, 136801 (2018).

[5] C. Calero and G. Franzese, Journal of Molecular Liquids **273**, 488 (2019).

[6] J. F. Nagle and S. Tristram-Nagle, Biochimica et Biophysica Acta (BBA)-Reviews on Biomembranes **1469**, 159 (2000).

[7] O. G. Mouritsen, *Life as a matter of fat* (Springer, 2005).

[8] E. Gorter and F. Grendel, Journal of experimental medicine **41**, 439 (1925).

[9] W. Chen, F. Duša, J. Witos, S. K. Ruokonen, and S. K. Wiedmer, Scientific Reports **8**, 14815 (2018).

[10] K. Damodaran and K. M. Merz Jr, Biophysical Journal **66**, 1076 (1994).

[11] S. W. Chiu, M. Clark, V. Balaji, S. Subramaniam, H. L. Scott, and E. Jakobsson, Biophysical Journal **69**, 1230 (1995).

[12] I. P. Sugár, T. E. Thompson, and R. L. Biltonen, Biophysical Journal **76**, 2099 (1999).

[13] S. E. Feller, Current Opinion in Colloid & Interface Science **5**, 217 (2000).

[14] A. M. Smondyrev and G. A. Voth, Biophysical Journal **82**, 1460 (2002).

[15] A. A. Gurtovenko, M. Patra, M. Karttunen, and I. Vattulainen, Biophysical Journal **86**, 3461 (2004).

[16] W. Shinoda, R. DeVane, and M. L. Klein, The Journal of Physical Chemistry B **114**, 6836 (2010).

[17] J. Yang, C. Calero, and J. Martí, The Journal of Chemical Physics **140**, 03B606_1 (2014).

[18] K. Tu, M. L. Klein, and D. J. Tobias, Biophysical Journal **75**, 2147 (1998).

[19] A. M. Smondyrev and M. L. Berkowitz, Biophysical Journal **77**, 2075 (1999).

[20] S. Chiu, E. Jakobsson, R. J. Mashl, and H. L. Scott, Biophysical Journal **83**, 1842 (2002).

[21] J. Hénin and C. Chipot, Chemical Physics Letters **425**, 329 (2006).

[22] F. de Meyer and B. Smit, Proceedings of the National Academy of Sciences **106**, 3654 (2009).

[23] F. J. M. de Meyer, A. Benjamini, J. M. Rodgers, Y. Misteli, and B. Smit, The Journal of Physical Chemistry B **114**, 10451 (2010).

[24] A. Rabinovich and A. P. Lyubartsev, Polymer Science Series C **55**, 162 (2013).

[25] A. Magarkar, V. Dhawan, P. Kallinteri, T. Viitala, M. Elmowafy, T. Róg, and A. Bunker, Scientific reports **4**, 5005 (2014).

[26] W. Knoll, G. Schmidt, K. Ibel, and E. Sackmann, Biochemistry **24**, 5240 (1985).

[27] P. F. Almeida, W. L. Vaz, and T. Thompson, Biochemistry **31**, 6739 (1992).

[28] F. Richter, G. Rapp, and L. Finegold, Physical Review E **63**, 051914 (2001).

[29] A. Léonard, C. Escrive, M. Laguerre, E. Pebay-Peyroula, W. Néri, T. Pott, J. Katsaras, and E. J. Dufourc, Langmuir **17**, 2019 (2001).

[30] X. Bin, S. L. Horswell, and J. Lipkowski, Biophysical Journal **89**, 592 (2005).

[31] N. Kučerka, J. F. Nagle, J. N. Sachs, S. E. Feller, J. Pencer, A. Jackson, and J. Katsaras, Biophysical Journal **95**, 2356 (2008).

[32] A. Leftin, T. R. Molugu, C. Job, K. Beyer, and M. F. Brown, Biophysical Journal **107**, 2274 (2014).

[33] S. Schurch, M. Lee, and P. Gehr, Pure and applied chemistry **64**, 1745 (1992).

[34] G. S. Ayton and G. A. Voth, Biophysical journal **87**, 3299 (2004).

[35] M. Abraham, G. Whiting, R. Fuchs, and E. Chambers, Chem. Soc., Perkin Trans **2**, 291 (1990).

[36] Y. Zhou, H. Liang, T. Rodkey, N. Ariotti, R. G. Parton, and J. F. Hancock, Molecular and cellular biology **34**, 862 (2014).

[37] X. Chen, F. Sa'Adedin, B. Deme, P. Rao, and J. Bradshaw, Biochimica et Biophysica Acta (BBA)-Biomembranes **1828**, 1982 (2013).

[38] T. Otosu and S. Yamaguchi, Physical Chemistry Chemical Physics (2020).

[39] E. Diamanti, E. Gutiérrez-Pineda, N. Politakos, P. Andreozzi, M. J. Rodriguez-Presa, W. Knoll, O. Azzaroni, C. A. Gervasi, and S. E. Moya, Soft matter **13**, 8922 (2017).

[40] C. Lütgebaucks, C. Macias-Romero, and S. Roke, The Journal of chemical physics **146**, 044701 (2017).

[41] M. Delcea, S. Moreno-Flores, D. Pum, U. B. Sleytr, and J. L. Toca-Herrera, arXiv preprint arXiv:0904.1662 (2009).

[42] H. Ohvo-Rekilä, B. Ramstedt, P. Leppimäki, and J. P. Slotte, Progress in lipid research **41**, 66 (2002).

[43] T. P. McMullen, R. N. Lewis, and R. N. McElhaney, Current opinion in colloid & interface science **8**, 459 (2004).

[44] S. Mohapatra, S. Ranjan, N. Dasgupta, R. Kumar, and S. Thomas, *Characterization and Biology of Nanomaterials for Drug Delivery: Nanoscience and Nanotechnology in Drug Delivery* (Elsevier, 2018).

[45] J. C. Lawrence, D. E. Saslowsky, J. M. Edwardson, and R. M. Henderson, Biophysical journal **84**, 1827 (2003).

[46] J. P. Litz, N. Thakkar, T. Portet, and S. L. Keller, Biophysical journal **110**, 635 (2016).

[47] D. Needham, T. J. McIntosh, and E. Evans, Biochemistry **27**, 4668 (1988).

[48] S. J. Johnson, T. M. Bayerl, D. C. McDermott, G. W. Adam, A. R. Rennie, R. K. Thomas, and E. Sackmann, Biophysical journal **59**, 289 (1991).

[49] A. M. Smondyrev and M. L. Berkowitz, Biophysical journal **80**, 1649 (2001).

[50] R. Ashkar, M. Doktorova, F. A. Heberle, H. Scott, E. Kelley, M. Nagao, R. Usery, F. N. Barrera, G. W. Feigenson, J. Katsaras, *et al.*, Biophysical Journal **116**, 328a (2019).

[51] C. Montour, T. S. Carpenter, and F. C. Lightstone, Biophysical Journal **112**, 74a (2017).

[52] A. Gorman, K. R. Hossain, F. Cornelius, and R. J. Clarke, Biochimica et Biophysica Acta (BBA)-Biomembranes **1862**, 183128 (2020).

[53] W. K. Man, A. De Simone, J. D. Barritt, M. Vendruscolo, C. M. Dobson, and G. Fusco, Frontiers in neuroscience **14**, 18 (2020).

[54] K. M. Sanchez, G. Kang, B. Wu, and J. E. Kim, Biophysical Journal **100**, 2121 (2011).

[55] E. Hartmann, Journal of psychiatric research **17**, 107 (1982).

[56] D. Schneider-Helmert and C. L. Spinweber, Psychopharmacology **89**, 1 (1986).

[57] A. Slominski, I. Semak, A. Pisarchik, T. Sweatman, A. Szczesniewski, and J. Wortsman, FEBS letters **511**, 102 (2002).

[58] A. J. de Jesus and T. W. Allen, Biochimica et Biophysica Acta (BBA)-Biomembranes **1828**, 864 (2013).

[59] J. Guo, Q. Xu, Z. Zheng, S. Zhou, H. Mao, B. Wang, and F. Yan, ACS Macro Letters **4**, 1094 (2015).

[60] F. Shima, T. Akagi, and M. Akashi, Colloid and Polymer Science **292**, 2663 (2014).

[61] A. Fikatas, P. Vervaeke, B. Martínez-Gualda, O. Martí-Marí, S. Noppen, E. Meyen, M. J. Camarasa, A. San-Félix, C. Pannecouque, and D. Schols, Antimicrobial Agents and Chemotherapy (2020).

[62] A. Chattopadhyay, S. S. Rawat, D. V. Greathouse, D. A. Kelkar, and R. E. Koeppe II, Biophysical journal **95**, 166 (2008).

[63] K. R. Park, E. C. Kim, J. T. Hong, and H. M. Yun, Theranostics **8**, 3087 (2018).

[64] S.-H. Jiang, J. Li, F.-Y. Dong, J.-Y. Yang, D.-J. Liu, X.-M. Yang, Y.-H. Wang, M.-W. Yang, X.-L. Fu, X.-X. Zhang, *et al.*, Gastroenterology **153**, 277 (2017).

[65] I. Kostoglou-Athanassiou, Therapeutic Advances in Endocrinology and Metabolism **4**, 13 (2013).

[66] E. A. Daubert and B. G. Condron, Trends in neurosciences **33**, 424 (2010).

[67] J. Peisach, P. Aisen, and W. E. Blumberg, in *Symposium on Copper in Biological Systems (1965: Harriman, NY)* (Academic Press, 1966).

[68] P. Freisinger, R. Horvath, C. Macmillan, J. Peters, and M. Jaksch, Journal of inherited metabolic disease **27**, 67 (2004).

[69] A. A. Ferguson, S. Roy, K. N. Kormanik, Y. Kim, K. J. Dumas, V. B. Ritov, D. Matern, P. J. Hu, and A. L. Fisher, PLoS genetics **9** (2013).

[70] S. S. Mousavi, M. Shohrati, E. Vahedi, M. Abdollahpour-Alitappeh, and Y. Panahi, Iranian journal of pharmaceutical research: IJPR **17**, 136 (2018).

[71] A. Savoca and D. Manca, ADMET and DMPK **7**, 44 (2019).

[72] A. T. Slominski, M. A. Zmijewski, I. Semak, T.-K. Kim, Z. Janjetovic, R. M. Slominski, and J. W. Zmijewski, Cellular and molecular life sciences **74**, 3913 (2017).

[73] A. T. Slominski, R. Hardeland, M. A. Zmijewski, R. M. Slominski, R. J. Reiter, and R. Paus, Journal of Investigative Dermatology **138**, 490 (2018).

[74] G. J. Maestroni, A. Sulli, C. Pizzorni, B. Villaggio, and M. Cutolo, Annals of the New York Academy of Sciences **966**, 271 (2002).

[75] C. M. Forrest, G. M. Mackay, N. Stoy, T. W. Stone, and L. G. Darlington, British journal of clinical pharmacology **64**, 517 (2007).

[76] J. Bang, H. W. Chang, H. R. Jung, C. H. Cho, J. A. Hur, S. I. Lee, T. H. Choi, S. H. Kim, and E. Ha, Rheumatology international **32**, 379 (2012).

[77] C. C. Huang, C. H. Chiou, S. C. Liu, S. L. Hu, C. M. Su, C. H. Tsai, and C. H. Tang, Journal of pineal research **66**, e12560 (2019).

[78] S. A. R. Hussain, Journal of pineal research **42**, 267 (2007).

[79] M. Wang, S. Duan, Z. Zhou, S. Chen, and D. Wang, Ecotoxicology and environmental safety **170**, 68 (2019).

[80] E. J. Costa, R. H. Lopes, and M. T. Lamy-Freund, Journal of Pineal Research **19**, 123 (1995).

[81] D. Bongiorno, L. Ceraulo, M. Ferrugia, F. Filizzola, A. Ruggirello, and V. T. Liveri, Journal of Pineal Research **38**, 292 (2005).

[82] D. Acufla-Castroviejo, G. Escames, M. Macks, A. M. Hoyos, A. M. Carballo, M. Arauzo, R. Montes, and F. Vives, Journal of Pineal Research **19**, 57 (1995).

[83] G. J. Maestroni, Expert Opinion on Investigational Drugs **10**, 467 (2001).

[84] H. Dies, L. Toppozini, and M. C. Rheinstädter, PLoS One **9**, e99124 (2014).

[85] F. Severcan, I. Sahin, and N. Kazancı, Biochimica et Biophysica Acta (BBA)-Biomembranes **1668**, 215 (2005).

[86] H. Dies, B. Cheung, J. Tang, and M. C. Rheinstädter, Biochimica et Biophysica Acta (BBA)-Biomembranes **1848**, 1032 (2015).

[87] E. J. Costa, C. S. Shida, M. H. Biaggi, A. S. Ito, and M. T. Lamy-Freund, FEBS Letters **416**, 103 (1997).

[88] W. M. Yau, W. C. Wimley, K. Gawrisch, and S. H. White, Biochemistry **37**, 14713 (1998).

[89] J. L. MacCallum, W. D. Bennett, and D. P. Tieleman, Biophysical journal **94**, 3393 (2008).

[90] Z. Varga, G. Panyi, M. Péter Jr, C. Pieri, G. Csécsei, S. Damjanovich, and R. Gáspár Jr, Biophysical Journal **80**, 1280 (2001).

[91] H. Lu and J. Martí, The Journal of chemical physics **149**, 164906 (2018).

[92] H. Yu, E. J. Dickson, S.-R. Jung, D.-S. Koh, and B. Hille, The Journal of General Physiology **147**, 63 (2016).

[93] D. Hevia, R. M. Sainz, D. Blanco, I. Quirós, D.-X. Tan, C. Rodríguez, and J. C. Mayo, Journal of Pineal Research **45**, 247 (2008).

[94] D. Hevia, P. González-Menéndez, I. Quiros-González, A. Miar, A. Rodríguez-García, D. X. Tan, R. J. Reiter, J. C. Mayo, and R. M. Sainz, Journal of Pineal Research **58**, 234 (2015).

[95] L. P. H. Andersen, I. Gögenur, J. Rosenberg, and R. J. Reiter, Clinical drug investigation **36**, 169 (2016).

[96] E. Drolle, N. Kučerka, M. Hoopes, Y. Choi, J. Katsaras, M. Karttunen, and Z. Leonenko, Biochimica et Biophysica Acta (BBA)-Biomembranes **1828**, 2247 (2013).

[97] Y. Choi, S. J. Attwood, M. I. Hoopes, E. Drolle, M. Karttunen, and Z. Leonenko, Soft Matter **10**, 206 (2014).

[98] H. I. Petrache, S. W. Dodd, and M. F. Brown, Biophysical journal **79**, 3172 (2000).

[99] P. Leyton, J. Brunet, V. Silva, C. Paipa, M. V. Castillo, and S. A. Brandán, Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy **88**, 162 (2012).

[100] H. Liu, H. Zhang, and B. Jin, Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy **106**, 54 (2013).

[101] M. A. Digman, V. R. Caiolfa, M. Zamai, and E. Gratton, Biophysical journal **94**, L14 (2008).

[102] L. Malacrida, D. M. Jameson, and E. Gratton, Scientific reports **7**, 1 (2017).

[103] R. Nussinov, C. J. Tsai, and H. Jang, Cancer research **78**, 593 (2018).

[104] P. F. Stouten, C. Sander, A. Wittinghofer, and A. Valencia, FEBS letters **320**, 1 (1993).

[105] A. G. Stephen, D. Esposito, R. K. Bagni, and F. McCormick, Cancer cell **25**, 272 (2014).

[106] H. R. Bourne, D. A. Sanders, and F. McCormick, Nature **349**, 117 (1991).

[107] A. Bernards and J. Settleman, Trends in cell biology **14**, 377 (2004).

[108] K. Wennerberg, K. L. Rossman, and C. J. Der, Journal of cell science **118**, 843 (2005).

[109] A. Schmidt and A. Hall, Genes & development **16**, 1587 (2002).

[110] I. Ntai, L. Fornelli, C. J. DeHart, J. E. Hutton, P. F. Doubleday, R. D. LeDuc, A. J. van Nispen, R. T. Fellers, G. Whiteley, E. S. Boja, *et al.*, Proceedings of the National Academy of Sciences **115**, 4140 (2018).

[111] S. A. Forbes, N. Bindal, S. Bamford, C. Cole, C. Y. Kok, D. Beare, M. Jia, R. Shepherd, K. Leung, A. Menzies, *et al.*, Nucleic acids research **39**, D945 (2010).

[112] G. A. Hobbs, C. J. Der, and K. L. Rossman, Journal of cell science **129**, 1287 (2016).

[113] F. D. Tsai, M. S. Lopes, M. Zhou, O. Ponce, J. J. Fiordalisi, J. J. Gierut, A. D. Cox, K. M. Haigis, M. R. Philips, *et al.*, Proceedings of the National Academy of Sciences **112**, 779 (2015).

[114] M. Gelabert-Baldrich, D. Soriano-Castell, M. Calvo, A. Lu, A. Viña-Vilaseca, C. Rentero, A. Pol, S. Grinstein, C. Enrich, and F. Tebar, The FASEB Journal **28**, 3023 (2014).

[115] T. S. Chavan, H. Jang, L. Khavrutskii, S. J. Abraham, A. Banerjee, B. C. Freed, L. Johannessen, S. G. Tarasov, V. Gaponenko, R. Nussinov, *et al.*, Biophysical journal **109**, 2602 (2015).

[116] K. j. Cho, D. van der Hoeven, Y. Zhou, M. Maekawa, X. Ma, W. Chen, G. D. Fairn, and J. F. Hancock, Molecular and cellular biology **36**, 363 (2016).

[117] M. C. Gregory, M. A. McLean, and S. G. Sligar, Biochemical and biophysical research communications **487**, 351 (2017).

[118] Y. Zhou, C. O. Wong, K. j. Cho, D. Van Der Hoeven, H. Liang, D. P. Thakur, J. Luo, M. Babic, K. E. Zinsmaier, M. X. Zhu, *et al.*, Science **349**, 873 (2015).

[119] K. L. Bryant, J. D. Mancias, A. C. Kimmelman, and C. J. Der, Trends in biochemical sciences **39**, 91 (2014).

[120] I. A. Prior, P. D. Lewis, and C. Mattos, Cancer research **72**, 2457 (2012).

[121] K. Marcus and C. Mattos, "Direct attack on ras: intramolecular communication and mutation-specific effects," (2015).

[122] R. Nussinov, C. J. Tsai, and C. Mattos, Trends in molecular medicine **19**, 695 (2013).

[123] J. Downward, Cancer cell **25**, 5 (2014).

[124] S. Lu, H. Jang, S. Muratcioglu, A. Gursoy, O. Keskin, R. Nussinov, and J. Zhang, Chemical reviews **116**, 6607 (2016).

[125] Y. Pylayeva-Gupta, E. Grabocka, and D. Bar-Sagi, Nature Reviews Cancer **11**, 761 (2011).

[126] J. M. Ostrem, U. Peters, M. L. Sos, J. A. Wells, and K. M. Shokat, Nature **503**, 548 (2013).

[127] K. M. Haigis, K. R. Kendall, Y. Wang, A. Cheung, M. C. Haigis, J. N. Glickman, M. Niwa-Kawakita, A. Sweet-Cordero, J. Sebolt-Leopold, K. M. Shannon, *et al.*, Nature genetics **40**, 600 (2008).

[128] K. Ilm, W. Kemmner, M. Osterland, S. Burock, G. Koch, P. Herrmann, P. M. Schlag, and U. Stein, Molecular cancer **14**, 38 (2015).

[129] A. D. Cox, S. W. Fesik, A. C. Kimmelman, J. Luo, and C. J. Der, Nature reviews Drug discovery **13**, 828 (2014).

[130] N. T. Ihle, L. A. Byers, E. S. Kim, P. Saintigny, J. J. Lee, G. R. Blumenschein, A. Tsao, S. Liu, J. E. Larsen, J. Wang, *et al.*, Journal of the National Cancer Institute **104**, 228 (2012).

[131] M. H. Yang, S. Nickerson, E. T. Kim, C. Liot, G. Laurent, R. Spang, M. R. Philips, Y. Shan, D. E. Shaw, D. Bar-Sagi, *et al.*, Proceedings of the National Academy of Sciences **109**, 10843 (2012).

[132] M. H. Yang, G. Laurent, A. S. Bause, R. Spang, N. German, M. C. Haigis, and K. M. Haigis, Molecular Cancer Research **11**, 1072 (2013).

[133] S. Lu, H. Jang, S. Gu, J. Zhang, and R. Nussinov, Chemical Society Reviews **45**, 4929 (2016).

[134] I. M. Ahearn, K. Haigis, D. Bar-Sagi, and M. R. Philips, Nature reviews Molecular cell biology **13**, 39 (2012).

[135] H. Abdelkarim, A. Banerjee, P. Grudzien, N. Leschinsky, M. Abushaer, and V. Gaponenko, International journal of molecular sciences **20**, 5718 (2019).

[136] I. Ahearn, M. Zhou, and M. R. Philips, Cold Spring Harbor perspectives in medicine , a031484 (2018).

[137] P. A. Konstantinopoulos, M. V. Karamouzis, and A. G. Papavassiliou, Nature reviews Drug discovery **6**, 541 (2007).

[138] S. Y. Zhang, B. Sperlich, F. Y. Li, S. Al-Ayoubi, H. X. Chen, Y. F. Zhao, Y. M. Li, K. Weise, R. Winter, and Y. X. Chen, ACS chemical biology **12**, 1703 (2017).

[139] B. Alvarez-Moya, C. Lopez-Alcala, M. Drosten, O. Bachs, and N. Agell, Oncogene **29**, 5911 (2010).

[140] T. G. Bivona, S. E. Quatela, B. O. Bodemann, I. M. Ahearn, M. J. Soskis, A. Mor, J. Miura, H. H. Wiener, L. Wright, S. G. Saba, *et al.*, Molecular cell **21**, 481 (2006).

[141] P. Kollár, J. Rajchard, Z. Balounová, and J. Pazourek, Pharmaceutical biology **52**, 237 (2014).

[142] C. Barceló, N. Paco, M. Morell, B. Alvarez-Moya, N. Bota-Rabassedas, M. Jaumot, F. Vilardell, G. Capella, and N. Agell, Cancer research **74**, 1190 (2014).

[143] H. Jang, S. J. Abraham, T. S. Chavan, B. Hitchinson, L. Khavrutskii, N. I. Tarasova, R. Nussinov, and V. Gaponenko, Journal of Biological Chemistry **290**, 9465 (2015).

[144] B. Alvarez-Moya, C. Barceló, F. Tebar, M. Jaumot, and N. Agell, Small GTPases **2**, 5911 (2011).

[145] S. Dharmaiah, L. Bindu, T. H. Tran, W. K. Gillette, P. H. Frank, R. Ghirlando, D. V. Nissley, D. Esposito, F. McCormick, A. G. Stephen, *et al.*, Proceedings of the National Academy of Sciences **113**, E6766 (2016).

[146] M. Schmick, N. Vartak, B. Papke, M. Kovacevic, D. C. Truxius, L. Rossmannek, and P. I. Bastiaens, Cell **157**, 459 (2014).

[147] G. Zimmermann, B. Papke, S. Ismail, N. Vartak, A. Chandra, M. Hoffmann, S. A. Hahn, G. Triola, A. Wittinghofer, P. I. Bastiaens, *et al.*, Nature **497**, 638 (2013).

[148] S. Muratcioglu, H. Jang, A. Gursoy, O. Keskin, and R. Nussinov, The Journal of Physical Chemistry B **121**, 5917 (2017).

[149] R. E. Salmas, M. Mestanoglu, M. Yurtsever, S. Y. Noskov, and S. Durdagi, Biophysical journal **109**, 1163 (2015).

[150] S. Murarka, P. Martín-Gago, C. Schultz-Fademrecht, A. Al Saabi, M. Baumann, E. K. Fansa, S. Ismail, P. Nussbaumer, A. Wittinghofer, and H. Waldmann, Chemistry–A European Journal **23**, 6083 (2017).

[151] A. M. Wahlstrom, B. A. Cutts, M. Liu, A. Lindskog, C. Karlsson, A.-K. M. Sjogren, K. M. Andersson, S. G. Young, and M. O. Bergo, Blood, The Journal of the American Society of Hematology **112**, 1357 (2008).

[152] R. Nussinov, H. Jang, C. J. Tsai, and F. Cheng, PLoS computational biology **15**, e1006658 (2019).

[153] R. Nussinov and P. G. Wolynes, Physical Chemistry Chemical Physics **16**, 6321 (2014).

[154] S. Kumar, B. Ma, C. J. Tsai, N. Sinha, and R. Nussinov, Protein science **9**, 10 (2000).

[155] B. Schuler, E. A. Lipman, and W. A. Eaton, Nature **419**, 743 (2002).

[156] I. Getov, M. Petukh, and E. Alexov, International journal of molecular sciences **17**, 512 (2016).

[157] S. V. Krivov, Journal of chemical theory and computation **14**, 3418 (2018).

[158] P. Prakash and A. A. Gorfe, The Journal of Physical Chemistry B **123**, 8644 (2019).

[159] T. J. Liao, H. Jang, D. Fushman, and R. Nussinov, Biophysical journal **115**, 629 (2018).

[160] H. Gohlke and D. A. Case, Journal of computational chemistry **25**, 238 (2004).

[161] B. R. Miller III, T. D. McGee Jr, J. M. Swails, N. Homeyer, H. Gohlke, and A. E. Roitberg, Journal of chemical theory and computation **8**, 3314 (2012).

[162] D. Casique-Aguirre, P. Briseño-Díaz, P. García-Gutiérrez, C. H. González-de la Rosa, R. S. Quintero-Barceinas, A. Rojo-Domínguez, I. Vergara, L. A. Medina, J. Correa-Basurto, M. Bello, *et al.*, BMC cancer **18**, 1299 (2018).

[163] P. Cruz-Nova, M. Schnoor, J. Correa-Basurto, M. Bello, P. Briseño-Diaz, A. Rojo-Domínguez, C. M. Ortiz-Mendoza, J. Guerrero-Aguirre, F. J. García-Vázquez, R. Hernández-Rivas, *et al.*, BMC cancer **18**, 1056 (2018).

[164] S. Muraoka, F. Shima, M. Araki, T. Inoue, A. Yoshimoto, Y. Ijiri, N. Seki, A. Tamura, T. Kumasaka, M. Yamamoto, *et al.*, FEBS letters **586**, 1715 (2012).

[165] S. Roet, F. Hooft, P. G. Bolhuis, D. W. Swenson, and J. Vreede, bioRxiv (2020).

[166] J. Yang, [Doctoral dissertation, "Ion binding landscapes and molecular dynamics of phospholipid membranes", Universitat Politècnica de Catalunya] (2015).

[167] S. Jo, J. B. Lim, J. B. Klauda, and W. Im, Biophysical journal **97**, 50 (2009).

[168] S. Jo, T. Kim, V. G. Iyer, and W. Im, Journal of computational chemistry **29**, 1859 (2008).

[169] S. Kim, J. Lee, S. Jo, C. L. Brooks III, H. S. Lee, and W. Im, Journal of computational chemistry **38**, 1879 (2017).

[170] G. E. Tamo, A. Maesani, S. Traeger, M. T. Degiacomi, D. Floreano, and M. Dal Peraro, Biophysical Journal **112**, 289a (2017).

[171] H. M. Berman, T. Battistuz, T. N. Bhat, W. Bluhm, P. E. Bourne, K. Burkhardt, L. Iype, S. Jain, P. Fagan, J. Marvin, *et al.*, ACTA Crystallographica Section D-Biological Crystallography **58** (2002).

[172] S. Jo, T. Kim, and W. Im, PloS one **2** (2007).

[173] E. L. Wu, X. Cheng, S. Jo, H. Rui, K. C. Song, E. M. Dávila-Contreras, Y. Qi, J. Lee, V. Monje Galvan, R. M. Venable, *et al.*, Journal of computational chemistry **35**, 1997 (2014).

[174] J. Lee, X. Cheng, J. M. Swails, M. S. Yeom, P. K. Eastman, J. A. Lemkul, S. Wei, J. Buckner, J. C. Jeong, Y. Qi, *et al.*, Journal of chemical theory and computation **12**, 405 (2016).

[175] T. H. Schmidt and C. Kandt, Journal of chemical information and modeling **52**, 2657 (2012).

[176] W. Humphrey, A. Dalke, K. Schulten, *et al.*, Journal of molecular graphics **14**, 33 (1996).

[177] B. Sommer, T. Dingersen, C. Gamroth, S. E. Schneider, S. Rubert, J. Kruger, and K. J. Dietz, Journal of chemical information and modeling **51**, 1165 (2011).

[178] M. G. Wolf, M. Hoefling, C. Aponte-Santamaría, H. Grubmüller, and G. Groenhof, Journal of computational chemistry **31**, 2169 (2010).

[179] L. Martínez, R. Andrade, E. G. Birgin, and J. M. Martínez, Journal of computational chemistry **30**, 2157 (2009).

[180] W. Im and B. Roux, Journal of molecular biology **319**, 1177 (2002).

[181] A. Pérez, I. Marchán, D. Svozil, J. Sponer, T. E. Cheatham III, C. A. Laughton, and M. Orozco, Biophysical journal **92**, 3817 (2007).

[182] J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, and D. A. Case, Journal of computational chemistry **25**, 1157 (2004).

[183] K. L. Meagher, L. T. Redman, and H. A. Carlson, Journal of computational chemistry **24**, 1016 (2003).

[184] E. Harder, W. Damm, J. Maple, C. Wu, M. Reboul, J. Y. Xiang, L. Wang, D. Lupyan, M. K. Dahlgren, J. L. Knight, *et al.*, Journal of chemical theory and computation **12**, 281 (2015).

[185] J. Huang, S. Rauscher, G. Nawrocki, T. Ran, M. Feig, B. L. de Groot, H. Grubmüller, and A. D. MacKerell Jr, Nature methods **14**, 71 (2017).

[186] P. Robustelli, S. Piana, and D. E. Shaw, Proceedings of the National Academy of Sciences **115**, E4758 (2018).

[187] J. B. Klauda, R. M. Venable, J. A. Freites, J. W. O'Connor, D. J. Tobias, C. Mondragon-Ramirez, I. Vorobyov, A. D. MacKerell Jr, and R. W. Pastor, The journal of physical chemistry B **114**, 7830 (2010).

[188] R. Pastor and A. MacKerell Jr, The journal of physical chemistry letters **2**, 1526 (2011).

[189] J. Huang and A. D. MacKerell Jr, Journal of computational chemistry **34**, 2135 (2013).

[190] S. E. Feller and A. D. MacKerell, The Journal of Physical Chemistry B **104**, 7510 (2000).

[191] J. B. Klauda, B. R. Brooks, A. D. MacKerell, R. M. Venable, and R. W. Pastor, The journal of physical chemistry B **109**, 5300 (2005).

[192] J. B. Klauda, R. W. Pastor, and B. R. Brooks, The Journal of Physical Chemistry B **109**, 15684 (2005).

[193] E. Fermi, J. Pasta, and S. Ulam, Collected papers of Enrico Fermi, edited by E. Segré,(University of Chicago, Chicago, 1965) **2**, 978 (1965).

[194] B. J. Alder and T. E. Wainwright, The Journal of Chemical Physics **31**, 459 (1959).

[195] A. Rahman, Physical review **136**, A405 (1964).

[196] D. P. Kroese, T. Brereton, T. Taimre, and Z. I. Botev, Wiley Interdisciplinary Reviews: Computational Statistics **6**, 386 (2014).

[197] D. Frenkel and B. Smit, "Understanding molecular simulation: From algorithms to applications," (2002).

[198] R. Petrenko and J. Meller, eLS (2010).

[199] L. Verlet, Physical review **159**, 98 (1967).

[200] S. A. Hollingsworth and R. O. Dror, Neuron **99**, 1129 (2018).

[201] G. B. Dawe, M. Musgaard, M. R. Aurousseau, N. Nayeem, T. Green, P. C. Biggin, and D. Bowie, Neuron **89**, 1264 (2016).

[202] L. Delemotte, M. Tarek, M. L. Klein, C. Amaral, and W. Treptow, Proceedings of the National Academy of Sciences **108**, 6109 (2011).

[203] R. O. Dror, H. F. Green, C. Valant, D. W. Borhani, J. R. Valcourt, A. C. Pan, D. H. Arlow, M. Canals, J. R. Lane, R. Rahmani, *et al.*, Nature **503**, 295 (2013).

[204] M. Ø. Jensen, V. Jogini, D. W. Borhani, A. E. Leffler, R. O. Dror, and D. E. Shaw, Science **336**, 229 (2012).

[205] H. E. Kato, Y. S. Kim, J. M. Paggi, K. E. Evans, W. E. Allen, C. Richardson, K. Inoue, S. Ito, C. Ramakrishnan, L. E. Fenno, *et al.*, Nature **561**, 349 (2018).

[206] A. Manglik, H. Lin, D. K. Aryal, J. D. McCorvy, D. Dengler, G. Corder, A. Levit, R. C. Kling, V. Bernat, H. Hübner, *et al.*, Nature **537**, 185 (2016).

[207] J. D. McCorvy, K. V. Butler, B. Kelly, K. Rechsteiner, J. Karpiak, R. M. Betz, B. L. Kormos, B. K. Shoichet, R. O. Dror, J. Jin, *et al.*, Nature chemical biology **14**, 126 (2018).

[208] T. I. Cheatham, J. Miller, T. Fox, T. Darden, and P. Kollman, Journal of the American Chemical Society **117**, 4193 (1995).

[209] G. Bryan, Nature **66**, 291 (1902).

[210] M. P. Allen and D. J. Tildesley, *Computer simulation in chemical physics*, Vol. 397 (Springer Science & Business Media, 2012).

[211] M. Tuckerman, *Statistical mechanics: theory and molecular simulation* (Oxford university press, 2010).

[212] B. Cheng and M. Ceriotti, Physical Review B **97**, 054102 (2018).

[213] P. G. Bolhuis, D. Chandler, C. Dellago, and P. L. Geissler, Annual review of physical chemistry **53**, 291 (2002).

[214] T. Huber, A. E. Torda, and W. F. Van Gunsteren, Journal of computer-aided molecular design **8**, 695 (1994).

[215] S. Park and K. Schulten, The Journal of chemical physics **120**, 5946 (2004).

[216] H. Grubmüller, Physical Review E **52**, 2893 (1995).

[217] E. M. Müller, A. de Meijere, and H. Grubmüller, The Journal of chemical physics **116**, 897 (2002).

[218] J. Kästner, Wiley Interdisciplinary Reviews: Computational Molecular Science **1**, 932 (2011).

[219] P. Virnau and M. Müller, The Journal of chemical physics **120**, 10925 (2004).

[220] A. Warmflash, P. Bhimalapuram, and A. R. Dinner, The Journal of chemical physics **127**, 114109 (2007).

[221] C. Matthews, J. Weare, A. Kravtsov, and E. Jennings, Monthly Notices of the Royal Astronomical Society **480**, 4069 (2018).

[222] S. Awasthi, V. Kapil, and N. N. Nair, Journal of computational chemistry **37**, 1413 (2016).

[223] H. Oshima, S. Re, and Y. Sugita, Journal of chemical theory and computation **15**, 5199 (2019).

[224] E. Darve and A. Pohorille, The Journal of chemical physics **115**, 9169 (2001).

[225] A. Barducci, M. Bonomi, and M. Parrinello, Wiley Interdisciplinary Reviews: Computational Molecular Science **1**, 826 (2011).

[226] A. Laio and M. Parrinello, Proceedings of the National Academy of Sciences **99**, 12562 (2002).

[227] A. Laio, A. Rodriguez-Fortea, F. L. Gervasio, M. Ceccarelli, and M. Parrinello, The journal of physical chemistry B **109**, 6714 (2005).

[228] F. L. Gervasio, A. Laio, and M. Parrinello, Journal of the American Chemical Society **127**, 2600 (2005).

[229] S. Angioletti-Uberti, M. Ceriotti, P. D. Lee, and M. W. Finnis, Physical Review B **81**, 125416 (2010).

[230] M. Bonomi, C. Camilloni, and M. Vendruscolo, Scientific reports **6**, 31232 (2016).

[231] A. Barducci, G. Bussi, and M. Parrinello, Physical Review Letters **100**, 020603 (2008).

[232] G. Bussi, A. Laio, and M. Parrinello, Physical review letters **96**, 090601 (2006).

[233] K. Hornik, Wiley Interdisciplinary Reviews: Computational Statistics **4**, 394 (2012).

[234] P. Murrell, *R graphics* (CRC Press, 2018).

[235] W. Chang, *R graphics cookbook: practical recipes for visualizing data* (O'Reilly Media, 2018).

[236] R. Ihaka and R. Gentleman, Journal of computational and graphical statistics **5**, 299 (1996).

[237] R. S. Bivand, E. J. Pebesma, V. Gomez-Rubio, and E. J. Pebesma, *Applied spatial data analysis with R*, Vol. 747248717 (Springer, 2008).

[238] C. Reimann, P. Filzmoser, R. Garrett, and R. Dutter, *Statistical data analysis explained: applied environmental statistics with R* (John Wiley & Sons, 2011).

[239] P. Hošek and V. Spiwok, Computer Physics Communications **198**, 222 (2016).

[240] R. Olsen, G. Kroes, G. Henkelman, A. Arnaldsson, and H. Jónsson, The Journal of chemical physics **121**, 9776 (2004).

[241] O. P. Koistinen, F. B. Dagbjartsdóttir, V. Ásgeirsson, A. Vehtari, and H. Jónsson, The Journal of chemical physics **147**, 152720 (2017).

[242] E. L. Kolsbjerg, M. N. Groves, and B. Hammer, The Journal of chemical physics **145**, 094107 (2016).

[243] D. Chen, L. L. Costello, C. B. Geller, T. Zhu, and D. L. McDowell, Acta Materialia **168**, 436 (2019).

[244] J. A. G. Torres, P. C. Jennings, M. H. Hansen, J. R. Boes, and T. Bligaard, Physical review letters **122**, 156001 (2019).

[245] A. Ghasemi, P. Xiao, and W. Gao, The Journal of Chemical Physics **151**, 054110 (2019).

[246] P. F. Bessarab, V. M. Uzdin, and H. Jónsson, Physical Review B **85**, 184409 (2012).

[247] P. F. Bessarab, V. M. Uzdin, and H. Jónsson, Zeitschrift für Physikalische Chemie **227**, 1543 (2013).

[248] T. E. Creighton, C. J. Bagley, L. Cooper, N. J. Darby, R. B. Freedman, J. Kemmink, and A. Sheikh, Journal of molecular biology **232**, 1176 (1993).

[249] J. Krol, A. Fiszer, A. Mykowska, K. Sobczak, M. de Mezer, and W. J. Krzyzosiak, Molecular cell **25**, 575 (2007).

[250] Y. Wang, C. B. Harrison, K. Schulten, and J. A. McCammon, Computational science & discovery **4**, 015002 (2011).

[251] D. E. Tanner, K. Y. Chan, J. C. Phillips, and K. Schulten, Journal of chemical theory and computation **7**, 3635 (2011).

[252] W. Jiang, J. C. Phillips, L. Huang, M. Fajer, Y. Meng, J. C. Gumbart, Y. Luo, K. Schulten, and B. Roux, Computer physics communications **185**, 908 (2014).

[253] J. Comer, J. C. Phillips, K. Schulten, and C. Chipot, Journal of chemical theory and computation **10**, 5276 (2014).

[254] M. C. Melo, R. C. Bernardi, T. Rudack, M. Scheurer, C. Riplinger, J. C. Phillips, J. D. Maia, G. B. Rocha, J. V. Ribeiro, J. E. Stone, *et al.*, Nature methods **15**, 351 (2018).

[255] S. Jo and W. Jiang, Computer Physics Communications **197**, 304 (2015).

[256] Y. T. Pang, Y. Miao, Y. Wang, and J. A. McCammon, Journal of chemical theory and computation **13**, 9 (2017).

[257] D. Hudiyanti, M. Radifar, T. J. Raharjo, N. Narsito, and S. Noegrohati, Journal of Chemistry (2014).

[258] W. Jiang, C. Chipot, and B. Roux, Journal of chemical information and modeling **59**, 3794 (2019).

[259] J. C. Gumbart, C. Balusek, H. Hwang, C. H. Lau, K. Lundquist, A. Hazel, A. Pavlova, D. Lynch, P. Reggio, and Y. Wang, Biophysical Journal **116**, 141a (2019).

[260] G. Fiorin, G. Brannigan, and J. Hénin, Biophysical Journal **116**, 142a (2019).

[261] L. Kalé, R. Skeel, M. Bhandarkar, R. Brunner, A. Gursoy, N. Krawetz, J. Phillips, A. Shinozaki, K. Varadarajan, and K. Schulten, Journal of Computational Physics **151**, 283 (1999).

[262] H. J. Berendsen, D. van der Spoel, and R. van Drunen, Computer physics communications **91**, 43 (1995).

[263] M. Bonomi, D. Branduardi, G. Bussi, C. Camilloni, D. Provasi, P. Raiteri, D. Donadio, F. Marinelli, F. Pietrucci, R. A. Broglia, *et al.*, Computer Physics Communications **180**, 1961 (2009).

[264] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, The Journal of chemical physics **79**, 926 (1983).

[265] J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kale, and K. Schulten, Journal of Computational Chemistry **26**, 1781 (2005).

[266] U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, and L. G. Pedersen, The Journal of chemical physics **103**, 8577 (1995).

[267] H. J. Berendsen, J. v. Postma, W. F. van Gunsteren, A. DiNola, and J. R. Haak, The Journal of chemical physics **81**, 3684 (1984).

[268] S. E. Feller, Y. Zhang, R. W. Pastor, and B. R. Brooks, The Journal of chemical physics **103**, 4613 (1995).

[269] E. Lindahl and O. Edholm, The Journal of chemical physics **115**, 4938 (2001).

[270] F. Jähnig, Biophysical journal **71**, 1348 (1996).

[271] F. Brochard, P. De Gennes, and P. Pfeuty, Journal de Physique **37**, 1099 (1976).

[272] H. Schindler, FEBS letters **122**, 77 (1980).

[273] K. Tu, D. J. Tobias, J. K. Blasie, and M. L. Klein, Biophysical journal **70**, 595 (1996).

[274] D. P. Tieleman and H. Berendsen, The Journal of chemical physics **105**, 4871 (1996).

[275] J. Yang, C. Calero, M. Bonomi, and J. Martí, Journal of chemical theory and computation **11**, 4495 (2015).

[276] J. Martı, The Journal of chemical physics **110**, 6876 (1999).

[277] D. Chandler, *Introduction to Modern Statistical Mechanics* (Oxford University Press, 1987).

[278] D. Trzesniak, A. P. E. Kunz, and W. F. van Gunsteren, ChemPhysChem **8**, 162 (2007).

[279] P. L. Geissler, C. Dellago, D. Chandler, J. Hutter, and M. Parrinello, Science **291**, 2121 (2001).

[280] J. Martí and F. S. Csajka, Physical Review E **69**, 061918 (2004).

[281] J. L. Thewalt and M. Bloom, Biophysical journal **63**, 1176 (1992).

[282] N. Kučerka, M. P. Nieh, and J. Katsaras, Biochimica et Biophysica Acta (BBA)-Biomembranes **1808**, 2761 (2011).

[283] J. Martí Rabassa and H. Lu, Procedia computer science **108**, 1242 (2017).

[284] G. W. Stockton and I. C. Smith, Chemistry and physics of lipids **17**, 251 (1976).

[285] C. Hofsäß, E. Lindahl, and O. Edholm, Biophysical journal **84**, 2192 (2003).

[286] T. P. Trouard, A. A. Nevzorov, T. M. Alam, C. Job, J. Zajicek, and M. F. Brown, The Journal of chemical physics **110**, 8802 (1999).

[287] O. Berger, O. Edholm, and F. Jähnig, Biophysical journal **72**, 2002 (1997).

[288] P. R. Pandey and S. Roy, The Journal of Physical Chemistry B **115**, 3155 (2011).

[289] O. Edholm and J. F. Nagle, Biophysical journal **89**, 1827 (2005).

[290] Y. Wang, P. Gkeka, J. E. Fuchs, K. R. Liedl, and Z. Cournia, Biochimica et Biophysica Acta (BBA)-Biomembranes **1858**, 2846 (2016).

[291] C. L. Armstrong, M. A. Barrett, A. Hiess, T. Salditt, J. Katsaras, A. C. Shi, and M. C. Rheinstädter, European Biophysics Journal **41**, 901 (2012).

[292] J. Yang, J. Martí, and C. Calero, Soft Matter **12**, 4557 (2016).

[293] J. Martı, F. S. Csajka, and D. Chandler, Chemical Physics Letters **328**, 169 (2000).

[294] J. Marti, Physical Review E **61**, 449 (2000).

[295] J. B. Klauda, B. R. Brooks, and R. W. Pastor, The Journal of chemical physics **125**, 144710 (2006).

[296] R. M. Venable, H. I. Ingólfsson, M. G. Lerner, B. S. Perrin Jr, B. A. Camley, S. J. Marrink, F. L. Brown, and R. W. Pastor, The Journal of Physical Chemistry B **121**, 3443 (2017).

[297] I. C. Yeh and G. Hummer, The Journal of Physical Chemistry B **108**, 15873 (2004).

[298] D. A. McQuarrie, *Statistical Mechanics* (University Science Books, 2000).

[299] J. Martí, J. Padró, and E. Guardia, Molecular Simulation **11**, 321 (1993).

[300] J. Martí, E. Guàrdia, and J. Padró, The Journal of chemical physics **101**, 10883 (1994).

[301] M. Praprotnik and D. Janežič, The Journal of chemical physics **122**, 174103 (2005).

[302] D. Eisenberg, W. Kauzmann, and W. Kauzmann, *The structure and properties of water* (Oxford University Press on Demand, 2005).

[303] I. Wood, M. F. Martini, and M. Pickholz, Journal of Molecular Structure **1045**, 124 (2013).

[304] M. W. Feyereisen, D. Feller, and D. A. Dixon, The Journal of Physical Chemistry **100**, 2993 (1996).

[305] G. H. Peters, M. Werge, M. N. Elf-Lind, J. J. Madsen, G. F. Velardez, and P. Westh, Chemistry and physics of lipids **184**, 7 (2014).

[306] A. Brzezinski, New England journal of medicine **336**, 186 (1997).

[307] R. Hardeland, S. R. Pandi-Perumal, and D. P. Cardinali, The international journal of biochemistry & cell biology **38**, 313 (2006).

[308] H. Lu and J. Marti, Chemical Physics Letters **712**, 190 (2018).

[309] A. Komornicki, K. Ishida, K. Morokuma, R. Ditchfield, and M. Conrad, Chemical Physics Letters **45**, 595 (1977).

[310] X. J. Zhang, C. Shang, and Z. P. Liu, Journal of Chemical Theory and Computation **9**, 5745 (2013).

[311] J. Yang, M. Bonomi, C. Calero, and J. Martí, Physical Chemistry Chemical Physics **18**, 9036 (2016).

[312] J. P. Jambeck and A. P. Lyubartsev, The Journal of Physical Chemistry Letters **4**, 1781 (2013).

[313] C. Chipot, Wiley Interdisciplinary Reviews: Computational Molecular Science **4**, 71 (2014).

[314] P. L. Yeagle, *The membranes of cells* (Academic Press, 2016).

[315] N. Ganesan, B. A. Bauer, T. R. Lucas, S. Patel, and M. Taufer, Journal of computational chemistry **32**, 2958 (2011).

[316] K. Modig, B. G. Pfrommer, and B. Halle, Physical review letters **90**, 075502 (2003).

[317] F. M. Ytreberg, R. H. Swendsen, and D. M. Zuckerman, The Journal of chemical physics **125**, 184114 (2006).

[318] H. M. Senn and W. Thiel, in *Atomistic approaches in modern biology* (Springer, 2006) pp. 173–290.

[319] C. Dellago and P. G. Bolhuis, in *Advanced Computer Simulation Approaches for Soft Matter Sciences III* (Springer, 2009) pp. 167–233.

[320] G. M. Florio and T. S. Zwier, The Journal of Physical Chemistry A **107**, 974 (2003).

[321] G. Singh, J. Abbas, S. D. Dogra, R. Sachdeva, B. Rai, S. Tripathi, S. Prakash, V. Sathe, and G. Saini, Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy **118**, 73 (2014).

[322] G. D. Fleming, R. Koch, J. M. Perez, and J. L. Cabrera, Vibrational Spectroscopy **80**, 70 (2015).

[323] E. Pięta, C. Paluszkiewicz, M. Oćwieja, and W. M. Kwiatek, Applied Surface Science **404**, 168 (2017).

[324] J. À. Padró and J. Martı, The Journal of Chemical Physics **120**, 1659 (2004).

[325] N. Vlachy, B. Jagoda-Cwiklik, R. Vácha, D. Touraud, P. Jungwirth, and W. Kunz, Advances in Colloid and Interface Science **146**, 42 (2009).

[326] S. Nosé, The Journal of Chemical Physics **81**, 511 (1984).

[327] W. G. Hoover, Physical review A **31**, 1695 (1985).

[328] H. Lu and J. Marti, Scientific Reports **10**, 1 (2020).

[329] B. Ensing, A. Laio, M. Parrinello, and M. L. Klein, The journal of physical chemistry B **109**, 6676 (2005).

[330] H. Lu and J. Martí, PloS one **14**, e0224624 (2019).

[331] D. Branduardi, F. L. Gervasio, and M. Parrinello, The Journal of Chemical Physics **126**, 054103 (2007).

[332] M. Chen and W. Yang, Journal of Computational Chemistry **30**, 1649 (2009).

[333] M. Moradi, G. Enkavi, and E. Tajkhorshid, Nature Communications **6**, 8393 (2015).

[334] K. J. Laidler, *Theories of chemical reaction rates* (McGraw-Hill New York, 1969).

[335] P. Pechukas, Annual Review of Physical Chemistry **32**, 159 (1981).

[336] W. C. Hung, M. T. Lee, F. Y. Chen, and H. W. Huang, Biophysical journal **92**, 3960 (2007).

[337] J. Peters, J. Marion, F. J. Becher, M. Trapp, T. Gutberlet, D. Bicout, and T. Heimburg, Scientific Reports **7**, 15339 (2017).

[338] Y. Wang, E. Gallagher, C. Jorgensen, E. P. Troendle, D. Hu, P. C. Searson, and M. B. Ulmschneider, Scientific Reports **9**, 6117 (2019).

[339] B. Roux, Biophysical journal **77**, 139 (1999).

[340] A. G. Turjanski, R. E. Rosenstein, and D. A. Estrin, Journal of Medicinal Chemistry **41**, 3684 (1998).

[341] A. G. Turjanski, D. A. Sáenz, F. Doctorovich, D. A. Estrin, and R. E. Rosenstein, Journal of Pineal Research **31**, 97 (2001).

[342] A. G. Turjanski, D. A. Estrin, R. E. Rosenstein, J. E. McCormick, S. R. Martin, A. Pastore, R. R. Biekofsky, and V. Martorana, Protein Science **13**, 2925 (2004).

[343] G. M. Florio, R. A. Christie, K. D. Jordan, and T. S. Zwier, Journal of the American Chemical Society **124**, 10236 (2002).

[344] D. S. Goodsell, Stem cells **17**, 235 (1999).

[345] E. D. Pleasance, R. K. Cheetham, P. J. Stephens, D. J. McBride, S. J. Humphray, C. D. Greenman, I. Varela, M. L. Lin, G. R. Ordóñez, G. R. Bignell, *et al.*, Nature **463**, 191 (2010).

[346] H. Jang, A. Banerjee, K. Marcus, L. Makowski, C. Mattos, V. Gaponenko, and R. Nussinov, Structure **27**, 1647 (2019).

[347] T. Yeung, G. E. Gilbert, J. Shi, J. Silvius, A. Kapus, and S. Grinstein, Science **319**, 210 (2008).

[348] E. Barklis, A. G. Stephen, A. O. Staubus, R. L. Barklis, and A. Alfadhli, Journal of molecular biology **431**, 3706 (2019).

[349] L. Janosi and A. A. Gorfe, Biophysical journal **99**, 3666 (2010).

[350] P. Liu, Y. Wang, and X. Li, Acta Pharmaceutica Sinica B (2019).

[351] J. T. Swarthout, S. Lobo, L. Farh, M. R. Croke, W. K. Greentree, R. J. Deschenes, and M. E. Linder, Journal of Biological Chemistry **280**, 31141 (2005).

[352] A. Kouranov, L. Xie, J. de la Cruz, L. Chen, J. Westbrook, P. E. Bourne, and H. M. Berman, Nucleic acids research **34**, D302 (2006).

[353] Y. Lyu, N. Xiang, J. Mondal, X. Zhu, and G. Narsimhan, The Journal of Physical Chemistry B **122**, 2341 (2018).

[354] A. Liu and X. Qi, Computational molecular bioscience **2**, 78 (2012).

[355] D. Poger and A. E. Mark, Journal of Chemical Theory and Computation **6**, 325 (2010).

[356] J. Pan, S. Tristram-Nagle, N. Kučerka, and J. F. Nagle, Biophysical journal **94**, 117 (2008).

[357] V. Chaban, Chemical Physics Letters **616**, 25 (2014).

[358] H. I. Petrache, S. Tristram-Nagle, K. Gawrisch, D. Harries, V. A. Parsegian, and J. F. Nagle, Biophysical journal **86**, 1574 (2004).

[359] J. Pan, S. Tristram-Nagle, and J. F. Nagle, Physical Review E **80**, 021931 (2009).

[360] E. Novakova, K. Giewekemeyer, and T. Salditt, Physical Review E **74**, 051911 (2006).

[361] Z. S. Hendsch and B. Tidor, Protein Science **3**, 211 (1994).

[362] B. H. Honig and W. L. Hubbell, Proceedings of the National Academy of Sciences **81**, 5412 (1984).

[363] D. Abankwa, A. A. Gorfe, K. Inder, and J. F. Hancock, Proceedings of the National Academy of Sciences **107**, 1130 (2010).

[364] K. Weise, S. Kapoor, C. Denter, J. Nikolaus, N. Opitz, S. Koch, G. Triola, A. Herrmann, H. Waldmann, and R. Winter, Journal of the American Chemical Society **133**, 880 (2011).

[365] H. Jang, A. Banerjee, T. S. Chavan, S. Lu, J. Zhang, V. Gaponenko, and R. Nussinov, The FASEB Journal **30**, 1643 (2016).

[366] Y. Zhou, P. Prakash, H. Liang, K. J. Cho, A. A. Gorfe, and J. F. Hancock, Cell **168**, 239 (2017).

[367] K. Gunasekaran, B. Ma, and R. Nussinov, Proteins: Structure, Function, and Bioinformatics **57**, 433 (2004).

[368] M. Bernetti, M. Masetti, M. Recanatini, R. E. Amaro, and A. Cavalli, Journal of chemical theory and computation **15**, 5689 (2019).

[369] I. Buch, T. Giorgino, and G. De Fabritiis, Proceedings of the National Academy of Sciences **108**, 10184 (2011).

[370] Y. Shan, E. T. Kim, M. P. Eastwood, R. O. Dror, M. A. Seeliger, and D. E. Shaw, Journal of the American Chemical Society **133**, 9181 (2011).

[371] A. Kazi, S. Xiang, H. Yang, L. Chen, P. Kennedy, M. Ayaz, S. Fletcher, C. Cummings, H. R. Lawrence, F. Beato, *et al.*, Clinical Cancer Research **25**, 5984 (2019).

[372] A. B. Nørholm, P. Francotte, L. Olsen, C. Krintel, K. Frydenvang, E. Goffin, S. Challal, L. Danober, I. Botez-Pop, P. Lestage, *et al.*, Journal of Medicinal Chemistry **56**, 8736 (2013).

[373] C. Miao, F. Dong, L. Jia, W. Li, M. Wang, Q.-H. Zheng, and Z. Xu, Bioorganic & medicinal chemistry letters **29**, 1177 (2019).

[374] X. Chen, C. Zhu, F. Guo, X. Qiu, Y. Yang, S. Zhang, M. He, S. Parveen, C. Jing, Y. Li, *et al.*, Journal of medicinal chemistry **53**, 8330 (2010).